JMIR mHealth and uHealth

Impact Factor (2024): 5.4
Volume 10 (2022), Issue 7 ISSN 2291-5222 Editor in Chief: Lorraine Buis, PhD, MSI

Contents

Reviews

Original Papers

Review

The Association Between Smartphone App–Based Self-monitoring of Hypertension-Related Behaviors and Reductions in High Blood Pressure: Systematic Review and Meta-analysis

Aikaterini Kassavou[1], BSc, MSc, PhD; Michael Wang[1], BSc, MPhil; Venus Mirzaei[1], BSc, MSc; Sonia Shpendi[1], BSc, MSc, PhD; Rana Hasan[1], BSc

The Primary Care Unit, Department of Public Health and Primary Care, The University of Cambridge, Cambridge, United Kingdom

Corresponding Author:
Aikaterini Kassavou, BSc, MSc, PhD
The Primary Care Unit
Department of Public Health and Primary Care
The University of Cambridge
East Forvie Building
Cambridge, CB2 0SR
United Kingdom
Phone: 44 1223330456
Email: aikaterini.kassavou@wbs.ac.uk

Abstract

Background: Self-monitoring of behavior can support lifestyle modifications; however, we do not know whether such interventions are effective in supporting positive changes in hypertension-related health behaviors and thus in reducing blood pressure in patients treated for hypertension.

Objective: This systematic literature review evaluates the extent to which smartphone app–based self-monitoring of health behavior supports reductions in blood pressure and changes in hypertension-related behaviors. It also explores the behavioral components that might explain intervention effectiveness.

Methods: A systematic search of 7 databases was conducted in August 2021. Article screening, study and intervention coding, and data extraction were completed independently by reviewers. The search strategy was developed using keywords from previous reviews and relevant literature. Trials involving adults, published after the year 2000, and in the English language were considered for inclusion. The random-effects meta-analysis method was used to account for the distribution of the effect across the studies.

Results: We identified 4638 articles, of which 227 were included for full-text screening. A total of 15 randomized controlled trials were included in the review. In total, 7415 patients with hypertension were included in the meta-analysis. The results indicate that app-based behavioral self-monitoring interventions had a small but significant effect in reducing systolic blood pressure (SBP), on average, by 1.64 mmHg (95% CI 2.73-0.55, n=7301; odds ratio [OR] 1.60, 95% CI 0.74-3.42, n=114) and in improving changes in medication adherence behavior (standardized mean difference [SMD] 0.78, 95% CI 0.22-1.34) compared to usual care or minimal intervention. The review found the intervention had a small effect on supporting improvements in healthy diet by changing habits related to high sodium food (SMD –0.44, 95% CI –0.79 to –0.08) and a trend, although insignificant, toward supporting smoking cessation, low alcohol consumption, and better physical activity behaviors. A subgroup analysis found that behavioral self-monitoring interventions combined with tailored advice resulted in higher and significant changes in both SBP and diastolic blood pressure (DBP) in comparison to those not providing tailored advice (SBP: –2.92 mmHg, 95% CI –3.94 to –1.90, n=3102 vs –0.72 mmHg, 95% CI –1.67 to 0.23, n=4199, χ^2=9.65, *P*=.002; DBP: –2.05 mmHg, 95% CI –3.10 to –1.01, n=968 vs 1.54 mmHg, 95% CI –0.53 to 3.61, n=400, χ^2=9.19, *P*=.002).

Conclusions: Self-monitoring of hypertension-related behaviors via smartphone apps combined with tailored advice has a modest but potentially clinically significant effect on blood pressure reduction. Future studies could use rigorous methods to explore its effects on supporting changes in both blood pressure and hypertension-related health behaviors to inform recommendations for policy making and service provision.

Trial Registration: PROSPERO CRD42019136158; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=136158

(JMIR Mhealth Uhealth 2022;10(7):e34767) doi:10.2196/34767

KEYWORDS

self-monitoring; smartphone apps; behavior change; hypertension; blood pressure; mobile health; mHealth; mobile app; self-management; lifestyle

Introduction

Hypertension, or high blood pressure, affects over 1 billion adults globally and is a leading risk factor for premature morbidity and mortality [1,2]. However, only about half of adults with hypertension achieve adequate blood pressure control, increasing both health care resources and the cost required for treatment [3]. In England, hypertension is estimated to cost the National Health Service an excess of £2 billion (US $2.4 billion) per year [4]. Although various risk factors contribute to poorly controlled blood pressure, nonadherence to prescribed health behaviors, like compliance to prescribed medications [5], improvements in physical activity [6,7], low salt intake [8,9], consumption of fruits and vegetables [10], low alcohol consumption [11], and smoking cessation [12], independently account for most of these uncontrolled cases.

Modifying health-related behaviors to address the underlying risk factors of hypertension could result in clinically significant health improvements and reduce morbidity, mortality, and treatment cost. Practitioners have an important role in prescribing lifestyle modifications; however, the time they can spend providing advice about and supporting adherence to health behavior change recommendations is limited and expensive [13], and there is currently limited evidence on effective interventions to support health behavior change in patients treated for hypertension [14-16].

There is growing interest in the potential of digital innovations as an inexpensive and scalable method to deliver personalized advice to people with long-term health conditions, enabling them to improve adherence to their recommended health behavior modifications and achieve health improvements [17-19]. Mobile apps, facilitated via digital technologies such as computers, smartphones, tablets, and other mobile devices are accessible to large numbers of people and in different settings [20]. Smartphone apps appear to be promising due to their potential to complement physician efforts and engage patients in decision-making processes regarding their health care [21,22]. Users of app-based interventions can receive real-time advice about patterns of health behaviors that impact their long-term health condition [23], with the potential to eliminate barriers that rely on memory and are prone to inaccuracies and recall bias, and to better inform shared decision-making during usual care consultations.

Moreover, reporting and monitoring health behaviors using apps could act as a behavior change strategy to support the individual in self-regulating health behaviors and thus lead to sustained improvements in clinical health indicators [23,24]. Self-monitoring of behavior could underpin individual behavior change by modifying self-regulation processes, for example, by enabling patients to reflect on and change their health behaviors informed by behavioral performance [24-26]. Interventions providing advice to support patients' self-regulatory processes might be more effective at improving long-term treatment adherence and thus might be a cost-effective solution for sustained health care.

While smartphone app–based self-monitoring of health behaviors has the potential to have a direct, positive effect on patients' health and an indirect effect on service provision, to date there is a lack of evidence on the clinical effectiveness of app-based behavioral self-monitoring to support patients treated for hypertension.

Previous systematic reviews have evaluated the impact of app-based interventions to support changes in behavioral or clinical outcomes, suggesting some promising evidence of their potential effectiveness [18,27-30]. Furthermore, content analysis of publicly available apps suggests that such interventions are complex and often consist of one or a combination of the following components: generic education about the health condition, provision of social support, reminders and feedback about the behavior, feedback on blood pressure measurements, or provision of clinical advice about medicine adjustments. However, previous reviews have neither investigated the impact of behavioral self-monitoring via smartphone apps on both clinical and behavioral effectiveness nor disentangled the components that account for clinical effectiveness in patients treated for hypertension.

This review investigates whether app-based self-monitoring of health behavior reduces blood pressure and improves health behaviors in patients treated for hypertension. The review also explores the intervention components combined with the behavioral self-monitoring interventions and estimates whether and to what extent they explain intervention clinical effectiveness.

Methods

Systematic Searches, Study Eligibility and Selection, and Data Coding

This systematic literature review involved searching the electronic databases MEDLINE via Ovid, Embase via Ovid, Web of Science, PsycINFO, Scopus, CINAHL, and the Cochrane Central Register of Controlled Trials (CENTRAL) in August 2021 to identify eligible studies. References for additional trials involved searches in 1 additional database: JMIR Publications [31].

The search strategy was developed using keywords from previous reviews and relevant literature (see an example of the search strategy in Multimedia Appendix 1). The review included randomized controlled trials testing intervention effects on behavior change and clinical effectiveness in people treated for hypertension. Studies on trials involving adults, published after the year 2000, and in the English language were considered for

inclusion. The review was preregistered on PROSPERO (CRD42019136158).

Screening of the title, abstract, and full text was conducted independently by 4 reviewers (MW, VM, SS, and RH), and disagreements were discussed by another reviewer (AK). Articles had to meet all of the following criteria to be eligible for full-text screening: (1) the population comprised adult individuals treated for hypertension; (2) the intervention consisted of self-monitoring of hypertension-related health behaviors via a mobile app; (3) the intervention aimed to support changes in both blood pressure and related health behaviors; (4) the comparator was usual care, enhanced usual care, or a minimal behavioral intervention; (5) the study included measurements of both blood pressure and health behaviors; and (6) the study design was a randomized controlled trial.

Outcome data were extracted for measurements of systolic and diastolic blood pressure (SBP and DSP, respectively), as well as health behaviors for medication adherence, physical activity, healthy diet, alcohol consumption, and smoking cessation. Outcome data for blood pressure and health behaviors were extracted for baseline and follow-up values for most of the studies; otherwise, only the follow-up values were extracted. When follow-up values were missing (eg, SD), the baseline values were selected to estimate intervention effects.

The Taxonomy of Behavior Change Techniques [32] was selected to conceptualize and guide the coding for the self-monitoring interventions. We also coded the intervention component "tailoring" for those interventions that delivered different messages to different participants based on information obtained about them [17,18], as well as the hypothesized mechanism of behavior change when these were reported. Authors of primary studies were contacted by email for missing information. Risk of bias was assessed using the Cochrane Risk of Bias tool, version 2, evaluating the risk introduced in the primary outcome of blood pressure [33,34]. Two reviewers independently coded study design and intervention components and extracted outcome data. Disagreement was discussed and resolved by a third reviewer.

Analysis

A random-effects meta-analysis was conducted to estimate the weighted, pooled effect for each of the blood pressure and behavioral outcomes to account for the true effect that may vary across the individual studies [35]. Effect sizes for continuous outcomes were calculated using the mean difference for blood pressure and the standardized mean difference (SMD) for behavioral outcome measurements. Mean difference was selected for blood pressure because measurements used similar units, whereas SMD was selected for behavioral outcomes because measurements were obtained using diverse methods, scales, units, or a composite of these. For example, for physical activity, these were minutes of physical activity per day and number of exercise sessions per week; for medication adherence, these were days of adherence per week, summary score responses to a 5-point scale, 8-item questionnaire. Effect sizes for dichotomous outcomes were calculated using the odds ratio (OR) for both blood pressure and behavioral outcomes [35,36]. In most cases, blood pressure outcomes were grouped based on an SBP threshold of 140 mmHg and a DBP threshold of 90 mmHg (exceeding the threshold indicates poorly controlled blood pressure whereas values below the threshold indicate controlled blood pressure) unless stratification variables were applied and reported (eg, age or gender-specific thresholds, clinic vs remote measurements thresholds, multimorbidity thresholds). Behavioral outcomes were grouped based on the corresponding guidelines for health behavior change adopted by individual studies.

Change-from-baseline outcomes were calculated unless baseline data were missing, in which case changes at follow-up were included in the analysis. The random-effects meta-analysis method was used to account for the distribution of the effect across studies [37].

The I^2 statistic was used to estimate the percentage of the variability in the effect estimates that is due to heterogeneity rather than chance [35]. Heterogeneity was explored further via subgroup analyses to investigate whether study-level variables could explain the observed heterogeneity.

Frequencies were used to summarize the behavioral strategies coded for each of the intervention and comparator groups [38]. Intervention strategies coded more than 3 times (frequency above 3) were considered for inclusion in the analysis. Subgroup analyses were performed to test for quantitative interactions, that is, whether intervention behavioral strategies could explain variation in the effect size.

Publication bias was examined by visual inspection of funnel plots and the Egger test. The meta-analysis was conducted using RevMan (version 5.4; The Cochrane Collaboration) [39].

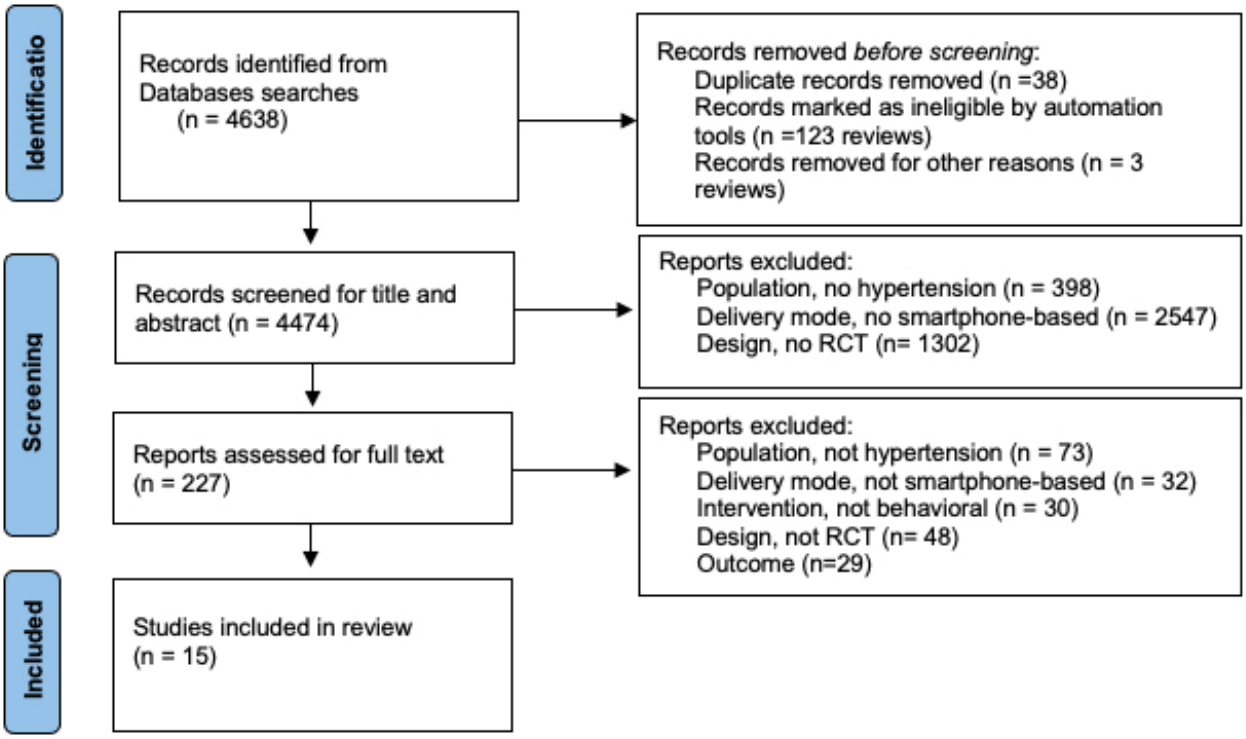
Results

Overview

The systematic search of the 7 databases identified 4638 articles, of which 227 were included for full-text screening. One additional trial was identified from another source. A total of 15 randomized controlled trials with 7415 participants met all the eligibility criteria and were included in the analysis (Figure 1).

The majority of the included trials were conducted in the United States [40-45], whereas 2 studies were conducted in Australia [46,47], and 1 study in each of the following countries: Canada [48], China [49], New Zealand [50], Ghana [51], India [52], China and India [53], and Norway [54]. Participants (adults aged >18 years) were recruited from primary and secondary health care settings (Multimedia Appendix 2).

Figure 1. The PRISMA flowchart.



Meta-analysis

Blood Pressure

The meta-analysis suggested that behavioral self-monitoring interventions via smartphone apps have a small but significant effect on reducing SPB by an average of 1.64 mmHg (95% CI 2.73-0.55, n=7301; Figure 2) across the studies among those in the intervention group compared to those in the control group. A similar but insignificant effect was found among studies measuring changes in SBP based on recommended thresholds; participants receiving the intervention were, on average, 60% more likely to achieve recommended levels of SBP (eg, SBP below 140 mmHg for measurements obtained in clinic) compared to those in the control group (OR 1.60, 95% CI 0.74-3.42, n=114; Figure 3).

A similar direction of effect, though not significant, was found for the impact of the app-based behavioral self-monitoring interventions in changing DBP. The interventions had a small effect in changing DBP by an average of 0.39 mmHg (95% CI –2.01 to 1.23, n=1368; Multimedia Appendix 3) compared to the control. The effect of the intervention in supporting reductions in DBP (eg, DBP below 90 mmHg) was, on average, 41% more likely in the intervention than in the control (OR 1.41, 95% CI 0.66-3.01, n=114; Multimedia Appendix 4), though the changes were not different between the 2 groups.

Heterogeneity between studies was low for most blood pressure outcome measurements (SBP, continuous: I^2=29%, T^2=0.87, P=.15; SBP, dichotomous: I^2=0%, T^2=0, P=.58; DBP, continuous: I^2=53%, T^2=2.56, P=.04; DBP, dichotomous: I^2=0%, T^2=0, P=.54), suggesting that there is potentially small, unimportant variation in the effect beyond chance between the 2 studies that operationalized the blood pressure outcome using categorical thresholds.

Figure 2. Meta-analysis of continuous outcome measurements for systolic blood pressure.

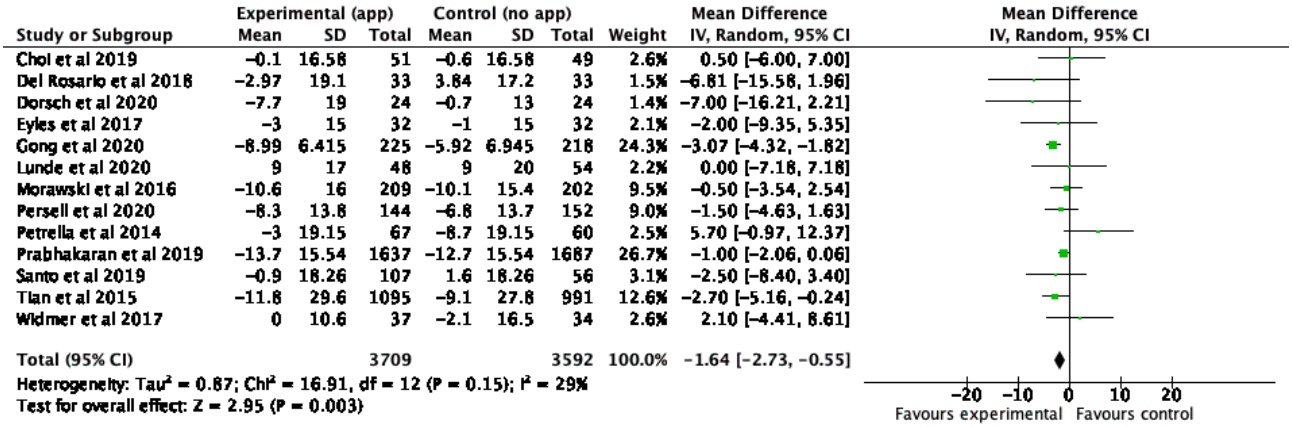
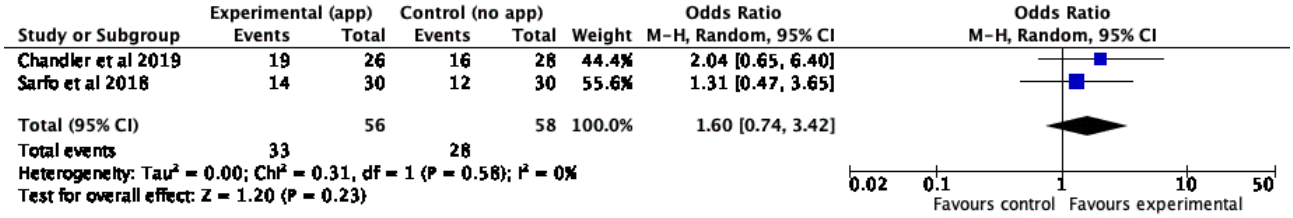


Figure 3. Meta-analysis of dichotomous outcome measurements for systolic blood pressure.



Medication Adherence

The SMD between the intervention and control groups was medium to large (SMD 0.78, 95% CI 0.22-1.34, n=688; Multimedia Appendix 5), suggesting that app-based behavioral self-monitoring is significantly more effective at supporting improvements in medication adherence behavior compared to the control. A similar direction of effect was found for the subsample of studies that used categorical operationalization for the intervention effect and suggested those receiving the app-based behavioral self-monitoring intervention are, on average, 3.8 times more likely to achieve clinically meaningful medication adherence than those not receiving the intervention (OR 3.83, 95% CI 1.25-11.76, n=6428; Multimedia Appendix 6).

Physical Activity

The review found a moderate but insignificant effect of the app-based behavioral self-monitoring interventions in improving physical activity (SMD 1.63, 95% CI –0.35 to 0.87, n=501; Multimedia Appendix 7) although only 4 studies provided data for this, with one of the included studies suggesting that intervention group patients were 1.6 times more likely to adhere to the lifestyle change.

Diet

The meta-analysis included 4 studies on healthy diet and suggested a moderate effect of behavioral self-monitoring on changing dietary habits by reducing the consumption of high sodium food by an SMD of 0.44 (95% CI 0.08-0.79, n=382; Multimedia Appendix 8). Objective measures of urinalysis suggested a positive but insignificant trend of the behavioral intervention in reducing salt intake. Although promising, these results should be interpreted with caution due to the small number of studies and sample size contributing to the meta-analyses.

Smoking and Alcohol

One study found, on average, a 53% improvement in smoking cessation among those receiving an app-based self-monitoring intervention compared to those in the control group (OR 1.53, 95% CI 0.76-3.09, n=3698) [52]. Effects on alcohol consumption were very small and not significant.

Subgroup Analyses

The most frequent behavior change technique coded in app-based behavioral self-monitoring interventions was feedback on behavior (n=13). Many app-based interventions (n=8) prompted participants to obtain advice from a health care provider following the behavioral measurements, and some (n=6) provided tailored advice to address the underlying mechanisms of behavior change. Goal setting of behavior, information about health consequences, and generic information about hypertension were strategies each coded in a small number of interventions (n=4). The most frequent strategy coded across both the intervention and control groups was reporting blood pressure and feedback on blood pressure (Multimedia Appendix 9).

Subgroup analysis found that tailored interventions resulted in higher and significant changes in both SBP and DBP in comparison to nontailored interventions (SBP: –2.92 mmHg, 95% CI –3.94 to –1.90, n=3102 vs –0.72 mmHg, 95% CI –1.67 to 0.23, n=4199, χ^2=9.65, *P*=.002; DBP: –2.05 mmHg, 95% CI –3.10 to –1.01, n=968 vs 1.54 mmHg, 95% CI –0.53 to 3.61, n=400, χ^2=9.19, *P*=.002). The differences between the 2 conditions were statistically significant and clinically meaningful (Multimedia Appendices 10 and 11).

Further investigation of the data revealed no effect of preselected variables that could influence blood pressure outcome (eg, sample size, time of follow-up, blood pressure outcome measurement obtained at clinic or remotely) on the observed effect.

Risk of Bias

The risk-of-bias analyses suggested that studies were of low risk of bias. Inspection of funnel plots and the Egger test suggested a low risk of publication bias (Multimedia Appendix 12).

Discussion

Principal Findings

This systematic literature review and meta-analysis included 15 randomized controlled trials with 7415 participants and found that patients treated for hypertension receiving an app-based behavioral self-monitoring intervention reduced SBP by an average of 1.64 mmHg (95% CI 2.73-0.55) and were, on average, 60% more likely to reduce SBP to <140 mmHg and DBP to <90 mmHg compared to those in the control group. Further subgroup analysis suggested that behavioral self-monitoring interventions combined with tailored advice had a higher and potentially clinically meaningful effect on reducing both SBP (mean reduction of 2.92 mmHg) and DBP (mean reduction of 2.05 mmHg) [55,56].

This study found that app-based self-monitoring of behavior interventions increased the odds of achieving medication adherence by 3 folds in the intervention group compared to the control. The significant effect of the app-based behavioral self-monitoring interventions in supporting improvements in

both blood pressure and medication adherence provides us with confidence that such interventions could be effective solutions to support health behavior change and thus reduce blood pressure in patients treated for hypertension during blood pressure checks or similar clinical consultations.

The behavioral interventions indicated positive effects for supporting improvements in healthy diet by reducing the consumption of high sodium food, as well as positive trends in supporting physical activity, smoking cessation, and alcohol consumption. Although promising, a small number of studies contributed to these meta-analyses, and thus the results should be treated with caution.

Strengths and Limitations

This review has several strengths and limitations. It did not include gray literature or unpublished studies and was limited to searching a few publicly accessible databases only. Nevertheless, this review summarizes the currently available evidence and suggests that behavioral tailored self-monitoring interventions are effective in changing SBP and DBP by –2.92 mmHg and –2.05 mmHg on average, respectively, compared to usual care, enhanced usual care, or a minimal behavioral intervention.

A limitation of the included studies is the use of self-reported measurements for the behavioral outcomes, which are inherent to bias. This might have diminished the validity of the observed intervention effect on health behaviors. Future trials should employ valid methods of measurement to assess behavioral outcomes and thus inform recommendations for policy making and practice.

This review has evaluated randomized controlled trials that compared behavioral self-monitoring interventions with usual care, enhanced usual care, or minimal behavioral interventions. We have used an extensive search strategy and identified all publicly available evidence. We have adopted a rigorous approach to data extraction and intervention coding to generate the results and form recommendations for best practices and future intervention development.

Implications for Practice and Intervention Development

The included trials had a duration of 1 to 12 months; thus, the evidence for the sustained effects of the intervention remains uncertain. However, a limited number of studies with long-term measurements had positive trends toward blood pressure reduction. Considering the wide reach and low-cost use of mobile technologies, this evidence indicates the potential impact of behavioral interventions on overall hypertension-related morbidity and mortality.

The comparator group included usual care (eg, clinic blood pressure checks), enhanced usual care (eg, regular blood pressure checks and medication adjustments), or minimal generic lifestyle interventions (eg, lifestyle tips and advice), suggesting that tailored behavioral self-monitoring is an acceptable addition to usual care and has a small, though clinically meaningful, effect on reducing blood pressure beyond and above usual care clinical practice.

Many studies involved clinicians in signposting participants to the app-based behavioral intervention, which might have influenced participants' engagement with the intervention and their health care. Moreover, the most frequent strategies reported being used with the behavioral self-monitoring interventions were feedback on health behaviors and prompts to obtain advice from a health care provider following behavioral measurements that require further support and monitoring. Although none of these strategies individually explained clinical effectiveness, they could have a synergistic effect in supporting patients' engagement with self-monitoring processes and thus in generating the observed improvements in health behaviors and reductions in blood pressure.

However, due to the limited information reported by primary studies, this review could not provide comprehensive, theory-based evidence on the mechanism by which the self-monitoring intervention achieved the observed clinical effectiveness [23-26]. Only a small number of studies explicitly reported on the theoretical concepts that informed the health behavior change intervention. For example, Chandler et al [40] and Dorsch et al [42] reported that the intervention aimed to modify beliefs and attitudes to support self-regulation processes and bring about changes in health behaviors. However, there is no evidence on the effects of the interventions with regard to modifying these theoretical influences to achieve changes in health behaviors and blood pressure. It would be useful if future studies report on the theoretical underpinnings and use valid measurements of engagement with the intervention strategies, as well as the underpinnings of health behaviors, to facilitate the generation of rigorous and replicable evidence on the mechanisms by which self-monitoring of behavior via the use of digital interventions support health behavior change and clinical effectiveness [57].

Conclusion

This systematic literature review suggested that tailored behavioral self-monitoring of hypertension-related behaviors facilitated via smartphone apps is effective in reducing blood pressure by an average of 2 mmHg above and beyond usual care, enhanced usual care, or minimal behavioral interventions. Thus, clinical practice should recommend behavioral self-monitoring combined with tailored behavioral advice to achieve clinical effectiveness. Considering the wide use of smartphone apps and their potential to reach large numbers of people, app-based behavioral self-monitoring interventions combined with tailored behavioral advice could potentially be a cost-effective addition to usual care blood pressure consultations. However, due to the limited quality of the trials included in this review, future research with rigorous methods is required to determine the direct impact of such interventions on both health behavior change and blood pressure, as well as their indirect effects on service provision and hypertension-related morbidity and mortality.

Acknowledgments

This paper presents independent research. AK, SS, and VM were funded by the National Institute for Health and Care Research (NIHR) under the Programme Grants for Applied Research (grant PR-PG-0615-20013). MW and RH were funded by studentships. The views expressed are those of the authors and are not necessarily those of the NIHR or the Department of Health and Social Care.

Data Availability

All data included in this review are reported in the multimedia appendices.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy for MEDLINE.
[DOCX File , 18 KB - mhealth_v10i7e34767_app1.docx]

Multimedia Appendix 2
Study characteristics.
[DOCX File , 23 KB - mhealth_v10i7e34767_app2.docx]

Multimedia Appendix 3
Meta-analysis of continuous outcome measurements for diastolic blood pressure.
[DOCX File , 904 KB - mhealth_v10i7e34767_app3.docx]

Multimedia Appendix 4
Meta-analysis of dichotomous outcome measurements for diastolic blood pressure.
[DOCX File , 602 KB - mhealth_v10i7e34767_app4.docx]

Multimedia Appendix 5
Meta-analysis of continuous outcome measurements for medication adherence.
[DOCX File , 660 KB - mhealth_v10i7e34767_app5.docx]

Multimedia Appendix 6
Meta-analysis of dichotomous outcome measurements for medication adherence.
[DOCX File , 725 KB - mhealth_v10i7e34767_app6.docx]

Multimedia Appendix 7
Meta-analysis of continuous outcome measurements for physical activity.
[DOCX File , 660 KB - mhealth_v10i7e34767_app7.docx]

Multimedia Appendix 8
Meta-analysis of continuous outcome measurements for healthy diet (consumption of low sodium food).
[DOCX File , 669 KB - mhealth_v10i7e34767_app8.docx]

Multimedia Appendix 9
Intervention coding of experimental and comparator groups.
[DOCX File , 33 KB - mhealth_v10i7e34767_app9.docx]

Multimedia Appendix 10
Subgroup analysis for the systolic blood pressure continuous outcome.
[DOCX File , 2091 KB - mhealth_v10i7e34767_app10.docx]

Multimedia Appendix 11
Subgroup analysis for the diastolic blood pressure dichotomous outcome.

[DOCX File , 1457 KB - mhealth_v10i7e34767_app11.docx]

Multimedia Appendix 12
Funnel plot of the systolic blood pressure continuous outcome.
[DOCX File , 957 KB - mhealth_v10i7e34767_app12.docx]

Multimedia Appendix 13
The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File (Adobe PDF File), 110 KB - mhealth_v10i7e34767_app13.pdf]

References

1. Zhou B, Carrillo-Larco RM, Danaei G, Riley LM, Paciorek CJ, Stevens GA, et al. . [doi: 10.1016/S0140-6736(21)01330-1]
2. Hypertension. World Health Organization. 2021. URL: https://www.who.int/health-topics/hypertension#tab=tab_1 [accessed 2021-08-23]
3. Mazzaglia G, Ambrosioni E, Alacqua M, Filippi A, Sessa E, Immordino V, et al. Adherence to antihypertensive medications and cardiovascular morbidity among newly diagnosed hypertensive patients. Circulation 2009 Oct 20;120(16):1598-1605. [doi: 10.1161/CIRCULATIONAHA.108.830299] [Medline: 19805653]
4. York Health Economics Consortium, School of Pharmacy, University of London. Evaluation of the Scale, Causes and Costs of Waste Medicines: Final Report. 2010 Nov. URL: https://discovery.ucl.ac.uk/id/eprint/1350234/1/Evaluation_of_NHS_Medicines_Waste__web_publication_version.pdf [accessed 2022-06-07]
5. Schroeder K, Fahey T, Ebrahim S. Interventions for improving adherence to treatment in patients with high blood pressure in ambulatory settings. Cochrane Database Syst Rev 2004(3). [doi: 10.1002/14651858.cd004804]
6. Diaz KM, Shimbo D. Physical activity and the prevention of hypertension. Curr Hypertens Rep 2013 Dec 20;15(6):659-668 [FREE Full text] [doi: 10.1007/s11906-013-0386-8] [Medline: 24052212]
7. Huai P, Xun H, Reilly KH, Wang Y, Ma W, Xi B. Physical Activity and Risk of Hypertension. Hypertension 2013 Dec;62(6):1021-1026. [doi: 10.1161/hypertensionaha.113.01965]
8. He FJ, Li J, Macgregor GA. Effect of longer term modest salt reduction on blood pressure: Cochrane systematic review and meta-analysis of randomised trials. BMJ 2013 Apr 03;346:f1325-f1325 [FREE Full text] [doi: 10.1136/bmj.f1325] [Medline: 23558162]
9. Strazzullo P, D'Elia L, Kandala N, Cappuccio FP. Salt intake, stroke, and cardiovascular disease: meta-analysis of prospective studies. BMJ 2009 Nov 24;339(nov24 1):b4567-b4567 [FREE Full text] [doi: 10.1136/bmj.b4567] [Medline: 19934192]
10. John J, Ziebland S, Yudkin P, Roe L, Neil H. Effects of fruit and vegetable consumption on plasma antioxidant concentrations and blood pressure: a randomised controlled trial. The Lancet 2002 Jun;359(9322):1969-1974. [doi: 10.1016/s0140-6736(02)98858-6]
11. Roerecke M, Kaczorowski J, Tobe SW, Gmel G, Hasan OSM, Rehm J. The effect of a reduction in alcohol consumption on blood pressure: a systematic review and meta-analysis. The Lancet Public Health 2017 Feb;2(2):e108-e120. [doi: 10.1016/s2468-2667(17)30003-8]
12. Huxley RR, Woodward M. Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies. The Lancet 2011 Oct;378(9799):1297-1305. [doi: 10.1016/s0140-6736(11)60781-2]
13. Quality and Outcomes Framework, 2020-21. NHS Digital. 2021 Sep 30. URL: https://digital.nhs.uk/data-and-information/publications/statistical/quality-and-outcomes-framework-achievement-prevalence-and-exceptions-data/2020-21 [accessed 2022-06-07]
14. Nieuwlaat R, Wilczynski N, Navarro T, Hobson N, Jeffery R, Keepanasseril A, et al. Interventions for enhancing medication adherence. Cochrane Database Syst Rev 2014 Nov 20(11):CD000011 [FREE Full text] [doi: 10.1002/14651858.CD000011.pub4] [Medline: 25412402]
15. Acin M, Rueda J, Saiz L, Parent MV, Alzueta N, Solà I, et al. Alcohol intake reduction for controlling hypertension. Cochrane Database Syst Rev 2020(9). [doi: 10.1002/14651858.cd010022.pub2]
16. Lee L, Mulvaney C, Wong Y, Chan E, Watson M, Lin H. Walking for hypertension. Cochrane Database Syst Rev 2021(2). [doi: 10.1002/14651858.cd008823.pub2]
17. Kassavou A, Sutton S. Automated telecommunication interventions to promote adherence to cardio-metabolic medications: meta-analysis of effectiveness and meta-regression of behaviour change techniques. Health Psychol Rev 2018 Mar 12;12(1):25-42. [doi: 10.1080/17437199.2017.1365617] [Medline: 28805162]
18. Armitage LC, Kassavou A, Sutton S. Do mobile device apps designed to support medication adherence demonstrate efficacy? A systematic review of randomised controlled trials, with meta-analysis. BMJ Open 2020 Jan 30;10(1):e032045 [FREE Full text] [doi: 10.1136/bmjopen-2019-032045] [Medline: 32005778]

19. Berry R, Kassavou A, Sutton S. Does self-monitoring diet and physical activity behaviors using digital technology support adults with obesity or overweight to lose weight? A systematic literature review with meta-analysis. Obes Rev 2021 Oct 30;22(10):e13306. [doi: 10.1111/obr.13306] [Medline: 34192411]
20. The text message is 20 years old today. Ofcom. 2012 Dec 03. URL: https://www.ofcom.org.uk/about-ofcom/latest/media/media-releases/2012/the-text-message-is-20-years-old-today [accessed 2021-08-22]
21. Medicines adherence: involving patients in decisions about prescribed medicines and supporting adherence. National Institute for Health and Care Excellence. 2009 Jan 28. URL: https://www.nice.org.uk/guidance/cg76 [accessed 2022-06-06]
22. Global strategy on digital health 2020-2025. World Health Organization. 2021. URL: https://apps.who.int/iris/bitstream/handle/10665/344249/9789240020924-eng.pdf [accessed 2021-08-23]
23. Spaulding EM, Marvel FA, Piasecki RJ, Martin SS, Allen JK. User Engagement With Smartphone Apps and Cardiovascular Disease Risk Factor Outcomes: Systematic Review. JMIR Cardio 2021 Feb 03;5(1):e18834 [FREE Full text] [doi: 10.2196/18834] [Medline: 33533730]
24. French D, Sutton S. Reactivity of measurement in health psychology: how much of a problem is it? What can be done about it? Br J Health Psy. 2010 2010 Dec 24:15-468. [doi: 10.1348/135910710x492341]
25. McCambridge J, Kypri K. Can simply answering research questions change behaviour? Systematic review and meta analyses of brief alcohol intervention trials. PLoS One 2011 Oct 5;6(10):e23748 [FREE Full text] [doi: 10.1371/journal.pone.0023748] [Medline: 21998626]
26. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. J Clin Epidemiol 2014 Mar;67(3):267-277 [FREE Full text] [doi: 10.1016/j.jclinepi.2013.08.015] [Medline: 24275499]
27. Haase J, Farris KB, Dorsch MP. Mobile Applications to Improve Medication Adherence. Telemed J E Health 2017 Feb;23(2):75-79. [doi: 10.1089/tmj.2015.0227] [Medline: 27248315]
28. Park JYE, Li J, Howren A, Tsao NW, De Vera M. Mobile Phone Apps Targeting Medication Adherence: Quality Assessment and Content Analysis of User Reviews. JMIR Mhealth Uhealth 2019 Jan 31;7(1):e11919 [FREE Full text] [doi: 10.2196/11919] [Medline: 30702435]
29. Alessa T, Hawley MS, Hock ES, de Witte L. Smartphone Apps to Support Self-Management of Hypertension: Review and Content Analysis. JMIR Mhealth Uhealth 2019 May 28;7(5):e13645 [FREE Full text] [doi: 10.2196/13645] [Medline: 31140434]
30. Xu H, Long H. The Effect of Smartphone App-Based Interventions for Patients With Hypertension: Systematic Review and Meta-Analysis. JMIR Mhealth Uhealth 2020 Oct 19;8(10):e21759 [FREE Full text] [doi: 10.2196/21759] [Medline: 33074161]
31. JMIR Publications. URL: http://www.jmir.org/search [accessed 2022-06-07]
32. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
33. Higgins J, Savović J, Page M, Elbers R, Sterne J. Chapter 8: Assessing risk of bias in a randomized trial. Cochrane Handbook for Systematic Reviews of Interventions, version 6.3 (updated February 2022). 2022. URL: https://training.cochrane.org/handbook/current/chapter-08 [accessed 2020-06-22]
34. RoB 2: A revised Cochrane risk-of-bias tool for randomized trials. Cochrane Methods Bias. URL: https://methods.cochrane.org/bias/resources/rob-2-revised-cochrane-risk-bias-tool-randomized-trials [accessed 2020-06-16]
35. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 06;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
36. Higgins J, Li T, Deeks J. Chapter 6: Choosing effect measures and computing estimates of effect. Cochrane Handbook for Systematic Reviews of Interventions, version 6.3 (updated February 2022). 2022. URL: https://training.cochrane.org/handbook/current/chapter-06 [accessed 2020-06-14]
37. Deeks J, Higgins J, Douglas A. Chapter 10: Analysing data and undertaking meta-analyses. Cochrane Handbook for Systematic Reviews of Interventions, version 6.3 (updated February 2022). 2022. URL: https://training.cochrane.org/handbook/current/chapter-10 [accessed 2022-06-07]
38. Michie S, Prestwich A, de Bruin M. Importance of the nature of comparison conditions for testing theory-based interventions: reply. Health Psychol 2010 Sep;29(5):468-470. [doi: 10.1037/a0020844] [Medline: 20836600]
39. Cochrane RevMan. Cochrane Training. URL: https://training.cochrane.org/online-learning/core-software-cochrane-reviews/revman [accessed 2021-11-14]
40. Chandler J, Sox L, Kellam K, Feder L, Nemeth L, Treiber F. Impact of a Culturally Tailored mHealth Medication Regimen Self-Management Program upon Blood Pressure among Hypertensive Hispanic Adults. Int J Environ Res Public Health 2019 Apr 06;16(7) [FREE Full text] [doi: 10.3390/ijerph16071226] [Medline: 30959858]
41. Choi BG, Dhawan T, Metzger K, Marshall L, Akbar A, Jain T, et al. Image-Based Mobile System for Dietary Management in an American Cardiology Population: Pilot Randomized Controlled Trial to Assess the Efficacy of Dietary Coaching Delivered via a Smartphone App Versus Traditional Counseling. JMIR Mhealth Uhealth 2019 Apr 23;7(4):e10755 [FREE Full text] [doi: 10.2196/10755] [Medline: 31012860]

42. Dorsch MP, Cornellier ML, Poggi AD, Bilgen F, Chen P, Wu C, et al. Effects of a Novel Contextual Just-In-Time Mobile App Intervention (LowSalt4Life) on Sodium Intake in Adults With Hypertension: Pilot Randomized Controlled Trial. JMIR Mhealth Uhealth 2020 Aug 10;8(8):e16696 [FREE Full text] [doi: 10.2196/16696] [Medline: 32663139]
43. Morawski K, Ghazinouri R, Krumme A, Lauffenburger JC, Lu Z, Durfee E, et al. Association of a Smartphone Application With Medication Adherence and Blood Pressure Control: The MedISAFE-BP Randomized Clinical Trial. JAMA Intern Med 2018 Jun 01;178(6):802-809 [FREE Full text] [doi: 10.1001/jamainternmed.2018.0447] [Medline: 29710289]
44. Persell SD, Peprah YA, Lipiszko D, Lee JY, Li JJ, Ciolino JD, et al. Effect of Home Blood Pressure Monitoring via a Smartphone Hypertension Coaching Application or Tracking Application on Adults With Uncontrolled Hypertension: A Randomized Clinical Trial. JAMA Netw Open 2020 Mar 02;3(3):e200255 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.0255] [Medline: 32119093]
45. Widmer RJ, Allison TG, Lennon R, Lopez-Jimenez F, Lerman LO, Lerman A. Digital health intervention during cardiac rehabilitation: A randomized controlled trial. Am Heart J 2017 Jun;188:65-72. [doi: 10.1016/j.ahj.2017.02.016] [Medline: 28577682]
46. Del Rosario MB, Lovell NH, Fildes J, Holgate K, Yu J, Ferry C, et al. Evaluation of an mHealth-Based Adjunct to Outpatient Cardiac Rehabilitation. IEEE J Biomed Health Inform 2018 Nov;22(6):1938-1948. [doi: 10.1109/jbhi.2017.2782209]
47. Santo K, Singleton A, Rogers K, Thiagalingam A, Chalmers J, Chow CK, et al. Medication reminder applications to improve adherence in coronary heart disease: a randomised clinical trial. Heart 2019 Feb 27;105(4):323-329. [doi: 10.1136/heartjnl-2018-313479] [Medline: 30150326]
48. Petrella RJ, Stuckey MI, Shapiro S, Gill DP. Mobile health, exercise and metabolic risk: a randomized controlled trial. BMC Public Health 2014 Oct 18;14:1082 [FREE Full text] [doi: 10.1186/1471-2458-14-1082] [Medline: 25326074]
49. Gong K, Yan Y, Li Y, Du J, Wang J, Han Y, et al. Mobile health applications for the management of primary hypertension: A multicenter, randomized, controlled trial. Medicine (Baltimore) 2020 Apr;99(16):e19715 [FREE Full text] [doi: 10.1097/MD.0000000000019715] [Medline: 32311957]
50. Eyles H, McLean R, Neal B, Jiang Y, Doughty RN, McLean R, et al. A salt-reduction smartphone app supports lower-salt food purchases for people with cardiovascular disease: Findings from the SaltSwitch randomised controlled trial. Eur J Prev Cardiol 2017 Sep;24(13):1435-1444. [doi: 10.1177/2047487317715713] [Medline: 28631933]
51. Sarfo FS, Treiber F, Gebregziabher M, Adamu S, Nichols M, Singh A, PINGS Team. Phone-based intervention for blood pressure control among Ghanaian stroke survivors: A pilot randomized controlled trial. Int J Stroke 2019 Aug 22;14(6):630-638. [doi: 10.1177/1747493018816423] [Medline: 30465630]
52. Prabhakaran D, Jha D, Prieto-Merino D, Roy A, Singh K, Ajay VS, Members of the Research Steering Committee, Investigators, Members of the Data SafetyMonitoring Board. Effectiveness of an mHealth-Based Electronic Decision Support System for Integrated Management of Chronic Conditions in Primary Care: The mWellcare Cluster-Randomized Controlled Trial. Circulation 2019 Jan 15;139(3):380-391 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.118.038192] [Medline: 30586732]
53. Tian M, Ajay VS, Dunzhu D, Hameed SS, Li X, Liu Z, et al. A Cluster-Randomized, Controlled Trial of a Simplified Multifaceted Management Program for Individuals at High Cardiovascular Risk (SimCard Trial) in Rural Tibet, China, and Haryana, India. Circulation 2015 Sep 01;132(9):815-824 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.015373] [Medline: 26187183]
54. Lunde P, Bye A, Bergland A, Grimsmo J, Jarstad E, Nilsson BB. Long-term follow-up with a smartphone application improves exercise capacity post cardiac rehabilitation: A randomized controlled trial. Eur J Prev Cardiol 2020 Nov 28;27(16):1782-1792 [FREE Full text] [doi: 10.1177/2047487320905717] [Medline: 32106713]
55. Lawes CM, Hoorn SV, Rodgers A. Global burden of blood-pressure-related disease, 2001. The Lancet 2008 May;371(9623):1513-1518. [doi: 10.1016/s0140-6736(08)60655-8]
56. Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. The Lancet 2002 Dec;360(9349):1903-1913. [doi: 10.1016/s0140-6736(02)11911-8]
57. Kassavou A, Court CA, Mirzaei V, Brimicombe J, Edwards S, Sutton S. Process Evaluation of MAPS: A Highly Tailored Digital Intervention to Support Medication Adherence in Primary Care Setting. Front Public Health 2021 Dec 20;9:806168 [FREE Full text] [doi: 10.3389/fpubh.2021.806168] [Medline: 34988058]

Abbreviations

CENTRAL: Cochrane Central Register of Controlled Trials
DBP: diastolic blood pressure
NIHR: National Institute for Health and Care Research
OR: odds ratio
SBP: systolic blood pressure
SMD: standardized mean difference

Edited by L Buis; submitted 08.11.21; peer-reviewed by C Basavarajappa, H Ayatollahi, A Joseph; comments to author 29.12.21; revised version received 05.02.22; accepted 02.05.22; published 12.07.22.

Please cite as:
Kassavou A, Wang M, Mirzaei V, Shpendi S, Hasan R
The Association Between Smartphone App–Based Self-monitoring of Hypertension-Related Behaviors and Reductions in High Blood Pressure: Systematic Review and Meta-analysis
JMIR Mhealth Uhealth 2022;10(7):e34767
URL: https://mhealth.jmir.org/2022/7/e34767
doi:10.2196/34767
PMID:35819830

Review

Wearing the Future—Wearables to Empower Users to Take Greater Responsibility for Their Health and Care: Scoping Review

Harjeevan Singh Kang[1,2], MD; Mark Exworthy[2], PhD

[1]College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom

[2]Health Services Management Centre, University of Birmingham, Birmingham, United Kingdom

Corresponding Author:
Harjeevan Singh Kang, MD
College of Medical and Dental Sciences
University of Birmingham
Edgbaston
Birmingham, B15 2TT
United Kingdom
Phone: 44 121 414 3344
Email: harjeevankangmedicine@gmail.com

Abstract

Background: Wearables refer to devices that are worn by individuals. In the health care field, wearables may assist with individual monitoring and diagnosis. In fact, the potential for wearable technology to assist with health care has received recognition from health systems around the world, including a place in the strategic Long Term Plan shared by the National Health Service in England. However, wearables are not limited to specialist medical devices used by patients. Leading technology companies, including Apple, have been exploring the capabilities of wearable health technology for health-conscious consumers. Despite advancements in wearable health technology, research is yet to be conducted on wearables and empowerment.

Objective: This study aimed to identify, summarize, and synthesize knowledge on how wearable health technology can empower individuals to take greater responsibility for their health and care.

Methods: This study was a scoping review with thematic analysis and narrative synthesis. Relevant guidance, such as the Arksey and O'Malley framework, was followed. In addition to searching gray literature, we searched MEDLINE, EMBASE, PsycINFO, HMIC, and Cochrane Library. Studies were included based on the following selection criteria: publication in English, publication in Europe or the United States, focus on wearables, relevance to the research, and the availability of the full text.

Results: After identifying 1585 unique records and excluding papers based on the selection criteria, 20 studies were included in the review. On analysis of these 20 studies, 3 main themes emerged: the potential barriers to using wearables, the role of providers and the benefits to providers from promoting the use of wearables, and how wearables can drive behavior change.

Conclusions: Considerable literature findings suggest that wearables can empower individuals by assisting with diagnosis, behavior change, and self-monitoring. However, greater adoption of wearables and engagement with wearable devices depend on various factors, including promotion and support from providers to encourage uptake; increased short-term investment to upskill staff, especially in the area of data analysis; and overcoming the barriers to use, particularly by improving device accuracy. Acting on these suggestions will require investment and constructive input from key stakeholders, namely users, health care professionals, and designers of the technology. As advancements in technology to make wearables viable health care devices have only come about recently, further studies will be important for measuring the effectiveness of wearables in empowering individuals. The investigation of user outcomes through large-scale studies would also be beneficial. Nevertheless, a significant challenge will be in the publication of research to keep pace with rapid developments related to wearable health technology.

(JMIR Mhealth Uhealth 2022;10(7):e35684) doi:10.2196/35684

KEYWORDS

wearable; device; tracker; activity tracker; fitness tracker; technology; MedTech; HealthTech; sensor; monitor; gadget; smartwatch; empowerment; self-care; management; behavior; responsibility; attitude; personalization; mobile phone; self-management; smartphone; wearable electronic devices; health promotion; health behavior; mHealth; digital health; health care wearables; scoping review

Introduction

Background

Wearable Health Technology

Wearables are "seamlessly embedded portable computers...worn on the body" [1]. Examples include consumer products marketed as wellness gadgets, such as smartwatches produced by Apple [2] or activity trackers from Fitbit [3], and more specialized medical devices, such as those that can detect electrolyte levels [4] or screen blood for cancer cells [5].

Wearable devices can be used in the medical field to monitor individuals and assist with diagnosis, thereby enabling individuals to contribute to their health [6] and gain greater control of their lives [7]. For example, certain wearables have been developed to recognize the symptoms of COVID-19 infection by measuring individuals' vital signs [8].

As technology advances, it may be expected that wearables will become more advanced in their health care capabilities. A future vision for wearables has been discussed [9], concerning the potential application of on-teeth sensors, smart contact lenses, electronic epidermal tattoos, smart patches, and smart textiles. Any data from wearables may be integrated with health systems and potentially inform care plans.

Empowerment

Patient empowerment has been well discussed in the literature, but the complexity of the concept is thought to be responsible for the "lack of a consensus definition" [10]. The most commonly cited definitions [11,12] indicate that "Patient empowerment starts from the principle of one's inherent capacity to be responsible for one's own life, and can be described as a complex experience of personal change, possibly facilitated by health care providers" [10]. Other researchers have proposed that patient empowerment encompasses activities that foster self-management [13].

Participatory health informatics (PHI) considers the role of technology in assisting individuals with self-management and decision-making by also improving health literacy and the physician-patient relationship so that individuals can become more involved in the aspects of their health and care [14]. Historically, research in the PHI field has predominantly been based on social media and internet-based applications, with patient empowerment having been identified as the most common theme in this body of research [14]. However, wearables are just beginning to be considered as part of PHI given recent technological advancements [14]. Therefore, similar research is now required to examine whether wearables can empower individuals in ways similar to those mentioned earlier regarding domains such as self-management, decision-making, and the physician-patient relationship.

There are several ways in which wearables may assist in empowering patients. First, wearables may minimize the impact of health care on the daily routine of patients. Wearables may offer greater convenience [15] if they reduce the need for patients to invest time in booking appointments with health care professionals, plan their schedule around such appointments, or commit time and money for appointment-related travel. Wearables have already been shown to reduce the need for certain in-person appointments [16].

Next, wearables collecting data throughout the day may provide a richer data set [17] than snapshot reading records obtained during visits to a health care facility. Such data may be collected more readily around individuals' normal daily activities, whether at rest or on exertion [18], which may be useful for heart rate readings, for example.

Furthermore, patients can take an electrocardiogram (ECG) and other readings multiple times each day over the course of months. This would add to the richness of the data set and potentially better inform diagnosis and treatment while also proving valuable in screening for COVID-19 infection, as Apple Watch could regularly monitor blood oxygen levels [19]. Attending appointments for taking such readings would neither allow the degree of frequency nor convenience of doing so at home and while on the move as with wearables.

Moreover, wearables may help preserve patient dignity when offering an alternative to more privacy-intrusive procedures. For example, an ECG taken by Apple Watch [19] may be preferred over a traditional ECG in a medical setting, which would require the removal of clothing to expose the patient's chest. Data from wearables may also flag early warning signs [2], prompting individuals to arrange appropriate medical consultations.

In addition, wearables may facilitate behavior change and potentially motivate patients to exercise, whether through daily step challenges, goal setting, or otherwise [20]. This could deliver associated health benefits [21] and help combat the obesity epidemic that faces health systems [22] and has been worsened by the COVID-19 pandemic [23].

Benefits for the Health System

The COVID-19 pandemic has exacerbated the pressure on the National Health Service (NHS) in England, as disruption to services has contributed to a backlog of care that is estimated to cost the NHS £2 billion (US $2.44 billion) to clear [24]. The NHS has been persistently overstretched, such that these additional pressures compound pre-existing problems of inadequate funding and understaffing [25]. As the NHS continues to face challenges, owing to resource constraints, care must be delivered more efficiently.

Innovative solutions are known to secure growth [26] by redefining care pathways [27] to improve patient satisfaction, teamwork, the provision of care, and clinical outcomes. In this way, wearables [28] can shift the burden of care from the NHS to the individual. Such a shift would represent greater convenience and independence for patients (as outlined earlier), while reducing costs and staff workloads. In fact, the NHS Long Term Plan has welcomed wearables from an efficiency standpoint [29], as the technology has the potential to revolutionize health care [28].

Remote patient monitoring, in the context of reducing the demand for health systems, has been of particular importance during the pandemic [30]. However, it should continue to retain

its relevance [31] by reducing patient consultations [32] because of the health care sector's focus on patient care and the versatility of wearables in catering to a wide spectrum of needs, from acting as a preventive tool in promoting fitness to managing chronic conditions [33].

Challenges Relating to Wearables

Although it has been stated that wearables can empower and emancipate patients [34] to manage their own care, the efficacy of these devices has attracted skepticism from some physicians [35], especially because the technology is emerging. However, change should be welcome, as patients are an "untapped resource" [7]. If patients were to take a more proactive role in their care, then the effects on the "quality and sustainability of health systems" could be transformative [7].

However, the accuracy of wearables is a concern that may deter their use, especially if they fail to produce reliable data. Therefore, regulatory oversight may be beneficial in ensuring that only accurate, tested devices are in circulation. Medical devices are regulated in the United Kingdom by the Medicines and Healthcare Products Regulatory Agency (MHRA) [36]. Nonetheless, certain wearables may not be regulated by the MHRA, as devices such as the Fitbit explicitly state that they are neither medical devices nor are "intended to diagnose, treat, cure or prevent any disease" [37]. Therefore, this may undermine the perceived efficacy of such devices and thereby fuel the skepticism of health care professionals. However, as wearables become more accurate, this is likely to change; some consumer-targeted wearables, such as Apple Watch, have already received Food and Drug Administration approval in the United States [38]. Consequently, it seems to be only a matter of time before approval is sought under the MHRA.

Furthermore, Accenture [39] advised that physicians should promote digital engagement and awareness of such devices among patients. This recommendation followed the findings that more than half of those surveyed [39] would take more responsibility for their care if their health care provider encouraged them to. However, only one-tenth of the respondents [39] reported having been recommended any digital tools to manage their care. It has been argued that despite initial reservations from patients, typically arising from a lack of confidence or knowledge, "it is incumbent on providers to foster [patients'] self-reliance" [7]. Clearly, with "self-management gaining ascendancy as a concept" [40,41], there is more to be done, including possibly reshaping the perceptions of providers and patients [7].

Objectives

This study aims to identify, summarize, and synthesize knowledge to answer the following research question: "How can wearable health technology empower individuals to take greater responsibility for their health and care?" To the researcher's knowledge, a review has yet to be conducted in this area; other reviews did not specifically focus their research on the concept of empowerment. Hence, research is needed to fill this gap and convey the importance of wearables to health care professionals.

Methods

Design

A scoping review design was chosen for its exploratory nature [42], which is useful when the international evidence base is heterogeneous [43]. In addition, this design enables the researcher to determine the range of available evidence and identify research gaps to guide future research [44].

Furthermore, the need to integrate research from a wide variety of sources and perspectives [43] across a broad area lends itself to a scoping review over alternative designs. A systematic review was found to be too restrictive and limited the materials considered [45,46], whereas research in the wearable field did not seem to place the same emphasis on theory as would be required for a realist review [47].

The 22-item PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [48] checklist was used, as it indicates what should be included in a scoping review. Background reading was conducted to ensure adherence to the latest guidelines. For example, there have been numerous additions [49-53] following the publication of a seminal paper by Arksey and O'Malley [44], which initially proposed a methodological framework for undertaking scoping reviews. The guidance document published by the Joanna Briggs Institute (JBI) [54] was also followed.

Selection Criteria

Selection criteria were set to ensure the coverage of evidence, while excluding irrelevant papers. Hence, the inclusion criteria were as follows: English-language articles, a focus on wearables rather than other digital health technologies, and relevance to the research objective (by offering information that may relate to empowerment, such as barriers to use or discussions of the efficacy of certain wearables, even if such information had not been explicitly linked to empowerment). The researcher was selective in only including sources where there was a substantive focus on wearables rather than those that only mentioned wearables in passing. Regarding the inclusion of literature reviews, the individual studies of the review were screened. If many of these met the inclusion criteria, the review was included instead of the individual studies.

Despite wearable technology being a fast-moving area, no limits were imposed on the publication year of articles. Studies were excluded if the full text was unavailable or the studies were published outside of Europe or the United States. The latter was determined after preliminary searches indicated the presence of sufficient evidence. At this point, it was necessary to refine the selection criteria during the literature search phase because of practical constraints. In fact, Arksey and O'Malley [44] encouraged an iterative approach to research by using broad searches to first gain a *sense* of the field and thereafter setting any search parameters more strictly to meet the research requirements. Such an approach is further supported by the fact that "reading is central to reviewing literature" [55] and informing literature searches.

Search Strategy

Database searches included MEDLINE, EMBASE, PsycINFO, HMIC, and Cochrane Library. Gray literature was also considered by searching OpenGrey, Google Scholar, and independent think tanks. The literature search was completed in early February 2021. A further search was conducted in May 2021 to account for any articles that may have been subsequently published.

The literature search involved relevant subject heading index terms, and subject headings were *exploded* as required. The search strategy (Multimedia Appendix 1) was adjusted to reflect variations in subject headings and syntax across the databases. For the breadth of coverage, a multipurpose search was used to search for keywords across numerous fields. A librarian was consulted to identify additional keywords.

Various strategies have been used to mitigate the risk of missing relevant evidence, including the use of synonymous terms, wildcard symbols, and truncation symbols. Boolean operators were used to combine the keywords and exclude others. Parentheses were used to group keywords joined by different Boolean operators, which yielded more relevant results than if a nesting approach had not been followed. In cases where quotation marks for phrase searching would potentially omit relevant results, proximity operators were used instead. The above-mentioned publication limits for language and location were also applied to the results. Furthermore, there was forward citation searching, and reference lists were snowballed for relevance to find studies that had not been identified in the initial literature search.

Duplicate records were identified using EndNote (Clarivate). The software-generated list of duplicates was manually reviewed to mitigate the risk of any records being incorrectly categorized as duplicates. The researcher then screened the remaining results and manually removed duplicates that had not been automatically flagged.

Data Collection

The single researcher screened the literature by using a 2-step process, with a review of the title and abstract before the full text. If neither the title nor the abstract seemed relevant to the research, the article was excluded. If the title and abstract appeared relevant, the full-text article was read. Papers that did not meet the inclusion criteria were excluded, and the main reasons for exclusion were noted.

Scoping reviews do not typically address the appraisal of sources [44]. However, this would have resulted in a much larger sample size of evidence of questionable quality. Therefore, the JBI critical appraisal tool was used because of its relatively greater sensitivity to validity [56] to help ensure that any emergent findings would be based on high-quality evidence. This involved considering the limitations of the evidence, while assessing the congruity between the research aims, methodology, and findings [57].

Data Charting

Key details were extracted to assess the relevance of a study [58], including publication details and study details relating to the objectives, findings, and type of wearable device. A data charting form (Multimedia Appendix 2) was adapted from the JBI [59] to incorporate other relevant details described elsewhere [60]. This form was piloted and updated with additional data that the researcher wished to chart.

Data Analysis

Oftentimes, reviews fail to go beyond a summary of the evidence. Hence, this research followed the 6-step process of thematic analysis by Braun and Clarke [61], which involved familiarization with the data, coding of the data, generation of themes based on the codes, refinement of the themes, naming and defining the themes, and final write-up.

NVivo (QSR International) was used for a more structured analysis, as each source was individually uploaded and coded, which enabled the identification of themes from a wide evidence base. Themes were refined, with the findings being presented in the style of a narrative synthesis and related to the research question.

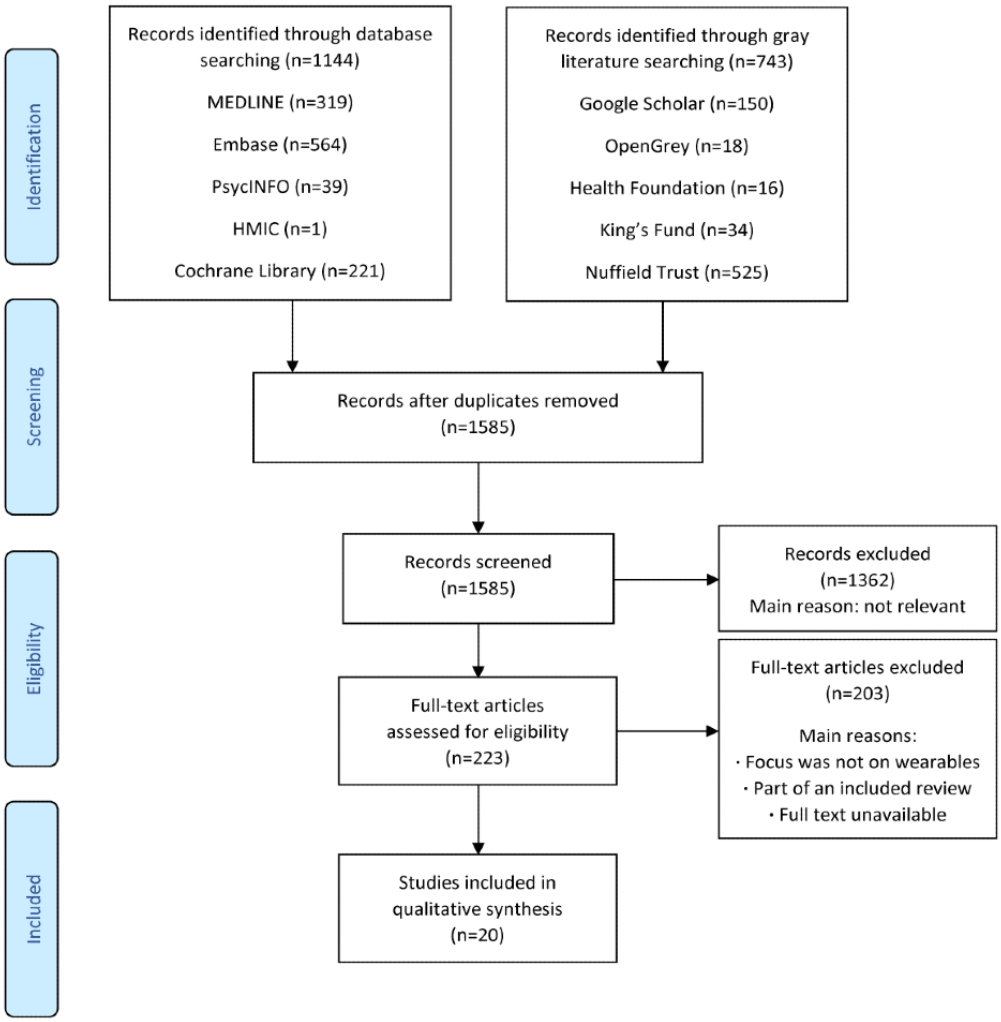
Such an approach to analysis and synthesis accords with guidance from Arksey and O'Malley [44], which stated the need for a scoping review to potentially use a "thematic construction in order to present a narrative account of existing literature." This has been reflected in the PRISMA-ScR [48] and guidance on advancing the methodology of scoping reviews [49]. There are also examples of scoping reviews incorporating such an approach to analysis [62-64].

Results

Literature Search

The search (Figure 1) identified 1887 records. Following screening, 20 studies were included in the final data set, as summarized in Multimedia Appendix 3 [65-84]. Some of these studies were identified for inclusion in gray literature searches [77,78,83] or snowballing the reference lists of the included studies [76,80].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram to illustrate the literature search.



Study Characteristics

The 20 included sources represent a significant body of literature, collectively accounting for >7000 participants. The studies were published between 2015 and 2021, with the number of studies appearing to have generally increased year-on-year. Most studies were published in the United States (12/20, 60%). The studies used quantitative (10/20, 50%), qualitative (8/20, 40%), and mixed methods (2/20, 10%; Multimedia Appendix 4). Funding details were provided by 65% (13/20) of the studies (Multimedia Appendix 5). Although Fitbit was the most common brand of wearable used (10/20, 50%), several studies (9/20, 45%) included multiple brands or discussed wearables in general.

Discussion

A total of 3 main themes, relevant to user empowerment, emerged from the literature, namely, *Health Care Providers—Benefits and Involvement*, *Behavior Change*, and *Barriers to Use*. Multimedia Appendix 6 [65-84] lists the contributions of the included studies to each theme.

Theme: Health Care Providers—Benefits and Involvement

Collaboration Between Providers and Patients

Health care providers are an important part of health care systems [85]. Therefore, it may be expected that providers would be considered as part of the literature on how wearables can empower patients.

Collaboration between research management and health care staff is imperative, especially during the study design process, as such a partnership may benefit patient compliance, particularly for those with cognitive impairments [65]. However, the role of clinicians may extend further. Outside of the research context, patients may rely on the clinicians' acceptance of their decision to use a wearable device for other purposes, including as part of rehabilitation; hence, it would be incorrect to limit the role of health care professionals to simply prescribing medication without considering their role in educating patients [66]. The significance of such support and backing from clinicians may be easily overlooked.

Users seem to appreciate that consumer wearables are not medically accurate devices and that clinicians would not solely rely on data from such devices to make clinical decisions [84]. An open-minded, supportive approach may encourage patients

to share data with their clinicians [84]. However, clinicians who are unwilling to engage with wearables and support their empowered patients, on the grounds of potential inaccuracies regarding data [84], may risk foregoing the benefits attributable to wearables.

Benefits to Providers and Patients

Wearables may offer several benefits to clinicians. First, wearables may offer objective, real-time patient data [66,76]. This would allow clinicians to remotely supervise progress [72,76] and provide comfort to patients who may otherwise feel that they are just communicating their subjective experiences and perceptions [76]. In such cases, it would be possible to use such data to inform clinicians of a patient's history, thereby enabling a more personalized approach to treatment tailored to individual needs that can be adjusted according to the management plan [66,76]. This should enable more timely feedback so that clinicians can be more responsive to situational changes [80]. Access to data such as nutrition and activity-related information over an extended period may offer a solution to the issues of conventional health measurements and tests, as clinicians would benefit from a more complete picture of a patient's health status [80]. In addition, data from wearables may eventually be used for risk stratification and early intervention [83], which should prevent further deterioration.

Furthermore, the accessibility of wearable data to patients may facilitate communication and assist with patient education [66]. Better-informed patients can offer more worthwhile contributions to any discussion, thereby promoting shared decision-making [66] and assisting with adherence to what is agreed [76,83]. In fact, a higher quality of life was associated with patients taking a more proactive role in their health [66]. There is the important caveat that to maximize these benefits, health care professionals should first identify patients with the willingness and ability to self-manage, especially because sustaining engagement can be challenging [83].

It is not difficult to imagine the potential for a large-scale rollout of wearables, which may help reduce the contact time and offer a more cost-effective approach to providers [75,76]. Such improvements in efficiency would likely free up resources, thereby alleviating the burden on health systems. The achievement of this is realistic, as supported by the Nuffield Trust [83], which has reported that "professional monitoring interventions for chronic conditions, whereby data is sent to the health care team, have had very positive results on health outcomes and resource use."

Data from wearables can also be integrated into medical records to facilitate care [70,80], which can help overcome current barriers to reporting and retrieving data for inpatients and remote monitoring [80]. Patients living with chronic conditions often feel undersupported in managing their conditions [83]; therefore, wearables may offer this support. This is largely why wearables and other patient-facing technologies have been praised as a "bright hope" in the health care sector horizon [83].

Challenges to Wearables Advocation in the Health Care Sector

Certain health care services do not have the best track record for the uptake of technology. For example, in the United Kingdom, the NHS has been portrayed as "one of the most backward industries in responding to digital technology" [83]. The Nuffield Trust has captured the fact that the NHS has the potential to capitalize on consumer wearables [83]. However, consumer wearables may not be suitable for use, in their current state, by health care professionals. In fact, poorly calibrated devices can work counterproductively by worsening health outcomes and increasing staff workload [83]. Nevertheless, care should be taken not to be overly critical about the lack of accuracy of certain wearables because of benefits associated with aspects such as the provision of insights over extended periods [17].

Staff may require further upskilling to encourage engagement with wearables and facilitate behavior change [83]. This may demand professional monitoring and the provision of feedback on an ongoing basis [83]. In the short term, this may impose greater pressure on staff as it will add to workers' responsibilities and may therefore appear unfeasible given the existing strain on staff. However, the short-term increase in workload may result in an overall reduction in workers' commitments over the long term because of benefits associated with self-measurement of readings and the consequential reduction in appointments for such purposes [86].

Ultimately, providers have much to gain from patients taking steps to monitor their own health. To realize these benefits, health care professionals should encourage patients by adopting a supportive attitude, recognizing that wearables offer a means for patients to take a more proactive role in managing their health rather than viewing the devices too critically. In fact, diffusion of innovations theory [87] classifies adopters into categories, ranging from those who easily embrace change to *laggards* who are more skeptical about the innovation. Applying this theory [87] to the adoption of wearables, providers can play an important role in seeking to convince laggards about the benefits of wearables.

Theme: Behavior Change

Overview

Breaking bad habits and establishing good ones, as part of a sustainable change to one's lifestyle, requires positive actions whereby attitudes or behaviors may need to shift. The potential for wearables to draw on various behavior change techniques to prompt positive behavior change [20] holds promise for individuals willing to take greater responsibility for their health and care. Behavior change through wearables can take many forms, from reminders and positive reinforcement associated with progress tracking and reporting to social group support for motivational purposes. However, such aspects, among others, can also give rise to negative outcomes if not carefully catered for, as discussed in the following sections.

Behavior Change Techniques and Support

Continually providing information to users through wearables may be useful for consolidating patients' understanding of their conditions and prompting behavior changes [66]. Furthermore, the ability of wearables to track progress and achievements could bolster adherence to exercise, which aligns with the behavior change theory [66]. It has also been suggested that introducing behavioral counseling based on feedback from wearables may lead to better results [70]. Another study has suggested the potential for activity trackers to complement behavioral counseling because of the behavior change techniques embedded in wearables, including those related to goal setting and social support [74]. These behavior change techniques have been leveraged by certain wearables [73,80] to help achieve positive changes, such as by promoting an active lifestyle [83]. Wearables seem to support behavior change, as another study has concluded that wearables further benefit patients in achieving their outcomes, as opposed to counseling alone [75].

Contrary to the position that has been taken in these studies, which have suggested that wearables can be effective, and the results for patients can be enhanced through the additional use of behavioral counseling, wearables' value as a positive behavior change strategy may be context dependent. This is supported by a study that found that activity tracking was insufficient for improving pain-related outcomes or daily activity without behavior change support [72]. Despite not tracking changes in variables linked to behavior change theories, it has been argued that wearables may not be effective from a behavior change standpoint when promoting physical activity in college students [73].

In one study, only a few participants recognized specific behavior changes arising from the use of wearables [71]. These participants were more disciplined and conscious about activity levels and which exercises were more effective [71]. Although only a few commented on any behavior changes, the subjective nature of these changes may mean that others made similar progress but did not recognize such progress. Another study stated that their effect size for behavioral outcomes ranged between small and medium but could not identify which aspects of the devices resulted in this finding; instead, they speculated that this was because of greater intrinsic motivation for exercise [74].

An analysis of behavior change techniques used by activity trackers suggested that wearables commonly have more *controlling* features than those that promote autonomy [69]. For some users, this focus on rewards or social comparison may only appear detrimental to their physical activity in the long term [69] and may not be reflected in the findings of relatively short studies.

Moreover, physical activity levels seem to affect users' perceptions of wearables, as those who are more active generally found the devices to have a higher number of *motivational affordances*, which refer to the features of technology that motivate and support users to meet their goals [79]. It has been suggested that this is because of greater familiarity with the motivational features of wearables, whereas novice exercisers may not understand or notice these features, such as the symbol denoting calorie burn [79]. Therefore, guided studies may not generalize to first-time, real-world use [79,80].

Self-efficacy

Self-efficacy refers to an individual's belief that they can perform a task [88]. The strength of self-efficacy is important in influencing behavior change and how the individual responds to adversity [88].

Wearables appeared to draw on 3 sources of self-efficacy proposed by Bandura [88]; these have been credited with increasing user compliance and positive behavior change [77]. The first source relates to personal accomplishments [88], which are encompassed by the various features of wearable devices, including awards, progress toward activity goals, and performance over time [77,82]. The use of activity reminders forms part of the second source of self-efficacy, related to verbal persuasion [88], as motivational notifications can encourage users to progress and meet their goals [77,84]. The third source is termed "vicarious experience" [88] and links to the social aspects of wearables, whereby seeing users of a similar ability complete activities motivate certain users to believe that they can execute the same tasks [77].

However, it may be detrimental to self-efficacy when users believe that they are significantly underachieving relative to their peers [77]. Therefore, individuals should be matched to fellow users with whom they identify and who are successfully achieving their goals, as otherwise they may be discouraged [84]. Of course, this must be balanced with the privacy implications associated with personal data use, as individuals must be provided with transparent information about how their data will be used, coupled with data minimization techniques to ensure that only data required for the particular objective are being used and shared [89]. Nevertheless, designers should continue to consider sources of self-efficacy when developing features for wearables [77].

Contextual Factors

Importantly, users' perceptions of self-efficacy seem context dependent [77]. The internal context comprises cognitive, behavioral, and emotional factors [77], whereas the external context considers factors outside the user's control, such as the weather or time of the day [77]. The internal context is particularly important for self-efficacy, as it can either neutralize or compound a negative external context, meaning that users will either persevere in the face of adversity or stop using the wearable device [77]. In the interest of long-term behavior change and compliance, users should be supported to develop positive internal processes. For example, it would be valuable for wearables to be capable of adjusting their feedback based on the momentary state of the user [84] to reinforce their successes while supporting them through any difficulties in meeting targets.

Wearables offer a safe environment, as users can try to meet their goals even after repeatedly falling short; this establishes the intrinsic motivation to stay committed [84]. However, the support offered by wearables may need to be individualized to reflect the uniqueness of users' personalities and priorities, which can factor into the affordances of wearables [80], as better

engagement may convert to positive steps for behavior change. In addition, it is believed to foster self-efficacy, thereby supporting self-management [84]. For example, less conscientious individuals may require additional motivational support to assist with goal setting [79]. In addition, because self-set targets may not aid motivation, it may be beneficial for wearables to suggest feasible goals after monitoring the user [84]. Less agreeable users may respond better to increased support for their autonomy or greater transparency to build trust in the technology [79]. Introverts may prefer greater privacy, whereas extroverts may be more receptive to social aspects, such as comparing activities with others [79].

Users comparing their own data against expected standards may prompt positive behavior change [82], as not meeting such standards may lead to discomfort, referred to as cognitive dissonance [80]. The companion app plays an important role in enabling users to process information, as it visualizes and contextualizes their data [82]; this positively affects self-reported health metrics [82]. Of course, the privacy implications, as discussed earlier, of identifying peer comparators with respect to expected standards must still be observed.

Incentivization

Economic incentives, such as offering discounts on insurance premiums or wellness products, also appear to increase the willingness of individuals to use wearables [81]. Such an approach, in terms of offering discounts, has been undertaken by health and life insurance providers who are motivated to minimize claims on their issued policies. For example, Vitality offers a discount on Apple Watch [90] and encourages members to track their activities via the app.

Consequently, incentivizing uptake may facilitate behavior change through regular use, but this would seem to be contingent on users' satisfaction with the data privacy and technical provisions of the wearable device. Therefore, it is important to address any barriers so that they do not hinder the use of wearables and prevent users from beginning the process of positive behavior change.

Motivational Profile

The subsequent discussion on barriers to use centers primarily on the design of the wearable, among other factors. However, there may be a case for considering the motivational profile (degree of autonomy and motivation) of users [69] and the motivational affordances of devices [79] when using wearables as a tool for empowerment, as is evident that there may be contextual factors that affect the ability of wearables to inspire behavior change.

For wearables to empower individuals, it would be worth undertaking a preliminary assessment of individuals who may require additional support in the form of behavioral counseling. This will help ensure that patients receive appropriate support, as individuals whose motivational profiles are not matched to the wearable device may become demotivated and experience negative emotions from persistently failing to meet goals [69].

Theme: Barriers to Use

Barriers to the adoption and use of wearables could have significant ramifications for empowerment.

Although individuals expressed willingness to use wearables, use seemed to be inconsistent; a study reported that >90% of the participants suspended use [65]. As this is not an isolated case, with the issue of compliance mentioned elsewhere [68], it is worth considering factors that may have contributed to this.

The barriers to use that were identified [65] include forgetting to apply, hospitalization, loss of interest, and temporary loss of the wearable device. Aside from the concerns of wearability, accuracy, and price, feelings of fatigue stemming from the use of technology highlight the need for wearables to constantly engage users, as loss of interest is a key reason for disuse [73]. It is perhaps surprising that losing wearables does not seem to be uncommon; this is evidenced by other studies [68,72,73], some of which have also reported malfunctioning devices that require replacement by the manufacturer [68].

Design-Related Aspects

In addition, although certain design aspects, such as color and size, may influence use [65], an aesthetically pleasing appearance may be a more important consideration for younger individuals [76].

Concerns regarding stigma arising from the use of certain wearables have also been raised. For instance, children who are overweight that wear the *badge* of an activity tracker may be bullied [91]. Similarly, this seems to factor into the decisions of patients who would prefer a sleek, discreet device rather than one that is overtly medical [76].

Technical Aspects

The technology itself may deter use. A study [67] has added the following to the list of potential barriers: health difficulties, technical difficulties, a lack of personalized advice, and an inability to track other types of physical activity such as strength exercises. These clearly represent barriers, as reported elsewhere [76]. Such concerns may also discourage regular use over a prolonged period [76], especially if individuals come to perceive that these issues are associated with all wearables.

Annoyances may also prevent users from engaging with the technology [71]. For example, users may be frustrated by the perceived inaccuracies of sleep or pulse monitors [71], as some have stopped using wearables for being unreliable [84]. Device inaccuracies have been cited elsewhere together with issues related to battery life [72].

Barriers That Are More Common for Older Users

In addition, a lack of familiarity [71] or not being tech-savvy [84] may mean that some individuals are put off by wearables that appear too complicated at first use. Such difficulties may be more common among the older generation [79], in the context of connecting wearables to smartphones and accessing metrics [71]. In fact, not owning a smartphone, through which many wearables tend to display such metrics, seemed to limit interest in tracking activities altogether [71].

Certain other barriers seem to apply to an older user base. Devices that require a high level of manual dexterity to operate proved unsuitable for older individuals to easily use [72,79]. Another complaint was that the displayed text was too small to read easily [79]. Furthermore, many users were frustrated by the lack of availability of instructions and guides for the execution of basic tasks. This may be more of an issue in research studies, as users typically have access to any device manual when they make a purchase themselves. However, technical issues are common and tend to be resolved by the staff leading the research study [72].

Cost

Cost may be another barrier, as even relatively low-cost trackers may be inaccessible to older adults [72]. For others, the cost is a nonissue, as it was suggested that if the device is beneficial, then it is a matter of answering the question, "What's my health worth to me?" [76]. This highlights the possible need for individuals to weigh the advantages offered by a wearable device against its shortcomings to ascertain whether the device is of value and justifies the investment in one's health.

Importantly, wearables should not seek to widen the health inequalities that have worsened during the COVID-19 pandemic [92], especially for the poorest in society who tend to be in the greatest need of care but least likely to receive such care [93]. Therefore, wearables should serve as an additional option for individuals to proactively manage their health care rather than acting as a replacement for any traditional mode of delivery.

Barriers Arising From Long-term Use

The nature of wearables, as a newly emerging technology that has gained traction in recent years, warrants further research and development [78] to allay concerns surrounding durability, comfort, power consumption, standardization, interoperability, accuracy, privacy, and confidentiality. These potential issues are more likely to arise from regular, long-term use of wearables; however, they are often missed in shorter clinical studies [78]. If the barriers and concerns that have been raised are deemed by users to outweigh the benefits offered by the wearable, then this may discourage individuals from using such devices to monitor their health, thereby potentially interfering with their ability to follow an active lifestyle [80].

Privacy

Moreover, privacy concerns have often been raised [84]. This is illustrated by the recent acquisition of Fitbit by Google [94], which gave rise to concerns about how personal and health data were going to be used by a *tech giant* that is active in the AdTech and data commercialization fields [95]. Consequently, it is necessary to balance privacy and security concerns with potential benefits to users and the health system [80].

Another high-profile example of significant privacy concerns from the use of portable technology in the context of health care has arisen from the development and use of COVID Track and Trace apps around the world [96-98]. Although this does not fall within the strict definition of a wearable, the privacy concerns raised [99] with respect to the apps with regard to location tracking of individuals and the sharing and aggregation of personal data are equally applicable to the use of wearables that capture and process such types of user data.

Technology-Specific or General Barriers

It must be acknowledged that some of the criticisms of wearables that seem to hinder use could be specific to the brand of wearables used in a study. Therefore, although the aforementioned concerns should be considered, it is important to distinguish the specific nature of some barriers rather than applying them to wearables in general. For example, the inability to measure strength exercises appears to be specific to the wearable used in a study as part of a review [67]. In reality, the availability of a range of wearables, some of which are designed to track strength exercises, may present less of a barrier to use.

However, the fact that the aforementioned barriers have been described in the literature seems to suggest that such issues are prevalent rather than being restricted to a single brand of wearable technology, as Multimedia Appendix 3 shows the diversity of wearables included in this review. In addition, the barriers are significant and clearly need to be overcome to avert any further negative effects on user perceptions, which may otherwise discourage the use of wearables. Failure to take appropriate steps for damage control may erode public trust in wearables, thereby limiting the potential to empower new users to manage their health more proactively. Therefore, although all technologies seem to have their own shortcomings or barriers, issues relating to wearable health technology may be viewed more critically, as such wearables can inform decisions related to one's health and care.

Principal Findings

A summary of the principal findings with respect to these themes is provided in the following sections.

Health Care Providers—Benefits and Involvement

Providers play an important role in empowering patients to use wearables. Therefore, providers require support because of the short-term resource constraints that they are likely to face. However, data from wearables may help create a more holistic understanding of a patient's health status, thereby accelerating the delivery of personalized advice. Better-informed patients should aid in communication and improve their adherence to advice.

Behavior Change

Wearables may lead to positive behavior changes. This may arise from the ability to set goals, receive motivational reminders, track progress, and contextualize user data via a companion app to facilitate understanding. Furthermore, peer comparison of activity data may benefit some in meeting their goals but may be detrimental to those who become discouraged from feeling that they are underperforming relative to their peers. Ultimately, wearables may better empower individuals by offering tailored support with positive reinforcement of users' successes while encouraging users when they fail to meet their targets.

Barriers to Use

Barriers to user empowerment include a perceived lack of accuracy and overly complex devices. However, lack of accessibility may be a greater issue, with concerns about pricing and how not owning a smartphone may mean that individuals miss out on the interpretation of data facilitated by the companion app. Another major concern relates to privacy, in which wearables collect sensitive health data. Consequently, strategies are required to mitigate the associated risks.

Strengths and Limitations

This review has its limitations. The nature of this research and its focus on wearable technology as a broad area may mean that relevant studies have been inadvertently missed. For example, although gray literature may reduce publication bias, it may give rise to selection bias because there is no gold standard method for retrieval [100]. Another potential source of bias may be the use of judgment when selecting studies for inclusion. Furthermore, the selection criteria may have excluded populations from low- and middle-income countries, where wearables can also be of benefit.

In certain circumstances, literature reviews have been included without including individual studies for review. It is important to note the reliance on the analysis undertaken as part of these reviews and that those reviews should be read alongside this review to see the full picture. Although this approach has its shortcomings from an analytic perspective, it was more practical as the individual studies that were screened met the inclusion criteria. It is also worth noting that although some studies that formed part of the literature review were identified from the initial literature searches for this scoping review, others were not. Although this only became apparent when the reference lists of the literature reviews were cross-checked against the records collated in the EndNote library, it gives rise to the question of how many other potential studies may have been missed and why?

The researcher took steps to minimize any bias and its effect on the research findings. The researcher consulted with senior academics throughout the research. A librarian guided the search strategy. Moreover, the researcher adhered to best practice recommendations from the PRISMA-ScR checklist and appraised the literature (which is not a requirement of scoping reviews) to further strengthen the rigor of this research. In fact, the very act of acknowledging these limitations has enabled the reader to contextualize the findings of the research within its limitations while demonstrating compliance with the recommended practice documented by the PRISMA-ScR [50].

The main consideration for this review was to balance the practicalities of research as a single researcher with the need to review representative, relevant evidence. This is where feedback on the research protocol and the availability of published scoping reviews (particularly those cited in the PRISMA-ScR "Tip Sheets" [48] as examples to illustrate good practice) have helped develop the methodology. Consequently, this review has been successful in meeting its aims and answering the research question; therefore, it should serve as a meaningful contribution to the literature in a dynamic, emerging area.

Conclusions

Although this scoping review has its limitations, its value is underscored by the fact that it fills a gap in the literature by addressing the research question and aims.

Considerable literature findings support the proposition that wearable health technology can empower users and, in turn, benefit providers and patients. Even if patients are unable to entirely self-manage their conditions, wearables have the potential to empower users to take more responsibility for their health and inspire positive behavior changes.

However, the ability of wearables to empower users may be limited by several factors. To maximize the potential for consumer wearables to integrate with the health system, support from health care professionals is critical. In addition, user feedback should be considered with respect to common barriers to use, such as technical issues and privacy concerns. As part of this process, designers of wearables should seek to incorporate more personalized support by way of positive reinforcement of any successes alongside encouragement for users who fail to meet their targets.

Future research may report whether there has been any progress in overcoming the barriers to use, including those mentioned earlier and others raised as part of this review. Further investigation of the long-term effects of wearables on individuals' outcomes through larger studies is warranted, as much of the literature revolves around small-scale studies. Moreover, despite the abundance of literature on wearables, what seems to be missing is the focus on the people who wear them. This may be because wearables, as viable instruments to assist with health care, have only been introduced in recent years. Specifically, future research may focus more closely on wearables and empowerment, especially as technology continues to evolve and advance over time. However, the challenge is for the publication of research to keep pace with rapid developments related to wearable health technology.

The adoption of wearables in the health sector may be gradual and fraught with challenges [101], but strategic change is certainly possible. In particular, any communication to relevant parties should emphasize the fact that although it may not be immediately apparent, each party has much to gain in the long run. Patients and users are expected to exercise greater control over their health and care decisions. Designers of the devices should benefit from having a more engaged user base. Similarly, individuals taking a more proactive role in their care should lessen the burden on clinicians and ease the pressure on the wider health system.

Acknowledgments

This study was supported by the University of Birmingham.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies for databases and gray literature.
[PDF File (Adobe PDF File), 121 KB - mhealth_v10i7e35684_app1.pdf]

Multimedia Appendix 2
Data charting form.
[PDF File (Adobe PDF File), 75 KB - mhealth_v10i7e35684_app2.pdf]

Multimedia Appendix 3
Summary of the included studies.
[PDF File (Adobe PDF File), 115 KB - mhealth_v10i7e35684_app3.pdf]

Multimedia Appendix 4
Characteristics of the included studies.
[PDF File (Adobe PDF File), 77 KB - mhealth_v10i7e35684_app4.pdf]

Multimedia Appendix 5
Funding details of the relevant included studies.
[PDF File (Adobe PDF File), 94 KB - mhealth_v10i7e35684_app5.pdf]

Multimedia Appendix 6
Contribution of the included studies to each of the themes.
[PDF File (Adobe PDF File), 82 KB - mhealth_v10i7e35684_app6.pdf]

References

1. Dehghani M, Kim K, Dangelico R. Will smartwatches last? factors contributing to intention to keep using smart wearable technology. Telematics Informatics 2018 May;35(2):480-490 [FREE Full text] [doi: 10.1016/j.tele.2018.01.007]
2. Apple Watch. Helping your patients identify early warning signs. Apple. URL: https://www.apple.com/uk/healthcare/apple-watch/ [accessed 2020-11-20]
3. Fitbit homepage. Fitbit. URL: https://www.fitbit.com/global/ie/home [accessed 2020-11-20]
4. Wearable sweat-sensor informs athletes of water and electrolyte loss. Doc Wire News. URL: https://www.docwirenews.com/docwire-pick/future-of-medicine-picks/wearable-sweat-sensor-informs-athletes-of-water-and-electrolyte-loss/ [accessed 2020-11-25]
5. Kim T, Wang Y, Oliver C, Thamm D, Cooling L, Paoletti C, et al. A temporary indwelling intravascular aphaeretic system for in vivo enrichment of circulating tumor cells. Nat Commun 2019 Apr 01;10(1):1478-1478 [FREE Full text] [doi: 10.1038/s41467-019-09439-9] [Medline: 30932020]
6. Coulter A. Engaging Patients in Health Care. Maidenhead, Berkshire, United Kingdom: Open University Press; 2011.
7. Walshe K, Smith J. Healthcare Management, 3rd Edition. London: Open University Press; 2016.
8. Steger A. Weighing the pros and cons of wearable health technology. HealthTech. URL: https://healthtechmagazine.net/article/2020/04/weighing-pros-and-cons-wearable-health-technology-perfcon [accessed 2020-11-27]
9. Ates H, Yetisen A, Güder F, Dincer C. Wearable devices for the detection of COVID-19. Nat Electron 2021 Jan 25;4(1):13-14 [FREE Full text] [doi: 10.1038/s41928-020-00533-1]
10. Werbrouck A, Swinnen E, Kerckhofs E, Buyl R, Beckwée D, De Wit L. How to empower patients? A systematic review and meta-analysis. Transl Behav Med 2018 Sep 08;8(5):660-674 [FREE Full text] [doi: 10.1093/tbm/iby064] [Medline: 29982675]
11. Aujoulat I, d'Hoore W, Deccache A. Patient empowerment in theory and practice: polysemy or cacophony? Patient Educ Couns 2007 Apr;66(1):13-20 [FREE Full text] [doi: 10.1016/j.pec.2006.09.008] [Medline: 17084059]
12. Funnell M. Patient empowerment. Crit Care Nurs Q 2004;27(2):201-204 [FREE Full text] [doi: 10.1097/00002727-200404000-00012] [Medline: 15137362]
13. Bravo P, Edwards A, Barr P, Scholl I, Elwyn G, McAllister M, Cochrane Healthcare Quality Research Group, Cardiff University. Conceptualising patient empowerment: a mixed methods study. BMC Health Serv Res 2015 Jul 01;15:252 [FREE Full text] [doi: 10.1186/s12913-015-0907-z] [Medline: 26126998]
14. Denecke K, Gabarron E, Petersen C, Merolli M. Defining participatory health informatics - a scoping review. Inform Health Soc Care 2021 Sep 02;46(3):234-243. [doi: 10.1080/17538157.2021.1883028] [Medline: 33622168]

15. Awotunde J, Jimoh R, AbdulRaheem M, Oladipo I, Folorunso S, Ajamu G. IoT-based wearable body sensor network for COVID-19 pandemic. In: Advances in Data Science and Intelligent Data Communication Technologies for COVID-19. Cham: Springer; 2022.
16. Cooper D, Bhuskute N, Walsh G. Exploring the impact and acceptance of wearable sensor technology for pre- and postoperative rehabilitation in knee replacement patients: a U.K.-based pilot study. JB JS Open Access 2022;7(2):e21.00154 [FREE Full text] [doi: 10.2106/JBJS.OA.21.00154] [Medline: 35506018]
17. Hasan A, Klintworth K, Hajat C. Wearable technology and health improvement. Occup Med (Lond) 2021 Apr 09;71(2):53-55 [FREE Full text] [doi: 10.1093/occmed/kqaa178] [Medline: 33836090]
18. Taffoni F, Rivera D, La Camera A, Nicolò A, Velasco J, Massaroni C. A wearable system for real-time continuous monitoring of physical activity. J Healthc Eng 2018;2018:1878354-1878316 [FREE Full text] [doi: 10.1155/2018/1878354] [Medline: 29849993]
19. Apple Watch Series 6. Apple. URL: https://www.apple.com/uk/apple-watch-series-6/ [accessed 2021-05-01]
20. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
21. Benefits of exercise. National Health Service England. URL: https://www.nhs.uk/live-well/exercise/exercise-health-benefits/ [accessed 2021-04-02]
22. Obesity and overweight. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2021-04-30]
23. Senthilingam M. Covid-19 has made the obesity epidemic worse, but failed to ignite enough action. BMJ 2021 Mar 04;372:n411 [FREE Full text] [doi: 10.1136/bmj.n411] [Medline: 33664084]
24. Oussedik S, MacIntyre S, Gray J, McMeekin P, Clement ND, Deehan DJ. Elective orthopaedic cancellations due to the COVID-19 pandemic: where are we now, and where are we heading? Bone Joint Open 2021 Feb 01;2(2):103-110. [doi: 10.1302/2633-1462.22.bjo-2020-0161.r1]
25. Pressure points in the NHS. British Medical Association. URL: https://www.bma.org.uk/advice-and-support/nhs-delivery-and-workforce/pressures/pressure-points-in-the-nhs [accessed 2020-12-02]
26. Innovation and growth: rationale for an innovation strategy. Organisation for Economic Co-operation and Development. URL: https://www.oecd.org/science/inno/39374789.pdf [accessed 2020-12-06]
27. Re-imagining care pathways: innovative solutions to improve clinical workloads. The Kings's Fund. URL: https://www.kingsfund.org.uk/events/re-imagining-care-pathways [accessed 2020-12-01]
28. Loncar-Turukalo T, Zdravevski E, Machado da Silva J, Chouvarda I, Trajkovik V. Literature on wearable technology for connected health: scoping review of research trends, advances, and barriers. J Med Internet Res 2019 Sep 05;21(9):e14017 [FREE Full text] [doi: 10.2196/14017] [Medline: 31489843]
29. NHS Long Term Plan. National Health Service England. URL: https://www.england.nhs.uk/long-term-plan/ [accessed 2020-11-28]
30. Liu N, Huang R, Baldacchino T, Sud A, Sud K, Khadra M, et al. Telehealth for noncritical patients with chronic diseases during the COVID-19 pandemic. J Med Internet Res 2020 Aug 07;22(8):e19493 [FREE Full text] [doi: 10.2196/19493] [Medline: 32721925]
31. Greiwe J, Nyenhuis SM. Wearable technology and how this can be implemented into clinical practice. Curr Allergy Asthma Rep 2020 Jun 06;20(8):36 [FREE Full text] [doi: 10.1007/s11882-020-00927-3] [Medline: 32506184]
32. Remote patient monitoring 2018. KLAS Research. URL: https://klasresearch.com/report/remote-patient-monitoring-2018/1273 [accessed 2020-12-05]
33. Scheid J, West S. Opportunities of wearable technology to increase physical activity in individuals with chronic disease: an editorial. Int J Environ Res Public Health 2019 Aug 28;16(17):1-6 [FREE Full text] [doi: 10.3390/ijerph16173124] [Medline: 31466226]
34. Bove L. Increasing patient engagement through the use of wearable technology. J Nurse Practitioners 2019 Sep;15(8):535-539 [FREE Full text] [doi: 10.1016/j.nurpra.2019.03.018]
35. Azodo I, Williams R, Sheikh A, Cresswell K. Opportunities and challenges surrounding the use of data from wearable sensor devices in health care: qualitative interview study. J Med Internet Res 2020 Oct 22;22(10):e19542 [FREE Full text] [doi: 10.2196/19542] [Medline: 33090107]
36. Guidance regulating medical devices in the UK. GOV.UK. URL: https://www.gov.uk/guidance/regulating-medical-devices-in-the-uk#overview [accessed 2021-04-27]
37. Important safety and product and services information. Fitbit. URL: https://tinyurl.com/277av7cb [accessed 2021-04-25]
38. Apple Watch 4 is now an FDA Class 2 medical device: detects falls, irregular heart rhythm. Forbes. URL: https://tinyurl.com/25t65jtz [accessed 2021-02-10]
39. How can leaders make recent digital health gains last? Accenture. URL: https://www.accenture.com/us-en/insights/health/leaders-make-recent-digital-health-gains-last [accessed 2020-11-28]
40. Exworthy M, Mannion R. Dismantling the NHS? Evaluating the Impact of Health Reforms. Bristol: Policy Press; 2016.

41. Dineen-Griffin S, Garcia-Cardenas V, Williams K, Benrimoj SI. Helping patients help themselves: a systematic review of self-management support strategies in primary health care practice. PLoS One 2019 Aug 1;14(8):e0220116 [FREE Full text] [doi: 10.1371/journal.pone.0220116] [Medline: 31369582]
42. Scoping reviews: what they are and how you can do them. Cochrane Training. URL: https://training.cochrane.org/resource/scoping-reviews-what-they-are-and-how-you-can-do-them [accessed 2020-12-07]
43. Peters M, Godfrey C, Khalil H, McInerney P, Parker D, Soares C. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146 [FREE Full text] [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
44. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Social Res Methodol 2005 Feb;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
45. Munn Z, Peters M, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143-147 [FREE Full text] [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
46. Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. BMC Med Res Methodol 2018 Jan 10;18(1):5 [FREE Full text] [doi: 10.1186/s12874-017-0468-4] [Medline: 29316881]
47. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review--a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy 2005 Jul;10 Suppl 1:21-34 [FREE Full text] [doi: 10.1258/1355819054308530] [Medline: 16053581]
48. PRISMA for scoping reviews. Preferred Reporting Items for Systematic Reviews and Meta-Analyses. URL: http://www.prisma-statement.org/Extensions/ScopingReviews [accessed 2021-02-12]
49. Levac D, Colquhoun H, O'Brien K. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
50. Tricco A, Lillie E, Zarin W, O'Brien K, Colquhoun H, Levac D. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018:467-473 [FREE Full text] [doi: 10.7326/m18-0850]
51. Sucharew H, Macaluso M. Progress notes: methods for research evidence synthesis: the scoping review approach. J Hosp Med 2019 Jul 01;14(7):416-418 [FREE Full text] [doi: 10.12788/jhm.3248] [Medline: 31251164]
52. Pham M, Rajić A, Greig J, Sargeant J, Papadopoulos A, McEwen S. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Res Synth Methods 2014 Dec;5(4):371-385 [FREE Full text] [doi: 10.1002/jrsm.1123] [Medline: 26052958]
53. Peters M, Marnie C, Tricco A, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126 [FREE Full text] [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
54. Resources. JBI. URL: https://jbi.global/scoping-review-network/resources [accessed 2020-11-28]
55. Robson C, McCartan K. Real World Research, 4th Edition. Chichester: John Wiley & Sons; 2016.
56. Long H, French D, Brooks J. Optimising the value of the critical appraisal skills programme (CASP) tool for quality appraisal in qualitative evidence synthesis. Res Method Med Health Sci 2020 Aug 06;1(1):31-42 [FREE Full text] [doi: 10.1177/2632084320947559]
57. Critical Appraisal Tools. JBI. URL: https://jbi.global/critical-appraisal-tools [accessed 2021-03-03]
58. Scott P, Nakkas H, Roderick P. Protocol for a scoping review to understand how interorganisational electronic health records affect hospital physician and pharmacist decisions. BMJ Open 2019 Jan 15;9(1):e023712 [FREE Full text] [doi: 10.1136/bmjopen-2018-023712] [Medline: 30647036]
59. JBI Manual for evidence synthesis. JBI. URL: https://wiki.jbi.global/display/MANUAL/11.2.7+Data+extraction [accessed 2020-12-04]
60. Booth A, Sutton A, Papaioannou D. Systematic Approaches to a Successful Literature Review 2nd Edition. Thousand Oaks, California: SAGE Publications Ltd; 2016.
61. Braun V, Clarke V. Successful Qualitative Research A Practical Guide for Beginners. Thousand Oaks, California: SAGE Publications; 2013.
62. Jesus TS, Bright F, Kayes N, Cott CA. Person-centred rehabilitation: what exactly does it mean? Protocol for a scoping review with thematic analysis towards framing the concept and practice of person-centred rehabilitation. BMJ Open 2016 Jul 19;6(7):e011959 [FREE Full text] [doi: 10.1136/bmjopen-2016-011959] [Medline: 27436670]
63. Safarpour H, Khorasani-Zavareh D, Mohammadi R. The common road safety approaches: a scoping review and thematic analysis. Chin J Traumatol 2020 Apr;23(2):113-121 [FREE Full text] [doi: 10.1016/j.cjtee.2020.02.005] [Medline: 32178997]
64. Finkelstein S, Sharma U, Furlonger B. The inclusive practices of classroom teachers: a scoping review and thematic analysis. Int J Inclusive Educ 2019 Jan 30;25(6):735-762 [FREE Full text] [doi: 10.1080/13603116.2019.1572232]
65. Auerswald T, Meyer J, von Holdt K, Voelcker-Rehage C. Application of activity trackers among nursing home residents-a pilot and feasibility study on physical activity behavior, usage behavior, acceptance, usability and motivational impact. Int J Environ Res Public Health 2020 Sep 14;17(18):1-21 [FREE Full text] [doi: 10.3390/ijerph17186683] [Medline: 32937840]

66. Belsi A, Papi E, McGregor AH. Impact of wearable technology on psychosocial factors of osteoarthritis management: a qualitative study. BMJ Open 2016 Feb 03;6(2):e010064 [FREE Full text] [doi: 10.1136/bmjopen-2015-010064] [Medline: 26842273]
67. Coughlin S, Caplan L, Stone R. Use of consumer wearable devices to promote physical activity among breast, prostate, and colorectal cancer survivors: a review of health intervention studies. J Cancer Surviv 2020 Jun;14(3):386-392 [FREE Full text] [doi: 10.1007/s11764-020-00855-1] [Medline: 31933148]
68. DiFrancisco-Donoghue J, Jung M, Stangle A, Werner W, Zwibel H, Happel P, et al. Utilizing wearable technology to increase physical activity in future physicians: a randomized trial. Prev Med Rep 2018 Dec;12:122-127 [FREE Full text] [doi: 10.1016/j.pmedr.2018.09.004] [Medline: 30234000]
69. Friel C, Garber C. An examination of the relationship between motivation, physical activity, and wearable activity monitor use. J Sport Exerc Psychol 2020 Mar 09:1-8 [FREE Full text] [doi: 10.1123/jsep.2019-0170] [Medline: 32150727]
70. Goode A, Hall K, Batch B, Huffman K, Hastings S, Allen K, et al. The impact of interventions that integrate accelerometers on physical activity and weight loss: a systematic review. Ann Behav Med 2017 Feb;51(1):79-93 [FREE Full text] [doi: 10.1007/s12160-016-9829-1] [Medline: 27565168]
71. Henriksen A, Sand A, Deraas T, Grimsgaard S, Hartvigsen G, Hopstock L. Succeeding with prolonged usage of consumer-based activity trackers in clinical studies: a mixed methods approach. BMC Public Health 2020 Aug 27;20(1):1300 [FREE Full text] [doi: 10.1186/s12889-020-09406-w] [Medline: 32854671]
72. Janevic M, Shute V, Murphy S, Piette J. Acceptability and effects of commercially available activity trackers for chronic pain management among older African American adults. Pain Med 2020 Feb 01;21(2):e68-e78 [FREE Full text] [doi: 10.1093/pm/pnz215] [Medline: 31509196]
73. Kim Y, Lumpkin A, Lochbaum M, Stegemeier S, Kitten K. Promoting physical activity using a wearable activity tracker in college students: a cluster randomized controlled trial. J Sports Sci 2018 Aug;36(16):1889-1896 [FREE Full text] [doi: 10.1080/02640414.2018.1423886] [Medline: 29318916]
74. Lewis Z, Ottenbacher K, Fisher S, Jennings K, Brown A, Swartz M, et al. Effect of electronic activity monitors and pedometers on health: results from the TAME health pilot randomized pragmatic trial. Int J Environ Res Public Health 2020 Sep 18;17(18):1-11 [FREE Full text] [doi: 10.3390/ijerph17186800] [Medline: 32961834]
75. Lugones-Sanchez C, Sanchez-Calavera M, Repiso-Gento I, Adalia E, Ramirez-Manent J, Agudo-Conde C, EVIDENT 3 Investigators. Effectiveness of an mHealth intervention combining a smartphone app and smart band on body composition in an overweight and obese population: randomized controlled trial (EVIDENT 3 study). JMIR Mhealth Uhealth 2020 Nov 26;8(11):e21771 [FREE Full text] [doi: 10.2196/21771] [Medline: 33242020]
76. Papi E, Belsi A, McGregor AH. A knee monitoring device and the preferences of patients living with osteoarthritis: a qualitative study. BMJ Open 2015 Sep 07;5(9):e007980 [FREE Full text] [doi: 10.1136/bmjopen-2015-007980] [Medline: 26346873]
77. Rieder A, Eseryel U, Lehrer C, Jung R. Why users comply with wearables: the role of contextual self-efficacy in behavioral change. Int J Human Comput Interact 2020 Sep 30;37(3):281-294 [FREE Full text] [doi: 10.1080/10447318.2020.1819669]
78. Rodgers M, Alon G, Pai V, Conroy R. Wearable technologies for active living and rehabilitation: current research challenges and future opportunities. J Rehabil Assist Technol Eng 2019;6:2055668319839607-2055668319839609 [FREE Full text] [doi: 10.1177/2055668319839607] [Medline: 31245033]
79. Rupp M, Michaelis J, McConnell D, Smither J. The role of individual differences on perceptions of wearable fitness device trust, usability, and motivational impact. Appl Ergon 2018 Jul;70:77-87 [FREE Full text] [doi: 10.1016/j.apergo.2018.02.005] [Medline: 29866329]
80. Shin G, Jarrahi M, Fei Y, Karami A, Gafinowitz N, Byun A, et al. Wearable activity trackers, accuracy, adoption, acceptance and health impact: a systematic literature review. J Biomed Inform 2019 May;93:103153 [FREE Full text] [doi: 10.1016/j.jbi.2019.103153] [Medline: 30910623]
81. Soliño-Fernandez D, Ding A, Bayro-Kaiser E, Ding E. Willingness to adopt wearable devices with behavioral and economic incentives by health insurance wellness programs: results of a US cross-sectional survey with multiple consumer health vignettes. BMC Public Health 2019 Dec 16;19(1):1649 [FREE Full text] [doi: 10.1186/s12889-019-7920-9] [Medline: 31839006]
82. Stiglbauer B, Weber S, Batinic B. Does your health really benefit from using a self-tracking device? Evidence from a longitudinal randomized control trial. Comput Human Behav 2019 May;94:131-139 [FREE Full text] [doi: 10.1016/j.chb.2019.01.018]
83. The digital patient: transforming primary care? The Nuffield Trust. 2016. URL: https://www.nuffieldtrust.org.uk/research/the-digital-patient-transforming-primary-care [accessed 2021-02-04]
84. Wulfovich S, Fiordelli M, Rivas H, Concepcion W, Wac K. " I must try harder": design implications for mobile apps and wearables contributing to self-efficacy of patients with chronic conditions. Front Psychol 2019;10:2388 [FREE Full text] [doi: 10.3389/fpsyg.2019.02388] [Medline: 31749733]
85. Frenk J. Dimensions of health system reform. Health Policy 1994 Jan;27(1):19-34 [FREE Full text] [doi: 10.1016/0168-8510(94)90155-4]

86. Smuck M, Odonkor C, Wilt J, Schmidt N, Swiernik M. The emerging clinical role of wearables: factors for successful implementation in healthcare. NPJ Digit Med 2021 Mar 10;4(1):45-48 [FREE Full text] [doi: 10.1038/s41746-021-00418-3] [Medline: 33692479]
87. Rogers E. Diffusion of innovations. London: Collier-Macmillan; 1962.
88. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215 [FREE Full text] [doi: 10.1037/0033-295x.84.2.191]
89. The principles. Information Commissioner's Office. URL: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/ [accessed 2021-05-02]
90. Get the Apple Watch deal. Vitality. URL: https://www.vitality.co.uk/rewards/partners/active-rewards/apple-watch/ [accessed 2021-05-02]
91. Jamruk K. The weight game: fighting childhood obesity with childhood video technology. J Leg Med 2017;37(1-2):175-194 [FREE Full text] [doi: 10.1080/01947648.2017.1303409] [Medline: 28910218]
92. Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. J Epidemiol Community Health 2020 Nov 13;74(11):964-968 [FREE Full text] [doi: 10.1136/jech-2020-214401] [Medline: 32535550]
93. Tudor Hart J. The inverse care law. Lancet 1971 Feb;297(7696):405-412 [FREE Full text] [doi: 10.1016/s0140-6736(71)92410-x]
94. Osterloh R. Google completes Fitbit acquisition. Google. URL: https://blog.google/products/devices-services/fitbit-acquisition/ [accessed 2021-05-02]
95. Google tries to allay Fitbit-deal privacy fears. BBC News. URL: https://www.bbc.co.uk/news/technology-55662659 [accessed 2021-05-02]
96. COVID-19 app support. National Health Service England. URL: https://www.covid19.nhs.uk/ [accessed 2021-05-02]
97. Privacy-preserving contact tracing. Apple. URL: https://covid19.apple.com/contacttracing [accessed 2021-05-02]
98. Criddle C, Kelion L. Coronavirus contact-tracing: world split between two types of app. BBC News. URL: https://www.bbc.co.uk/news/technology-52355028 [accessed 2021-05-02]
99. Lehmann C. Privacy concerns hindering digital contact tracing. WebMD. URL: https://www.webmd.com/lung/news/20200928/privacy-concerns-hindering-digital-contact-tracing [accessed 2021-05-02]
100. Paez A. Gray literature: an important resource in systematic reviews. J Evid Based Med 2017 Aug;10(3):233-240 [FREE Full text] [doi: 10.1111/jebm.12266] [Medline: 28857505]
101. Kotter J. Force For Change How Leadership Differs from Management. London: Free Press; 1990.

Abbreviations

ECG: electrocardiogram
JBI: Joanna Briggs Institute
MHRA: Medicines and Healthcare Products Regulatory Agency
NHS: National Health Service
PHI: participatory health informatics
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by L Buis; submitted 13.12.21; peer-reviewed by S Mantalvanos, S Subramaniam; comments to author 06.01.22; revised version received 20.01.22; accepted 08.06.22; published 13.07.22.

Please cite as:
Kang HS, Exworthy M
Wearing the Future—Wearables to Empower Users to Take Greater Responsibility for Their Health and Care: Scoping Review
JMIR Mhealth Uhealth 2022;10(7):e35684
URL: https://mhealth.jmir.org/2022/7/e35684
doi:10.2196/35684
PMID:35830222

Original Paper

Impact of a Mobile Application for Tracking Nausea and Vomiting During Pregnancy (NVP) on NVP Symptoms, Quality of Life, and Decisional Conflict Regarding NVP Treatments: MinSafeStart Randomized Controlled Trial

Elin Ngo[1], MSc; Maria Bich-Thuy Truong[1], PhD; David Wright[2,3], PhD; Hedvig Nordeng[1,4], PhD

[1]PharmacoEpidemiology and Drug Safety Research Group, Department of Pharmacy, University of Oslo, Oslo, Norway

[2]School of Allied Health Professions, University of Leicester, England, United Kingdom

[3]Centre for Pharmacy, University of Bergen, Bergen, Norway

[4]Department of Child Health and Development, National Institute of Public Health, Oslo, Norway

Corresponding Author:
Elin Ngo, MSc
PharmacoEpidemiology and Drug Safety Research Group
Department of Pharmacy
University of Oslo
Postbox 1068 Blindern
Oslo, 0316
Norway
Phone: 47 93849866
Email: e.t.p.ngo@farmasi.uio.no

Related Article:

This is a corrected version. See correction statement: https://mhealth.jmir.org/2022/9/e41927

Abstract

Background: Pregnant women are active users of mobile apps for health purposes. These apps may improve self-management of health-related conditions. Up to 70% of pregnant women experience nausea and vomiting (NVP). Even mild NVP can significantly reduce quality of life (QoL), and it can become an economic burden for both the woman and society. NVP often occurs before the first maternal care visit; therefore, apps can potentially play an important role in empowering pregnant women to recognize, manage, and seek appropriate treatment for NVP, when required.

Objective: This study investigated whether the MinSafeStart (MSS) mobile app could impact NVP-related symptoms, QoL, and decisional conflict regarding NVP treatment.

Methods: This randomized controlled trial enrolled 268 pregnant women with NVP in Norway from 2019 to 2020. The intervention group had access to the MSS app, which could be used to track NVP symptoms and access tailored advice. NVP severity was rated with the Pregnancy Unique Quantification of Emesis (PUQE) score. The control group followed standard maternal care. We collected data on maternal baseline characteristics, NVP severity, QoL, and decisional conflict using 2 sets of online questionnaires. One set of questionnaires was completed at enrollment, and the other was completed after 2 weeks. We performed linear regression analyses to explore whether the use of the MSS app was associated with NVP severity, QoL, or decisional conflict.

Results: Among the 268 women enrolled in the study, 192 (86.5%) completed the baseline questionnaires and were randomized to either the intervention (n=89) or control group (n=103). In the intervention group, 88 women downloaded the app, and 468 logs were recorded. In both groups, women were enrolled at a median of 8 gestational weeks. At baseline, the average PUQE scores were 4.9 and 4.7; the average QoL scores were 146 and 149; and the average DCS scores were 40 and 43 in the intervention and control groups, respectively. The app had no impact on NVP severity (aβ 0.6, 95% Cl −0.1 to 1.2), QoL (aβ −5.3, 95% Cl −12.5 to 1.9), or decisional conflict regarding NVP treatment (aβ −1.1, 95% Cl −6.2 to 4.2), compared with standard care.

Conclusions: Tracking NVP symptoms with the MSS app was not associated with improvements in NVP symptoms, QoL, or decisional conflict after 2 weeks, compared with standard care. Future studies should include a process evaluation to improve our understanding of how pregnant women use the app and how to optimize its utility within maternity care. Specifically, studies should focus on how digital tools might facilitate counseling and communication between pregnant women and health care providers regarding NVP management during pregnancy.

Trial Registration: ClinicalTrails.gov (NCT04719286): https://www.clinicaltrials.gov/ct2/show/NCT04719286

(JMIR Mhealth Uhealth 2022;10(7):e36226) doi:10.2196/36226

KEYWORDS

eHealth; mHealth; decision support tool; nausea and vomiting; pregnancy; RCT

Introduction

Background

Pregnant women and women of reproductive age are active users of mobile apps for health purposes [1]. Available apps are designed for promoting self-management of chronic diseases, such as migraine and diabetes; tracking gestational weeks, weight, and belly measurements during pregnancy; and keeping track of pregnancy development in general [1,2]. These apps are often used to supplement routine care, because women tend to search for health-related information early in pregnancy, before and after health consultations, and when making decisions [1,3-5]. Often, the primary motivation for using apps is the need for easily accessible health information [6]. Our recent systematic review on decision support tools in pregnancy revealed that few studies had investigated the effect of digital tools on the course of pregnancy and pregnancy-related ailments. However, available studies have shown that apps could have a positive impact on the knowledge level of pregnant women, when integrated as part of patient care. Pregnant women also seemed to appreciate and were satisfied with digital tools [7].

Nausea and vomiting in pregnancy (NVP) is one of the most common pregnancy-related conditions. NVP affects up to 70% of pregnant women worldwide [8,9]. NVP symptoms often occur during the first few weeks of pregnancy, on average, at around gestational week 4 [10]. The etiology of NVP is not clearly understood, but it is thought to be multifactorial and complex [10]. The severity of NVP can range from mildly uncomfortable to hyperemesis gravidarum (HG), which is the most severe form of NVP. HG affects 1%-3% of all pregnant women, and it is the most common reason for hospitalization in early pregnancy [8]. Although HG is a relatively rare condition, it is essential to recognize the burden of NVP in general. Previous studies have shown that even mild NVP symptoms significantly reduce quality of life (QoL) of pregnant women and their willingness to become pregnant again [11,12]. Moreover, as the severity of NVP increases, the costs for society increase due to increased hospital and emergency room admissions, health care visits, prescribed medications, and income loss for both the woman and her partner [13].

NVP treatment guidelines recommend early recognition and treatment to prevent or reduce more severe symptoms. The first-line management of mild symptoms consists of nonpharmacologic measures, including lifestyle and dietary changes (Multimedia Appendix 1). Pharmacological treatment is indicated when NVP symptoms are moderate to severe or when symptoms significantly impact the women's daily activities [14,15]. The first NVP symptoms typically occur early in pregnancy and, often, before the first maternal care visit. Therefore, it is important to empower pregnant women to ensure that they can optimally manage NVP symptoms [15,16].

Digitalization, eHealth initiatives, and the wide use of the internet have opened up new possibilities for using digital tools in maternal care [17]. Mobile apps can enable pregnant women to take a more active role in self-care and disease management during pregnancy. Moreover, these apps can provide large amounts of patient-generated data during pregnancy for research purposes [17,18]. The Pregnancy Unique Quantification of Emesis (PUQE) score is an internationally validated tool for categorizing the severity of NVP based on 3 questions regarding vomiting, nausea, and retching symptoms [19,20]. In the latest (2009) version of the PUQE score, women are asked to rate the severity of symptoms that occurred in the last 24 hours [19]. A translated and validated Norwegian version of the PUQE score became available in 2015 [21]. Incorporating the PUQE score into an app could potentially empower women by improving their management of NVP. The app could allow women to track symptoms over time and record responses to interventions. Because 99%-100% of women of reproductive age use smartphones [22] and most women use health-related apps [23,24], digital tools should be particularly suitable for maternal care.

A recent review pointed out that, although there is a growing number of apps available for monitoring and managing health-related issues, the majority are never tested nor clinically validated [25]. That finding implied that it remains largely unknown whether available apps are beneficial or whether they even have an effect on clinical outcomes. A prior study showed that integrating apps into professional clinical services could potentially improve the effectiveness of health care [26]. Our previous review concluded that the innovative use of eHealth initiatives and digitalization could potentially empower pregnant patients and improve maternal care [7]. However, at the same time, a more scientific approach is needed for testing and evaluating these apps and other digital tools. Indeed, health care providers should encourage patients to use only tools that are beneficial and effective as a supplement to routine maternity care.

Objective

The primary aim of this study was to investigate whether the MinSafeStart (MSS) mobile app could impact NVP severity in pregnant women. The secondary aims were to assess whether the MSS app could affect the QoL of pregnant women and improve their ability to make decisions regarding NVP treatment.

Specifically, the primary research question was: Will women who use the MSS app for 2 weeks have different NVP symptoms, based on PUQE scores, compared with women who follow standard maternal care without the MSS app?

The specific secondary research questions were: (1) Will women who use the MSS app for 2 weeks have different QoL, based on Health-related Quality of Life for Nausea and Vomiting during Pregnancy (NVPQOL) scores, compared with women who follow standard maternal care without the MSS app? (2) Will women who use the MSS app for 2 weeks have different decisional conflict scale (DCS) scores regarding NVP treatment, compared with women who follow standard maternal care without the MSS app? (3) Will the use of the MSS app modify the association between the PUQE score and the NVPQOL score (ie, is the MSS app an effect modifier)?

Methods

Study Design, Study Population, Recruitment, and Sample Size

The MinSafeStart study was a randomized controlled trial. We recruited pregnant women in Norway between September 2019 and June 2020. All pregnant women over 18 years old who were currently experiencing NVP, owned a smartphone (iOS or Android), and could speak and understand Norwegian were eligible for inclusion.

Participants were primarily recruited through social media advertisements. Invitations to participate in the study were available on the study Facebook page, the Norwegian Hyperemesis Gravidarum Patient Organization's Facebook page, and other pregnancy-related web pages or forums, such as "altformamma.no" (all for mommy) and "tryggmammamedisin.no" (safe mother medications). Invitations were additionally accessible through the Helseoversikt app. Helseoversikt is a digital platform used by health care centers all over Norway that provides relevant health information to pregnant women and parents.

All invitations to participate contained a link to the online consent form. When the women signed the consent form and responded to the baseline questionnaire, they were automatically randomized to either the intervention or control group. Both groups received emails with information about the study group to which they were assigned. The intervention group also received an email with instructions on how to download and use the app.

Results from the power analysis suggested that we would need a total of 250 pregnant women (n=125 in each group, 2-tailed hypothesis) to detect a mean difference of 3 points in the PUQE score between the groups, with a power of 80% (Cohen d=0.5). This total sample size included a 25% dropout rate.

Randomization

An automated software program was specifically developed for the project. The software automatically managed participant enrollment, randomization to study groups, and email distributions of electronic information and online questionnaires to the study participants. This software was developed for the project by the University Center for Information Technology (USIT) at the University of Oslo.

Development of the MinSafeStart Mobile Application

The MSS app was a patient-centered app for women with NVP. Our research group developed the MSS app in collaboration with interaction designers, programmers, and researchers from USIT. The app utilized the daily PUQE score (Multimedia Appendix 2) to categorize NVP severity (ie, mild, moderate, or severe), and it displayed the fluctuations over time in a graph (Figures 1 and 2). The aim of the app was to assist pregnant women in identifying and managing NVP. The app tracked their NVP symptoms every day and provided tailored advice according to the severity of their symptoms. All women with NVP symptoms received lifestyle and dietary advice (eg, stay hydrated, eat small meals frequently, and get some rest). Women that experienced severe NVP also received information about medical treatments. The app alerted the woman to seek appropriate treatment when she logged PUQE scores >13 for more than 3 consecutive days. The app was user tested in July 2018. The user test included 9 women who completed a structured interview with a set of tasks and questions regarding the app. Of these 9 women, 5 also participated in a focus group to discuss and share their experiences and opinions about the app. The user test results showed that the app was user-friendly and had the potential to empower women who experienced NVP to improve their management skills and treatment decisions. Nevertheless, some minor issues were mentioned in the user test and focus group that could be improved (ie, explanations of terminologies, an opportunity to change the due date, links to external information, an overview of previously logged scores, and the layout and design). These suggestions were incorporated into the app to make it as user-friendly as possible before it was launched for iOS and Android smartphones.

Figure 1. Front page of the MinSafeStart application (in Norwegian) for pregnant women to track nausea and vomiting, showing the user's gestational week at the top, text in the center (“How do you feel? Use the button below to log your NVP symptoms”), and button to log nausea and vomiting in pregnancy (NVP) symptoms.

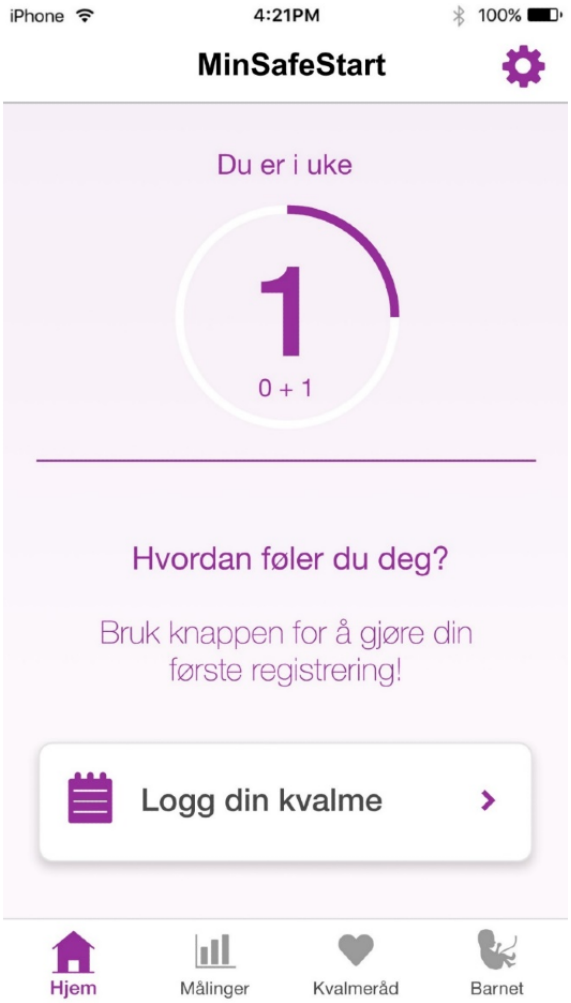
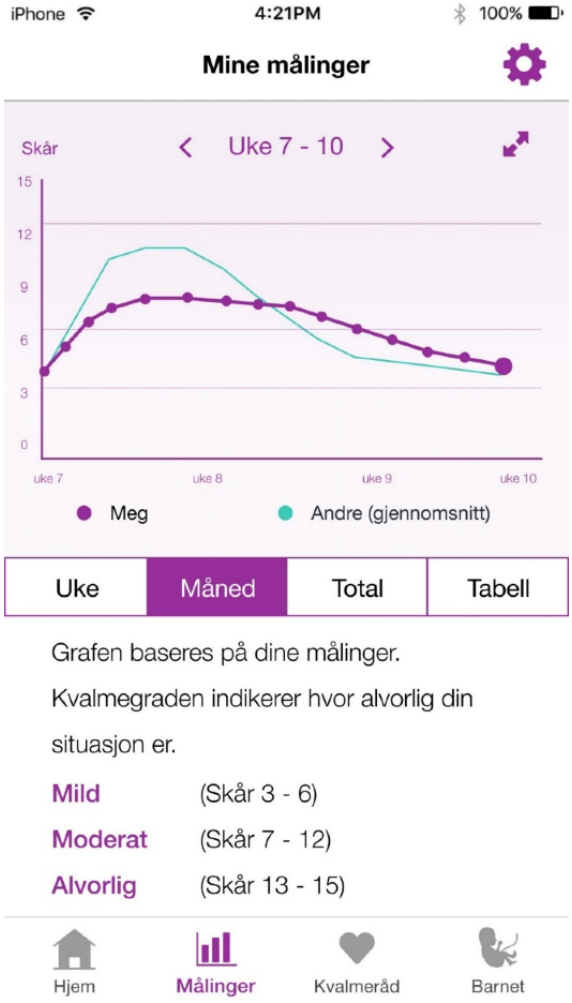


Figure 2. The MinSafeStart app (in Norwegian) for pregnant women with nausea and vomiting (NVP) shows the women's NVP loggings (Mine Malinger) as the user's NVP scores (purple) as a graph over time (week [Uke], month [Måned], for all data recorded in the app [Total]), compared with the mean Pregnancy Unique Quantification of Emesis (PUQE) score of other pregnant women (blue line), or as a table (Tabell). The bottom section shows the numeric rating scale for NVP symptoms. Alvorlig: severe; Moderat: moderate; Skår: Score.



Data Collection

In this MinSafeStart study, we collected data from the MSS app and from 4 sets of questionnaires (Q1-Q4) that were completed electronically. Q1 was administered to participants at enrollment (baseline), and Q2 was administered 2 weeks later. Q3 and Q4 were additional follow-up questionnaires administered at 4 weeks and 6 weeks after baseline, respectively. All questionnaires were sent to participants by email with the automated software developed for the study. This study only analyzed data from the Q1 and Q2 sets of questionnaires. We selected a 2-week follow-up for this study because we considered that 2 weeks were sufficient to become familiar with the app.

All data collected from the app and questionnaires were automatically encrypted and stored at the Service for Sensitive Data at the University of Oslo (TSD). The TSD platform is available to collect, store, and analyze sensitive data [27]. The platform is protected by a 2-step password system and meets all the necessary requirements to maintain compliance with Norwegian regulations regarding individual privacy. The data are not accessible outside of the TSD. Only registered researchers within the project had access to the data and the encryption key.

The study is reported in accordance with the CONSORT-EHEALTH checklist (Multimedia Appendix 3).

Intervention Group

All women in the intervention group were given access to the MSS app in addition to standard maternal care. They were free to log their NVP symptoms into the app whenever convenient. Standard maternity care in Norway is free of charge. It includes 9 routine checkups with a midwife or physician and 1 ultrasound scan at gestational week 18 [28].

The app recommended logging symptoms every 24 hours because the PUQE score was calculated based on NVP symptoms over the past 24 hours. Users could also compare their symptoms to the expected population average NVP score. Thus, women received individual treatment advice based on their PUQE scores (Multimedia Appendix 1). Women also received general dietary and lifestyle advice (eg, get some rest, stay hydrated, eat small meals frequently, and avoid fatty and spicy foods [29]) independent of their PUQE score. Women with moderate or severe symptoms received additional advice about antiemetic medications. When a woman scored ≥13 points

(ie, severe NVP) for more than 3 consecutive days, she would see a pop-up message that encouraged her to see the doctor.

Control Group

The control group received only standard maternal care.

Outcome Measures

NVP Severity

The PUQE score was internationally validated for rating the severity of NVP symptoms over the past 24 hours (Multimedia Appendix 2) [19,21]. The scale consists of 3 questions. Each question is rated from 1 to 5. The total score ranges from 3 to 15 points, where ≤6 points indicate mild NVP, 7-12 points indicate moderate NVP, and 13 or more points indicate severe NVP. This study utilized the translated and validated Norwegian version of the PUQE [21]. We evaluated the change in PUQE scores from Q1 to Q2 (ie, after 2 weeks).

Quality of Life

The NVPQOL was used to rate QoL [30] over the past week (Multimedia Appendix 2). The score includes 30 items covering 4 general domains: physical symptoms and aggravating factors, fatigue, emotions, and limitations. Each item is rated on a Likert scale that ranges from 1 (never) to 7 (all the time). The total score ranges from 30 to 210 points, and lower scores indicate a better QoL. The NVPQOL score is significantly associated with the SF-12 health-related QoL questionnaire [30]. We evaluated the change in NVPQOL scores from Q1 to Q2.

Decisional Conflict

Decisional conflict was measured with the decisional conflict scale (DCS). The DCS measures the individual's perception of uncertainty in choosing options, modifiable factors that contributed to uncertainty, and decision-making effectiveness [31,32] (Multimedia Appendix 2). The DCS has been widely used in previous studies among pregnant women to evaluate their decision-making abilities regarding the use of antidepressants and the choice between vaginal birth or cesarean section [33,34]. The DCS consists of 16 items and 5 response categories (strongly agree, agree, neither agree nor disagree, disagree, and strongly disagree). The total score ranges from 0 to 100 points. Scores below 25 points indicate low decisional conflict, scores of 25 to 37.5 points indicate moderate decisional conflict, and scores above 37.5 points indicate high decisional conflict. We evaluated the change in DCS scores from Q1 to Q2.

Statistical Analyses

Descriptive Analysis

Categorical variables (ie, relationship status, education level, work situation, parity, and prior NVP symptoms) are presented as percentages for each group (intervention and control groups). Continuous variables are presented as the median and range (eg, gestational week) or the mean and SD (eg, maternal age). We performed a Pearson Chi-squared test to compare categorical variables, except when the expected cell count was less than 5; in those cases, we performed a Fisher exact test. We performed a Student *t* test to compare continuous variables. All analyses were performed with Stata/MP v.16.1. *P* values <.05 were considered statistically significant.

Primary and Secondary Analyses

We performed univariate and multivariable linear regression analyses to estimate associations between the use of the MSS app and (1) NVP severity, (2) QoL, and (3) decisional conflict. All results are presented as the crude and adjusted beta-coefficients (β) with 95% CIs. We adjusted the multivariable linear regression model with predefined covariates (ie, baseline PUQE score, baseline NVPQOL score, and baseline DCS) [35].

Subanalyses

We performed a prespecified stratified analysis to assess whether employment in the health sector modified the association between the use of the MSS app and the PUQE score. We reasoned that women employed in the health sector might have better access to information and advice regarding NVP management, and thus, they may have less need for an app to track their NVP symptoms, compared with women employed in other settings. Alternatively, they may have received more support or information from co-workers in the field that allowed them to capitalize on the information provided by the app, compared with women employed in other settings.

Ethical Approval

This study was approved by the Regional Committees for Medical and Health Research Ethics in Norway (Ref: 2018/2298). Informed consent to participate in the study was obtained from all participants.

Results

Study Population

Overall, 268 women consented to participate in the study (Figure 3). Of these, 192 (86.5%) responded to the baseline questionnaires (Q1) and were randomized to either the intervention group (n=89) or the control group (n=103). In total, 137 women responded to the follow-up questionnaires 2 weeks later (Q2). The dropout rates were 34% (30/89) for the intervention group and 24.3% (25/103) for the control group. The main reason for dropout was "lack of response."

At enrollment, the median stage of pregnancy was the same in both groups: 8 (range 4-36) gestational weeks in the intervention group and 8 (range 4-39) gestational weeks in the control group. These groups had the same mean age at enrollment: 32 (SD 4.6) years and 32 (SD 3.9) years, respectively. Most women had been pregnant previously (65/89, 73%, and 76/103, 73.8%, respectively). In both groups, 80% (52/89 and 61/103, respectively) had experienced NVP in at least one previous pregnancy. None of the women reported severe NVP (ie, PUQE score ≥13) at baseline. A comparison of baseline characteristics using the Student *t* test, Chi-squared test, or Fisher exact test indicated no statistical difference (all *P*<.05) between the 2 study groups (Table 1).

Figure 3. Flowchart of the study participants in the enrolled group, allocation groups, and follow-up groups. app: MinSafeStart mobile app; PUQE: Pregnancy Unique Quantification of Emesis; Q1: Questionnaire 1.

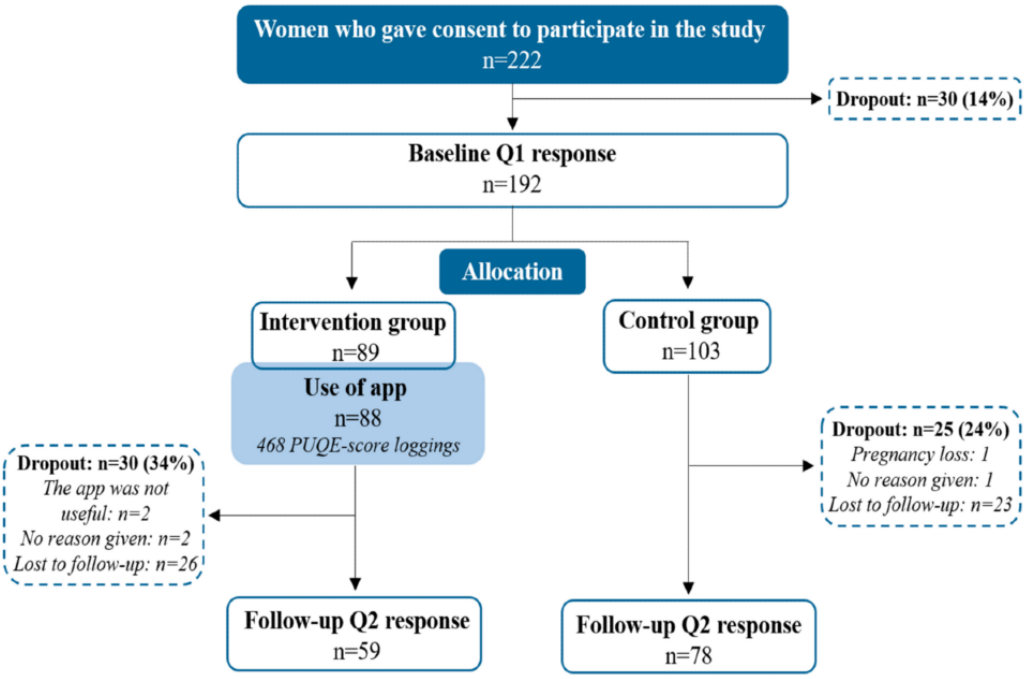


Table 1. Baseline characteristics of the study population (n=192), stratified by whether they used the MinSafeStart (MSS) app (intervention) or received standard maternity care (control).

Characteristics	Intervention group (n=89)	Control group (n=103)
Gestational week at enrollment, median (range)	8 (4-36)	8 (4-39)
Age (years), mean (SD)	32 (4.6)	32 (3.9)
Relationship status, n (%)		
Married/cohabitation	85 (95.5)	100 (97.1)
Other[a]	4 (4.5)	3 (2.9)
Higher education, n (%)		
Yes	69 (77.5)	85 (82.5)
No	20 (22.5)	18 (17.5)
Working situation, n (%)		
Employed	55 (61.8)	60 (58.2)
Employed in the health sector	19 (21.4)	31 (30.1)
Other[b]	15 (16.8)	12 (11.7)
Primigravida, n (%)		
Yes	24 (27.0)	27 (26.2)
No	65 (73.0)	76 (73.8)
NVP[c] during previous pregnancy/pregnancies, n (%)		
Yes	52 (80.0)	61 (80.3)
No	13 (20.0)	15 (19.7)

[a]Includes single/unmarried and divorced/separated women.

[b]Includes students and unemployed women.

[c]NVP: nausea and vomiting during pregnancy.

The Intervention

Of the 89 women randomized to the intervention group, 88 downloaded the MSS app. These women performed a total of 468 logs. Because they were not satisfied with the app, 2 women dropped out of the study. They reported no benefit in using the MSS app.

Impact on NVP Severity

The groups showed no differences in the change in PUQE scores between Q1 and Q2 (adjusted β 0.6, 95% Cl −0.1 to 1.2). Among women employed in the health sector, those who used the MSS app had a significantly higher PUQE score (adjusted β 2.1, 95% Cl 0.9 to 3.2) after 2 weeks than those who did not use the app. However, among women employed in other sectors, the PUQE scores were not significantly different between the intervention and control groups (Table 2).

Table 2. Associations between the use of the MinSafeStart (MSS) app and the Pregnancy Unique Quantification of Emesis (PUQE) score.

Analysis	Baseline (Q1) PUQE score[a], mean (SD)	Follow-up (Q2) PUQE score, mean (SD)	Change in PUQE score (Q2-Q1)		
			Mean change (SD)	Crude difference in mean changes, β (95% CI)	Adjusted difference in mean changes[b], β (95% CI)
Primary analysis					
Intervention group (n=88)	4.9 (2.0)	5.6 (1.8)[c]	0.8 (2.0)	0.4 (−0.3 to 1.2)	0.6 (−0.1 to 1.2)
Control group (n=103)	4.7 (1.9)	4.9 (1.8)[d]	0.4 (2.3)	Reference	Reference
Subanalyses by employment: women employed in the health sector					
Intervention group (n=19)	4.6 (1.9)	6.6 (1.7)[e]	1.8 (2.5)	2.1 (0.3 to 3.9)	2.1 (0.9 to 3.2)
Control group (n=31)	4.5 (1.9)	4.6 (1.6)[f]	−0.3 (2.7)	Reference	Reference
Subanalyses by employment: women employed in other sectors					
Intervention group (n=55)	4.9 (2.1)	5.2 (1.7)[g]	0.4 (1.7)	−0.1 (−0.8 to 0.7)	0.0 (−0.7 to 0.7)
Control group (n=60)	4.7 (1.9)	5.1 (1.8)[h]	0.5 (1.9)	Reference	Reference

[a]This score ranges from 3 to 15 points, and symptoms are rated as follows: mild: ≤6 points; moderate: 7-12 points; severe ≥13 points.

[b]Adjusted for the baseline PUQE score.

[c]n=59.

[d]n=78.

[e]n=14.

[f]n=23.

[g]n=38.

[h]n=45.

Impact on Quality of Life

The adjusted primary analysis showed that the changes in NVPQOL scores from baseline to Q2 were not significantly different between the intervention and control groups (adjusted β −5.3, 95% Cl −12.5 to 1.9; Table 3).

Table 3. Association between the use of the MinSafeStart (MSS) app and quality of life.

Group	Baseline (Q1) NVPQOL[a,b] score, mean (SD)	Follow-up (Q2) NVPQOL score, mean (SD)	Change in NVPQOL score (Q2-Q1)		
			Mean change (SD)	Crude difference in mean changes, β (95% CI)	Adjusted difference in mean changes[c], β (95% CI)
Intervention group (n=88)	145.7 (34.0)	143.8 (29.7)[d]	−4.5 (22.4)	−4.2 (−11.9 to 3.5)	−5.3 (−12.5 to 1.9)
Control group (n=103)	148.5 (28.8)	151.6 (28.9)[e]	−0.3 (22.9)	Reference	Reference

[a]NVPQOL: Health-Related Quality of Life for Nausea and Vomiting during Pregnancy scale.

[b]This score ranges from 30 to 210 points, and lower scores indicate better quality of life.

[c]Adjusted for the baseline NVPQOL score.

[d]n=59.

[e]n=78.

Impact on Decisional Conflict Scale Score

The mean changes in the DCS between Q1 and Q2 were −5.9 (SD 16.4) for the intervention group and −5.3 (SD 15.5) for the control group (Table 4). The changes in DCS were not significantly different between the women in the intervention group and the women in the control group (adjusted β −1.1, 95% Cl −6.2 to 4.2).

Table 4. Association between the use of the MinSafeStart (MSS) app and the decisional conflict scale (DCS).

Group	Baseline (Q1) DCS, mean (SD)		Follow-up (Q2) DCS, mean (SD)	Change in DCS[a] (Q2-Q1)		
				Mean change (SD)	Crude difference in mean changes, β (95% CI)	Adjusted difference in mean changes[b], β (95% CI)
Intervention group (n=88)	88	40.3 (17.9)	36.2 (21.6)[c]	−5.9 (16.4)	−0.7 (−6.1 to 4.7)	−1.1 (−6.2 to 4.2)
Control group (n=103)	103	42.5 (20.9)	38.1 (20.3)[d]	−5.3 (15.5)	Reference	Reference

[a]This score ranges from 0 points (no decisional conflict) to 100 points (extremely high decisional conflict).

[b]Adjusted for the baseline decisional conflict score.

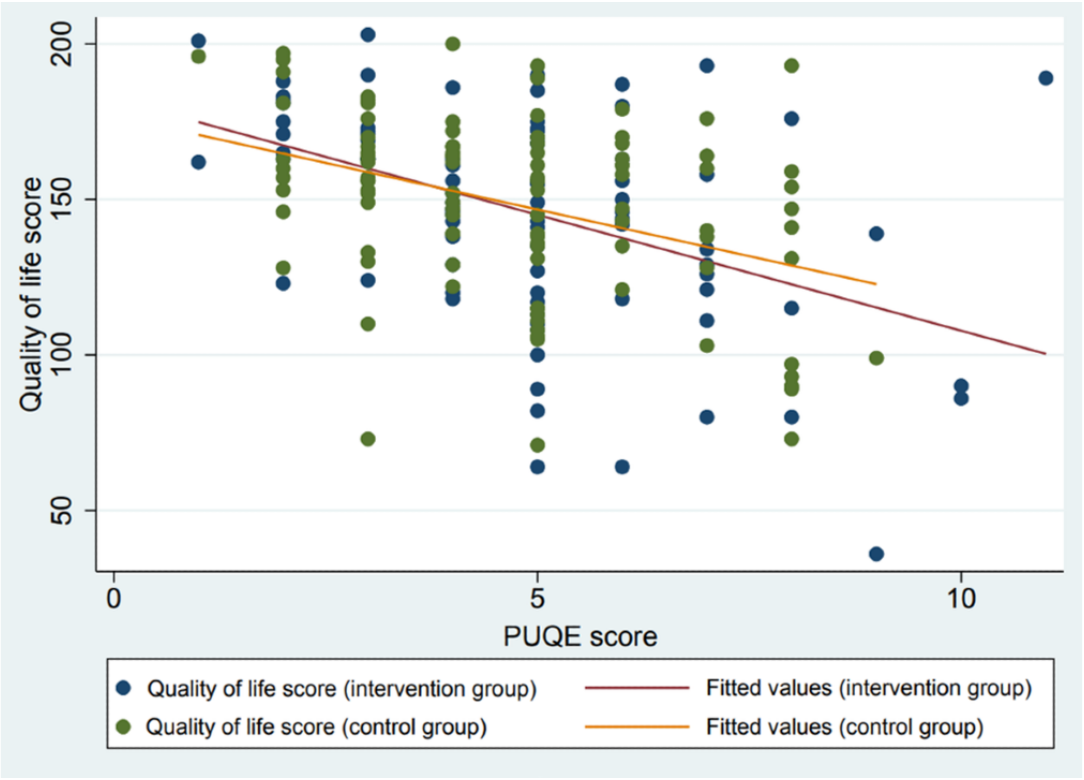
[c]n=59.

[d]n=78.

Association Between NVP Severity and Quality of Life

Women with more severe NVP (higher PUQE scores) had lower NVPQOL scores than women with less severe NVP (lower PUQE scores; Figure 4).

Figure 4. Association between the Health-Related Quality of Life for Nausea and Vomiting during Pregnancy score (NVPQOL) score and the Pregnancy Unique Quantification of Emesis (PUQE) score. MSS app: MinSafeStart mobile application.



Discussion

Main Findings

The MinSafeStart trial was the first to investigate the effectiveness of a patient-centered mobile app that was designed to empower pregnant women to optimally manage their NVP symptoms. We found no significant associations between the use of the MSS app and the severity of NVP symptoms, QoL, or decisional conflict, compared with standard maternal care. These results should be interpreted with caution because the study was slightly underpowered, due to a higher dropout rate than expected.

Earlier studies have shown that the majority of the pregnant population owns a smartphone and over 50% use apps related to pregnancy [36]. Studies that have investigated the use of health-related apps have shown that the apps could improve the knowledge levels of pregnant women and the apps were perceived as tools during pregnancy [7,24]. Except for user satisfaction, our results were not consistent with those from previous studies. We found no associations between the use of the MSS app and NVP symptoms at 2 weeks after baseline. This may be explained by several factors related to our study population and study design. First, we included women at any gestational stage in pregnancy. In fact, 15% of the women included were beyond the first trimester, which is the most relevant time window for NVP. On average, NVP occurs during gestational week 4 [10] and peaks during gestational weeks 10-16 [37,38]. However, our intervention group had completed a median of 8 gestational weeks at enrollment, with a range of 4-36 weeks. Therefore, in many cases, it may have been too late for women to benefit from the app. Moreover, we included women with mild NVP, and this group may not derive the most benefit from the app. Second, a 2-week follow-up may not have been optimal for evaluating the effect of the intervention. The rationale for choosing a 2-week follow-up was based on earlier studies that showed that PUQE scores decreased by 4.7 points when treated within 1 week [39]. We could not exclude the possibility that natural fluctuations in NVP severity could have affected the results or that a shorter follow-up time before the app assessment might have been a better choice. In fact, there might not be a particular time that is optimal for measuring the effects of the app. Indeed, NVP severity varies from morning to evening and from day to day. Therefore, selecting a specific time point for follow-up and reporting the PUQE score in Q2 may not have fully captured the changes in NVP severity over time. Future studies should consider these elements when designing a trial to evaluate the effect of using a digital tool during pregnancy.

Another factor that may have affected the results was that the study included a high proportion of parous women with a prior NVP history. Moreover, most were in a relationship with a partner, which may have provided emotional support. Therefore, these women may have already been informed about optimal NVP management and treatment, and consequently, they may not have felt they needed more information from an NVP tool. Many earlier studies have shown that women with a higher sociodemographic status and women who are pregnant for the first time are more likely to search for information online [40-42]. In their first pregnancy, women often search for information about concerns and symptoms related to the first period of pregnancy [6,40,43-45]. Therefore, our study may not have targeted the appropriate subgroup of pregnant women.

Strengths and Limitations

The main strength of this study was that very few studies have been conducted to assess the effectiveness of mobile apps for disease management among pregnant women. This study provided new insights in this regard. An important strength of this study was the use of the randomized controlled trial study design, which is considered the gold standard in evidence-based medicine [46]. Another strength of this study included our use of the internet for recruitment and electronic data collection. The main benefit of social media recruiting is that it is convenient for sampling. Indeed, pregnant women in their first trimester are not given any routine care, and there is no ideal place to reach out to this group, outside of social media. This approach facilitated the participation of pregnant women all over Norway, which may have increased the representativeness of the study sample and, thus, the generalizability of the results. In addition, the NVPQOL may have provided an advantage over other QoL scales because the NVPQOL is more specific [40].

The major limitation of this study was that we did not reach our targeted number of participants, which was 250 women, including a 25% dropout rate. Furthermore, as in all studies based on voluntary patient recruitment, there might have been a self-selection bias, where more motivated and resourceful women are included in the study compared with the general population. Participants who were parous women with higher sociodemographic status than the general birthing population in Norway might also have contributed to a selection bias. Because these women might have been more informed about optimal NVP management, they might have had less use for the app. We could not exclude the possibility that this selection bias might explain why we did not find any significant beneficial effect of the app on NVP severity in this study.

Last, 15% of the women in the intervention group were beyond the first trimester when the app was introduced. It may have been too late for many of these women to take advantage of the app because NVP often occurs in week 4 [10] and it peaks around weeks 10-16 [37,38].

Future Research

Digitalization and eHealth have provided opportunities to develop innovative apps that support pregnant women. These mobile applications must be tested in clinical studies to establish evidence for health efficacy before they can be included in the health care system or recommended by health care personnel [47]. Our review from 2020, consistent with previous studies [48], demonstrated that decision support tools could potentially provide benefit to pregnant women. However, the tools were mainly useful when relevant information was assembled into one digital tool and when the woman could share her recordings with her health care provider [7]. Based on the results of this study, future research should focus on how to design trials to determine the effect of digital tools on the pregnancy outcomes that are most important to pregnant patients. Future studies should also investigate whether digital tools and apps might be more effective when developed as part of a more extensive health intervention. Specific focus should be placed on how digital tools might facilitate counseling and communication between pregnant women and health care providers regarding NVP management in pregnancy.

Conclusion

This study showed that tracking NVP symptoms with a mobile application was not associated with reduced NVP symptoms, less decisional conflict, or improved QoL after 2 weeks of use. These findings may have been influenced by study design–related factors, such as the gestational week of enrollment, women's parity, time to follow-up, and sample size. Future studies should include a process evaluation to improve our understanding of how pregnant women use the app and how to optimize its utility within maternity care.

Acknowledgments

The authors would like to thank Pål Fugelli, Dagfinn Bergsager, and their team at the University Center for Information Technology (USIT) for the technical development, technical support, and maintenance of the MinSafeStart app. We also thank Hyperemesis Gravidarum Norway, tryggmammamedisin.no, altformamma.no, and Helseoversikt for their contribution to the recruitment for the project and all of the women that participated in the study.

EN is funded by the Dam Foundation, Norwegian Women's Public Health Association.

Authors' Contributions

EN, MBTT, and HN designed the study. EN conducted the main analysis. EN drafted the first version of the manuscript. EN, MBTT, DW, and HN contributed to the interpretation of the results and the critical appraisal of the manuscript. All authors approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Management of nausea and vomiting in pregnancy (NVP), according to treatment guidelines. PUQE= Pregnancy Unique Quantification of Emesis score; this score ranges from 3 to 15 points.

[DOCX File , 104 KB - mhealth_v10i7e36226_app1.docx]

Multimedia Appendix 2
The questions in the PUQE score, NVPQOL scale, and the decisional conflict scale. PUQE= Pregnancy Unique Quantification of Emesis; NVPQOL=.
[DOCX File , 16 KB - mhealth_v10i7e36226_app2.docx]

Multimedia Appendix 3
CONSORT-eHEALTH checklist (V 1.6.1)
[PDF File (Adobe PDF File), 1164 KB - mhealth_v10i7e36226_app3.pdf]

References

1. Hughson JP, Daly JO, Woodward-Kron R, Hajek J, Story D. The rise of pregnancy apps and the implications for culturally and linguistically diverse women: narrative review. JMIR Mhealth Uhealth 2018 Nov 16;6(11):e189 [FREE Full text] [doi: 10.2196/mhealth.9119] [Medline: 30446483]
2. Agarwal P, Gordon D, Griffith J, Kithulegoda N, Witteman HO, Sacha Bhatia R, et al. Assessing the quality of mobile applications in chronic disease management: a scoping review. NPJ Digit Med 2021 Mar 10;4(1):46 [FREE Full text] [doi: 10.1038/s41746-021-00410-x] [Medline: 33692488]
3. Lynch MM, Squiers LB, Kosa KM, Dolina S, Read JG, Broussard CS, et al. Making decisions about medication use during pregnancy: implications for communication strategies. Matern Child Health J 2018 Jan 12;22(1):92-100 [FREE Full text] [doi: 10.1007/s10995-017-2358-0] [Medline: 28900803]
4. Song H, Cramer EM, McRoy S, May A. Information needs, seeking behaviors, and support among low-income expectant women. Women Health 2013 Nov 11;53(8):824-842. [doi: 10.1080/03630242.2013.831019] [Medline: 24215275]
5. Tripp N, Hainey K, Liu A, Poulton A, Peek M, Kim J, et al. An emerging model of maternity care: smartphone, midwife, doctor? Women Birth 2014 Mar;27(1):64-67. [doi: 10.1016/j.wombi.2013.11.001] [Medline: 24295598]
6. Bert F, Gualano MR, Brusaferro S, De Vito E, de Waure C, La Torre G, et al. Pregnancy e-health: a multicenter Italian cross-sectional study on Internet use and decision-making among pregnant women. J Epidemiol Community Health 2013 Dec 01;67(12):1013-1018. [doi: 10.1136/jech-2013-202584] [Medline: 24072743]
7. Ngo E, Truong MB, Nordeng H. Use of decision support tools to empower pregnant women: systematic review. J Med Internet Res 2020 Sep 14;22(9):e19436 [FREE Full text] [doi: 10.2196/19436] [Medline: 32924961]
8. Austin K, Wilson K, Saha S. Hyperemesis gravidarum. Nutr Clin Pract 2019 Apr 18;34(2):226-241. [doi: 10.1002/ncp.10205] [Medline: 30334272]
9. Einarson T, Piwko C, Koren G. Quantifying the global rates of nausea and vomiting of pregnancy: a meta analysis. J Popul Ther Clin Pharmacol 2013;20(2):e171-e183. [Medline: 23863575]
10. Bustos M, Venkataramanan R, Caritis S. Nausea and vomiting of pregnancy - What's new? Auton Neurosci 2017 Jan;202:62-72 [FREE Full text] [doi: 10.1016/j.autneu.2016.05.002] [Medline: 27209471]
11. Heitmann K, Nordeng H, Havnen GC, Solheimsnes A, Holst L. The burden of nausea and vomiting during pregnancy: severe impacts on quality of life, daily life functioning and willingness to become pregnant again - results from a cross-sectional study. BMC Pregnancy Childbirth 2017 Feb 28;17(1):75 [FREE Full text] [doi: 10.1186/s12884-017-1249-0] [Medline: 28241811]
12. Tan A, Lowe S, Henry A. Nausea and vomiting of pregnancy: Effects on quality of life and day-to-day function. Aust N Z J Obstet Gynaecol 2018 Jun 26;58(3):278-290. [doi: 10.1111/ajo.12714] [Medline: 28949009]
13. Piwko C, Ungar WJ, Einarson TR, Wolpin J, Koren G. The weekly cost of nausea and vomiting of pregnancy for women calling the Toronto Motherisk Program. Curr Med Res Opin 2007 Apr 06;23(4):833-840. [doi: 10.1185/030079907x178739] [Medline: 17407640]
14. Clark SM, Costantine MM, Hankins GDV. Review of NVP and HG and early pharmacotherapeutic intervention. Obstet Gynecol Int 2012;2012:252676-252678 [FREE Full text] [doi: 10.1155/2012/252676] [Medline: 22190950]
15. Committee on Practice Bulletins-Obstetrics. ACOG practice bulletin No. 189: nausea and vomiting of pregnancy. Obstet Gynecol 2018 Jan;131(1):e15-e30. [doi: 10.1097/AOG.0000000000002456] [Medline: 29266076]
16. Shehmar M, Maclean MA, Nelson-Piercy C, Gadsby R, O'Hara M. The Management of Nausea and Vomiting of Pregnancy and Hyperemesis Gravidarum (Green-top Guideline No.69). Royal College of Obstetricians & Gynaecologists. 2016. URL: https://www.rcog.org.uk/globalassets/documents/guidelines/green-top-guidelines/gtg69-hyperemesis.pdf [accessed 2021-12-23]
17. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
18. Iyawa GE, Dansharif AR, Khan A. Mobile apps for self-management in pregnancy: a systematic review. Health Technol 2021 Feb 04;11(2):283-294. [doi: 10.1007/s12553-021-00523-z]

19. Koren G, Piwko C, Ahn E, Boskovic R, Maltepe C, Einarson A, et al. Validation studies of the Pregnancy Unique-Quantification of Emesis (PUQE) scores. J Obstet Gynaecol 2005 Apr 02;25(3):241-244. [doi: 10.1080/01443610500060651] [Medline: 16147725]
20. Koren G, Cohen R. Measuring the severity of nausea and vomiting of pregnancy; a 20-year perspective on the use of the pregnancy-unique quantification of emesis (PUQE). J Obstet Gynaecol 2021 Apr 19;41(3):335-339. [doi: 10.1080/01443615.2020.1787968] [Medline: 32811235]
21. Birkeland E, Stokke G, Tangvik RJ, Torkildsen EA, Boateng J, Wollen AL, et al. Norwegian PUQE (Pregnancy-Unique Quantification of Emesis and nausea) identifies patients with hyperemesis gravidarum and poor nutritional intake: a prospective cohort validation study. PLoS One 2015 Apr 1;10(4):e0119962 [FREE Full text] [doi: 10.1371/journal.pone.0119962] [Medline: 25830549]
22. Bruk av IKT i husholdningene. Statistics Norway. URL: https://www.ssb.no/statbank/table/12344/tableViewLayout1/ [accessed 2021-08-30]
23. Ford EA, Roman SD, McLaughlin EA, Beckett EL, Sutherland JM. The association between reproductive health smartphone applications and fertility knowledge of Australian women. BMC Womens Health 2020 Mar 04;20(1):45 [FREE Full text] [doi: 10.1186/s12905-020-00912-y] [Medline: 32131809]
24. Wang N, Deng Z, Wen LM, Ding Y, He G. Understanding the use of smartphone apps for health information among pregnant Chinese women: mixed methods study. JMIR Mhealth Uhealth 2019 Jun 18;7(6):e12631 [FREE Full text] [doi: 10.2196/12631] [Medline: 31215516]
25. Wang K, Varma DS, Prosperi M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. J Psychiatr Res 2018 Dec;107:73-78. [doi: 10.1016/j.jpsychires.2018.10.006] [Medline: 30347316]
26. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct 09;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
27. About TSD. University of Oslo. URL: https://www.uio.no/english/services/it/research/sensitive-data/about/index.html [accessed 2022-03-22]
28. Pregnancy consultations. The Norwegian Directorate of Health. URL: https://www.helsenorge.no/en/pregnancy-and-maternity-care-in-norway/antenatal-checks-and-tests/ [accessed 2022-04-04]
29. Smith JA, Fox KA, Clark SM. Patient education: Nausea and vomiting of pregnancy (Beyond the Basics). UpToDate. URL: https://www.uptodate.com/contents/nausea-and-vomiting-of-pregnancy-beyond-the-basics [accessed 2021-12-06]
30. Lacasse A, Bérard A. Validation of the nausea and vomiting of pregnancy specific health related quality of life questionnaire. Health Qual Life Outcomes 2008 May 09;6(1):32 [FREE Full text] [doi: 10.1186/1477-7525-6-32] [Medline: 18471301]
31. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making 2016 Jul 02;15(1):25-30. [doi: 10.1177/0272989x9501500105]
32. Garvelink MM, Boland L, Klein K, Nguyen DV, Menear M, Bekker HL, et al. Decisional conflict scale use over 20 years: the anniversary review. Med Decis Making 2019 May 29;39(4):301-314. [doi: 10.1177/0272989x19851345]
33. Walton GD, Ross LE, Stewart DE, Grigoriadis S, Dennis C, Vigod S. Decisional conflict among women considering antidepressant medication use in pregnancy. Arch Womens Ment Health 2014 Dec 9;17(6):493-501. [doi: 10.1007/s00737-014-0448-1] [Medline: 25104244]
34. Hadizadeh-Talasaz F, Ghoreyshi F, Mohammadzadeh F, Rahmani R. Effect of shared decision making on mode of delivery and decisional conflict and regret in pregnant women with previous cesarean section: a randomized clinical trial. BMC Pregnancy Childbirth 2021 Feb 17;21(1):144 [FREE Full text] [doi: 10.1186/s12884-021-03615-w] [Medline: 33596854]
35. Kahan BC, Jairath V, Doré CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. Trials 2014 Apr 23;15(1):139 [FREE Full text] [doi: 10.1186/1745-6215-15-139] [Medline: 24755011]
36. Frid G, Bogaert K, Chen KT. Mobile health apps for pregnant women: systematic search, evaluation, and analysis of features. J Med Internet Res 2021 Oct 18;23(10):e25667 [FREE Full text] [doi: 10.2196/25667] [Medline: 34524100]
37. Lee NM, Saha S. Nausea and vomiting of pregnancy. Gastroenterol Clin North Am 2011 Jun;40(2):309-34, vii [FREE Full text] [doi: 10.1016/j.gtc.2011.03.009] [Medline: 21601782]
38. Furneaux EC, Langley-Evans AJ, Langley-Evans SC. Nausea and vomiting of pregnancy: endocrine basis and contribution to pregnancy outcome. Obstet Gynecol Surv 2001 Dec;56(12):775-782. [doi: 10.1097/00006254-200112000-00004] [Medline: 11753180]
39. Lehmann AS, Renbarger JL, McCormick CL, Topletz AR, Rouse C, Haas DM. Pharmacogenetic predictors of nausea and vomiting of pregnancy severity and response to antiemetic therapy: a pilot study. BMC Pregnancy Childbirth 2013 Jun 20;13(1):132 [FREE Full text] [doi: 10.1186/1471-2393-13-132] [Medline: 23786674]
40. Kavlak O, Atan US, Güleç D, Oztürk R, Atay N. Pregnant women's use of the internet in relation to their pregnancy in Izmir, Turkey. Inform Health Soc Care 2012 Dec 07;37(4):253-263. [doi: 10.3109/17538157.2012.710686] [Medline: 22958087]

41. Sayakhot P, Carolan-Olah M. Internet use by pregnant women seeking pregnancy-related information: a systematic review. BMC Pregnancy Childbirth 2016 Mar 28;16(1):65 [FREE Full text] [doi: 10.1186/s12884-016-0856-5] [Medline: 27021727]
42. Kamali S, Ahmadian L, Khajouei R, Bahaadinbeigy K. Health information needs of pregnant women: information sources, motives and barriers. Health Info Libr J 2018 Mar 13;35(1):24-37 [FREE Full text] [doi: 10.1111/hir.12200] [Medline: 29131537]
43. Lagan BM, Sinclair M, Kernohan WG. Internet use in pregnancy informs women's decision making: a web-based survey. Birth 2010 Jun;37(2):106-115. [doi: 10.1111/j.1523-536X.2010.00390.x] [Medline: 20557533]
44. Larsson M. A descriptive study of the use of the Internet by women seeking pregnancy-related information. Midwifery 2009 Feb;25(1):14-20. [doi: 10.1016/j.midw.2007.01.010] [Medline: 17408822]
45. Bakhireva LN, Young BN, Dalen J, Phelan ST, Rayburn WF. Patient utilization of information sources about safety of medications during pregnancy. J Reprod Med 2011;56(7-8):339-343. [Medline: 21838165]
46. Hariton E, Locascio JJ. Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. BJOG 2018 Dec 19;125(13):1716 [FREE Full text] [doi: 10.1111/1471-0528.15199] [Medline: 29916205]
47. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. Transl Behav Med 2014 Dec 24;4(4):363-371 [FREE Full text] [doi: 10.1007/s13142-014-0293-9] [Medline: 25584085]
48. Whybrow R, Webster LM, Seed PT, Sandall J, Chappell LC. The effectiveness of decision aids for pregnancy related decision-making in women with pre-pregnancy morbidity; systematic review and meta-analysis. BMC Pregnancy Childbirth 2022 Jan 29;22(1):81 [FREE Full text] [doi: 10.1186/s12884-022-04402-x] [Medline: 35093017]

Abbreviations

DCS: decisional conflict scale
HG: hyperemesis gravidarum
MSS: MinSafeStart
NVP: nausea and vomiting in pregnancy
NVPQOL: Health-Related Quality of Life for Nausea and Vomiting during Pregnancy scale
PUQE: Pregnancy Unique Quantification of Emesis
Q1: Questionnaire 1
Q2: Questionnaire 2
Q3: Questionnaire 3
Q4: Questionnaire 4
QoL: quality of life
TSD: Service for Sensitive Data at the University of Oslo
USIT: University Center for Information Technology

Edited by L Buis; submitted 06.01.22; peer-reviewed by B Nievas Soriano, C Jacob; comments to author 24.02.22; revised version received 11.04.22; accepted 19.04.22; published 05.07.22.

Please cite as:
Ngo E, Truong MBT, Wright D, Nordeng H
Impact of a Mobile Application for Tracking Nausea and Vomiting During Pregnancy (NVP) on NVP Symptoms, Quality of Life, and Decisional Conflict Regarding NVP Treatments: MinSafeStart Randomized Controlled Trial
JMIR Mhealth Uhealth 2022;10(7):e36226
URL: https://mhealth.jmir.org/2022/7/e36226
doi:10.2196/36226
PMID:35787487

Original Paper

The Efficacy of "Foundations," a Digital Mental Health App to Improve Mental Well-being During COVID-19: Proof-of-Principle Randomized Controlled Trial

Silvina Catuara-Solarz[1], PhD; Bartlomiej Skorulski[1], PhD; Iñaki Estella-Aguerri[1], PhD; Claudia Bibiana Avella-Garcia[1], PhD; Sarah Shepherd[1], MA; Emily Stott[1], BA; Nicola R Hemmings[1], MSc; Aleix Ruiz de Villa[1], PhD; Laura Schulze[1], PhD; Sophie Dix[1], PhD

Koa Health, Barcelona, Spain

Corresponding Author:
Sophie Dix, PhD
Koa Health
Carrer de la Ciutat de Granada 121
Barcelona, 08018
Spain
Phone: 44 7702 500882
Email: sophie.dix@koahealth.com

Abstract

Background: Against a long-term trend of increasing demand, the COVID-19 pandemic has led to a global rise in common mental disorders. Now more than ever, there is an urgent need for scalable, evidence-based interventions to support mental well-being.

Objective: The aim of this proof-of-principle study was to evaluate the efficacy of a mobile-based app in adults with self-reported symptoms of anxiety and stress in a randomized control trial that took place during the first wave of the COVID-19 pandemic in the United Kingdom.

Methods: Adults with mild to severe anxiety and moderate to high levels of perceived stress were randomized to either the intervention or control arm. Participants in the intervention arm were given access to the Foundations app for the duration of the 4-week study. All participants were required to self-report a range of validated measures of mental well-being (10-item Connor-Davidson Resilience scale [CD-RISC-10], 7-item Generalized Anxiety Disorder scale [GAD-7], Office of National Statistics Four Subjective Well-being Questions [ONS-4], World Health Organization-5 Well-Being Index [WHO-5]) and sleep (Minimal Insomnia Scale [MISS]) at baseline and at weeks 2 and 4. The self-reported measures of perceived stress (10-item Perceived Stress Score [PSS-10]) were obtained weekly.

Results: A total of 136 participants completed the study and were included in the final analysis. The intervention group (n=62) showed significant improvements compared to the control group (n=74) on measures of anxiety, with a mean GAD-7 score change from baseline of –1.35 (SD 4.43) and –0.23 (SD 3.24), respectively (t_{134}=1.71, P=.04); resilience, with a mean change in CD-RISC score of 1.79 (SD 4.08) and –0.31 (SD 3.16), respectively (t_{134}=–3.37, P<.001); sleep, with a mean MISS score change of –1.16 (SD 2.67) and –0.26 (SD 2.29), respectively (t_{134}=2.13, P=.01); and mental well-being, with a mean WHO-5 score change of 1.53 (SD 5.30) and –0.23 (SD 4.20), respectively (t_{134}=–2.16, P=.02), within 2 weeks of using Foundations, with further improvements emerging at week 4. Perceived stress was also reduced within the intervention group, although the difference did not reach statistical significance relative to the control group, with a PSS score change from baseline to week 2 of –2.94 (SD 6.84) and –2.05 (SD 5.34), respectively (t_{134}= 0.84, P=.20).

Conclusions: This study provides a proof of principle that the digital mental health app Foundations can improve measures of mental well-being, anxiety, resilience, and sleep within 2 weeks of use, with greater effects after 4 weeks. Foundations therefore offers potential as a scalable, cost-effective, and accessible solution to enhance mental well-being, even during times of crisis such as the COVID-19 pandemic.

Trial Registration: OSF Registries osf.io/f6djb; https://osf.io/vm3xq

(JMIR Mhealth Uhealth 2022;10(7):e30976) doi:10.2196/30976

KEYWORDS

mental well-being; digital health; cognitive behavioral therapy; positive psychology; insomnia; COVID-19; mental health; mobile app; anxiety; health app

Introduction

Background

Mental illness is a highly prevalent and complex public health issue. The total number of people with any mental health disorder reached 792 million in 2017 [1]. Moreover, according to the World Health Organization, the number of people with common mental disorders (CMD) such as mild to severe depression and anxiety is globally increasing over time [2]. These figures are worrying since mental health conditions account for a greater burden of disease based on years lived with disability [3] and they also are the largest cost driver in health care, estimated to reach over US $2.2 trillion and to rise to nearly US $6 trillion by 2030 [4].

Within the context of the COVID-19 pandemic, the prospects of the mental health status of society have become even more concerning. Numerous reports have highlighted that during the COVID-19 pandemic, poor mental health has been exacerbated globally [5,6]. In the United Kingdom, the population prevalence of clinically significant levels of mental distress rose from 18.9% (2018-2019) to 27.3% in April 2020 [7]. A recent study in the United States showed a 3-fold increase in depressive symptoms during the COVID-19 epidemic compared with the previous rate [8]. More recently, a meta-analysis of 66 studies with 221,970 participants reported an overall pooled prevalence of depression, anxiety, distress, and insomnia of 31.4%, 31.9%, 41.1%, and 37.9%, respectively [9]. Taken together, the magnitude of the socioeconomic burden of CMD illustrates that the current model of mental health care has yet to be fully optimized. Thus, better, more easily scalable interventions are urgently needed.

Digital technologies have shown great potential to offer scalable, easy-to-access, and timely solutions to increase the delivery of psychotherapeutic interventions and evidence-based recommendations for self-care and self-management [10,11]. During the COVID-19 pandemic, demand for these technologies has increased on a global scale and the use of mental health apps has risen exponentially. Although some apps have been shown to have positive effects on mental well-being [10,12,13], concerns remain regarding the level of credibility and the robustness of evidence underlying the majority of the thousands of available apps on the market [10,14]. More efforts toward the proper demonstration of their efficacy are needed if digital apps are to offer an adjunctive method to reduce the prevalence and impact of CMD across the world, and to support the growing mental health crisis surrounding the COVID-19 pandemic.

Objectives

The aim of this 4-week proof-of-principle study was to test the efficacy of a digital intervention delivered via a mobile app, named Foundations, in comparison to a no-intervention control group on a range of psychological measures, including anxiety, mental well-being, resilience, sleep, and stress. More specifically, we aimed to assess the efficacy of Foundations in improving mental well-being. The study took place during the months of April and May of 2020 throughout the first outbreak of the COVID-19 pandemic in the United Kingdom.

Foundations includes a plethora of interventions and psychoeducational content that are scientifically robust (ie, cognitive behavioral therapy [CBT], meditation). Based on this, in combination with preliminary user research, we hypothesized that participants in the intervention group would show significant improvements compared with the control group in the following targeted areas of well-being after 4 weeks: (1) anxiety, (2) mental well-being, (3) resilience, (4) sleep, and (5) stress.

Methods

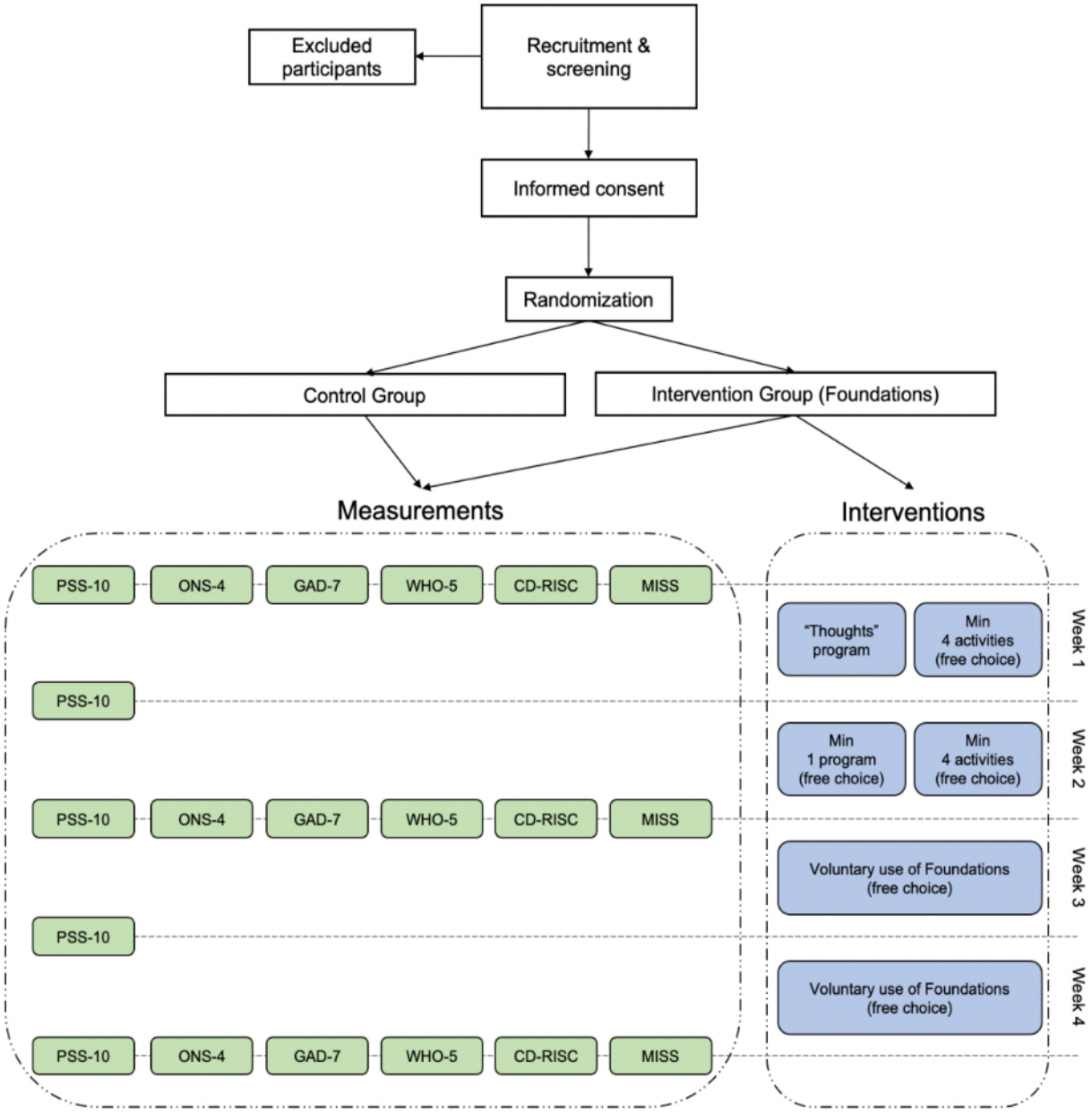
Ethics Approval

This study protocol involving human participants was General Data Protection Regulation–compliant and developed in accordance with Alpha Health's ethics framework and principles [15], which are compliant with the World Medical Association's Declaration of Helsinki [16]. All participants provided informed, electronic consent to share their data, including for publication of the results, prior to their enrollment in the study. Due to unforeseen logistical reasons, the study was not preregistered; however, all statistical testing was performed by a data analyst who remained blind to treatment assignment. Furthermore, the authors confirm that all ongoing studies for the Foundations app are registered.

Trial Design

This proof-of-concept study was a 2-armed randomized controlled trial (RCT) comparing an app-based intervention (Foundations) to a nonintervention control group. At the time of the study, the app was named Evermind. It was rebranded in November 2020 and we refer to it as its current name (Foundations) throughout this paper. Anxiety, sleep, resilience, and mental well-being were assessed at baseline, and at weeks 2 and 4 of the study, and perceived stress was assessed weekly (at baseline through weeks 1 and 4; Figure 1). The study started on April 22, 2020, and ended on May 20, 2020. Of note, at the start of the study, the United Kingdom was at the first peak of the COVID-19 pandemic, and the reported numbers of positive cases and deaths were in decline at the end of the study [17].

Figure 1. Overview of the study design. PSS-10: 10-item Perceived Stress Score; ONS-4: Office of National Statistics Four Subjective Well-being Questions; GAD-7: 7-item Generalized Anxiety Disorder scale; WHO-5: World Health Organization-5 Well-Being Index; CD-RISC: Connor-Davidson Resilience Scale; MISS: Minimal Insomnia Scale.



Participants

Recruitment

Study participants were recruited through a research company specializing in study recruitment between the first 2 weeks of April, according to a screening based on the following predefined criteria for inclusion and exclusion. Eligible participants were between 30 and 50 years old, owned a smartphone and regularly used apps, were fluent in English, and had been employed for at least 3 months in the United Kingdom. In addition, participants were considered eligible for the study if they showed moderate to high levels of perceived stress (10-item Perceived Stress Score [PSS-10]>13 [18,19]), mild to severe anxiety (7-item Generalized Anxiety Disorder scale [GAD-7] score of 5-18 [20]), and no to moderate sleep problems (Minimal Insomnia Scale [MISS] score of 0-8 [21]).

Candidates were excluded from the study in case of current pregnancy, current high-stress event (eg, family bereavement with the exception of a COVID-19 crisis), current psychotherapeutic treatment or counselling, current diagnosis of psychiatric illness (with the exception of depression) by a specialist/secondary care, regular use of mental health apps, or a recent (3 months) change in medication for mood disorders.

Assessment of Eligibility and Randomization

Eligible participants were randomly allocated (1:1) to either the intervention group or control group. Assignment of participants to the groups was performed with a computer-based algorithm (Python Software Foundation) that generated randomly

permuted blocks, which were stratified by gender (male or female) and age (30-40 years or 40-50 years), and were balanced regarding the degree of perceived stress ("moderate perceived stress," PSS-10 score of 13-26, or "high perceived stress," PSS-10 score of 27-40) and sleep disturbances ("no sleep disturbances," MISS score of 0-4, or "moderate sleep disturbances," MISS score of 5-8). Furthermore, participants' baseline scores on the GAD-7 were analyzed for statistical differences and were rerandomized if needed. The randomization analysis was performed by a statistician who remained blinded to the identity of the groups.

Study Conditions

Intervention Group

After randomization, the intervention group received access and instructions to download the Foundations app. The intervention app (Foundations) comprised interventions and psychoeducational content aimed at decreasing stress and promoting mental well-being (Figure 2). Content was organized into activities (ie, units of content) and could take a variety of formats, which are listed in Table 1. They all have a brief (typically 1-2 sentences) introduction and end with a closing sentence. Activities were either *in the moment* or part of a *program*. A *program* is a locked sequence of activities that is delivered in daily steps designed to teach a skill such as healthy sleep behaviors, positive psychology, working with thoughts, or relaxation techniques (see Table 2). *In the moment* activities were not part of a locked sequence of activities and could be accessed at any time. These activities included sleep meditations, articles, and mindfulness. Programs and activities were organized into a library of themed modules.

Figure 2. Screenshots of the Foundations app.

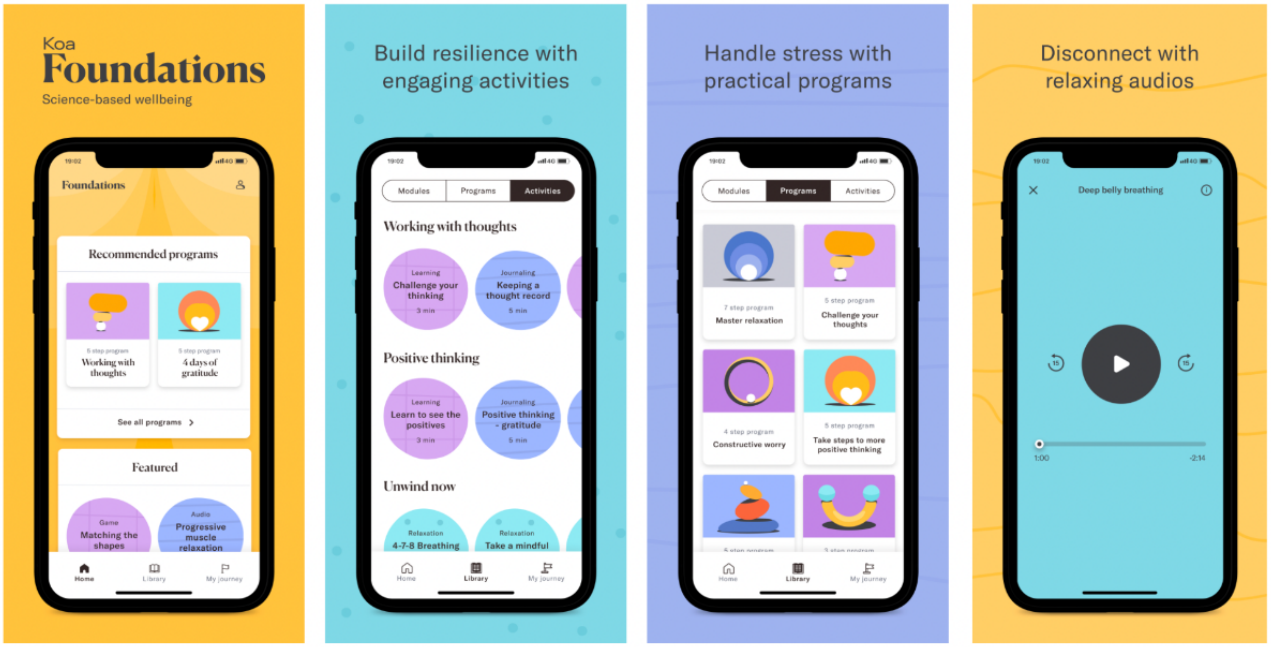


Table 1. Activity formats and descriptions of the content delivered in Foundations.

Activity	Number of activities	Type of content	Description
Slides	12	Psychoeducation	Comprises individual screens of 1-2 sentences that the user swipes through. Usually 10-20 screens
Article/blog	13	Psychoeducation and tips	User scrolls through to read. Typically 0.25-1 of A4 page in length
Add record	10	Journaling/reflection	User can add free text (eg, thought records or gratitude journaling)
Label record	5	Journaling/reflection	User selects a record created by the "add record" feature and can choose a theme label
Question record	5	Journaling/reflection	User selects a previous record and is asked a series of questions about the record. There is a free-text box for the user to write in
Record review	4	Journaling/reflection	Log of record entries
Audios	17	Mindfulness/meditations	Mindful meditations (5-8 min), sleep meditations (30 min), relaxation techniques
Ambient sounds	8	Ambient sounds	30-min relaxation sounds and soundscapes (eg, waves, rain)
Quiz	6	Interactive psychoeducation	Reinforces psychoeducation with two-choice answers
Game	1	Spatial working memory game	The user has to recall spatial sequences

Table 2. Description of programs and the number of activities within each program.

Program name	Number of activities	Description
Become a breathing master	3	Teaches the skill of diaphragmatic breathing
Relax your body and mind	2	Teaches the skill of progressive muscle relaxation
Working with thoughts	5	CBT[a]-based journaling and reflection. Includes psychoeducation on cognitive distortions and questions to balance unhelpful thoughts
Positive thinking	5	Gratitude journaling on achievements, reasons to be thankful, and people
Healthy sleep habits	9	CBT psychoeducation on healthy sleep habits and sleep hygiene
Break your bad sleep habits	4	CBT interactive psychoeducation on breaking bad habits
Take control of your sleep	9	CBT sleep scheduling
Constructive worry for sleep	4	Introduces constructive worry to put worries aside before bed
3 days to improve your self-esteem	10	CBT-based journaling and psychoeducation on automatic thoughts and balancing thoughts about oneself
Boost your confidence	3	Identify strengths

[a]CBT: cognitive behavioral therapy.

The study lasted 4 weeks and all participants started the study at the same time. During the first 2 weeks of the study, the participants were required to perform a minimum number of activities and programs. In week 1, participants were instructed to complete the "Working with thoughts" program, consisting of journaling and CBT interventions focused on ameliorating cognitive distortions and unhelpful ruminative patterns, and a minimum of 4 free-choice activities. During week 2, participants had to complete 1 program of their choosing and a minimum of 4 other activities (free choice). During weeks 3 and 4, participants were free to use the app as little or as much as they wished with complete free choice of activities and programs.

For the first 2 weeks of the research period, participants were encouraged to use Foundations via daily text messages. The messages were written and sent by the research manager via the messaging service WhatsApp and included wording such as:

> *Good morning! If you were able to start the programme yesterday then please do another activity of your choice from the Library page. If you weren't able to start the programme yesterday, please do the first day of the Working with thoughts programme.*
>
> *Remember, if you need to contact us at any point, the best way to get in touch is via email at: hello@evermind.health.*

Control Group

The control group completed the same questionnaires at the same time points as the intervention group. At the end of the study, all participants were provided optional access to the Foundations app (waitlist control condition).

Incentives

At the end of the study, each participant received a monetary incentive as compensation for their involvement in the trial. As compensation was tied to the participants' time, participants in the control group received £35 (US $50) and those in the intervention group received £85 (US $115).

Outcome Measures

Overview

All participants were invited to fill out questionnaires via an online platform (Google Form) to assess their mental well-being on a weekly basis starting with the first day of the study (baseline). All participants, regardless of group, were sent the link to the online platform via the messaging service WhatsApp at the same time. Perceived stress was assessed weekly, whereas all other measures were assessed every 2 weeks (ie, baseline, 2 weeks, and 4 weeks). All outcome measures were treated as continuous variables. Each measure and the associated questionnaire are detailed below.

Anxiety

Anxiety levels were assessed through the GAD-7 scale [20], a 7-item questionnaire that measures the severity of the subject's anxiety over the previous week.

Sleep Problems

The MISS was used to examine sleep problems in the sample and their evolution across the study [21]. The MISS includes 3 items that cover issues of initiating sleep, waking up in the night, and not feeling refreshed in the morning.

Resilience

Resilience levels were assessed by the Connor-Davidson Resilience Scale (CD-RISC-10), which includes 10 items that assess the individual's ability to cope successfully with adversity [22].

Mental Well-being

To assess current mental well-being, the World Health Organization-5 Well-Being Index (WHO-5) questionnaire was administered to the study participants. Each of the 5 items is scored with a Likert scale ranging from 0 to 5 (at no time, some of the time, less than half of the time, more than half of the time, most of the time, all of the time, respectively) [23].

In addition, the United Kingdom Office of National Statistics questions on well-being (ONS-4) scale was used to measure subjective well-being [24]. Each of the questions in this scale is aimed at measuring a different aspect of well-being: life satisfaction, worthwhileness, happiness, and anxiety, which are each rated by the subject from 0 (not at all) to 10 (completely). These questions are not designed to provide an aggregate score, but rather to illustrate different aspects of perceived well-being.

Perceived Stress

The degree to which participants perceived their life situations as stressful was assessed using the PSS-10 [18,19]. The time frame selected for questions was the past week, which enabled examination of weekly effects of the intervention. It should be noted that participants were able to access the PSS-10 within the app whenever they liked. These additional data were not evaluated as part of the study.

Statistical Analyses

Power

A power analysis based on published data and previous pilot studies [25] was performed (using the "pwr" R package) to estimate the required sample size for the study. The estimated sample size was at least 78 participants in each arm providing 0.8 power to detect an effect size of Cohen d=0.4 with an α of .05.

Data Analysis

Two principal sets of analyses were performed on each of the outcome measures. The first set of analyses, which we denote by *within-group analyses*, sought to determine whether there was a significant change within each group compared to the measure at the start of the study (baseline). Within-group paired two-tailed t tests were used for pre-post intervention assessments for each group. Statistical significance was set at P<.05. Bonferroni correction for multiple comparisons was performed adjusting the significance level to 1.25% for PSS-10 (significance level of 5% divided by 4 measures in time, baseline to weeks 1-4) and 2.5% (significance level of 5% divided by 2 measures in time, baseline to week 4) for the rest of the outcome measures.

The second set of analyses, denoted as *between-group analyses*, sought to determine whether the change from baseline (Δ) was equivalent in both groups. The analysis was performed using linear mixed models (LMMs) incorporating group (intervention or control), time, and group×time interaction terms, including a change score of 0 at the baseline time point and modeling participant as a random effect. Significance of the group variable was assessed using the likelihood ratio test. A confirmatory set of analyses was performed using an independent two-tailed t test on the differences of each group's scores in each measure at a given time point from their baseline scores (Δ). Statistical significance was set at P<.05.

Results

Participants

From April 2020 to May 2020, 190 participants were enrolled in the study and randomized to either the intervention group (n=95) or control group (n=95). Of the 95 participants in the intervention group, 7 failed to complete the study (ie, did not use the app as required) and were excluded from the primary analysis. A further 9 participants (5 in the intervention group and 4 in the control group) were excluded due to missing data (failed to complete the outcome measure questionnaires). An additional 38 participants (21 in the intervention group and 17 in the control group) were excluded from the analysis due to a calculation error of the PSS-10 at screening (PSS-10 score<13). Due to the study's single-blind design, this error was not identified until after the study was completed. Of the remaining participants, 74 (79%) from the control group and 62 (65%) from the intervention group completed the study and were

analyzed for the primary outcome and secondary outcome measures (Figure 3).

Table 3 provides the baseline demographics of the study participants (N=136). Participants in the intervention (n=62) and control (n=74) groups did not differ significantly with respect to gender, age, nor their levels of mental well-being at baseline.

The values for the within-group and between-group analyses at different points in time for each of the outcome measures are shown in Table 4 and Table 5, respectively. In the following, we discuss the results for each outcome metric as well as for the subgroup analysis and engagement. As described in the Methods section, for within-group analyses, significance thresholds were adjusted to account for multiple comparisons (Bonferroni correction) using a significance threshold of α=.025 for all scales excluding the PSS-10, which had a significance threshold of α=.0125.

Figure 3. CONSORT (Consolidated Standards of Reporting Trials) flowchart of participants. PSS-10: 10-item Perceived Stress Score.

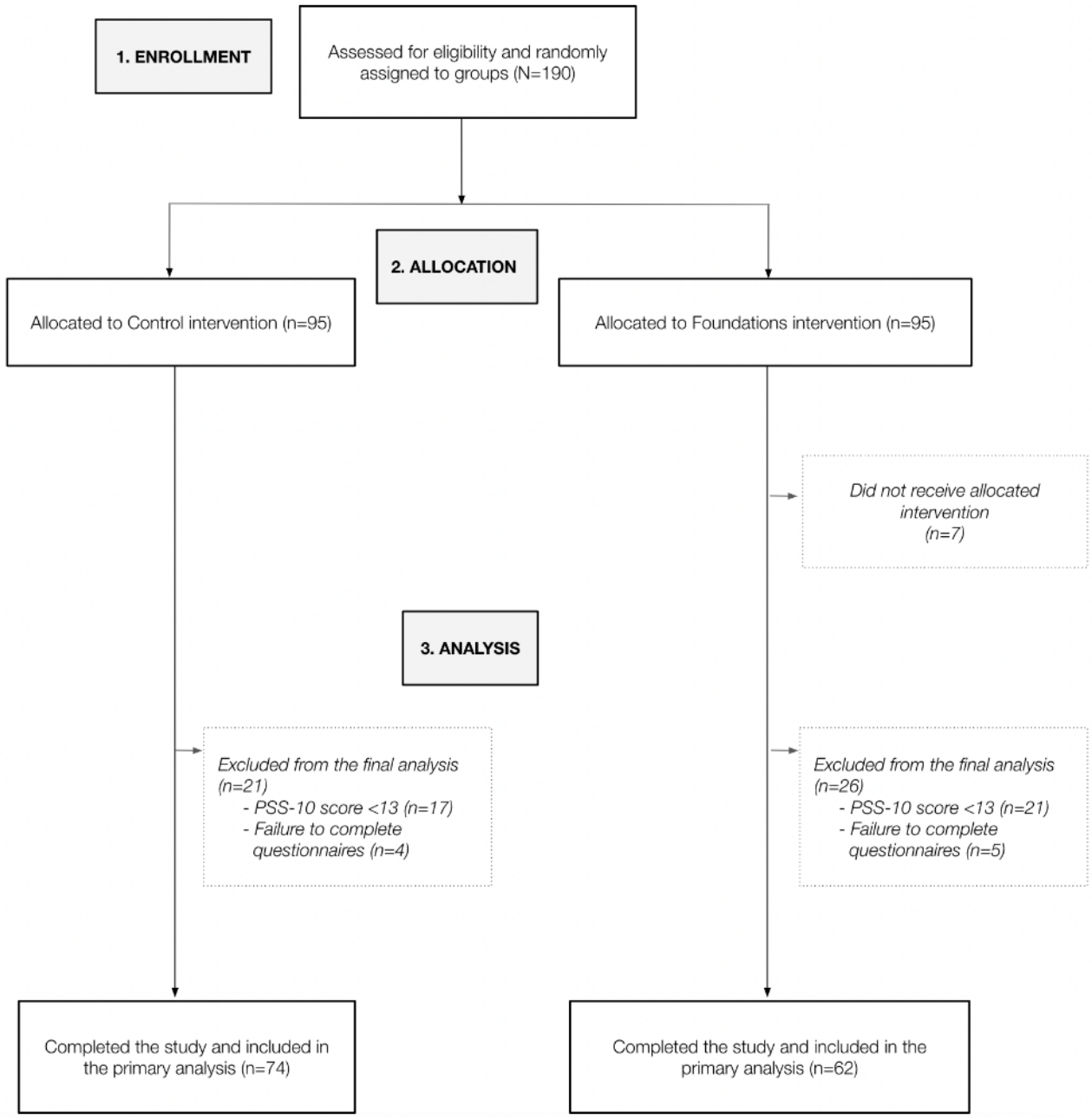


Table 3. Characteristics of study participants at baseline.

Characteristic	Foundation group (n=62)	Control group (n=74)	Effect size (*d*)	P value	Clinical interpretation
Gender					
Females, n	33	40	N/A[a]	.32	N/A
Males, n	29	34	N/A	.48	N/A
Age (years), mean (SD)	40.58 (6.08)	39.49 (6.13)	N/A	.30	N/A
PSS-10[b], mean (SD)	19.19 (4.12)	19.05 (3.58)	–0.04 (n[c])	0.58	moderate stress
GAD-7[d], mean (SD)	8.06 (3.83)	6.91 (3.28)	–0.32 (small)	0.97	mild anxiety
WHO-5[e], mean (SD)	11.35 (4.29)	12 (4.23)	0.15 (n)	0.81	poor well-being
ONS-4[f], mean (SD)	22.63 (5.14)	23.64 (5.82)	0.18 (n)	0.85	N/A
CD-RISC[g], mean (SD)	23.27 (5.76)	24.91 (4.96)	0.31 (small)	0.96	problems in coping with stress or bouncing back from adversity
MISS[h], mean (SD)	4.82 (2.3)	4.6 (2.51)	–0.09 (n)	0.7	no sleep problems

[a]N/A: not applicable.

[b]PSS-10: 10-item Perceived Stress Score.

[c]n: negligible or no effect.

[d]GAD-7: 7-item Generalized Anxiety Disorder scale.

[e]WHO-5: World Health Organization-5 Well-Being Index.

[f]ONS-4: Office of National Statistics Four Subjective Well-being Questions.

[g]CD-RISC: Connor-Davidson Resilience Scale.

[h]MISS: Minimal Insomnia Scale.

Table 4. Within-group analyses showing the change from baseline on the outcome measures at each time point (week 2 and week 4).

Outcome measure	Intervention group				Control group			
	Mean (SD)	*t* (*df*=61)	P value[a]	Effect size (*d*)	Mean (SD)	*t* (*df*=73)	P value[a]	Effect size (*d*)
GAD-7[b]								
Baseline	8.06 (3.83)	__[c]	—	—	6.92 (3.28)	—	—	—
Week 2	6.71 (4.79)	–2.41	<.001	–0.31	6.69 (3.92)	–0.61	.27	–0.06
Week 4	6.02 (4.29)	–3.69	<.001	–0.50	6.14 (3.88)	–1.92	.03	–0.22
MISS[d]								
Baseline	4.82 (2.30)	—	—	—	4.61 (2.51)	—	—	—
Week 2	3.66 (2.28)	–3.43	<.001	–0.51	4.35 (2.61)	–0.96	.17	–0.10
Week 4	3.15 (1.90)	–5.31	<.001	–0.79	4.05 (1.72)	–2.03	.02	–0.21
CD-RISC-10[e]								
Baseline	23.27 (5.76)	—	—	—	24.91 (4.97)	—	—	—
Week 2	25.06 (5.47)	3.45	<.001	0.31	24.59 (5.02)	–0.84	.80	–0.06
Week 4	25.66 (5.48)	4.07	<.001	0.42	25.19 (5.78)	0.58	.28	0.05
WHO-5[f]								
Baseline	11.35 (4.29)	—	—	—	12.00 (4.23)	—	—	—
Week 2	12.89 (5.10)	2.28	.01	0.32	11.77 (4.38)	–0.47	.68	–0.05
Week 4	13.95 (4.45)	–3.85	<.001	0.59	12.12 (4.64)	0.25	.40	0.03
ONS-4[g] Life satisfaction								
Baseline	5.69 (1.53)	—	—	—	6.08 (1.76)	—	—	—
Week 2	6.66 (1.64)	3.79	<.001	0.61	6.69 (3.92)	1.33	.09	0.15
Week 4	6.95 (1.32)	5.41	<.001	0.88	6.61 (1.57)	2.60	.006	0.32
ONS-4 Worth								
Baseline	6.34 (1.94)	—	—	—	6.30 (1.70)	—	—	—
Week 2	6.82 (1.89)	1.69	.05	0.25	6.42 (1.78)	0.64	.26	0.07
Week 4	7.06 (1.56)	2.64	.005	0.41	6.59 (1.70)	1.48	.07	0.17
ONS-4 Happiness								
Baseline	5.95 (1.51)	—	—	—	6.16 (1.72)	—	—	—
Week 2	6.55 (1.87)	2.32	.01	0.35	6.35 (1.53)	0.99	.16	0.12
Week 4	7.08 (1.55)	4.99	<.001	0.74	6.47 (1.78)	1.39	.08	0.18
ONS-4 Anxiety								
Baseline	4.65 (1.66	—	—	—	5.09 (2.00)	—	—	—
Week 2	5.13 (2.50)	1.64	.05	0.22	5.42 (2.25)	1.38	.09	0.15
Week 4	5.08 (2.77)	1.24	.11	0.19	5.12 (2.42)	0.09	.46	0.01
PSS-10[h]								
Baseline	19.19 (4.12)	—	—	—	19.05 (3.58)	—	—	—
Week 1	18.32 (5.27)	–1.35	.09	–0.18	17.59 (5.03)	–2.95	.002	–0.32
Week 2	16.26 (6.20)	–3.38	<.001	–0.55	17.00 (5.50)	–3.31	<.001	–0.43
Week 3	15.65 (5.55)	–4.79	<.001	–0.72	17.01 (5.93	–3.04	.002	–0.40
Week 4	15.53 (5.82)	–4.86	<.001	–0.72	16.38 (5.72)	–3.99	<.001	–0.55

[a]Comparisons were made for each time point relative to the baseline level; significance thresholds have been adjusted to account for multiple comparisons (Bonferroni): P<.025 for all outcomes except for PSS-10, which was set to P<.012.

[b]GAD-7: 7-item Generalized Anxiety Disorder scale.
[c]not applicable.
[d]MISS: Minimal Insomnia Scale.
[e]CD-RISC-10: Connor-Davidson Resilience Scale.
[f]WHO-5: World Health Organization-5 Well-Being Index.
[g]ONS-4: Office of National Statistics Four Subjective Well-being Questions.
[h]PSS-10: 10-item Perceived Stress Score.

Table 5. Between-group analyses showing the change of outcome measures (Δ) from baseline at each time point (week 2 and week 4).

Outcome measure	Intervention Δ, mean (SD)	Control Δ, mean (SD)	Mean difference	*t* (*df*=134)	P value	Effect size (*d*)
GAD-7[a]						
Week 2	–1.35 (4.43)	–0.23 (3.24)	1.13	1.71	.04	0.29
Week 4	–2.05 (4.37)	–0.78 (3.52)	1.26	1.87	.03	0.32
MISS[b]						
Week 2	–1.16 (2.67)	–0.26 (2.29)	0.90	2.13	.02	0.37
Week 4	–1.68 (2.49)	–0.55 (2.35)	1.12	2.70	.004	0.47
CD-RISC[c]						
Week 2	1.79 (4.08)	–0.31 (3.16)	–2.10	–3.37	<.001	–0.58
Week 4	2.39 (4.62)	0.28 (4.24)	–2.10	–2.78	.003	–0.48
WHO-5[d]						
Week 2	1.53 (5.30)	–0.23 (4.20)	–1.76	–2.16	.02	–0.37
Week 4	2.59 (5.32)	0.12 (4.20)	–2.46	–3.05	.001	–0.52
ONS-4[e] life satisfaction						
Week 2	0.97 (2.01)	0.24 (1.58)	–0.72	–2.36	.01	–0.41
Week 4	1.26 (1.53)	0.53 (1.75)	–0.73	–2.38	.009	–0.41
ONS-4 worth						
Week 2	0.48 (2.25)	0.12 (1.63)	–0.36	–1.09	.14	–0.19
Week 4	0.73 (2.17)	0.297 (1.76)	–0.43	–1.28	.10	–0.22
ONS-4 happiness						
Week 2	0.596 (2.02)	0.19 (1.64)	–0.41	–1.29	.10	–0.22
Week 4	1.13 (1.78)	0.31 (1.92)	–0.82	–2.56	.006	–0.44
ONS-4 anxiety						
Week 2	0.48 (2.32)	0.32 (2.02)	–0.16	–0.43	.33	–0.07
Week 4	0.44 (2.77)	0.03 (2.48)	–0.41	–0.91	.18	–0.16
PSS-10[f]						
Week 1	–0.87 (5.09)	–1.45 (4.25)	–0.58	–0.73	.77	–0.13
Week 2	–2.94 (6.84)	–2.05 (5.34)	0.88	0.84	.20	0.15
Week 3	–3.55 (5.84)	–2.04 (5.77)	1.51	1.51	.07	0.26
Week 4	–3.66 (5.93)	–2.68 (5.77)	0.99	0.98	.16	0.17

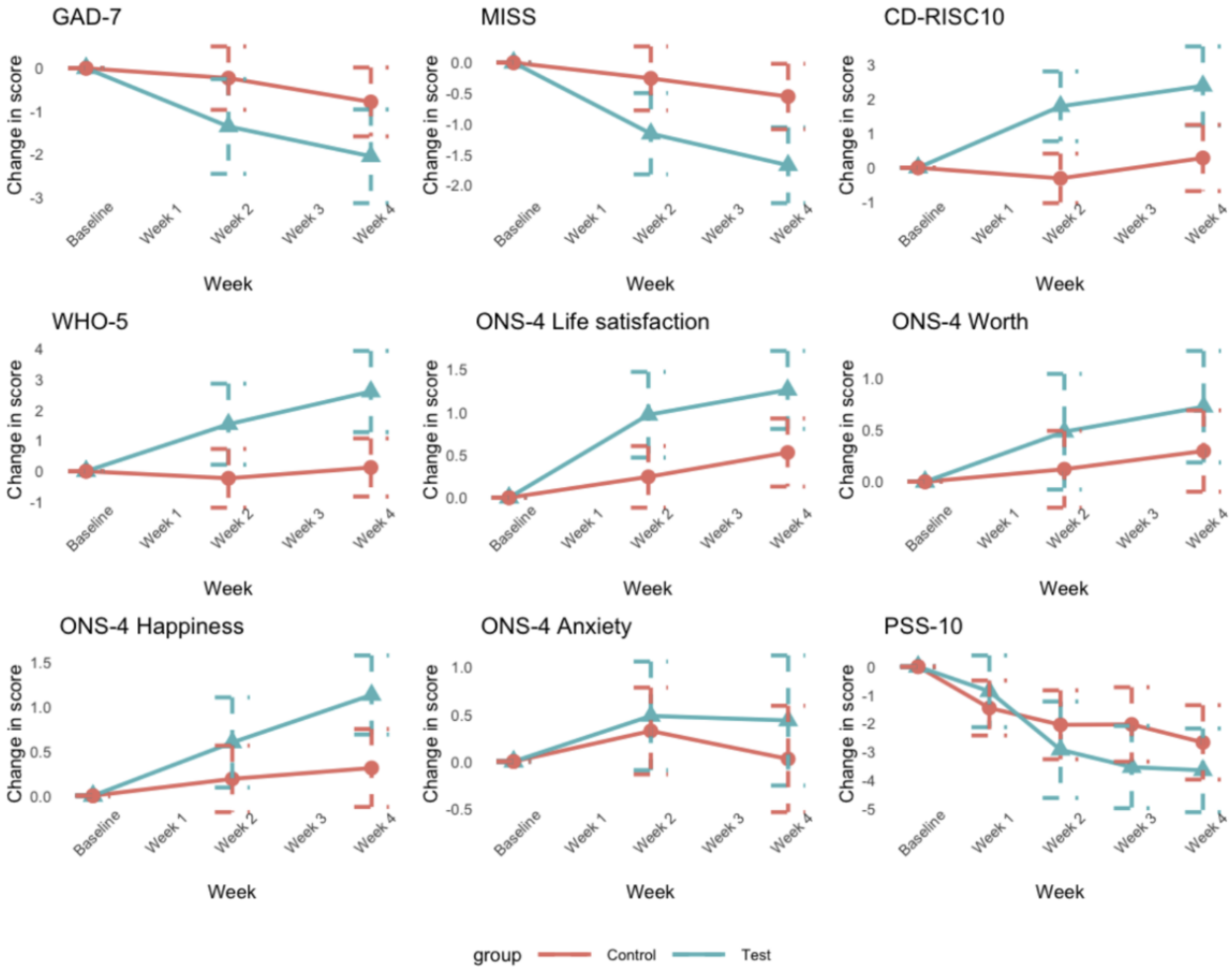
[a]GAD-7: 7-item Generalized Anxiety Disorder scale.
[b]MISS: Minimal Insomnia Scale.
[c]CD-RISC: Connor-Davidson Resilience Scale.
[d]WHO-5: World Health Organization-5 Well-Being Index
[e]ONS-4: Office of National Statistics Four Subjective Well-being Questions.
[f]PSS-10: 10-item Perceived Stress Score.

Anxiety (GAD-7)

Within-group *t* tests for those in the intervention arm (Foundations) showed that GAD-7 scores reduced significantly compared with those at baseline at both week 2 and week 4. In contrast, there was no change in GAD-7 scores in the control group when comparing baseline to week 2 (Table 4).

For the between-group comparisons, LMM analysis of the change from baseline (Δ) showed a significant global effect of the intervention (P=.03), where the main impact seems to be at week 4 with P=.05 for the interaction term (see Tables S1 and S2 in Multimedia Appendix 1). Posthoc comparisons confirmed significant differences at weeks 2 and 4 (Figure 4, Table 5), such that the intervention group had lower GAD-7 scores than the control group. These data suggest that Foundations reduced anxiety within 4 weeks of use.

Figure 4. Between-group analyses showing the change from baseline on outcome measures at each time point. Data points represent the mean and bars are standard errors. PSS-10: 10-item Perceived Stress Score; ONS-4: Office of National Statistics Four Subjective Well-being Questions; GAD-7: 7-item Generalized Anxiety Disorder scale; WHO-5: World Health Organization-5 Well-Being Index; CD-RISC: Connor-Davidson Resilience Scale; MISS: Minimal Insomnia Scale.



Sleep (MISS)

Within-group analysis of the MISS score showed that the score in the intervention group significantly improved both at week 2 and week 4. In comparison, within-group analysis of the control group showed no significant change in the MISS score from baseline to week 2 or from baseline to week 4 (see Table 4).

Between-group LMM analyses of the change from baseline revealed a significantly greater improvement in MISS in the intervention group compared to the control group (P=.01). Even though there was a global decrease of MISS at week 4, interaction terms showed significant results within weeks 2 and 4 (see Tables S3 and S4 in Multimedia Appendix 1). Posthoc comparisons confirmed significant differences at weeks 2 and 4 (Figure 4, Table 5), such that the intervention group had lower MISS scores (higher sleep quality) than the control group (also see Tables S3 and S4 in Multimedia Appendix 1). These data suggest that Foundations improves sleep within 2 weeks of use.

Resilience (CD-RISC-10)

Within-group analyses of the two groups showed significant improvement in resilience scores for the intervention group at both weeks 2 and 4, but no effect in the control group (Table 4).

There were significant between-group effects in the analysis of the change from baseline, such that the intervention group showed a greater change in score than the control group

(P<.001). Moreover, we found a significant positive effect (interaction terms) within the intervention group at both weeks 2 and 4 (see Tables S5 and S6 in Multimedia Appendix 1). Posthoc comparisons confirmed significant differences at weeks 2 and 4 (Figure 4, Table 5), such that the intervention group had higher CD-RISC-10 scores (higher resiliency) compared with those of the control group. These data suggest that Foundations enhances resilience within 2 weeks of use.

Mental Well-being

WHO-5 Scores

Within-group analyses of each group by time showed that the WHO-5 score significantly increased in the intervention group at week 4 but not at week 2. Among controls, there was no significant change at either time point (Table 4).

Group differences were revealed in the LMM between-group analysis (see Tables S7 and S8 of Multimedia Appendix 1) of the change from baseline: the intervention group showed a greater change from baseline than the control group (P=.03) in both weeks 2 and 4. Posthoc comparisons confirmed these significant differences at weeks 2 and 4 (Figure 4, Table 5), such that the intervention group had higher WHO-5 scores (higher quality of life) compared with those of the control group. These data suggest that Foundations can enhance well-being within 2 weeks of use, as measured by the WHO-5.

ONS-4 Questionnaire

The ONS-4 questionnaire comprises four independent questions on life satisfaction, worth, happiness, and anxiety (the scale for the latter has been reversed in the analysis, so that a larger score means a greater positive impact on the participants). The within-group analyses of the intervention group showed improvements in life satisfaction at week 2 and week 4, improvements in worth at week 4, and improvements in happiness at week 2 and week 4. No improvement in anxiety was found at either time point. In contrast, in the control group, only the measure of life satisfaction at week 4 showed a significant improvement compared to baseline (Table 4).

All LMM analyses revealed only overall significant effects on ONS satisfaction and happiness (P=.002 and P=.008, respectively; see Tables S9 and S13 in Multimedia Appendix 1), with inconclusive effects on ONS on worth and anxiety (see Tables S11 and S15 in Multimedia Appendix 1). On ONS satisfaction, we found a significant effect at both 2 and 4 weeks, whereas on ONS happiness, we only found significant effects at week 4 (see interaction terms in Tables S9 and S13 in Multimedia Appendix 1). These results were confirmed by a posthoc *t* test, which showed higher scores in the intervention group than in the control group on the measures of life satisfaction and happiness (Figure 4, Table 5). These data suggest that the intervention group experienced improvement in the sense of life satisfaction and happiness compared to the control group.

Perceived Stress (PSS-10)

In contrast to the other measures, perceived stress was measured weekly. Within-group analyses showed that the PSS-10 score was significantly lower at 2, 3, and 4 weeks, but not at week 1, compared to that measured at baseline for the intervention group (Table 4). However, the control group also showed significantly lower perceived stress levels at each of the four time points (Table 4).

In the same way, between-group analyses of the change from baseline showed no statistically significant result. The likelihood ratio test on LMM showed a P value of .19 (see Tables S17 and S18 in Multimedia Appendix 1), and no significant group differences were observed using *t* tests (Table 5). These data suggest that perceived stress was reduced across the course of the study, but this reduction was no greater in the intervention group than in the control group.

Subclass Analyses

Additional analyses were performed with the factors of gender and age included on each of the outcome measures for both within-group and between-group comparisons. No statistically significant differences were observed, due to the fact that the size of groups was small. See Table S19 in Multimedia Appendix 1 for the age and gender interaction terms for each of the metrics.

Engagement With the App

The participants in the intervention group presented an average usage of the app of 18 days out of the 28 (SD 5.1) total days of the study period (median days active: 17). On average, participants engaged with the app 15.25 minutes per day (median minutes per day 13.93, SD 9.66). During the study period, users tried an average of 29 distinct activities and 3.9 programs. However, no correlation was found between the total amount of engagement with the app and the difference from baseline to 4 weeks in any of the outcome measures (all P>.05).

Research suggests that the level of engagement with a digital intervention impacts the outcomes [26]. However, this was not the case in our study, which may be due to the mechanisms of action taking place outside of the app in which the user is encouraged by the app to complete a healthy behavior. Once a behavior is learned or a skill is developed, the user may not need to engage with the app to experience the benefits (eg, a user can maintain a healthy sleep routine without accessing the app every night).

Discussion

Principal Findings

The aim of the study was to investigate the efficacy of a mobile app, Foundations, in improving mental well-being in adults with moderate to high levels of stress and anxiety. Given the timing of the study (April 2020 to May 2020), our secondary aim was to assess the efficacy of Foundations in mitigating the mental health challenges surrounding the COVID-19 pandemic.

Encouragingly, results of this proof-of-principle study confirmed four out of five of our hypotheses by demonstrating that the use of Foundations can significantly improve measures of (1) anxiety, (2) resilience, (3) well-being, and (4) sleep relative to a control group within 2 weeks of use, with greater effects after 4 weeks. In contrast to our final hypothesis, perceived stress

was reduced within the intervention group, although the results did not reach statistical significance relative to the control group.

Comparison With Prior Work

The results of this study contribute to a limited number of RCTs examining the efficacy of mental health apps that are commercially available. Encouragingly, our findings are consistent with published meta-analyses that have affirmed the efficacy of digital interventions over control conditions in improving mental well-being. For example, a meta-analysis of 18 RCTs reported that smartphone interventions were significantly more efficacious in reducing depressive symptoms in comparison to both waitlist and active control groups [27]. Congruent with results from a recent meta-analysis of 66 RCTs of mobile apps for mental health, we found a greater reduction in anxiety among participants in the intervention versus control group as well as greater improvements in resilience and overall mental well-being [28].

However, our results are at odds with previous RCTs reporting greater reductions in stress in active versus control interventions [29,30]. These results were surprising given the positive results on measures of anxiety, resilience, and well-being. This lack of a statistically significant effect is due to an improvement overall in the control group across the period of the study. There are several potential factors that may have contributed to this finding. First, the study started at the peak of the COVID-19 pandemic in the United Kingdom (April 2020). The reported numbers of positive cases and deaths were reduced by approximately 40% during the 4 weeks of the study [17]. Further, as the population adjusted to lockdown measures, perceived stress levels may have naturally reduced. However, it is not clear why this would impact perceived stress to a greater extent to, for example, anxiety and resilience. Another factor may be the frequency of testing. The PSS-10 was assessed on a weekly basis as opposed to every 2 weeks in the case of all other measures. Foundations' users were also able to take the PSS-10 whenever they liked within the app; these were voluntary assessments, and the data were not evaluated. It is therefore possible that survey fatigue or more frequent insight into stress levels impacted perceived stress.

With regard to intervention features, Foundations differentiates from many of the apps that have published efficacy data in that it offers a breadth of interactive content delivered. For example, Sleepio [31] provides an internet-based CBT course targeted at individuals with insomnia, and Woebot [32] provides CBT-directed mental health via an artificial intelligence–powered chatbot for individuals struggling with poor mental health symptoms. Foundations provides a wider array of support across the mental health spectrum via a number of methodologies (CBT, acceptance and commitment training, positive psychology, and sleep science). Although it is unclear whether single- versus multi-intervention apps differ in efficacy, at the very least, multi-intervention apps such as Foundations offer users flexibility and variety in the content and functionalities they can choose from, which has been shown to increase app engagement and likability [33].

Despite the rapid upscaling of digital mental health interventions during the COVID-19 pandemic, literature on the efficacy of these technologies is rather sparse at present. To the best of our knowledge, this study is one of the few RCTs investigating the efficacy of a mental health app during the COVID-19 pandemic [34,35]. The importance of this research is two-fold. First, the results from this study demonstrate that Foundations could add a scalable efficacious digital intervention to support mental well-being that can be accessed at any time. Second, these findings may help to inform the rationale and design of future studies and digital health technologies. Going forward, it will be critical to maintain this momentum as the mental health consequences of the pandemic could be severe and long-lasting.

Limitations

Several important limitations of this study should be recognized. Perhaps the most important limitation is the notable number of participants excluded from the primary analysis due to a calculation error of the PSS-10 at screening. It remains uncertain whether a larger sample of patients would result in outcomes that differ from those observed; however, the sample size for the final analysis (N=136) is comparable to those reported in the literature on digital interventions for mental well-being [27,28].

Second, it is possible that the daily scheduled broadcast messages received by participants in the Foundations group contributed to a placebo effect or to results that do not reflect real-world usage of the app. Conversely, it is also possible that the frequency of the notifications *hindered* engagement in some cases, as the amount of contact for optimizing user engagement likely varies across participants. However, it is notable that the size of the effect was greater at 4 weeks than at 2 weeks for all measures. Participants received no messages during weeks 3 and 4 other than the links to fill out the questionnaires. It therefore seems unlikely that the daily messaging could account for the efficacy observed in the study. However, the impact of the frequency and content of notifications on well-being outcomes remains to be fully elucidated and further studies may be warranted.

Another limitation of the study is the use of a monetary incentive for participation. Previous studies have shown that monetary incentives can increase engagement with wellness apps yet have no impact on the outcome of the study [12]. However, the true influence of the monetary incentive in this case is unknown.

Finally, it is necessary to acknowledge the limitations of a single-blind design, as used here. Only the data analysts were blind to the group assignment. The control group participants were aware they were taking part in an intervention study, but were only given access to the app after the study (and were not informed they would have access until study completion). The design of the study and the passive nature of the control group do not allow ruling out digital placebo effects in the intervention group derived from their expectations about the interventions [36]. Both groups, however, received the same messages and invitations to fill out the outcome questionnaires. It remains possible that insights into mental well-being through completion of the questionnaires, along with knowledge that they were participating in an intervention trial may have impacted outcome measures. Future studies may explore the use of alternative

designs such as an active control or psychoeducation control group.

Future Research

Before we can optimize the efficacy of digital health technologies such as Foundations, we must build a richer understanding of which interventions are most effective for which individual needs. It was not feasible with the sample size of the current study to evaluate the efficacy of individual components of the app with respect to symptom type and severity, and the content that the participant engaged with during the course of the study. An important question for future research is whether personalization of these interventions makes care more effective.

Similarly, it is unclear whether there was an effect of the level of engagement (ie, a dose effect). Although a satisfactory level of engagement was observed in this study, establishing a dose-response relationship between usage and improvement in mental well-being could prove useful for intervention personalization and bears further investigation in future.

A specific domain that has gained attention in recent years is that of work-related mental health. This is due to an increase of awareness of the magnitude and costs of this issue. Concretely, it has been reported that an outstanding 72% of employees of large organizations in the United Kingdom have disclosed an increase of CMD during 2019 [37]. Ill mental health in the workplace is associated with decreased productivity, early retirement, increased sickness absence, presenteeism (not working at capacity while at work), and staff turnover. All this translates into an estimated cost of over US $45 billion per year for companies, which has increased by 16% in the last few years [37]. Although this study examined the efficacy of Foundations in a working population, a truly rigorous investigation of its effects on workplace mental health would require a future study employing randomized, controlled allocation of participants to an intervention arm (Foundations) or placebo arm, both in the absence and presence of employment. At the very least, these preliminary observations provide evidence that working adults using the Foundations app can experience significant improvement in their mental well-being during a 2-week period.

Future research is also required to evaluate the long-term effects of Foundations on mental well-being both in terms of postintervention follow-up and longer sustained use of the app. Looking forward and if replicated in further studies, these results may have important implications for addressing the treatment gap in mental health care through the use of evidence-based digital interventions.

Overall, results from this study may help to propel the use of mobile apps such as Foundations to assume a more widespread role in both the promotion and maintenance of mental health. This could offer new possibilities to further optimize the efficacy of these technologies while removing obstacles for evidence-based mental health care.

Conclusions

This proof-of-principle study demonstrates that Foundations can improve measures of anxiety, sleep, resilience, and mental well-being within 2 weeks of use, with a greater effect after 4 weeks. Foundations may therefore offer potential as a scalable, cost-effective intervention to enhance mental well-being even during a period of crisis such as the COVID-19 pandemic.

Acknowledgments

The authors are grateful to Oliver Harrison (CEO, Koa Health) and Aleksandar Matic (Research and Development Director, Koa Health) for support with this research.

Authors' Contributions

SD, SC, and CBA conceptualized the study and guided the study design. ES, BS, and IE provided input on the study design. Participant recruitment was supervised by SS. ARV, IE, and BS developed the plan for the statistical analyses. ARV analyzed the data. ARV, BS, IE, and SD contributed to the interpretation of the data. SC, LS, NRH, and SD drafted the manuscript. All authors made revisions to the manuscript. LS had final responsibility for submission of the manuscript.

Conflicts of Interest

All authors, excluding NRH and LS, were employees of Telefónica Innovación Alpha at the time of the study. The app and the authors have now transferred from Telefonica Innovation Alpha to Koa Health.

Editorial Notice

This randomized study was only retrospectively registered, as due to unforeseen logistical reasons, authors were unable to preregister the study. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials, because the risk of bias appears low. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
Supplementary Tables S1-S19.
[DOCX File , 31 KB - mhealth_v10i7e30976_app1.docx]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.2).
[PDF File (Adobe PDF File), 114 KB - mhealth_v10i7e30976_app2.pdf]

References

1. Dattani S, Ritchie H, Roser M. Mental health. Our World in Data. 2021. URL: https://ourworldindata.org/mental-health [accessed 2021-05-13]
2. Depression and other common mental disorders: global health estimates. World Health Organization. 2017. URL: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf [accessed 2021-02-05]
3. Wittchen H, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011 Sep;21(9):655-679. [doi: 10.1016/j.euroneuro.2011.07.018] [Medline: 21896369]
4. Gustavsson A, Svensson M, Jacobi F, Allgulander C, Alonso J, Beghi E, CDBE 2010 Study Group. Cost of disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011 Oct;21(10):718-779. [doi: 10.1016/j.euroneuro.2011.08.008] [Medline: 21924589]
5. Czeisler M, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic - United States, June 24-30, 2020. MMWR Morb Mortal Wkly Rep 2020 Aug 14;69(32):1049-1057. [doi: 10.15585/mmwr.mm6932a1] [Medline: 32790653]
6. Vindegaard N, Benros ME. COVID-19 pandemic and mental health consequences: systematic review of the current evidence. Brain Behav Immun 2020 Oct 30;89:531-542 [FREE Full text] [doi: 10.1016/j.bbi.2020.05.048] [Medline: 32485289]
7. Pierce M, Hope H, Ford T, Hatch S, Hotopf M, John A, et al. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. Lancet Psychiatry 2020 Oct;7(10):883-892 [FREE Full text] [doi: 10.1016/S2215-0366(20)30308-4] [Medline: 32707037]
8. Ettman CK, Abdalla SM, Cohen GH, Sampson L, Vivier PM, Galea S. Prevalence of depression symptoms in US adults before and during the COVID-19 pandemic. JAMA Netw Open 2020 Sep 01;3(9):e2019686 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.19686] [Medline: 32876685]
9. Wu T, Jia X, Shi H, Niu J, Yin X, Xie J, et al. Prevalence of mental health problems during the COVID-19 pandemic: a systematic review and meta-analysis. J Affect Disord 2021 Feb 15;281:91-98 [FREE Full text] [doi: 10.1016/j.jad.2020.11.117] [Medline: 33310451]
10. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013 Nov 15;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
11. Naslund JA, Aschbrenner KA, Araya R, Marsch LA, Unützer J, Patel V, et al. Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. Lancet Psychiatry 2017 Jun;4(6):486-500 [FREE Full text] [doi: 10.1016/S2215-0366(17)30096-2] [Medline: 28433615]
12. Kawadler JM, Hemmings NR, Ponzo S, Morelli D, Bird G, Plans D. Effectiveness of a smartphone app (BioBase) for reducing anxiety and increasing mental well-being: pilot feasibility and acceptability study. JMIR Form Res 2020 Nov 10;4(11):e18067 [FREE Full text] [doi: 10.2196/18067] [Medline: 32969341]
13. Ponzo S, Morelli D, Kawadler JM, Hemmings NR, Bird G, Plans D. Efficacy of the digital therapeutic mobile app BioBase to reduce stress and improve mental well-being among university students: randomized controlled trial. JMIR Mhealth Uhealth 2020 Apr 06;8(4):e17767 [FREE Full text] [doi: 10.2196/17767] [Medline: 31926063]
14. Larsen ME, Huckvale K, Nicholas J, Torous J, Birrell L, Li E, et al. Using science to sell apps: Evaluation of mental health app store quality claims. NPJ Digit Med 2019 Mar 22;2(18):18-1-. [doi: 10.1038/s41746-019-0093-1] [Medline: 31304366]
15. Ethical principles. Koa Health. URL: https://koahealth.com/legal/ethical-principles/ [accessed 2013-05-21]
16. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
17. United Kingdom Coronavirus. Worldometer. 2020. URL: https://www.worldometers.info/coronavirus/country/uk/ [accessed 2021-05-13]
18. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
19. Lee E. Review of the psychometric evidence of the perceived stress scale. Asian Nurs Res (Korean Soc Nurs Sci) 2012 Dec;6(4):121-127 [FREE Full text] [doi: 10.1016/j.anr.2012.08.004] [Medline: 25031113]
20. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
21. Broman J, Smedje H, Mallon L, Hetta J. The Minimal Insomnia Symptom Scale (MISS): a brief measure of sleeping difficulties. Ups J Med Sci 2008;113(2):131-142. [doi: 10.3109/2000-1967-221] [Medline: 18509808]
22. Campbell-Sills L, Stein MB. Psychometric analysis and refinement of the Connor-davidson Resilience Scale (CD-RISC): validation of a 10-item measure of resilience. J Trauma Stress 2007 Dec;20(6):1019-1028. [doi: 10.1002/jts.20271] [Medline: 18157881]

23. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015 Mar 28;84(3):167-176 [FREE Full text] [doi: 10.1159/000376585] [Medline: 25831962]
24. Dolan P, Metcalfe R. Measuring subjective wellbeing: recommendations on measures for use by national governments. J Soc Pol 2012 Jan 12;41(2):409-427. [doi: 10.1017/s0047279411000833]
25. Baik SH, Fox RS, Mills SD, Roesch SC, Sadler GR, Klonoff EA, et al. Reliability and validity of the Perceived Stress Scale-10 in Hispanic Americans with English or Spanish language preference. J Health Psychol 2019 Apr;24(5):628-639 [FREE Full text] [doi: 10.1177/1359105316684938] [Medline: 28810432]
26. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
27. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298. [doi: 10.1002/wps.20472] [Medline: 28941113]
28. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct;18(3):325-336. [doi: 10.1002/wps.20673] [Medline: 31496095]
29. Coelhoso CC, Tobo PR, Lacerda SS, Lima AH, Barrichello CRC, Amaro E, et al. A new mental health mobile app for well-being and stress reduction in working women: randomized controlled trial. J Med Internet Res 2019 Nov 07;21(11):e14269 [FREE Full text] [doi: 10.2196/14269] [Medline: 31697244]
30. Fuller-Tyszkiewicz M, Richardson B, Little K, Teague S, Hartley-Clark L, Capic T, et al. Efficacy of a smartphone app intervention for reducing caregiver stress: randomized controlled trial. JMIR Ment Health 2020 Jul 24;7(7):e17541 [FREE Full text] [doi: 10.2196/17541] [Medline: 32706716]
31. Elison S, Ward J, Williams C, Espie C, Davies G, Dugdale S, et al. Feasibility of a UK community-based, eTherapy mental health service in Greater Manchester: repeated-measures and between-groups study of 'Living Life to the Full Interactive', 'Sleepio' and 'Breaking Free Online' at 'Self Help Services'. BMJ Open 2017 Jul 20;7(7):e016392 [FREE Full text] [doi: 10.1136/bmjopen-2017-016392] [Medline: 28729322]
32. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun 06;4(2):e19 [FREE Full text] [doi: 10.2196/mental.7785] [Medline: 28588005]
33. Alqahtani F, Orji R. Insights from user reviews to improve mental health apps. Health Informatics J 2020 Sep;26(3):2042-2066 [FREE Full text] [doi: 10.1177/1460458219896492] [Medline: 31920160]
34. Boucher EM, McNaughton EC, Harake N, Stafford JL, Parks AC. The impact of a digital intervention (Happify) on loneliness during COVID-19: qualitative focus group. JMIR Ment Health 2021 Feb 08;8(2):e26617 [FREE Full text] [doi: 10.2196/26617] [Medline: 33498011]
35. Fiol-DeRoque MA, Serrano-Ripoll MJ, Jiménez R, Zamanillo-Campos R, Yáñez-Juan AM, Bennasar-Veny M, et al. A mobile phone-based intervention to reduce mental health problems in health care workers during the COVID-19 pandemic (PsyCovidApp): randomized controlled trial. JMIR Mhealth Uhealth 2021 May 18;9(5):e27039 [FREE Full text] [doi: 10.2196/27039] [Medline: 33909587]
36. Torous J, Firth J. The digital placebo effect: mobile mental health meets clinical psychiatry. Lancet Psychiatry 2016 Feb;3(2):100-102. [doi: 10.1016/S2215-0366(15)00565-9] [Medline: 26851322]
37. Hampson E, Jacob A. Mental health and employers: refreshing the case for investment. Deloitte. 2020. URL: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/consultancy/deloitte-uk-mental-health-and-employers.pdf [accessed 2021-05-13]

Abbreviations

CBT: cognitive behavioral therapy
CD-RISC: Connor-Davidson Resilience Scale
CMD: common mental disorder
GAD-7: 7-item Generalized Anxiety Disorder scale
LMM: linear mixed model
MISS: Minimal Insomnia Scale
ONS-4: Office of National Statistics Four Subjective Well-being Questions
PSS-10: 10-item Perceived Stress Score
RCT: randomized controlled trial
WHO-5: World Health Organization-5 Well-Being Index

Edited by L Buis; submitted 04.06.21; peer-reviewed by A Teles, D Pan; comments to author 29.09.21; revised version received 19.10.21; accepted 06.12.21; published 01.07.22.

Please cite as:
Catuara-Solarz S, Skorulski B, Estella-Aguerri I, Avella-Garcia CB, Shepherd S, Stott E, Hemmings NR, Ruiz de Villa A, Schulze L, Dix S
The Efficacy of "Foundations," a Digital Mental Health App to Improve Mental Well-being During COVID-19: Proof-of-Principle Randomized Controlled Trial
JMIR Mhealth Uhealth 2022;10(7):e30976
URL: https://mhealth.jmir.org/2022/7/e30976
doi:10.2196/30976
PMID:34978535

Original Paper

Awareness, Acceptability, and Perceived Effectiveness of Text-Based Therapy Among Graduate Students: Cross-sectional Study

Samari A Blair[1,2], MPH; Andrea N Brockmann[3], MA; Kelsey M Arroyo[3], MS; Chelsea A Carpenter[3], MPH, PhD; Kathryn M Ross[2,3], MPH, PhD

[1]College of Medicine, University of Florida, Gainesville, FL, United States

[2]Social and Behavioral Sciences Program, College of Public Health and Health Professions, University of Florida, Gainesville, FL, United States

[3]Department of Clinical and Health Psychology, College of Public Health and Health Professions, University of Florida, Gainesville, FL, United States

Corresponding Author:
Kathryn M Ross, MPH, PhD
Department of Clinical and Health Psychology
College of Public Health and Health Professions
University of Florida
PO Box 100165
Gainesville, FL, 32610
United States
Phone: 1 352 294 8433
Email: kmross@phhp.ufl.edu

Abstract

Background: Research has suggested that there is a mental health crisis occurring among graduate students in the United States. Moreover, many students go without effective treatment owing to the limited availability of mental and behavioral health resources on college campuses. Text-based therapy may represent a viable method for increasing access to mental health support for graduate students, but little is known regarding its acceptability in this population.

Objective: The purpose of this study was to assess how graduate students perceive text-based therapy and their likelihood of seeking out this form of therapy.

Methods: In total, 265 graduate students completed a cross-sectional web-based survey that included multiple-choice and open-ended questions assessing their perceptions of text-based therapy and the likelihood of seeking out this form of therapy. Chi-square tests, ANOVAs, and nonparametric Wilcoxon signed-rank tests were used to examine differences in multiple-choice questions. The constant comparative method was used for qualitative analyses of the open-ended question responses.

Results: Participants (n=265) were predominately non-Hispanic White (166/265, 62.6%) and female (167/265, 63%) with a mean age of 28.3 (SD 5.1) years. Over half of the participants (139/265, 52.5%) were not aware that text-based therapy existed; however, 65.3% (173/265) reported that they would consider using text-based services, if available. In comparison to face-to-face therapy, participants reported being less likely to seek out text-based therapy and perceived it as less effective (P<.001). Qualitative results indicated that participants were concerned about the ability to effectively communicate and build rapport through text-based therapy and thought that this modality may be more effective for some mental and behavioral health concerns than others. Moreover, participants noted that text-based therapy would be best implemented as a way to supplement, rather than replace, face-to-face services.

Conclusions: Altogether, the results of this study suggest that text-based therapy holds the potential to increase access to and use of mental and behavioral health services; however, graduate students remain concerned about its effectiveness and the optimal methods of implementation. Future research should investigate how therapeutic processes (eg, effective communication and rapport-building) can be facilitated in digital environments and how text-based therapy could be best implemented to supplement and extend, rather than replace, face-to-face services.

(JMIR Mhealth Uhealth 2022;10(7):e34102) doi:10.2196/34102

KEYWORDS

mental health; text-based therapy; graduate students

Introduction

Approximately 18.9% of all adults in the United States live with a mental illness, with the highest prevalence rates observed among young adults aged 18 to 25 years [1,2]. Mental health disorders account for approximately one half of the disease burden among young adults [3] and, when compared with neurological and substance use disorders, constitute the highest proportion of disability-adjusted life years [4]. Graduate students represent a particularly vulnerable population, with data indicating greater risk of depression and anxiety compared with the general population [5,6]. Therefore, many researchers have argued that there is a mental health crisis occurring in this population [5].

In-person therapy remains the most widely used method to treat individuals with mental and behavioral health conditions [7]. Across several one-on-one or group sessions, trained therapists aim to help patients understand their thoughts, feelings, and patterns of behavior while identifying ways to positively cope with psychosocial stressors [8,9]. Therapy has been proven to be effective in treating common mental and behavioral health disorders, with substantial evidence supporting its effectiveness for decreasing symptoms of major depressive disorder, generalized anxiety disorder, and panic disorder in adult populations [7,10].

There are several barriers experienced by undergraduate and graduate students related to the use of mental and behavioral health services [11,12]. Individual-level barriers include scheduling conflicts, lack of time, negative self-stigma associated with seeking therapy, financial constraints, lack of perceived need, and perceived ineffectiveness of treatment [13]. Organizational-level barriers include a lack of insurance coverage for mental and behavioral health services and extended waitlists for appointments on campus [11]. Innovative delivery formats, such as delivery of treatment through the internet or mobile devices, have the potential to address several of these barriers and thus increase graduate students' use of mental health services. These approaches may be particularly well suited for this population as technology has a strong presence among young adults and students on university campuses [11,14].

The delivery of therapy via SMS text message, in particular, has been gaining momentum [15]. Text-based therapy relies on asynchronous communication and allows individuals to send written messages to a certified counselor or a therapist using a mobile device. The therapist then responds (via the same platform) at their earliest convenience [15]. Despite its increasing popularity [16], research on the efficacy and acceptability of text-based therapy has been limited. One study demonstrated that adding SMS text messages to ongoing in-person therapy did not improve clinical outcomes in adult populations [17]; however, another found that the use of a text-based therapy service for 15 weeks produced significant reductions in symptoms and increased work productivity in adults with depression and anxiety [18]. In terms of acceptability, 1 study found that asynchronous text therapy with a licensed therapist was viewed as acceptable and clinically beneficial for adults with various diagnoses and histories of psychological distress [19]. To date, no studies have investigated the acceptability of text-based therapy in graduate students.

To address this gap, this study aimed to survey current graduate students at a large southeastern university in the United States regarding their awareness, acceptability, and perceived effectiveness of text-based therapy for treating mental and behavioral health challenges. Moreover, this survey aimed to identify which factors would promote or prevent the use of text-based therapy within this specific population. Finally, this study examined whether there were differences in the likelihood of seeking out face-to face and text-based therapy by age, race, gender, and sexual orientation. On the basis of the results from a previous study, which found that White individuals were more likely than Hispanic, Black, or African American individuals to receive mental health treatment [20], it was hypothesized that individuals who were younger, female, heterosexual, and non-Hispanic White would be more likely to seek out therapy than their respective counterparts.

Methods

Recruitment

This study aimed to recruit current graduate students at both the master and doctoral level who were employed by their departments at the University of Florida in exchange for a stipend (considered *graduate assistants* at the University of Florida). Potential participants were eligible to complete the survey if they were a current graduate assistant, were aged ≥18 years, and were enrolled in GatorGradCare, an employer-sponsored health plan specific to graduate assistants at the University of Florida. At the time that the current survey was conducted (spring 2019), GatorGradCare had started offering free asynchronous text-based therapy (via TalkSpace) as part of their health insurance plan benefits.

To recruit participants for this study, the first author (SAB, who at the time was completing a public health internship at GatorCare) sent an email on January 21, 2019, to all current University of Florida graduate assistants through the GatorGradCare listserv. This email introduced the purpose of the study and provided participants with a link to complete the study survey through Qualtrics, a web-based survey platform. Participants were asked to read a statement that outlined the purpose of the project; instructions describing how to complete the survey; and information about potential risks and benefits, confidentiality, voluntary participation, and participant rights (eg, the right to refuse to answer individual questions and to withdraw participation at any time). Individuals who provided web-based consent to participate were prompted to complete the remainder of the survey questions. On February 1, 2019, 2 weeks after the initial email was sent, the first author sent a second follow-up email to the GatorGradCare listserv, to serve as a reminder for those who may have still wanted to participate.

The survey closed 1 month following initial contact (February 15, 2019). No compensation was provided for completing the survey.

Measures

The survey used in this project consisted of 19 self-report items (with a mix of multiple-choice and open-ended questions) developed by the first author (SAB). Questions are described by content area in subsequent sections.

Demographics

Participants were asked to report their age, race, gender, and sexual orientation. No personal identifiers (eg, name, address, or email address) were collected from the survey.

Experiences With Traditional Face-to-Face Therapy

Participants were asked if they had previously used mental and behavioral health services (defined as "therapy or counseling") and, if so, for what concerns. Participants could check all that applied from the following options: substance abuse, alcohol abuse, anxiety, depression, stress management, attention-deficit/hyperactivity disorder, gender dysphoria, relationships, eating disorders, sexual assault, smoking cessation, autism spectrum disorder, or bereavement (these terms were used as they were listed as example conditions for treatment by the TalkSpace text-based therapy service [21]). There were also options of "prefer not to answer" and "other," which included a text box that enabled participants to elaborate further. Participants were also asked about barriers that prevented them from using therapy or counseling services in the past, perceived effectiveness of face-to-face therapy, and their likelihood of seeking out face-to-face therapy. Perceived effectiveness of face-to-face therapy was assessed using a 5-point Likert scale ranging from 1=*extremely effective* to 5=*not effective at all*. The likelihood of seeking out face-to-face therapy was also assessed on a 5-point Likert scale ranging from 1=*very likely* to 5=*not likely at all*.

Experiences With Text-Based Therapy

Participants were asked if they were aware of text-based therapy before the onset of the study, if they would consider using text-based therapy if available, the factors that would promote them to use text-based therapy, the factors that would promote them to not use text-based therapy, their likelihood of seeking out text-based therapy, and perceived effectiveness of text-based therapy. Perceived effectiveness of text-based therapy was assessed using a 5-point Likert scale ranging from 1=*extremely effective* to 5=*not effective at all*. The likelihood of seeking out text-based therapy for behavioral health issues was also assessed on a 5-point Likert scale ranging from 1=*very likely* to 5=*not likely at all*.

Open-Ended Survey Questions

Participants were asked to provide free-text responses to three open-ended questions: (1) "Do you think text-based counseling can be effective? Why or why not? What factors do you think limit its effectiveness?" (2) "Do you see any advantages to using text-based counseling over face-to-face counseling? Disadvantages?" (3) "Please provide any additional comments regarding your perception of text-based therapy."

Data Analysis

All quantitative analyses were conducted using SPSS Statistics (version 25; IBM Corp). Descriptive statistics were used to summarize the demographic and quantitative data. Chi-square tests and ANOVAs were used to examine whether there was an association between past use of therapy or counseling and willingness to use text-based therapy in the future, and if this association varied by age, race, gender or sex, or sexual orientation. Nonparametric Wilcoxon signed-rank tests were conducted to assess whether there were significant differences between the perceived effectiveness and likelihood of seeking therapy between text-based and face-to-face therapy services.

The constant comparative method, which is used by researchers to uncover similarities, differences, and patterns in data [22], was used for qualitative analyses of the responses from the 3 open-ended questions. Open coding methods [22] were used to examine, compare and contrast, and categorize the raw data. Data were then compiled into subthemes, and these subthemes were then compiled into larger, overarching themes. Coding was done individually by two authors (SAB and ANB) to limit subjectivity and bias. SAB and ANB then compared results of this coding process, with disagreements resolved through a consensus discussion and review by KMR.

Ethics Approval

This project was approved by the University of Florida Institutional Review Board (IRB#201802501).

Results

Participants

Of the 3609 graduate assistants who were contacted via the GatorGradCare listserv, 297 provided consent to participate in this study (reflecting a response rate of 8.2%). Responses from 32 participants were excluded owing to extensive missing data (ie, completing less than half of the survey questions), resulting in 265 participants being included in the current analyses. Demographic characteristics of survey participants are provided in Table 1. On average, participants were aged 28.3 (SD 5.1) years, and the majority identified as non-Hispanic White (166/265, 62.6%), female (167/265, 63%), and heterosexual or straight (215/265, 81.1%).

Table 1. Participant demographics (N=265).

Characteristic	Value
Age (years; n=245), mean (SD)	28.3 (5.1)
Race and ethnicity, n (%)	
Asian, non-Hispanic	32 (12.1)
Black or African American, non-Hispanic	12 (4.5)
Hispanic or Latinx	28 (10.6)
White, non-Hispanic	166 (62.6)
Another race or multiple races selected	22 (8.3)
Prefer not to answer or missing	5 (1.9)
Gender, n (%)	
Female	167 (63)
Male	87 (32.8)
Transgender	3 (1.1)
Other	1 (0.4)
Prefer not to answer or missing	7 (2.6)
Sexual orientation, n (%)	
Heterosexual or straight	215 (81.1)
Bisexual	20 (7.6)
Homosexual	12 (4.5)
Another category	9 (3.4)
Prefer not to answer or missing	9 (3.4)

Past Utilization of Behavioral Health Services and Awareness of Text-Based Therapy

Approximately 60.4% (160/265) of survey participants reported previous use of some form of mental or behavioral health services. The top three concerns individuals sought face-to-face therapy for were anxiety, depression, and stress management (Table 2). Additional concerns that participants mentioned previously seeking face-to-face therapy for included eating disorders, bereavement, parental divorce, posttraumatic stress disorder, bipolar disorder, and suicidal ideation. Individuals who identified with historically marginalized racial or ethnic groups were significantly less likely to report having used mental or behavioral health services compared with non-Hispanic White participants (45/94, 48%, vs 112/166, 67.5%, respectively; n=260, χ^2_1=9.6; *P*=.002). Moreover, female students were more likely to report having used these services in the past than male students (113/167, 67.7%, vs 38/87, 44%, respectively; n=254, χ^2_1=13.7; *P*<.001). There were no significant differences in reports of past use by sexual orientation (n=256, χ^2_2=2.3; *P*=.34) or age ($F_{1,245}$=2.6; R^2=.01; *P*=.10).

Table 2. Top 5 concerns for which individuals sought face-to-face therapy.

Condition	Participants who sought out face-to-face therapy for this condition, n (%)
Anxiety	110 (41.5)
Depression	105 (39.6)
Stress management	70 (26.4)
Relationships	45 (16.9)
Sexual assault	20 (7.5)

Approximately half of participants (139/265, 52.5%) were not aware that text-based counseling existed before participating in this study; however, 65.3% (173/265) of participants reported they would consider using text-based counseling if available. Participants reported they would most likely seek text-based therapy for concerns such as stress management, anxiety, and depression (Table 3).

Table 3. Top 5 concerns for which individuals would most likely seek text-based therapy.

Condition	Participants who reported that they would likely seek text-based therapy for this condition, n (%)
Stress management	136 (51.3)
Anxiety	131 (49.4)
Depression	110 (41.5)
Relationships	81 (30.5)
Sexual assault	23 (8.6)

Barriers to Accessing Face-to-Face and Text-Based Therapy Services

The top barriers that prevented participants from seeking mental and behavioral health services included cost (115/265, 43.4%), scheduling conflicts (107/265, 40.4%), that it was not needed (70/265, 26.4%), and that a therapist could not be found (66/265, 24.9%). In total, 50/265 (18.9%) participants also reported “other” barriers and wrote responses including anxiety about starting therapy, stigma, long wait times owing to an insufficient number of counselors, and insufficient coverage by insurance.

The top factors that would promote participants to seek text-based therapy were convenience (206/265, 77.7%), low cost (141/265, 53.2%), anonymity from other individuals in person (76/265, 28.7%), and a better sense of privacy and security (70/265, 26.4%). In total, 29/265 (10.9%) participants described “other” reasons promoting the use of text-based therapy, including preferences for communicating via text versus in person and anxiety regarding face-to-face communication. Barriers that would prevent participants from seeking text-based therapy were a preference for face-to-face contact (189/265, 71.3%), that they were not interested or did not need therapy (58/265, 21.9%), cost (40/265, 15.1%), and a preference for anonymity related to privacy and security (20/265, 7.5%). In total, 42/265 (15.8%) “other” barriers were reported, including the impersonal nature of text-based communications, perceptions that text-based communication would prevent development of a deep relationship between therapist and client, and a perceived lack of effectiveness.

Perceived Effectiveness and Likelihood of Seeking Face-to-Face Versus Text-Based Psychotherapy Services

Figure 1 provides ratings of perceived effectiveness of therapy by modality. Most participants (169/265, 63.8%) reported beliefs that face-to-face therapy was "very effective" in treating mental and behavioral health concerns. In comparison, only 13/265 (4.9%) participants reported that text-based therapy was “very effective”; the remaining participants reported beliefs that text-based therapy was “somewhat effective” or that they remained “undecided.” Figure 2 provides responses related to the likelihood of seeking therapy by each delivery modality. Although most participants (221/265, 83.4%) reported that they were “extremely likely” or “somewhat likely” to seek face-to-face therapy, only 34.4% (90/262) of participants responded similarly for text-based therapy. Participants rated the perceived effectiveness of face-to-face therapy significantly higher than text-based therapy (z=−12.33; n=264; P<.001) and reported a significantly higher likelihood of seeking face-to-face therapy than text-based therapy (z=10.69; n=262; P<.001). There was no significant association between past use of behavioral health services and willingness to use text-based therapy (N=265, χ^2_1=0.0; P=.91). Moreover, no significant differences in likelihood of seeking text-based therapy were found by age ($F_{1,243}$=1.70; R^2=.01; P=.19), race (n=260, χ^2_4=6.0; P=.20), sexual orientation (n=247, χ^2_2=1.2; P=.54), or gender (n=257, χ^2_2=0.0; P=.99).

Figure 1. Participants' perceived effectiveness of face-to-face versus text-based therapy.

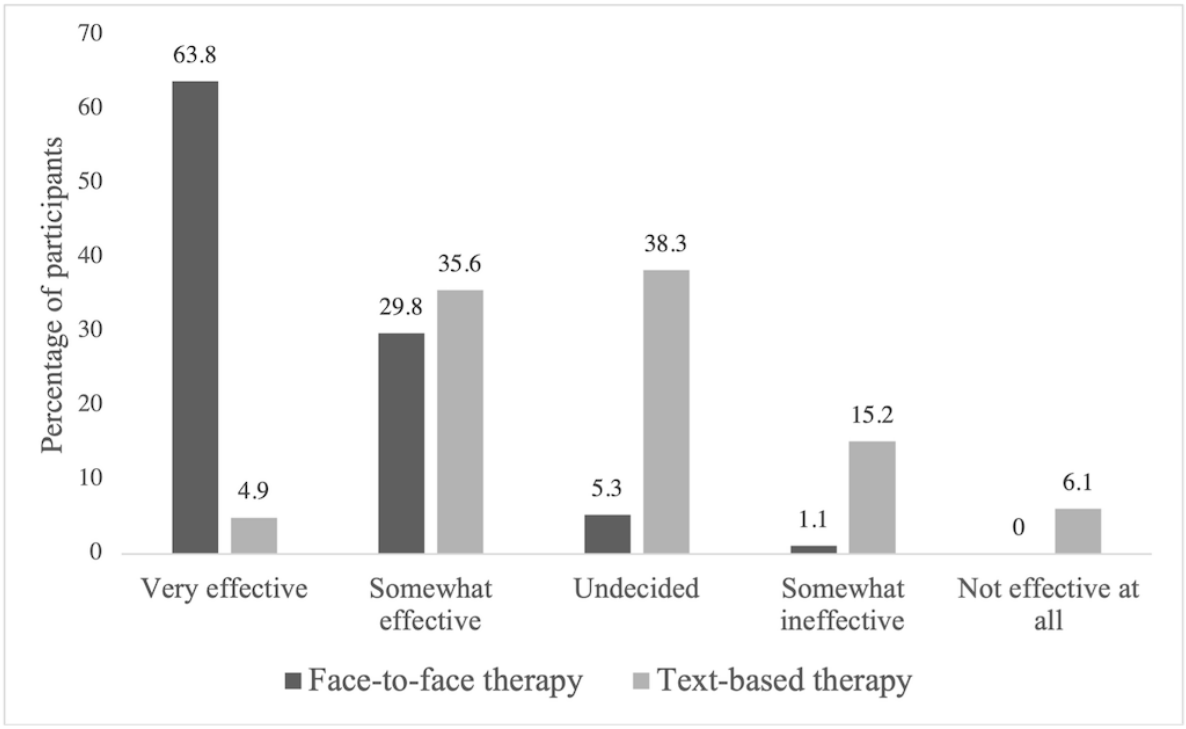
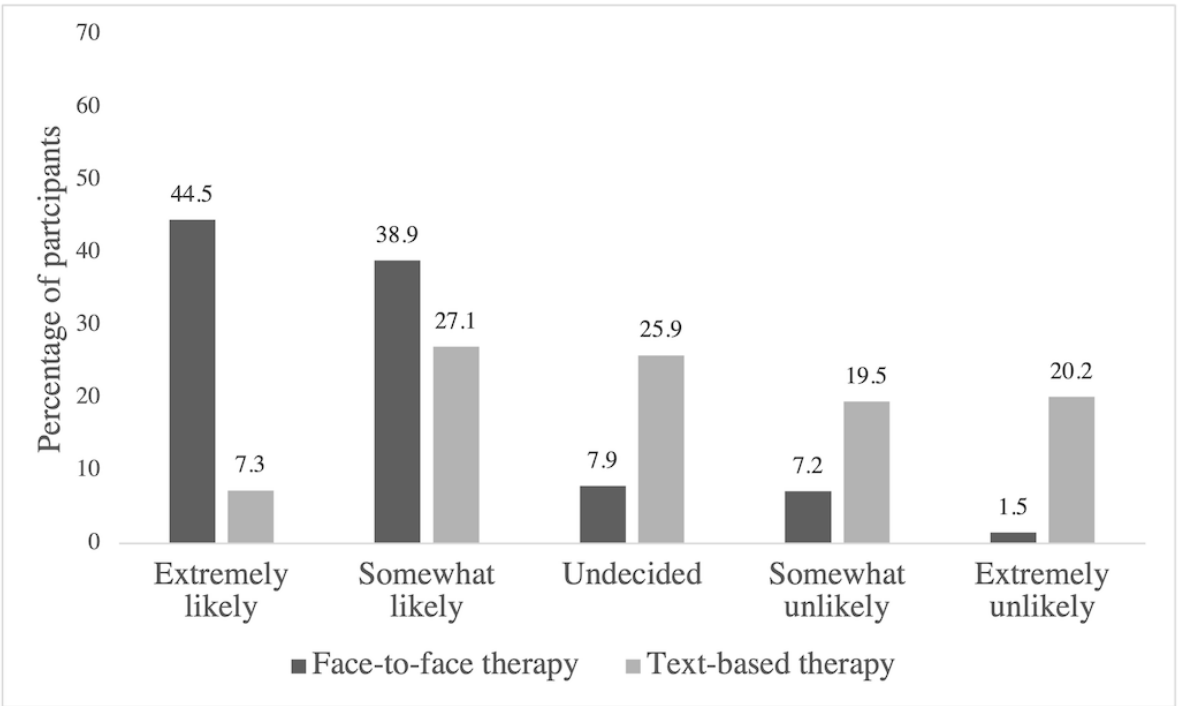


Figure 2. Participants' reported likelihood of seeking face-to-face versus text-based therapy.



Qualitative Results

Participants provided 533 responses across the 3 open-ended survey questions, from which 1090 data points were coded, reduced to 36 subthemes, and then 6 broader overarching themes (Table 4). The 6 themes that emerged are described in the following sections.

Table 4. Themes and subthemes developed from open-ended questions assessing participants' perceptions of text-based therapy.

Theme	Mentions, n (%)
Factors that promote the use of text-based therapy (advantages)	398[a] (36.51)
Convenience	215 (19.72)
Provides anonymity	42 (3.85)
Less costly	39 (3.57)
Mitigates social anxiety	37 (3.39)
Technological features	23 (2.11)
Easier to express thoughts	18 (1.65)
Time efficient	16 (1.46)
Less stigma than face-to-face therapy	4 (0.36)
Past positive experiences with text-based therapy	4 (0.36)
Factors that dissuade the use of text-based therapy (disadvantages)	325[a] (29.81)
Does not provide robust therapy experience for patients	159 (14.58)
Stunted communication	141 (12.93)
Privacy and security risks	9 (0.82)
Unsure of how it works	6 (0.55)
Technological barriers	5 (0.45)
Past negative experiences with text-based therapy	5 (0.45)
Perceived effectiveness of text-based therapy	210 (19.26)
Potential to be effective	61 (5.59)
Face-to-face therapy is more effective	43 (3.94)
Not an effective mode of communication	37 (3.39)
Effective in some scenarios	30 (2.75)
Unsure about effectiveness	23 (2.11)
Effective for younger populations	11 (1.00)
More effective than no counseling	5 (0.45)
Overall perception of text-based therapy	94 (8.62)
Increases access to mental health services	32 (2.93)
Avoids providing adequate and effective mental health care	24 (2.2)
Support implementation	23 (2.11)
Potential to be helpful	10 (0.91)
Oppose implementation	5 (0.45)
Recommendations to improve effectiveness of text-based therapy	34 (3.11)
Use as a supplement to face-to-face therapy	21 (1.92)
Various options to communicate electronically with therapist	8 (0.73)
Have consistent counselors	2 (0.18)
Use as a way to triage patients	2 (0.18)
Need to establish clear boundaries	1 (0.09)
Likelihood of using text-based therapy	29 (2.66)
Would not use it	19 (1.74)
Would use it	7 (0.64)
Would want a test trial before committing	2 (0.18)

Theme	Mentions, n (%)
Would use as a last resort	1 (0.09)

[a]Each participant could answer more than once; thus, the total may exceed 265.

Factors That Promote the Use of Text-Based Therapy

The most common overarching theme focused on factors that promote the use of text-based therapy (398/1090, 36.51%), which had 9 subthemes. Ease of finding a therapist, ability to use text-based therapy anywhere, and avoidance of long wait times at the counseling center or other providers' offices were all elements mentioned by participants when referring to the convenience of text-based therapy. For example, 1 participant explained:

> *Text-based is far more immediate. If you're struggling with something, you can get a response much more quickly. There's also a lower activation energy to setting up that service than to getting an appointment with a counselor face-to-face.*

Another said, "It would be much easier to schedule around daily life activities (as a grad student, I don't have much free time)."

Factors That Dissuade the Use of Text-Based Therapy

The second most common theme represented factors that would dissuade participants from using text-based therapy (325/1090, 29.81%), which included 6 subthemes. Responses noted that face-to-face therapy offers a "robust" experience including a personal setting, engagement from both parties, and a strong therapist-client relationship. In contrast, it was mentioned that these important characteristics are lost within text-based therapy, with one participant saying:

> *I believe that the key to the counselor/patient relationship is face-to-face empathy and rapport building. This might be more difficult to achieve via text.*

The lack of nonverbal communication was also frequently mentioned as a barrier to text-based approaches, as both therapists and patients would be unable to express and read nonverbal cues including facial expressions, emotions, and body language, leading to misinterpretation of text messages and stunted communication. This sentiment was reflected in the following participant response:

> *Text messaging simply does not convey all the same cues, signals, and emotions as face-to-face communication (e.g., voice inflexion; body language). It is not as effective a mode of communication. I am not convinced that a therapist would be able to make an accurate diagnosis or to have as deep of insights over text as they would if they were able to observe their patients face-to-face. I also do not think that I would be able to express myself and my feelings as well via text as I would in person, even if I tried.*

Perceived Effectiveness of Text-Based Therapy

The third most common theme represented participants' perceptions regarding the effectiveness of text-based therapy (210/1090, 19.26%) and included 7 subthemes. The effectiveness of text-based therapy was described as being dependent on the complexity of the issue and personal preferences of the individual seeking care. For example, 1 participant said, "For basic, straightforward issues such as stress management techniques it might be fine, but for more complex issues I think the connection of direct face-to-face communication would be better and more effective."

Similarly, another participant stated:

> *When it comes down to mental health issues, I would like to speak to someone in person. So, in this case, technology get in the way in this case.*

Overall Perception of Text-Based Therapy

The fourth most common theme focused on how participants perceived the implementation of text-based therapy as a new and potentially viable option (94/1090, 8.62%). Some participants perceived text-based therapy as a step in the right direction for increasing access to and reach of services, filling a gap in areas where there is a lack of available counselors. One participant said:

> *For students who do not have access to counseling in person, it will be very advantageous and for students who do not like to meet in person. I see its advantages and support the effort on campus.*

Conversely, many participants viewed text-based therapy as a cost-cutting measure that the university and other organizations may be using to avoid the problem of not having enough counselors available for use:

> *I appreciate that this can broaden the scope of services for graduate students and other populations (e.g., rural populations) but I also have concerns it can be a band-aid that allows people to disregard the larger problem of major disparities in access to adequate healthcare.*

Recommendations to Improve Effectiveness of Text-Based Therapy

The fifth most common theme included various strategies participants recommended to improve the effectiveness of text-based therapy (34/1090, 3.1%). Participants stated that text-based therapy should be used to augment, and not replace, face-to-face therapy. For example, 1 participant stated:

> *I think it has potential to help but that's not to say that we should be trying to do away with face-to-face counseling. This should just be something supplemental. Like if therapists/psychologists have "on-call hours" when their clients can text.*

Another stated:

> *Something is better than nothing. Not a big fan of the idea, but if that is the way to facilitate mental health support throughout the system, it is a good idea to*

> *implement it side by side with the traditional face-to-face resources.*

In addition, some participants reported they would like to see text-based therapy platforms include different features to communicate electronically, including video or audio chat with a therapist and group-based counseling.

Likelihood of Using Text-Based Therapy

The least common theme that emerged illustrated participants' likelihood of using text-based therapy (29/1090, 2.7%), which had 4 subthemes. When asked why they would not use text-based therapy, some participants referenced that they "hated texting" and others noted that they would not want to use their phone when experiencing behavioral health issues. One participant said, "I would never use text-based counseling because I don't feel comfortable putting private, sensitive information in writing."

Conversely, other participants noted that they would use text-based therapy because they used technology for everything else in their lives, with 1 participant stating:

> *I think it is a great idea especially as younger generations are becoming more and more dependent on technology and prefer communicating via text message. In addition, it allows the opportunity for individuals with very busy schedules to be able to get help.*

Discussion

Principal Findings

The aim of this study was to assess awareness and acceptability of text-based therapy for the treatment of mental and behavioral health concerns in graduate students. Approximately half of the participants reported not being aware that text-based therapy services existed, despite existing coverage by their insurance plan. Promisingly, two-thirds of participants said they would consider trying text-based therapy. Qualitative results demonstrated that participants found text-based therapy services convenient and less costly than traditional face-to-face services and that participants thought that these services provided beneficial anonymity.

Despite most participants reporting that they would consider using text-based therapy, participants were significantly less likely to report that they would seek these services compared with face-to-face therapy. This may have been driven by differences in perceived effectiveness, as participants perceived text-based therapy as significantly less effective than face-to-face therapy. Qualitative results revealed that participants had concerns over the ability of patients and therapists to develop rapport given the inability to express emotion and mood through nonverbal cues (eg, eye contact, body posture, and tone of voice). Given that rapport, or therapeutic alliance, has been established as one of the most influential factors in the success of therapy [23], this could be a key driver of the lower rates of perceived effectiveness of text-based therapy. However, there is emerging evidence that rapport can be successfully developed through text-based therapy [24] because of factors unique to text-based communication. For example, the web-based disinhibition effect may allow individuals to feel safe sharing things through digital means that they would not otherwise feel comfortable sharing in person [25,26]. Interestingly, participants in web-based therapy have reported both similar [27] and significantly higher [28] levels of therapeutic alliance than levels reported by participants in face-to-face therapy. Future work should continue to examine how the digital environment can be used to better facilitate therapeutic processes.

Qualitative results also indicated that the type of mental and behavioral health concern may also affect likelihood to engage with text-based therapy services, as text-based therapy was perceived by some respondents as more effective for some concerns versus others. Nevertheless, participants reported similar concerns for which they would seek face-to-face and text-based therapy services. Overall, the top three concerns for which participants said they would seek text-based therapy were stress management, depression, and anxiety. Given that these conditions are the most common mental health concerns in graduate students [5,29] and research demonstrating that text-based therapy can reduce anxiety and depression levels in young adults [30,31], the current results suggest that text-based therapy has potential to increase access to and the use of mental and behavioral health services among graduate students.

Approximately 60% of respondents reported previous use of behavioral health services before the onset of this study, with a significantly higher proportion of female (vs male) and non-Hispanic White (vs racial or ethnic minority) participants reporting previous use of behavioral health services. These results were consistent with lower rates of health care use observed among men and individuals from racial or ethnic minority groups in previous studies [32-34]. In contrast, there were no differences in likelihood of seeking text-based therapy by age, race or ethnicity, sexual orientation, or gender (and no association was found between past use of mental or behavioral health services and likelihood of using text-based therapy), suggesting this modality may also increase access to and use of mental and behavioral health services by individuals in groups that have historically demonstrated lower rates of use of face-to-face services [35,36].

Overall, qualitative results demonstrated that participants perceived text-based therapy as having potential to increase access to mental and behavioral health services; however, caution was suggested at the use of text-based therapy to replace, rather than supplement, face-to-face services. Participants noted that text-based therapy would be best implemented as a method of extending access to services (eg, providing services to individuals who experience barriers to attending traditional face-to-face sessions), but that face-to-face services should still be offered, especially for students experiencing more complex mental or behavioral health challenges. Some concern was also expressed regarding the use of text-based therapy as a "Band-Aid" in place of larger systemic changes that may be necessary to address existing disparities in access to mental and behavioral health services.

Strengths and Limitations

There were several strengths of this study. First, this study was the first to assess perceptions of text-based therapy services in

graduate students, a key population at high risk for the development (and worsening) of mental and behavioral health conditions [5]. Second, the mixed methods approach provided a deeper understanding of participants' perceptions of text-based therapy than would have been possible using only quantitative analysis approaches, allowing us to not only observe differences in perceived effectiveness between the services but also to provide potential avenues for future research and implementation. Finally, data were coded independently by two researchers (SAB and ANB), which minimized the impact of individual bias on thematic coding.

This study also had important limitations. The sample was comprised primarily of heterosexual, non-Hispanic White women, which limits the generalizability of results. The University of Florida has a status as a predominately White institution, with 56.6% of students identifying as White in 2019 [37]. Thus, the results may not be generalizable to institutions of other classifications, such as historically Black colleges or universities. In addition, the sample included only graduate assistants at the University of Florida. These graduate students have a position that funds their education costs and provides a health insurance plan. It is possible that there are differences in perceptions of text-based counseling between graduate assistants who have the GatorGradCare insurance plan versus other graduate students who do not work for the university, who have other forms of insurance, or who have no health insurance at all. It is also unclear how these perceptions may differ by socioeconomic status, which is a variable worth exploring in future studies. Finally, the low rates of awareness of text-based therapy may be attributed to the novelty of the modality and the recent availability of the service; text-based therapy services became available within GatorCare only 3 months before the initiation of the study [38]. GatorCare used a variety of different methodologies to increase awareness surrounding this benefit, including tabling at the new graduate student orientation and sending an email campaign to GatorGradCare subscribers; however, it appears that these initiatives were only partially successful in increasing awareness of the benefit and of text-based therapy as a therapy modality.

Conclusions

This study provided novel insights regarding perceptions of text-based therapy among graduate students. Most participants reported that they would consider using text-based therapy; however, participants reported that they were less likely to seek these services compared with traditional face-to-face therapy. Although participants felt that text-based therapy had several advantages (eg, the ability to increase access to mental and behavioral health services in both a convenient and cost-efficient way), most perceived text-based therapy as less effective than traditional face-to-face therapy. Altogether, results suggest that text-based therapy may hold potential to increase access to and use of mental and behavioral health services but that concerns remain regarding its effectiveness and how it should best be implemented. Future research should focus on methods to facilitate effective therapeutic processes in digital environments (eg, using aspects of communication unique to digital mediums to improve communication of emotion, mood, and tone) and on how text-based therapy services could be best implemented to supplement and extend, rather than replace, face-to-face services.

Acknowledgments

The authors would like to thank GatorCare at the University of Florida, along with all GatorGradCare participants who completed the study survey.

Conflicts of Interest

None declared.

References

1. Bose J. Key Substance Use and Mental Health Indicators in the United States: Results from the 2017 National Survey on Drug Use and Health. Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration; 2017:124.
2. Mental illness. National Institute of Mental Health. URL: https://www.nimh.nih.gov/health/statistics/mental-illness [accessed 2022-02-02]
3. Mathers C. The Global Burden of Disease: 2004 Update. Geneva, Switzerland: World Health Organization; 2008:34-35.
4. Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The global burden of mental, neurological and substance use disorders: an analysis from the Global Burden of Disease Study 2010. PLoS One 2015;10(2):e0116820 [FREE Full text] [doi: 10.1371/journal.pone.0116820] [Medline: 25658103]
5. Evans TM, Bira L, Gastelum JB, Weiss LT, Vanderford NL. Evidence for a mental health crisis in graduate education. Nat Biotechnol 2018 Mar 06;36(3):282-284. [doi: 10.1038/nbt.4089] [Medline: 29509732]
6. Levecque K, Anseel F, De Beuckelaer A, Van der Heyden J, Gisle L. Work organization and mental health problems in PhD students. Res Policy 2017 May;46(4):868-879. [doi: 10.1016/j.respol.2017.02.008]
7. Khoury B, Ammar J. Cognitive behavioral therapy for treatment of primary care patients presenting with psychological disorders. Libyan J Med 2014 Jan;9(1):24186 [FREE Full text] [doi: 10.3402/ljm.v9.24186] [Medline: 24690496]
8. Craighead WE, Craighead LW. The role of psychotherapy in treating psychiatric disorders. Med Clin North Am 2001 May;85(3):617-629. [doi: 10.1016/s0025-7125(05)70332-1] [Medline: 11349476]

9. Wampold BE. The Basics of Psychotherapy: An Introduction to Theory and Practice, 2nd ed. Washington, D.C., United States: American Psychological Association; 2019:x, 176.
10. Linde K, Sigterman K, Kriston L, Rücker G, Jamil S, Meissner K, et al. Effectiveness of psychological treatments for depressive disorders in primary care: systematic review and meta-analysis. Ann Fam Med 2015;13(1):56-68 [FREE Full text] [doi: 10.1370/afm.1719] [Medline: 25583894]
11. Eisenberg D, Hunt J, Speer N, Zivin K. Mental health service utilization among college students in the United States. J Nerv Ment Dis 2011 May;199(5):301-308. [doi: 10.1097/NMD.0b013e3182175123] [Medline: 21543948]
12. Harvey AG, Gumport NB. Evidence-based psychological treatments for mental disorders: modifiable barriers to access and possible solutions. Behav Res Ther 2015 May;68:1-12 [FREE Full text] [doi: 10.1016/j.brat.2015.02.004] [Medline: 25768982]
13. Czyz EK, Horwitz AG, Eisenberg D, Kramer A, King CA. Self-reported barriers to professional help seeking among college students at elevated risk for suicide. J Am Coll Health 2013;61(7):398-406 [FREE Full text] [doi: 10.1080/07448481.2013.820731] [Medline: 24010494]
14. Sutton B. The effects of technology in society and education. Digital Common. URL: https://digitalcommons.brockport.edu/cgi/viewcontent.cgi?article=1196&context=ehd_theses [accessed 2022-02-08]
15. Hull T. A preliminary study of talkspace's text-based psychotherapy. Talkspace. 2015. URL: https://terapygoperu.com/static/files/Talkspace-Preliminary-Research-Study.pdf [accessed 2022-01-25]
16. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of synchronous text-based dialogue systems in mental health interventions: systematic review. J Med Internet Res 2017 Jul 21;19(8):e267 [FREE Full text] [doi: 10.2196/jmir.7023] [Medline: 28784594]
17. Furber G, Jones GM, Healey D, Bidargaddi N. A comparison between phone-based psychotherapy with and without text messaging support in between sessions for crisis patients. J Med Internet Res 2014 Oct 08;16(10):e219 [FREE Full text] [doi: 10.2196/jmir.3096] [Medline: 25295667]
18. DellaCrosse M, Mahan K, Hull TD. The effect of messaging therapy for depression and anxiety on employee productivity. J technol behav sci 2018 Jun 26;4(1):1-5. [doi: 10.1007/s41347-018-0064-4]
19. Hull TD, Mahan K. A study of asynchronous mobile-enabled SMS text psychotherapy. Telemed J E Health 2017 Mar;23(3):240-247. [doi: 10.1089/tmj.2016.0114] [Medline: 27797646]
20. Cook BL, Trinh N, Li Z, Hou SS, Progovac AM. Trends in racial-ethnic disparities in access to mental health care, 2004-2012. Psychiatr Serv 2017 Jan 01;68(1):9-16 [FREE Full text] [doi: 10.1176/appi.ps.201500453] [Medline: 27476805]
21. Personal match. Talkspace. URL: https://tinyurl.com/excjdsye [accessed 2022-01-30]
22. Glaser BG. The constant comparative method of qualitative analysis. Soc Problems 1965 Apr;12(4):436-445. [doi: 10.2307/798843]
23. Flückiger C, Del Re AC, Wampold B, Horvath A. The alliance in adult psychotherapy: a meta-analytic synthesis. Psychotherapy (Chic) 2018 Dec;55(4):316-340 [FREE Full text] [doi: 10.1037/pst0000172] [Medline: 29792475]
24. Schwartzman CM, Boswell JF. A narrative review of alliance formation and outcome in text-based telepsychotherapy. Pract Innov 2020 Jun;5(2):128-142. [doi: 10.1037/pri0000120]
25. Bambling M, King R, Reid W, Wegner K. Online counselling: the experience of counsellors providing synchronous single-session counselling to young people. Counsel Psychother Res 2008 Jun;8(2):110-116. [doi: 10.1080/14733140802055011]
26. Dowling M, Rickwood D. Online counseling and therapy for mental health problems: a systematic review of individual synchronous interventions using chat. J Technol Human Service 2013 Jan;31(1):1-21. [doi: 10.1080/15228835.2012.728508]
27. Preschl B, Maercker A, Wagner B. The working alliance in a randomized controlled trial comparing online with face-to-face cognitive-behavioral therapy for depression. BMC Psychiatry 2011 Dec 06;11(1):189. [doi: 10.1186/1471-244x-11-189]
28. Cook JE, Doyle C. Working alliance in online therapy as compared to face-to-face therapy: preliminary results. Cyberpsychol Behav 2002 Apr;5(2):95-105. [doi: 10.1089/109493102753770480] [Medline: 12025884]
29. American College Health Association. American College Health Association-National College Health Assessment II: Graduate and Professional Students Reference Group Executive Summary Spring 2018. Silver Spring, MD: American College Health Association; 2018:13-16.
30. Benton SA, Heesacker M, Snowden SJ, Lee G. Therapist-assisted, online (TAO) intervention for anxiety in college students: TAO outperformed treatment as usual. Prof Psychol Res Pract 2016 Oct;47(5):363-371. [doi: 10.1037/pro0000097]
31. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun 06;4(2):e19 [FREE Full text] [doi: 10.2196/mental.7785] [Medline: 28588005]
32. Vaidya V, Partha G, Karmakar M. Gender differences in utilization of preventive care services in the United States. J Womens Health (Larchmt) 2012 Feb;21(2):140-145. [doi: 10.1089/jwh.2011.2876] [Medline: 22081983]
33. Bourdon JL, Moore AA, Long EC, Kendler KS, Dick DM. The relationship between on-campus service utilization and common mental health concerns in undergraduate college students. Psychol Serv 2020 Feb;17(1):118-126 [FREE Full text] [doi: 10.1037/ser0000296] [Medline: 30299149]

34. Turner JC, Keller A. College health surveillance network: epidemiology and health care utilization of college students at US 4-year universities. J Am Coll Health 2015;63(8):530-538 [FREE Full text] [doi: 10.1080/07448481.2015.1055567] [Medline: 26086428]
35. Sussman LK, Robins LN, Earls F. Treatment-seeking for depression by black and white Americans. Soc Sci Med 1987 Jan;24(3):187-196. [doi: 10.1016/0277-9536(87)90046-3]
36. Nam SK, Chu HJ, Lee MK, Lee JH, Kim N, Lee SM. A meta-analysis of gender differences in attitudes toward seeking professional psychological help. J Am Coll Health 2010;59(2):110-116. [doi: 10.1080/07448481.2010.483714] [Medline: 20864437]
37. Facts. Institutional Planning and Research- University of Florida. URL: https://ir.aa.ufl.edu/uffacts/enrollment-1/ [accessed 2022-02-05]
38. Talkspace - a free text-based therapy option for UF graduate assistants. News Notes Info ABE Graduate Students. URL: http://uf-abegrad.blogspot.com/2018/12/talkspace-free-text-based-therapy.html [accessed 2022-02-03]

Edited by L Buis; submitted 06.10.21; peer-reviewed by S Benton, L Günthner; comments to author 30.01.22; revised version received 25.03.22; accepted 02.05.22; published 07.07.22.

Please cite as:
Blair SA, Brockmann AN, Arroyo KM, Carpenter CA, Ross KM
Awareness, Acceptability, and Perceived Effectiveness of Text-Based Therapy Among Graduate Students: Cross-sectional Study
JMIR Mhealth Uhealth 2022;10(7):e34102
URL: https://mhealth.jmir.org/2022/7/e34102
doi:10.2196/34102
PMID:35797098

Original Paper

Comparison of Accelerometry-Based Measures of Physical Activity: Retrospective Observational Data Analysis Study

Marta Karas[1*], MS, PhD; John Muschelli[1*], MS, PhD; Andrew Leroux[2], MS, PhD; Jacek K Urbanek[3], MEng, PhD; Amal A Wanigatunga[4], MPH, PhD; Jiawei Bai[1], MS, PhD; Ciprian M Crainiceanu[1], MS, PhD; Jennifer A Schrack[3,4], MS, PhD

[1]Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

[2]Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado, Aurora, CO, United States

[3]Center on Aging and Health, Division of Geriatric Medicine and Gerontology, Department of Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[4]Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

[*]these authors contributed equally

Corresponding Author:
Jennifer A Schrack, MS, PhD
Department of Epidemiology
Bloomberg School of Public Health
Johns Hopkins University
615 N Wolfe St
Baltimore, MD, 21205
United States
Phone: 1 410 303 1723
Email: jschrac1@jhu.edu

Abstract

Background: Given the evolution of processing and analysis methods for accelerometry data over the past decade, it is important to understand how newer summary measures of physical activity compare with established measures.

Objective: We aimed to compare objective measures of physical activity to increase the generalizability and translation of findings of studies that use accelerometry-based data.

Methods: High-resolution accelerometry data from the Baltimore Longitudinal Study on Aging were retrospectively analyzed. Data from 655 participants who used a wrist-worn ActiGraph GT9X device continuously for a week were summarized at the minute level as ActiGraph activity count, monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, and activity intensity. We calculated these measures using open-source packages in R. Pearson correlations between activity count and each measure were quantified both marginally and conditionally on age, sex, and BMI. Each measures pair was harmonized using nonparametric regression of minute-level data.

Results: Data were from a sample (N=655; male: n=298, 45.5%; female: n=357, 54.5%) with a mean age of 69.8 years (SD 14.2) and mean BMI of 27.3 kg/m2 (SD 5.0). The mean marginal participant-specific correlations between activity count and monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, and activity were r=0.988 (SE 0.0002324), r=0.867 (SE 0.001841), r=0.913 (SE 0.00132), and r=0.970 (SE 0.0006868), respectively. After harmonization, mean absolute percentage errors of predicting total activity count from monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, and activity intensity were 2.5, 14.3, 11.3, and 6.3, respectively. The accuracies for predicting sedentary minutes for an activity count cut-off of 1853 using monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, and activity intensity were 0.981, 0.928, 0.904, and 0.960, respectively. An R software package called SummarizedActigraphy, with a unified interface for computation of the measures from raw accelerometry data, was developed and published.

Conclusions: The findings from this comparison of accelerometry-based measures of physical activity can be used by researchers and facilitate the extension of knowledge from existing literature by demonstrating the high correlation between activity count and monitor-independent movement summary (and other measures) and by providing harmonization mapping.

(JMIR Mhealth Uhealth 2022;10(7):e38077) doi:10.2196/38077

KEYWORDS

accelerometry; actigraphy; activity counts; wearable computing; monitor-independent movement summary; MIMS; physical activity; aging; older adult population; wearable device; health monitoring; digital health; wearable technology; health technology

Introduction

The use of accelerometry-based activity monitors has become increasingly popular in research studies because they provide noninvasive objective measures of physical activity, and with these monitors, physical activity data can be collected continuously for extended periods of time [1]. Modern wearable accelerometers measure acceleration of a body at a high frequency (typically 10-100 Hz). These raw data are then typically aggregated into fixed-time epochs. Yet, the choice of epoch-based measures varies across studies. For example, the Baltimore Longitudinal Study on Aging [2] used wrist-worn accelerometers and summarized data using activity counts, a measure proposed and implemented by ActiGraph [3]. Monitor-independent movement summary [4] was used for wrist-worn accelerometry data collected for the National Health and Nutrition Examination Survey (NHANES) 2011-2014 [5]. The UK Biobank study [6] used wrist-worn accelerometers and Euclidean norm minus one [7]. Additional summary measures of acceleration are mean amplitude deviation [8] and activity intensity [9].

Given the evolution of processing and analysis methods for accelerometry data over the past decade, it is important to know how new summary measures compare with established measures. Harmonizing, or mapping, values of physical activity summaries derived from different algorithms enables knowledge from the thousands of manuscripts that have been published using ActiGraph activity count [10] (and for which no repository or access to raw accelerometry data is currently available).

In this study, we aimed to (1) provide simple summaries of associations between pairs of minute-level measures (ActiGraph activity count and monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, activity intensity) and a guide for the strength of these associations in subgroups defined by demographic information; (2) provide a mapping between any 2 physical activity summary measures considered; (3) derive cut-points of open-source physical activity measures that correspond to established cut-points to estimate time spent in different physical activity intensities for activity count.

Methods

Study Design and Population

We conducted a retrospective data analysis study using data collected as part of the National Institute on Aging's Baltimore Longitudinal Study of Aging (BLSA) from participants who were community-dwelling volunteers free of all major chronic conditions and cognitive and functional impairment at the time of enrollment [2]. The data used in this work were from participants who agreed to wear an accelerometer between July 2015 and January 2019 .

Ethics Approval and Consent to Participate

The BLSA study protocol has ongoing approval from the Institutional Review Board (IRB) of the National Institute of Environmental Health Science, National Institutes of Health ("Early Markers of Alzheimer's Disease [BLSA]", IRB No. 2009-074). Informed written consent was obtained from all participants.

Accelerometry Data Collection and Export

Data had been collected with a triaxial accelerometer (ActiGraph GT9X Link; range: ±8 *g*; frequency: 80 Hz). Participants had been instructed to wear the accelerometer on their nondominant wrist for 7 days, except for periods of extended swimming or bathing. The ActiLife software (version 6.13.4) was used to (1) export data into GT3X file format, (2) derive and export minute-level ActiGraph activity count as CSV files, and (3) export raw acceleration data (in *g*) as three-dimensional time series with subsecond-level timestamps into CSV files. The ActiLife's low-frequency extension (a filtering option that decreases the lower end of the intensity threshold to increase sensitivity to low-intensity movements) was used based on recommendations and findings of greater comparability with older ActiGraph devices (model 7164) [11]. Hereon, *activity count* is used to denote ActiGraph activity count.

Raw Accelerometry Data Quality Control

We used 3 raw data quality check flags (Multimedia Appendix 1) adapted from a set of 9 flags in the NHANES protocol [12]. The selected flags subset represents intuitive flags that are meant to "determine signal patterns that were unlikely to be a result of human movement" but are not aimed at identifying nonwear [12]. A raw data observation was valid if none of the 3 flags were triggered and invalid otherwise.

Summary Measures of Raw Accelerometry Data

Commonly used minute-level measures—monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, and activity intensity (Multimedia Appendix 2 [4, 7-9])—were calculated using raw accelerometry data. With R software (version 3.6.3; The R Project), we developed and used SummarizedActigraphy R package to compute the measures. SummarizedActigraphy is a package that provides a unified data interface to compute a range of measures; it references original software for computing monitor-independent movement summary (R package: MIMSunit [13], version 0.9.2) and calibrating data for computation of Euclidean norm minus one (R package: GGIR [14], version 2.3).

Minute-Level Accelerometry Data Preprocessing

We defined minute-level data flags that represented whether the device was being worn or not using the *get_wear_flag* method (R package: arctools [15]; version 1.1.4), which

implements a wear status detection algorithm based on activity count data [16]. A given minute was classified as nonwear if it belonged to a 90-minute interval with consecutive 0-values in activity count data; otherwise, the minute was classified as wear. A given minute was valid if no raw data-level quality control flags had been triggered within the minute and it had been classified as wear, and invalid otherwise. A valid day was defined as a day (12:00 AM to 11:59 PM) with no more than 10% (144 minutes) [17] invalid minutes. Only data from participants who had at least 3 valid days of data, and only data from valid days, were included in further preprocessing and analyses.

Activity count, monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, and activity intensity data were winsorized [18] to reduce the effect of extreme values in the data set, by computing the measure-specific 0.999 quantile and then using it to replace values that exceeded this quantile.

A separate data set was constructed with imputed data, using a method described in [19]. Imputation was conducted separately for each measure: invalid minutes were replaced with corresponding values from smoothed time series produced using functional principal component analysis of the original participant- and day-specific minute-level time series (in which invalid minutes data had been denoted by *NA*). We used the *fpca.face* method (R package: refund [20], version 0.1.23) for functional principal component analysis due to its computational speed and given the large volume of data. The resulting data set was used in the summary of daily sums of measures values and in our application example where data without missing values were needed.

Statistical Data Analysis

The mean daily sums of minute-level measures were computed for each participant and then aggregated (mean and SD; median and range) across participants.

Pearson correlation coefficients for 4 pairs of measures—activity count and monitor-independent movement summary, activity count and Euclidean norm minus one, activity count and mean amplitude deviation, and activity count and activity intensity—were computed for each participant. For each pair, mean correlations and standard errors were quantified using intercept-only linear regression with participant-specific correlation as the outcome. The effects of demographic characteristics (covariates: age, BMI, and sex) on correlations were estimated using adjusted linear regression with participant-specific correlation as the outcome and α=.05 to determine the statistical significance of coefficients. This procedure was repeated for secondary analyses with a subsample (participants' age ≤65 years).

Harmonization

Mapping

To derive the harmonization mapping, relationships were estimated using generalized additive modeling for each pair of measures. The generalized additive models were chosen to allow flexible adaptation to the data rather than imposing a particular functional form of the fit. In each model, the outcome was a minute-level measure (monitor-independent movement summary, or Euclidean norm minus one, or mean amplitude deviation, or activity intensity), and a smooth term of minute-level activity count was set as a predictor. For the smooth term, cubic regression splines with a basis dimension equal to 30 were used to allow a flexible relationship between the measure and activity count. Models were estimated with nonparametric smoothing (method: *gam*; R package: mgcv [21], version 1.8.34). Smoothness of the nonlinear effects was enforced via a second derivative penalty, and parameter selection was performed using cross-validation [22]. Data from all participants' valid minutes were used in the model fitting except for minutes, which had activity count values equal 0. The activity count=0 exclusion was motivated by a large proportion of zero values, and the need to estimate the relation for small activity count values without it being inflated by the large number of zeros. Relationships were estimated as strictly monotonic (without monotonicity having been constrained explicitly). The generalized additive model was used to provide values for 2-way mapping between activity count and each measure. All measurements were mapped into activity count, where $\hat{f}(x)$ represents the activity count value estimated by mapping the x value of a measure, where *measure* represents monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, or activity intensity.

Evaluation

To assess mapping accuracy in estimating physical activity volume statistics, total activity count (the sum of minute-level activity count values from a day) was computed for each participant, using activity count data and $\hat{f}$, and the difference was defined the estimation error. Estimation error was summarized by calculating mean percentage error (MPE), mean absolute percentage error (MAPE), median percentage error, and median absolute percentage error for each participant and aggregated across participants (mean and SD).

To assess whether mapping accuracy depended on participant activity level, MPE values were plotted against the participant's average total activity count.

The utility of the mapping for classifying minutes into various activity intensity classes was assessed. We used activity count cut-offs derived to (1) separate sedentary and active minutes in data collected with a sensor worn on nondominant wrist in older adults [23], (2) separate sedentary from light and (3) light from moderate-to-vigorous activity intensity levels in data collected with a sensor worn on a nondominant wrist in young to older adults [24]. In the classification task, for each minute, the true value was defined based on whether activity count > cut-off, and the predicted value was defined based on whether $\hat{f}$ > cut-off. Accuracy, sensitivity, and specificity were computed for each participant and aggregated across participants (mean and SD).

Minute-Level Patterns of Daily Physical Activity

Minute-level activity count and $\hat{f}$ were used to estimate smoothed 24-hour time series of median activity count for age

groups <60 years, 60-67 years, 68-74 years, and ≥75 years, for which 24-hour time series of median activity count have previously been published [25]. Activity count–based and -based estimates were compared by calculating MAPE defined as sum of absolute value of the difference between a pair of estimates divided by sum of activity count-based estimates.

Results

Population Characteristics

Data from 655 individuals (Table 1) were included in the analyses. The mean age was 69.8 (SD 14.2, range 22-97) years. There was a higher proportion of women (357/655, 54.5%) than men (298/655, 45.5%). The racial composition reflected that of the BLSA enrollment [2]. Of the 655 participants, 445 participants (67.9%) were White, 157 (24%) were Black, 44 (6.7%) were classified as other race, and 9 participants (1.4%) did not provide this information. Almost 96% of participants (628/655, 95.9%) self-reported good, very good, or excellent health. The prevalences of hypertension, high blood cholesterol levels, and osteoarthritis were 43.5% (285/655), 52.8% (346/655), and 48.2% (316/655), respectively. Participants had a median of 6 (range 3-7) days of valid accelerometry data; for valid days, participants had a mean of 1438 (SD 8) valid minutes (out of 1440 possible minutes per day).

The mean participant daily sums (Table 2) were 2,204,169 (SD 600,965) for activity count, 11,299.7 (SD 2766.0) for monitor-independent movement summary, 47.7 (SD 13.3) for mean amplitude deviation, 30.9 (SD 9.1) for Euclidean norm minus one, and 4157.6 (SD 1068.8) for activity intensity.

Table 1. Study sample (N=655) characteristics.

Characteristic	Value
Sociodemographic	
Age	
Mean (SD)	69.8 (14.2)
Median (range)	72.0 (22.0-97.0)
Weight (kg)	
Mean (SD)	77.4 (17.1)
Median (range)	76.3 (41.1-142.7)
Height (cm)	
Mean (SD)	168.0 (9.2)
Median (range)	167.3 (143.8-196.2)
BMI	
Mean (SD)	27.3 (5.0)
Median (range)	26.6 (17.1-52.5)
Sex	
Female count (%)	357 (54.5)
Male count (%)	298 (45.5)
Race	
White count (%)	445 (67.9)
Black count (%)	157 (24.0)
Chinese count (%)	30 (4.6)
Hawaiian count (%)	11 (1.7)
Other non-White count (%)	3 (0.5)
Not reported count (%)	9 (1.4)
Sensor wear	
Valid days	
Mean (SD)	5.9 (0.4)
Median (range)	6.0 (3.0, 7.0)
Nonwear minutes (/day)	
Mean (SD)	2.0 (7.8)
Median (range)	0.0 (0.0, 77.0)
Valid minutes (/day)	
Mean (SD)	1437.8 (8.0)
Median (range)	1440.0 (1361.7-1440.0)
Health	
Self-reported health	
Good, very good, or excellent count (%)	628 (95.9)
Fair or poor count (%)	22 (3.4)
Not reported count (%)	5 (0.8)
Medical history	
Myocardial infarction, congestive heart failure, ischemic chest pain, vascular procedure, or peripheral artery disease count (%)	55 (8.4)
Hypertension count (%)	285 (43.5)

Characteristic	Value
High blood cholesterol count (%)	346 (52.8)
Stroke or transient ischemic attack count (%)	34 (5.2)
Pulmonary disease count (%)	74 (11.3)
Diabetes count (%)	95 (14.5)
Cancer count (%)	191 (29.2)
Osteoarthritis count (%)	316 (48.2)

Table 2. Mean daily sum values for physical activity measures.

Measure	Value
Activity count	
Mean (SD)	2,204,169 (600,965)
Median (range)	2,157,496 (731,945-5,071,196)
Monitor-independent movement summary	
Mean (SD)	11,299.7 (2766.0)
Median (range)	11,195.2 (4252.3-23,931.5)
Mean amplitude deviation	
Mean (SD)	47.7 (13.3)
Median (range)	46.3 (16.1-108.1)
Euclidean norm minus one	
Mean (SD)	30.9 (9.1)
Median (range)	29.6 (11.8-75.3)
Activity intensity	
Mean (SD)	4157.6 (1068.8)
Median (range)	4085.5 (1529.7-9418.6)

Correlations Between Minute-Level Summary Statistics

Monitor-independent movement summary was most correlated with activity count (estimated mean 0.988, SE 0.0002), closely followed by activity intensity (estimated mean 0.970, SE 0.0007, mean amplitude deviation (estimated mean 0.913, SE 0.0013), and Euclidean norm minus one (estimated mean 0.867, SE 0.0018) (Table 3).

The estimated effects of age (with female as the reference level) were not statistically significant in the models for activity count and monitor-independent movement summary (P=.97), activity count and mean amplitude deviation (P=.64), and activity count and activity intensity (P=.64), and were statistically significant in the model for activity count and Euclidean norm minus one (P<.001). The estimated effects of BMI on correlations were statistically significant for correlations between activity count and mean amplitude deviation (estimate 0.001, SE 0.0003, P=.001) and those between activity count and activity intensity (estimate 0.000278, SE 0.0001, P=.04). The estimated effects of sex (with female as the reference level) were statistically significant in the models for activity count and monitor-independent movement summary (estimate –0.002, SE 0.0005, P<.001), activity count and mean amplitude deviation (estimate –0.01, SE 0.0026, P<.001), and activity count and activity intensity (estimate –0.01, SE 0.0013, P<.001).

The results of secondary analysis (Table S1 in Multimedia Appendix 3) closely follow the results obtained from the full sample (Table 3) for both unadjusted (activity count and monitor-independent movement summary: difference 0; activity count and Euclidean norm minus one: difference –0.06; activity count and mean amplitude deviation: difference 0.02; activity count and activity intensity: difference 0.01) and adjusted models.

Table 3. Summary of intercept-only linear regression and adjusted linear regression with outcome defined as participant-specific correlation between activity count and other measures (monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, or activity intensity).

Model and response variable[a]	Intercept	Age		BMI		Sex[b]	
	Estimate (SE)	Estimate (SE)	*P* value	Estimate (SE)	*P* value	Estimate (SE)	*P* value
Unadjusted							
Monitor-independent movement summary	0.988042 (0.000232)	—[c]	—	—	—	—	—
Euclidean norm minus one	0.867158 (0.001841)	—	—	—	—	—	—
Mean amplitude deviation	0.913412 (0.001320)	—	—	—	—	—	—
Activity intensity	0.969984 (0.000687)	—	—	—	—	—	—
Adjusted for age, BMI, and sex							
Monitor-independent movement summary	0.987969 (0.001744)	0.000001 (0.000016)	.97	0.000032 (0.000046)	.48	–0.001859 (0.000466)	<.001
Euclidean norm minus one	0.886566 (0.013766)	–0.000532 (0.000129)	<.001	0.000653 (0.000363)	.07	–0.000206 (0.003678)	.96
Mean amplitude deviation	0.892177 (0.009852)	0.000044 (0.000092)	.64	0.000840 (0.000260)	.001	–0.010410 (0.002632)	<.001
Activity intensity	0.962364 (0.005016)	0.000063 (0.000047)	.18	0.000278 (0.000132)	.04	–0.009576 (0.001340)	<.001

[a]Correlation with activity count.

[b]Female was used as the reference.

[c]Not included in the model.

Mapping Between Minute-Level Summary Measures

Model Fit

Figure 1 shows the estimated association between minute-level activity count (x-axis) and minute-level monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, and activity intensity (y-axis). The black solid line represents fitted values obtained from generalized additive models.

For a widely used activity count cut-off 1853 [23], the corresponding cut-offs (Table 4) were 10.558 (monitor-independent movement summary), 0.022 (Euclidean norm minus one), 0.039 (mean amplitude deviation), and 3.620 (activity intensity).

Figure 1. Estimated minute-level mapping. A black solid line shows generalized additive model–fitted values of a measure (monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, activity intensity) given the activity count value. The points represent a subset of the data created by taking every 100th observations from all participant- and minute-specific observations; this subset is the same for all 4 plots. AC: activity count; AI: activity intensity; ENMO: Euclidean norm minus one; MAD: mean amplitude deviation; MIMS: monitor-independent movement summary.

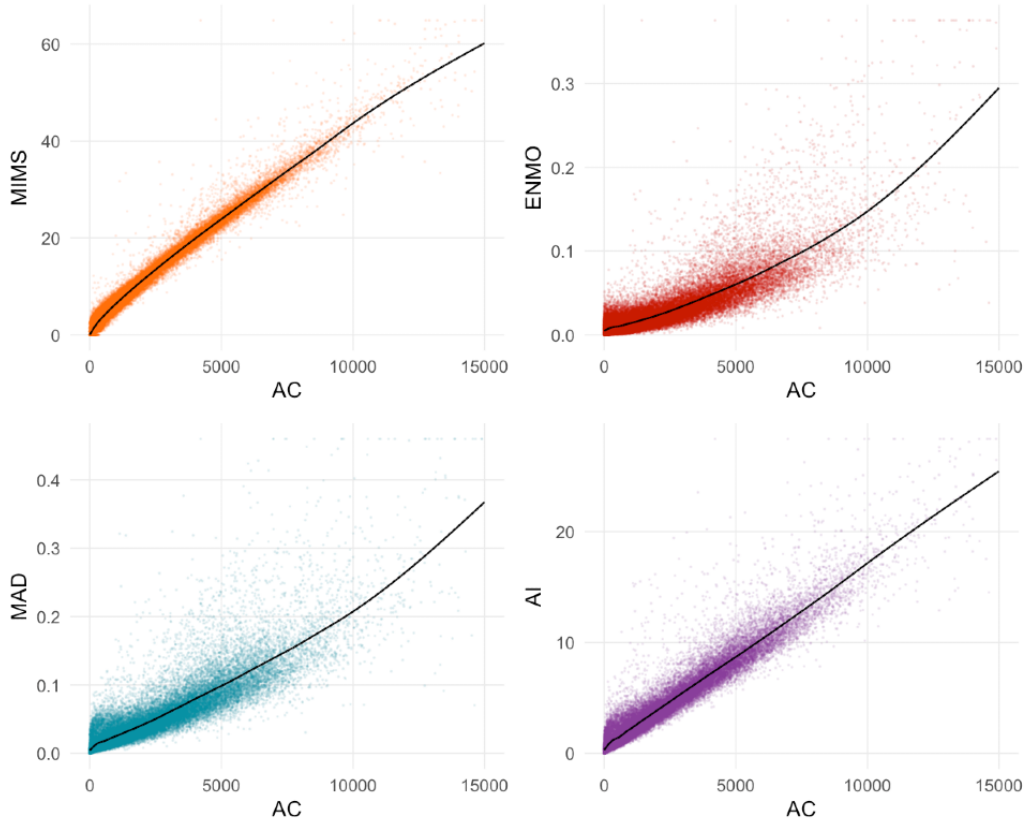


Table 4. Corresponding values of each measure for activity count cut-off values.

Method	Activity count cut-off value	Corresponding value			
		Monitor-independent movement summary	Euclidean norm minus one	Mean amplitude deviation	Activity intensity
Separate sedentary and active in older adults [23]	1853	10.558	0.022	0.039	3.620
Separate sedentary and light activity in young to older adults [24]	2860	15.047	0.033	0.057	5.273
Separate light and moderate-to-vigorous activity in young to older adults [24]	3940	19.614	0.046	0.078	7.025

Mapping Evaluation

In the task of estimating total activity count, MAPE values were lowest for monitor-independent movement summary (mean 2.5, SD 2.4), followed by activity intensity (mean 6.3, SD 5.1), mean amplitude deviation (mean 11.3, SD 8.4), and Euclidean norm minus one (mean 14.3, SD 10.3). MPE values were similar for monitor-independent movement summary (mean 0.2, SD 3.2), activity intensity (mean 0.3, SD 7.6), mean amplitude deviation (mean –0.3, SD 13.3), and Euclidean norm minus one (mean 4.6, SD 16.1). The findings for median absolute percentage error and median percentage error were similar to those for MAPE and MPE, respectively (Table S2 in Multimedia Appendix 3).

Based on visual inspection, there was larger variability in MPE values among participants with smaller mean total activity count values, but there was no apparent tendency for lower or higher MPE values based on participants' average total activity counts (Figure S1 in Multimedia Appendix 3).

In the task of predicting whether the activity count for a given minute was above a certain cut-off, for the cut-off equal 1853, participant-specific classification accuracy (Table S3 in Multimedia Appendix 3) was the highest for monitor-independent movement summary (mean 0.981, SD 0.005), followed by activity intensity (mean 0.960, SD 0.012), mean amplitude deviation (mean 0.928, SD 0.021), and Euclidean norm minus one (mean 0.904, SD 0.028). Overall, the accuracy of predicting whether the activity count for a given

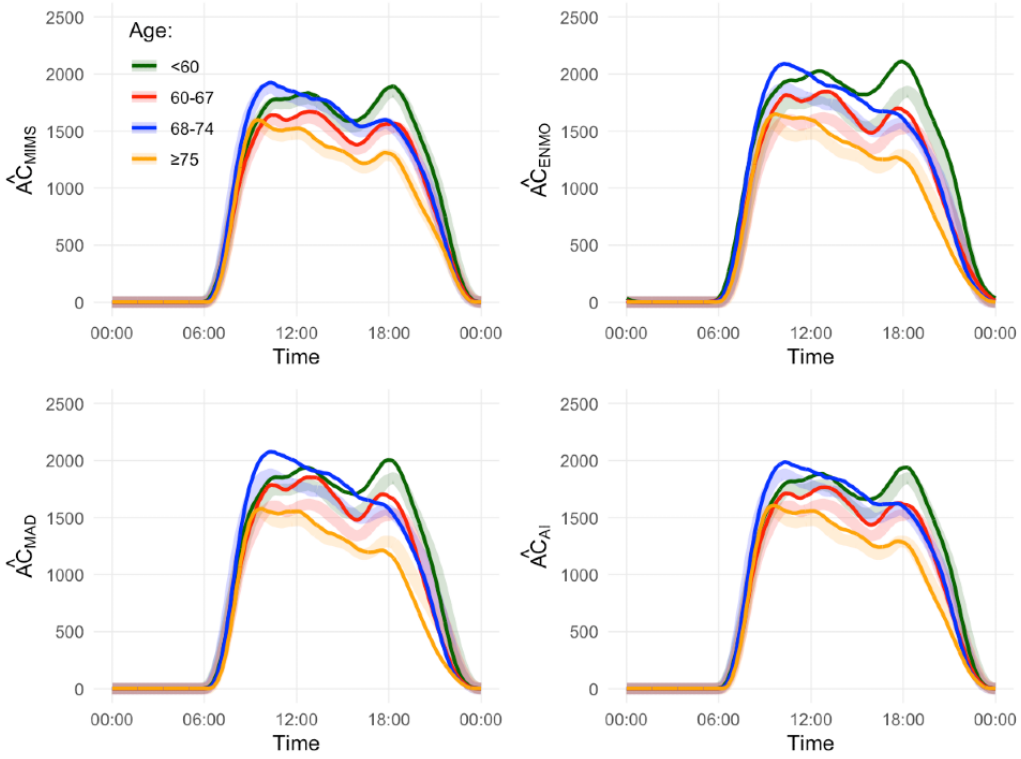
minute was above a certain cut-off was better for higher activity count cut-off values (ie, accuracy was higher for predicting whether a given minute has activity count >3940 than for predicting whether a given minute activity count >2860; Table S3 in Multimedia Appendix 3). This is consistent with our observation that the variability along the estimated mapping is lower for higher activity values (Figure S1 in Multimedia Appendix 3).

Minute-Level Patterns of Daily Physical Activity

Figure 2 shows the estimated smoothed 24-hour median activity counts across the previously published age groups: <60-year old (green; N = 140), 60- to 67-year old (red; N = 102), 68- to 74-year old (blue; N = 129), ≥ 75-year old (orange; N = 284). Semi-transparent thick colour lines represent results obtained with activity count.Solid thin colour lines represent results obtained with .

The -based curves yielded roughly the same information as the activity count-based curves [25] for each age group (<60 years: n=140; 60-67 years: n=102; 68-74 years: n=129; ≥75 years: n=284). MAPE for activity count-based and -based estimates was the lowest for monitor-independent movement summary (MAPE 3.2), followed by activity intensity (MAPE 6.7), mean amplitude deviation (MAPE 11.1), and Euclidean norm minus one (MAPE 12.5).

Figure 2. Smoothed 24-hour median activity counts per minute for each age group: <60 years (green), 60-67 years (red), 68-74 years (blue), and ≥75 years (orange). Semitransparent thick colored lines represent results obtained with activity count; they are the same for all 4 plots. Solid thin colored lines represent results obtained with values mapped into activity count from monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, or activity intensity. AC: activity count; AI: activity intensity; ENMO: Euclidean norm minus one; MAD: mean amplitude deviation; MIMS: monitor-independent movement summary.



Discussion

Principal Results

Correlations between activity count and the other raw data summary metrics were all large (mean $r \geq 0.87$) and were especially high for monitor-independent movement summary and activity intensity (mean $r \geq 0.97$) (Table 3). After harmonization, monitor-independent movement summary allowed for excellent accuracy in predicting total activity count and sedentary minutes using a cut-off that corresponded to an activity count cut-off determined using [23]. Our analysis is especially timely given the recent release of physical activity data from NHANES 2011-2014 that uses the open-source monitor-independent movement summary measure.

To the best of our knowledge, the correlation between activity count and monitor-independent movement summary in continuous data collected in the free-living environment has not been previously explored. The activity count measure had the highest mean participant-specific correlation with monitor-independent movement summary (mean $r = 0.988$), closely followed by activity intensity (mean $r=0.97$), and mean amplitude deviation (mean $r=0.913$) and Euclidean norm minus one (mean $r=0.867$). Both monitor-independent movement summary and activity intensity measures are based on variability within each dimension, whereas mean amplitude deviation and Euclidean norm minus one are based on the Euclidean norm of three-dimensional data; therefore, it is consistent with expectations that monitor-independent movement summary and activity intensity behave similarly and demonstrate similar

correlations with activity count. While we found there were statistically significant effects of age (in the model for correlation between activity count and Euclidean norm minus one: $P<.001$), BMI (in the model for correlation between activity count and mean amplitude deviation: $P=.001$; in the model for correlation between activity count and activity intensity: $P=.04$), and sex (in the model for correlation between activity count and monitor-independent movement summary: $P<.001$; in the model for activity count and mean amplitude deviation: $P<.001$; in the model for activity count and activity intensity: $P<.001$), the effect sizes were of very small magnitude. In particular, the analysis showed that monitor-independent movement summary had a correlation with activity count that did not differ significantly for age ($P=.97$) or BMI ($P=.48$), and differed significantly ($P<.001$) between men and women by a magnitude of 0.002. The results from secondary analysis, with a subsample of the youngest participants (participants of age 65 years or less; 31.9% of the full sample), were similar to those from the full sample.

Harmonization mapping can be particularly useful to translate commonly used cut-off values of physical activity intensity levels from activity count into measures implemented in open-source software. For the tasks of predicting sedentary minutes for an activity count cut-off of 1853 [23], we observed excellent accuracy for monitor-independent movement summary (accuracy 0.981) and activity intensity (accuracy 0.960). The utility of the derived mapping was demonstrated in the example in which previous findings [25] were replicated. The physical activity volume daily trajectories for age groups obtained with activity count were closely matched with those from the measures, with monitor-independent movement summary yielding visually almost identical results (MAPE 3.2), followed by activity intensity (MAPE 6.7), mean amplitude deviation (MAPE 11.1), and Euclidean norm minus one (MAPE 12.5).

To the best of our knowledge, we are the first to provide freely available R software (SummarizedActigraphy R package) with a unified interface for computation of the 4 open-source measures from raw accelerometry data, with complicated mathematical formulas distilled into a reader-friendly text (Multimedia Appendix 2).

Limitations

First, the data were from a sample that consisted of predominantly middle-aged to older adults (Table 1). However, we observed that (1) the level of activity of adults in the sample ranged from sedentary to moderate and vigorous activity, (2) mapping results did not exhibit any trend based on the average level of the participant's physical activity, and (3) the variability of estimates was lower for higher activity values, which suggests that mapping could prove useful in future studies with younger (more active) populations [25].

Second, physical activity measures were computed using raw accelerometry data collected at a frequency of 80 Hz. While this frequency matches that of physical activity data from NHANES 2011-2014 [12] that uses the monitor-independent movement summary measure, caution should be used in adapting our harmonization mapping to raw data collected at a different frequency.

Third, data had been collected with sensors worn on the nondominant wrist only. While we expect the results to be generalizable to data from sensors worn on the dominant wrist, we presume that correlations and mapping would not be applicable to chest- or hip-worn sensors, because physical activity volume statistics (eg, total activity count) calculated from raw data collected by these devices are expected to be substantially lower than when measured at wrist.

Fourth, harmonization mapping was estimated using generalized additive modeling, which does not offer an easy, closed-form formula of the transformation. While a closed-form formula could be obtained using polynomial regression models, the choice of generalized additive models allowed for thorough estimation of a relationship between activity count and other measures in a more flexible way.

Finally, our results may be conditional upon the data preprocessing methods used; however, we believe that the steps we performed are commonly done [17,19] and are reasonable given the obtained data summary statistics and visual quality checks performed.

Conclusions

Activity count was highly correlated with monitor-independent movement summary, Euclidean norm minus one, mean amplitude deviation, and activity intensity. Mapping provides a way to harmonize accelerometry data sets with different summary measures; however, further research is warranted to test the validity of mapping with data collected at a different frequency or from different body locations.

Acknowledgments

This work was funded by National Institute on Aging (grant U01AG057545) and a Johns Hopkins University Catalyst Award.

Data Availability

Data used in this manuscript were collected as part of the National Institute on Aging's Baltimore Longitudinal Study of Aging. A project repository with all data preprocessing and analysis code is publicly available on GitHub (muschellij2/blsa_mims).

Authors' Contributions

MK implemented the final version of the analyses and took the lead in finalizing the manuscript. JM co-initiated the work, wrote the first draft of this manuscript and the analyses, and authored the *SummarizedActigraphy* package. CMC provided scientific

input and provided major edits to the manuscript draft. JAS co-initiated the work, provided scientific input, and led the interpretation of results. AL, JKU, AAW, and JB provided scientific input and helped with the interpretation of results. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

Conflicts of Interest

CMC is a consultant for Bayer and Johnson and Johnson. Both these consulting contracts have been disclosed through the Johns Hopkins University Edisclose system. The current manuscript is not related to or influenced by any of these contracts.

Multimedia Appendix 1
Raw accelerometry data quality control.
[DOCX File , 24 KB - mhealth_v10i7e38077_app1.docx]

Multimedia Appendix 2
Open-source summary measures of raw accelerometry data.
[DOCX File , 30 KB - mhealth_v10i7e38077_app2.docx]

Multimedia Appendix 3
Results.
[DOCX File , 835 KB - mhealth_v10i7e38077_app3.docx]

References

1. Karas M, Bai J, Strączkiewicz M, Harezlak J, Glynn NW, Harris T, et al. Accelerometry data in health research: challenges and opportunities. Stat Biosci 2019 Jul;11(2):210-237 [FREE Full text] [doi: 10.1007/s12561-018-9227-2] [Medline: 31762829]
2. Kuo P, Schrack JA, Shardell MD, Levine M, Moore AZ, An Y, et al. A roadmap to build a phenotypic metric of ageing: insights from the Baltimore Longitudinal Study of Aging. J Intern Med 2020 Apr;287(4):373-394 [FREE Full text] [doi: 10.1111/joim.13024] [Medline: 32107805]
3. Neishabouri A, Nguyen J, Samuelsson J, Guthrie T, Biggs M, Wyatt J, et al. Quantification of acceleration as activity counts in ActiGraph wearables. Research Square Preprint posted online on February 22, 2022. [doi: 10.21203/rs.3.rs-1370418/v1]
4. John D, Tang Q, Albinali F, Intille S. An open-source monitor-independent movement summary for accelerometer data processing. J Meas Phys Behav 2019 Dec;2(4):268-281. [doi: 10.1123/jmpb.2018-0068] [Medline: 34308270]
5. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Br J Sports Med 2014 Jul;48(13):1019-1023 [FREE Full text] [doi: 10.1136/bjsports-2014-093546] [Medline: 24782483]
6. Doherty A, Jackson D, Hammerla N, Plötz T, Olivier P, Granat MH, et al. Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. PLoS One 2017;12(2):e0169649 [FREE Full text] [doi: 10.1371/journal.pone.0169649] [Medline: 28146576]
7. van Hees VT, Gorzelniak L, Dean León EC, Eder M, Pias M, Taherian S, et al. Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. PLoS One 2013 Apr 23;8(4):e61691 [FREE Full text] [doi: 10.1371/journal.pone.0061691] [Medline: 23626718]
8. Vähä-Ypyä H, Vasankari T, Husu P, Suni J, Sievänen H. A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. Clin Physiol Funct Imaging 2015 Jan 07;35(1):64-70. [doi: 10.1111/cpf.12127] [Medline: 24393233]
9. Bai J, He B, Shou H, Zipunnikov V, Glass TA, Crainiceanu CM. Normalization and extraction of interpretable metrics from raw accelerometry data. Biostatistics 2014 Jan 01;15(1):102-116 [FREE Full text] [doi: 10.1093/biostatistics/kxt029] [Medline: 23999141]
10. ActiGraph Corp. URL: https://actigraphcorp.com/about-us/#:~:text=Appearing%20in%20more%20than%2017%2C000,grade%20device%20of%20its%20kind [accessed 2022-06-27]
11. Cain KL, Conway TL, Adams MA, Husak LE, Sallis JF. Comparison of older and newer generations of ActiGraph accelerometers with the normal filter and the low frequency extension. Int J Behav Nutr Phys Act 2013 Apr 25;10(1):51 [FREE Full text] [doi: 10.1186/1479-5868-10-51] [Medline: 23618461]
12. National Health and Nutrition Examination Survey 2011-2012 data documentation, codebook, and frequencies. Centers for Disease Control and Prevention. 2022. URL: https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/PAXMIN_G.htm [accessed 2022-06-13]
13. John D, Tang Q, Albinali F, Intille S. MIMSunit: algorithm to compute monitor independent movement summary unit. The Comprehensive R Archive Network. 2019. URL: https://CRAN.R-project.org/package=MIMSunit [accessed 2022-07-04]
14. van Hees VT, Fang Z, Zhao JH, Heywood J, Mirkes E, Sabia S, et al. GGIR: raw accelerometer data analysis. The Comprehensive R Archive Network. 2022. URL: https://CRAN.R-project.org/package=GGIR [accessed 2022-07-04]

15. Karas M, Schrack J, Urbanek J. arctools: processing and physical activity summaries of minute level activity data. The Comprehensive R Archive Network. 2021. URL: https://CRAN.R-project.org/package=arctools [accessed 2022-07-04]
16. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364 [FREE Full text] [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]
17. Wanigatunga A, Wang H, An Y, Simonsick EM, Tian Q, Davatzikos C, et al. Association between brain volumes and patterns of physical activity in community-dwelling older adults. J Gerontol A Biol Sci Med Sci 2020 Jul 13;76(8):1504-1511. [doi: 10.1093/gerona/glaa294] [Medline: 33230557]
18. Hastings C, Mosteller F, Tukey JW, Winsor CP. Low moments for small samples: a comparative study of order statistics. Ann Math Statist 1947 Sep;18(3):413-426. [doi: 10.1214/aoms/1177730388]
19. Leroux A, Xu S, Kundu P, Muschelli J, Smirnova E, Chatterjee N, et al. Quantifying the predictive performance of objectively measured physical activity on mortality in the UK Biobank. J Gerontol A Biol Sci Med Sci 2021 Jul 13;76(8):1486-1494. [doi: 10.1093/gerona/glaa250] [Medline: 33000171]
20. Goldsmith J, Scheipl F, Huang L, Wrobel J, Di C, Gellar J, et al. refund: regression with functional data. The Comprehensive R Archive Network. 2021. URL: https://CRAN.R-project.org/package=refund [accessed 2022-07-04]
21. Wood SN. mgcv: mixed GAM computation vehicle with automatic smoothness estimation. The Comprehensive R Archive Network. 2011. URL: https://CRAN.R-project.org/package=mgcv [accessed 2022-07-04]
22. Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J R Stat Soc Series B Stat Methodol 2011;73(1):3-36. [doi: 10.1111/j.1467-9868.2010.00749.x]
23. Koster A, Shiroma EJ, Caserotti P, Matthews CE, Chen KY, Glynn NW, et al. Comparison of sedentary estimates between activPAL and hip- and wrist-worn ActiGraph. Med Sci Sports Exerc 2016 Aug;48(8):1514-1522. [doi: 10.1249/MSS.0000000000000924] [Medline: 27031744]
24. Montoye AH, Clevenger KA, Pfeiffer KA, Nelson MB, Bock JM, Imboden MT, et al. Development of cut-points for determining activity intensity from a wrist-worn ActiGraph accelerometer in free-living adults. J Sports Sci 2020 Nov 17;38(22):2569-2578. [doi: 10.1080/02640414.2020.1794244] [Medline: 32677510]
25. Schrack JA, Zipunnikov V, Goldsmith J, Bai J, Simonsick EM, Crainiceanu C, et al. Assessing the "physical cliff": detailed quantification of age-related differences in daily patterns of physical activity. J Gerontol A Biol Sci Med Sci 2014 Aug 14;69(8):973-979 [FREE Full text] [doi: 10.1093/gerona/glt199] [Medline: 24336819]

Abbreviations

BLSA: Baltimore Longitudinal Study on Aging
MAPE: mean absolute percentage error
MPE: mean percentage error
NHANES: National Health and Nutrition Examination Survey

Edited by L Buis; submitted 17.03.22; peer-reviewed by V Yadav, X Li; comments to author 08.04.22; revised version received 04.05.22; accepted 10.05.22; published 22.07.22.

Please cite as:
Karas M, Muschelli J, Leroux A, Urbanek JK, Wanigatunga AA, Bai J, Crainiceanu CM, Schrack JA
Comparison of Accelerometry-Based Measures of Physical Activity: Retrospective Observational Data Analysis Study
JMIR Mhealth Uhealth 2022;10(7):e38077
URL: https://mhealth.jmir.org/2022/7/e38077
doi:10.2196/38077
PMID:35867392

Original Paper

Data Management and Privacy Policy of COVID-19 Contact-Tracing Apps: Systematic Review and Content Analysis

Marco Bardus[1,2], BA, MA, PhD; Melodie Al Daccache[3], BSc, MSc; Noel Maalouf[4,5], BE, PhD; Rayan Al Sarih[5], BE; Imad H Elhajj[5], MSc, PhD

[1]Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom

[2]Department of Health Promotion and Community Health, Faculty of Health Sciences, American University of Beirut, Beirut, Lebanon

[3]Center for Research on Population and Health, Faculty of Health Sciences, American University of Beirut, Beirut, Lebanon

[4]Department of Electrical and Computer Engineering, School of Engineering, Lebanese American University, Byblos, Lebanon

[5]Department of Electrical and Computer Engineering, Maroun Semaan Faculty of Engineering and Architecture, American University of Beirut, Beirut, Lebanon

Corresponding Author:
Marco Bardus, BA, MA, PhD
Institute of Applied Health Research
College of Medical and Dental Sciences
University of Birmingham
Edgbaston
Birmingham, B15 2TT
United Kingdom
Phone: 44 0121 414 3344
Email: marco.bardus@gmail.com

Abstract

Background: COVID-19 digital contact-tracing apps were created to assist public health authorities in curbing the pandemic. These apps require users' permission to access specific functions on their mobile phones, such as geolocation, Bluetooth or Wi-Fi connections, or personal data, to work correctly. As these functions have privacy repercussions, it is essential to establish how contact-tracing apps respect users' privacy.

Objective: This study aimed to systematically map existing contact-tracing apps and evaluate the permissions required and their privacy policies. Specifically, we evaluated the type of permissions, the privacy policies' readability, and the information included in them.

Methods: We used custom Google searches and existing lists of contact-tracing apps to identify potentially eligible apps between May 2020 and November 2021. We included contact-tracing or exposure notification apps with a Google Play webpage from which we extracted app characteristics (eg, sponsor, number of installs, and ratings). We used Exodus Privacy to systematically extract the number of permissions and classify them as *dangerous* or *normal*. We computed a Permission Accumulated Risk Score representing the threat level to the user's privacy. We assessed the privacy policies' readability and evaluated their content using a 13-item checklist, which generated a Privacy Transparency Index. We explored the relationships between app characteristics, Permission Accumulated Risk Score, and Privacy Transparency Index using correlations, chi-square tests, or ANOVAs.

Results: We identified 180 contact-tracing apps across 152 countries, states, or territories. We included 85.6% (154/180) of apps with a working Google Play page, most of which (132/154, 85.7%) had a privacy policy document. Most apps were developed by governments (116/154, 75.3%) and totaled 264.5 million installs. The average rating on Google Play was 3.5 (SD 0.7). Across the 154 apps, we identified 94 unique permissions, 18% (17/94) of which were dangerous, and 30 trackers. The average Permission Accumulated Risk Score was 22.7 (SD 17.7; range 4-74, median 16) and the average Privacy Transparency Index was 55.8 (SD 21.7; range 5-95, median 55). Overall, the privacy documents were difficult to read (median grade level 12, range 7-23); 67% (88/132) of these mentioned that the apps collected personal identifiers. The Permission Accumulated Risk Score was negatively associated with the average App Store ratings (r=−0.20; P=.03; 120/154, 77.9%) and Privacy Transparency Index (r=−0.25; P<.001; 132/154, 85.7%), suggesting that the higher the risk to one's data, the lower the apps' ratings and transparency index.

Conclusions: Many contact-tracing apps were developed covering most of the planet but with a relatively low number of installs. Privacy-preserving apps scored high in transparency and App Store ratings, suggesting that some users appreciate these apps.

Nevertheless, privacy policy documents were difficult to read for an average audience. Therefore, we recommend following privacy-preserving and transparency principles to improve contact-tracing uptake while making privacy documents more readable for a wider public.

(JMIR Mhealth Uhealth 2022;10(7):e35195) doi:10.2196/35195

KEYWORDS

COVID-19; mobile applications; contact tracing

Introduction

Strategies to Contain COVID-19

Since the beginning of the COVID-19 pandemic in early 2020, in the absence of vaccines or pharmacological treatments for the SARS-CoV-2 virus, some researchers have urged governments and the global public health community to speed up the response to contain SARS-CoV-2, pushing for the implementation of integrated nonpharmaceutical or nonpharmacological interventions (NPIs) [1]. Traditionally, NPIs adopted to curb an epidemic or pandemic such as influenza include mandating personal protective measures among health care professionals and citizens (eg, wearing masks), environmental measures such as isolating or quarantining positive cases [2], physical distancing, lockdowns, and travel restrictions [3]. However, more than two years into the pandemic, even after several vaccines were developed and rolled out worldwide, many countries have struggled to effectively and efficiently implement NPIs. In the absence of aggressive testing, contact tracing, and quarantining, an early study suggested that the only way to control COVID-19 would have included intermittent lockdowns until *herd immunity* was built up, resulting in unnecessary deaths [4]. Unfortunately, this seems to have been the case in many countries of the world, which observed alarming surges in cases of the SARS-CoV-2 virus. As of April 2021, the virus had infected >130 million individuals and claimed the lives of >2.8 million worldwide [5]. NPIs require significant investments in human resources and equipment and a level of coordination that might not be feasible in all contexts. For example, many low- and middle-income countries do not have the resources to enforce containment and testing policies [6] or purchase large amounts of vaccines. In addition, quarantining and physical distancing may not work among underprivileged and vulnerable segments of the population relying on daily wages for survival [7].

Contact tracing is one of the most cost-efficient NPIs available to break the chain of viral transmission [8]. According to the interim guidance of the World Health Organization, contact tracing consists of systematically identifying individuals exposed to confirmed positive cases, quarantining them, following up with them to ensure rapid isolation, and, finally, testing and treating them in case they develop symptoms [9]. This approach effectively controlled COVID-19 as long as quick and efficient processes were followed [10]. A way to guarantee such efficiency was to include digital technologies, particularly mobile phone–based tools, which are widely available worldwide [11]. In the last year, a few systematic literature reviews of COVID-19 apps mentioned contact-tracing apps as an essential type of app used in the context of the pandemic to curb virus transmission [12-17]. On the basis of the experience with the Ebola [18] and H1N1 [19,20] viruses, digital technologies have been increasingly used to support governments in carrying out manual contact-tracing activities. Several conceptual papers and overviews exist on mobile apps for COVID-19 contact tracing [21-27]. There are also a few systematic reviews on the topic [28-30], including a Cochrane review [30] and a literature review [29], focusing on digital contact tracing. The Cochrane review analyzed technologies used in epidemics and was updated in May 2020 to include new COVID-19–related studies. This review showed that such technologies are most effective when used to complement rather than substitute manual contact-tracing activities [30]. A literature review by Jalabneh et al [29] identified 17 apps that could be used for contact tracing and mentioned the use of these apps to help governments contain the pandemic.

Digital Contact Tracing

According to the Centers for Disease Control and Prevention (CDC), 2 main types of digital contact-tracing tools are used for case management and proximity-tracing or exposure notification apps [31]. Case management tools involve apps and devices that health workers involved in contact-tracing activities can use to capture data and manage contact databases of people tested for the virus. When a person tests positive, contact tracers interview them to recall where, when, and with whom they have been. The contacts are then triaged for assignment to case managers who call and notify contacts, providing options for testing, self-isolation, and referral to a health care provider if necessary. This activity can be done manually and on paper-and-pencil forms, so the technology allows for the streamlining of the process of data entry and management.

Proximity-tracing or exposure notification apps are designed for citizens who voluntarily download and activate such apps to assist in contact-tracing efforts. These apps rely on Bluetooth technology or location-based information stored on the phone to estimate the distance and duration of an encounter between users [12]. The phones exchange alphanumeric strings or keys via Bluetooth that contain such information. This information can be stored on the phone only (decentralized framework) or on a central server (centralized framework) and retained for a limited amount of time [12,32]. Depending on the type of framework, once a positive case is identified, the user or the central server flag their profile as positive, triggering the network and urging them to take action and get tested, self-isolate, or seek the help of health care professionals. This way, exposure notification apps can expand the reach of traditional manual contact tracers, who may fail to identify cases. The apps can reduce the burden on public health staff by

allowing for the electronic self-reporting by cases and contacts or by using location data or other features to identify community contacts unknown to the case to look at possible exposure to the virus. This study focuses on proximity-tracing or exposure notification apps as these are designed for citizens.

Use and Application of Contact-Tracing Apps

Many governments have developed digital contact-tracing apps following international guidelines (eg, the World Health Organization [33], CDC [31], or European CDC [34]) and Google and Apple exposure notification frameworks [35]. For example, as of May 28, 2020, when we started working on this project, we identified 36 apps by searching in the Google Play and Apple App Stores. On the same date, the page entitled *COVID-19 apps* on Wikipedia [36]—which was created on April 1, 2020—included information on 37 contact-tracing apps. As of June 15, 2020, there were already 68 apps and, by December 2, 2020, the page included approximately 100 apps. A recent review of COVID-19 apps in the Google Play and Apple App Stores identified 51 contact-tracing apps available until May 2, 2020 [13]. In the same period, the Technology Review of the Massachusetts Institute of Technology (MIT) launched the *Covid Tracing Tracker* project [37] with the purpose of monitoring and evaluating existing contact-tracing apps. A recent literature review published in July 2020 identified 17 apps in 15 countries [29], whereas Wen et al [38] analyzed 51 apps.

When can contact-tracing apps be considered effective? According to a seminal conceptual paper by Ferretti et al [25], contact-tracing apps can be deemed effective when at least 60% of the population uses them. More than a year after the COVID-19 pandemic was declared, some reviews asked the following question: Are these apps used by individuals [39]? Although many calls for evaluations of contact-tracing apps have been made [40], the evidence about contact-tracing app adoption and effectiveness is scant. A scoping review by Thorneloe et al [41] reported only a couple of examples of apps used by 10% to 20% of the population using data reported in news media outlets. Similar numbers were presented in an overview of contact-tracing apps [16] that provided descriptive information on 14 apps based on publicly available information. The authors focused on technical characteristics (eg, centralized or decentralized frameworks, tracing technology, and technical flaws) and the proportion of the country's population that used the apps, showing wide ranges (between 0.1% for *BlueZone* Vietnam and 60% for the Chinese *Health Code* used on Alipay and WeChat) [16]. In another conceptual paper, Seto et al [42] argued that the concept of privacy is context-specific and that there is a trade-off between privacy and public health value. To the best of our knowledge, the only comprehensive evaluation of contact-tracing apps is a longitudinal study involving the German *Corona-Warn-App* [43], one of the most downloaded contact-tracing apps in Europe totaling 26.5 million downloads as of March 25, 2021 [44].

How can this low global uptake be explained? The study on the German *Corona-Warn-App* by Munzert et al [43] reported a differential app uptake depending on the users' self-reported sociodemographic and behavioral profiles. For example, app use was positively associated with older age (≥50 years), education, socioeconomic status, health preconditions, and other preventive behaviors (eg, hand hygiene and mask wearing). App uptake was also higher among those who reported positive cases in their social network or who lived in areas of known outbreaks [43]; it was also higher among users who trusted the national government, the health care system, and science in general, and among those with a strong digital literacy who were less concerned about privacy [43].

Privacy and Transparency in Data Protection

Privacy, data protection, and the problem of trust in the government appear to be issues of concern, as reported in the aforementioned scoping review [41] and Cochrane review [30]. Numerous conceptual papers in the system design literature have discussed the issue of privacy [45-47], mainly focusing on the use of tracing techniques (eg, location-based vs Bluetooth [48]) and on the use of centralized versus decentralized frameworks, urging some researchers to develop their privacy-preserving apps and frameworks [49]. Decentralized models are privacy-preserving by design; however, they are generally inefficient in responding to the needs of public health systems as they rely on individual users' willingness to notify the network, which might never occur or might happen with delays that cannot be sustained when dealing with a highly transmissible virus such as SARS-CoV-2 [50]. Conversely, a recent simulation study showed that centralized models could be effective only when 80% of the population uses these technologies [51]. However, centralized models might discourage uptake among users who do not trust the organizations managing the centralized database. In their seminal paper, Ferretti et al [25] argued that app designers and governments supporting contact-tracing apps should be guided by ethical principles (eg, beneficence, reducing misery, equity, and social justice) and follow transparent practices to generate trust in citizens and promote app uptake. Transparency could be achieved, for example, by creating independent oversight advisory boards, publishing the code of the app and the algorithms used, integrating evaluation and research by third parties, and clearly communicating privacy and data protection principles. A way to express such principles is to use the apps' privacy policy documents, whose availability is requested by the main app stores and recommended by numerous institutions, including the Privacy Trust Framework [52]; the US Federal Trade Commission [53]; and the General Data Protection Regulation (GDPR) of the European Union, which entered into force as of May 2018 [54]. The general recommendation for developers is to produce privacy policy documents that are clear and easy to understand. A way to ensure clarity and comprehensibility of documents is to provide a low readability level, which has been previously considered an element for evaluating apps' privacy policies [55]; for instance, the Privacy Trust Framework recommends a reading grade level of ≤12 and a Flesch reading ease of 45 [52]. A recent paper investigating contact-tracing apps [56] reported that transparency in the documentation was perceived as an essential element of trust in the apps and developers.

Beyond the conceptual and normative debates among scholars, are citizens' concerns about privacy real? Are contact-tracing

apps truly invasive of privacy? Are the developers or governments behind the apps able to provide transparent and clear information about data protection and treatment? The answers to these questions do not appear in the existing literature on contact-tracing apps. There are a few systematic reviews of COVID-19 apps that mention contact-tracing apps as a type used in the context of the pandemic [12-17]. There are also a number of reviews of COVID-19 contact-tracing apps. For example, the aforementioned MIT *Covid Tracing Tracker* project [37] provides some descriptive information on the technological infrastructure and uptake of these apps. The other 2 overviews of COVID-19 contact-tracing apps [12,16] describe general vulnerabilities instead of considering privacy concerns using the information included in the privacy policy documents. These reviews do not provide a comprehensive, specific analysis of the permissions and data protection [16]. A more recent review of COVID-19 contact-tracing apps [38], published as a conference proceeding, focused on the user privacy aspects, potential data leakage, and other technical features of a sample of 41 apps. The authors mentioned the role of transparency to ensure uptake but did not investigate app characteristics that could enhance transparency beyond publishing the source code, an element present in a few open-source apps analyzed. Another content analysis of contact-tracing apps [57] looked at the public perception of these apps through user reviews and at the number of downloads, tackling the issue of privacy-by-design. Another review analyzed which permissions are needed to allow tracking and tracing and whether the apps have embodied principles of privacy and data protection by design [58]. Another review focused on apps developed in the United States and on usability and qualitative features [59]. Finally, another review looked at the readability of contact-tracing apps [60] without looking at privacy aspects. In conclusion, none of these reviews of contact-tracing apps includes a combined analysis of privacy and data protection principles.

Furthermore, in April 2020, our research group embarked on a project that a few months later resulted in the creation of a nationwide contact-tracing app (Ma3an) [61] in collaboration with the local Ministry of Public Health. Parallel to this project, we searched app databases to identify benchmark apps and used them as a reference for privacy-preserving contact-tracing apps. This was one of the main drivers urging us to undertake a comprehensive systematic review of contact-tracing apps and focus on data protection and privacy aspects.

This study aimed to identify, map, and evaluate all available COVID-19 contact-tracing apps developed worldwide in a systematic way. The specific objectives of this study were to (1) identify and map existing contact-tracing apps; (2) evaluate the type of data collected to define the risks to users' privacy based on the permissions required; and (3) evaluate the readability and content of privacy policy documents to establish whether these documents transparently communicate details about privacy, data protection, management, and retention. Finally, after more than a year of implementation of the search protocols, data extraction, and assessment, we decided that it was time to respond to the recent call for COVID-19 contact-tracing app evaluations launched by Colizza et al [40] on Nature Medicine, February 15, 2021.

Methods

We conducted a systematic review of information about existing COVID-19 contact-tracing apps following a rigorous process of app identification, selection, data extraction, and analysis as used in a previously published app review by MB [62] and similar studies targeting different kinds of apps [63,64]. In addition, to address the research objectives, we performed a content analysis of contact-tracing apps' publicly available Google Play pages and associated privacy policy documents.

Searches and Sources of Information

We used two main strategies to identify contact-tracing apps: (1) searching for keywords in the Google Play and on the Apple App Store using Google and (2) scanning the list of apps included in 5 websites identified via Google search.

For the first search strategy, we applied the following two search queries: (1) "allintext:COVID-19|covid|covid19|coronavirus AND tracing|exposure site:play.google.com" and (2) "allintext:COVID-19|covid|covid19|coronavirus AND tracing|exposure site:play.google.com." We conducted the initial searches on May 7, 2020, and updated them almost monthly, on June 1, 2, and 24, 2020; August 18, 2020; November 27, 2020; April 8, 2021; August 7, 2021; and October 31, 2021.

The second search strategy consisted of scanning 5 webpages containing lists of COVID-19 apps, such as the Wikipedia page on COVID-19 apps (first published on April 1, 2020, and last edited on October 20, 2021) [36]; the MIT *Covid Tracing Tracker* project (first published on May 7, 2020, and last updated on January 25, 2021) [37]; the database of contact-tracing apps of the Council of Europe (last updated on June 10, 2020, and then discontinued) [65]; an article on *COVID tracing app roundup* on Android Police (published on September 1, 2020, and last updated on November 21, 2020) [66], with 26 US states using Google Exposure Notification System (ENS), 37 international apps using the same ENS system, and 30 apps not using the ENS framework; and the *List of countries using Google and Apple's COVID-19 Contact Tracing API* on the XDA Developers website (published on June 24, 2020, and updated on February 25, 2021) [67]. All of these sources were last checked on October 31, 2021.

Inclusion Criteria

To be included, the apps had to (1) be explicitly aimed at COVID-19 *contact tracing* or *exposure notification*, (2) have a publicly available page on the Google Play or Apple App Stores, and (3) have information on permissions and a privacy policy document available from Google Play. Therefore, we excluded apps designed for contact tracing not explicitly made for COVID-19 that provided general information on COVID-19 or that were *symptom checkers* without mentioning contact-tracing features. We also excluded apps that had an available page only on the Apple App Store as the pages do not include information on permissions as in the Google Play. We also excluded apps if their privacy policy documents were not available (eg, through a broken link) or that did not include a privacy policy explicitly related to the app.

App Selection Process

We followed a multistage selection process. MB exported the Google search results by looking at the Apple App Store and Google Play in Microsoft Excel. MB then screened the links for relevance, and MAD confirmed the selection. Next, we resolved all disagreements through discussions. Finally, we entered the Google Play links in the Exodus Privacy database [68], which is the auditing platform for Android apps. The Exodus platform looks for embedded trackers (a software meant to collect user data) and permissions requested by each app. An app was excluded if the link to the Exodus database was not working.

Data Extraction

App Characteristics

We extracted the following information from Google Play pages: number of installs, a link to the privacy policy document, 5-star reviews, number of reviews, version of the app, version of the operating system, sponsor, and permission designations. From the Apple App Store page (if available), we extracted the following information: 5-star ratings and number of ratings, app version, seller, operating system version, and language. MB extracted the information, and MAD double-checked it. Any discrepancies were flagged and resolved through discussion.

Permission Data

In total, 4 authors (MB, MAD, NM, and RAS), in pairs and independently, extracted the information on permissions using a standardized web-based extraction form based on Exodus reports [68]. All permission items were entered as binary values (1=yes; 0=no). Overall, the raters achieved excellent interrater agreement (percentage of agreement=98.3%; Cohen κ=0.954; Krippendorff α=0.953). All disagreements were resolved through discussion. The Exodus reports [68] label permissions according to 2 levels of risk, as described on the Android developers' page [69]: *Normal* or *Dangerous*, including *Signature* and *SignatureOrSystem*. As described in the book *Android Application Security Essentials* [70], normal permissions are those that "cannot do much harm to the user. They generally do not cost users money, but they might cause users some annoyance...These permissions are automatically granted to the app." Dangerous permissions are always shown to the user as "they can cause user privacy or financial loss." Signature permissions allow 2 apps authored by the same developer to access each other's components. This type of permission is automatically granted to the app if it has the same certificate as the app that declared the permission. Signature or system permissions are "granted to applications with the same certificate as the app that defined the permission. In addition, this protection level includes an app with the same certificate as the Android system image. This permission level is mainly used for applications built by handset manufacturers, carriers, and system apps. These permissions are not allowed for third-party apps. These permissions let apps perform some very powerful functions" [70].

Privacy Policy Data Extraction

For apps with available privacy policy information, if the document was in a language other than English, it was translated using Google Translate and saved in PDF format with a timestamp. Similar to the procedure for extracting permission information, 4 authors (MB, NM, MAD, and RAS) independently completed a privacy policy assessment using a standardized web-based checklist. The checklist was adapted from a similar study focusing on data security and privacy in mobile apps addressing depression [71]. The inventory contained a total of 13 specific items (Table 1), which we grouped into 3 main categories: 4 items were in the *privacy* category; 6 items were in the *data management* category; and 3 items were in the *legal framework* category for data protection (eg, the GDPR for European countries or any other framework) explicitly mentioned the right to delete or edit the data, which should be clarified in the legislative framework. We rated each item on a nominal scale (yes=1, no=0, or not applicable, depending on the item).

For this data extraction task, we conducted first a calibration exercise with a sample of 15 randomly selected apps to ensure sufficient reliability and adjust the instrument before applying it to the remaining set of apps. The aforementioned 4 authors individually and independently completed the same checklist. The exercise yielded a sufficient level of agreement (84.8%) as well as reliability indexes (Cohen κ=0.823; Krippendorff α=0.696). Disagreements were resolved through discussion, which allowed for the clarification of a few interpretation issues. After we resolved the disagreements, the 4 raters independently completed the checklist for other apps. The interrater reliability notably improved (percentage of agreement=87.5%; Cohen κ=0.749; Krippendorff α=0.749). Finally, the reviewers completed the data extraction for the remaining apps in pairs. As in the previous task, we resolved disagreements through discussion until we reached a consensus.

Table 1. Privacy policy checklist and rubric used to calculate the Privacy Transparency Index (0-100).

Domain and items	Score
Privacy (25 points)	
Does the app collect personally identifiable information?	Yes=0; partial[a]=5; no=10
Does the privacy policy mention that the app can be used without entering identifiable information?	Yes[b]=5; no=0
Does the privacy policy mention that the app collects identifiable information such as full name, email, and phone number?	Yes or (N/A[c])[b]=5; no=0
Does the privacy policy mention that the app provides the option of a personal identification number, password, or log-in process to view and enter user data?	Yes or N/A[b]=5; no=0
Data management (50 points)	
Does the privacy policy explicitly state which type of data are processed?	Yes=15; no=0
Does the privacy policy contain a section on "how the app works" explicitly?	Yes=5; no=0
Does the privacy policy state that the app or server encrypts the entered data?	Yes=10; no=0
Does the privacy policy describe the process of data exchange and communication between server and phone related to user-entered information?	Yes=5; no=0
Does the privacy policy state that the user information is stored on the phone or device?	Yes=10; no=0
Does the privacy policy mention data retention?	Yes=5; no=0
Legal framework (25 points)	
Does the privacy policy mention the GDPR[d]? If not, does the privacy policy mention other legislative frameworks?	Yes=15; no=0
Does the privacy policy state whether users can delete entered information?	Yes=5; no=0
Does the privacy policy state whether users can edit entered information?	Yes=5; no=0

[a]In this context, partial information is related to the use of location services only.

[b]Not applicable options for apps that do not collect personal or identifying information.

[c]N/A: not applicable.

[d]GDPR: General Data Protection Regulation.

Data Elaboration

App Characteristics

On the basis of the information reported on the Google Play page or on the other sources we used, we categorized the apps by country and continent according to the NationsOnline classification [72]. We also categorized the apps by the type of coverage (country or state, county, or city, depending on their geographical coverage) and type of sponsor (government; nonprofit organization; profit organization; and multistakeholder, involving a combination of the previous categories). Finally, we grouped the apps with the associated Google Play information according to the number of installs (ranging from ≥50 to ≥100 million) as a relative measure of popularity.

Permission Data

We counted the number of trackers and permissions identified through the Exodus platform [68]. Next, we assigned numeric values to each protection level: normal permissions=1; dangerous, signature, or system permissions=2; and trackers=3 as they constitute a higher level of danger to users' privacy. We then multiplied the number of permissions and trackers by the protection level to calculate a *Permission Accumulated Risk Score*. The higher this score, the higher the risk.

Privacy Policy Data

We assigned different points to the aforementioned checklist (Table 1) to calculate a *Privacy Transparency Index*, which could range from 0 to 100. Similar to the Permission Accumulated Risk Score, the higher the Privacy Transparency Index, the more transparent the privacy policy.

We also assessed the readability needed to understand the policy using a combined estimate of readability indexes provided by the Automatic Readability Checker, a web-based readability calculator [73]. This tool outputs an estimate based on 7 popular readability formulas: the Flesch Reading Ease Formula, Flesch-Kincaid Grade Level, Gunning fog index, Simple Measure of Gobbledygook Index, Coleman-Liau Index, Automated Readability Index, and Linsear Write Formula. Naturally, the lower the grade, the easier it is to understand. For example, the Privacy Trust Framework recommends a reading grade level of ≤12 for policy documents [52].

Analyses

We used descriptive statistics to summarize the apps' characteristics. For apps with permission and privacy policy data, we summarized continuous variables (eg, number of ratings, Permission Accumulated Risk Score, and Privacy Transparency Index) using mean and SD or median and IQR for count variables where appropriate. In addition, we investigated potential associations between app characteristics

(such as type of sponsors, number of installs, and number of reviews or ratings), Permission Accumulated Risk Score, and Privacy Transparency Index using ANOVAs, chi-square tests, and Pearson correlation tests (significance level was assumed at $P<.05$).

Results

Search Results

The selection process is illustrated in Figure 1. We applied the search queries to Google on May 7, 2020; June 1, 2, and 24, 2020; August 18, 2020; November 27, 2020; April 8, 2021; August 7, 2021; and October 31, 2021. We exported 1055 records from Google Play and 1027 records from the Apple App Store in Microsoft Excel. After removing duplicate links, we screened 15.64% (165/1055) of unique app links available from Google Play and 16.85% (173/1027) of links available from the Apple App Store. In this first screening stage, we excluded 11.5% (19/165) and 25.4% (44/173) of apps from Google Play and the Apple App Store, respectively, that were deemed irrelevant as they were not related to COVID-19. The remaining 88.5% (146/165) and 74.6% (129/173) of apps from Google Play and the Apple App Store, respectively, were assessed for eligibility together with 152 apps from the Wikipedia page [36], 81 from the MIT *Covid Tracing Tracker* project [37], 52 from the Council of Europe database of contact-tracing apps [65], 93 from the Android Police page [66], and 65 from the XDA Developers page [67]. Finally, we excluded apps that were not designed for contact tracing (52/146, 35.6% from Google Play; 42/129, 32.6% from the App Store; and 59/152, 38.8% from Wikipedia).

The final list included 180 unique COVID-19 contact-tracing apps that were potentially eligible for review. Of these, 85.6% (154/180) had a Google Play link to generate an Exodus platform permission report [68], and 76.1% (137/180) had an associated privacy policy document. A total of 14.4% (26/180) of the apps did not have a permission report either because a Google Play link was not available (19/26, 73%) or because it was no longer available at the time of the analysis (7/26, 27%). The other 4.5% (7/154) of apps did not have a privacy policy document available, and 11% (17/154) had privacy policy documents that were not app-specific. Of the 180 selected apps, 132 (73.3%) contained data related to both permissions and privacy policies. A complete list of all 180 identified apps up to October 31, 2021, is included in Multimedia Appendix 1. The list contains links to the Google Play pages and to the Exodus platform reports. The list is also publicly available on Tableau Public from the link [74].

The identified 180 apps covered 152 geographical units (countries or regions, states or provinces, counties, territories, or cities) in 90 different countries spanning all 5 continents. Most apps came from the Americas (53/180, 29.4%), Asia (53/180, 29.4%), and Europe (46/180, 25.6%). The African continent had 6.1% (11/180) of the apps, Oceania had 5% (9/180), and 4.4% (8/180) of the apps covered multiple continents or were developed to cover different countries. The world map in Figure 2 represents the global distribution of the COVID-19 contact-tracing apps. The larger the bubble, the higher the number of apps for each country.

The United States had the highest absolute number of apps as 20% (36/180) were developed to cover different states. This number does not include 1.1% (2/180) of the apps, which came from the US unincorporated territories of Guam (*Guam Covid Alert*) and Puerto Rico (*Rastrea el Virus*). A total of 1.1% (2/180) of the apps (*Care19 Alert* and *Care19 Diary*) covered the states of North Dakota, South Dakota, and Wyoming. The country with the second-highest number of apps was Australia (6/180, 3.3%). Germany, Great Britain, India, and the Philippines had 2.8% (5/180) of the apps each; Brazil and Italy had 2.2% (4/180) of the apps each; France, Malaysia, Mexico, Nepal, Russia, Singapore, South Africa, and Spain had 1.7% (3/180) of the apps each; and Canada, Iran, the Netherlands, Oman, Switzerland, and the United Arab Emirates had 1.1% (2/180) of the apps each.

Most contact-tracing apps were sponsored by governments (132/180, 73.3%), followed by private organizations (28/180, 15.6%) and nonprofit organizations (14/180, 7.8%). A small number of apps involved multiple stakeholders, including consortia of private, nonprofit, and governmental organizations (6/180, 3.3%).

Figure 1. App selection process. COE: Council of Europe; MIT: Massachusetts Institute of Technology.

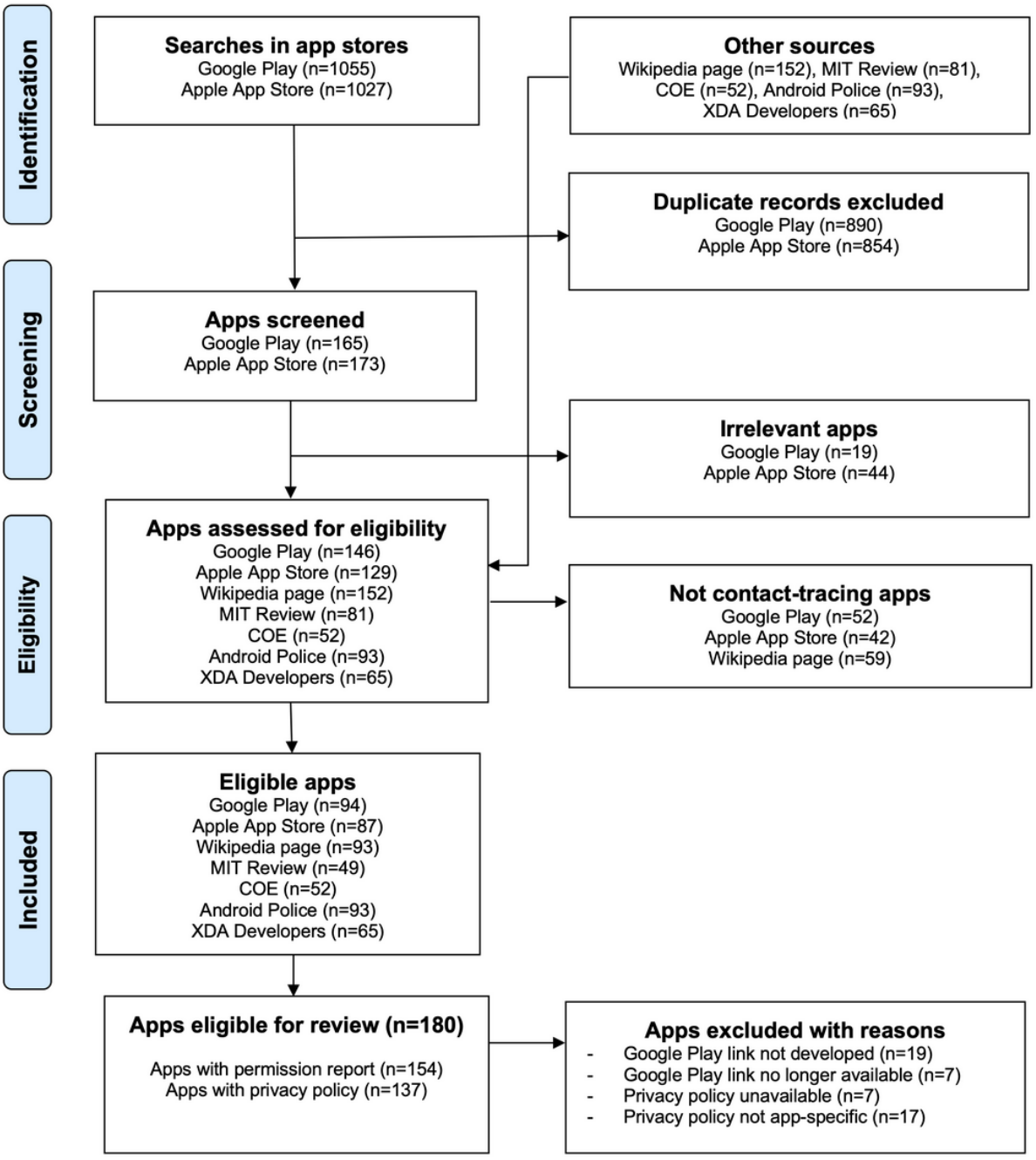
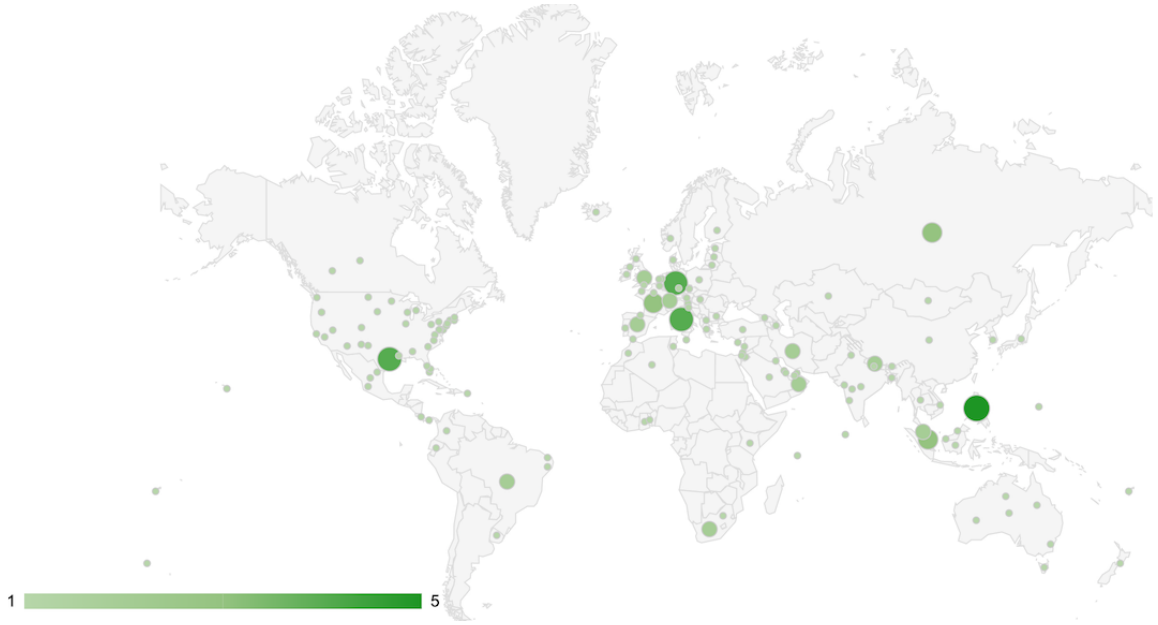


Figure 2. Geographic distribution of the identified contact-tracing apps, aggregated by country.



App Characteristics

The list of 154 apps with permission data is provided in Multimedia Appendix 2 (tab 1). In Table 2, we report the basic descriptive information for the sample of apps grouped according to the number of installs. As of October 8, 2021, based on Google Play install categories, the 154 apps totaled 264.5 million installs (1.7 million on average), ranging from 10 to 100 million. The most installed app was *AarogyaSetu*, developed by the Indian National Informatics Centre eGov Mobile Apps department. The least installed app was *Aggie-COVID-19*, which was designed for New Mexico University. Most apps were installed between 100,000 and 1 million times (106/154, 68.8%), with approximately one-third being installed <100,000 times (48/154, 31.2%).

Most apps were developed by governments (116/154, 75.3%), private organizations (17/154, 11%), nonprofit organizations (11/154, 7.1%), and multistakeholder consortia (10/154, 6.5%). No significant association between the type of sponsor and number of install categories was detected. The average rating was 3.5 (SD 0.7) on Google Play and 3.6 (SD 0.9) on the Apple App Store based on a subsample of 120 apps with a valid App Store page. The average number of reviews was 26,412 (SD 143,803) on Google Play and 5120 (SD 28,826) on the App Store.

The average number of reviews on Google Play was significantly associated with the number of installs (F_3=5.04; P<.001; η^2=0.07), with apps installed ≥1 million times receiving more reviews than those installed between 100,000 and 500,000 times and more reviews than those installed <100,000 times. We detected a similar difference in the number of ratings on the Apple App Store (F_3=3.59; P=.02; η^2=0.05); in this case, the average number of ratings was significantly higher in apps installed <100,000 times and those installed between 500,000 and 1 million times.

Table 2. Characteristics of the sample of apps organized by the number of installs (N=154).

Variable	Total	Number of installs				*P* value[a]
		≥1 million (n=49)	500,000 to 1 million (n=12)	100,000 to 500,000 (n=45)	<100,000 (n=48)	
Type of sponsor, n (%)						.05
Government	116 (75.3)	42 (85.7)	10 (83.3)	36 (80)	28 (58.3)	
Private	17 (11)	2 (4.1)	1 (8.3)	3 (6.7)	11 (22.9)	
Nonprofit	11 (7.1)	4 (8.2)	1 (8.3)	3 (6.7)	3 (6.3)	
Multistakeholder	10 (6.5)	1 (2)	0 (0)	3 (6.7)	6 (12.5)	
Average app ratings, mean (SD; range)						
Google Play	3.49 (0.69; 1.00-5.00)	3.53 (0.72; 1.00-4.70)	3.16 (0.65; 1.80-4.10)	3.41 (0.62; 1.30-4.40)	3.61 (0.71; 1.70-5.00)	.17 (.13)
Apple App Store[b]	3.59 (0.94; 1.00-5.00)	3.46 (0.97; 1.00-4.90)	3.42 (1.15; 1.20-5.00)	3.69 (0.91; 1.60-5.00)	3.74 (0.86; 1.50-5.00)	.52 (.64)
Average number of ratings, median (IQR; range)						
Google Play	972 (5094; 1-1,600,000)	16,373 (32,382; 960-1,600,000)	2362 (1200; 1553-4094)	750 (569; 155-3275)	114 (137; 1-1135)	<.001 (<.001)
Apple App Store[b]	103 (591; 1-287,200)	1047 (3553; 1-287,200)	464 (763; 4-1400)	112 (226; 1-2200)	25 (47; 1-595)	.02 (<.001)
Permission data, median (IQR; range)						
Average number of permissions	9 (10; 2-44)	10 (9; 6-40)	7 (6; 6-44)	8 (7; 2-42)	12 (12; 4-41)	.27 (.19)
Average percentage of dangerous permissions	13 (21; 0-63)	13 (21; 0-36)	0 (20; 0-63)	11 (19; 0-50)	15 (21; 0-44)	.87 (.50)
Average number of trackers	1 (2; 0-11)	2 (2; 0-7)	0 (1.3; 0-4)	0 (2; 0-5)	1 (1.25; 0-11)	.38 (.23)
Permission Accumulated Risk Score	16 (26; 4-74)	14 (26; 6-63)	10 (13; 6-74)	14 (22; 4-70)	23.5 (26; 4-65)	.34 (.11)
Privacy policy data, n (%)						.76
Privacy policy available	132 (85.7)	44 (89.8)	10 (83.3)	37 (82.2)	41 (85.4)	
Privacy policy unavailable	22 (14.3)	5 (10.2)	2 (16.7)	8 (17.8)	7 (14.6)	
Readability[b]						
Grade level, median (IQR; range)	12 (3; 7-23)	12 (4; 8-23)	11 (2; 7-16)	12 (3; 7-19)	12 (2; 8-18)	.07 (.13)
Readability level, n (%)[b]						.50
Very difficult to read	14 (9.1)	7 (15.9)	1 (10)	4 (10.8)	2 (4.9)	
Difficult to read	67 (43.5)	24 (54.6)	4 (40)	18 (48.7)	21 (51.2)	
Fairly difficult to read	47 (30.5)	11 (25)	4 (40)	14 (37.8)	18 (43.9)	
Standard or average	4 (2.6)	2 (4.6)	1 (10)	1 (2.7)	0 (0)	
Policy—transparency index, median (IQR; range)[b]	55 (30; 5-95)	60 (31.3; 25-95)	60 (36.3; 20-90)	60 (35; 5-85)	50 (20; 5-90)	.65 (.68)

[a]*P* value for independent sample *t* tests (2-tailed), chi-square tests, or *F* tests comparing the number of install categories and the other variables. The *P* value for the Kruskal-Wallis test, the nonparametric equivalent of an ANOVA, is indicated in parentheses.

[b]The calculations are available from a total of 132 apps with privacy policy documents.

Permission Data

The typology of permissions, identified through the Exodus platform automatic permission extraction, is presented in Multimedia Appendix 3 (tab 2). Across the 154 apps with valid permission data, there were 94 different types of permissions, of which 17 (18%) were dangerous or special.

Among the normal permissions, the one used in all apps was *Internet, have full network access* (154/154, 100%). The permissions used by more than half of the apps were *view network connections* (150/154, 97.4%); *wake lock, prevent*

phone from sleeping (142/154, 92.2%); *run in foreground* (137/154, 89%); *run at startup* (131/154, 85.1%); and the permissions related to Bluetooth as *pair with Bluetooth devices* (118/154, 76.6%). The most frequently used dangerous permission was *access precise location (GPS and network-based)*, which was used by approximately half of the apps (73/154, 47.4%). Other dangerous permissions used by approximately one-third of the sample included *access approximate location (network-based)* (57/154, 37%), *take pictures and videos* (51/154, 33.1%), and *modify or delete the contents of your SD card* (44/154, 28.6%). On average, each app collected 9 permissions (IQR 10, range 2-44). Only 0.6% (1/154) of the apps collected 2 permissions (*TRACE Taguig*, the Philippines), and only 0.6% (1/154) collected 44 permissions (*Shlonik*, Kuwait); 46.1% (71/154) of the apps required fewer permissions.

The average proportion of dangerous permissions was 13% (IQR 21%, range 0%-63%). A total of 39% (60/154) of the apps did not use any dangerous permissions, and 0.6% (1/154) reported using the most dangerous permissions (*Corona Watch*, Karnataka province, India).

In addition, the Exodus platform extracted approximately 30 different trackers (Multimedia Appendix 3, tab 3). Google Firebase Analytics was the most frequently used tracker (80/154, 51.9%), followed by Google CrashLytics for crash reporting (48/154, 31.2%). Although some apps had analytics and app statistic information trackers, others had trackers used to profile users (eg, Facebook log-in, Segment, AltBeacon, and DOV-E) or for advertising (Google AdMob; 6/154, 3.9%). On average, each app used 1 tracker (IQR 2, range 0-11). Although 41.6% (64/154) of the apps did not use any trackers, 0.6% (1/154) used the most trackers (*Citizen SafePass*).

On the basis of the number and type of permissions and trackers, the average Permission Accumulated Risk Score was 16 (IQR 26, range 4-74). Of the 154 apps, 2 (1.3%) scored the lowest—*TRACE Taguig* (the Philippines) and *Beat COVID Gibraltar*—and 1 (0.6%) scored the highest—*Shlonik* (Kuwait). Approximately one-fifth of the sample (40/154, 26%) obtained the second- and third-lowest Permission Accumulated Risk Score (score of 6: 23/154, 14.9%; score of 7: 17/154, 11%).

Privacy Policy Data

Privacy policy data extraction was available for 85.7% (132/154) of the apps, as 14.3% (22/154) did not have a working privacy policy link or document. A spreadsheet containing the privacy policy data extraction for each app is available in Multimedia Appendix 4.

Regarding readability, the privacy documents required a median grade level of 12 (IQR 3, range 7-23). We found the lowest level in the privacy policy documents of *Stopp Corona* (Austria) and *The Territory Check-In* (Northern Territory, Australia) and the highest level in the policy document of the *Taiwan Social Distancing* app.

Most of the privacy policy documents were *difficult* or *very difficult to read* (81/132, 61.4%), with approximately one-third being *fairly difficult to read* (47/132, 35.6%). Only 3% (4/132) of the apps had a *standard or average* reading level. In addition to *Stopp Corona* and *The Territory Check-In*, the other 2 apps were *COVID Alert* (South Africa) and *COVID Alert* (Canada).

The sample distribution according to the privacy policy checklist items is shown in Table 3. Notable strengths in terms of privacy included the fact that most policy documents explicitly stated when personal identifiers were collected (116/132, 87.9%) and what type of data was collected and for how long (100/132, 75.8%). In addition, privacy policy documents mentioned that these data were protected through a personal identification number or password (78/132, 59.1%). Nevertheless, most apps collected or partially collected personally identifiable information (89/132, 67.4%). Other limitations of data management included the fact that most privacy policies did not have a section clearly explaining how the app worked (86/132, 65.2%), did not state or explain how the app or server encrypted the data, or did not describe the process of data exchange (105/132, 79.5%).

In terms of the legal framework used, most policy documents mentioned that they abided by the GDPR or other national-level legislative data protection frameworks (82/132, 62.1%). Another notable strength of the right to be forgotten is that most of the policy documents stated that the users had the right to delete the app or their profile (90/132, 68.2%). Nevertheless, a few policies mentioned the right to rectify or edit the profile (50/132, 37.9%).

On the basis of the privacy policy checklist, the average Privacy Transparency Index was 56 (SD 22, range 5-95), which can be considered moderate as it is slightly above the median value of 50. Of the 132 apps, 4 (3%) scored the lowest Privacy Transparency Index—*The Territory Check-In* (Australia); *Bardghat Municipality - COVID-19/Disaster Response* and *Bharatpur Metropolitan|COVID-19 Response System* (both from Nepal); and *Check On the other hand, oneTAS* (Australia)—1 (0.8%) scored the highest Privacy Transparency Index—*COVID Tracker Ireland*—and 5 (3.8%) scored the second-highest Privacy Transparency Index (90/100)—*Corona-Warn-App* (Germany), *NHS COVID-19 App* (the United Kingdom; 2 versions, one pilot and one national), *SwissCovid* (Switzerland), and *Protect Scotland*.

Table 3. Completed checklist of the Privacy Transparency Index applied to 132 apps.

Domain, item, and score	Apps, n (%)
Privacy	
Does the app collect personally identifiable information?	
Yes=0	79 (59.8)
Partial[a]=5	10 (7.6)
No=10	43 (32.6)
Does the privacy policy mention that the app can be used without entering identifiable information?	
Yes or N/A[b]=5	48 (36.4)
No=0	84 (63.6)
Does the privacy policy mention that the app collects identifiable information such as full name, email, and phone number?	
Yes or N/A=5	116 (87.9)
No=0	16 (12.1)
Does the privacy policy mention that the app provides the option of a personal identification number, password, or log-in process to view and enter user data?	
Yes or N/A=5	78 (59.1)
No=0	54 (40.9)
Data management	
Does the privacy policy explicitly state which type of data are processed?	
Yes=15	100 (75.8)
No=0	32 (24.2)
Does the privacy policy contain a section on "how the app works" explicitly?	
Yes=5	46 (34.8)
No=0	86 (65.2)
Does the privacy policy state that the app or server encrypts the entered data?	
Yes=10	57 (43.2)
No=0	75 (56.8)
Does the privacy policy describe the process of data exchange and communication between server and phone related to user-entered information?	
Yes=5	27 (20.5)
No=0	105 (79.5)
Does the privacy policy state that the user information is stored on the phone or device?	
Yes=10	51 (38.6)
No=0	81 (61.4)
Does the privacy policy mention data retention?	
Yes=5	100 (75.8)
No=0	32 (24.2)
Legal framework	
Does the privacy policy mention the GDPR[c]? If not, does the privacy policy mention other legislative frameworks?	
Yes=15	82 (62.1)
No=0	50 (37.9)
Does the privacy policy state whether users can delete entered information?	
Yes=5	90 (68.2)
No=0	42 (31.8)

Domain, item, and score	Apps, n (%)
Does the privacy policy state whether users can edit entered information?	
Yes=5	50 (37.9)
No=0	82 (62.1)

[a]Partial score when the app used location services only.

[b]N/A: not applicable.

[c]GDPR: General Data Protection Regulation.

Correlations

The correlations among continuous variables representing app characteristics, Permission Accumulated Risk Score, readability, and Privacy Transparency Index are shown in Table 4. There was a small significant correlation between the average app ratings in the 2 app stores (r=0.21; P=.02; 116/154, 75.3%). Similarly, there was a larger, highly significant correlation between the number of ratings reported in the Google Play and Apple App Stores (r=0.87; P<.001; 116/154, 75.3%), which, in turn, was significantly correlated with the number of installs (Google Play ratings: r=0.96; P<.001; 150/154, 97.4%; Apple App Store ratings: r=0.90; P<.001; 120/154, 77.9%). This finding is consistent with the ANOVA reported at the end of the *App Characteristics* section. The Permission Accumulated Risk Score had a small negative correlation with the average rating on the Apple App Store (r=−0.20; P=.03; 120/154, 77.9%), suggesting that, the lower the rating, the higher the risk to the users' privacy. The Privacy Transparency Index was negatively associated with the Permission Accumulated Risk Score (r=−0.25; P<.001; 132/154, 85.7%), suggesting that, the higher the risk to one's data, the lower the transparency index of the related policy document.

Figure 3 is a screenshot of a map representing the relationship between the Permission Accumulated Risk Score and Privacy Transparency Index. The map is publicly available on Tableau Public [74].

Table 4. Correlation table for continuous variables.

Variables	1	2	3	4	5	6	7
1. Average rating (Google Play)	—[a]	—	—	—	—	—	—
2. Average rating (Apple App Store)	0.21[b]	—	—	—	—	—	—
3. Number of ratings (Google Play)	0.04	0.03	—	—	—	—	—
4. Number of ratings (Apple App Store)	0.11	0.13	0.87[c]	—	—	—	—
5. Number of installs	0.02	0.04	0.96[c]	0.90[c]	—	—	—
6. PARS[d]	0.13	−0.20[b]	0.05	0.04	0.01	—	—
7. Grade level (readability)	0.04	0.10	0.05	0.11	0.08	0.02	—
8. PTI[e]	−0.03	0.10	0.00	<0.00	0.02	−0.25[f]	−0.15

[a]Not applicable.

[b]P<.05.

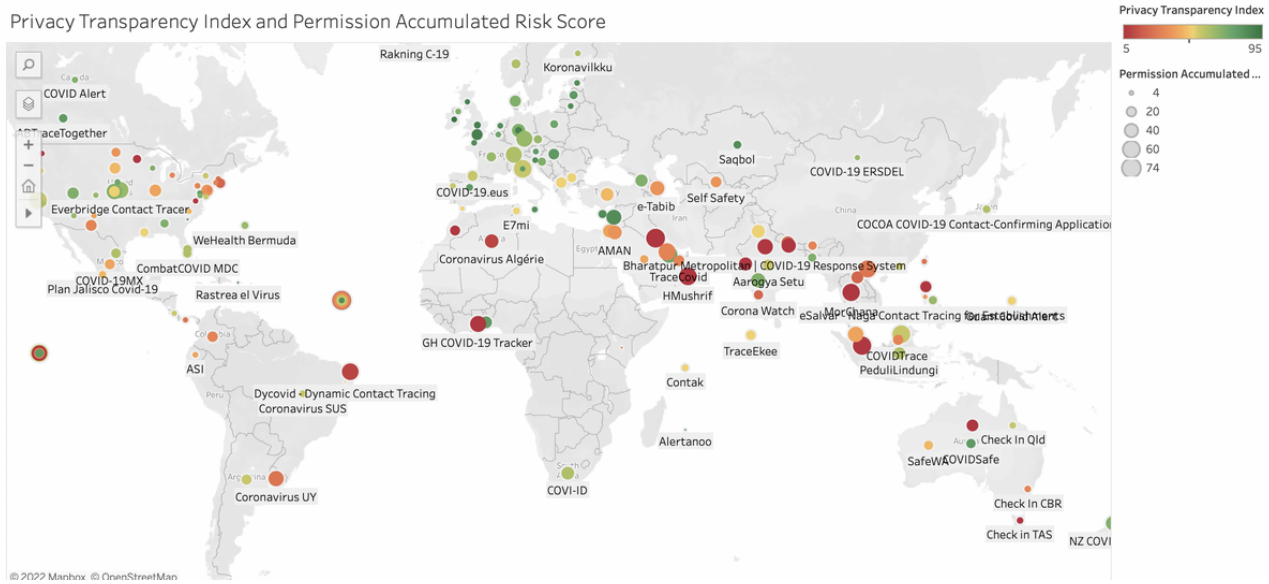
[c]P<.001.

[d]PARS: Permission Accumulated Risk Score.

[e]PTI: Privacy Transparency Index.

[f]P<.01.

Figure 3. Map representing the Permission Accumulated Risk Score (size of the bubble) and the Privacy Transparency Index (color).



Discussion

Principal Findings

This paper presents the first systematic review of COVID-19 contact-tracing apps developed worldwide that explored the apps' approach to data protection and privacy. In addition, we assessed the number and type of permissions requested by the apps and how transparent the privacy policy documents were about data protection rules. This systematic review aimed to (1) identify and map existing contact-tracing apps, (2) evaluate the potential risks to users' privacy through the assessment of the type and quantity of permissions requested, and (3) evaluate the readability and level of transparency of related privacy policy documents.

We adopted a systematic search, selection, and identification process using different sources [62]. This systematic approach allowed us to identify 180 COVID-19 contact-tracing apps covering 90 countries. Of these 180 apps, 154 (85.6%) had valid links to extract permission data, and 132 (73.3%) had privacy policy documents that could be analyzed. Our search strategy allowed us to generate a much larger sample than those reported in recent COVID-19–related app audits [13,29,38]. Furthermore, the selection of apps we analyzed is more extensive than the one included in the MIT *Covid Tracing Tracker* project [37], which currently comprises 81 apps. The most updated source of information to date is Wikipedia's *COVID-19 apps* page [36], which lists 152 apps. Although the number of apps for COVID-19 contact tracing might grow over time with more governments embarking on digital contact-tracing efforts, some researchers believe that the momentum is now over considering how the pandemic has evolved. In the absence of zero–COVID-19 strategies, mitigation strategies and vaccination campaigns might take priority over contact tracing and other NPIs [75]. Nevertheless, we hope that this review will spark the interest of the public health and global health community, who might want to contribute to the enlargement and maintenance of the app database, which is already accessible on Tableau Public [74].

Permission Data and Privacy Risk

To achieve the second objective, we analyzed publicly available information from Google Play webpages and extracted it using the Exodus platform scanner [68]. This objective assessment and data extraction allowed us to systematically identify and classify the types of permissions and their relative risk to users' privacy. We developed a Permission Accumulated Risk Score to qualify the level of risk, accounting for some dangerous permissions and the presence of invasive trackers. The wide variability in the number and type of permissions and trackers identified across the sample of 154 apps included in this study suggests that there is no single approach to privacy-preserving app development. Consistent with the conclusions of Azad et al [58], many apps seem to collect more permissions than needed, some of which have the potential to breach users' privacy. Although the number and type of permissions varied across the apps, it seems that some governments are particularly interested in collecting more data than others. On the one hand, most of the apps requested nondangerous permissions such as allowing for full network access, preventing the devices from sleeping, and asking to pair Bluetooth devices. The use of Bluetooth technology for contact tracing seems to be almost ubiquitous [58,76] and has been deemed a privacy-preserving approach [45,48,51]; nevertheless, some apps included very invasive permissions or required constant internet connectivity, which might not be available at all times, making real-time exposure notification difficult or impractical [77].

Moreover, some apps require read-and-write privileges to access the phone storage and camera to use QR codes, an approach that seems appropriate for some types of offline self-check actions for digital contact tracing [77]. Other apps require access to the microphone, GPS location, and phone identity to allow for government operations of contact tracing and network exposure notification. Although it can be efficient from a public health perspective, this approach might generate some general privacy concerns. Our findings show a negative correlation between the Permission Accumulated Risk Score and the average rating of the selected apps on the Apple App Store,

which might indicate that users did not like the design or usability or did not trust these apps, expressing a lower rating [78].

Readability and Transparency of Privacy Policies

Most apps (81/132, 61.4%) included privacy policies that were very difficult to read, suggesting that only educated users could interpret the information presented. This finding is consistent with some studies evaluating the readability of contact-tracing app privacy policies [60] and with other apps for other health domains such as mental health [71,79], health and fitness [63], and general health for young generations [54,55].

When it comes to transparency, of the 180 contact-tracing apps identified, 24 (13.3%) did not include a valid link to a privacy policy document or included a link to a policy document that was not specific to the app. Although not many users might read a privacy policy before or after installing an app, not having such a document available can raise concerns about the developers' transparency, negligence, or incompetence [79]. Another notable finding was the inverse relationship between the Permission Accumulated Risk Score and Privacy Transparency Index, suggesting that, the higher the risk of violating one's privacy through app permissions, the lower the level of transparency of the policy document. Although this relationship is based on our *expert assessment* of the documents and the permission data, the data make sense. The data suggest that some developers might collect more data than necessary without feeling the need to communicate this to the users [80].

As trust in governments seems to be dwindling worldwide, it would be expected that contact-tracing apps would follow a truly decentralized framework and be based on transparency and openness principles [25]. Of the 132 privacy policy documents analyzed, most (81/132, 61.4%) achieved an above-average rating in the bespoke Privacy Transparency Index. Most policy documents indicated that the apps collected personal identifiers. Although it provides helpful information about data management, this suggests that a genuinely privacy-preserving and completely anonymous approach to contact tracing may be unfeasible in real-life scenarios [49]. Nevertheless, the privacy-preserving apps (ie, those with low Permission Accumulated Risk Score) had higher ratings on the Apple App Store. Their privacy policy documents had a higher Privacy Transparency Index, suggesting that transparency and privacy can go well together with positive app reviews, which may indicate better user engagement and sustained use.

Strengths and Limitations

This is the first systematic review and evaluation of COVID-19 contact-tracing apps that combines an assessment of the privacy risk and the privacy policies' transparency and readability. An essential strength of this study is the methodological approach following a specific protocol for selection, data extraction, and analysis. Another strength is the availability of the data collected across 154 apps developed worldwide. The limitations of this study include the use of bespoke measures to quantify the level of risk (the Permission Accumulated Risk Score) and the level of transparency (the Privacy Transparency Index). Although these instruments require formal validation, we tried to minimize the potential subjectivity and errors by completing a series of trainings and assessing interrater agreement and reliability indexes to establish a good level of agreement in evaluating the apps. Another limitation was the use of data generated from Google Play as some apps were developed only for iOS and were not included in the study. Unfortunately, the App Store for iOS does not include information about the permissions that the apps require; this is due to the different software architecture between iOS and Android. Another limitation is related to the extreme volatility of the mobile app market and its characteristics. We provided a global snapshot of all available contact-tracing apps as of October 31, 2021, after having monitored the market for approximately a year. Considering that the pandemic is still ongoing, existing contact-tracing apps might disappear, new ones could be developed, or different technological solutions could be adopted to provide exposure notifications (eg, merging databases or aligning data exchange protocols between European or US states). This would imply that the existing apps might have different software permissions and privacy policies. Our database provides a historical classification of contact-tracing apps that were developed over more than a year, and we made such a list of apps available from the Tableau link.

Conclusions

COVID-19 contact-tracing app developers should find a balance between following privacy-preserving frameworks and collecting personal information to serve the needs of public health institutions to ensure efficient and practical support for manual contact-tracing efforts. Developers should reduce the amount of data collected and relate it to the sole purpose of contact tracing. They should also put more effort into making privacy policy documents more accessible and easier to read and providing the information needed to foster trust in governments and institutions for the fight against COVID-19. Better and more useful digital contact-tracing apps would help governments undertake contact-tracing efforts more efficiently and effectively.

Acknowledgments

The authors would like to acknowledge the help of Ms Dalia Sarieldine, a graduate research assistant in the Department of Health Promotion and Community Health, Faculty of Health Sciences, the American University of Beirut, who helped verify the inclusion of selected apps and update the data on the selected apps.

Authors' Contributions

MB conceived and designed the review with intellectual input from IHE, NM, and RAS. MB developed the search strategy and performed the searches. MB coordinated the selection process with the help of NM, RAS, and MAD. MB, NM, RAS, and MAD independently extracted the data. MB performed the data analyses, and NM, RAS, IHE, and MAD contributed to data interpretation. MB drafted the manuscript, which all authors then edited. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of contact-tracing apps identified (N=180).
[XLSX File (Microsoft Excel File), 28 KB - mhealth_v10i7e35195_app1.xlsx]

Multimedia Appendix 2
Characteristics of selected contact-tracing apps with permission data available (n=154).
[XLSX File (Microsoft Excel File), 46 KB - mhealth_v10i7e35195_app2.xlsx]

Multimedia Appendix 3
Permission data for selected contact-tracing apps (n=154).
[XLSX File (Microsoft Excel File), 46 KB - mhealth_v10i7e35195_app3.xlsx]

Multimedia Appendix 4
Privacy policy assessment of selected contact-tracing apps with an existing valid policy document (n=132).
[XLSX File (Microsoft Excel File), 446 KB - mhealth_v10i7e35195_app4.xlsx]

References

1. Fisher D, Wilder-Smith A. The global community needs to swiftly ramp up the response to contain COVID-19. Lancet 2020 Apr 04;395(10230):1109-1110 [FREE Full text] [doi: 10.1016/S0140-6736(20)30679-6] [Medline: 32199470]
2. Heymann DL, Shindo N. COVID-19: what is next for public health? Lancet 2020 Feb 22;395(10224):542-545 [FREE Full text] [doi: 10.1016/S0140-6736(20)30374-3] [Medline: 32061313]
3. Non-pharmaceutical public health measures for mitigating the risk and impact of epidemic and pandemic influenza: annex: report of systematic literature reviews. World Health Organization. 2019. URL: https://apps.who.int/iris/handle/10665/329439 [accessed 2021-04-03]
4. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science 2020 May 22;368(6493):860-868 [FREE Full text] [doi: 10.1126/science.abb5793] [Medline: 32291278]
5. Coronavirus Update (Live) - Worldometer. Worldometer. 2022. URL: https://www.worldometers.info/coronavirus/ [accessed 2021-04-03]
6. Chowdhury R, Luhar S, Khan N, Choudhury SR, Matin I, Franco OH. Long-term strategies to control COVID-19 in low and middle-income countries: an options overview of community-based, non-pharmacological interventions. Eur J Epidemiol 2020 Aug;35(8):743-748 [FREE Full text] [doi: 10.1007/s10654-020-00660-1] [Medline: 32656618]
7. Smith JA, Judd J. COVID-19: vulnerability and the power of privilege in a pandemic. Health Promot J Austr 2020 Apr;31(2):158-160 [FREE Full text] [doi: 10.1002/hpja.333] [Medline: 32197274]
8. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Global Health 2020 Apr;8(4):e488-e496 [FREE Full text] [doi: 10.1016/S2214-109X(20)30074-7] [Medline: 32119825]
9. Contact tracing in the context of COVID-19. World Health Organization. URL: https://www.who.int/publications-detail-redirect/contact-tracing-in-the-context-of-covid-19 [accessed 2021-04-04]
10. Juneau C, Briand A, Pueyo T, Collazzo P, Potvin L. Effective contact tracing for COVID-19: a systematic review. BMJ 2020 Jul 25 (forthcoming). [doi: 10.1101/2020.07.23.20160234]
11. Verhagen LM, de Groot R, Lawrence C, Taljaard J, Cotton M, Rabie H. COVID-19 response in low- and middle-income countries: don't overlook the role of mobile phone communication. Int J Infect Dis 2020 Oct;99:334-337 [FREE Full text] [doi: 10.1016/j.ijid.2020.07.069] [Medline: 32763447]
12. Ahmed N, Michelin RA, Xue W, Ruj S, Malaney R, Kanhere SS, et al. A survey of COVID-19 contact tracing apps. IEEE Access 2020;8:134577-134601. [doi: 10.1109/access.2020.3010226]

13. Collado-Borrell R, Escudero-Vilaplana V, Villanueva-Bueno C, Herranz-Alonso A, Sanjurjo-Saez M. Features and functionalities of smartphone apps related to COVID-19: systematic search in app stores and content analysis. J Med Internet Res 2020 Aug 25;22(8):e20334 [FREE Full text] [doi: 10.2196/20334] [Medline: 32614777]
14. Davalbhakta S, Advani S, Kumar S, Agarwal V, Bhoyar S, Fedirko E, et al. A systematic review of smartphone applications available for corona virus disease 2019 (COVID19) and the assessment of their quality using the mobile application rating scale (MARS). J Med Syst 2020 Aug 10;44(9):164 [FREE Full text] [doi: 10.1007/s10916-020-01633-3] [Medline: 32779002]
15. Kondylakis H, Katehakis DG, Kouroubali A, Logothetidis F, Triantafyllidis A, Kalamaras I, et al. COVID-19 mobile apps: a systematic review of the literature. J Med Internet Res 2020 Dec 09;22(12):e23170 [FREE Full text] [doi: 10.2196/23170] [Medline: 33197234]
16. Li J, Guo X. COVID-19 contact-tracing apps: a survey on the global deployment and challenges. arXiv 2020 May 07. [doi: 10.48550/arXiv.2005.03599]
17. Ming LC, Untong N, Aliudin NA, Osili N, Kifli N, Tan CS, et al. Mobile health apps on COVID-19 launched in the early days of the pandemic: content analysis and review. JMIR Mhealth Uhealth 2020 Sep 16;8(9):e19796 [FREE Full text] [doi: 10.2196/19796] [Medline: 32609622]
18. Danquah LO, Hasham N, MacFarlane M, Conteh FE, Momoh F, Tedesco AA, et al. Use of a mobile application for Ebola contact tracing and monitoring in northern Sierra Leone: a proof-of-concept study. BMC Infect Dis 2019 Sep 18;19(1):810 [FREE Full text] [doi: 10.1186/s12879-019-4354-z] [Medline: 31533659]
19. Hashemian M, Stanley K, Osgood N. Leveraging H1N1 infection transmission modeling with proximity sensor microdata. BMC Med Inform Decis Mak 2012 May 02;12:35 [FREE Full text] [doi: 10.1186/1472-6947-12-35] [Medline: 22551391]
20. Hashemian M, Stanley K, Osgood N. Flunet: automated tracking of contacts during flu season. In: Proceedings of the 8th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks. 2010 Presented at: 8th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks; May 31 - Jun 04, 2010; Avignon, France.
21. Alwashmi MF. The use of digital health in the detection and management of COVID-19. Int J Environ Res Public Health 2020 Apr 23;17(8):2906 [FREE Full text] [doi: 10.3390/ijerph17082906] [Medline: 32340107]
22. Support in the UK for app-based contact tracing of COVID-19. OSF Home. URL: https://osf.io/huqtr [accessed 2022-07-04]
23. McKendry RA, Rees G, Cox IJ, Johnson A, Edelstein M, Eland A, et al. Share mobile and social-media data to curb COVID-19. Nature 2020 Apr;580(7801):29. [doi: 10.1038/d41586-020-00908-6] [Medline: 32235939]
24. Owusu PN. Digital technology applications for contact tracing: the new promise for COVID-19 and beyond? Glob Health Res Policy 2020 Aug 3;5(1):36 [FREE Full text] [doi: 10.1186/s41256-020-00164-1] [Medline: 32766453]
25. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491):eabb6936 [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
26. The Lancet Digital Health. Contact tracing: digital health on the frontline. Lancet Digit Health 2020 Nov;2(11):e561 [FREE Full text] [doi: 10.1016/S2589-7500(20)30251-X] [Medline: 33103094]
27. Fagherazzi G, Goetzinger C, Rashid MA, Aguayo GA, Huiart L. Digital health strategies to fight COVID-19 worldwide: challenges, recommendations, and a call for papers. J Med Internet Res 2020 Jun 16;22(6):e19284 [FREE Full text] [doi: 10.2196/19284] [Medline: 32501804]
28. Braithwaite I, Callender T, Bullock M, Aldridge R. Automated and partially-automated contact tracing: a rapid systematic review to inform the control of COVID-19. medRxiv 2020 [FREE Full text] [doi: 10.1101/2020.05.27.20114447]
29. Jalabneh R, Zehra Syed H, Pillai S, Hoque Apu E, Hussein MR, Kabir R, et al. Use of mobile phone apps for contact tracing to control the COVID-19 pandemic: a literature review. SSRN J 2020. [doi: 10.2139/ssrn.3641961]
30. Anglemyer A, Moore T, Parker L, Chambers T, Grady A, Chiu K, et al. Digital contact tracing technologies in epidemics: a rapid review. Cochrane Database Syst Rev 2020 Aug 18;8:CD013699 [FREE Full text] [doi: 10.1002/14651858.CD013699] [Medline: 33502000]
31. Contact tracing: using digital tools. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/index.html [accessed 2022-06-25]
32. Martin T, Karopoulos G, Hernández-Ramos J, Kambourakis G, Fovino I. Demystifying COVID-19 digital contact tracing: a survey on frameworks and mobile apps. arXivcs 2020 [FREE Full text] [doi: 10.1155/2020/8851429]
33. Digital tools for COVID-19 contact tracing. World Health Organization. 2020. URL: https://www.who.int/publications-detail-redirect/WHO-2019-nCoV-Contact_Tracing-Tools_Annex-2020.1 [accessed 2020-11-04]
34. Mobile applications in support of contact tracing for COVID-19: a guidance for EU/EEA Member States. European Centre for Disease Prevention and Control. URL: https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-mobile-applications-contact-tracing.pdf [accessed 2021-04-03]
35. Michael K, Abbas R. Behind COVID-19 contact trace apps: the Google–Apple partnership. IEEE Consumer Electron Mag 2020 Sep 1;9(5):71-76. [doi: 10.1109/mce.2020.3002492]
36. COVID apps. Wikipedia. URL: https://en.wikipedia.org/wiki/COVID-19_apps [accessed 2021-10-31]
37. A flood of coronavirus apps are tracking us. Now it's time to keep track of them. MIT Technology Review. URL: https://www.technologyreview.com/2020/05/07/1000961/launching-mittr-covid-tracing-tracker/ [accessed 2020-06-15]

38. Wen H, Zhao Q, Lin Z, Xuan D, Shroff N. A study of the privacy of COVID-19 contact tracing apps. In: Security and Privacy in Communication Networks. Cham: Springer International Publishing; 2020.
39. Chiusi F. Digital contact tracing apps: do they actually work? A review of early evidence. AlgorithmWatch. URL: https://algorithmwatch.org/en/analysis-digital-contact-tracing-apps-2021/ [accessed 2021-10-31]
40. Colizza V, Grill E, Mikolajczyk R, Cattuto C, Kucharski A, Riley S, et al. Time to evaluate COVID-19 contact-tracing apps. Nat Med 2021 Mar 15;27(3):361-362. [doi: 10.1038/s41591-021-01236-6] [Medline: 33589822]
41. Thorneloe R, Epton T, Fynn W, Daly M, Stanulewicz N, Kassianos A, et al. Scoping review of mobile phone app uptake and engagement to inform digital contract tracing tools for COVID-19. PsyArXiv 2020. [doi: 10.31234/osf.io/qe9b6]
42. Seto E, Challa P, Ware P. Adoption of COVID-19 contact tracing apps: a balance between privacy and effectiveness. J Med Internet Res 2021 Mar 04;23(3):e25726 [FREE Full text] [doi: 10.2196/25726] [Medline: 33617459]
43. Munzert S, Selb P, Gohdes A, Stoetzer LF, Lowe W. Tracking and promoting the usage of a COVID-19 contact tracing app. Nat Hum Behav 2021 Feb;5(2):247-255. [doi: 10.1038/s41562-020-01044-x] [Medline: 33479505]
44. Corona-Warn-App Open Source Project homepage. Corona-Warn-App Open Source Project. URL: https://www.coronawarn.app/en/ [accessed 2021-04-04]
45. Raskar R, Schunemann I, Barbar R, Vilcans K, Gray J, Vepakomma P, et al. Apps gone rogue: maintaining personal privacy in an epidemic. arXivcs 2020 [FREE Full text]
46. Sharma T, Bashir M. Use of apps in the COVID-19 response and the loss of privacy protection. Nat Med 2020 Aug 26;26(8):1165-1167. [doi: 10.1038/s41591-020-0928-y] [Medline: 32457443]
47. Cho H, Ippolito D, Yu Y. Contact tracing mobile apps for COVID-19: privacy considerations and related trade-offs. arXiv 2020 [FREE Full text] [doi: 10.48550/arXiv.2003.11511]
48. Berke A, Bakker M, Vepakomma P, Raskar R, Larson K, Pentland AS. Assessing disease exposure risk with location data: a proposal for cryptographic preservation of privacy. arXiv 2020 [FREE Full text] [doi: 10.48550/arXiv.2003.14412]
49. Yasaka TM, Lehrich BM, Sahyouni R. Peer-to-peer contact tracing: development of a privacy-preserving smartphone app. JMIR Mhealth Uhealth 2020 Apr 07;8(4):e18936 [FREE Full text] [doi: 10.2196/18936] [Medline: 32240973]
50. Scantamburlo T, Cortés A, Dewitte P, Van Der Eycken, D, Billa V, Duysburgh P, et al. Covid-19 and contact tracing apps: a review under the European legal framework. arXiv 2020 [FREE Full text] [doi: 10.48550/arXiv.2004.14665]
51. Hernández-Orallo E, Calafate CT, Cano J, Manzoni P. Evaluating the effectiveness of COVID-19 Bluetooth-based smartphone contact tracing applications. Applied Sci 2020 Oct 13;10(20):7113. [doi: 10.3390/app10207113]
52. Peel D. Privacy trust framework. SSRN J 2013:13 [FREE Full text] [doi: 10.2139/ssrn.2231667]
53. Federal Trade Commission. Mobile Privacy Disclosures Building Trust Through Transparency. Collingdale, PA: DIANE Publishing Company; 2013.
54. Mulder T. Health apps, their privacy policies and the GDPR. Eur J Law Technol 2019;10(1):3 [FREE Full text]
55. Das G, Cheung C, Nebeker C, Bietz M, Bloss C. Privacy policies for apps targeted toward youth: descriptive analysis of readability. JMIR Mhealth Uhealth 2018 Jan 04;6(1):e3 [FREE Full text] [doi: 10.2196/mhealth.7626] [Medline: 29301737]
56. Oldeweme A, Märtins J, Westmattelmann D, Schewe G. The role of transparency, trust, and social influence on uncertainty reduction in times of pandemics: empirical study on the adoption of COVID-19 tracing apps. J Med Internet Res 2021 Feb 08;23(2):e25893 [FREE Full text] [doi: 10.2196/25893] [Medline: 33465036]
57. Elkhodr M, Mubin O, Iftikhar Z, Masood M, Alsinglawi B, Shahid S, et al. Technology, privacy, and user opinions of COVID-19 mobile apps for contact tracing: systematic search and content analysis. J Med Internet Res 2021 Feb 09;23(2):e23467 [FREE Full text] [doi: 10.2196/23467] [Medline: 33493125]
58. Azad MA, Arshad J, Akmal SM, Riaz F, Abdullah S, Imran M, et al. A first look at privacy analysis of COVID-19 contact-tracing mobile applications. IEEE Internet Things J 2021 Nov 1;8(21):15796-15806. [doi: 10.1109/jiot.2020.3024180]
59. Blacklow S, Lisker S, Ng M, Sarkar U, Lyles C. Usability, inclusivity, and content evaluation of COVID-19 contact tracing apps in the United States. J Am Med Inform Assoc 2021 Aug 13;28(9):1982-1989 [FREE Full text] [doi: 10.1093/jamia/ocab093] [Medline: 34022053]
60. Zhang M, Chow A, Smith H. COVID-19 contact-tracing apps: analysis of the readability of privacy policies. J Med Internet Res 2020 Dec 03;22(12):e21572 [FREE Full text] [doi: 10.2196/21572] [Medline: 33170798]
61. Elhajj I, Abou ML, Germani A, Bardus M, Dohna H, Ghattas H, et al. Ma3an-together against corona, the official Lebanese contact tracing app: experience and lessons learned. In: Emerging science, frontier technologies, and the SDGs - Perspectives from the UN system and science and technology communities. New York: United Nations Interagency Task Team on Science, Technology and Innovation for the Sustainable Development Goals; 2021.
62. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016 Mar 10;13:35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
63. Rowan M, Dehlinger J. A privacy policy comparison of health and fitness related mobile applications. Procedia Comput Sci 2014;37:348-355. [doi: 10.1016/j.procs.2014.08.051]
64. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. J Am Med Inform Assoc 2015 Apr;22(e1):e28-e33. [doi: 10.1136/amiajnl-2013-002605] [Medline: 25147247]

65. Contact tracing apps. Council of Europe. 2022. URL: https://www.coe.int/en/web/data-protection/contact-tracing-apps [accessed 2020-07-13]
66. Sholtz M. COVID tracing app roundup: all of the countries and US states that currently offer exposure notification apps. Android Police. URL: https://www.androidpolice.com/2020/11/21/covid-tracing-apps-ens-android/ [accessed 2020-11-27]
67. Rahman M. List of countries using Google and Apple's COVID-19 Contact Tracing API. XDA developers. 2021. URL: https://xda-developers.com/google-apple-covid-19-contact-tracing-exposure-notifications-api-app-list-countries/ [accessed 2021-08-06]
68. Exodus privacy homepage. Exodus Privacy. 2022. URL: https://exodus-privacy.eu.org/en/ [accessed 2021-08-07]
69. Manifest permission. Android Developer. URL: https://developer.android.com/reference/android/Manifest.permission [accessed 2020-07-13]
70. Rai P. Android Application Security Essentials. Birmingham, UK: Packt Publishing; 2013.
71. O'Loughlin K, Neary M, Adkins EC, Schueller SM. Reviewing the data security and privacy policies of mobile apps for depression. Internet Interv 2019 Mar;15:110-115 [FREE Full text] [doi: 10.1016/j.invent.2018.12.001] [Medline: 30792962]
72. Continents of the World. Nations Online. URL: https://www.nationsonline.org/oneworld/continents.htm [accessed 2020-12-06]
73. Automatic readability checker. Readability Formula. URL: https://readabilityformulas.com/free-readability-formula-tests.php [accessed 2020-12-02]
74. Bardus M. Systematic review of C19-Contact tracing apps: App List - Marco Bardus. Tableau Public. 2022. URL: https://public.tableau.com/views/SR-C19-ContactTracingApps/AppList?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link [accessed 2021-09-02]
75. Phillips N. The coronavirus is here to stay - here's what that means. Nature 2021 Feb 16;590(7846):382-384. [doi: 10.1038/d41586-021-00396-2] [Medline: 33594289]
76. Al Jutail M, Al-Akhras M, Albesher A. Associated risks in mobile applications permissions. J Inf Security 2019;10(02):69-90. [doi: 10.4236/jis.2019.102004]
77. Min-Allah N, Alahmed BA, Albreek EM, Alghamdi LS, Alawad DA, Alharbi AS, et al. A survey of COVID-19 contact-tracing apps. Comput Biol Med 2021 Oct;137:104787 [FREE Full text] [doi: 10.1016/j.compbiomed.2021.104787] [Medline: 34482197]
78. Wang H, Wang L, Wang H. Market-level analysis of government-backed COVID-19 contact tracing apps. In: Proceedings of the 35th IEEE/ACM International Conference on Automated Software Engineering Workshops. 2020 Presented at: ASE '20: 35th IEEE/ACM International Conference on Automated Software Engineering; Sep 21 - 25, 2020; Virtual Event Australia. [doi: 10.1145/3417113.3422186]
79. Robillard JM, Feng TL, Sporn AB, Lai J, Lo C, Ta M, et al. Availability, readability, and content of privacy policies and terms of agreements of mental health apps. Internet Interv 2019 Sep;17:100243 [FREE Full text] [doi: 10.1016/j.invent.2019.100243] [Medline: 30949436]
80. Huckvale K, Prieto JT, Tilney M, Benghozi P, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med 2015 Sep 25;13(1):214 [FREE Full text] [doi: 10.1186/s12916-015-0444-y] [Medline: 26404673]

Abbreviations

CDC: Centers for Disease Control and Prevention
ENS: Exposure Notification System
GDPR: General Data Protection Regulation
MIT: Massachusetts Institute of Technology
NPI: nonpharmacological intervention

Edited by L Buis, A Mavragani; submitted 24.11.21; peer-reviewed by A Thongprasert, F Folkvord; comments to author 23.12.21; revised version received 04.01.22; accepted 17.02.22; published 12.07.22.

Please cite as:
Bardus M, Al Daccache M, Maalouf N, Al Sarih R, Elhajj IH
Data Management and Privacy Policy of COVID-19 Contact-Tracing Apps: Systematic Review and Content Analysis
JMIR Mhealth Uhealth 2022;10(7):e35195
URL: https://mhealth.jmir.org/2022/7/e35195
doi:10.2196/35195
PMID:35709334

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org