Running title: Randomized trial of app use and fertility

Article title: A randomized trial of web-based fertility-tracking software and fecundability

Lauren A. Wise, Sc.D.[1], Tanran R. Wang, M.P.H.[1], Joseph B. Stanford, M.D.[2], Amelia K. Wesselink, Ph.D.[1], Collette N. Ncube, Dr.P.H.[1], Kenneth J. Rothman, Dr.P.H.[1], Eleanor J. Murray, Ph.D.[1]

[1] Department of Epidemiology, Boston University School of Public Health, Massachusetts
[2] Office of Cooperative Reproductive Health, Division of Public Health, Department of Family and Preventive Medicine, University of Utah, Salt Lake City, Utah.

Corresponding author:
Dr. Lauren A. Wise
Department of Epidemiology
Boston University School of Public Health
715 Albany Street
Boston, MA 02118
617-358-3424
lwise@bu.edu

Article type: Clinical trial

Funding: This research was supported by grants R21HD072326 and R01HD086742, from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, USA.

Disclosures: Dr. Wise reports in-kind donations from FertilityFriend.com (2013-2019) and Kindara.com (2019+) (free fertility apps for PRESTO participants), Swiss Precision Diagnostics (free home pregnancy tests for PRESTO participants), and Sandstone Diagnostics (discounted semen test kits for PRESTO participants) during the conduct of the study, and consulting fees from AbbVie.com for work unrelated to this paper. Dr. Stanford reports personal fees from Swiss Precision Diagnostics, outside the submitted work. Dr. Ncube is supported by National Institute on Minority Health and Health Disparities (NIMHD) Award Number K01-MD013911.

Attestation Statement:
• We do not have any knowledge about whether participants in our study were concomitantly involved in other randomized trials. Because our trial did not involve randomization of a medical treatment, supplement, or device, if our participants were indeed concomitantly involved in other randomized trials, their involvement in other trials should not have introduced any concerns with respect to human subjects research.
• Data from this particular analysis have not been previously published

Data sharing: We do not have permission from our participants to share individual-level data.

Word count: abstract=357; text=4,879.

Trial Registration Number: Not applicable

Capsule: No appreciable associations were observed in intent-to-treat analyses. In secondary per-protocol analyses that accounted for adherence, randomization to FF was associated with slightly greater fecundability among selected subgroups of participants.

Structured Abstract

Objective: To assess the effect of randomization to a premium subscription of FertilityFriend.com (FF), a mobile computing fertility-tracking app, on fecundability.

Design: Parallel non-blinded randomized controlled trial.

Setting: Pregnancy Study Online (PRESTO), a North American preconception cohort study.

Patients: Female-identified participants aged 21-45 years who had been attempting conception for ≤6 menstrual cycles at enrollment (baseline).

Methods: At baseline, 5,532 of participants were randomized with 50% probability to receive a premium FF subscription. Participants completed bimonthly follow-up questionnaires until pregnancy or a censoring event, whichever came first. We first performed an intent-to-treat analysis of the effect of randomization to FF on fecundability. In secondary analyses, we used a per-protocol approach that accounted for adherence in each trial arm. In both analyses, we used proportional probabilities regression models to estimate fecundability ratios (FR) and 95% confidence intervals (CI) comparing those randomized vs. not randomized to FF, and applied inverse probability weights to account for loss-to-follow-up (intent-to-treat and per-protocol analyses) and adherence (per-protocol analyses only).

Results: Using life-table methods, 64% of the 2,775 participants randomized to FF and 63% of the 2,767 participants not randomized to FF conceived during 12 cycles; these respective percentages were each 70% among those with 0-1 cycles of attempt time at enrollment. Of those randomized to FF, 72% were defined as adherent (68% of observed menstrual cycles). In intent-to-treat analyses, there was no appreciable association overall (FR=0.97, 95% CI: 0.90-1.04) or within strata of pregnancy attempt time at enrollment, age, education, or other characteristics. In per-protocol analyses, we observed little association overall (FR=1.06, 95% CI: 0.99-1.14), but weak-to-moderate positive associations among participants who had longer attempt times at enrollment (FR=1.15, 95% CI: 0.98-1.35 for 3-4 cycles; 1.14, 95% CI: 0.87-1.48 for 5-6 cycles), were aged <25 years (FR=1.29, 95% CI: 1.01-1.66), had ≤12 years of education (FR=1.32, 95% CI: 0.92-1.89), or were non-users of hormonal contraception within 3 months before enrollment (FR=1.10, 95% CI: 1.02-1.19).

Conclusions: No appreciable associations were observed in intent-to-treat analyses. In secondary per-protocol analyses that accounted for adherence, randomization to FF was associated with slightly greater fecundability among selected subgroups of participants; however, these results are susceptible to unmeasured confounding.

Key Words: fecundability; fertility; time-to-pregnancy; preconception; randomized trial

Introduction

In recent years, there has been a substantial increase in the use of mobile computing applications ("apps") to track menstrual cycles and record fertility signs (1, 2). In 2020, an estimated 50 million individuals worldwide used menstrual-tracking apps (3). Many apps estimate the days when intercourse is most likely to lead to conception (i.e., fertile window) (4); however, the extent to which use of such apps influences fecundability—the per-cycle probability of conception—is unclear (5, 6). In an observational study, we reported previously that self-reported use of some fertility apps—particularly those that track cycle days, cervical fluid, cervix position, basal body temperature (BBT), and urine luteinizing hormone (LH)—was associated with 12-20% greater fecundability (7). These results are consistent with data from most (8-13) but not all (14) trials and cohort studies, including ours (13), which indicate that use of fertility indicators to identify the fertile window increases fecundability. They also agree with results from a 2020 randomized trial in which use of the ClearBlue E1C/LH test app to time intercourse within the fertile window increased the likelihood of conceiving within two menstrual cycles (test arm: 36.2% vs. control arm: 28.6%) (15). To date, however, no randomized trial has evaluated the effect of using a fertility app that tracks multiple fertility indicators on fecundability. Observational studies of this association are prone to selection bias and confounding because fertility app users may differ with respect to underlying fecundity and other socio-demographic characteristics (e.g., age, education, income, and fertility awareness).

In this study, we conducted a parallel non-blinded randomized controlled trial to assess the extent to which randomization of a free premium subscription to FertilityFriend.com (FF), a popular mobile computing app that permits users to track their menstrual cycles and fertility signs, was associated with fecundability.

Methods

Study design and population

This trial was nested within Pregnancy Study Online (PRESTO), an ongoing web-based preconception cohort study of pregnancy planners. Study methods have been described in detail previously (16). Eligible participants identify as female and are aged 21-45 years, reside in the United States or Canada, and are attempting to conceive without the use of fertility treatment at cohort entry. Participants complete a comprehensive baseline questionnaire with items on socio-demographics, behaviors, and reproductive and medical histories, and medication use. Participants complete follow-up questionnaires every 8 weeks for up to 12 months to ascertain pregnancy status and update any factors that may have changed over time.

Fertility Friend (FF)

Fertility Friend (FF) is a mobile app integrated with an online web-based platform for users to track their fertility awareness indicators. It provides a graphical summary of each menstrual cycle, with the ability for the user to enter daily observations for basal body temperature, urine LH and/or estrogen testing, cervical mucus observations, vaginal bleeding observations, and intercourse. The app database applies algorithms based on these features to identify the fertile window, focused on the most fecund (fertile) days, which precede and include the estimated day of ovulation (17). It identifies the days of the estimated fertile window for the user prospectively on its interface. It also allows the user to output graphical summaries of each cycle, or long-term summaries over many cycles (e.g., means, minimum, maximum of cycle lengths, day of

ovulation, luteal lengths), as pdf documents, which the user can then give to a clinician for review.

Randomization scheme

PRESTO participants who enrolled between June 2013 through March 2019 were considered for inclusion in this trial. All participants who reported never use of FF at baseline and never had an FF account associated with their email address were eligible for randomization. Eligible participants were randomized immediately after enrollment (defined as completion of the baseline questionnaire) by an independent programmer who implemented a simple randomization scheme via computer algorithm using the Mersenne Twister random number generator (18). The procedure randomized 8,397 participants with 50% probability to receive a free premium FF subscription (N=4,162) or the standard study procedures only (N=4,235). All participants had the opportunity to contribute the full 12 months of follow-up after enrollment.

Immediately after randomization, investigators and participants were made aware of the group to which participants were randomized. Participants assigned to receive FF had to click on a link embedded within the invitation email to access the subscription. Other than the standard information found in the FF app, no additional information was provided to participants about the app, nor was there any stated requirement or additional incentive to use the app. In the app, participants could record day-specific data on menstrual flow (none, spotting, light, medium, heavy), BBT, cervical fluid consistency (dry, sticky, creamy, watery, egg white), sexual intercourse (no vs. yes; if yes, AM, PM, or both), secondary signs (e.g., cervical position), and results from testing devices (ovulation predictor kit for urine luteinizing hormone (OPK); progesterone test; fern microscope; and pregnancy test). Using a secure password-protected

server, study investigators downloaded de-identified FF app data from PRESTO participants who were randomized to receive the premium FF subscription.

The study protocol was approved by the Institutional Review Board at Boston University Medical Campus, and online informed consent was obtained from all participants. There was no separate consent form for the trial. In the main study's consent form, all participants were informed that a random subset of never users of FF would be offered a premium subscription after enrollment. The trial was not registered with clinicaltrials.gov because it did not involve randomization of a medical treatment, supplement, or device.

Exclusions

From June 2013 through March 2019 (i.e., time period when FF was randomized), 10,405 participants completed the baseline questionnaire. Of these, 2,008 participants were not eligible for randomization because they were former or current users of FF. Of the remaining 8,397 participants, we randomized 4,162 to FF and 4,235 to the standard protocol. As part of our standard fecundability analyses, we excluded 120 participants with missing/implausible last menstrual period (LMP) data (defined as having baseline LMP >6 months before the baseline questionnaire completion date or any time after the baseline questionnaire completion date), and 1,158 participants who were pregnant at study entry as identified using dates of LMP and first positive pregnancy test (47% randomized and 53% not randomized to FF). We then excluded 1,577 participants who had been attempting pregnancy for >6 cycles at enrollment, to reduce the potential for reverse causation bias. We also anticipated that the effect among participants who had been attempting pregnancy for >6 cycles at enrollment might differ from the effect among those attempting for ≤6 cycles. Our results should thus be interpreted as relevant primarily to this

latter group. The final analytic population comprised 5,542 participants: 2,775 randomized to FF and 2,767 not randomized to FF (Supplemental Figure 1).

Assessment of time-to-pregnancy

We estimated time-to-pregnancy (TTP) using data from the baseline and follow-up questionnaires. At baseline, participants reported their LMP date, usual cycle length (only for those with regular cycles), and the number of cycles they had attempted conception. On each follow-up questionnaire, participants reported their most recent LMP date, whether they had conceived since the previous questionnaire, and the method of pregnancy confirmation. In these analyses, we accepted confirmation of pregnancy via positive home pregnancy test, blood test from doctor's office, or ultrasound. Among those with irregular cycles, defined as those who reported not being able to predict from one menstrual period to the next about when their next menstrual period would start (when not using hormonal contraception), we estimated cycle length based on date of LMP at baseline and prospectively-reported LMP dates during follow-up. We estimated intervening LMP dates between questionnaires by subtracting most recent cycle length or typical cycle length from the LMP reported on follow-up questionnaires and repeated this process until the calculated LMP was within 15 days of the LMP reported on the previous questionnaire. We then calculated TTP as follows: menstrual cycles of attempt time at study entry + total number of self-reported and calculated LMPs between baseline and end of observation.

Assessment of covariates

At baseline, we collected demographic and clinical information, including age, height, weight, relationship duration, marital status, race/ethnicity, income, education, hours of sleep per

night, parity, gravidity, vitamin use, caffeine intake, smoking status, marijuana use, alcohol consumption, sugar-sweetened soda intake in the past month, total metabolic equivalents of task (METs) from vigorous and moderate exercise in the past week (19), depressive symptoms via the Major Depression Inventory (20), the 10-item Perceived Stress Scale (PSS) in the past month (21), average intercourse frequency in the past month, contraception history, and infertility. We calculated body mass index (BMI) as weight (kilograms) divided by height (meters) squared.

On the baseline and follow-up questionnaires, we asked participants whether they were currently doing anything to improve their chances of conception (e.g., recording BBT, monitoring cervical mucus, using OPKs) and if they used "any software program and/or web-based or phone "app" to record menstrual cycle data and/or fertility signs?" If they responded yes, they were asked to write the name of the program or app using a free text response.

Data analysis

Couples that did not conceive within 12 cycles of attempted conception were censored at 12 cycles, the time after which couples typically seek infertility treatment (22). Couples contributed menstrual cycles to the analysis from enrollment until reported pregnancy (49.4%) or a censoring event (initiation of fertility treatment: 7.5%; cessation of pregnancy attempts: 3.3%; loss to follow-up: 18.7%; or 12 cycles of attempt: 21.1%), whichever came first. Participants who were and were not lost-to-follow-up were similar with respect to baseline age, last method of contraception used, parity, sleep duration, and depressive symptoms, but differed according to randomization to FF or free home pregnancy tests, BMI, educational attainment, income, marital status, geographic region of residence, use of prenatal supplements, infertility history, pregnancy attempt time at enrollment, and male partner participation (Supplemental Table 1).

We used life-table methods to compute the percentage of couples that conceived during follow-up, after accounting for censoring (23). To account for variation in attempt times at study entry (range: 0-6 cycles) and to reduce bias due to left truncation, we analyzed observed cycles using the Andersen-Gill data structure (i.e., we outputted one observation per menstrual cycle at risk, excluding earlier cycles that were unobserved) (24). For example, if a participant enrolled in the study after having tried to conceive for 2 cycles and was followed for an additional 4 observed cycles until reported pregnancy, that participant would contribute a total of 4 observed cycles starting at cycle 3. We used proportional probabilities regression models (25) to estimate fecundability ratios (FRs) and 95% confidence intervals (CIs) for the association between trial arm and fecundability. The FR represents the ratio of fecundability in each group compared with the reference category. Our models included cycle-specific indicator variables for "menstrual cycle at risk" to account for the decline in fecundability in the study population over time (25).

We first performed an intent-to-treat analysis to compare fecundability across the trial arms. In these analyses, those who showed no evidence of adherence to their assigned intervention (e.g., randomized to FF but showed no evidence of using the app) were analyzed in their assigned group. We then performed a separate per-protocol analysis in which we sought to analyze the relationship between initiating use of FF and fecundability. Exposure was defined as adherence to the assigned protocol *during or before* the estimated follicular phase of a given menstrual cycle: [(date LMP_n), (date LMP_{n+1} -14 days)]. We defined adherence for those not randomized to FF as the continued absence of any reported FF use on a follow-up questionnaire. We defined adherence for those randomized to FF as initiation of FF use, beginning on the earlier of two dates:

1) Date participant first clicked on the invitation link to access the FF app (68% of all

menstrual cycles) or the date of earliest reported data in FF app (40% of all cycles), whichever occurred later (because those who logged into the app could have entered data retrospectively in calendar time).

2) Date of first reported FF use on a follow-up questionnaire (13% of all cycles).

Once a participant moved from "non-adherent" to "adherent" in the FF arm or "adherent" to "non-adherent" in the control arm (i.e., user of FF in both scenarios), they remained coded in their latter status for all subsequent observed menstrual cycles. The per-protocol analyses discarded observed menstrual cycles for which participants were not adherent to their assigned protocol and applied weights to account for non-adherence. In each cycle of follow-up, we created weights for the probability of adherence based on potential determinants of FF use (Supplemental Table 2). We conducted sensitivity analyses in which we 1) redefined adherence as "not using any fertility app" in the control arm and 2) redefined adherence as "use of any fertility app" in the intervention arm and as "not using any fertility app" in the control arm.

In both intention-to-treat and per-protocol analyses, we derived and applied weights to account for bias due to differential loss-to-follow-up. Briefly, we used inverse probability weighting to create a pseudo-population of participants who, had they not been lost to follow-up, would have contributed more complete follow-up to the study. Using data from all participants enrolled at the start of follow-up, we developed logistic regression models for the probability of continuing in the study at each follow-up cycle, conditional on remaining uncensored at the previous follow-up. The model contained a set of variables, some of which were time-varying, hypothesized to predict loss to follow-up (Supplemental Table 1). We fit separate logistic regression models that included only time-invariant variables as independent variables. We computed stabilized weights by dividing the predicted probability of loss to follow-up from the

second model (containing time-invariant variables only) by the predicted probability of loss to follow-up from the first model (time-varying and time-invariant variables) and multiplying by stabilized weights from previous cycles. The resulting weights were inversely proportional to the probability of remaining under study at each cycle. We then applied these weights to our analyses at each time point (i.e., follow-up cycle). We used a similar approach to account for non-adherence (Supplemental Table 2). In the per-protocol analyses only, we multiplied the weights for adherence with the weights for loss-to-follow-up.

In both analyses, we ran additional models in which we discarded all pregnancies and person-time that occurred in each participant's first contributed menstrual cycle of observation. We reasoned that it would take at least one cycle for those randomized to FF to use the software in a way that could have meaningfully influenced fecundability. Finally, we also stratified by factors that could potentially modify the association between FF use and fecundability: attempt time at entry, age, education, parity, infertility history, recency of hormonal contraceptive use, and non-use of fertility app at baseline.

Missingness for covariates ranged from <0.1% (prior pregnancy, history of subfertility, caffeine use, and history of anxiety) to 3.4% for household income. We used the fully conditional specification (FCS) method to multiply impute missing data for exposures, covariates, and pregnancy status (26). To reduce selection bias from differential loss to follow-up, we assigned one cycle of follow-up for the 13% of participants with no data from follow-up questionnaires (N=726) and then imputed their pregnancy status (yes vs. no) via multiple imputation. The imputation models included >100 covariates. To ensure validity without compromising computing efficiency (27), we created twenty imputed datasets using SAS PROC MI and then combined coefficient and standard error estimates from the datasets using SAS

PROC MIANALYZE (28). We used linear regression to impute continuous variables, the discriminant function method to impute nominal categorical variables, and logistic regression to impute dichotomous and ordinal categorical variables. We examined the trace plots for continuous variables and compared the frequency distribution of dichotomous and categorical variables before and after imputation to ensure the quality of the imputation. Analyses were performed using SAS software version 9.4 (28).

Results

Table 1 presents demographic, reproductive, and behavioral characteristics of the participants at baseline, stratified by randomization status. The median age of participants was 30 years (interquartile range: 27-33 years). The majority of participants were married (89%) and college-educated (70%). Fewer than 10% reported a history of infertility; 32% were parous; 76% reported taking folic acid, multivitamins, or prenatal supplements; 32% had a BMI ≥30 kg/m^2; and nearly 40% had used a hormonal method as their most recent form of contraception.

Of the 2,775 participants randomized to FF, 1,990 (72%) of participants and 9,567 (68%) of cycles were defined as adherent. Of the 2,767 participants not randomized to FF, 2,766 (99.9%) of participants and 13,199 (97%) of cycles were defined as adherent. Figure 1 shows descriptive data among participants assigned to the FF arm and who entered data into the FF app at any point during follow-up [n=1,102 (40%)]. Median time to first FF app use was 0 days (interquartile range: 0-4 days). Among participants who entered the start and end date of at least one cycle into FF, on average 52% of the days in a cycle were logged (interquartile range: 23%-93%). Among the 1,102 (40%) participants who entered some data into the FF app during follow-up (5,956 total cycles), 1074 participants (97% of FF users; 4,021 cycles; 68% of total

cycles) entered data on menstrual bleeding dates; 897 participants entered data on intercourse (81% of FF users; 3,147 cycles; 53% of total cycles); 793 participants entered data on cervical fluid (72% of FF users; 2409 cycles; 40% of total cycles); 573 participants entered data on BBT (52% of FF users; 1839 cycles; 31% of total cycles); 490 participants entered data on OPK (44% of FF users; 1459 cycles; 25% of total cycles); and 184 participants entered data on cervical position/openness/texture (17% of FF users; 412 cycles; 7% of total cycles).

Overall, 81.3% of participants in the analytic cohort completed follow-up (i.e., conceived or reached another study endpoint). Using life-table methods, we identified pregnancy during 12 cycles of follow-up for 64% of the 2,775 participants randomized to FF and 63% of the 2,767 participants not randomized to FF. When restricting to participants with 0-1 cycles of attempt time at enrollment, those respective cumulative pregnancy rates were 70% in each arm.

Table 2 reports FRs from the intent-to-treat analysis, stratified by potential effect measure modifiers. The overall FR comparing those who were and were not randomized to FF was 0.97 (95% CI: 0.90-1.04); results were identical after removing the first cycle of observation contributed by each participant. While there was a slightly elevated but imprecise FR among participants aged <25 years (FR=1.18, 95% CI: 0.91-1.53), we did not observe any consistent patterns of association when we stratified results by attempt time at cohort entry, parity, history of infertility, cycle regularity, or education. Results were slightly inverse among non-users of a different fertility app at baseline (FR=0.89, 95% CI: 0.78-1.01). There was little evidence of an association among non-users of hormonal contraception within 3 months before study enrollment.

Table 3 presents FRs from the per-protocol analysis, which accounts for adherence, stratified by potential effect measure modifiers. The overall FR comparing those who did and did

not initiate FF use according to their randomization assignment was 1.06 (95% CI: 0.99-1.14); results were null after removing the first cycle of adherence contributed by each participant (FR=0.99, 95% CI: 0.93-1.07). There were weak to moderate positive associations among participants with longer attempt times at enrollment (FR=1.15, 95% CI: 0.98-1.35 for 3-4 cycles; 1.14, 95% CI: 0.87-1.48 for 5-6 cycles), were aged <25 years (FR=1.29, 95% CI: 1.01-1.66), had ≤12 years of education (FR=1.32, 95% CI: 0.92-1.89), or were non-users of hormonal contraception within 3 months before enrollment (FR=1.10, 95% CI: 1.02-1.19). In addition, when we stratified results by attempt cycle, we did not observe any time trends in the estimated per-cycle FR (data not shown). Use of inverse probability weights to account for loss to follow-up did not make a large difference in the results of either the intent-to-treat or the per-protocol analyses (Supplemental Tables 3 and 4). Finally, when we re-ran the per-protocol analyses in which we 1) redefined adherence as "not using any fertility app" in the control arm or 2) redefined adherence as "use of any fertility app" in the intervention arm and as "not using any fertility app in the control arm, effect estimates were generally stronger than or similar to, respectively, the original per-protocol results (Supplemental Table 5).

Discussion

In this randomized controlled trial conducted within a North American prospective cohort study of pregnancy planners, we did not find any appreciable association between randomization to FertilityFriend.com (FF), a popular fertility tracking app, and fecundability in intent-to-treat analyses. Low adherence in the intervention arm likely contributed to these null findings, thereby limiting our ability to rely on the intent-to-treat analyses for causal inference. The most reasonable explanation for low adherence in the intervention arm was that 69% of participants

were already using other fertility apps at entry into the study and participants were likely satisfied with their current apps. Participants may have also already collected several cycles' worth of data and did not want to transfer their data to a new app, even if FF was perceived to be of better quality.

In secondary per-protocol analyses that accounted for adherence, we found evidence for a weak positive association between initiating use of FF and fecundability. In the latter analyses, we observed slightly stronger associations between FF use and fecundability among participants who were younger, had lower attained levels of education, had longer pregnancy attempt times at study enrollment, and who were non-users of hormonal contraception within the 3 months before study enrollment.

The analysis of randomized exposure is an improvement over previous studies of observational data because users of fertility-tracking apps may be different from non-users in ways that are difficult to measure. For example, users of fertility-tracking apps may be more likely to be older and more educated than non-users, and have greater awareness of their fertility signs. On the other hand, users of fertility-tracking apps may also have a higher prevalence of fertility problems that resulted in app use. We reasoned that we would still have potential for bias in our analyses if non-users of apps at baseline who experienced difficulties conceiving during follow-up were more likely to initiate app use over time, potentially biasing FRs downward. Accounting for adherence in the per-protocol analyses helped address this concern. In addition, when we stratified the per-protocol estimates by attempt cycle, we did not observe any time trends in the estimated per-cycle FR.

Our definition of adherence was limited in that, for some participants, we did not know the precise date during which data were first entered into the FF app. We relied on data from date

of first click of the invitation link to the app, and we used that or any later date of actual FF data entry or reported first use on a follow-up questionnaire as the date of adherence. Data on app use were self-reported on all questionnaires, in addition to being downloaded directly from the app itself for participants randomized to receive FF. Thus, for the FF users who were not randomized to the FF intervention arm, we did not have any comparable information about the features of FF use. As a result, our trial could not assess the impact of perfect, consistent, or continuous use of FF on fecundability and had to instead estimate the effect of initiation of use only. Sensitivity analyses in which we varied the definition of adherence in the control arm (i.e., no use of any fertility app) or both the intervention arm (i.e., use of any fertility app) and control arm (i.e., no use of any fertility app) showed generally similar or stronger results. Some studies have documented inconsistent recording of data among >50% of users of similar apps (17, 29). Sporadic app use would be expected to attenuate results if more consistent use improves fecundability. While our protocol included allocation concealment, which was achieved by having an independent computer programmer code and implement the randomization scheme, it did not include participant or investigator "blinding" in randomizing the intervention. Stratification of data by multiple factors may have introduced potential for chance findings. Finally, the prospective assessment of FF use was based on questionnaires completed every 2 months, which would not capture more frequent changes; this could have resulted in non-differential misclassification of FF use obtained outside of the app itself.

In the per-protocol analyses, we cannot rule out potential for unmeasured or residual confounding, whereby use of FF reflects behaviors that have not been fully accounted for by measured covariates. We did, however, account for many covariates known to be associated with fecundability, including age, income, education, parity, BMI, infertility history, and use of

prenatal supplements. The intent-to-treat analyses, in which investigators randomized FF (in a 1:1 ratio) to participants, are robust to measured and unmeasured confounders at baseline. However, the intent-to-treat effect is expected to be closer to the null than the true effect of app use in this trial because adherence to FF use was low (30). Moreover, many participants were already using other fertility-tracking apps (69%), as the prevalence of app use among pregnancy planners tends to be higher than the general population. Despite the limitations of the per-protocol analysis, we believe the per-protocol effect asks a clinically relevant question: what is the impact on fecundability of initiating use of the FF app, had everyone adhered to their treatment assignment? Previous research has demonstrated that per-protocol effect estimates are of greatest interest to individuals for making treatment decisions when they intend to adhere (31). Since it is reasonable to expect that individuals attempting to conceive would know *a priori* whether they expect to initiate the use of a fertility app, the effect of initiation will likely be useful information. In comparison, while other usage questions, such as the impact of consistent or prolonged use, may also be of clinical interest, it may be less likely that an individual would know *a priori* whether they are likely to use a new app consistently.

PRESTO enrolled participants who were trying to conceive spontaneously. All participants completed self-administered questionnaires via the Internet. The study population overrepresents non-Hispanic White participants (>80%) with higher educational attainment and household income than the general population. For these reasons, the results of this study may not extend to the general population.

Previous studies of observational data on use of fertility-tracking apps and fertility indicators, including our own (7, 13), have found positive associations between the use of fertility indicators and fecundability (8, 9, 11-13, 32). While a 2020 randomized trial in which

use of an ovulation testing app to time intercourse within the fertile window increased the likelihood of conceiving within two menstrual cycles (15), other randomized trials of use of fertility indicators have produced inconsistent results, some positive and some null, likely reflecting difficulties in conducting trials among pregnancy planners (10, 12, 14, 33). Our overall effect estimates for the association between FF use and fecundability in this trial are weaker than those reported in a prior publication based on an observational analysis from PRESTO, in which we found effect sizes of 20% (95% CI 1.13-1.28) for use of selected apps including FF and other similar apps at baseline vs. non-use; and 12% (95% CI 1.04-1.21) for time-varying use of selected apps vs. non-use (7). Nevertheless, our present results remain consistent with the possibility that app use may improve fecundability and reduce TTP for some subgroups. Our observation of stronger effect among younger individuals and those with lower educational attainment is not entirely surprising given that these individuals may have lower awareness of their fertility indicators (34). App usage may have led to substantial improvements in their knowledge about fertility signs and identification of the fertile window. We did not stratify by age in the prior report (7) and we are not aware of any previous publications that have done so, which limits our ability to corroborate the possibility of a differential effect by the user's age. Similarly, those with longer pregnancy attempt times at enrollment may have also been less aware of their fertility signs, which may have contributed to their delays in conception before FF assignment. Stronger effects among non-users of hormonal contraception in the 3 months before enrollment are plausible because recent use of hormonal contraception temporarily reduces fecundity (35, 36). Thus, recent hormonal contraceptive use may have obscured any effect of app use on fecundability.

In conclusion, we found no appreciable association between randomization to FF and fecundability in intent-to-treat analyses. The high percentage of non-adherence in the intervention arm may have contributed to these null results. In secondary per-protocol analyses that accounted for adherence, initiating use of FF—a fertility-tracking app that allows the user to record multiple indicators (e.g., BBT, cervical fluid, or OPK testing)—was associated with a small increase in fecundability (shortening of TTP) among selected subgroups. Our study was not able to identify which FF features were most important, if any, in promoting fecundability. Results based on the per-protocol analyses are prone to unmeasured confounding and other sources of bias. Nevertheless, if the per-protocol results are causal, they might extend to a broader population of pregnancy planners, which is the population most likely to use a fertility-tracking app.

Acknowledgements: We acknowledge the contributions of PRESTO participants and staff. We thank Mr. Michael Bairos for technical support in developing the study's web-based infrastructure.

Authors' roles

LAW is the PI for the PRESTO study, obtained funding for this work, and took the lead in drafting the manuscript. LAW and JBS designed the questions related to app use and fertility indicator use within PRESTO. TRW performed data management, quality control, and statistical analysis. EJM directed the per-protocol and intent-to-treat analyses, as well as the development of inverse probability weights. All authors made key contributions to the conduct of the analysis and the interpretation of results. All authors revised the manuscript critically for intellectual content. All authors made meaningful contributions to the manuscript in accordance with the ICMJE guidelines.

References

1. Worsfold L, Marriott L, Johnson S, Harper JC. Period tracker applications: What menstrual cycle information are they giving women? Womens Health (Lond). 2021;17:1-8.

2. Schantz JS, Fernandez CSP, Anne Marie ZJ. Menstrual Cycle Tracking Applications and the Potential for Epidemiological Research: A Comprehensive Review of the Literature. Curr Epidemiol Rep. 2021;8(1):9-19.

3. https://www.consumerreports.org/health-privacy/what-your-period-tracker-app-knows-about-you-a8701683935/.

4. Favaro C, Pearson JT, Rowland SP, Jukic AM, Chelstowska M, Berglund Scherwitzl E, et al. Time to Pregnancy for Women Using a Fertility Awareness Based Mobile Application to Plan a Pregnancy. J Womens Health (Larchmt). 2021;30(11):1538-45.

5. Earle S, Marston HR, Hadley R, Banks D. Use of menstruation and fertility app trackers: a scoping review of the evidence. BMJ Sex Reprod Health. 2021;47(2):90-101.

6. Duane M, Stanford JB, Porucznik CA, Vigil P. Fertility Awareness-Based Methods for Women's Health and Family Planning. Front Med (Lausanne). 2022;9:858977.

7. Stanford JB, Willis SK, Hatch EE, Rothman KJ, Wise LA. Fecundability in relation to use of mobile computing apps to track the menstrual cycle. Hum Reprod. 2020;35(10):2245-52.

8. Evans-Hoeker E, Pritchard DA, Long DL, Herring AH, Stanford JB, Steiner AZ. Cervical mucus monitoring prevalence and associated fecundability in women trying to conceive. Fertil Steril. 2013;100(4):1033-8.

9. Mu Q, Fehring RJ. Efficacy of achieving pregnancy with fertility-focused intercourse. MCN The American journal of maternal child nursing. 2014;39(1):35-40.

10. Robinson JE, Wakelin M, Ellis JE. Increased pregnancy rate with use of the Clearblue Easy Fertility Monitor. Fertil Steril. 2007;87(2):329-34.
11. Tiplady S, Jones G, Campbell M, Johnson S, Ledger W. Home ovulation tests and stress in women trying to conceive: a randomized controlled trial. Hum Reprod. 2013;28(1):138-51.
12. Johnson S, Stanford JB, Warren G, Bond S, Bench-Capon S, Zinaman MJ. Increased Likelihood of Pregnancy Using an App-Connected Ovulation Test System: A Randomized Controlled Trial [Advance publication]. J Womens Health (Larchmt). 2019:1-7.
13. Stanford JB, Willis SK, Hatch EE, Rothman KJ, Wise LA. Fecundability in relation to use of fertility awareness indicators in a North American preconception cohort study. Fertil Steril. 2019;112(5):892-9.
14. Stanford JB, Smith KR, Varner MW. Impact of instruction in the Creighton model fertilitycare system on time to pregnancy in couples of proven fecundity: results of a randomised trial. Paediatr Perinat Epidemiol. 2014;28(5):391-9.
15. Johnson S, Stanford JB, Warren G, Bond S, Bench-Capon S, Zinaman MJ. Increased Likelihood of Pregnancy Using an App-Connected Ovulation Test System: A Randomized Controlled Trial. J Womens Health (Larchmt). 2020;29(1):84-90.
16. Wise LA, Rothman KJ, Mikkelsen EM, Stanford JB, Wesselink AK, McKinnon C, et al. Design and Conduct of an Internet-Based Preconception Cohort Study in North America: Pregnancy Study Online. Paediatr Perinat Epidemiol. 2015;29(4):360-71.
17. Faust L, Bradley D, Landau E, Noddin K, Farland LV, Baron A, et al. Findings from a mobile application-based cohort are consistent with established knowledge of the menstrual cycle, fertile window, and conception. Fertil Steril. 2019;112(3):450-7 e3.

18. Matsumoto M, Nishimura T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Trans Model Comput Simul. 1998;8:3-30.

19. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc. 2000;32(9 Suppl):S498-504.

20. Olsen LR, Jensen DV, Noerholm V, Martiny K, Bech P. The internal and external validity of the Major Depression Inventory in measuring severity of depressive states. Psychol Med. 2003;33:351-6.

21. Cohen S, Williamson G. Perceived stress in a probability sample of the United States In: Spacapan S, Oskamp S, editors. The social psychology of health: Claremont Symposium on applied social psychology. Newbury Park, CA: Sage; 1988.

22. Baird DD, Wilcox AJ, Weinberg CR. Use of time to pregnancy to study environmental exposures. Am J Epidemiol. 1986;124(3):470-80.

23. Cox DR. Regression models and life-tables. R Stat Soc J (B). 1972;34:187-220.

24. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York, NY (USA): Springer-Verlag; 2000.

25. Weinberg CR, Wilcox AJ, Baird DD. Reduced fecundability in women with prenatal exposure to cigarette smoking. Am J Epidemiol. 1989;129(5):1072-8.

26. Zhou XH, Eckert GJ, Tierney WM. Multiple imputation in public health research. Stat Med. 2001;20(9-10):1541-9.

27. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.

554 28. SAS. SAS Institute Inc. 2014. SAS/STAT® 9.4 User's Guide. Cary, NC. Cary, NC: SAS
555 Institute; 2014.
556 29. Symul L, Wac K, Hillard P, Salathe M. Assessment of Menstrual Health Status and
557 Evolution through Mobile Apps for Fertility Awareness. NPJ Digit Med. 2019;2(64):1-10.
558 30. Murray EJ, Swanson SA, Hernán MA. Guidelines for estimating causal effects in pragmatic
559 randomized trials (pre-print). arXiv. 2019.
560 31. Murray EJ, Caniglia EC, Swanson SA, Hernandez-Diaz S, Hernan MA. Patients and
561 investigators prefer measures of absolute risk in subgroups for pragmatic randomized trials. J
562 Clin Epidemiol. 2018;103:10-21.
563 32. Bigelow JL, Dunson DB, Stanford JB, Ecochard R, Gnoth C, Colombo B. Mucus
564 observations in the fertile window: a better predictor of conception than timing of intercourse.
565 Hum Reprod. 2004;19(4):889-92.
566 33. Manders M, McLindon L, Schulze B, Beckmann MM, Kremer JA, Farquhar C. Timed
567 intercourse for couples trying to conceive. Cochrane Database Syst Rev. 2015(3):CD011345.
568 34. Lundsberg LS, Pal L, Gariepy AM, Xu X, Chu MC, Illuzzi JL. Knowledge, attitudes, and
569 practices regarding conception and fertility: a population-based survey among reproductive-age
570 United States women. Fertil Steril. 2014;101(3):767-74.
571 35. Hassan MA, Killick SR. Is previous use of hormonal contraception associated with a
572 detrimental effect on subsequent fecundity? Hum Reprod. 2004;19(2):344-51.
573 36. Yland JJ, Bresnick KA, Hatch EE, Wesselink AK, Mikkelsen EM, Rothman KJ, et al.
574 Pregravid contraceptive use and fecundability: prospective cohort study. BMJ. 2020;371:m3966.
575
576

577 Table 1. Baseline characteristics of participants by randomization status, 2013-2019, PRESTO

	Randomized to FF	
Characteristic[a]	Yes (N=2,775)	No (N=2,767)
Age at baseline, years (mean)	29.8	29.9
Race/ethnicity (%)		
Non-Hispanic, White	82.6	82.6
Hispanic	6.5	7.6
Non-Hispanic mixed/other race	5.0	4.5
Non-Hispanic Black	3.3	3.6
Non-Hispanic Asian	2.6	1.8
Education (%)		
≤High school	6.3	5.9
Some college	24.0	24.8
College degree	34.3	33.0
Graduate school	35.3	36.4
Annual household income (%)		
<$50,000	23.1	23.5
$50,000-$99,999	38.0	39.0
$100,000-$149,999	23.6	23.8
≥$150,000	15.4	13.7
Body mass index, kg/m^2 (%)		
<25	43.2	42.5
25-29	24.2	25.1
30-34	13.9	13.2
≥35	18.7	19.3
Last method of contraception - hormonal (%)	39.6	39.7
History of infertility (%)	9.1	9.7
History of miscarriage (%)	24.4	24.3
Sleep duration, hours (%)		
<7	26.9	27.6
7-8	66.5	66.0
≥9	6.6	6.4
Major Depression Inventory score (%)		
<20	82.9	82.1
20-24	7.1	7.7
25-29	4.8	4.5
≥30	5.3	5.7
History of sexually transmitted infections (%)	14.4	16.4
History of polycystic ovary syndrome (%)	9.1	9.2
History of uterine leiomyomata (%)	2.2	2.2
History of endometriosis (%)	3.8	2.5
Parous (%)	32.0	31.4
Not married (%)	10.4	10.9
Geographic region of residence (%)		
U.S. Northeast	21.9	22.8
U.S. South	23.6	25.1
U.S. Midwest	21.2	21.0
U.S. West	16.0	15.2
Canada	17.3	15.8
Unemployed (%)	3.9	4.6

Hours/week of work (%)		
<30	25.1	27.2
30-49	64.9	63.8
≥50	10.0	9.0
Had Pap test in past 3 years (%)	91.8	92.8
History of abnormal Pap test (%)	33.0	32.6
Current smoker (%)	11.9	12.3
Daily use of folic acid/multivitamins (%)	75.7	76.6
Intercourse frequency, times/week (mean)	2.3	2.3
Doing something to improve chances (%)	72.7	72.9
Sugar-sweetened soda intake, drinks/wk (%)		
≤1	74.1	73.8
2-6	18.5	18.9
≥7	7.4	7.3
Alcohol intake, drinks/week (%)		
0	31.2	30.1
1-6	54.3	55.1
≥7	14.5	14.8
Perceived stress scale score (%)		
<10	12.0	11.3
10-14	27.6	28.8
15-19	30.8	31.5
20-24	20.3	18.3
≥25	9.3	10.0
Physical activity, hours/week (%)		
<5	13.4	14.7
5-9	20.4	20.9
10-14	33.6	32.8
≥15	32.6	31.6
Current marijuana use (%)	13.7	14.0
Use other fertility tracking app (%)	68.3	69.0
Randomized to receive home pregnancy tests (%)	12.1	12.4
Completed dietary questionnaire (%)	64.2	61.3
Invited male partner (%)	52.3	54.0
Attempt time at study entry, cycles (mean)	2.1	2.1

578 [a]All characteristics except for age are standardized to baseline age of cohort.
579

580 Table 2. Intent-to-treat analysis of randomization to FertilityFriend.com (FF) and fecundability, 2013-
581 2020, PRESTO

	Pregnancies	Cycles	FR (95% CI)[a]
All cycles			
Randomized to FF	1,373	14,033	0.97 (0.90-1.04)
Not randomized to FF	1,357	13,661	Reference
Omit first contributed cycle			
Randomized to FF	1,373	11,258	0.97 (0.90-1.04)
Not randomized to FF	1,357	10,894	Reference
Attempt time at entry <3 cycles			
Randomized to FF	1,038	9,580	0.95 (0.88-1.03)
Not randomized to FF	1,040	9,219	Reference
Attempt time at entry 3-4 cycles			
Randomized to FF	242	3,022	1.00 (0.84-1.18)
Not randomized to FF	231	2,967	Reference
Attempt time at entry 5-6 cycles			
Randomized to FF	93	1,431	1.09 (0.83-1.43)
Not randomized to FF	86	1,475	Reference
Age <25 years			
Randomized to FF	122	1,331	1.18 (0.91-1.53)
Not randomized to FF	92	1,188	Reference
Age 25-34 years			
Randomized to FF	1,105	10,597	0.95 (0.88-1.02)
Not randomized to FF	1,118	10,353	Reference
Age ≥35 years			
Randomized to FF	146	2,105	0.99 (0.79-1.22)
Not randomized to FF	147	2,120	Reference
Parous			
Randomized to FF	434	4,245	0.98 (0.86-1.11)
Not randomized to FF	411	4,026	Reference
Nulliparous			
Randomized to FF	939	9,788	0.96 (0.89-1.05)
Not randomized to FF	946	9,635	Reference
No history of infertility			
Randomized to FF	1,314	12,784	0.97 (0.91-1.05)
Not randomized to FF	1,291	12,443	Reference
History of infertility			
Randomized to FF	59	1,249	0.87 (0.62-1.21)
Not randomized to FF	66	1,218	Reference

582

(Table 2 continued...)

	Pregnancies	Cycles	FR (95% CI)
Regular menstrual cycles			
Randomized to FF	1,181	11,477	0.97 (0.90-1.05)
Not randomized to FF	1,171	11,235	Reference
Irregular menstrual cycles			
Randomized to FF	192	2,556	0.97 (0.80-1.18)
Not randomized to FF	186	2,426	Reference
Education ≤12 years			
Randomized to FF	54	773	0.99 (0.67-1.45)
Not randomized to FF	42	627	Reference
Education 13-15 years			
Randomized to FF	257	3,314	0.95 (0.81-1.12)
Not randomized to FF	259	3,257	Reference
Education ≥16 years			
Randomized to FF	1,062	9,946	0.98 (0.90-1.06)
Not randomized to FF	1,056	9,777	Reference
Non-use of fertility app at baseline			
Randomized to FF	413	4,800	0.89 (0.78-1.01)
Not randomized to FF	424	4,413	Reference
Non-user of hormonal contraception within 3 months before baseline			
Randomized to FF	1,054	11,117	0.99 (0.91-1.07)
Not randomized to FF	1,028	10,901	Reference

[a] All models included cycle-specific indicator variables for "menstrual cycle at risk" to adjust for the decline in fecundability in the study population over time.

587 Table 3. Per-protocol analysis of randomization to and actual use of FertilityFriend.com (FF)
588 and fecundability, 2013-2020, PRESTO

	Pregnancies	Cycles	FR (95% CI)[a]
All cycles			
Randomized to FF	1,059	9,567	1.06 (0.99-1.14)
Not randomized to FF	1,301	13,199	Reference
Omits first contributed cycle			
Randomized to FF	895	7,571	0.99 (0.92-1.06)
Not randomized to FF	1,301	10,433	Reference
Attempt time at entry <3 cycles			
Randomized to FF	800	6,664	1.00 (0.92-1.08)
Not randomized to FF	997	8,864	Reference
Attempt time at entry 3-4 cycles			
Randomized to FF	190	1,952	1.15 (0.98-1.35)
Not randomized to FF	220	2,881	Reference
Attempt time at entry 5-6 cycles			
Randomized to FF	69	951	1.14 (0.87-1.48)
Not randomized to FF	84	1,454	Reference
Age <25 years			
Randomized to FF	95	905	1.29 (1.01-1.66)
Not randomized to FF	92	1,182	Reference
Age 25-34 years			
Randomized to FF	858	7,326	1.04 (0.96-1.12)
Not randomized to FF	1,068	9,957	Reference
Age ≥35 years			
Randomized to FF	106	1,336	1.10 (0.89-1.36)
Not randomized to FF	141	2,060	Reference
Parous			
Randomized to FF	316	2,731	1.06 (0.94-1.20)
Not randomized to FF	398	3,941	Reference
Nulliparous			
Randomized to FF	743	6,836	1.06 (0.98-1.15)
Not randomized to FF	903	9,258	Reference
No history of infertility			
Randomized to FF	1,012	8,749	1.07 (0.99-1.14)
Not randomized to FF	1,237	11,999	Reference
History of infertility			
Randomized to FF	47	818	0.99 (0.72-1.37)
Not randomized to FF	64	1,200	Reference

589

(Table 3 continued...)

	Pregnancies	Cycles	FR (95% CI)
Regular menstrual cycles			
Randomized to FF	905	7,718	1.07 (0.99-1.15)
Not randomized to FF	1,121	10,824	Reference
Irregular menstrual cycles			
Randomized to FF	154	1,849	1.06 (0.88-1.28)
Not randomized to FF	180	2,375	Reference
Education ≤12 years			
Randomized to FF	42	437	1.32 (0.92-1.89)
Not randomized to FF	41	626	Reference
Education 13-15 years			
Randomized to FF	190	2,224	1.01 (0.86-1.19)
Not randomized to FF	252	3,205	Reference
Education ≥16 years			
Randomized to FF	827	6,906	1.07 (0.99-1.15)
Not randomized to FF	1,008	9,368	Reference
Non-use of fertility app at baseline			
Randomized to FF	313	3,259	0.94 (0.83-1.06)
Not randomized to FF	411	4,282	Reference
Non-user of hormonal contraception within 3 months before baseline			
Randomized to FF	809	7,467	1.10 (1.02-1.19)
Not randomized to FF	984	10,517	Reference

[a] All models included cycle-specific indicator variables for "menstrual cycle at risk" to adjust for the decline in fecundability in the study population over time.

Figure 1. Use of app features among those randomized to FF who entered data into app, 2013-2020, PRESTO

[SEE ATTACHED HIGH RESOLUTION TIF DOCUMENT]

Abbreviations: FF=FertilityFriend.com, OPK=ovulation predictor kit (urine LH), BBT=basal body temperature, Cervix Other=position/openness/texture.

Notes: Dark shaded bar displays percentage of participants randomized to FF who entered data into FF app at any point during follow-up (unit of analysis: participant). Light shaded bar displays the percentage of menstrual cycles during which participants randomized to FF entered data into FF app (unit of analysis: menstrual cycle).

Supplemental Table 1. Predictors of loss to follow-up, adjusted for all other variables, 2013-2020, PRESTO

	# of participants	Odds ratio (95% CI)[a]
Randomized to FF		
No	2,767	Reference
Yes	2,775	0.86 (0.75, 0.98)
Age at baseline (years, continuous)	5,542	1.02 (1.00, 1.04)
Race/ethnicity		
Non-Hispanic, White	4,580	Reference
Hispanic	387	1.18 (0.93, 1.49)
Non-Hispanic mixed/other race	263	1.40 (1.05, 1.85)
Non-Hispanic Black	191	1.13 (0.84, 1.52)
Non-Hispanic Asian	121	1.81 (1.18, 2.77)
Education		
≤High school	338	1.83 (1.38, 2.44)
Some college	1,350	1.48 (1.19, 1.83)
College degree	1,864	1.18 (0.97, 1.44)
Graduate school	1,990	Reference
Annual household income		
<$50,000	1,290	1.45 (1.07, 1.97)
$50,000-$99,999	2,132	1.09 (0.83, 1.44)
$100,000-$149,999	1,314	1.22 (0.92, 1.63)
≥$150,000	806	Reference
Body mass index (kg/m^2)		
<25	2,376	Reference
25-29	1,365	1.15 (0.95, 1.38)
30-34	749	1.17 (0.95, 1.45)
≥35	1,052	1.69 (1.40, 2.04)
Last method of contraception		
Hormonal	2,198	0.91 (0.79, 1.04)
Non-hormonal	3,344	Reference
History of infertility		
No	5,021	Reference
Yes	521	1.64 (1.36, 1.98)
History of miscarriage		
No	4,194	Reference
Yes	1,348	1.17 (1.01, 1.36)
Sleep duration (hours)		
<7	1,509	1.05 (0.90, 1.22)
7-8	3,672	Reference
≥9	361	0.80 (0.60, 1.08)
Major Depression Inventory score		
<20	4,574	Reference
20-24	408	1.03 (0.80, 1.31)
25-29	259	1.00 (0.74, 1.34)
≥30	301	0.98 (0.75, 1.29)
History of sexually transmitted infections		
No	4,691	Reference
Yes	851	1.06 (0.89, 1.27)
History of polycystic ovary syndrome		
No	5,035	Reference
Yes	507	1.28 (1.05, 1.57)

History of uterine leiomyomata		
No	5,422	Reference
Yes	120	1.03 (0.65, 1.62)
History of endometriosis		
No	5,370	Reference
Yes	172	1.10 (0.78, 1.55)
Parity		
Nulliparous	3,785	Reference
Parous	1,757	0.94 (0.80, 1.10)
Marital status		
Not married	588	1.58 (1.32, 1.90)
Married	4,954	Reference
Geographic region of residence		
U.S. Northeast	1,240	Reference
U.S. South	1,351	1.60 (1.28, 1.98)
U.S. Midwest	1,170	1.55 (1.23, 1.94)
U.S. West	863	1.99 (1.57, 2.53)
Canada	915	1.90 (1.50, 2.42)
Employment status		
Unemployed	237	1.19 (0.89, 1.59)
Employed	5,305	Reference
Hours/week of work		
<30	1,449	1.03 (0.86, 1.22)
30-49	3,568	Reference
≥50	525	0.83 (0.63, 1.09)
Had Papanicolaou test in past three years		
No	429	Reference
Yes	5,113	0.99 (0.79, 1.25)
History of abnormal Papanicolaou test		
No	3,725	Reference
Yes	1,817	0.90 (0.77, 1.04)
Current smoker[b]		
No	4,873	Reference
Yes	669	0.98 (0.81, 1.18)
Daily use of folic acid, prenatal vitamins, or multivitamins[b]		
No	1,321	Reference
Yes	4,221	0.71 (0.61, 0.82)
Intercourse frequency (times/week)[b]	5,542	0.97 (0.94, 1.01)
Doing something to improve chances[b]		
No	1,510	Reference
Yes	4,032	1.10 (0.95, 1.29)
Sugar-sweetened soda intake (drinks/week)[b]		
≤1	4,100	Reference
2-6	1,035	1.05 (0.89, 1.23)
≥7	407	1.08 (0.85, 1.36)
Alcohol intake (drinks/week)[b]		
0	1,697	Reference
1-6	3,032	1.00 (0.86, 1.17)
≥7	813	0.81 (0.64, 1.03)
Perceived stress scale score[b]		

<10	647	Reference
10-14	1,566	1.06 (0.82, 1.38)
15-19	1,726	1.09 (0.85, 1.41)
20-24	1,070	1.07 (0.81, 1.41)
≥25	533	1.09 (0.78, 1.51)
Physical activity (MET hours/week)[b]		
<10	777	Reference
10-19	1,147	1.02 (0.82, 1.26)
20-39	1,839	0.97 (0.80, 1.19)
≥40	1,779	0.95 (0.77, 1.17)
Current marijuana use[b]		
No	4,775	Reference
Yes	767	1.12 (0.93, 1.35)
Randomized to receive home pregnancy tests		
No	4,862	Reference
Yes	680	1.46 (1.18, 1.79)
Completed dietary questionnaire		
No	2,064	Reference
Yes	3,478	0.19 (0.16, 0.22)
Invited male partner		
No	2,596	Reference
Yes	2,946	0.78 (0.68, 0.89)
Attempt time at study entry	5,542	1.42 (1.35, 1.49)
Time-varying cycle of attempt time	5,542	0.72 (0.70, 0.75)

Abbreviations: FF=FertilityFriend.com, MET=metabolic equivalent of task.

[a] Odds ratio (OR) and 95% confidence interval (CI) for probability of being lost to follow-up. C-statistic: 0.81.

[b] Time-varying variables.

608 Supplemental Table 2. Predictors of adherence in intervention arm and non-adherence in control arm, adjusted for all other variables, 2013-2020,
609 PRESTO

	Intervention arm[a]		Control arm[a]	
	# of women	Odds ratio (95% CI)[b]	# of women	Odds ratio (95% CI)[c]
Age at baseline (years, continuous)	2,775	0.98 (0.97, 0.99)	2,767	0.96 (0.93, 1.00)
Race/ethnicity				
Non-Hispanic, White	2,293	Reference	2,287	Reference
Hispanic	179	1.19 (1.01, 1.39)	208	Not estimable
Non-Hispanic mixed/other race	140	1.07 (0.90, 1.27)	123	Not estimable
Non-Hispanic Black	92	1.44 (1.13, 1.82)	99	Not estimable
Non-Hispanic Asian	71	1.07 (0.84, 1.35)	50	Not estimable
Education				
≤High school	177	0.52 (0.43, 0.63)	161	0.15 (0.02, 1.14)
Some college	668	0.90 (0.80, 1.01)	682	0.65 (0.44, 0.98)
College degree	952	1.13 (1.02, 1.24)	912	0.77 (0.59, 1.01)
Graduate school	978	Reference	1,012	Reference
Annual household income				
<$50,000	644	1.18 (1.00, 1.38)	646	0.27 (0.15, 0.46)
$50,000-$99,999	1,052	1.07 (0.94, 1.21)	1,080	0.68 (0.49, 0.95)
$100,000-$149,999	654	0.98 (0.87, 1.12)	660	0.48 (0.34, 0.68)
≥$150,000	425	Reference	381	Reference
Body mass index (kg/m^2)				
<25	1,200	Reference	1,176	Reference
25-29	671	1.20 (1.09, 1.32)	694	0.57 (0.42, 0.77)
30-34	385	1.11 (0.98, 1.25)	364	0.74 (0.50, 1.08)
≥35	519	1.19 (1.06, 1.34)	533	0.81 (0.54, 1.22)
Last method of contraception				
Hormonal	1,101	1.04 (0.96, 1.13)	1,097	0.59 (0.46, 0.75)
Non-hormonal	1,674	Reference	1,670	Reference
History of infertility				
No	2,522	Reference	2,499	Reference
Yes	253	0.89 (0.77, 1.03)	268	0.82 (0.46, 1.45)
History of miscarriage				
No	2,098	Reference	2,096	Reference
Yes	677	1.06 (0.97, 1.17)	671	0.46 (0.32, 0.67)

610

Sleep duration (hours)				
<7	745	0.90 (0.82, 0.99)	764	1.43 (1.06, 1.91)
7-8	1,846	Reference	1,826	Reference
≥9	184	1.14 (0.97, 1.33)	177	0.95 (0.56, 1.61)
Major Depression Inventory score				
<20	2,299	Reference	2,275	Reference
20-24	196	0.76 (0.65, 0.89)	212	0.74 (0.42, 1.28)
25-29	134	0.89 (0.73, 1.08)	125	0.87 (0.41, 1.83)
≥30	146	0.59 (0.49, 0.72)	155	0.34 (0.14, 0.87)
History of sexually transmitted infections				
No	2,377	Reference	2,314	Reference
Yes	398	1.00 (0.90, 1.12)	453	1.46 (1.04, 2.03)
History of polycystic ovary syndrome				
No	2,522	Reference	2,513	Reference
Yes	253	1.40 (1.20, 1.62)	254	1.09 (0.71, 1.67)
History of uterine leiomyomata				
No	2,715	Reference	2,707	Reference
Yes	60	0.98 (0.76, 1.25)	60	Not estimable
History of endometriosis				
No	2,671	Reference	2,699	Reference
Yes	104	1.28 (1.04, 1.57)	68	0.18 (0.06, 0.52)
Parity				
Nulliparous	1,888	Reference	1,897	Reference
Parous	887	0.85 (0.77, 0.93)	870	0.55 (0.40, 0.75)
Marital status				
Not married	2,488	1.23 (1.07, 1.41)	2,466	1.02 (0.62, 1.67)
Married	287	Reference	301	Reference
Geographic region of residence				
U.S. Northeast	608	Reference	632	Reference
U.S. South	655	1.07 (0.96, 1.20)	696	0.91 (0.66, 1.27)
U.S. Midwest	588	1.01 (0.90, 1.13)	582	1.12 (0.79, 1.59)
U.S. West	443	1.10 (0.97, 1.25)	420	1.48 (1.06, 2.08)
Canada	480	0.95 (0.84, 1.08)	435	0.47 (0.32, 0.70)
Employment status				
Unemployed	2,665	1.88 (1.47, 2.41)	127	1.63 (0.83, 3.20)

Employed	110	Reference	2,640	Reference
Hours/week of work				
<30	699	1.11 (1.00, 1.25)	750	0.93 (0.64, 1.35)
30-49	1,800	Reference	1,768	Reference
≥50	276	1.18 (1.03, 1.35)	249	0.70 (0.46, 1.08)
Had Papanicolaou test in past three years				
No	230	Reference	199	Reference
Yes	2,545	1.17 (1.02, 1.35)	2,568	1.14 (0.68, 1.91)
History of abnormal Papanicolaou test				
No	1,862	Reference	1,863	Reference
Yes	913	0.91 (0.84, 0.99)	904	1.22 (0.94, 1.58)
Current smoker [d]				
No	2,445	Reference	2,428	Reference
Yes	330	1.13 (0.99, 1.28)	339	0.41 (0.21, 0.79)
Daily use of folic acid, prenatal vitamins, or multivitamins [d]				
No	676	Reference	645	Reference
Yes	2,099	1.05 (0.95, 1.15)	2,122	2.63 (1.77, 3.91)
Intercourse frequency (times/week) [d]	2,775	1.02 (1.00, 1.04)	2,767	1.11 (1.03, 1.19)
Doing something to improve chances [d]				
No	719	Reference	751	Reference
Yes	2,056	1.32 (1.20, 1.44)	2,016	10.90 (6.91, 17.18)
Sugar-sweetened soda intake (drinks/week) [d]				
≤1	2,056	Reference	2,044	Reference
2-6	513	0.90 (0.81, 1.00)	522	1.16 (0.82, 1.64)
≥7	206	0.86 (0.74, 1.01)	201	3.13 (1.86, 5.27)
Alcohol intake (drinks/week) [d]				
0	865	Reference	832	Reference
1-6	1,507	1.03 (0.94, 1.12)	1,525	1.04 (0.79, 1.38)
≥7	403	0.94 (0.83, 1.06)	410	1.23 (0.83, 1.83)
Perceived stress scale score [d]				
<10	333	Reference	314	Reference
10-14	766	1.13 (1.00, 1.28)	800	0.89 (0.64, 1.24)
15-19	855	1.28 (1.13, 1.45)	871	0.98 (0.69, 1.38)

20-24	564	1.29 (1.13, 1.49)	506	1.13 (0.76, 1.69)
≥25	257	1.17 (0.98, 1.40)	276	0.79 (0.43, 1.44)
Physical activity (MET hours/week) [d]				
<10	371	Reference	406	Reference
10-19	568	1.23 (1.08, 1.41)	579	0.72 (0.49, 1.06)
20-39	931	1.10 (0.97, 1.25)	908	0.72 (0.51, 1.03)
≥40	905	1.06 (0.93, 1.20)	874	0.46 (0.31, 0.68)
Current marijuana use [d]				
No	2,395	Reference	2,380	Reference
Yes	380	1.13 (1.00, 1.27)	387	0.61 (0.39, 0.97)
Use other fertility app				
No	881	Reference	857	Reference
Yes	1,894	0.73 (0.68, 0.80)	1,910	0.04 (0.03, 0.05)
Randomized to receive home pregnancy tests				
No	2,440	Reference	2,422	Reference
Yes	335	1.02 (0.90, 1.15)	345	1.07 (0.74, 1.55)
Completed dietary questionnaire				
No	995	Reference	1,069	Reference
Yes	1,780	1.42 (1.31, 1.54)	1,698	5.76 (4.07, 8.15)
Invited male partner				
No	1,323	Reference	1,273	Reference
Yes	1,452	1.09 (1.01, 1.18)	1,494	0.84 (0.66, 1.08)
Attempt time at study entry	2,775	0.82 (0.80, 0.84)	2,767	0.71 (0.65, 0.77)
Time-varying cycle of attempt time	2,775	1.21 (1.19, 1.23)	2,767	1.38 (1.33, 1.44)

Abbreviations: FF=FertilityFriend.com, MET=metabolic equivalent of task.

[a] Adherence in intervention arm is defined as evidence of FF app use or reporting of FF use on a questionnaire. Non-adherence in control arm is defined as reporting of FF use on a questionnaire. Data in this table thus reflect predictors of FF app use for both trial arms.

[b] Odds ratio (OR) and 95% confidence interval (CI) for probability of using FF app in a given menstrual cycle. C-statistic: 0.68.

[c] Odds ratio (OR) and 95% confidence interval (CI) for probability of using FF app in a given menstrual cycle. C-statistic: 0.93.

[d] Time-varying variables.

Supplemental Table 3. Intent-to-treat analysis of randomization to FF and fecundability without applying weights for loss to follow-up, 2013-2020, PRESTO

	Pregnancies	Cycles	FR (95% CI)[a]
All cycles			
Randomized to FF	1,373	14,033	0.97 (0.91-1.04)
Not randomized to FF	1,357	13,661	Reference
Omit first contributed cycle			
Randomized to FF	1,373	11,258	0.97 (0.91-1.04)
Not randomized to FF	1,357	10,894	Reference
Attempt time at entry <3 cycles			
Randomized to FF	1,038	9,580	0.95 (0.88-1.03)
Not randomized to FF	1,040	9,219	Reference
Attempt time at entry 3-4 cycles			
Randomized to FF	242	3,022	1.00 (0.85-1.18)
Not randomized to FF	231	2,967	Reference
Attempt time at entry 5-6 cycles			
Randomized to FF	93	1,431	1.09 (0.83-1.44)
Not randomized to FF	86	1,475	Reference
Age <25 years			
Randomized to FF	122	1,331	1.18 (0.91-1.53)
Not randomized to FF	92	1,188	Reference
Age 25-34 years			
Randomized to FF	1,105	10,597	0.95 (0.88-1.03)
Not randomized to FF	1,118	10,353	Reference
Age ≥35 years			
Randomized to FF	146	2,105	0.99 (0.80-1.23)
Not randomized to FF	147	2,120	Reference
Parous			
Randomized to FF	434	4,245	0.98 (0.87-1.12)
Not randomized to FF	411	4,026	Reference
Nulliparous			
Randomized to FF	939	9,788	0.97 (0.89-1.05)
Not randomized to FF	946	9,635	Reference
No history of infertility			
Randomized to FF	1,314	12,784	0.98 (0.91-1.05)
Not randomized to FF	1,291	12,443	Reference
History of infertility			
Randomized to FF	59	1,249	0.87 (0.63-1.22)
Not randomized to FF	66	1,218	Reference

(Supplemental Table 3 continued...)

	Pregnancies	Cycles	FR (95% CI)
Regular menstrual cycles			
Randomized to FF	1,181	11,477	0.97 (0.90-1.05)
Not randomized to FF	1,171	11,235	Reference
Irregular menstrual cycles			
Randomized to FF	192	2,556	0.97 (0.80-1.18)
Not randomized to FF	186	2,426	Reference
Education ≤12 years			
Randomized to FF	54	773	0.98 (0.66-1.43)
Not randomized to FF	42	627	Reference
Education 13-15 years			
Randomized to FF	257	3,314	0.96 (0.82-1.13)
Not randomized to FF	259	3,257	Reference
Education ≥16 years			
Randomized to FF	1,062	9,946	0.98 (0.90-1.06)
Not randomized to FF	1,056	9,777	Reference
Non-use of fertility app at baseline			
Randomized to FF	413	4,800	0.89 (0.78-1.01)
Not randomized to FF	424	4,413	Reference
Non-user of hormonal contraception within 3 months before baseline			
Randomized to FF	1,054	11,117	0.99 (0.91-1.07)
Not randomized to FF	1,028	10,901	Reference

[a] All models included cycle-specific indicator variables for "menstrual cycle at risk" to adjust for the decline in fecundability in the study population over time.

Supplemental Table 4. Per-protocol analysis of randomization to and actual use of FF and fecundability without applying weights for loss to follow-up, 2013-2020, PRESTO

	Pregnancies	Cycles	FR (95% CI)[a]
All cycles			
Randomized to FF	1,059	9,567	1.06 (0.99-1.14)
Not randomized to FF	1,301	13,199	Reference
Omits first contributed cycle			
Randomized to FF	895	7,571	0.99 (0.92-1.07)
Not randomized to FF	1,301	10,433	Reference
Attempt time at entry <3 cycles			
Randomized to FF	800	6,664	1.00 (0.93-1.08)
Not randomized to FF	997	8,864	Reference
Attempt time at entry 3-4 cycles			
Randomized to FF	190	1,952	1.15 (0.97-1.35)
Not randomized to FF	220	2,881	Reference
Attempt time at entry 5-6 cycles			
Randomized to FF	69	951	1.14 (0.87-1.49)
Not randomized to FF	84	1,454	Reference
Age <25 years			
Randomized to FF	95	905	1.29 (1.00-1.65)
Not randomized to FF	92	1,182	Reference
Age 25-34 years			
Randomized to FF	858	7,326	1.04 (0.96-1.12)
Not randomized to FF	1,068	9,957	Reference
Age ≥35 years			
Randomized to FF	106	1,336	1.11 (0.89-1.37)
Not randomized to FF	141	2,060	Reference
Parous			
Randomized to FF	316	2,731	1.07 (0.94-1.21)
Not randomized to FF	398	3,941	Reference
Nulliparous			
Randomized to FF	743	6,836	1.06 (0.98-1.15)
Not randomized to FF	903	9,258	Reference
No history of infertility			
Randomized to FF	1,012	8,749	1.07 (1.00-1.15)
Not randomized to FF	1,237	11,999	Reference
History of infertility			
Randomized to FF	47	818	0.99 (0.72-1.37)
Not randomized to FF	64	1,200	Reference

(Supplemental Table 4 continued...)

	Pregnancies	Cycles	FR (95% CI)
Regular menstrual cycles			
Randomized to FF	905	7,718	1.07 (0.99-1.15)
Not randomized to FF	1,121	10,824	Reference
Irregular menstrual cycles			
Randomized to FF	154	1,849	1.06 (0.88-1.28)
Not randomized to FF	180	2,375	Reference
Education ≤12 years			
Randomized to FF	42	437	1.29 (0.90-1.84)
Not randomized to FF	41	626	Reference
Education 13-15 years			
Randomized to FF	190	2,224	1.02 (0.87-1.20)
Not randomized to FF	252	3,205	Reference
Education ≥16 years			
Randomized to FF	827	6,906	1.07 (0.99-1.15)
Not randomized to FF	1,008	9,368	Reference
Non-use of fertility app at baseline			
Randomized to FF	313	3,259	0.94 (0.83-1.07)
Not randomized to FF	411	4,282	Reference
Non-user of hormonal contraception within 3 months before baseline			
Randomized to FF	809	7,467	1.10 (1.02-1.19)
Not randomized to FF	984	10,517	Reference

[a] All models include cycle-specific indicator variables for "menstrual cycle at risk" to adjust for the decline in fecundability in the study population over time.

631 Supplemental Table 5. Per-protocol analyses using alternative definitions of adherence, 2013-2020, PRESTO

	Adherence in control arm redefined as not using any app			Adherence in intervention arm redefined as using any app and in control arm as not using any app		
	Pregnancies	Cycles	FR (95% CI)[a]	Pregnancies	Cycles	FR (95% CI)[a]
All cycles						
Randomized to FF	1059	9567	1.11 (1.01-1.23)	1276	12531	1.03 (0.93-1.15)
Not randomized to FF	365	3929	Reference	365	3929	Reference
Omits first contributed cycle						
Randomized to FF	895	7571	1.07 (0.97-1.19)	1233	9989	1.06 (0.95-1.17)
Not randomized to FF	362	3041	Reference	362	3041	Reference
Attempt time at entry <3 cycles						
Randomized to FF	800	6664	1.06 (0.95-1.18)	958	8552	1.01 (0.90-1.13)
Not randomized to FF	305	2905	Reference	305	2905	Reference
Attempt time at entry 3-4 cycles						
Randomized to FF	190	1952	1.20 (0.90-1.59)	231	2657	1.11 (0.83-1.49)
Not randomized to FF	45	630	Reference	45	630	Reference
Attempt time at entry 5-6 cycles						
Randomized to FF	69	951	1.58 (0.96-2.60)	87	1322	1.47 (0.89-2.42)
Not randomized to FF	15	394	Reference	15	394	Reference
Age <25 years						
Randomized to FF	95	905	1.71 (1.10-2.65)	115	1156	1.59 (1.02-2.47)
Not randomized to FF	20	355	Reference	20	355	Reference
Age 25-34 years						
Randomized to FF	858	7326	1.05 (0.94-1.17)	1033	9608	0.97 (0.86-1.09)
Not randomized to FF	298	2838	Reference	298	2838	Reference
Age ≥35 years						
Randomized to FF	106	1336	1.19 (0.88-1.60)	128	1767	1.10 (0.81-1.50)
Not randomized to FF	47	736	Reference	47	736	Reference

632

633 (Supplemental Table 5 continued..)

	Adherence in control arm redefined as not using any app			Adherence in intervention arm redefined as using any app and in control arm as not using any app		
	Pregnancies	Cycles	FR (95% CI)[a]	Pregnancies	Cycles	FR (95% CI)[a]
Parous						
Randomized to FF	316	2731	1.09 (0.92-1.29)	387	3660	1.00 (0.84-1.18)
Not randomized to FF	140	1431	Reference	140	1431	Reference
Nulliparous						
Randomized to FF	743	6836	1.13 (1.00-1.29)	889	8871	1.06 (0.93-1.20)
Not randomized to FF	225	2498	Reference	225	2498	Reference
No history of infertility						
Randomized to FF	1012	8749	1.13 (1.02-1.26)	1221	11473	1.05 (0.94-1.17)
Not randomized to FF	341	3553	Reference	341	3553	Reference
History of infertility						
Randomized to FF	47	818	0.82 (0.54-1.26)	55	1058	0.77 (0.50-1.20)
Not randomized to FF	24	376	Reference	24	376	Reference
Regular menstrual cycles						
Randomized to FF	905	7718	1.14 (1.02-1.27)	1095	10202	1.05 (0.94-1.18)
Not randomized to FF	306	3169	Reference	306	3169	Reference
Irregular menstrual cycles						
Randomized to FF	154	1849	1.01 (0.77-1.32)	181	2329	0.94 (0.72-1.23)
Not randomized to FF	59	760	Reference	59	760	Reference
Education ≤12 years						
Randomized to FF	42	437	1.38 (0.78-2.45)	50	648	1.09 (0.60-1.98)
Not randomized to FF	12	195	Reference	12	195	Reference
Education 13-15 years						
Randomized to FF	190	2224	1.11 (0.86-1.43)	237	2935	1.04 (0.81-1.35)
Not randomized to FF	64	897	Reference	64	897	Reference

634

(Supplemental Table 5 continued...)

	Adherence in control arm redefined as not using any app			Adherence in intervention arm redefined as using any app and in control arm as not using any app		
	Pregnancies	Cycles	FR (95% CI)[a]	Pregnancies	Cycles	FR (95% CI)[a]
Education ≥16 years						
Randomized to FF	827	6906	1.11 (0.99-1.25)	989	8948	1.04 (0.93-1.17)
Not randomized to FF	289	2837	Reference	289	2837	Reference
Non-use of fertility app at baseline						
Randomized to FF	313	3259	0.97 (0.85-1.10)	322	3345	0.96 (0.84-1.09)
Not randomized to FF	358	3854	Reference	358	3854	Reference
Non-user of hormonal contraception within 3 months before baseline						
Randomized to FF	809	7467	1.16 (1.03-1.31)	985	9946	1.06 (0.94-1.20)
Not randomized to FF	252	2894	Reference	252	2894	Reference

[a] All models included cycle-specific indicator variables for "menstrual cycle at risk" to adjust for the decline in fecundability in the study population over time.

Supplemental Figure 1. Flow chart of exclusions in FF randomization trial, PRESTO (2013-2019). Abbreviations: FF=FertilityFriend.com, LMP=last menstrual period.

[SEE ATTACHED HIGH RESOLUTION TIF DOCUMENT]