human reproduction	ORIGINAL ARTICLE *Infertility*

Fecundability in relation to use of mobile computing apps to track the menstrual cycle

Joseph B. Stanford[1], Sydney K. Willis[2,*], Elizabeth E. Hatch[2], Kenneth J. Rothman[2,3], and Lauren A. Wise[2]

[1]Office of Cooperative Reproductive Health, Division of Public Health, Department of Family and Preventive Medicine, University of Utah, Salt Lake City, UT, USA [2]Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA [3]RTI International, Research Triangle Park, NC 27709, USA

*Correspondence address. Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA.
E-mail: siwillis@bu.edu

STUDY QUESTION: To what extent does the use of mobile computing apps to track the menstrual cycle and the fertile window influence fecundability among women trying to conceive?

SUMMARY ANSWER: After adjusting for potential confounders, use of any of several different apps was associated with increased fecundability ranging from 12% to 20% per cycle of attempt.

WHAT IS KNOWN ALREADY: Many women are using mobile computing apps to track their menstrual cycle and the fertile window, including while trying to conceive.

STUDY DESIGN, SIZE, DURATION: The Pregnancy Study Online (PRESTO) is a North American prospective internet-based cohort of women who are aged 21–45 years, trying to conceive and not using contraception or fertility treatment at baseline.

PARTICIPANTS/MATERIALS, SETTING, METHODS: We restricted the analysis to 8363 women trying to conceive for no more than 6 months at baseline; the women were recruited from June 2013 through May 2019. Women completed questionnaires at baseline and every 2 months for up to 1 year. The main outcome was fecundability, i.e. the per-cycle probability of conception, which we assessed using self-reported data on time to pregnancy (confirmed by positive home pregnancy test) in menstrual cycles. On the baseline and follow-up questionnaires, women reported whether they used mobile computing apps to track their menstrual cycles ('cycle apps') and, if so, which one(s). We estimated fecundability ratios (FRs) for the use of cycle apps, adjusted for female age, race/ethnicity, prior pregnancy, BMI, income, current smoking, education, partner education, caffeine intake, use of hormonal contraceptives as the last method of contraception, hours of sleep per night, cycle regularity, use of prenatal supplements, marital status, intercourse frequency and history of subfertility. We also examined the impact of concurrent use of fertility indicators: basal body temperature, cervical fluid, cervix position and/or urine LH.

MAIN RESULTS AND THE ROLE OF CHANCE: Among 8363 women, 6077 (72.7%) were using one or more cycle apps at baseline. A total of 122 separate apps were reported by women. We designated five of these apps before analysis as more likely to be effective (Clue, Fertility Friend, Glow, Kindara, Ovia; hereafter referred to as 'selected apps'). The use of any app at baseline was associated with 20% increased fecundability, with little difference between selected apps versus other apps (selected apps FR (95% CI): 1.20 (1.13, 1.28); all other apps 1.21 (1.13, 1.30)). In time-varying analyses, cycle app use was associated with 12–15% increased fecundability (selected apps FR (95% CI): 1.12 (1.04, 1.21); all other apps 1.15 (1.07, 1.24)). When apps were used at baseline with one or more fertility indicators, there was higher fecundability than without fertility indicators (selected apps with indicators FR (95% CI): 1.23 (1.14, 1.34) versus without indicators 1.17 (1.05, 1.30); other apps with indicators 1.30 (1.19, 1.43) versus without indicators 1.16 (1.06, 1.27)). In time-varying analyses, results were similar when stratified by time trying at study entry (<3 vs. 3–6 cycles) or cycle regularity. For use of the selected apps, we observed higher fecundability among women with a history of subfertility: FR 1.33 (1.05–1.67).

LIMITATIONS, REASONS FOR CAUTION: Neither regularity nor intensity of app use was ascertained. The prospective time-varying assessment of app use was based on questionnaires completed every 2 months, which would not capture more frequent changes. Intercourse frequency was also reported retrospectively and we do not have data on timing of intercourse relative to the fertile window.

Although we controlled for a wide range of covariates, we cannot exclude the possibility of residual confounding (e.g. choosing to use an app in this observational study may be a marker for unmeasured health habits promoting fecundability). Half of the women in the study received a free premium subscription for one of the apps (Fertility Friend), which may have increased the overall prevalence of app use in the time-varying analyses, but would not affect app use at baseline. Most women in the study were college educated, which may limit application of results to other populations.

WIDER IMPLICATIONS OF THE FINDINGS: Use of a cycle app, especially in combination with observation of one or more fertility indicators (basal body temperature, cervical fluid, cervix position and/or urine LH), may increase fecundability (per-cycle pregnancy probability) by about 12–20% for couples trying to conceive. We did not find consistent evidence of improved fecundability resulting from use of one specific app over another.

Introduction

In recent years, there has been a rapid growth of mobile computing apps for women's health in general, and menstrual cycle health in particular (Bull *et al.*, 2019; Faust *et al.*, 2019; Stanford, 2019; Symul *et al.*, 2019). Many apps purport to identify the day of ovulation and/or the days when intercourse is most likely to result in conception (i.e. the fecund window, usually called the fertile window). Independent evaluations have suggested that many such apps provide inaccurate information about the timing of ovulation and the fertile window (Duane *et al.*, 2016; Freis *et al.*, 2018).

The actual fertile window of the menstrual cycle includes approximately the day of ovulation and the preceding 5 days, with some variability and shorter windows in subfertile women (Dunson *et al.*, 1999; Colombo and Masarotto, 2000; Keulers *et al.*, 2007; Faust *et al.*, 2019). In women with normal length menstrual cycles, the timing of onset of the fertile window in any particular cycle may vary by up to 15 days (Day 6 to 21) (Wilcox *et al.*, 2000). A number of biomarkers or fertility indicators can be observed by women to help identify the onset of the fertile window, either prospectively or retrospectively (American Society for Reproductive Medicine, 2017). The most common algorithms are based on simple counts of previous cycle lengths (Brayer *et al.*, 1969; Arevalo *et al.*, 1999; Li *et al.*, 2016); other algorithms can be based on markers of ovulation in previous cycles (such as temperature shift) (Berglund Scherwitzl *et al.*, 2015) or the current cycle (such as cervical fluid changes and measures of oestrogen metabolites or LH in urine) (Ecochard *et al.*, 2001, 2015; Stanford *et al.*, 2002).

We have previously reported that use of fertility indicators, including tracking cycle days, cervical fluid, cervix position, basal body temperature and urine LH, is associated with increased fecundability (Stanford *et al.*, 2019). This finding is consistent with data from a number of cohort studies and trials, which show that use of one or more fertility indicators to identify the fertile window may shorten time to pregnancy (TTP) (Robinson *et al.*, 2007; Evans-Hoeker *et al.*, 2013; Tiplady *et al.*, 2013; Mu and Fehring, 2014; Johnson *et al.*, 2019). Not all

studies fully support this finding, however (Stanford *et al.*, 2014). Mobile computing apps can manage the collection, recording and interpretation of fertility indicators for identifying the fertile window. However, there is scant published evidence about which, if any, mobile apps are effective in timing intercourse and shortening TTP (Manders *et al.*, 2015).

The aim of this analysis is to measure the effect on fecundability of the use of mobile computing apps that track the menstrual cycle and the fertile window. A secondary aim is to estimate the prevalence of use of these apps among an internet-based volunteer sample of couples trying to conceive in Canada and the USA.

Materials and methods

Study design and population

Pregnancy Study Online (PRESTO) is a prospective web-based preconception cohort study of pregnancy planners, conducted through Boston University. Study methods have been described in detail previously (Wise *et al.*, 2015). PRESTO enrols women and then invites their male partners to enrol as well. Eligibility criteria for women are age 21–45 years, living in the USA or Canada, not pregnant, and not using contraception or fertility treatment at baseline. Female participants complete an online baseline questionnaire with items on demographics, behaviours, reproductive and medical histories and medication use. Females complete follow-up questionnaires every 8 weeks or until reported conception, initiation of fertility treatment, cessation of pregnancy attempt, withdrawal, loss to follow up or 12 months, whichever occurs first. Online informed consent was obtained from all participants and this study was approved by the Institutional Review Board at Boston University Medical Campus.

For this analysis, we excluded women who had missing or implausible data for last menstrual period (LMP) at entry to the study, or whose LMP was more than 6 months before participants' study entry.

We also excluded women had been attempting pregnancy for more than 6 cycles at study entry.

Use of apps

On the baseline and follow-up questionnaire, women were asked, 'Do you use any software program and/or web-based or phone "app" to record your menstrual cycle data and/or fertility signs (e.g. FertilityFriend.com, Taking Charge of Your Fertility (TCOYF))?' If they responded yes, they were asked to write the name of the program or app, with a free-text response. We compiled these responses and content-coded them to generate a list of 122 apps, which we verified by searching online. There were 161 responses that were too non-specific to code (e.g. 'fertility app' or 'other app') or that we could not verify; we coded these as 'other'.

We designated five apps a priori as being more likely to be accurate (Clue, Fertility Friend, Glow, Kindara, Ovia) based on prior published data and their reported recording and use of fertility indicators (basal body temperature, cervical fluid, cervix position and/or urine LH) to identify the fertile window (Duane et al., 2016). We grouped these together as 'selected apps', and all other apps were grouped as 'other apps'. We also analysed individually each app that had at least 100 users (eight apps: Clue, Fertility Friend, Flo, Glow, Kindara, My Days, Ovia, Period Tracker).

Outcome: TTP

To assess fecundability (the per-cycle probability of pregnancy, as identified by a home pregnancy test and reported by the woman on a follow-up questionnaire), we estimated TTP, based on questionnaire responses (Baird et al., 1986; Eijkemans et al., 2019; Wesselink et al., 2020). On the baseline questionnaire, women reported their LMP, usual menstrual cycle length (if regular) and the number of cycles attempting pregnancy at study entry. For women who reported irregular cycles, we estimated cycle length based on their LMP start date at baseline and the consecutive menstruation dates reported at follow up. On the follow-up questionnaires, women reported their LMP and whether they have conceived since the prior questionnaire, and the method of pregnancy confirmation (e.g. home pregnancy test, blood test and ultrasound). In these analyses, we relied on confirmation via a positive home pregnancy test. TTP was estimated based on the total discrete cycles at risk of pregnancy, calculated as follows: cycles of attempt at study entry + [(LMP date from most recent follow-up questionnaire − date of baseline questionnaire completion)/usual cycle length] +1 (Wise et al., 2010).

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, multivitamin use, caffeine intake, smoking status, alcohol use, contraception history, intercourse frequency, use of fertility indicators, menstrual regularity and history of subfertility. Menstrual regularity was defined through asking 'within the past couple of years, has your menstrual period been regular (regular in a way so you can usually predict about when the next period will start? Please think about those times you were not using hormonal contraceptives')', and history of subfertility was defined as having tried to conceive for more than 6 months without success for any prior pregnancy attempt. BMI was calculated as weight (kilograms) divided by height (metres) squared.

Analysis

We calculated descriptive statistics for baseline demographic and health characteristics of participants. We stratified participants by use of 'selected' and 'other' apps or no cycle apps, as reported at baseline.

We calculated proportional probability regression models to estimate fecundability ratios (FRs) and 95% confidence intervals (CI). In primary analyses, we focused on the effect of using any of the a priori selected apps. In additional analyses, we evaluated the effect of using an app that was not part of the a priori group, and in further analyses, we evaluated separately any app with 100 or more users. In stratified analyses, we assessed the effect of selected apps, other apps or no apps, with or without concurrent use of fertility indicators (basal body temperature, cervical fluid, cervix position and/or urine LH). Couples contributed menstrual cycles until pregnancy, initiation of fertility treatment, cessation of pregnancy attempts, withdrawal, loss to follow-up or completion of 12 cycles, whichever came first. We performed analyses based on: (i) apps reported at baseline only and (ii) apps reported on each follow-up questionnaire ('time-varying'). In the time-varying analyses, we replaced the former values of exposure with the updated information reported on each bi-monthly questionnaire.

We estimated the FR, the ratio of the average per-cycle probability of conception within a specific exposure category in comparison with the average per-cycle probability of conception of a designated reference group. FRs of <1 indicate a longer TTP among exposed compared with unexposed participants. The model incorporates each observed cycle at risk, which accounts for the baseline decline in fecundability over time (Weinberg et al., 1989). The Andersen–Gill data structure outputs a single menstrual cycle per observation to accommodate time-varying variables and to account for left truncation from delayed entry into the cohort (Therneau and Grambsch, 2000).

We assessed potential confounders a priori based on a literature review and the consideration of directed acyclic graphs. Final models were adjusted for female age (<25, 25–29, 30–34, ≥35 years), BMI (<18, 18.5–24.9, 25–29.9, ≥30 kg/m^2), smoking status (current vs. non-current), education ($<$ college degree vs. $\geq$ college degree), partner education ($<$ college degree vs. $\geq$ college degree), income ($<\$50\,000$, $\$50–99\,999$, $\$100–149\,999$, $\$\geq150\,000$ US dollars/ year), use of hormonal contraceptives as last method of contraception (yes vs. no), hours of sleep per night (<7, 7–8, ≥9 h), prior pregnancy (yes vs. no), use of prenatal multivitamin supplements (yes vs. no), race/ethnicity (White, non-Hispanic vs. other race/ethnicity), caffeine intake (mg/day), intercourse frequency (<2 vs. ≥2 times per week), marital status (married vs. not married), history of subfertility (yes, no, never attempted to conceive) and cycle regularity (regular vs. irregular). We conducted supplementary exploratory analyses stratified by time trying at study entry, cycle regularity, and history of subfertility (trying for ≥6 months to conceive in previous pregnancy attempt).

Missing data

We imputed missing values for exposures, covariates and pregnancy status using multiple imputation. We created five imputed datasets,

analysing each dataset separately, and combining coefficient and standard error estimates across the imputed datasets (Sterne *et al.*, 2009). To reduce selection bias from differential loss to follow-up, we assigned one cycle of follow-up for the 15% of women with no data from follow-up questionnaires (N = 1285) and then imputed their pregnancy status (yes vs. no) by multiple imputation. Fewer than 0.2% of participants were missing data for use of an app to record menstrual cycle and/or fertility data. Missingness for covariates ranged from <0.1% (prior pregnancy, history of subfertility, caffeine use and history of anxiety) to 3.4% for income. There were no missing values for age.

Results

From June 2013 until May 2019, 10 599 eligible women completed the baseline questionnaire. Of these, we excluded 120 women, because the start date of their LMP at baseline was more than 6 months before the participants entered the study, and 28 women, because of missing LMP data. We excluded an additional 2088 women who had been attempting to achieve pregnancy for more than 6 cycles at baseline. After these exclusions, the final study population for this analysis comprised 8363 women. These women reported 4858 pregnancies during 31 572 cycles of observation.

Demographic, reproductive and behavioural characteristics of the participants at baseline are shown in Table I, stratified by reported use of no cycle app (2286 women; 27%), selected cycle app (3339 women, 40%) or other cycle app (2738 women, 33%). Overall at baseline, 72.7% of women were using a cycle app (95% CI: 71.7%, 73.6%).

The majority of women (over 70%) were married and college educated. Fewer than 20% reported a history of subfertility. Among women using selected or other apps, compared with women not using apps, a higher proportion were taking folic acid (83.3% and 73.7% vs. 66.8%), and a lower proportion had used a hormonal method as their most recent form of contraception (37.4% and 34.6% vs. 44.0%).

Table II reports FRs for each app category, and for each specific app with 100 or more users each. Analyses are reported for baseline app use, and for time-varying use (i.e. use updated every 2 months based on each follow-up questionnaire). At baseline, compared with no app use, each of the apps was associated with higher fecundability (FR 1.13–1.46), with similar results after adjustment for covariates (FR 1.14–1.42). In time-varying analyses, the associations were attenuated

Table I Baseline characteristics of PRESTO based on cycle app use.*

	All women	Non-users of cycle apps	Users of selected cycle apps	Users of other cycle apps
Number of women	8363	2286	3339	2738
Age, years (mean)	29.9	30.2	29.7	30.0
Partners age, years (mean)	31.8	32.0	31.7	31.8
BMI, kg/m^2 (mean)	28.0	28.0	27.6	28.3
Partners BMI, kg/m^2 (mean)	28.2	28.1	28.3	28.3
Relationship duration, 5 years (%)	49.3	48.2	53.1	45.7
Attempt time at study entry, cycles (mean)	2.1	1.8	2.2	2.2
Non-Hispanic White (%)	84.2	80.9	87.0	83.6
Income >50 000 (%)	59.1	62.2	54.8	61.9
Education ≥ college (%) education	72.0	70.2	75.8	69.0
Sleep <7 h (%)	25.4	26.0	23.4	27.3
Prior pregnancy (%)	51.2	53.3	50.8	50.1
Took folic acid/multivitamin (%)	75.6	66.8	83.3	73.7
Prior birth (%)	33.3	37.5	31.3	32.1
Last form of contraception hormonal (%)	38.3	44.0	37.4	34.6
Current smoker (%)	6.9	8.2	5.0	8.2
Intercourse ≥2 times per week (%)	60.5	59.5	60.9	60.8
Partner education ≥ College education (%)	54.4	53.6	58.0	50.5
Married (%)	90.0	88.3	91.1	90.0
Alcohol intake, drinks/week (mean)	3.2	3.1	3.2	3.2
Caffeine intake, mg/day	124.1	120.6	128.2	123.7
History of subfertility (%)**	13.3	15.9	11.5	13.4
Irregular periods (%)	17.3	19.9	15.8	17.0
Cycle length (median days)	29	29	29	29

*All characteristics, with the exception of age, are standardised to age distribution of cohort at baseline.
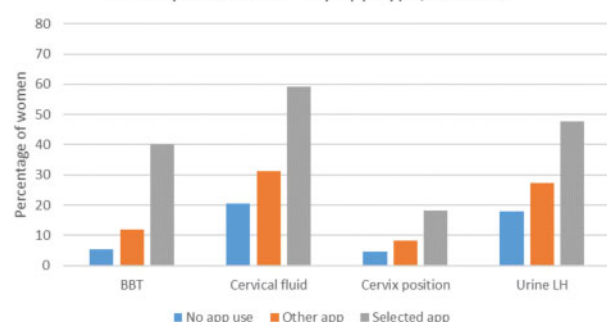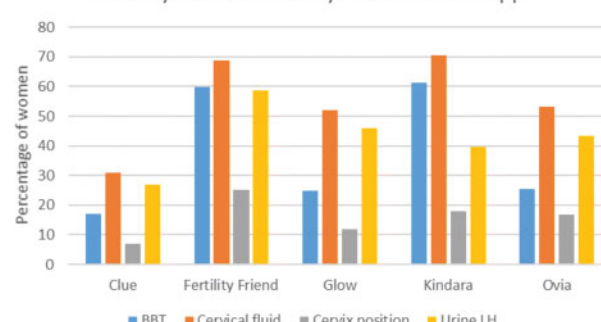**Among individuals who had previously attempted to conceive.

Table II Associations between app utilisation and fecundability.

	Baseline app use						Time-varying app use					
	No. of pregnancies	No. of cycles	Unadjusted		Adjusted[a]		No. of pregnancies	No. of cycles	Unadjusted		Adjusted[a]	
			FR	95% CI	FR	95% CI			FR	95% CI	FR	95% CI
No app	1265	9578	1.00	Ref	1.00	Ref	933	6638	1.00	Ref	1.00	Ref
Selected apps[b]	2028	12124	1.25	1.18–1.34	1.20	1.13–1.28	2289	14705	1.18	1.10–1.27	1.12	1.04-.1.21
Clue	235	1154	1.46	1.29–1.65	1.42	1.25–1.60	231	1092	1.41	1.24–1.60	1.36	1.20–1.55
Fertility friend	703	4378	1.24	1.14–1.34	1.15	1.05–1.26	962	6778	1.14	1.05–1.24	1.07	0.98–1.16
Glow	453	2638	1.28	1.16–1.41	1.28	1.16–1.42	460	2777	1.22	1.10–1.35	1.21	1.09–1.35
Kindara	167	953	1.29	1.12–1.50	1.14	0.98–1.32	175	1013	1.24	1.07–1.44	1.11	0.96–1.29
Ovia	642	3858	1.24	1.13–1.37	1.21	1.10–1.33	666	4138	1.19	1.08–1.31	1.13	1.03–1.25
Other apps[c]	1565	9870	1.22	1.14–1.31	1.21	1.13–1.30	1636	10229	1.16	1.08–1.25	1.15	1.07–1.24
Flo	353	1939	1.32	1.19–1.47	1.33	1.19–1.48	357	1885	1.30	1.16–1.45	1.29	1.15–1.45
My Days	121	804	1.23	1.03–1.46	1.24	1.04–1.48	123	762	1.26	1.05–1.50	1.28	1.08–1.53
Period Tracker	311	2203	1.13	1.01–1.27	1.15	1.02–1.29	300	2132	1.08	0.96–1.22	1.09	0.97–1.24

[a]Adjusted for female age, BMI, smoking status, education, partner education, income, hormonal method for most recent contraception, sleep, prior pregnancy, multivitamin use, race/ethnicity, caffeine intake, intercourse frequency, marital status, prior history of subfertility and cycle regularity.
[b]Women could be in more than one app group.
[c]Other apps includes use of any app(s) that is not one of the selected apps. Those with over 100 users are also analysed individually.

Figure 1. Fertility indicator use by no app, other app or selected app use.

Figure 2. Fertility indicator use for selected apps.

slightly but remained after adjustment. In these analyses, the FR ranged from 1.07 to 1.36.

For each of the fertility indicators (basal body temperature, cervical fluid, cervix position, urinary LH 'ovulation predictor' testing), use was highest among women with the a priori selected apps, followed by the other apps and lowest for women not using any app (Fig. 1). Details of use of fertility indicators for all apps reported by at least 100 women (which included all selected apps) are shown in Figs 2 and 3.

We assessed the interaction between use of fertility indicators (basal body temperature, urine LH, cervical fluid or cervix position) and use of apps in relation to fecundability (Table III). The combined use of an app with one or more fertility indicators at baseline was associated with the highest fecundability; i.e. selected apps and fertility indicators FR (95% CI) 1.23 (1.14, 1.34)); other apps and fertility

indicators 1.30 (1.19, 1.43), as compared with when apps were used without fertility indicators (selected apps FR (95% CI): 1.17 (1.05, 1.30); other apps 1.16 (1.06, 1.27)). Use of fertility indicators without apps was not appreciably associated with fecundability compared with use of neither at baseline: FR 1.03 (0.92–1.16). There was slight attenuation of FRs in the time-varying analyses, but the patterns were the same, e.g. selected apps and fertility indicators FR (95% CI): 1.21 (1.11, 1.31); other apps and fertility indicators 1.26 (1.15, 1.38).

We conducted several stratified analyses, using time-varying data on app use. There was no consistent pattern of effect modification when we stratified by attempt time at cohort entry (<3 cycles vs. 3–6 cycles; Supplementary Table SI) or by cycle regularity (Supplementary Table SII). However, there was a suggestive stronger association among selected apps for women with a history of subfertility

(FR = 1.40, 95% CI: 1.11, 1.77), as compared with women without a history of subfertility (FR = 1.15, 95% CI: 1.07, 1.24); while little difference was seen for other apps: history of subfertility (aFR = 1.16, 95% CI: 0.91, 1.48) and no history of subfertility (FR = 1.15, 95% CI: 1.06, 1.25) (Supplementary Table SIII).

Discussion

In this large prospective cohort study of pregnancy planners, we found consistently positive associations between mobile cycle app use and fecundability. There were stronger associations when the apps were used simultaneously with one or more fertility indicators (basal body temperature, urine LH 'ovulation predictor' testing, cervical fluid or cervix position). While fertility indicators were used more often among women with selected apps, fecundability was similar for the selected apps and other apps, and among all the specific apps with at least 100 users. Positive associations were somewhat stronger for app use assessed at baseline, versus time-varying assessment of use. Speculatively, it may be possible that baseline users were on average more consistent or effective users, but we have no evidence to evaluate this possibility directly.
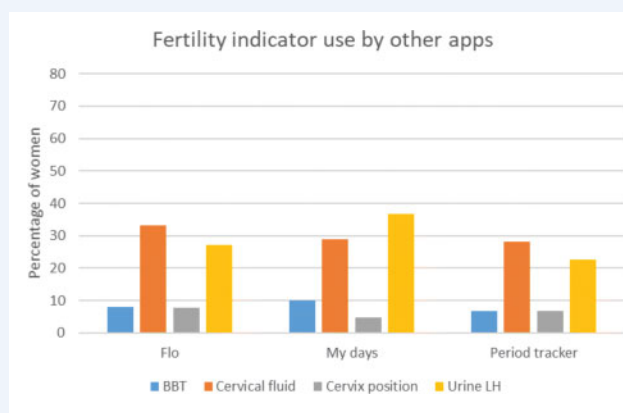

Figure 3. Fertility indicator use for other apps.

We selected five menstrual cycle apps a priori as being more likely to record fertility indicators and therefore presumably more likely to provide accurate information about the fertile window, which in theory would be associated with increased fecundability. In fact, the selected apps had a higher concurrent use of fertility indicators (basal body temperature, cervical fluid, cervix position or urine LH 'ovulation predictor' testing). However, the increase in fecundability was similar for selected apps versus other apps (at baseline, FR = 1.20, 95% CI: 1.13, 1.28, vs. FR = 1.21, 95% CI: 1.13, 1.30, respectively). In supplementary analyses, we did find higher fecundability with the selected apps versus the other apps among women with a history of subfertility (trying to conceive for 6 months or more in the past).

All apps had a stronger association with increased fecundability when they were used simultaneously with fertility indicators (basal body temperature, cervical fluid, cervix position or urine LH 'ovulation predictor' testing). This finding is consistent with our prior report that use of fertility indicators, including 'charting cycles', was associated with increased fecundability in PRESTO (Stanford *et al.*, 2019). It should be noted that all cycle apps require the identification of the first day of the menstrual cycle and thus cycle length, which is in itself a fertility indicator that can be used to estimate the fertile window (Arevalo *et al.*, 1999; Li *et al.*, 2016). In addition, there were 114 apps with fewer than 100 users at baseline; although these apps are included among the 'other' apps in our results, we cannot necessarily apply our results to any of them individually.

Limitations

Data on participant app use were self-reported on the PRESTO baseline and follow-up questionnaires, as opposed to downloaded directly from the apps themselves; thus, we have no information about the regularity, accuracy or intensity of their use. Some studies have found inconsistent recording of data in more than half of women using some of these same cycle apps (Faust *et al.*, 2019; Symul *et al.*, 2019). However, inconsistent use would reasonably be expected to attenuate our results; more consistent use could perhaps enhance fecundability. Women reported information on intercourse frequency averaged over the previous month; thus, we do not have day-specific data on intercourse frequency or timing, which would be necessary to assess directly the influence on the timing of intercourse (Stanford and Dunson,

Table III Associations between use of apps with or without fertility indicators and fecundability.

	Baseline							Time-varying						
	No. of pregnancies	No. of cycles	Unadjusted		Adjusted[a]			No. of pregnancies	No. of cycles	Unadjusted		Adjusted[a]		
No app and no fertility indicator	831	6414	1.00	Ref	1.00	Ref		601	4263	1.00	Ref	1.00	Ref	
No app and fertility indicator*	434	3164	1.08	0.96–1.21	1.03	0.92–1.16		332	2375	1.04	0.92–1.19	1.01	0.89–1.15	
Selected app and no fertility indicator	494	3060	1.19	1.08–1.32	1.17	1.05–1.30		568	3597	1.13	1.02–1.25	1.11	1.01–1.23	
Selected app and fertility indicator*	1534	9064	1.32	1.22–1.43	1.23	1.14–1.34		1721	11108	1.21	1.11–1.31	1.13	1.04–1.23	
Other app and no fertility indicator	760	4991	1.16	1.06–1.28	1.16	1.06–1.27		793	5032	1.09	0.99–1.19	1.10	1.00–1.20	
Other app and fertility indicator*	805	4879	1.35	1.23–1.47	1.30	1.19–1.43		843	5197	1.26	1.15–1.38	1.22	1.11–1.34	

[a]Adjusted for female age, BMI, smoking status, education, partner education, income, hormonal method for most recent contraception, sleep, prior pregnancy, multivitamin use, race/ethnicity, caffeine intake, intercourse frequency, marital status, prior history of subfertility and cycle regularity.
*Fertility indicator: basal body temperature, urine LH, cervical fluid, or cervix position.

2007). The prospective time-varying assessment of app use was based on questionnaires completed every 2 months, which would not capture more frequent changes; this may result in some non-differential misclassification of app use, which would generally bias results towards the null. In addition, after completion of the baseline questionnaire, half of the women were randomised to receive a free subscription to Fertility Friend, unless they were already using it. In the baseline analysis, Fertility Friend was used in 14% of cycles (4378/31 572), whereas based on time-varying analysis, Fertility Friend was used in 21% of cycles (6778/31 572). Furthermore, the proportion of women using a specific app at baseline who discontinued that app by the first follow-up questionnaire was only 10% for Fertility Friend, but ranged from 23–35% for the other apps. However, the FRs for the time-varying analyses were close to those from the baseline analyses.

Our study is based on observational data, rather than randomised assignment; we urge caution in making any causal inferences. The increased fecundability, consistent across all app categories and individual apps with over 100 users, could reflect a causal effect of using the apps to time intercourse during the fertile window. On the other hand, we cannot exclude the possibility that some or possibly all of the observed associations may arise from residual confounding, whereby use of a cycle app is a marker for engagement, interest or behaviour that has not been fully controlled by measured covariates. We did, however, adjust for many covariates known to be associated with fecundability, including age, education, socioeconomic status, use of folic acid and others. A trial in which investigators randomised cycle app use (in a 1:1 ratio) to participants would more effectively control for measured and unmeasured confounders at baseline; however, recruitment of pregnancy planners not already using apps may be challenging, as prevalence of app use tends to be high in this population.

Our study is unique in examining the use of multiple cycle apps across a diverse population of pregnancy planners, with assessment of a wide variety of relevant factors for fecundability. The use of apps and of fertility indicators was one among many components of the study questionnaire, which should reduce any influence of social desirability in responses. Our results are broadly consistent with our own prior analysis of fertility indicators and other cohort studies, which have found positive impact of fertility indicators on fecundability, particularly for cervical fluid or urine LH testing (Robinson *et al.*, 2007; Evans-Hoeker *et al.*, 2013; Tiplady *et al.*, 2013; Mu and Fehring, 2014; Johnson *et al.*, 2019; Stanford *et al.*, 2019). Randomised trials of fertility indicators have generated mixed results, some positive and some null, which may reflect difficulties in conducting trials among women trying to conceive, including enrolling women early in their attempts, heterogeneity in levels of motivation to conceive and self-adoption of interventions among women who are highly motivated to conceive (Robinson *et al.*, 2007; Stanford *et al.*, 2014; Manders *et al.*, 2015; Johnson *et al.*, 2019). It is clear from this study that use of one or more cycle apps is highly prevalent among educated women planning pregnancy. Our results indicate that the combined use of fertility indicators and apps may be more effective than use of either alone.

Most women in the study were college educated, and we cannot say whether these results would extrapolate to less-educated women. However, some prior studies of fertility awareness methods used to avoid pregnancy have found similar effectiveness among women of both high- and low-educational backgrounds (World Health Organization, 1981).

Conclusion

Our results indicate that use of mobile cycle tracking apps may help women trying to conceive, potentially increasing fecundability (per-cycle probability of pregnancy) by 12–20%. Based on observational data, we found that the use of fertility indicators (basal body temperature, cervical fluid, cervix position or urine LH 'ovulation predictor' testing) together with use of an app appears to be more effective than using apps that only track the start date of each menstrual cycle. We saw no consistent results that would distinguish among the eight specific apps that were each used by at least 100 women in the study.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Acknowledgements

Authors' roles

J.B.S. and L.A.W. designed the questions related to app use and fertility indicator use within the PRESTO study. L.A.W. serves as the PI for the PRESTO study. J.B.S. directed the analysis. S.K.W. contributed to the analytic plan and conducted the analysis. E.E.H. and K.J.R. contributed key insights to design of the study and conduct and interpretation of the analysis. J.B.S. wrote the manuscript. All authors revised the manuscript for critical intellectual content. All authors made significant contributions to the manuscript in accordance with the ICMJE guidelines.

Funding

Conflict of interest

In the last 3 years, Dr L.A.W. has served as a fibroid consultant for AbbVie.com. Dr L.A.W. has also received in-kind donations from Sandstone Diagnostics, Swiss Precision Diagnostics, FertilityFriend.com and Kindara.com for primary data collection and participant incentives in the PRESTO cohort. Dr J.B.S. reports personal fees from Swiss Precision Diagnostics, outside the submitted work. The remaining authors have nothing to declare.

References

American Society for Reproductive Medicine. Optimizing natural fertility: a committee opinion. *Fertil Steril* 2017;**107**:52–58.

Arevalo M, Sinai I, Jennings V. A fixed formula to define the fertile window of the menstrual cycle as the basis of a simple method of natural family planning. *Contraception* 1999;**60**:357–360.

Baird DD, Wilcox AJ, Weinberg CR. Use of time to pregnancy to study environmental exposures. *Am J Epidemiol* 1986;**124**:470–480.

Berglund Scherwitzl E, Linden Hirschberg A, Scherwitzl R. Identification and prediction of the fertile window using NaturalCycles. *Eur J Contracept Reprod Health Care* 2015;**20**:403–408.

Brayer FT, Chiazze L Jr, Duffy BJ. Calendar rhythm and menstrual cycle range. *Fertil Steril* 1969;**20**:279–288.

Bull JR, Rowland SP, Scherwitzl EB, Scherwitzl R, Danielsson KG, Harper J. Real-world menstrual cycle characteristics of more than 600,000 menstrual cycles. *NPJ Digit Med* 2019;**2**:1–8.

Colombo B, Masarotto G. Daily fecundability: first results from a new data base. *Demographic research* 2000;**3**: 1–39.

Duane M, Contreras A, Jensen ET, White A. The performance of fertility awareness-based method apps marketed to avoid pregnancy. *J Am Board Fam Med* 2016;**29**:508–511.

Dunson DB, Baird DD, Wilcox AJ, Weinberg CR. Day-specific probabilities of clinical pregnancy based on two studies with imperfect measures of ovulation. *Hum Reprod* 1999;**14**:1835–1839.

Ecochard R, Boehringer H, Rabilloud M, Marret H. Chronological aspects of ultrasonic, hormonal, and other indirect indices of ovulation. *BJOG* 2001;**108**:822–829.

Ecochard R, Duterque O, Leiva R, Bouchard T, Vigil P. Self-identification of the clinical fertile window and the ovulation period. *Fertil Steril* 2015;**103**:1319–1325.e3.

Eijkemans MJC, Leridon H, Keiding N, Slama R. A systematic comparison of designs to study human fecundity. *Epidemiology* 2019;**30**:120–129.

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**: 1033–1038.

Faust L, Bradley D, Landau E, Noddin K, Farland LV, Baron A, Wolfberg A. Findings from a mobile application-based cohort are consistent with established knowledge of the menstrual cycle, fertile window, and conception. *Fertil Steril* 2019;**112**:450–457.e3.

Freis A, Freundl-Schutt T, Wallwiener LM, Baur S, Strowitzki T, Freundl G, Frank-Herrmann P. Plausibility of menstrual cycle apps claiming to support conception. *Front Public Health* 2018;**6**:98.

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)* 2019:29:84–90.

Keulers MJ, Hamilton CJCM, Franx A, Evers JLH, Bots RSGM. The length of the fertile window is associated with the chance of spontaneously conceiving an ongoing pregnancy in subfertile couples. *Hum Reprod* 2007;**22**:1652–1656.

Li D, Heyer L, Jennings VH, Smith CA, Dunson DB. Personalised estimation of a woman's most fertile days. *Eur J Contracept Reprod Health Care* 2016;**21**:323–328.

Manders M, McLindon L, Schulze B, Beckmann MM, Kremer JA, Farquhar C. Timed intercourse for couples trying to conceive. *Cochrane Database Syst Rev* 2015;**3**:CD011345.

Mu Q, Fehring RJ. Efficacy of achieving pregnancy with fertility-focused intercourse. *MCN Am J Matern Child Nurs* 2014;**39**: 35–40.

Robinson JE, Wakelin M, Ellis JE. Increased pregnancy rate with use of the Clearblue Easy Fertility Monitor. *Fertil Steril* 2007;**87**: 329–334.

Stanford JB. Big data meets the menstrual cycle. *Fertil Steril* 2019; **112**:464–465.

Stanford JB, Dunson DB. Effects of sexual intercourse patterns in time to pregnancy studies. *Am J Epidemiol* 2007;**165**: 1088–1095.

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**:391–399.

Stanford JB, White GL, Hatasaka H. Timing intercourse to achieve pregnancy: current evidence. *Obstet Gynecol* 2002;**100**: 1333–1341.

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**: 892–899.

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

Symul L, Wac K, Hillard P, Salathe M. Assessment of menstrual health status and evolution through mobile apps for fertility awareness. *NPJ Digit Med* 2019;**2**:1–10.

Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. New York City, New York: Springer Verlag, 2000.

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**:138–151.

Weinberg CR, Wilcox AJ, Baird DD. Reduced fecundability in women with prenatal exposure to cigarette smoking. *Am J Epidemiol* 1989;**129**:1072–1078.

Wesselink AK, Wise LA, Hatch EE, Mikkelsen EM, Sorensen HT, Riis AH, McKinnon CJ, Rothman KJ. Seasonal patterns in fecundability in North America and Denmark: a preconception cohort study. *Hum Reprod* 2020;**35**:565–572.

Wilcox AJ, Dunson D, Baird DD. The timing of the "fertile window" in the menstrual cycle: day specific estimates from a prospective study. *Br Med J* 2000;**321**:1259–1262.

Wise LA, Rothman KJ, Mikkelsen EM, Sorensen HT, Riis A, Hatch EE. An internet-based prospective study of body size and time-to-pregnancy. *Hum Reprod* 2010;**25**:253–264.

Wise LA, Rothman KJ, Mikkelsen EM, Stanford JB, Wesselink AK, McKinnon C, Gruschow SM, Horgan CE, Wiley AS, Hahn KA *et al.* Design and conduct of an internet-based preconception cohort study in North America: pregnancy study online. *Paediatr Perinat Epidemiol* 2015;**29**:360–371.

World Health Organization. A prospective multicentre trial of the ovulation method of natural family planning. II. The effectiveness phase. *Fertil Steril* 1981;**36**:591–598.