Journal of Obstetrics and Gynaecology

ISSN: (Print) (Online) Journal homepage: informahealthcare.com/journals/ijog20

Time to pregnancy recognition among users of an FDA-cleared fertility application

Agathe van Lamsweerde, Jack T. Pearson, Rachel Urrutia, Kristina Gemzell-Danielsson, Helena Kopp Kallner, Anita Nelson, Eleonora Benhar, Carlotta Favaro, Elina Berglund Scherwitzl & Raoul Scherwitzl

To cite this article: Agathe van Lamsweerde, Jack T. Pearson, Rachel Urrutia, Kristina Gemzell-Danielsson, Helena Kopp Kallner, Anita Nelson, Eleonora Benhar, Carlotta Favaro, Elina Berglund Scherwitzl & Raoul Scherwitzl (2024) Time to pregnancy recognition among users of an FDA-cleared fertility application, Journal of Obstetrics and Gynaecology, 44:1, 2337687, DOI: 10.1080/01443615.2024.2337687

To link to this article: https://doi.org/10.1080/01443615.2024.2337687

JOURNAL OF OBSTETRICS AND GYNAECOLOGY
2024, VOL. 44, NO. 1, 2337687
https://doi.org/10.1080/01443615.2024.2337687

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Time to pregnancy recognition among users of an FDA-cleared fertility application

Agathe van Lamsweerde[a], Jack T. Pearson[a], Rachel Urrutia[b], Kristina Gemzell-Danielsson[c], Helena Kopp Kallner[d], Anita Nelson[e], Eleonora Benhar[a], Carlotta Favaro[a], Elina Berglund Scherwitzl[a] and Raoul Scherwitzl[a]

[a]Natural Cycles Nordic AB, Stockholm, Sweden; [b]Department of Obstetrics and Gynaecology, UNC-Chapel Hill, NC, USA; [c]Department of Women's and Children's Health, Division for Neonatology,Obstetrics and Gynecology and Reproductive Health, Karolinska Institutet, and Karolinska University Hospital, Stockholm, Sweden; [d]Department of Clinical Sciences at Danderyd Hospital, Karolinska Institutet, Sweden and Department of Obstetrics and Gynecology, Danderyd Hospital, Stockholm, Sweden; [e]Western University of Health Sciences, Pomona, CA, USA

ABSTRACT

Background: Previous investigations of time-to-pregnancy recognition have analysed data from national surveys and clinics, but this has not been investigated in the context of digital fertility applications. Timely pregnancy recognition can help individuals in health and pregnancy management, reducing maternal and foetal risk and costs, whilst increasing treatment options, availability, and cost. **Methods:** This dataset contained 23,728 pregnancies (conceived between June 2018 and December 2022) from 20,429 participants using a Food and Drug Administration (FDA) cleared fertility app in the United States. Most participants (with non-missing information) identified as Non-Hispanic White, and one-third reported obtaining a university degree. We used two-tailed Welch's t-test, Mann–Whitney U-test, and two-tailed Z-tests to compare time to pregnancy recognition between those using the app to conceive or contracept. **Results:** Participants using an app to conceive recognised pregnancy on average at 31.3 days from last menstrual period (LMP) compared to 35.9 days among those using the app to prevent pregnancy. **Conclusion:** Generalisability is limited, as all participants were using a fertility app and had relatively homogenous sociodemographic characteristics.

PLAIN LANGUAGE SUMMARY

People who recognise pregnancy early may benefit, as earlier recognition can reduce costs and risks, and make more treatment options available. In the past, researchers have studied the time it takes for an individual to recognise that they are pregnant by asking them in national surveys or when they attend a clinic. However, with the advent of digital fertility tracking apps, we investigated the time it takes to recognise pregnancy when using such an app. We analysed data from 23,728 pregnancies from 20,429 users of the Natural Cycles app between June 2018 and December 2022. We found that participants using the app to try to get pregnant recognised pregnancy an average of 4.6 days earlier than those using the app to prevent pregnancy.

KEYWORDS
Contraception; pregnancy; pregnancy recognition; fertility application; conception

Introduction

Scholars have called for research on pregnancy recognition trajectories across different contexts (Strong *et al.* 2023). Studies (with variable methodological approaches and definitions of pregnancy recognition) show inconsistent results concerning time to pregnancy recognition (TTPR) (Table 1), but generally, later recognition was more common among youth (Finer *et al.* 2006), those with lower education attainment, lower socioeconomic status, unintended pregnancy (Dott et al. 2010, Swanson *et al.* 2014, Ayoola 2015), or identifying as Black and Latino (Branum and Ahrens 2017). A 2021 study found that among women seeking second-trimester abortions in the United States (US), most recognised pregnancy more than 8 weeks after LMP; more than one in five recognised pregnancy after 20 weeks (Foster *et al.* 2021).

Timely recognition of pregnancy can facilitate health and pregnancy management. Those who will seek an abortion may benefit from earlier pregnancy detection (Ralph *et al.* 2022) which can broaden method options to include medication abortion, reduce cost of care, and reduce risk of complications (Jones *et al.* 2017; Upadhyay *et al.* 2022, 2015). Early pregnancy detection and first-trimester prenatal care may decrease the risk of adverse pregnancy outcomes (American College of Obstetricians and Gynecologists,2005 Ayoola *et al.* 2009, Floyd *et al.* 2013). Attempts to predict pregnancy go back decades, often using Bayesian methods (Lum *et al.*

CONTACT Jack T. Pearson, jack.pearson@naturalcycles.com, Natural Cycles Nordic AB, Stockholm, Sweden
Supplemental data for this article can be accessed online at https://doi.org/10.1080/01443615.2024.2337687.

Table 1. Summary of the literature investigating time to pregnancy detection.

Paper	#participants	location	sampling date	sample	Timing/ definition of pregnancy detection
(Ayoola 2015)	143,303	29 US states	2000–2004	**live births** within 2 to 6 months before being contacted	Self-reported survey of women that had had a live birth 2–6 months before being contacted. Participants were asked whe n they were 'sure' they were pregnant (pregnancy test or nurse/doctor visit), which is analysed as pregnancy '**recognition**'.
(Branum and Ahrens 2017)	17,406	US	1995–2013	at least one pregnancy in the 4 or 5 years prior to interview that did **not** result in induced **abortion or adoption**	Self-reported timing of pregnancy '**awareness'** in a national survey asking women about past pregnancies.
(Finer *et al.*2006)	1209	US	2004	**abortion** patients at 11 large providers	In-clinic questionnaire. Self-reported pregnancy **'suspicion'**, pregnancy test timing.
(Foster *et al.*2021)	956	US	2008–2010	**abortion** patients at 30 abortion facilities contacted 1 week after either receiving or being denied an abortion	Telephone interviews. Participants were asked when they first discovered they were pregnant, which is analysed as pregnancy '**recognition**' or '**discovery**' used interchangeably.
(Finer *et al.*2006, McCarthy *et al.*2018)	458	Utah	2013–2014	**abortion** patients at 4 family planning facilities	Self-reported state-mandated survey from abortion information visit patients are required to complete prior to an abortion. Pregnancy '**discovery**'.
(Ralph *et al.*2022)	259	6 US states	2016–2017	visitors to reproduction clinics for **mixed reasons** (34% for abortion, 25% for prenatal care, 31% for pregnancy tests)	In-clinic questionnaire. Self-reported pregnancy '**suspicion**', home test timing or '**confirmation**' and clinic visit.
(Swanson *et al.*2014)	592	6 US states	2010	**abortion** patients at 6 large abortion clinics	In-clinic computer questionnaire. Participants were asked when they took their first pregnancy test, which was gauged against gestational age at time of visit evaluated by ultrasound, which is analysed as pregnancy '**recognition**'.

2016). Other women's health tracking mobile applications have been used to predict pregnancy (Liu *et al.* 2019). Many of the hundred of fertility apps on the market are heterogeneous in their underlying methods of predicting fertile days and the benefit of using them to reduce time to pregnancy needs to be evaluated (Ali *et al.* 2020). Fecundability in relation to use of mobile computing apps to track the menstrual cycle has also been investigated, finding that they can increase fecundability from 12–20% (Stanford *et al.* 2020). Overall, individuals using fertility apps want the apps to be accurate, evidence based regardless of whether they are using them as contraceptives or for pregnancy planning (Earle *et al.* 2021). Better understanding of TTPR in the United States, as well as its variation according to multiple physiological or social variables, is needed to guide clinical counselling and legal guidelines (Watson and Angelotta 2022).

We aimed to investigate self-reported gestational age at pregnancy detection in a cohort of Natural Cycles (NC°) users, and to understand factors associated with TTPR.

Methods

We conducted a retrospective cohort analysis of prospectively collected data on pregnancy recognition, defined as the date participants reported the first positive pregnancy test result in the app. Natural Cycles users track menstrual cycle and basal body temperatures (BBT), and an algorithm predicts the most fertile days of their menstrual cycle, either for contraception or to aid conception. We focused this analysis on users residing in the United States who registered on the application after June 1, 2018 and were 18 to 45 years old. We studied all pregnancies (as defined below) indicated on the application between June 1, 2018 and January 1, 2023.

FDA-cleared fertility application

Natural Cycles is a mobile-based application for fertility monitoring and contraception cleared by the Food and Drug Administration (FDA) in the United States and marked by the Conformité Européenne (CE) in the European Union. Users must record BBT and menstruation dates and can, optionally, add results from home-based urine luteinizing hormone tests. The proprietary algorithm then estimates and shows the user's most likely upcoming ovulation date, fertile window, and upcoming menstruation dates. Participants can indicate pregnancy by entering a positive urine pregnancy test result, in real time or retroactively.

The user indicates their intent to either plan or avoid a pregnancy by selecting between two modes: NC° Plan Pregnancy or NC° Birth Control. The selected mode does not influence estimates for ovulation and menstruation dates but affects the application's urine pregnancy test recommendations encouraging NC° Plan Pregnancy users to test for pregnancy 1–5 days after their predicted menstruation, depending on their cycle regularity.

As a proxy for participant's intention towards pregnancy, we relied on the NC° mode being used by participants at conception (e.g. NC° Plan Pregnancy or NC° Birth Control). The number of unintended pregnancies in this analysis was aligned with the published 13-cycle cumulative typical use pregnancy probability of 7.2% (Pearson *et al.* 2021).

Quantitative variables

We took the self-reported age for each conception cycle as their age on the first day of the cycle and classified the sample into five categories (18–24, 25–29, 30–34, 35–39, or 40–45 years). We computed the body mass index (BMI) from self-reported height and weight at the time of registration and classified the sample into four categories (<18.5, underweight; 18.5–25, normal weight; >25–30, overweight; >30, obese).

Menstrual cycle length was based on cycles recorded in the app from registration until the cycle prior to conception. We calculated average cycle length and classified these into six categories (0–20, >20–25, >25–30, >30–35, >35–40, >40 days). We defined two categories for the standard deviation in cycle length: <5 and ≥5 days. If no cycles were recorded on the application prior to conception, these variables were reported as missing data.

The application periodically requests sociodemographic information on race/ethnicity, parity, education level, relationship status, and the number of young children living with the participant but answering these questions in the application is voluntary. Sociodemographic variables that were never given by the participant are reported as *No information*. Furthermore, we only included information on the time-varying socio-demographic variables (i.e. parity, relationship status, and the number of young children living with the participant) if this information was updated within the year prior to the start of the conception cycle; otherwise we classified this information as *No information*. Socio-demographic information that was time-invariant or reasonably stable over time (race/ethnicity, education level) was not filtered according to the time it was indicated by the participant.

Outcomes

Our primary outcome was the number of days between the first day of the last recorded menstrual period (LMP) and the first positive self-reported pregnancy urine test result. Our secondary outcomes were the proportion of pregnancies in which recognition time since LMP was => 6 and = >7 weeks.

Sensitivity analyses

We did not restrict the analysis according to the frequency of temperature measurement in the main analysis. Computation of TTPR relies on the correct tracking of the LMP and the first positive pregnancy result.

We performed a sensitivity analysis restricted to participants who added at least 10 BBT measurements between the 1st and 30th day of the conception cycle as a proxy for those most likely to accurately record menstruation or pregnancy test data in a timely manner.

Statistical analysis

We stratified the analytical population into two cohorts according to mode used in the application (NC° Plan Pregnancy vs. NC° Birth Control). For each cohort, we computed the sample mean and 95% confidence interval for TTPR.

We used two-tailed Welch's t-tests to compare average TTPR between cohorts, Mann-Whitney U-tests to compare the probability distributions of TTPR between cohorts, and two-tailed Z-tests to compare proportions of late recognition time between cohorts.

We compared TTPR stratifying for one covariate at a time among sociodemographic variables and menstrual characteristics prior to pregnancy. We computed Kruskal–Wallis H-tests to compare mean TTPR without the assumption of equal variances among categories.

We conducted data extraction and analysis using Kotlin and Python, and we considered values of $P < 0.01$ as statistically significant.

Data privacy

All participants included in this analysis authorised use of their pseudonymized data for research purposes in the application. At any time, participants could remove this consent (including for retrospective data). Researchers only had access to pseudonymized copies of data and were unable to derive the identity of any participant. Natural Cycles is compliant with General Data Protection Regulation (GDPR) standards. We obtained an ethics waiver for analysis of pseudonymized data (Reading Independent Ethics Committee, 10022023).

Results

The dataset contains 23,728 pregnancies from 20,429 participants (Figure 1). Beforehand, we excluded 34 pregnancies for being logged over 40 weeks after last logged menstruation, which would exceed the expected biological length of pregnancy. Participants originated from 50 states, the District of Columbia, and the four territories of the United States.

Overall, two-thirds (75.6%) of pregnancies occurred in individuals aged 25–35 (Table 2). Among the 27% who indicated race/ethnicity, 82.0% identified as non-Hispanic White. A total of 45.0% reported obtaining a university or doctoral degree, whereas 47.5% did not receive or answer the educational level question. We had substantial fractions of missing data for all variables except age (Table 2), for example, BMI remains unknown for 41.1% of participants.

NC° Plan Pregnancy users had significantly higher temperature measurement frequencies versus NC° Birth control users (56.2% of days during first 30 days of the conception cycle, versus 48.8% of days (*p*-value <0.001, data not shown)). Additionally, the mean number of urine LH tests within the studied conception cycle was 3.0 for NC° Plan Pregnancy and 1.2 for NC° Birth control (*p*-value <0.001, data not shown).

Time to pregnancy recognition

Pregnancies occurring on NC° Plan Pregnancy were recognised an average of 4.6 days earlier than those on NC° Birth Control; respectively, 31.3 days versus 35.9 days from LMP (Table 3, Figure 2, *p*-value <0.001).

A significantly smaller proportion of NC° Plan Pregnancy users (6.5%) than NC° Birth Control users (15.7%) recognised

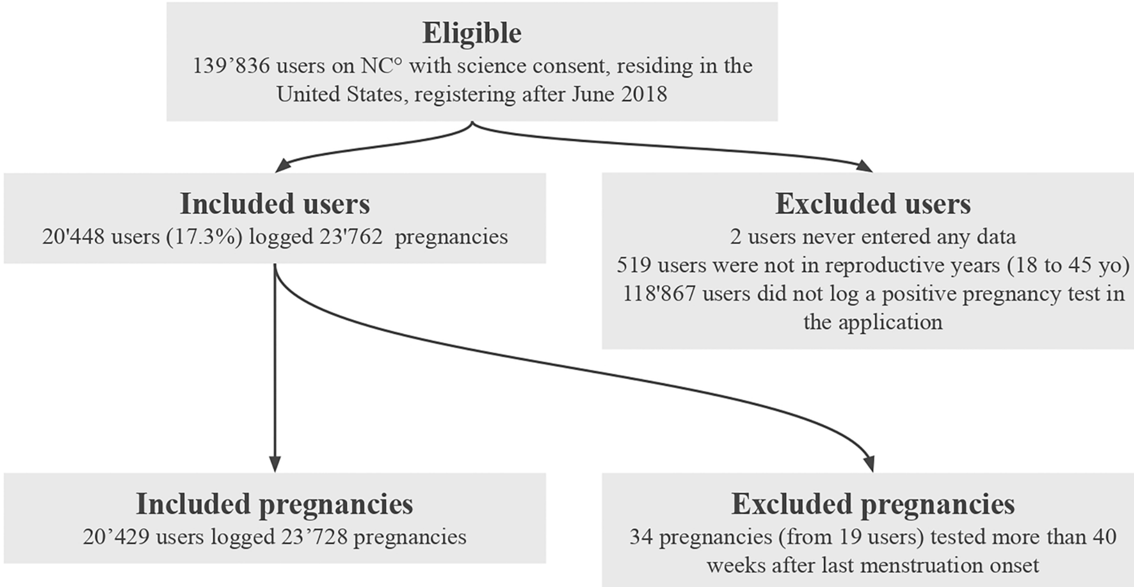

Figure 1. Selection of pregnancies on the Natural Cycles application.

Table 2. Demographic characteristics: distribution of the pregnancies ($n = 23728$) and stratified mean time from the last menstrual period to the first positive pregnancy test (NC° Plan Pregnancy vs NC° Birth Control).

Sociodemographic variable	distribution within the dataset – number, percentage of total	LMP to pregnancy test [days] – mean (with 95% confidence interval)		LMP to pregnancy test >= 6 weeks – percentage of total	
		NC° Plan Pregnancy	NC° Birth Control	NC° Plan Pregnancy	NC° Birth Control
Age at the start of the conception cycle					
18–24	3111 (13.11%)	31.8 (95% CI: 31.2–32.4)	36.9 (95% CI: 35.9–37.8)	8.7%	18.1%
25–29	9423 (39.71%)	31.4 (95% CI: 31.1-31.7)	35.9 (95% CI: 35.3-36.5)	6.9%	15.8%
30–34	8518 (35.90%)	31.2 (95% CI: 31.0–31.5)	35.5 (95% CI: 34.7–36.2)	6.3%	14.6%
35–39	2373 (10.00%)	30.8 (95% CI: 30.3–31.3)	35.2 (95% CI: 33.6–36.7)	4.6%	13.0%
40–45	303 (1.28%)	30.6 (95% CI: 29.3–32.0)	32.8 (95% CI: 31.3–34.3)	4.8%	10.5%
Body mass index					
0–18.5	309 (1.30%)	32.6 (95% CI: 31.2–34.1)	39.2 (95% CI: 33.9–44.4)	8.5%	23.3%
18.5–25	8015 (33.78%)	31.0 (95% CI: 30.7–31.3)	35.3 (95% CI: 34.8–35.9)	5.7%	15.8%
25–30	3567 (15.03%)	30.8 (95% CI: 30.4–31.2)	34.9 (95% CI: 34.1–35.8)	6.0%	13.9%
>30	2095 (8.83%)	31.6 (95% CI: 31.0–32.2)	37.6 (95% CI: 35.7–39.5)	7.6%	16.7%
None	9742 (41.06%)				
Race/ethnicity					
Black or African American or African	224 (0.94%)	32.0 (95% CI: 29.1–35.0)	35.1 (95% CI: 33.1–37.2)	5.4%	13.2%
Hispanic/Latina	562 (2.37%)	32.3 (95% CI: 31.1–33.5)	34.4 (95% CI: 32.3–36.5)	10.5%	16.1%
Non-Hispanic White	5256 (22.15%)	31.0 (95% CI: 30.7–31.3)	34.8 (95% CI: 34.1–35.6)	5.8%	13.0%
Middle Eastern or North African	17 (0.07%)	28.2 (95% CI: 25.9–30.5)	32.5 (95% CI: 28.1–36.9)	0.0%	0.0%
Central Asian	6 (0.03%)	34.0 (95% CI: 29.6–38.4)	–	0.0%	0.0%
South Asian	25 (0.11%)	33.3 (95% CI: 29.9–36.7)	–	12.5%	0.0%
Southeast Asian	49 (0.21%)	32.2 (95% CI: 27.6–36.8)	35.6 (95% CI: 32.7–38.5)	3.0%	12.5%
East Asian	41 (0.17%)	31.6 (95% CI: 29.5–33.6)	51.8 (95% CI: 13.5–90.1)	9.4%	11.1%
Native Hawaiian or Other Pacific Islander	22 (0.09%)	33.4 (95% CI: 29.2–37.7)	36.7 (95% CI: 30.4–42.9)	15.8%	33.3%
American Indian or Alaska Native or First Peoples or Indigenous or Aboriginal	33 (0.14%)	29.2 (95% CI: 27.8–30.7)	31.3 (95% CI: 28.1–34.6)	0.0%	0.0%
Other	172 (0.72%)	31.2 (95% CI: 30.1–32.2)	33.6 (95% CI: 32.1–35.0)	2.8%	9.4%
None	17321 (73.00%)				
Age of the youngest child living at home					
0–12 months	1495 (6.30%)	30.9 (95% CI: 30.4–31.4)	35.9 (95% CI: 34.2–37.7)	5.1%	16.1%
1–3 years	1214 (5.12%)	30.4 (95% CI: 30.0–30.8)	33.9 (95% CI: 32.7–35.1)	5.2%	12.1%
3–6 years	223 (0.94%)	30.6 (95% CI: 28.7–32.6)	33.1 (95% CI: 31.9–34.3)	5.1%	5.7%
Older than 6 years	186 (0.78%)	31.3 (95% CI: 29.1–33.5)	32.6 (95% CI: 31.0–34.1)	7.1%	9.6%
I don't have children living at home	3848 (16.22%)	31.2 (95% CI: 30.9–31.6)	34.6 (95% CI: 34.0–35.3)	6.5%	12.9%
None	16762 (70.64%)				
Parity, number of previous pregnancies					
No, never	6305 (26.57%)	31.7 (95% CI: 31.5–32.0)	36.4 (95% CI: 35.7–37.2)	8.1%	17.0%
Yes, once	1873 (7.89%)	30.9 (95% CI: 30.3–31.5)	35.7 (95% CI: 34.3–37.1)	6.0%	16.3%
Yes, twice	701 (2.95%)	31.0 (95% CI: 30.1–31.9)	34.3 (95% CI: 32.6–36.0)	5.7%	11.8%
Yes, 3 times or more	575 (2.42%)	31.0 (95% CI: 29.4–32.5)	32.6 (95% CI: 31.6–33.7)	4.7%	10.8%

(*Continued*)

Table 2. Continued.

Sociodemographic variable	distribution within the dataset – number, percentage of total	LMP to pregnancy test [days] – mean (with 95% confidence interval)		LMP to pregnancy test >= 6 weeks – percentage of total	
		NC° Plan Pregnancy	NC° Birth Control	NC° Plan Pregnancy	NC° Birth Control
None	14274 (60.16%)				
Highest education level					
Elementary school	13 (0.05%)	27.9 (95% CI: 26.5–29.2)	31.3 (95% CI: 27.8–34.8)	0.0%	0.0%
High school degree	905 (3.81%)	32.3 (95% CI: 31.2–33.3)	36.2 (95% CI: 34.6–37.8)	7.5%	14.0%
Trade/technical/vocational training	856 (3.61%)	32.0 (95% CI: 30.8–33.1)	35.9 (95% CI: 34.5–37.3)	7.6%	17.4%
University degree	9928 (41.84%)	31.2 (95% CI: 30.9–31.4)	35.2 (95% CI: 34.6–35.7)	6.2%	14.9%
PhD	752 (3.17%)	31.3 (95% CI: 30.3–32.4)	39.0 (95% CI: 34.8–43.2)	4.5%	17.1%
None	11274 (47.51%)				
Relationship status					
In a relationship	187 (0.79%)	31.2 (95% CI: 29.6–32.9)	33.0 (95% CI: 32.0–34.1)	0.0%	10.4%
Engaged or married	3176 (13.39%)	30.7 (95% CI: 30.3–31.1)	34.8 (95% CI: 34.2–35.4)	4.6%	13.6%
It's complicated	36 (0.15%)	28.3 (95% CI: 18.7–38.0)	34.4 (95% CI: 30.2–38.6)	0.0%	12.1%
Single	49 (0.21%)	27.0 (95% CI: 22.9–31.1)	34.5 (95% CI: 31.6–37.5)	0.0%	6.5%
None	20280 (85.47%)				
Average cycle length prior to conception					
<20	82 (0.35%)	31.8 (95% CI: 29.1–34.6)	37.8 (95% CI: 30.6–45.0)	13.7%	19.4%
>20-25	1184 (4.99%)	27.9 (95% CI: 27.3–28.5)	34.6 (95% CI: 32.4–36.7)	3.1%	11.7%
>25-30	12495 (52.66%)	29.3 (95% CI: 29.1–29.5)	33.4 (95% CI: 32.9–33.8)	2.5%	8.7%
>30-35	7099 (29.92%)	32.8 (95% CI: 32.5–33.1)	37.4 (95% CI: 36.7–38.0)	7.2%	19.6%
>35-40	1705 (7.19%)	36.5 (95% CI: 35.9–37.0)	42.8 (95% CI: 41.2–44.5)	20.9%	40.0%
>40	317 (1.34%)	41.5 (95% CI: 39.3–43.6)	43.1 (95% CI: 40.3–45.9)	38.9%	48.2%
None	846 (3.57%)				
Variation in cycle length prior to conception					
<5	17644 (74.36%)	30.5 (95% CI: 30.3–30.6)	34.6 (95% CI: 34.2–34.9)	4.4%	12.5%
≥5	3038 (12.80%)	33.8 (95% CI: 33.3–34.3)	38.0 (95% CI: 36.8–39.2)	13.1%	23.9%
None	3046 (12.84%)				

Table 3. Time to pregnancy confirmation (days from the first day of the last menstrual period to the first positive pregnancy test), average and right tail of the distribution: NC° Plan Pregnancy vs NC° Birth Control.

	LMP to pregnancy test [days] – mean (with 95% confidence interval)			LMP to pregnancy test ≥6 weeks – percentage of total			LMP to pregnancy test ≥7 weeks – percentage of total		
	NC° Plan Pregnancy	NC° Birth Control	statistical p value	NC° Plan Pregnancy	NC° Birth Control	statistical p value	NC° Plan Pregnancy	NC° Birth Control	statistical p value
main analytical population - no conditions	31.3 (95% CI: 31.1–31.5)	35.9 (95% CI: 35.5–36.3)	WT <0.001 MWU = <0.001	6.5%	15.7%	Z < 0.001	3.0%	7.6%	Z < 0.001
tracking regularly – 10+ basal body temperatures tracked before the 30th day of the conception cycle	30.6 (95% CI: 30.5–30.7)	34.2 (95% CI: 33.9– 34.5)	WT <0.001 MWU <0.001	5.6%	13.1%	Z < 0.001	2.1%	5.6%	Z < 0.001
statistical p value- between *main analytical population* and *tracking regularly*	WT <0.001 MWU = 0.004	WT <0.001 MWU = 0.02		Z < 0.001	Z < 0.001		Z < 0.001	Z < 0.001	

The statistical *p* values for comparisons are given for the two-tailed Welch's T-tests (WT), the Mann–Whitney U-tests (MWU) or the two-tailed Z-test (Z).

the pregnancy later than six weeks after LMP (Table 3, *p*-value <0.001). Similarly, 3.0% of NC° Plan Pregnancy users and 7.6% of NC° Birth Control users recognised pregnancy later than seven weeks after LMP (*p*-value <0.001). We compared TTPR with observations from prior studies (Table S1).

Sensitivity analyses

We observed faster recognition times among participants with at least ten BBTs within the first 30 days of the conception cycle (Table 3), 0.8 and 1.7 days for participants on NC° Plan Pregnancy and NC° Birth Control, (*p*-value <0.001). In this subgroup, we also found a smaller proportion with later pregnancy recognition, especially for participants on NC° Birth Control: there was an absolute decrease of 2.6% and 2.0% in the proportion of participants recording a first positive pregnancy test at 6 and 7 weeks, respectively (*p*-value <0.001). We performed a second sensitivity analysis to evaluate the impact of in-application communication about late menstruation to participants on NC° Birth Control but found

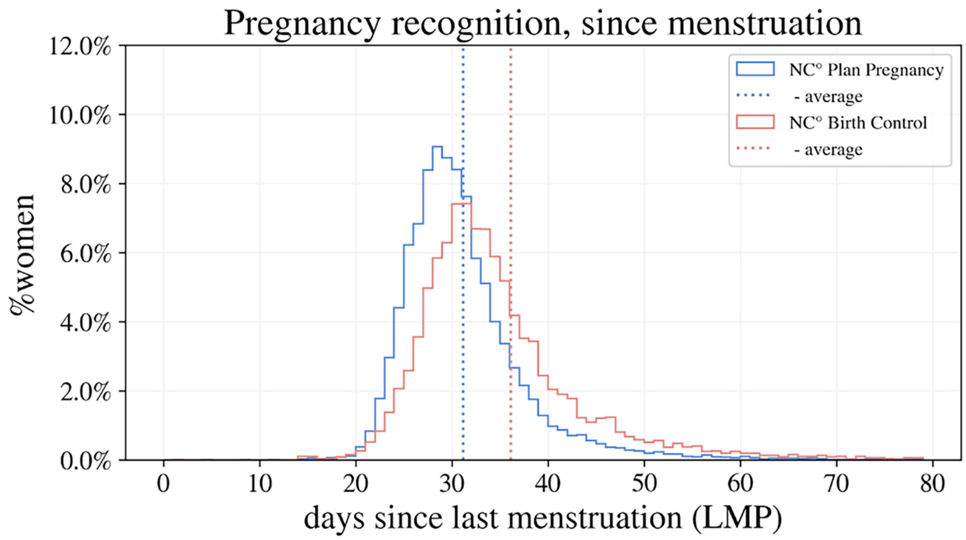

Figure 2. Mean time from last menstrual period to first positive pregnancy test: NC° Plan Pregnancy vs NC° Birth Control.

no significant difference in mean pregnancy recognition time between NC° Birth Control users receiving or not receiving warning messages on late menstruation (p-value = 0.53), see supplementary material (Table S2).

Effect of demographic and menstrual characteristics on pregnancy recognition

BMI, parity, and the presence of young children at home was associated with TTPR (Table 2). Among participants on NC° Plan Pregnancy, we observed longer pregnancy recognition time among NC° Plan Pregnancy users who registered a BMI above 30.0 (obese) or below 18.5 (underweight), (p-value = 0.004). On the contrary, parity (versus nulliparity) was associated with shorter mean recognition time and lower likelihood of late pregnancy recognition (p-value <0.001). Maternal age (p-value = 0.06) and living with young children (p-value = 0.06) were not significantly correlated to pregnancy recognition time. No significant differences across other demographic categories (race/ethnicity, relationship status, and education level) were observed ($p > 0.05$).

Both average cycle length and cycle length regularity had a significant influence on TTPR. Participants with irregular cycles (cycle-length standard deviation in previous cycles ≥ 5 days) or longer cycles took longer, on average, to recognise pregnancy (p-value <0.001 for both analyses).

Discussion

In this cohort of nearly thirty thousand individuals, pregnancies logged on NC° Plan Pregnancy and on NC° Birth Control were recognised at 31.3 and 35.9 days from the LMP, respectively. A significantly smaller proportion of NC° Plan Pregnancy users (6.5%) versus NC° Birth Control users (15.7%) recognised the pregnancy later than six weeks after the LMP.

This cohort of digital fertility application users (including both those trying to conceive and prevent pregnancy) appears to detect pregnancies earlier on average than populations studied in previous publications (Table S1). However, sociodemographic characteristics of our analytic population and those of previous studies differ substantially. For example, in Ralph *et al.* (2022), study participants identified mostly as Latina (42%) or Non-Latina Black (22%), largely reported a high school diploma as their highest education level (74%), and nearly half received public assistance (42%). In comparison with the Ralph cohort, the Natural Cycles participants (the majority of whom, where data were provided, identified as non-Hispanic White and holding a university degree) recognised pregnancy on average 12 days sooner when trying to avoid it. Furthermore, the proportion of individuals recognising pregnancy over six or seven weeks after LMP in our study was less than one-third of that reported in previous studies (Ayoola,2015 Branum and Ahrens 2017, McCarthy *et al.* 2018, Ralph *et al.* 2022).

Several factors may have contributed to the shorter pregnancy recognition time in our study. Natural Cycles relies on daily tracking of menstrual data; hence, users may be more alert to missed menstruation than the general population. In addition, the application provides estimations of the participant's ovulation date and fertile window; therefore, participants may be comparatively more conscious of the likelihood of conception from unprotected intercourse on a day indicated by the app as fertile. Furthermore, in addition to the unique sociodemographic characteristics of our sample, our cohort may differ in other unmeasured ways from those who do not use an app for fertility tracking. For example, the application of Natural Cycles offers regular in-application content on sexual and reproductive health, which may improve knowledge of the signs and symptoms of pregnancy, which in turn may accelerate timely pregnancy recognition (Watson and Angelotta 2022).

The user-selected fertility mode had a significant impact on the TTPR, with participants who conceived in the NC° Plan Pregnancy mode logging a first positive pregnancy test on average five days earlier than those on NC° Birth Control. This result correlates closely with previous studies in which unintended or unwanted pregnancies were suspected, recognised, or confirmed later (Ayoola 2015, Branum and Ahrens 2017, Ralph *et al.* 2022). Participants actively planning a pregnancy are more prone to testing before or on the first day of missed menstruation, whereas those trying to avoid pregnancy suspect pregnancy only after late menstruation.

Sensitivity analyses suggest that more active app users recognised pregnancy earlier, particularly those using NC° Birth Control. Participants who use the application more regularly may have a higher interest in tracking their fertility, may receive more accurate menstruation predictions from the application, may have more knowledge on fertility and pregnancy, and are prompted earlier to test for pregnancy. Those using the application more sporadically may be less likely to correctly log their first positive pregnancy test on the actual date of testing, which would bias the pregnancy recognition time of this cohort towards longer durations.

The relationship between pregnancy recognition time and various demographic variables in our study showed some similarities and differences with previous studies. While age was not correlated with TTPR in our study, older age was associated with earlier pregnancy recognition elsewhere (Ayoola 2015, Branum and Ahrens 2017, Ralph *et al.* 2022). We also identified a relationship between parity and the presence of young children at home with faster pregnancy recognition, as in some (Foster *et al.* 2021) but not all other studies (Ayoola 2015, Branum and Ahrens 2017, Ralph *et al.* 2022). We observed no significant differences in pregnancy recognition time by relationship status, whereas other studies found that pregnancies were identified later if individuals were not married or argued more with their partners (Ayoola,2015 Ralph *et al.* 2022). We observed no significant relationship between race/ethnicity and TTPR in our cohort. Previous cohorts among lower income, lower educated populations have shown an increase in TTPR among Hispanic and Non-Hispanic Black compared to Non-Hispanic White individuals. Since race is a sociologic and not a biologic construct, these differences, if present, may be due to systemic inequities in healthcare access literacy and body literacy. (Ayoola 2015, Branum and Ahrens 2017, Ralph *et al.* 2022). Our findings may be different because this group had fewer other adverse social determinants (e.g. education, health insurance) than previous cohorts. Additionally, our sample was relatively homogenous, and had substantial missing data for several socio-demographic variables, thus limiting our ability to detect potential differences.

We also observed later pregnancy recognition among participants with longer or irregular menstrual cycles, which can be easily explained by physiological and behavioural factors. A longer follicular phase delays fertilisation, implantation, and hCG increase. Consequently, pregnancy recognition is naturally farther away from the LMP. An individual with a more irregular cycle may take longer to suspect pregnancy as delayed menses are 'normal' for them. The average cycle length, average follicular phase length, and cycle length variation did not vary significantly between the pregnancies on NC° Plan Pregnancy and NC° Birth Control, and hence did not contribute to the observed gap in pregnancy recognition time between the two cohorts.

Our study had several limitations, including substantial missing data for several sociodemographic characteristics. This is due to the fact that this is a pragmatic study conducted amongst a real-world population of contraceptive app users not in the context of a clinical trial. Users were not required to input all data into the app in order to use it. The benefit to this approach is studying how real-world fertility awareness app users interact with pregnancy testing.

We also did not collect data on all social determinants of health which have been observed to be correlated with the timing of pregnancy recognition, such as food insecurity, (Ralph *et al.* 2022). However, the sociodemographic information collected shows the cohort to be a relatively homogeneous sample, not representative of the general population in the US. In our cohort, participants predominantly identified as Non-Hispanic White (82.0% compared to 59.3% in the general population (United States Census Bureau 2023)) and highly educated (85.8% of those who answered the question had obtained a university degree). Our dataset, representing only women who use a fertility awareness-based method, is likely different from the general population in ways that would impact TTPR. Some misclassification of pregnancy intentions is also possible in our data. We cannot draw meaningful conclusions for the cohort with the shortest cycles with an average cycle length below 20 days due to the small size of the sample. Finally, this study focused solely on the recognition of pregnancy through a positive pregnancy urine test result. Pregnancy suspicion based on physical symptoms and pregnancy confirmation by a healthcare professional were not captured.

Furthermore, this study could only compare the TTPR in cycles where a positive urine pregnancy test was logged into the application. However, there could be a difference in the frequency of unreported pregnancy tests between participants planning to conceive and those intending to avoid pregnancy, which could influence the observed difference in TTPR between those cohorts.

Timely recognition of pregnancy can aid individuals in their health and pregnancy management. Not all fertility-tracking apps are regulated, and some may not have robust data privacy policies in place for the benefit of the user (Mozilla 2022). In such cases, any gains in the speed of pregnancy awareness may need to be balanced against the potential risks of uploading pregnancy-related information, particularly in places where abortions are criminalised.

Acknowledgements

We would like to express our sincere gratitude to the participants, who generously consented to sharing their anonymized menstrual and fertility data on the Natural Cycles application for research. A special thank you to Chelsea Polis who provided substantial input into prior versions of this manuscript; she has not received any financial support from Natural Cycles.

Author contributions

AvL and EB designed, performed and interpreted the analysis and wrote the manuscript, JTP, RU, KGD, HKK, CF, AN, EB and RS interpreted the analysis, wrote and reviewed the manuscript.

Disclosure statement

AvL, JP, EB, and CF are employees of Natural Cycles with shares or stock warrants in the company. EBS and RS are the CEO's and Co-founders of Natural Cycles and have shares in the company. KGD received honorarium for the support of manuscript writing and as a speaker at meetings for Natural Cycles. HKK received a past honorarium from Natural Cycles. AN had previously attended an advisory board for Natural Cycles. CBP and RU did not receive any financial support from Natural Cycles.

Funding

Natural Cycles provided access to anonymized data collected on their digital fertility application as well as salaries for AvL, JTP, CF, EB, EBS, and RS as full-time employees.

Data availability statement

Data is available upon reasonable request to the corresponding author.

References

Ali, R., Gürtin, Z. B. and Harper, J. C., 2020. Do fertility tracking applications offer women useful information about their fertile window? *Reproductive Biomedicine Online*. doi: 10.1016/j.rbmo.2020.09.005.

American College of Obstetricians and Gynecologists. 2005. ACOG Committee Opinion Number 313, September 2005. The importance of preconception care in the continuum of women's health care. *Obstetrics and Gynecology*, *106* (3), 1–9.

Ayoola, A. B., 2015. Late recognition of unintended pregnancies. *Public Health Nursing (Boston, Mass.)*, *32* (5), 462–470.

Ayoola, A. B., Stommel, M. and Nettleman, M. D., 2009. Late recognition of pregnancy as a predictor of adverse birth outcomes. *American Journal of Obstetrics and Gynecology*, *201* (2), 156.e1–6–156.e6.

Branum, A. M. and Ahrens, K. A., 2017. Trends in timing of pregnancy awareness among US women. *Maternal and Child Health Journal*, *21* (4), 715–726.

Dott, M., National Birth Defects Prevention Study., *et al*., 2010.Association between pregnancy intention and reproductive-health related behaviors before and after pregnancy recognition, national birth defects prevention study, 1997-2002. *Maternal and Child Health Journal*, *14* (3), 373–381

Earle, S., *et al*., 2021. Use of menstruation and fertility app trackers: a scoping review of the evidence. *BMJ Sexual and Reproductive Health*, *47* (2), 90–101.

Finer, L. B., *et al*., 2006. Timing of steps and reasons for delays in obtaining abortions in the United States. *Contraception*, *74* (4), 334–344.

Floyd, R. L., *et al*., 2013. A national action plan for promoting preconception health and health care in the United States (2012-2014). *Journal of Women's Health (2002)*, *22* (10), 797–802.

Foster, D. G., Gould, H. and Biggs, M. A., 2021. Timing of pregnancy discovery among women seeking abortion. *Contraception*, *104* (6), 642–647.

Jones, H. E., *et al*., 2017. First trimester medication abortion practice in the United States and Canada. *PLoS One*, *12* (10), e0186487.

Liu, B., *et al*., 2019. Predicting pregnancy using large-scale data from a women's health tracking mobile application. *Proceedings of the. International World-Wide Web Conference. International* WWW *Conference*, *2019*, 2999–3005.

Lum, K. J., *et al*., 2016. A Bayesian joint model of menstrual cycle length and fecundity. *Biometrics*, *72* (1), 193–203.

McCarthy, M., *et al*., 2018. Predictors of timing of pregnancy discovery. *Contraception*, *97* (4), 303–308.

Mozilla 2022. *In Post Roe v. Wade Era, Mozilla Labels 18 of 25 Popular Period and Pregnancy Tracking Tech With *Privacy Not Included Warning*. Accessed 27th April 2023, https://foundation.mozilla.org/en/blog/in-post-roe-v-wade-era-mozilla-labels-18-of-25-popular-period-and-pregnancy-tracking-tech-with-privacy-not-included-warning/.

Pearson, J. T., *et al*., 2021. Contraceptive effectiveness of an FDA-cleared birth control app: results from the natural cycles U.S. Cohort. *Journal of Women's Health (2002)*, *30* (6), 782–788.

Ralph, L. J., *et al*., 2022. Home pregnancy test use and timing of pregnancy confirmation among people seeking health care. *Contraception*, *107*, 10–16.

Stanford, J. B., *et al*., 2020. Fecundability in relation to use of mobile computing apps to track the menstrual cycle. *Human Reproduction (Oxford, England)*, *35* (10), 2245–2252.

Strong, J., *et al*., 2023. Pregnancy recognition trajectories: a needed framework. *Sexual and Reproductive Health Matters*, *31* (1), 2167552.

Swanson, M., *et al*., 2014. Delayed pregnancy testing and second-trimester abortion: can public health interventions assist with earlier detection of unintended pregnancy? *Contraception*, *89* (5), 400–406.

United States Census Bureau 2023. Quick facts, United states. Available at: https://www.census.gov/quickfacts/fact/table/US/RHI825221#RHI825221

Upadhyay, U. D., *et al*., 2015. Incidence of emergency department visits and complications after abortion. *Obstetrics and Gynecology*, *125* (1), 175–183.

Upadhyay, U. D., *et al*., 2022. Trends in self-pay charges and insurance acceptance for abortion in the United States, 2017-20. *Health Affairs* (Millwood), *41* (4), 507–515.

Watson, K. and Angelotta, C., 2022. The frequency of pregnancy recognition across the gestational spectrum and its consequences in the United States. *Perspectives on Sexual and Reproductive Health*, *54* (2), 32–37.