REPRODUCTIVE AND PERINATAL EPIDEMIOLOGY (K FERGUSON, SECTION EDITOR)

Menstrual Cycle Tracking Applications and the Potential for Epidemiological Research: a Comprehensive Review of the Literature

Joelle S. Schantz [1] · Claudia S. P. Fernandez [1] · Anne Marie Z. Jukic [2]

Abstract

Purpose of Review We reviewed published studies on menstrual cycle tracking applications (MCTAs) in order to describe the potential of MCTAs for epidemiologic research.

Recent Findings A search of PubMed, Web of Science, and Scopus for MCTA literature yielded 150 articles. After exclusions, there were 49 articles that addressed the primary interest areas: 1) characteristics of MCTA users in research, 2) reasons women use or continue using MCTAs, 3) accuracy of identifying ovulation and utility at promoting and preventing pregnancy, and 4) quality assessments of MCTAs across several domains.

Summary MCTAs are an important tool for the advancement of epidemiologic research on menstruation. MCTA studies should describe the characteristics of their user base and missing data patterns. Describing the motivation for using MCTAs throughout a user's life and validating the data collected should be prioritized in future research.

Keywords Menstrual cycle · Data collection · Ovulation · Mobile application · Fertility · Epidemiologic research design

Introduction

Digital health applications, or "apps," have become a popular method of tracking important health indicators for clinical and personal use [1]. Apps can be useful tools for tracking and accomplishing health goals [1, 2]. A segment of these health apps is used to track menstrual cycles. Menstrual cycle tacking apps (MCTAs) assist users in observing their menstrual cycle and related signs and symptoms, as well as managing their fertility [3]. MCTAs give users more control over their own personal health [3]. MCTAs have features that can increase users' knowledge about the menstrual cycle in general, and the act of tracking cycles can help users learn the patterns of

their own bodies [4]. This makes a sample of women using an MCTA a useful source of data for menstrual cycle research.

MCTAs are a valuable potential tool for epidemiologic research [1]. With MCTAs, menstrual cycle study samples can expand from hundreds of participants to thousands or more. For some MCTAs, all users agree to share their data anonymously, which avoids the need for study "recruitment" and may decrease volunteer bias [5–9]. Some MCTAs facilitate tracking of ovulation, providing researchers with access to population-level data on ovulation timing, which is not easily accessible. The use of apps is increasing as mobile users switch from web browsing to app use and smart phone use continues to grow worldwide. Three quarters of smartphone subscription growth came from Africa and Asia in the first quarter of 2015 [10]. This suggests that apps are globally available, increasing the feasibility of including a diverse group of users in menstrual cycle research [10]. Use of data from MCTAs could improve our understanding of the menstrual cycle [11].

While the promise of MCTAs for epidemiologic research is exciting, there are potential limitations to these data. As of yet, it is not clear how representative or accurate MCTAs are, or how susceptible they are to missing data and loss to follow-up. The purpose of this literature review was to synthesize

---

This article is part of the Topical Collection on *Reproductive and Perinatal Epidemiology*

✉ Anne Marie Z. Jukic
jukica@niehs.nih.gov

[1] Department of Maternal and Child Health, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

[2] Epidemiology Branch, MD A3-05, National Institute of Environmental Health Sciences, Durham, NC 27709, USA

published literature on MCTAs with respect to their utility for epidemiologic research. For this review, we examined all published studies that included MCTA-collected data and extracted from those studies information related to several primary areas of interest, chosen for their relevance for epidemiologic research. These areas were 1) selection: who uses MCTAs and why, 2) misclassification: is MCTA-collected data accurate, and 3) overall, what is the potential for using MCTAs in epidemiologic research.

Methods

PubMed, Web of Science, and Scopus were searched for literature on the topic of MCTAs. Search terms used were phrases about menstruation (menstrual cycle, menses, menstrual), contraception (natural family planning methods, fertilization, fertility, fertile, conception, conceiving, ovarian cycle, endometrial cycle, family planning, time-to-pregnancy), and applications (smartphone, mobile application, app(s), portable software, portable electronic, app-based). References from retrieved papers were also examined for additional literature. One additional article was included after reviewing these references. These searches yielded a total of 150 articles as of September 2020.

Articles were screened to identify those that address the epidemiologic characteristics of app data; in broad categories, these were 1) accuracy of identifying ovulation or measuring time to pregnancy; 2) demographic or behavioral characteristics of users of MCTAs and reported reasons for using apps, which may further describe population characteristics; and 3) degree of missingness and loss to follow-up in app data. Articles were initially screened on their title and abstract. Publications were excluded either if they did not contain original data or if the purpose of the described apps was solely clinical decision-making or patient education. The remaining 81 articles were subject to a full-text review. The same criteria were applied along with these additional exclusions: meeting or poster abstracts, articles not in English, and studies focusing only on a wearable device. This left 49 articles.

While we had established broad categories of interest prior to the full-text review, we modified these categories, if for example no studies addressed them, or to accommodate categories we had not considered. After reviewing these full-text articles, the specific domains for this review were 1) characteristics of MCTA users in research on menstrual cycles, fertility, and contraception; 2) reasons users use or continue using MCTAs; 3) accuracy of identifying ovulation and utility at promoting and preventing pregnancy; and 4) previously published quality assessments of MCTAs. Table 1 summarizes the domains chosen for this review and the primary objectives of each of the papers included within that domain. We aim to review these specific domains in the published

literature, not make direct comparisons across apps. Our review synthesizes the literature within each of these domains independent of the research objective of each paper.

Results

Characteristics of MCTA Users in Research

This section focuses on published scientific papers that used MCTA data, and as such, the samples of users described in our review are those who volunteered or provided enough data to be included in those publications. Some MCTAs consent women separately for research studies, a separate recruitment step, while others obtain consent from all users when the MCTA is first downloaded. Thus, MCTAs have a population of users, and the published data from any one MCTA may not be representative of all users of that MCTA; however, we do not have any means of describing the characteristics of all users of a given MCTA. Recruitment techniques may help to diversify study participants recruited from an MCTA. For example, a US study of users of the app Dot (Dynamic Optimal Timing) found that self-guided enrollment was preferred, and that the percentage of Black and Hispanic participants increased when recruitment changed from enrollment with a study representative to a self-guided process [12•].

Understanding the demographic characteristics of MCTA users is important for several reasons. First, it may help researchers select which app they want to use for their research, for example, an app primarily used by teenagers, or an app primarily used by those trying to conceive. Second, describing MCTA users will help to determine the generalizability of any analytic results derived from their recorded data. Third, it would help to describe the potential for selection bias when addressing a specific research question. For instance, in a study looking at the behavior of users of two different apps, Kindara users mainly resided in the USA and used the app to promote pregnancy, while Sympto users were mainly European and used the app to prevent pregnancy [46•]. This section highlights the demographic data of papers from the literature review in which the primary research objective was menstrual cycles, fertility, or contraception. Table 2 summarizes MCTAs discussed in this paper.

MCTA Users in Menstrual Cycle Research We found ten studies that describe menstrual cycle research using MCTAs. Six were descriptive studies of menstrual cycle length or ovulation, and the remaining four will be described individually. One study addressed menstrual bleeding intensity and was focused on adolescents [13]. This study required a history of regular menstrual cycles (21 to 45 days long) and menses lasting less than 8 days [13]. Since teens are more likely to have irregular cycles, and the time since menarche influences

Table 1 Primary objective of each paper included in this review and the section(s) in which it is cited

Primary objective	Publication	MCTAs included in these papers
Characteristics of MCTA users in menstrual cycle research		
To highlight decision points during the recruitment–enrollment process and the effect of modifications on enrollment numbers and demographics in a contraceptive efficacy study of an MCTA.	[12•]	Dot
To assess MCTA use for women or girls with heavy menstrual disorders.	[13]	Not named
To track premenstrual symptoms and sexually transmitted infections.	[14]	Clue
To examine an acupressure intervention on dysmenorrhea.	[15]	AKUD
To evaluate whether an MCTA was associated with reduced depression and dysmenorrhea.	[16]	Karada-no-kimochi
To assess the efficacy of an MCTA at pregnancy prevention.	[17]	Dot
To compare self-reported fertility data generated by an MCTA cohort in comparison with existing data.	[18]	Ovia
To explore variations of menstrual cycle length and ovulation day in women seeking to conceive.	[19]	Clearblue Connected Ovulation System
To clarify how MCTA data can be used to improve the ovulation prediction accuracy.	[20]	LunaLuna
To describe menstrual cycle characteristics observed with an MCTA and investigate associations of menstrual cycle characteristics with cycle length, age, and body mass index (BMI).	[21•]	Natural Cycles
Characteristics of MCTA users in fertility research		
To compare self-reported fertility data generated by an MCTA cohort in comparison with existing data.	[18]	Ovia
To evaluate the association between menstrual cycle characteristics and time to pregnancy.	[22]	Ovia
To evaluate the difference in conception rates between women trying to conceive using an MCTA ovulation test system, versus those trying without ovulation testing.	[23]	Clearblue Connected Ovulation Test System
To predict a woman's probability of becoming pregnant using data from an MCTA.	[24]	Clue
To examine the extent to which MCTAs influence fecundability in women trying to conceive.	[25]	Multiple
Characteristics of MCTA users of contraceptive apps		
To retrospectively evaluate the effectiveness of a fertility awareness–based MCTA to prevent unwanted pregnancies.	[8]	Natural Cycles
To assess whether women in a variety of settings were interested in and able to use a fertility app, learn the profile of women who chose to use an app and their purpose for using it, assess what their prior experience with family planning was, and determine how satisfied they were with the app.	[26••]	CycleBeads
To assess whether the CycleBeads app brings new users to family planning, to understand the experience of its users in Kenya, and to determine how user experience varies by the channel through which women learned about the app.	[27]	CycleBeads
Reasons users choose or continue using a particular MCTA		
To propose an approach for a period tracking app with an adaptive user interface that takes the users goal and context into account.	[28]	Multiple
To assess whether the CycleBeads app brings new users to family planning, to understand the experience of its users in Kenya, and to determine how user experience varies by the channel through which women learned about the app.	[27]	CycleBeads
To investigate how people engaging in period tracking use and experience these apps through qualitative research.	[29]	Multiple
To explore women's uses of and relationships with MCTAs.	[30]	Multiple
To determine ovulation prediction accuracy of MCTAs and published calendar methods compared with the LH surge.	[31]	Multiple
To retrospectively evaluate the effectiveness of a fertility awareness–based MCTA to prevent unwanted pregnancies.	[8]	Natural Cycles
Among women with dysmenorrhea and premenstrual syndrome: can MCTAs selected based on users' needs lead to changes in health-related factors?	[32]	Multiple
To identify women who currently use or intend to use an MCTA to prevent pregnancy, and explore their preferences and perceptions about MCTAs for this purpose.	[33]	Multiple
To develop an evidence-based evaluative framework to support a systematic examination of menstrual health concepts, and to evaluate MCTA features and functions in terms of these menstrual health concepts.	[34]	Multiple
To consider why and how women track their menstrual cycles, examining their experiences to uncover design opportunities and extend the field's understanding of personal informatics tools.	[35••]	Multiple
To examine how MCTAs targeted at female reproductive health are offering novel forms and practices of knowledge production about reproductive bodies and processes.	[36]	Multiple
To explore how health consumers use apps for health monitoring, their perceived benefits from use of health apps, and suggestions for improvement of health apps.	[37]	MCTAs were one type of included health app
To develop an MCTA for monitoring the menstrual cycle for clinical treatment planning.	[38]	Infertility Handling
Accuracy of ovulation and fertile window prediction		
To identify, describe, and evaluate mobile phone apps that purport to help users prevent unintended pregnancy.	[39]	Multiple
To identify smartphone MCTAs and evaluate their accuracy, features, and functionality.	[40]	Multiple
To review and appraise Apple Store MCTAs for ovulation prediction, and other topics related to fertility or the management of infertility.	[41]	Multiple
To determine ovulation prediction accuracy of MCTAs and published calendar methods compared with the LH surge.	[31]	Multiple
	[42]	DOT

Table 1 (continued)

Primary objective	Publication	MCTAs included in these papers
To develop a new, personalized approach of estimating a woman's most fertile days that only requires recording the first day of menses in an MCTA.		
To clarify how MCTA data can be used to improve the ovulation prediction accuracy.	[20]	LunaLuna
To assess whether the CycleBeads app brings new users to family planning, to understand the experience of its users in Kenya, and to determine how user experience varies by the channel through which women learned about the app.	[27]	CycleBeads
To examine the extent to which MCTAs influence fecundability in women trying to conceive.	[25]	Multiple
To evaluate the difference in conception rates between women trying to conceive using an MCTA ovulation test system, versus those trying without ovulation testing.	[23]	Clearblue Connected Ovulation Test System
To provide the first theoretical review and evaluation of BBT-based ovulation detection.	[43]	Not focused on one app, but on BBT in general
To retrospectively evaluate the effectiveness of a fertility awareness–based MCTA to prevent unwanted pregnancies.	[8]	Natural Cycles
To investigate the contraceptive efficacy of Natural Cycles through the perfect- and typical-use Pearl Index.	[9]	Natural Cycles
To evaluate the ability of a novel web and mobile application to identify a woman's ovulation day and fertile window, in order to use it as a method of natural birth control.	[7]	Natural Cycles
To compare self-reported fertility data generated by an MCTA cohort in comparison with existing data.	[18]	Ovia
To develop a tool to evaluate and rate fertility apps specifically designed to help couples avoid pregnancy.	[44]	Multiple
To explore differences in the use of CycleProGo between those exposed to it as a part of natural family planning (NFP) instruction versus those who find it on their own.	[45]	CycleProGo
Missing data and accuracy		
To compare the accuracy of fertile window identification with the contraceptive app Natural Cycles against the Rhythm Method and Standard Days Method (SDM).	[6]	Natural Cycles
To examine user tracking behaviors.	[46•]	Sympto, Kindara
To predict a woman's probability of becoming pregnant using data from an MCTA.	[24]	Clue
To describe menstrual cycle characteristics observed with an MCTA and investigate associations of menstrual cycle characteristics with cycle length, age, and BMI.	[21]	Natural Cycles
To propose an approach for a period tracking app with an adaptive user interface that takes the users goal and context into account.	[28]	Multiple
To investigate how people engaging in period tracking use and experience these apps through qualitative research.	[29]	Multiple
To examine how MCTAs targeted at female reproductive health are offering novel forms and practices of knowledge production about reproductive bodies and processes.	[36]	Multiple
To explore women's uses of and relationships with MCTAs.	[30]	Multiple
To consider why and how women track their menstrual cycles, examining their experiences to uncover design opportunities and extend the field's understanding of personal informatics tools.	[35••]	Multiple
To investigate the MCTA DOT's effectiveness, calculating perfect- and typical-use failure rates.	[47]	DOT
Quality assessments of MCTAs		
To identify smartphone MCTAs and evaluate their accuracy, features, and functionality.	[40]	Multiple
To develop a tool to evaluate and rate fertility apps specifically designed to help couples avoid pregnancy.	[44]	Multiple
To develop a scoring system for rating available apps for determining the fertile window and to pilot test 12 apps currently available in both German and English.	[48]	Multiple
To identify and evaluate apps intended to aid women in conception.	[49•]	Multiple

their regularity, this study selected for girls who are further from menarche or achieved regularity quickly [51]. The study found that teenagers prefer an MCTA to paper cycle tracking, but this may only be applicable to teens with regular cycles [13]. One MCTA study reported an association between sexually transmitted infection and increased premenstrual symptoms (headache, cramps, and sadness) in younger users (median age: 26) [14]. Another study used an app to deliver an acupressure intervention to German users with dysmenorrhea aged 18–34 and reported a reduction in menstrual pain over six menstrual cycles [15]. Finally, in a randomized controlled trial of workers in Japan aged 20–45, use of an MCTA was associated with reduced depression and dysmenorrhea after 3 months of use [16].

The populations from the remaining six studies are from the USA [17, 18], the UK and the USA [19], Japan [20], or a combination of the USA, UK, and Sweden [21•]. In the sixth study, the full user base includes 150 countries, 5 continents, and 8 languages, yet most users reside in the USA and Europe [46•]. These app studies primarily represent the USA and Europe. In one of these studies, most participants reported their race as White (78%) [18]. However, four menstrual cycle studies do not present other race/ethnicity data outside of country of residence [19, 20, 21•, 46•]. In total, it appears that most studies include populations of European descent or do not present race/ethnicity data at all.

Three of the menstrual cycle studies had a mean age of 30 or higher [20, 21•, 46•] and one study reported that 70% of

Table 2 Summary of the most commonly mentioned MCTAs in this review

MCTA	Primary motivation for use	Location of users in publications	Data availability/privacy	Additional information of note for research
Clue	Track cycles [14, 24]	USA [14]	Authors were employed at company that produced Clue, so they had access to data [24].	Is gender inclusive [14, 35••], has a lot of features for tracking behavior [14, 24, 40]. Most users do not log pregnancy test or other biomarker test predictive of pregnancy [24].
CycleBeads	Prevent pregnancy [26••], observe cycle [27]	Egypt, Ghana, India, Jordan, Kenya, Nigeria, and Rwanda [26••] Kenya [27]	Signing up for the app allows for de-identified user data to be utilized for research for these specific researchers [26••]. Requires in-app recruitment for research [27]	
DOT	Prevent pregnancy [47]	USA [47]	Requires in-app recruitment for research [47]. A written consent for was required after enrollment [31].	Effective for women with longer and shorter than average cycle lengths [42]
Kindara	Promote pregnancy [46•]	Mostly USA [46•]	N/A	
LunaLuna	Plan pregnancy [20]	Japan [20]	Authors were employed at company that produced LunaLuna, so they had access to data [20].	
Natural Cycles	Prevent pregnancy [5–9, 21, 44]	Sweden and Switzerland [7]; UK, USA, and Sweden [6, 21]; Sweden [5, 8, 9]	Signing up for the app allows for de-identified user data to be utilized for research [8].	Certified in Europe [43] and the USA [50] as a form of contraception
Ovia	Plan pregnancy [18, 22]	USA [18, 22]	Signing up for the app allows for de-identified user data to be utilized for research [18, 22].	More representative of a subfertile population [18]
Sympto	Prevent pregnancy [46•]	Mostly Europe [46•]	N/A	

users were aged 25–34 [18]. One study did not provide any demographic data [19]. While the literature is small, most MCTA menstrual cycle studies included older users with adolescents and young adults less well-represented. In one study, users tended to have a college degree or higher education (71%) [18]. Many studies did not report the education level of their participants [19, 20, 21•, 46•].

Body mass index was centered on the normal range for two of the studies (mean BMI = 23) [21•, 46•], while one included 31% overweight or obese users [18]. In the latter study, 30% of users were missing BMI information [18]. Some studies did not report the BMI distribution [19, 20]. In total, MCTA menstrual cycle research shows varying distributions of BMI although other MCTA studies did not report BMI data at all.

MCTA studies sometimes impose limits on the average cycle length or regularity of the participants in their analyses. For example, one study observed cycles of 23–67 days but excluded those that were 1.5 times longer or shorter than the user's reported cycle length (approximately 70% of participants reported a cycle length of 25 to 30 days) [19]. Some studies have included a wide range of cycle lengths: 19 to 60 days [18], 10 to 90 days [21•], 20 to 45 days [20]. Another study based on fertility awareness methods did not have strict exclusion criteria based on length, but did require

that cycles > 40 days did not have any mid-cycle bleeding and that the total cycle length was 4 days longer than the number of bleeding days reported [46•]. Overall, menstrual cycle research using MCTAs has incorporated a wide range of average cycle lengths.

In conclusion, MCTA menstrual cycle studies focus on users of White race or European residence, although some studies did not describe their study sample. MCTA menstrual cycle studies tend to include older users, and researchers interested in younger populations may need to seek out an app that targets that population. MCTA-based studies include a wide range of BMI and cycle lengths. Exclusions of cycle lengths should be carefully considered with regard to how this may balance misclassification and generalizability. MCTA studies have an opportunity through their large user base to describe menstrual cycle characteristics across a diverse sample. Without the careful reporting of demographic characteristics, it is a challenge to ascertain if samples are representative of the MCTA user population, or of wider country or even global populations. We suggest that publications using MCTA data for menstrual cycle research thoroughly describe their sample's demographic characteristics, and the peer-review process should request this. This is fundamental for understanding generalizability and potential for selection bias.

MCTA Users in Fertility Research Two of the apps that were previously described as contributing to menstrual cycle research [18, 19] also conduct research on the probability of conception; these include the Clearblue Connected Ovulation Test System (a combination of app and ovulation tests) [23] and Ovia Fertility [22]. One additional study using data from the Clue app has also been used to develop a model for conception; however demographics of the user base were not described [24]. The two studies with demographic information were based in either the UK [23] or the USA [22], and both samples had a mean age of 30 and a wide range of BMI (mean = 26 and SD = 5 [23] or 43% overweight or obese [22]). MCTAs have the potential to over-select for people who are subfertile, for example, women who know they are subfertile are more likely to use an app to help them time their intercourse or track their cycles to facilitate conception. However, data to evaluate this issue are limited. Lower conception rates and a higher rate of endometriosis than those of the general population were found in one MCTA study [18]. In a cohort study of women attempting to conceive, the prevalence of subfertility was not higher in MCTA users [25]. However, users of "selected" MCTAs and "other" MCTAs were more likely to take folic acid than non-users (83.3% and 73.7% vs 66.8%) and were less likely to have recently used contraceptive hormones (37.4% and 34.6% vs 44.0%) [25]. The fertility profile of MCTA users should be further investigated especially when the research objective is fertility related because this is important for generalizing the results.

MCTA Users of Contraceptive Apps MCTAs can be used to prevent pregnancy, and some of the previously described apps used for fertility research or menstrual cycle research can also be used by women who wish to avoid pregnancy. The characteristics of women who use an app for contraception may differ from those who choose to use the same app to aid conception, so we focus on describing users studied in MCTA contraception research in this section. The app with the most published research, the Natural Cycles contraceptive app, is FDA-approved as a contraceptive [50] and provided data in the previous section on menstrual cycle research. Users of the app were for the most part young (aged 20–35), of a healthy BMI, and from Sweden [8]. The remaining two studies that specifically address contraceptive research focused on low-income countries [26••, 27] and used the Cycle Beads contraceptive app. The Cycle Beads users were young (69.7% are 18–25), 45.4% were students, 27% were in a relationship but not married, and 49.7% attended/completed university/postgraduate school [26••]. One important observation from the Cycle Beads studies was that a third of the users were not previously using another form of contraceptive, suggesting that MCTAs could fill an unmet need for contraception [26••].

Reasons Users Choose or Continue Using a Particular MCTA

A person may select a specific MCTA for a variety of reasons, including accuracy, a referral from a friend, inclusiveness, and tracking features. MCTA users continue to engage with an app if they are satisfied with their experience. One important consideration for choosing and continuing an MCTA is its perceived accuracy. Those planning to use an app for contraception rated accurate ovulation prediction as highly important (90.7%). Users lose trust in an MCTA [28] and discontinue use when menstrual cycle milestones are miscalculated, which can be due to the method of calculation or to user characteristics, such as irregular cycles [27, 29, 30]. For contraceptive MCTAs, the more accurate methods include a wider fertile window [31], yet discontinuation is also more likely with a wide fertile window as it allows for fewer days with unprotected intercourse [8]. This highlights users' desire for an MCTA that accurately identifies the exact fertile window which allows for more days of unprotected intercourse. Research using MCTAs should consider that the sample is selected for users who have been satisfied with the accuracy of the app.

Users refer their friends to a well-liked MCTA [27, 32, 33]. Fifty percent of Cycle Bead's users were referred by a friend [27]. Furthermore, 68.4% of 1000 survey respondents reported that word of mouth was a somewhat or very important reason for choosing an app [33]. This suggests that the users of a given app may be clustered in meaningful ways for epidemiologic research.

Using MCTAs can be discouraging for users whose identities are not represented by the app. There is a tendency for MCTAs to focus on fertility, and users who are outside of the gender binary, are infertile, are new menstruators, are celibate, or are not heterosexual, may feel excluded [29, 34]. My Period Tracker and Glow have received feedback for being gendered, and therefore are more likely to select for cis-gendered users [35••]. Users describe Clue as gender neutral and not focused on pregnancy, despite imagery suggesting a male partner [35••].

MCTAs are used to observe cycles independent of fertility. In two online surveys, most respondents used MCTAs to track their cycle [28, 30]. Likewise, in two qualitative studies, themes of cycle observation for health purposes emerged, regardless of fertility goals [29, 35••]. Another qualitative study found that people trying to conceive reported learning more about personal fertility patterns from using MCTAs [36]. This suggests that app users who track their cycles are interested in learning about their health, and apps that address this motivation may better retain users. Furthermore, capitalizing on this interest may lead to more complete data in MCTA datasets. Users with irregular cycles find it beneficial and report better control of their condition when using an MCTA [37]. Keeping

track of symptoms could be useful for both providers and patients for managing menstrual disorders [38]. However, in a study with 72 participants who had dysmenorrhea and premenstrual syndrome (PMS), only 24% said it helped them understand PMS patterns [32]. This points to an area in which MCTAs can grow and potentially retain a group of women who otherwise may not continue with an app.

Accuracy of Ovulation and Fertile Window Prediction

Here, we review the published literature that has evaluated the accuracy of fertile window and ovulation prediction of MCTAs. Many MCTAs are based on fertility awareness methods and some are marketed for pregnancy prevention [39]. For some MCTAs, it is unclear whether health professionals or the published scientific literature have contributed to their development [40, 41]. A calendar method refers to tracking a menstrual cycle on a calendar, and different MCTAs have different variations of this method. Apps that used one of four calendar methods all had an ovulation day prediction accuracy of lower than 90%, ranging from 17 to 89% [31]. The wider the estimated fertile window, the more likely the app was to identify the true fertile days [31]. However, a wider fertile window has implications for user adherence, as less days with unprotected intercourse will be allowed. In a different study, the MCTA Dot was found to be more effective than other calendar-based MCTAs for users with non-average cycle lengths because of the app's ability to change predictions based on an individual's data [42]. Still, the authors note that Dot is most effective for people with regular cycles [42]. Similarly, the LunaLuna app was compared with existing calendar-based methods (Ogina and HCL) and was found to be more accurate at predicting ovulation, particularly as the number of menstrual cycles per user increased, and at the extremes of cycle length [20]. The standard day method, a method used for the app Cycle Beads, was only effective for users with cycles between 26 and 32 days and does not adjust predictions over time [27, 42]. The standard day method, a type of calendar method, is traditionally used with a color-coded strand of beads that is representative of a menstrual cycle and identifies days 8 through 19 as fertile [27].

Ovulation testing in combination with an MCTA may improve accurate identification of the fertile window leading to increased conception rates. In a prospective cohort study of women trying to become pregnant (the Pregnancy Study Online or PRESTO), women who used an MCTA had a higher probability of conceiving and the associations were stronger when used with fertility indicators such as basal body temperature, ovulations tests, and cervical fluid [25]. Similarly, in another study, using ovulation tests in combination with a study app was associated with twice the odds of conception compared with only using the app [23]. On its

own, basal body temperature methods are sensitive to misclassification due to multiple temperature peaks in a cycle (not necessarily due to multiple ovulations) and fever [43]. The Natural Cycles contraceptive app incorporates basal body temperature measurements into the algorithm that predicts the fertile window. In two studies of the accuracy of the Natural Cycles app (over 4000 users), the Pearl Index typical-use score (the number of contraceptive failures per 100 person-years of exposure) was 7.0 [8, 9, 52], which was a conservative estimate [7–9], and was still an improvement on other fertility awareness–based methods which have a typical-use Pearl Index of 24 [8].

The Ovia Fertility app allows users to input self-detected ovulation information, and the resulting data have been compared with published studies of ovulation timing to determine their consistency [18]. For example, the probability of pregnancy was highest on the 5 days before and on the day of ovulation, which is consistent with previously published biomarker-based studies of the fertile window [18]. Symptoms of ovulation such as cervical fluid changes also changed around the estimated ovulation day. Taken together, the authors suggest that the Ovia Fertility app contains accurate ovulation and fertile window data. While ovulation testing may improve the accuracy of fertile window identification, it does increase user burden and may decrease user engagement or lead to discontinuation.

The accuracy of MCTAs may also be improved by incorporating an educational or training component [44]. A study on the app CycleProGo found fewer missing data and increased long-term use when users took a natural family planning course [45]. Researchers should determine whether an app they are considering employing includes a training component, and the implications of that training for accuracy and user burden.

Missing Data and Accuracy Missing data can affect MCTA prediction accuracy [6]. MCTA studies acknowledge that missing data are a problem [46•], but do not always fully describe the frequency of missingness or the implications of missing data for interpretation [24]. Of 1.4 million eligible menstrual cycles, ovulation could not be assigned in the Natural Cycles app in 665,603 cycles, which may have been largely due to missing temperatures: 75% of the cycles without a day of ovulation were missing temperatures for at least half of the cycle [21•]. Users who record more intercourse track more data in the Kindara and Sympto apps; these users are also trying to conceive (40% of cycles had recordings every single day when the user was trying to conceive) [46•]. This suggests that a goal of conception encourages users to record data. MCTAs are designed to reduce missingness by incorporating reminders, which users report as helpful [53], but also annoying [53] and patronizing [36]. Users have a clear preference for making reminders optional [30, 35••].

MCTAs with a low burden for users could promote data entry and improve app accuracy [47]. However, even with innovations designed to target simplicity, over half of 196 MCTA users in an online survey cited app complexity as a reason for switching, and 22% reported having switched apps before [28]. Future research and innovation will be necessary to address missingness in MCTA data. Moreover, innovations must balance the decrease in missingness with the increase in user annoyance that may lead to app switching. For epidemiologists, a longitudinal cohort might not be feasible if many users switch apps.

Quality Assessments of MCTAs

Previous studies have evaluated MCTAs on disparate measures of "quality." While accuracy of identifying the fertile window and scientific quality are typically included [40, 44, 48, 49•], other criteria include access to technical support [40, 44], password protection [40], privacy policy clarity [54], third party advertising [40], cost [44], and ease of use [44, 49•]. However, perception of quality may differ based on intended use of the app, and studies comparing quality have been inconsistent in their scoring criteria [44, 48]. For example, one study developed ten criteria applicable to apps in general (other than accuracy), and another developed eight criteria related to fertility specifically [44, 48]. Thus, quality assessments of apps are difficult to compare, having been based on different criteria. Efforts have been made to develop a standardized method of evaluating quality for MCTAs. Moglia et al. adapted the APPLICATIONS Scoring System, which is used to evaluate mobile apps in general, for use with MCTAs; one additional study followed suit [40, 49•]. None of the reviewed apps was considered perfect by the applied scoring system in either study. Although these two studies applied similar criteria, they did not review the same applications so direct comparisons cannot be made [40, 49•]. One of these two studies reported the highest score for Ovia Fertility Period Tracker with 13/15 points and the lowest to Pregnancy Tracker Baby Center with 9/15 points [49•], while in the second of these studies, Clue received the highest score with 13/15 points and Free Girl Cal received the lowest score with 6/15 points [40]. Reasons for lower scores included an unclear link to published scientific research, and a potential lack of health professional involvement in their design [49•]. Both Clue and Ovia received high scores for addressing all reviewed areas except "involvement of a health professional" [40, 49•]. Ovia scored particularly well in comprehensiveness, which was a measure of the diversity of tracking features and educational information [49•]. Clue included "other" features that set it apart including having a medical disclaimer and health education material, data security features like a backup and the ability to export data, availability in Spanish, custom reminders, the ability to track many menstrual characteristics,

and alerts for the next menses and the fertile window [40]. These assessments show that "quality" can include many facets of apps beyond scientific accuracy and will therefore depend somewhat on the priorities of the user or the researcher. Furthermore, if a user shifts their goals, for example, from wanting to avoid pregnancy to trying to achieve it, the relative quality of the app they are using may change, which the adapted APPLICATIONS Scoring System does not account for [49•].

Discussion

The objective of this review was to provide an epidemiologic perspective on the current MCTA literature. We found that the existing literature fell into four relevant categories which are described in our "Results" section: characteristics of MCTA users, reasons women use or continue using MCTAs, accuracy of identifying ovulation or the fertile window, and comparisons of MCTAs across differing measures of "quality." These four categories inform our understanding of the potential for selection bias and misclassification when using MCTA data independent of the research goals of the currently published literature. We included 48 studies in this review—a small but burgeoning literature. Regarding the first category, we found a tendency for published MCTA menstrual cycle studies to include users of White race or European residence, while most included a wide range of BMI and cycle lengths. Several studies did not report the demographic information of their users at all. MCTA-based studies of fertility should investigate the prevalence of subfertility in their user base to determine if MCTAs designed to aid conception are used predominantly by those with fertility concerns. Fewer studies exist describing characteristics of individuals who use MCTAs for contraception, despite there being some evidence that MCTAs can fill unmet contraceptive needs.

We found that the reasons that people use MCTAs vary and can change over time which is consistent with one other review article [55•]. Users who feel an MCTA meets their needs are more likely to continue to use it and to recommend it to others. MCTAs that employ non-binary or non-gendered environments will appeal to these underrepresented groups. Additionally, cycle tracking, independent of pregnancy planning, is an important motivation for using MCTAs. Features that appeal to MCTA users who are not planning a pregnancy, such as a focus on tracking symptoms rather than ovulation, may improve their satisfaction with the app, and therefore their data completeness.

The accuracy of apps depends upon the algorithm used, and apps may benefit from the inclusion of biological monitoring rather than just calendar methods. Missing data can contribute to inaccuracy of the MCTA's prediction algorithm, as well as selection or generalizability issues; however, few

MCTA-based studies have described the potential impact of missing data on their results, and this is likely to be an important area of future research.

Several reviews have already been conducted that have identified MCTAs that are of higher quality based on both scientific and non-scientific criteria. These assessments show that "quality" can include many facets of apps beyond scientific accuracy and will therefore depend somewhat on the priorities of the user or the researcher. Without a standardized method to evaluate quality, users, providers, and researchers will need to make their own decisions about an app on an individual basis.

The literature on incorporating MCTAs into research is currently small, but the availability of large datasets (hundreds of thousands of menstrual cycles) and rich data collection (daily diaries, questionnaires, geolocation data) will lead to increasing interest from researchers. Our recommendations for future research using MCTAs include describing the diversity of the user base and assessing how the MCTA addresses the needs of its users. Different strategies have been applied to promote user engagement, which is important for reducing missing data and maintaining longitudinal cohorts. Furthermore, MCTA studies should thoroughly describe their sample's demographic and behavioral characteristics and previous fertility experience, and the peer-review process should request this. This is fundamental for understanding generalizability and the potential for selection bias. Finally, the choice of MCTA for research will depend on the research question and the ability of the MCTA to address the population, exposure, or health condition of interest.

Limitations

A limitation of this review is the inability to directly compare MCTAs because studies used different validation methods to evaluate accuracy. Furthermore, not all of the literature disclosed the name of the particular app being studied. Literature is still focusing on evaluating the algorithms and methods that an MCTA uses, and most apps do not share their method. This compounds the difficulty of direct app comparisons.

Conclusion

MCTAs are an important tool for the advancement of epidemiologic research on menstruation. MCTA studies should describe the demographic and behavioral characteristics of their user base and the patterns of missing data. Describing the motivation for using MCTAs over time and validating the data collected should be prioritized in future research. The ubiquity of MCTAs provides an opportunity for epidemiologic research that is demographically diverse. At the same time,

MCTA data could also be leveraged to address an underrepresented health condition or exposure.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance
•• Of major importance

1. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. J Med Internet Res. 2017;19:1–9.
2. Fox S, Duggan M. Mobile Health. Pew Research Center: Internet, Science & Tech. 2012. Available from: https://www.pewresearch.org/internet/2012/11/08/mobile-health-2012/. Accessed 20 Apr 2020
3. Lupton D. Quantified sex: a critical analysis of sexual and reproductive self-tracking using apps. Cult Health Sex. 2015;17:440–53.
4. Ford EA, Roman SD, McLaughlin EA, Beckett EL, Sutherland JM. The association between reproductive health smartphone applications and fertility knowledge of Australian women. BMC Womens Health. 2020;20:1–10.
5. Bull J, Rowland S, Lundberg O, Berglund-Scherwitzl E, Gemzell-Danielsson K, Trussell J et al. Typical use effectiveness of Natural Cycles: postmarket surveillance study investigating the impact of previous contraceptive choice on the risk of unintended pregnancy. BMJ Open. 2019;9:e026474. https://doi.org/10.1136/bmjopen-2018-026474.
6. Kleinschmidt TK, Bull JR, Lavorini V, Rowland SP, Pearson JT, Scherwitzl EB, et al. Advantages of determining the fertile window with the individualised Natural Cycles algorithm over calendar-based methods. European Journal of Contraception and Reproductive Health Care. 2019;24:457–63.
7. Berglund Scherwitzl E, Lindén Hirschberg A, Scherwitzl R. Identification and prediction of the fertile window using NaturalCycles. Eur J Contracept Reprod Health Care. 2015;20:403–8.
8. Berglund Scherwitzl E, Gemzell Danielsson K, Sellberg JA, Scherwitzl R. Fertility awareness-based mobile application for contraception. Eur J Contracept Reprod Health Care. 2016;21:234–41.
9. Berglund Scherwitzl E, Lundberg O, Kopp Kallner H, Gemzell Danielsson K, Trussell J, Scherwitzl R. Perfect-use and typical-use Pearl Index of a contraceptive mobile app. Contraception. 2017;96:420–5.
10. Ericson (2015) Mobility report: on the pulse of the networked society.
11. Vaz F, Silva RR, Bernardino J. Using data mining in a mobile application for the calculation of the female fertile period. In: A. F, J. F, editor. 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2018. Rua Pedro Nunes, Coimbra, Portugal: SciTePress, Polytechnic of Coimbra, ISEC; 2018. p. 359–66.

12.• Shattuck D, Haile L, Simmons R. Lessons from the dot contraceptive efficacy study: analysis of the use of agile development to improve recruitment and enrollment for mhealth research. JMIR mHealth and uHealth. 2018;6:e99. https://doi.org/10.2196/mhealth.9661. **This paper demonstrates the importance of a strategic and iterative recruitment technique in order to have a diverse study sample.**

13. Jacobson AE, Vesely SK, Haamid F, Christian-Rancy M, O'Brien SH. Mobile application vs paper pictorial blood assessment chart to track menses in young women: a randomized cross-over design. J Pediatr Adolesc Gynecol. 2018;31:84–8.

14. Alvergne A, Vlajic Wheeler M, Hogqvist Tabor V. Do sexually transmitted infections exacerbate negative premenstrual symptoms? Insights from digital health. Evolution, medicine, and public health. 2018;2018:138–50.

15. Blödt S, Pach D, Eisenhart-Rothe S v, Lotz F, Roll S, Icke K, et al. Effectiveness of app-based self-acupressure for women with menstrual pain compared to usual care: a randomized pragmatic trial. Am J Obst Gynecol. 2018;218:227.e1–227.e9.

16. Song M, Kanaoka H. Effectiveness of mobile application for menstrual management of working women in Japan: randomized controlled trial and medical economic evaluation. J Med Econ. 2018;21:1131–8.

17. Simmons R, Shattuck D, Jennings V. Assessing the efficacy of an app-based method of family planning: the dot study protocol. JMIR Res Protocols. 2017;6:e5. https://doi.org/10.2196/resprot.8829.

18. 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. Fertility and Sterility. 2019;112:450–457.e3.

19. Soumpasis I, Grace B, Johnson S. Real-life insights on menstrual cycles and ovulation using big data. Human Reproduction Open. 2020;2020:1–9.

20. Sohda S, Suzuki K, Igari I. Relationship between the menstrual cycle and timing of ovulation revealed by new protocols: analysis of data from a self-tracking health app. J Med Internet Res. 2017;19:e391.

21.• 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 Digital Med. 2019;2:83. **This paper has menstrual cycle data from more than 600,000 menstrual cycles with wide age and BMI ranges, showing how large the potential sample size can be with MCTAs.**

22. Bradley D, Landau E, Jesani N, Mowry B, Chui K, Baron A et al. Time to conception and the menstrual cycle: an observational study of fertility app users who conceived. Hum Fertil. 2019;1–9.

23. 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 Women's Health. 2020;29:84–90.

24. Liu B, Thomas D, Shi S, Symul L, Leskovec J, Wu Y, et al. Predicting pregnancy using large-scale data from a women's health tracking mobile application. In: 2019 World Wide Web Conference, WWW 2019. Stanford, United States: Association for Computing Machinery, Inc, Dept. of Computer Science; 2019. p. 2999–3005.

25. 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 (Oxford, England). 2020;35: 2245–52.

26.•• Haile LT, Fultz HM, Simmons RG, Shelus V. Market-testing a smartphone application for family planning: assessing potential of the CycleBeads app in seven countries through digital monitoring. mHealth. 2018;4:27–27. **This paper highlights the potential for MCTAs to fill an unmet need for contraceptive in low-income**

countries. Additionally, the authors find that a social media campaign is an effective way to recruit a diverse sample.

27. Shelus V, Ashcroft N, Burgess S, Giuffrida M, Jennings V. Preventing pregnancy in Kenya through distribution and use of the CycleBeads mobile application. Int Perspect Sex Reprod Health. 2017;43:131–41.

28. Bretschneider RA. A goal- and context-driven approach in mobile period tracking applications. In: Antona M, Stephanidis C, editors. Universal access in human-computer interaction: access to learning, health and well-being, uahci 2015, pt III; 2015. p. 279–87.

29. Levy J, Romo-Avilés N. "A good little tool to get to know yourself a bit better": a qualitative study on users' experiences of app-supported menstrual tracking in Europe. BMC Public Health. 2019;19:1213.

30. Gambier-Ross K, McLernon DJ, Morgan HM. A mixed methods exploratory study of women's relationships with and uses of fertility tracking apps. Digital Health. 2018;4:205520761878507.

31. Johnson S, Marriott L, Zinaman M. Can apps and calendar methods predict ovulation with accuracy? Curr Med Res Opin. 2018;34: 1587–94.

32. Lee J, Kim J. Can menstrual health apps selected based on users' needs change health-related factors? A double-blind randomized controlled trial. J Am Med Inform Assoc. 2019;26:655–66.

33. Starling MS, Kandel Z, Haile L, Simmons RG. User profile and preferences in fertility apps for preventing pregnancy: an exploratory pilot study. mHealth. 2018;4:21–21.

34. Eschler J, Menking A, Fox S, Backonja U. Defining menstrual literacy with the aim of evaluating mobile menstrual tracking applications. Computers, Informatics, Nursing : CIN. 2019;37:638–46.

35.•• Epstein DA, Lee NB, Kang JH, Agapie E, Schroeder J, Pina LR, et al. Examining menstrual tracking to inform the design of personal informatics tools. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17. New York, New York, USA, United States: ACM Press; 2017. p. 6876–88. **This paper provides qualitative data valuable for understanding user experiences with MCTAs.**

36. Hamper J. 'Catching ovulation': exploring women's use of fertility tracking apps as a reproductive technology. Body & Society. 2020;26:3–30. https://doi.org/10.1177/1357034X19898259.

37. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. PLoS One. 2016;11:e0156164.

38. Vanya M, Jakó M, Füle G, Fidrich M, Surányi A, Bitó T, et al. Use of infertility handling among women of reproductive age. International Summit on eHealth. 2017;360°, 2016 181 LNICST: 497–501.

39. Mangone ER, Lebrun V, Muessig KE. Mobile phone apps for the prevention of unintended pregnancy: a systematic review and content analysis. JMIR mHealth and uHealth. 2016;4:e6.

40. Moglia ML, Nguyen HV, Chyjek K, Chen KT, Castaño PM. Evaluation of smartphone menstrual cycle tracking applications using an adapted applications scoring system. Obstet Gynecol Surv. 2016;71:713–4.

41. Zwingerman R, Chaikof M, Jones C. A critical appraisal of fertility and menstrual tracking apps for the iPhone. J Obstet Gynaecol Canada. 2020;42:583–590.

42. Li D, Heyer L, Jennings VH, Smith CA, Dunson DB. Personalised estimation of a woman's most fertile days. European Journal of Contraception and Reproductive Health Care. 2016;21:323–8.

43. Handel P, Wahlstrom J. Digital contraceptives based on basal body temperature measurements. BIOMEDICAL SIGNAL PROCESSING AND CONTROL. 2019;52:141–51.

44. 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 : JABFM. 2016;29:508–11.

45. Manhart M. A comparison of user behaviors for a fertility-tracking app: does training in an NFP method improve persistence and use? The Linacre Quarterly. 2019;87:53–59.

46.• Symul L, Wac K, Hillard P, Salathé M. Assessment of menstrual health status and evolution through mobile apps for fertility aware- ness. npj Digital Med. 2019;2:64. **This study is selected as impor- tant because of the focus on capturing user tracking behavior, which is key to understanding in order to promote good-quality data from MCTAs.**

47. Jennings V, Haile LT, Simmons RG, Spieler J, Shattuck D. Perfect- and typical-use effectiveness of the Dot fertility app over 13 cycles: results from a prospective contraceptive effectiveness trial. Eur J Contracept Reprod Health Care. 2019;24:148–53.

48. Freis A, Freundl-Schütt T, Wallwiener L-M, Baur S, Strowitzki T, Freundl G, et al. Plausibility of menstrual cycle apps claiming to support conception. Front Public Health. 2018;6:98.

49.• Hutcherson T, Cieri-Hutcherson N, Donnelly P, Feneziani M, Grisanti K. Evaluation of mobile applications intended to aid in conception using a systematic review framework. Annals of Pharmacotherapy. 2019;54:178–186. **This paper scored MCTAs based on an adapted tool for evaluation of applications, thus contributing to the literature on evidence-based strategy for comparing MCTAs.**

50. U.S. Food and Drug Administration. FDA allows marketing of first direct-to-consumer app for contraceptive use to prevent pregnancy. 2018. Available from: https://www.fda.gov/news-events/press- announcements/fda-allows-marketing-first-direct-consumer-app- contraceptive-use-prevent-pregnancy. Accessed 14 Oct 2020.

51.• ACOG. Menstruation in girls and adolescents: using the menstrual cycle as a vital sign. https://www.acog.org/clinical/clinical- guidance/committee-opinion/articles/2015/12/menstruation-in- girls-and-adolescents-using-the-menstrual-cycle-as-a-vital-sign. Accessed 4 Apr 2020.

52. Medical Dictionary Pearl Index. https://medical-dictionary. thefreedictionary.com/Pearl+index. Accessed 20 Apr 2020.

53. Levy J. "It's your period and therefore it has to be pink and you are a girl": users' experiences of (de-)gendered menstrual app design. In: 4th Conference on Gender and IT, GenderIT 2018. Granada, Spain: Association for Computing Machinery, University of Granada; 2018. p. 63–5.

54. Fowler L, Gillard C, Morain S. Readability and accessibility of terms of service and privacy policies for menstruation-tracking smartphone applications. Health Promotion Pract. 2020;21:679– 683.

55.• Earle S, Marston H, Hadley R, Banks D. Use of menstruation and fertility app trackers: a scoping review of the evidence. BMJ Sex Reprod Health. 2020. https://doi.org/10.1017/ CBO9781107415324.004. **This literature review strengthens our findings because of the similar conclusions reached by reviewing a similar body of literature.**