

Reliability of Cycle Applications for Pregnancy Planning and Contraception: A Systematic Review

Isabell Rabe, DMD, MPH, and Jan P. Ehlers, DVM, MA

Abstract

Objective: To show the effectiveness of cycle applications in both areas of application—contraception and intended pregnancy.

Methods: A systematic review based on the PubMed and Google Scholar databases, with the addition of a hand search, was conducted from May 11, 2023, through April 11, 2024, to objectively answer this question. Of 1539 sources with matching search terms, 19 sources remained after checking for inclusion criteria according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses scheme. These were analyzed according to an evaluation scale regarding their quality in various areas. The average quality ratings and pregnancy probabilities of the studies were compared.

Results: Comparability within and between the subquestions was hardly possible owing to different presentation of results, bias risks, and mostly uncontrolled study designs. Applications for those wishing to become pregnant provided better quality ratings in some cases. There were indications that cycle applications shorten the time to achieving a desired pregnancy in cases of reduced fertility. In addition, some seem to have a similar contraceptive safety as the contraceptive pill but require significantly more compliance.

Conclusion: Independent, controlled studies with a diverse clientele of test subjects are necessary for a scientific classification. In addition, social, structural, and political adjustments are needed to enable individuals to make informed decisions about the use of cycle and fertility applications.

© 2025 THE AUTHORS. Published by Elsevier Inc on behalf of Mayo Foundation for Medical Education and Research. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>) ■ Mayo Clin Proc Digital Health 2025;3(4):100239

The compatibility of contraceptives is becoming increasingly important, especially for young people.¹ In the German health care system, these informed and empowered individuals with a desire for alternatives to hormonal contraception often encounter practitioners who are fundamentally sceptical about information obtained from the internet and who attach great importance to contraceptive safety, which is why concerns about possible side effects tend to be met with a placatory response.^{2,3} Consequently, patients frequently look for alternative, self-determined contraceptive methods, just as couples with an unfulfilled desire to have children look for support. Fertility applications for determining the fertile period are attracting customers for both areas of application. However, to the best of our knowledge, there is no legal classification in Germany that is safe for users.⁴ The Association of the

Medical Societies in Germany has established a guideline addressing nonhormonal contraception, which was published in 2024.⁵ The varying effectiveness of various methods of natural family planning were mentioned, as well as the lack of sufficient evidence supporting the effectiveness of contraception applications, regardless of their symptothermal basis. It is not uncommon for users to be misled by the false sense of security that is created by relying on the known certifications of the applications in question.⁶ However, according to Frank-Herrmann et al,⁷ an approval of some contraceptive applications by Conformance Européenne (European Conformity) classification, Technische Überwachungsverein (Technical Inspection Association) seal, or US Food and Drug Administration is not a guarantee of their safety of use. This is because these organizations do not test their effectiveness in their own studies; instead, they rely



From the Didactics and Educational Research in Healthcare, Medicine Department, Faculty of Health, Witten/Herdecke University, Germany (I.R., J.P.E.); and Public Health, Professional School, Leuphana University Lueneburg, Germany (I.R.).

on the information and documents provided by the manufacturers. Furthermore, numerous applications are not designed for contraceptive purposes but are used for this function based on the stated fertile period.⁷

Natural Family Planning

The term natural contraception is not always unambiguous for users and professionals alike.⁸ According to the latest guideline,⁵ in order to achieve a high level of contraceptive effectiveness, only symptothermal methods with a high level of effectiveness should be recommended. One evidence-based method is marketed under the brand name Sensiplan.⁹ The beginning and end of the fertile phase can be determined by observing the hormonally triggered, ovulation-related rise in basal body temperature and the consistency and color of the cervical mucus (or the consistency and position of the cervix), through a double check.¹⁰ However, it should be noted that other measurement times or types, alcohol, illnesses, physical activity, traveling, heat, or hypothermia can influence this.¹¹ Other factors that may influence the menstrual cycle include physiological cycle fluctuations¹² and the impact of age, body mass index (BMI), and ethnicity on the cycle.¹³ Most users express a preference for evidence-based contraceptive applications.^{14,15} However, although many applications claim to predict fertile days and thus be usable as a safe contraceptive method,¹⁶ only a few are based on evidence.^{14,15}

Possible Ineffectiveness

Cycle applications provide users with a high degree of empowerment. However, users frequently lack the requisite health literacy to reliably categorize the information they receive, for example, the reliability of the specified fertile period, which can result in unintended pregnancies.¹⁷ Furthermore, numerous cycle applications lack substantial medical sources and are not developed by medical professionals,¹⁸ and the prediction is often inaccurate.^{19,20}

Objective and Research Question

To show the effectiveness of cycle applications in both areas of application, the research

question was as follows: “How reliable is the prediction of the fertile window by menstrual cycle applications for users who are planning a pregnancy by using such an application, and how reliable is the prediction of the fertile window by menstrual cycle applications for users who are using such an application for contraception?”

METHODS

The method of the systematic review was selected to ensure the greatest possible objectivity in answering the research question.²¹ This systematic review was conducted from May 11, 2023, through April 11, 2024. This was achieved by minimizing potential bias through a meticulous, preplanned procedure in accordance with the *Cochrane Handbook for Systematic Reviews of Interventions*. To guarantee a systematic and reproducible search for pertinent sources, the literature search was meticulously aligned with the criteria set forth in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement.²²

As inclusion criteria, we searched for studies that dealt with smartphone cycle applications that support users based on the cycle information entered and the fertile time calculated from it, either for contraception or in the event of pregnancy. Because both achieving pregnancy and contraception based on cycle monitoring are closely linked to determining the correct time of ovulation, the broader concept of the time of ovulation was also included in the inclusion criteria.

The exclusion criteria were related areas with a different focus as follows:

- Studies on animals, microorganisms, plants, or fungi
- Applications that record the measurement of luteinizing hormone in urine, for example, as the main measurement parameter
- Stand-alone cycle computers without applications
- Applications for preconception care
- Applications for pure knowledge transfer without cycle evaluation
- Applications that provide information on pregnancy (eg, visualization of growth)
- Cycle monitoring only in relation with certain diseases (eg, sexually transmitted

diseases, premenstrual syndrome and premenstrual dysphoric disorder, polycystic ovary syndrome, diabetes, and tumors)

- Cycle monitoring only in relation with certain medications
- Applications specifically for people undergoing fertility treatment
- Applications that are explicitly aimed at people in menopause or puberty
- Studies that summarize applications and do not break them down individually by name
- (Partial) publications that refer to a study that has already been included
- Studies that examined applications purely technically on hypothetical cycles

Based on these defined inclusion and exclusion criteria, the first database search was conducted in PubMed (May 12, 2023) and Google Scholar (June 9, 2023), supplemented by a hand search and repeated search until the end of the study. The search terms included the keywords *cycle application* and *contraception*, as well as *cycle application* and *desire to have children* (and the German translations thereof). The languages selected for the search were German and English. As the term *cycle tracking* applications only emerged in 2015,²³ no temporal constraints were imposed on the search. The databases searched and the respective result figures are shown in [Supplemental Appendix 2](#) for reference (available online at <https://www.mcpdigitalhealth.org/>). [Supplemental Appendix 3](#) (available online at <https://www.mcpdigitalhealth.org/>) presents the search generation process for PubMed search.

This was followed by duplicate removal and screening by title and abstract according to the inclusion and exclusion criteria in accordance with the PRISMA statement as shown in the PRISMA Checklist in [Supplemental Appendix 1](#). Full texts were searched for the remaining studies, and these were also screened individually by both authors ([Figure 1](#)).²² The included publications identified in this way were analyzed in a data extraction table [Figure 2](#)^{24–45} using previously defined quality characteristics and then evaluated in tables on the topics contraception, pregnancy, and time of ovulation according to a defined quality assessment key. The resulting average quality

scores and pregnancy probabilities were then compared.

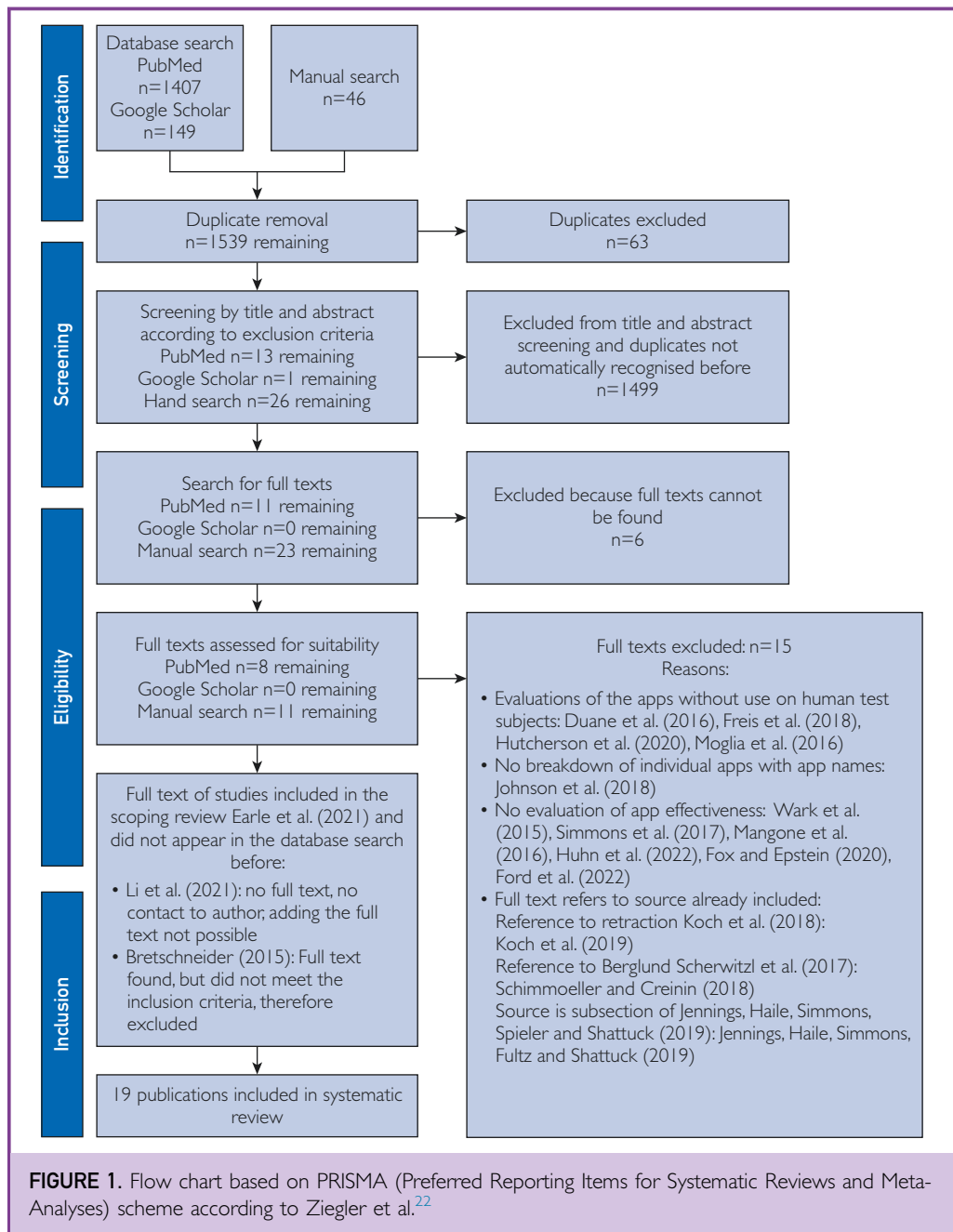
Included Publications

The categories delineated in the columns are based on the Cochrane PICOS (population, intervention, comparison, outcomes, study) scheme. The population section includes the columns “No. of cycles at start/No. of cycles after 3 mo/∅ included cycles,” “Participant selection,” and “Participant requirements.” The columns “Application,” “Technology,” and “Parameters analyzed by application,” on the contrary, pertain to intervention. The column designated “Control group” is categorized at comparison. The outcome is delineated in the “Study outcome” column. The study type is described in the column “Study design.” Abbreviations used in the tables in [Figures 2 to 4](#)^{24–45} are explained in the list of abbreviations in [Supplemental Appendix 4](#) (available online at <https://www.mcpdigitalhealth.org/>). The data extraction table can be found in [Figure 2](#). The assessments in [Figure 4](#) were performed according to the legend in [Figure 3](#)

General Issues Encountered During the Course of the Evaluation

Owing to the lack of randomized controlled trials and the heterogeneity of the included studies in terms of methods, calculation methods, follow-up conditions, and lost-to-follow-up rates, as well as the inadequate sample sizes and considerable risk of bias, the quality of most of the included studies was found to be inadequate. That is why, most studies (with the exception of 3) found deficiencies in internal and/or external validity, resulting in the inability to attain objective comparison. As a result, the results of 16 of the 19 included and evaluated studies could not be considered valid because of various potential biases or limitations in generalizability and were excluded ([Figure 4](#) “Results not included due to poor quality”).

It is also important to note that of the 19 included studies, 8 studies addressed the application Natural Cycles, 5 studies assessed the contraceptive effectiveness, and 3 assessed the topic of pregnancy. Although not all of them are written by the same authors, these studies appear at least similar and low in quality owing



to strong risks of bias because of manufacturer funding, several conflicts of interests, and deviating methods of calculations.

Two studies^{33,45} were identified with adequate quality. These studies were controlled and not funded by producers, addressing the topic of pregnancy. However, given that both were part of the Pregnancy Study Online, a cohort study from North

America, it is not reasonable to compare them. A third study²⁴ also reported good quality, but a direct comparison was not possible owing to a lack of data regarding pregnancy rates. Instead, the discussion centred on the potential of a bracelet that measures body temperature to predict ovulation.

Nevertheless, it is important to articulate these limitations because they are not always

Author (year)	App	Re-search Question	Country	Duration of Study	Study Design	N Start/ N After 3 Months/ 0 Included Cycles	Participant Selection	Participation Requirements	Reasons for Dropout	Comparison Parameters in Study	Technology	Parameters Analysed by the App	Control Group	Study Outcome	Other Potential Bias / Confounding
Shilaih et al. (2018)	Ava Fertility	OV	Switzerland	n/a	Prospective study	194/136/321 cycles	from University Hospital Zurich	Language: n/a Lead-Time-Bias: none Age: 20-40 Cycle: regularly, 24-36 days Further: no pregnancy, no health problems, no cycle-impairing medication, no regular flights between time zones, no sleep disorders	Insufficient compliance, anovulatory cycle Exclusion of 186 cycles with less than 80% daily recording between fertile window and early luteal phase	HM	App+ Wearable	Mens., Temp., Peripheral	0	Determination OV: in 82% of cycles 2 temp phases were detected, nadir of skin temp did not show robust prospective prediction of OV, retrospective determination of OV showed 2 temp phases in 86 % of cycles	Funding: Swiss Commission for Technology and Innovation CTI Conflict of interest: none declared Detection bias: use of the same LH test kits for comparison with OV displayed by app Performance bias: not recognizable Confounding age: ECO, \bar{Q} 33.66 years. Confounding hormonal C: NoECO Adjusted models: yes (late eating, BMI, coffee, sport, sex, alcohol, age, spotting) Further confoundings: sex, late eating alcohol, BMI \bar{Q} 22.97
Halle et al. (2018)	Cycle Beads	C	Egypt, Ghana, India, Jordan, Kenya, Nigeria, Rwanda	6 Cycles	Prospective market test	18,591/18,591/ n/a	Social media advertising	Language: Q in English, Arabic, App in English, Arabic, French, Spanish, Turkish, Hindi Lead-Time-Bias: none Age: 18-39 Cycle length: 26-32 days Further: n/a	n/a	Q	App	Mens.	0	Satisfaction: after 3 months of use: 60% of respondents would definitely recommend the app to others, 22% would probably recommend it to others	Funding: US Agency for International Development Conflict of interest: LT Halle and HM Fultz work for Institute for Reproductive Health, Georgetown University (owns patent for CycleBeads) Detection bias: none Performance bias: none Confounding age: noECO, \bar{Q} 20.34 Confounding hormonal C: noECO, hormonal C last 3 months Ghana (21.3%), Kenya (28.2%), Egypt (22.2%), India (8.4%) Nigeria (8.6%), Jordan (14.2%), Rwanda (8.6%) Adjusted models: none Further confoundings: none
Shelus et al. (2017)	Cycle Beads	C	Kenia	3 Months	Prospective study with quantitative and qualitative IN	185/115 (active users End-line)/ n/a	NGOs, Social Media, Word-of-Mouth, Marketing, Willingness to participate via app itself by P	Language: English and Swahili Lead-Time-Bias: none Age: from 18 Cycle: no exclusion from study, but app is for regular cycles 26-32 days, this was explained to users in app or in IN Further: n/a	Dropout on request P (especially with higher education level) If the cycle is too short, refer to the dot app is displayed in the app 38% Dropout	IN	App	Mens.	0	Contraceptive effect: 1 unplanned pregnancy when trying to conceive (due to cycle being too short) Health literacy: P know when fertile days: Baseline 88.4%, Midline 100%; P know minimum required cycle length (26 days): Midline 97.5% Abstinence or condom use (last 45 days): midline 59%, endline 63.2% Satisfaction: 70% BL, 93% ML, 97% EL (probably dropout, who is dissatisfied)	Funding: Bill and Melinda Gates Foundation Conflict of interest: Victoria Jennings developer CycleBeads, Georgetown University owns Patent Detection bias: interviews of P by employees of the Institute for Reproductive Health or the manufacturer's app, also explanation of standard daily method and use of app, 11% explanations baseline, midline 6%, endline 3% Performance bias: high number who understood the method well (due to explanations in IN) Confounding age: noECO, \bar{Q} 25 Confounding hormonal C: noECO, low proportion, not clear when before (but explanation in app regarding handling) Adjusted models: none Further confoundings: many students (45%) and university graduates (50%), very high level of education
(Koch et al. 2018) Study withdrawn (see also (Koch et al. 2019; Poliss, 2018)	Daisy View	C	n/a	13 Cycles	Retrospective efficiency study	798/125/ only P with at least 13 cycles included in PI calculation (rest in Kaplan-Meier analysis)	Mail to already registered users of DaisyView	Language: n/a Lead-Time-Bias: n/a Cycle: n/a Further: Replied to mail and completed questionnaire	Complete Case Analysis: 84% excluded from PI calculation (because of less than 13 cycles), possibly excludes those with <13 cycles due to unintentional Prg Kaplan-Meier Analysis: 10 cumulative Prg after 13 cycles), Cf. 2. Prg in PI typical use, 1 Prg in PI Perfect Use (Polis, 2018) Dropout: n/a	Prg, Q	App+ Daisys	BT daily (oral thermometer)	0	Fertility monitor Days can be used individually or linked to the DaisyView app. Here only use of Daisys and DaisyView Contraceptive effect: Typical Use PI 1.3, Perfect Use PI 0.8 (unprotected SI only on green days with continuous App use) Kaplan-Meier Analysis: 2.19% Perfect Use, 2.71% Typical Use probability of pregnancy after 13 cycles Method Failure: PI 0.6 (Pregnancy due to SI on falsely labelled infertile day) Cycle variation: 24% had irregular cycles (more than 3-4 days variation) Satisfaction: 99% would recommend the app to others (even in case of unintentional Prg), 65% use the app daily, 29% weekly, 44% monthly, frequency of use decreased with time of use, but Daisys can also be used without DaisyView	Funding: Manufacturer of Daisys Conflict of interest: One author works for the manufacturer Detection bias: Due to retrospective observation unclear when and how often SI occurred in the individual cycles, instead Perfect Use Prg rate was calculated by including all cycles in which the fertility monitor was used correctly, although it is unclear whether SI occurred Performance bias: 64% of P. use additional C, of which 22% on fertile and infertile days Confounding age: NoECO, \bar{Q} 29 Confounding hormonal C: NoEco Adjusted models: none Further confounding: BMI \bar{Q} 23, cycle length \bar{Q} 28.9 days, 90 of the 798 had experience of abortion or miscarriage. Unclear whether in study or before. Unclear how this influence s the study, if during, use the app daily, 29% weekly, 44% monthly, frequency of use decreased with time of use, but Daisys can also be used without DaisyView 70% used Daisy for C and planning of a Prg and it is unclear when intention changed

FIGURE 2. Data Extraction Table (Sorted by App Name).

immediately obvious on initial inspection of the studies' documentation available on the internet. Our quality rating is displayed in the following sections and in Figures 3 and 4. At this point, it is important to acknowledge

that a comparison of the included studies was not possible owing to these huge differences in the quality of the studies. Because of the heterogeneity of the included studies, it was not possible to address and assess all potential

Author (year)	App	Re-search Question	Country	Duration of Study	Study Design	N Start/ N After 3 Months/ Ø Included Cycles	Participant Selection	Participation Requirements	Reasons for Dropout	Comparison Parameters in Study	Technology	Parameters Analysed by the App	Control Group	Study Outcome	Other Potential Bias / Confounding
Jennings et al. (2019)	Dot	C	USA	13 Cycles	Prospective efficiency study	718/382/49 Cycles	Selection from new Dot users (reaction to pop-up message in app)	Language: English Lead-Time-Bias: none Age: 18-39 Cycle: 20-40 days, <10 days variability Further: at least 1 year C planned, heterosexual relationship, at least 3 cycles since last Prg, no hormonal C	Insufficient compliance, dropout at the request of P, cycle length >40 days, <20 days, variability <9 days, technical problems of P Exclusion of cycles in which no SI took place Dropouts after 12 months 53%	Prg	App	Mens	0	Contraceptive effect: Perfect Use failure 1%, Typical Use failure 5.8% Worst Case: if possible Prg are counted for dropouts: Worst Case Typical Use 12.8%	Funding: U.S. Agency for International Development Conflict of interest: 2 lead authors employed by Institute for Reproductive Health, Georgetown University, Washington, DC, USA, family member university director owns Cycle Technologies (manufacturer Dot) Detection bias: Various methods of pregnancy measurement (doctor, urine test or indication by P herself), Sensitivity analysis (probably pregnant if SI in fertile window and stop of use after OV; probably not pregnant if no SI in fertile window or stop of use before OV) Performance bias: Not detectable Confounding age: Eco, Ø 29 Confounding hormonal C: NoEco, but no hormonal contraception in the last 3 cycles (inclusion criteria) Adjusted models: none Further confoundings: possibly higher compliance than normal users due to the fact that P report themselves for the study via the app, 55% light-skinned, Dot was the first app to disclose its algorithm
Earle et al. (2021)					Scoping Review										Funding: School of Life and others. Conflict of interest: none Detection bias: n/a Performance bias: n/a Confounding age: n/a Confounding hormonal C: n/a Adjusted models: n/a Further confoundings: n/a
Li et al. (2016)	Dot	C	USA	n/a	Retrospective analysis simulation study	Data from 2 existing studies 68 and 221 P	n/a	Language: n/a Lead-Time-Bias: n/a Age: n/a Cycle: n/a Further: n/a	n/a	n/a	App	unclear	n/a	No full text, abstract in PubMed differs from analysis Earle et al. (2021) Not analysable, as little information available Benefit possibly overrated, as it is assumed that no unprotected SI on fertile days (Earle et al. (2021))	Funding: Cycle Technologies (manufacturer Dot App) Conflict of interest: Author works for manufacturers Detection bias: n/a Performance Bias: n/a Confounding age: n/a Confounding hormonal C: n/a adjusted models: n/a Further confoundings: young, white, well-educated P
Wise et al. (2023)	Fertility Friend (FF)	Prg	USA, Canada	12 months	Parallel non-blinded randomised control study embedded in Pregnancy Study Online (PRESTO)	8397/ (2775 intervention group + 2767 control group)/ n/a	By PRESTO	Language: n/a Lead-Time-Bias: none Age: 21-45 Cycle: No restrictions Further: no previous use of Fertilityfriend.com	Prg, >6 months since last period, trying to get pregnant for >6 months, Exclusion of P who were not pregnant after 12 months, initiation of fertility treatment, termination of attempt to become pregnant	Prg	App or App+Thermometer	Mens, BT daily, (cervical mucus), (cervical position) (LH), (SI)	2767	Does random distribution of free premium membership FertilityFriend.com (50% of P) shorten TTP? (use of an app vs. no use of an app) Intervention group 64% Probability of Prg in 12 Cycles, in control group 63%, both 70% in the first cycle. FR=1.06 Intervention group vs. control group, without first cycle FR=0.99, Stronger differences in the probability of pregnancy with age <25 (FR=1.29), longer trying to get pregnant before the start of the study (FR=1.15), with <12 years of schooling (FR=1.32), non-use of hormonal C in the last 3 months before the start of the study (FR=1.10)	Funding: partly from charitable funds, donations in kind from FertilityFriend.com and Kindara.com Conflict of interest: Lead author Consulting income from Abbvie.com Detection bias: Various measurement methods for Prg (urine test, ultrasound, blood test, test by doctor) Performance Bias: not recognisable Confounding Age: ECO, Ø 29.9 Confounding hormonal C: ECO 39.6% hormonal C before adjusted models: yes (e.g. age, births, education) Further confoundings: high level of education, high income, predominantly white population (80.6%), 70% college degree, 89% married, ≤10% previously reported infertility, 32 % had children, 76% took multivitamins, folic acid or prenatal vitamins, 43% BMI under 25, Ø 2.1 months trying to get pregnant at the start of the study, 52% of days were documented by active P in the intervention group (97% of which were menses times, 81% SI data, 72% cervical mucus, 52% BT, 44% LH, 17% cervical texture), 81.3% participated in the follow-up
Sohda et al. (2017)	Luna Luna	OV	Japan	8 Cycles	Retrospective study – secondary data analysis	7043/ n/a/ n/a	Data records of the app on server	Language: n/a Lead-Time-Bias: none Age: 20-45 Cycle: 20-45 days Further: at least one cycle recorded with OV day	n/a	n/a	App	Mens. + OV day detected by P	0	Comparison of the "Optimised Method" Luna Luna app calculation with Ogino Method and Half-Cycle Method calculations Optimised method: best prediction of OV Correlation between cycle length and follicular phase length stronger than correlation between cycle length and luteal phase. Correlation between length of past cycles and length of future cycles. strong correlation between length of past cycles and length of follicular phase. significant correlation between length of past cycles and length of luteal phase	Funding: Producer-paid salary for 2 of the authors Conflict of interest: Luna Luna data belongs to MIT Ltd., 2 of the 3 authors are employees of MIT Ltd. and have applied for a patent. Detection bias: OV detected by clin. Examination (usually ultrasound or blood test) or LH urine test: However, methods vary in accuracy Performance bias: not recognisable Confounding age: NoEco, Ø 32.94, higher than Øage of first-time mothers in Japan (30.1 years) Confounding hormonal C: NoEco Adjusted models: no Further confoundings: none mentioned

FIGURE 2. Continued

Author (year)	App	Re-search Question	Country	Duration of Study	Study Design	N Start/ N After 3 M onths/ 0 Included Cycles	Participant Selection	Participation Requirements	Reasons for Dropout	Comparison Parameters in Study	Technology	Parameters Analysed by the App	Control Group	Study Outcome	Other Potential Bias / Confounding
Berglund Scherwitzl et al. (2015)	Natural Cycles (NC)	C OV	Switzerland, Sweden	Not clear 1501 cycles (161 with LH measurement)/ 317 P/ 4.74 cycles per P	Retrospective study	317/ 317/ 4.74 Cycles	Online and offline advertising, co-operation with birth control clinics	Language: n/a Lead-Time-Bias: none Age: 18-40 Cycle: No restriction Further: sexually active, not pregnant, no Prg planned	not at least 30 days temp measurement, other C on green (infertile) days 161 cycles with LH measurement of 1501 cycles	Prg	App+Thermometer	Mens., BT daily, LH-Measurement, (SI)	0	Comparison of 161 cycles with LH measurement vs. OV provided by app C-effect: no Prg with unprotected SI only on green (infertile) days, one Prg with unprotected SI on red day Method Failure: only 0.05% of fertile days labelled as infertile Proportion of infertile days in app: first cycle 44%, with previous use of hormonal C 32%, after 3 months of use 55% (51% after hormonal C), with >70% days with BT measurements 61%, increase of 4% when entering OViTest OV day: First positive LH-OVtest Ø1.9 days before BT-based OV Cycle phase length: Luteal phase varies in Ø 1.25 days per P	Funding: n/a Conflict of interest: main author and one author are founders and shareholders NC Detection bias: unclear which LH tests were used by users and whether they were used correctly, not sure how reliable the information on pregnancies is (no questioning of P) Performance bias: not recognisable Confounding age: noECO, 30.1 Confounding hormonal C: ECO, 22% of P have discontinued hormonal C in the last 2 months Adjusted models: none Further confoundings: diseases, medication, No restrictions on P with regard to BMI, smoking, cycle length, cycle regularity, illnesses
Berglund Scherwitzl et al. (2016)	Natural Cycles	C	Sweden	5 Months	Retrospective study	4054/ 1186/ 6.3 Cycles	PR, online & offline advertising regarding natural C Mail to already registered users of NC	Language: n/a Lead-Time-Bias: P at least 3 months app access beforehand Age: from 18 Cycle: no restriction Further: potential Prg would not be a danger to P or child, at least 20 days data entered	Insufficient compliance, dropout on request P, Prg Dropout 34% (but short study)	HM	App + Thermometer	Mens., BT daily, (LH-measurement, (SI)	0	Contraceptive effect: 0.5 PI Perfect Use (unprotected SI only on green days), 7.0 PI Typical Use, Kaplan-Meier: 7.5% probability of Prg in 1 year Worst Case: if all P who have stopped app use and who have unclear pregnancy status are counted as pregnant, then 9.8 PI with typical use Proportion of infertile days in app: 41% at the beginning, after learning period and with BT input for regular cycles Ø 58% green day (5% more with LH measurement) Satisfaction: 83% more satisfied than before, 88% recommendation,	Funding: NC Conflict of interest: main author and one author are shareholders of NC, another author is an employee of NC Detection bias: Use of the same basal thermometer specified, but Prg was only recorded by P, thus unclear method of pregnancy detection Performance bias: not recognisable Confounding age: NoECO, 92% der P 20-35 Jahre Confounding hormonal C: NoECO Adjusted models: none Further confoundings: none information on SI was only entered on 8% of the days, making it difficult to calculate perfect use and typical use, therefore perfect use PI determined by when algorithm incorrectly indicated non-fertile (green) days, although these were fertile (red) and Prg occurred (regardless of whether P protocol was followed or not)
Berglund Scherwitzl et al. (2017)	Natural Cycles	C	37 Countries, 79% Sweden	18 Cycles	Prospective observational study	22,785/ after 3 months 19 534 (2684 after 18 months)/ 9.8 cycles	End user marketing techniques Mail to already registered users of NC	Language: n/a Lead-Time-Bias: none Age: n/a Cycle: n/a Further: 20 days of data stored in the app, no pre-existing Prg, paying users of the app (not just registration)	Indication of desire of pregnancy despite using the app for C 5.4% Dropout after 12 months	Prg	App + Thermometer	Mens., BT daily, (LH-measurement, (SI)	0	Contraceptive effect: Perfect Use PI 1.0 (only 9.6% Perfect Use Cycles), Typical Use PI 6.9 in 100 Women years, (6.8 in 12 months), failure rate Typical Use to 13 cycles 8.3% Worst Case: (all women with unknown pregnancy status counted as pregnant) PI 9.0 after 1 year and 9.3 after 13 cycles Method Failure: Prg rate of cycles where App displays days incorrectly as infertile: 0.5 per 100 woman-years	Funding: NC Conflict of interest: lead author and one author are founders and shareholders of NC, one co-author works for NC, 3 co-authors receive fees for e.g. presentations by NC Detection bias: study discontinuation between menses and OV -> counted as probably not pregnant (thus possible downward bias of results) Performance bias: not recognisable Confounding age: noECO, Ø 29.2 Confounding hormonal C: NoECO Adjusted models: none Further confoundings: none possibly intention to get pregnant without changing this before Prg in App, seen as Failure in Typical Use (Typical Use PI possibly higher than actual)
Berglund Scherwitzl et al. (2019)	Natural Cycles	Prg	86% Sweden, 4.3% Norway, 3.0% USA	14 Cycles	Prospective observational study	2934/ 2874/ n/a	End user marketing techniques Mail to already registered users of NC	Language: n/a Lead-Time-Bias: hormonal C <2 months ago or C with NC at least 60 days of data Age: from 20 Cycle: No restriction Further: at least 90 daily data before the end of the programme + at least 1 BT	n/a	Prg	App+Thermometer	Mens., BT daily, (LH-measurement, (SI)	0	TTP: for previous users of NC (n=1284) 2.3 cycles vs. 3.7 cycles with previous use of hormonal C (n=1590), no difference between the two groups over a period of 13 cycles Kaplan-Meier analysis: significantly lower probability of becoming pregnant after hormonal C, values equalise after about 10 months 1656 P pregnant, test after Ø 120 days	Funding: NC Conflict of interest: main author and one author are founders and shareholders of NC, 4 of the 8 authors are employed by NC, 3 co-authors receive fees for e.g. lectures from NC Detection bias: Input of positive pregnancy test by user Performance bias: not recognisable Confounding age: ECO, Ø 28.1 Confounding hormonal C: ECO, in one of the comparison groups Adjusted models: yes (age) Further confoundings: Ø BMI 23.7

FIGURE 2. Continued

shortcomings and biases we found in this study. It is also important to emphasize that the effectiveness of the studies outlined in the following sections, being calculated

differently, are subjects to various limitations, including limitations in generalizability and potential biases. Therefore, a comparison was not possible, and the effectiveness described

Author (year)	App	Research Question	Country	Duration of Study	Study Design	N Start/ N After 3 Months/ Ø Included Cycles	Participant Selection	Participation Requirements	Reasons for Dropout	Comparison Parameters in Study	Technology	Parameters Analysed by the App	Control Group	Study Outcome	Other Potential Bias / Confounding
Bull, Rowland, Lundberg et al. (2019)	Natural Cycles	C	Sweden	13 Cycles	Prospective observational study	16,331/ (14,889 >3 months data, 11 515 >6 months, 5683 >9 months, 2206 >12 months)/ 8 months data per P	End user marketing techniques Mail to already registered users of NC	Language: n/a Lead-Time-Bias: not mentioned Age: 18-45 Cycle: No restriction Further: NC as primary C, at least 20 days of data	Dropout on request P, Prg 207 (1.3%) Dropout	Prg, Q	App+ Thermometer	Mens, BT daily, (LH-measurement), (SI)	0	C effect depending on ages used C (< 2 months before) Best: Condom PI 3.5 Worst: Pill PI 8.1 Whole cohort: PI 6.1 Previously hormonal C: 7.6 Previously non-hormonal C: 4.8 Difference partly explained by more frequent unprotected SI fertile days (compliance better with e.g. condom use before)	Funding: NC Conflict of interest: main author and 4 other authors work for NC (including 2 founders and shareholders), 1 co-author receives fees for e.g. lectures from NC Detection bias: Input by P, not recognisable if additional Q Performance bias: not mentioned Confounding age: noECO, Ø 30.0 Confounding hormonal C: ECO, 41% hormonal C ≤2 months before Adjusted models: none Further confoundings: none
Bull, Rowland, Berglund Scherwitzl et al. (2019)	Natural Cycles	OV	Sweden, UK	At least 6 Cycles	Retrospective study	1,24,648 (1.4 Mio. Cycles)/ 6,12,613 Cycles, Cycles with BT and LH: 50,270	End user marketing techniques User data of app users on server	Language: n/a Lead-Time-Bias: none Age: 18-45 Cycle: 10-90 days Further: BMI 15-50, no hormonal C before <12 months, no PCOS, no menopause, 10 non-deviating BT (deviating temp, is when P characterises this as abnormal, e.g. due to fever)	<6 Cycles in App, Prg, Cycle length outside 10-90 days, BT <50% of days, with 56% exclusion of individual cycles, as in some cases no OV day can be determined due to missing BT, if only cycles with LH measurement + BT are considered 96.4% exclusion of cycles (due to fever)	n/a	App+ Thermometer	Mens, BT daily, (LH-measurement), (SI)	0	Length of cycle phases: mean length of the follicular phase 1.69 days, of the luteal phase 1.24 days Change in cycle by age: mean cycle length shortened by 0.18 days per year between the ages of 25 and 45, the follicular phase by 0.19 days Change due to BMI: 0.4 days or 14% higher cycle variability in P with BMI over 35 compared to P with BMI between 18.5 and 25	Funding: NC Conflict of interest: lead author and 3 other authors work at NC (including 2 founders and shareholders), one other author received salary from NC Detection bias: OV determined by app was not compared with any measured parameter Performance bias: not mentioned Confounding age: ECO, Ø 30.3 Confounding hormonal C: NoECO, but exclusion of PB with hormonal C in the last 12 months Adjusted models: yes (age, BMI) Further confoundings: BMI Ø 23.6 (Range: 18-45), only 8% obese compared to 15% in the population 80% of P had target C, 20% Prg
Favaro et al. (2021)	Natural Cycles	Prg	Sweden, USA	max. 14 Cycles	Prospective observational study	5376/ 3736/ n/a	End user marketing techniques Mail to already registered users of NC	Language: n/a Lead-Time-Bias: first at least 2 cycles of NC use in C mode without getting pregnant, then planning mode for at least one fertile window Age: 18-45 Cycle: no restriction Further: at least 20 days of data in planning mode, Prg intention	Insufficient compliance, dropout at P's request, exclusion of all cycles in P who switched back to C mode of the app during the study, regardless of whether pregnant or not Dropout 30.5%	Prg, Q	App+ Thermometer	Mens, BT daily, (LH-measurement), (SI)	0	During 13 cycles 84.7% pregnant Pregnancy rate after 6 months 61%, after 12 months 74%, TTP Ø 4 cycles Factors influencing TTP: highest probability (6 months: 88%, 12 months 95%; TTP Ø 2 cycles) with: Age under 35 years, variability cycle length <5 days, SI on ≥20% of days, lowest probability: older than 35 years (after 6 months 54%, after 12 months 75%), irregular cycles (after 6 months 58%, after 12 months 79%), only 3.3% of P in group with worst probability	Funding: NC Conflict of interest: 9 authors; Lead author, 7 other authors work at NC (including shareholders and the 2 founders and shareholders), 1 co-author receive honorarium from NC Detection bias: cycles counted as attempts to get pregnant where P is in the planning mode of the app on the 4th day of the cycle at the latest (as NC assumes fertile window in the first cycle from the first day after the end of menses and Ø menses of P lasts 4 days) -> as a result, TTP might be included one cycle shorter than actual, input Prg directly by P Performance bias: none Confounding age: ECO, Ø 31.8 Confounding hormonal C: NoECO, but 2 months C mode before study (lead-time bias) Adjusted models: BMI, age Further confounding: higher education and lower BMI (BMI Ø 23.3) than total population, in Ø 6 months NC before
Pearson et al. (2021)	Natural Cycles	C	UK	13 Cycles	Prospective observational study	12,247/ 12,247 in 1. 4375 at 13. cycles 9.9 months	End user marketing techniques Mail to already registered users of NC	Language: n/a Lead-Time-Bias: n/a Age: 18-45 Cycle: no restrictions Further: Paying users of NC, C intention, at least 20 daily data records, app registered in the UK	Insufficient compliance, dropout on request Pb. 35.7% delivered 13 cycles, 41.9% were still using the app at the end of the study Cycles without data were discarded Dropout at the end 58.1%	Prg, Q	App+ Thermometer	Mens, BT daily, (LH-measurement), (SI)	0	C effect: Perfect Use PI 2.0, Typical Use PI 6.1, probability of pregnancy at 13 months 7.1%, probability of pregnancy at age 19-34 8.9%, 34-45 years 5.0% Worst Case: PI 7.6 (incl. P who were assessed as probably pregnant, as discontinuation of use in luteal phase) Proportion of infertile days in app: for P with 50-70% indication of BT 34.1%, in cycle 13 even 46.5%; with over 70% indication of BT 36.8% in the first cycle and 54.5% in the 13th cycle	Funding: not recognisable Conflict of interest: 8 authors; Lead author and 4 other authors work at NC (including the 2 founders and shareholders) Detection bias: Indication of pregnancy by P by entering a pregnancy or by permanently increased BT after OV detected by app, if no confirmation or refutation and P stopped app use in the follicular phase or later in the cycle (without OV being detected), classified as probably not pregnant (included in effectiveness analysis). Stop of app use during fertile window or in luteal phase, categorised as possibly pregnant (included in Worst Case Scenario). Performance bias: not recognisable Confounding age: ECO, Ø 30.3 Confounding hormonal contraception: NoECO, 56% had hormonal C (in the last 12 months) Adjusted models: yes, age Further confoundings: Ø BMI 23.4 (12-55), 83.2% in stable partnership, 83% university degree, 84.5% had not yet had children

FIGURE 2. Continued

in these studies might not be transferable to the heterogeneity and diversity of potential users. In the following sections, we outline these values and the limitations of these studies.

Summary of Evaluations on Contraception

The average quality ratings were formed based on the evaluations. The lowest quality rating was evaluated at the study done by Shelus et al.²⁶ Primary factors contributing to this

Author (year)	App	Re-search Question	Country	Duration of Study	Study Design	N Start/ N After 3 Months/ 0 Included Cycles	Participant Selection	Participation Requirements	Reasons for Dropout	Comparison Parameters in Study	Technology	Parameters Analysed by the App	Control Group	Study Outcome	Other Potential Bias / Confounding
Bradley et al. (2021)	Ovia Fertility	Prg	USA	12 Months	Prospective observational study	45,360/ 8835 (in adjusted Model)/ n/a	User data of app users on server	Language: n/a Lead-Time-Bias: none Age: 18-45 Cycles: no restrictions Further: have just started trying to get pregnant.	Insufficient compliance, dropout on request P. only P considered who also reported Prg in max. 12 months	Prg, Q	App	Mens. (SI)	0	TTP , $\bar{\mu}$ 3.94 months Relationship between cycle characteristics and TTP : P under 25-26 days cycle length shortest TTP with irregular cycles, otherwise: P with normal cycle length (27-29 days) pregnant faster than with other cycle lengths, with normal cycle variability (<9 days) faster than with more variability, Longest TTP with very long cycles over 34 days Smokers had shorter TTPs than non-smokers (however, only P who became pregnant were analysed in the study)	Funding : Ovia Health Conflict of interest : Lead author works for manufacturers Detection bias : report Prg by P Performance bias : not recognisable Confounding age : ECO, $\bar{\mu}$ 30.8 Confounding hormonal C : NoECO Adjusted models : yes, 8835 P (age, smoking, prg) Further confoundings : BMI was asked directly (not calculated in app from height and weight), 17% did not know what BMI means, therefore BMI was excluded from final analysis, 38% had previously given birth, 15% smokers, 41% had cycle of 27-29 days, 86% P cycle variability of less than 9 days
Regidor et al. (2018)	Ovula Ring	Prg OV	Germany	15 Months	Prospective non-randomised clinical study	158 (55 aiming Prg) / n/a/ n/a	n/a	Language: n/a Lead-Time-Bias: n/a Age: 18-45 Cycle: 20-50 days Further: BMI (18.2-36.9). Wide range of cycle duration (20 to over 50 days) also for abnormal cycle length and shift work	n/a	In a sub cohort: HM US	App+ Ovula Ring	Mens, BT every 5 minutes	0	TTP , 3.79 months usage (Median), 67.7% TTP 3 months Detection OV , 96.31% (96.5% retrospective and 88.5% prospective) Accuracy of fertile window (3 days before OV, OV day, 3 days after OV) : after 3 months learning phase 88.8% prospective	Funding : n/a Conflict of interest : Main author works for a company (Exeltis West Europe & Germany), which has had a sales cooperation with the manufacturer OvulaRing since 2017 Detection bias : Type of pregnancy detection not recognisable, ovulation in part of the P compared with ultrasound examination and LH measurement Performance bias : not recognisable Confounding age : NoECO Confounding hormonal C : NoECO Adjusted models : no Further confoundings : BMI ($\bar{\mu}$ 22.2), PCOS (indication of TTP in the median, also BMI in the median)
Stanford et al. (2020)		Prg	USA, Canada	1 Year	Prospective cohort study	10,599/ (6077 intervention group + 2286 control group)/ n/a	Online study	Language: n/a Lead-Time-Bias: n/a Age: 21-45 Cycle: no restriction Further: desire to conceive for no more than 6 months, currently no C or fertility treatment	no / implausible data last menses, >6 months since last menses. Distortions possible due to loss-to-follow-up were reduced by assigning the 15% of P without data from follow-up questionnaire to the follow-up cycle and assuming the pregnancy status by multiple extrapolation	Prg	App or App+ Ther-mometer	Mens. or (Mens. + BT daily) or (Mens + e.g cervical mucus, cervical position) or (Mens+ LH-measurement), (SI)	2286	Probability of conception when using a corresponding app vs. without , 20% higher probability of conception per cycle when using an app, little difference between the apps Influence of apps with fertility indicators vs. without : with fertility indicators (BT, cervical mucus, cervical position and/or LH in urine) higher fertility per cycle than without fertility indicators	Funding : Eunice Kennedy Shriver National Institute of Child Health and Human Development, USA; one author worked for AbbVie.com (pharmaceutical manufacturer) received in-kind donations from Sandstone Diagnostics, Swiss Precision Diagnostics, FertilityFriend.com and Kindara.com, another received salary from Swiss Precision Diagnostics, outside of this study Detection bias : frequency of SI recorded, but no time when in the cycle, usage behaviour was not directly checked (when which values were entered in the app) Performance bias : partial use of several apps, discontinuation of use of one app, change which apps are used 23-35%, half of the P received free premium membership of the FertilityFriend app, increase in use of FertilityFriend from baseline: 14% to 21% over the course of the study Confounding age : ECO, $\bar{\mu}$ 29.9 Confounding hormonal C : ECO, 38% used hormonal C as their last form of C Adjusted models : yes (including folic acid intake, smoking behaviour, BMI) Further confoundings : most P have a college degree (70%), height, relationship status, (many confounding factors analysed) randomised allocation of the FertilityFriend premium membership to people who had not yet used FF (cf. Wise et al. (2023)), this group categorisation was used in Stanford et al. (2020) not considered in the final evaluation (which app was used)
	Clue					1092 Cycles								1.36	
	Fertility Friend					6778 Cycles								1.07	
	Glow					2777 Cycles								1.21	
	Kindara					1013 Cycles								1.11	
	Ovia					4138 Cycles								1.13	
	Flo					1885 Cycles								1.29	
	My Days					762 Cycles								1.28	
	Period Tracker					2132 Cycles								1.09	

FIGURE 2. Continued

were potential bias resulting from the use of CycleBeads being explained to study participants during interviews and only 37% of the

185 participants used CycleBeads for contraception, as well as the short study duration of 3 months.

Legend for the evaluation

Duration of study	Study design	P selection	Selection bias	Attrition bias	Parameters analysed by the app	Results not included due to poor quality	Other possible bias/ distortion
Duration of study C or OV + at least 13 cycles or 12 months 0 at least 6 cycles or months - less than 6 cycles or months	Prospective or retrospective study + prospective study 0 not recognisable - retrospective study Type of study + randomised controlled trial 0 non-randomised controlled study - non-randomised, non-controlled study	+ selection by university clinic, institute or as part of other large studies (e.g. PRESTO) 0 selection based on expression of interest by potential P in response to advertising (e.g. social media) - selection from already registered app users	Language barrier + study languages largely cover national languages, including e.g. migrants 0 study in the language of the country where P come from - no coverage of the languages of the population, especially in studies with more countries or countries with many languages Lead-time bias + start of treatment coincides with start of study 0 selection of P from existing users of the app - prior use of the app for a certain period of time is a prerequisite for participation Age + no or little limitation of P due to a restricted age group (within 18-45 years for C and OV and 21-45 for Prg) - narrow restriction due to age (no P over 40) Cycle + no exclusion of certain cycle lengths and variabilities 0 exclusion of individual cycles with deviating cycle lengths or variability - exclusion of test subjects with deviating cycle lengths or variability	Attrition bias C or OV + no abnormalities or exclusion of individual cycles due to missing data 0 dropout over 30% after one year or over 30% after 6 months or less - Complete Case Study Attrition bias Prg + no abnormalities, or exclusion of individual cycles due to missing data 0 exclusion of all cycles when P switched to contraceptive mode in the app - exclusion of all cycles if P did not become pregnant within the study period	+ Mens, BT, cervical mucus, HM, (+SI) 0 Mens, BT (+SI) - Mens. only (+SI)	X Results not included	Funding + no third-party funding or funding by donors independent of the manufacturer 0 low price funding (e.g. premium app memberships) - study funded by the manufacturer Conflict of interest + no recognisable conflict of interest 0 minor conflicts of interest - main authors have a financial relationship with manufacturers or are co-owners Detection bias Prg or C + standardised detection of the results 0 several methods for detecting the results - strong possibility of influencing the result through the type of result detection Detection bias OV + comparison of ovulation displayed by the app with ovulation day from hormone measurement, blood test or ultrasound - no comparison, only ovulation displayed by the app Performance bias + no treatment deviating from the basic intervention can be recognised 0 additional treatment is possible - directly described (and thus planned) additional treatment Confounding age + influence of confounding age on outcome reported - no influence of confounding age on outcome reported Confounding hormonal contraception + influence of confounding previous hormonal contraception on outcome reported 0 exclusion of users who have used hormonal contraception during the last 2 months - no influence of confounding prior hormonal contraception on outcome reported Customised models + models adjusted for confounding exist - models adjusted for confounding do not exist

FIGURE 3. Legend for Tabular Analysis of the Topics Contraception, Pregnancy and Ovulation Time.

The second lowest quality rating was found in the study by Koch et al,²⁷ which was retracted.²⁹ A further 4 studies reported mixed ratings with more negative ones. Of the 8 studies included in this article that address the application Natural Cycles, 5 studies assessed the contraceptive effectiveness of this application. A mere 3 studies yielded positive averages, led by Pearson et al,⁴² but as already mentioned, the quality was still poor (see General issues encountered during the course of the evaluation section). A comparison of the values described as the contraceptive effect was only possible for 5 sources with Pearl Index (PI) information. The 4 sources referring to Natural Cycles^{36,37,39,42} reported similar typical-use PI between 7.0 and 6.1, with different quality scores ranging from 0.35 to 0.24. The study by Koch et al²⁷ with the typical-use PI of 1.3 (which differs greatly from the other 4 studies) was withdrawn by the editor-in-chief of *Reproductive*

Health owing to suspected selection bias and retrospective self-reporting by the participants as to whether pregnancy was intended.²⁹ This study found even more limitations: the low response rate (only 13%) and the fact that this was a complete case study, which excluded women who participated for less than 13 cycles (potentially resulting in the exclusion of their pregnancies).

Bull et al³⁹ compared the typical-use PI of 7.6 after previous hormonal contraception (2 months before the start of the study) with previous nonhormonal contraception (PI=4.8). The entire cohort resulted in a PI of 6.1. For the comparison, the typical-use PI of all test subjects was considered. The lowest PI (of 3.5) was found after previous condom use, the highest after use of the pill (PI=8.1). According to the authors, this could be due to the fact that condoms, similar to natural contraception, require better compliance compared with hormonal contraception. The

Analysis contraception												
Author (year)	App	Study duration	Study design	N Start/ N after 3 months/ 0 incl. cycles	P selection	Selection bias	Attrition bias	Parameters analysed by the app	Study outcome	Other possible bias/distortion	Results not included due to poor quality	Short summarised evaluation
Haile et al. (2018)	Cycle Beads	0	+	18,591/ 18,591/ n/a	0	Language: - Lead-time bias: + Age: - Cycle: -	n/a	-	Satisfaction: after 3 months, 60% of respondents would definitely recommend the app to others, 22% would probably recommend it to others	Funding: + Conflict of interest: 0 Detection bias: + Performance bias: + Confounding age: - Confounding hormonal contraception: - Adjusted models: -	X	Average quality rating: -0.18 Outcome: Satisfaction
Shelus et al. (2017)	Cycle Beads	-	+	185/ 115/ n/a	0	Language: 0 Lead-time bias: + Age: - Cycle: -	0	-	C effect: 1 unplanned pregnancy due to cycle being too short Health Literacy: know when fertile days: Baseline 88.4%, Midline 100%, know minimum cycle length (26 days): Midline 97.5% Satisfaction: 70% BL, 93% ML, 97% EL	Funding: + Conflict of interest: 0 Detection bias: - Performance bias: - Confounding age: - Confounding hormonal contraception: - Adjusted models: -	X	Average quality rating: -0.54 Outcome: health literacy, satisfaction, no PI, 1 unplanned pregnancy (cycle fell into exclusion criteria)
Jennings, Haile, Simmons, Spieler and Shattuck (2019)	Dot	+	+	718/ 382/ 4.9 cycles	-	Language: 0 Lead-time bias: - Age: - Cycle: -	0	-	C effect: Perfect Use failure 1%, Typical Use failure 5.8% Worst Case: Typical Use 12.8%	Funding: + Conflict of interest: 0 Detection bias: 0 Performance bias: + Confounding Age: + Confounding hormonal contraception: 0 Adjusted models: -	X	Average quality rating: -0.12 Outcome: Perfect Use 1%, Typical Use 5.8%, Worst Case 12.8%
Earle et al. (2021)										Funding: + Conflict of interest: + Detection bias: n/a Performance bias: n/a Confounding age: n/a Confounding hormonal contraception: n/a Adjusted models: n/a		See below
Li et al. (2016) mentioned in Earle et al. (2021), no full text, not analysable	Dot	n/a	-	Data from 2 existing studies: 68 and 221 P	n/a	Language: n/a Lead-time bias: n/a Age: n/a Cycle: n/a	n/a	n/a	Not assessable, see data extraction table	Funding: - Conflict of interest: - Detection bias: n/a Performance bias: n/a Confounding Age: n/a Confounding hormonal contraception: n/a Adjusted models: n/a	X (due to the fact that no full text)	Average quality rating: not assessable Outcome: not assessable
Author (year)	App	Study duration	Study design	N Start/ N after 3 months / 0 incl. cycles	P selection	Selection bias	Attrition bias	Parameters analysed by the app	Study outcome	Other possible bias/distortion	Results not included due to poor quality	Short summarised evaluation
Koch et al. (2018) Study withdrawn (Koch et al., 2019; Polis, 2018)	Days View	+	-	798/ 125/ only Pb, with at least 13 cycles included in PI calculation	-	Language: n/a Lead-time bias: 0 Age: n/a Cycle: n/a	-	0	C effect: Perfect Use PI 0.8, Typical Use PI 1.3 Kaplan-Meier analysis: 2.19% perfect use, 2.71% typical use probability after 13 cycles, Method Failure: PI 0.6 Satisfaction: 99% recommendation (even in the case of unintentional pregnancy)	Funding: - Conflict of interest: - Detection bias: 0 Performance bias: 0 Confounding age: - Confounding hormonal contraception: - Adjusted models: -	X	Average quality rating: -0.47 Outcome: (study withdrawn) Satisfaction, PI Perfect Use 0.8, PI Typical Use 1.3, PI Method Failure 0.6
Berglund Scherwitzl et al. (2015)	Natural Cycles	0	-	317/ 317/ 4.74 cycles	(+) also advertising	Language: n/a Lead-time bias: + Age: - Cycle: +	+	+	C effect: no pregnancy because of unprotected intercourse only on green (infertile) days, one pregnancy with unprotected intercourse on red days Method Failure: 0.05%	Funding: n/a Conflict of interest: - Detection bias: 0 Performance bias: + Confounding age: - Confounding hormonal contraception: + Adjusted models: -	X	Average quality rating: +0.06 Outcome: 0 Prg Perfect Use, 1 Prg Typical Use, Method Failure 0.05%
Berglund Scherwitzl et al. (2016)	Natural Cycles	-	-	4054/ 1186/ 6.3 cycles	-	Language: n/a Lead-time bias: - Age: + Cycle: +	0	+	C effect: 0.5 PI with Perfect Use, 7.0 PI Typical Use Kaplan-Meier: 7.5% Prg in 1 year Worst Case: 9.8 PI Typical Use Satisfaction: 83% more satisfied than before, 88% recommendation	Funding: - Conflict of interest: - Detection bias: 0 Performance bias: + Confounding age: - Confounding hormonal contraception: - Adjusted models: -	X	Average quality rating: -0.35 Outcome: Satisfaction, PI Perfect Use 0.5, PI Typical Use 7.0, PI Worst Case Typical Use 9.8
Berglund Scherwitzl et al. (2017)	Natural Cycles	+	+	22,785/ 19,534/ 9.8 cycles	-	Language: n/a Lead-time bias: 0 Age: n/a Cycle: n/a	0	+	C effect: Perfect Use PI 1.0, Typical Use PI 6.9 Worst case scenario PI 9.0 Method Failure: PI 0.5	Funding: - Conflict of interest: - Detection bias: 0 Performance bias: + Confounding age: - Confounding hormonal contraception: - Adjusted models: -	X	Average quality rating: -0.18 Outcome: PI Typical Use 6.9, PI Perfect Use 1.0, PI Worst Case 9.0, PI Method Failure: 0.5
Bull, Rowland, Lundberg et al. (2019)	Natural Cycles	+	+	16,331/ 14,889/ 8 months	-	Language: n/a Lead-time bias: 0 Age: + Cycle: +	+	+	C effect: Best: Condom PI 3.5 Worst: Pill PI 8.1 Whole cohort: PI 6.1 Before hormonal C: 7.6 Previously non-hormonal C: 4.8	Funding: - Conflict of interest: - Detection bias: 0 Performance bias: + Confounding age: - Confounding hormonal contraception: + Adjusted models: -	X	Average quality rating: +0.12 Outcome: Typical Use PI 6.1 previously hormonal C: PI 7.6 previously non-hormonal C: PI 4.8
Pearson et al. (2021)	Natural Cycles	+	+	12,247/ 12,247 at 1st cycle vs 4375 at 13th cycle / 9.9 months	-	Language: n/a Lead-time bias: 0 Age: + Cycle: +	0	+	C effect: Perfect Use PI 2.0, Typical Use PI 6.1 Worst case: PI 7.6	Funding: n/a Conflict of interest: - Detection bias: 0 Performance bias: + Confounding Age: + Confounding hormonal contraception: - Adjusted models: +	X	Average quality rating: +0.24 Outcome: PI Perfect Use 2.0, PI Typical Use 6.1, PI Worst Case 7.6

FIGURE 4. Tabular Evaluation of the Topics Contraception, Pregnancy and Ovulation Time

more frequent unprotected sexual intercourse on potentially fertile days after hormonal contraception speaks in favor of this.

No PI data were available for 4 studies, so a comparison was not possible. Haile et al.²⁵ reported user satisfaction as an outcome,

Analysis pregnancy											
Author (year)	App	Study design	N Start/ N after 3 months/ Ø incl. cycles	P selection	Selection bias	Attrition bias	Parameters analysed by the app	Study outcome	Other possible bias/distortion	Results not included due to poor quality	Short summarised evaluation
Wise et al. (2023)	Fertility Friend.com	+	8397/ (2775 intervention group + 2767 control group)/ n/a	+	Language: n/a Lead-time bias: + Age: + Cycle: +	-	+	Use of an app vs. no use of an app - intervention group 64% probability of pregnancy in 12 cycles, in control group 63%, in the first cycle both 70 % - FR intervention group vs. control group 1.06, without first cycle FR=0.99, stronger differences in the probability of pregnancy with age <25 (FR=1.29), longer attempt, before the start of the study (FR=1.15), with <12 years of education (FR=1.32), non-use of hormonal C in the last 3 months before the start of the study (FR=1.10)	Funding: 0 Conflict of interest: 0 Detection bias: 0 Performance bias: + Confounding Age: + Confounding hormonal contraception: + Adjusted models: +		Average quality rating: +0.63 Outcome: FR intervention group vs. control group 1.06, without first cycle FR=0.99
Sohda et al. (2017)	Luna Luna	-	7043/ n/a/ n/a	-	Language: n/a Lead-time bias: + Age: + Cycle: -	n/a	(*)	Optimised method provides better prediction of the OV time compared to half cycle method and Ogino method	Funding: - Conflict of interest: - Detection bias: 0 Performance Bias: + Confounding age: - Confounding hormonal contraception: - Adjusted models: -	X	Average quality rating: -0.31 Outcome: Optimised method provides better prediction of the time of ovulation compared to half cycle method and Ogino method
Berglund Schen-witz et al. (2019)	Natural Cycles (NC)	+	2934/2874/ n/a	-	Language: n/a Lead-time bias: - Age: + Cycle: +	n/a	+	TTP: previous users of NC 2.3 cycles vs. 3.7 cycles after previous use of hormonal C; no difference between the two groups after a period of 13 cycles	Funding: - Conflict of interest: - Detection bias: 0 Performance bias: + Confounding Age: + Confounding hormonal contraception: + Adjusted models: +	X	Average quality rating: +0.19 Outcome: TTP after C with NC: 2.3 cycles, TTP after hormonal C: 3.7 cycles
Favaro et al. (2021)	Natural Cycles	+	5376/ 3736/ n/a	-	Language: n/a Lead-time bias: - Age: + Cycle: +	0	+	TTP: Ø after 4 cycles. Highest probability of TTP Ø 2 cycles with: Age under 35 years, variability cycle length <5 days, 5l on ≥20% of days Lowest probability: older than 35 years, irregular cycles	Funding: - Conflict of interest: - Detection bias: 0 Performance bias: + Confounding Age: + Confounding hormonal contraception: 0 Adjusted models: +	X	Average quality rating: +0.13 Outcome: TTP Ø 4 cycles
Author (year)	App	Study design	N Start/ N after 3 months/ Ø included cycles	P selection	Selection bias	Attrition bias	Parameters analysed by the app	Study outcome	Other possible bias/distortion	Results not included due to poor quality	Short summarised evaluation
Bradley et al. (2021)	Ovia Fertility	+	45,360/ *8835 in adjusted model	-	Language: n/a Lead-time bias: + Age: + Cycle: +	-	-	TTP: Ø 3.94 months Relationship between cycle characteristics and TTP. P under 25 days or 25-26 days cycle length shortest TTP with irregular cycles, otherwise: P with normal cycle length (27-29 days) pregnant faster than with other cycle lengths, with normal cycle variability (<9 days) faster than with more variability, longest TTP with very long cycles over 34 days Smokers had shorter TTPs than non-smokers (however, the study only analysed people who became pregnant)	Funding: - Conflict of interest: - Detection bias: 0 Performance bias: + Confounding Age: + Confounding hormonal contraception: - Adjusted models: +	X	Average quality rating: 0 Outcome: TTP: Ø 3.94 months
Regidor et al. (2018)	OvulaRing	+	158 (55 in relation to Prg)/ n/a / n/a	n/a	Language: n/a Lead-time bias: n/a Age: + Cycle: -	n/a	0	TTP: 3.79 months of use (median)	Funding: n/a Conflict of interest: - Detection bias: 0 Performance bias: + Confounding age: + Confounding hormonal contraception: - Adjusted models: -	X	Average quality rating: -0.19 Outcome: TTP 3.79 months (median) <i>(Median may be incorrect, average may be meant)</i>
Stanford et al. (2020)		+	10,599/ (6077 intervention group and 2286 control group)/ n/a	+	Language: n/a Lead-time bias: n/a Age: + Cycle: +	+	+	Probability of pregnancy with use of app vs. without, with use of app 20% higher probability of pregnancy per cycle, little difference between the apps Influence of apps with fertility indicators vs. without, with fertility indicators (BT, cervical mucus, cervical position and/or LH in urine) higher fertility per cycle than without.	Funding: + Conflict of interest: 0 Detection bias: 0 Performance bias: 0 Confounding age: + Confounding hormonal contraception: + Adjusted models: +		Average quality rating: +0.56 Outcome: 20% higher probability of pregnancy per cycle with app use (especially with fertility indicators)
	Clue		1092 cycles					1.36			
	Fertility Friend		6778 cycles					1.07			
	Glow		2777 cycles					1.21			
	Kindara		1013 cycles					1.11			
	Ovia		4138 cycles					1.13			
	Flo		1885 Cycles					1.29			
	My Days		762 cycles					1.28			
	Period Tracker		2132 cycles					1.09			

FIGURE 4. Continued

which was also mentioned by Shelus et al.²⁶ However, the topic of user satisfaction was not included in the research question. Furthermore, user satisfaction can be biased by influencing of the participants answers (intentionally or not) or simply by social desirability. Shelus et al.²⁶ recorded an improvement in health literacy over the course of the study, although this could be due to the way the results were recorded in interviews. As

with the other studies with poor quality, the results had to be excluded because only 37% of the 185 participants used CycleBeads for contraception and the duration of this study was only 3 months. Jennings et al.³⁰ described the perfect-use failure rate (ie, the proportion of probands with unplanned pregnancies) as 1%, whereas in typical use, it was 5.8%. However, only 24% of the cycles were reported as perfect-use cycles. This lacks generalizability

Analysis ovulation time											
Author (year)	App	Study duration	Study design	N Start/ N after 3 months/ Ø included cycles	PB selection	Selection bias	Attrition bias	Parameters analysed by the app	Other possible bias/distortion	Results not included due to poor quality	Short summarised evaluation
Shilah et al. (2018)	Ava Fertility	n/a	+ –	194/ 136/ 3.21 Cycles	+	Language: n/a Lead-time bias: + Age: – Cycle: –	+	(0)	Funding: + Conflict of interest: + Detection Bias: + Performance Bias: + Confounding Age: + Confounding hormonal contraception: – Adjusted models: +		Average quality rating: +0.35 Outcome: Determination of OV; in 82% of cycles detection of 2 temp phases, nadir of skin temp did not show robust prospective prediction of OV, retrospective determination of OV in 86% of cycles where 2 temp phases were detected.
Sohda et al. (2017)	Luna Luna	0	– –	7043/ n/a/ n/a	–	Language: n/a Lead-time bias: + Age: – Cycle: –	n/a	(+)	Funding: – Conflict of interest: – Detection bias: + Performance bias: + Confounding age: – Confounding hormonal contraception: – Adjusted models: –	X	Average quality rating: –0.35 Outcome: Correlation between length of cycle and length of follicular phase stronger than correlation between cycle length and luteal phase. Correlation between length of past cycles and length of future cycles. Strong correlation between length of past cycles and length of follicular phase. Significant correlation between length of past cycles and length of luteal phase
Berglund Scherwitzl et al. (2015)	Natural Cycles	0	– –	317/ 317/ 4.74 cycles	(+)	Language: n/a Lead-time bias: + Age: – Cycle: +	+	+	Funding: n/a Conflict of interest: – Detection bias: – Performance bias: + Confounding age: – Confounding hormonal contraception: + Adjusted models: –	X	Average quality rating: 0 Outcome: OV day: First positive LH ovulation test in Ø 1.9 days before temp based OV displayed in app Cycle phase length: Luteal phase varies in Ø 1.25 days per P
Bull, Rowland, Berglund Scherwitzl et al. (2019)	Natural Cycles	0	– –	1,24,648 (1.4 million cycles)/ 6,12,613 cycles, cycles with BT and LH: 50,270/ n/a	–	Language: n/a Lead-time bias: + Age: + Cycle: –	0	+	Funding: – Conflict of interest: – Detection bias: – Performance bias: + Confounding Age: + Confounding hormonal contraception: 0 Adjusted models: +	X	Average quality rating: –0.06 Outcome: Length of cycle phases: mean length of follicular phase: 16.9 days, luteal phase: 12.4 days Change in cycle with age: mean cycle length shortened by 0.18 days per year between the ages of 25 and 45, the follicular phase by 0.19 days Change with BMI: 0.4 days or 14% higher cycle variability with BMI >35 compared to BMI 18.5-25
Regidor et al. (2018)	OvulaRing	+	+ –	158/ n/a/ n/a	n/a	Language: n/a Lead-time bias: n/a Age: + Cycle: –	n/a	0	Funding: n/a Conflict of interest: – Detection bias: + Performance bias: + Confounding age: – Confounding hormonal contraception: – Adjusted models: –	X	Average quality rating: –0.06 Outcome: Detection of OV: 96.31% (96.5 % retrospectively and 88.5 % prospectively) Accuracy of fertile window: after 3 months learning phase (3 days before OV, OV day, 3 days after OV) 88.8% prospective.

FIGURE 4. Continued

and reports the low inherent effectiveness of these methods of contraception because the discrepancy between typical and perfect use is obvious owing to possible errors in use. The low percentage of perfect-use cycles documented in the included studies indicates that there is a significant degree of compliance necessary to ensure the proper utilization of these methods. Berglund Scherwitzl et al³⁵ reported no pregnancy (n=0) in participants who practiced unprotected sexual intercourse only on green (infertile) days and 1 pregnancy in unprotected sexual intercourse on days labeled as fertile.

As a worst-case scenario, that is, typical-use including test participants with unclear pregnancy status at dropout assigned as part of the sensitivity analysis, Jennings et al³⁰ described the probability of unintended pregnancies at 12.8%, Berglund Scherwitzl et al³⁷ at a PI of 9.0, and Pearson et al⁴² at a PI of 7.6. However, the sensitivity analyses were not comparable with each other (similar most of the other reported results of the

included studies). As method failure, that is, pregnancies due to fertile days incorrectly labeled as infertile by the application, Koch et al²⁷ described a PI of 0.6, Berglund Scherwitzl et al³⁵ a pregnancy probability of 0.05% and Berglund Scherwitzl et al³⁷ a PI of 0.5. Unfortunately, there was no information in the other studies on method failure.

Summary of Evaluations on Pregnancy

Because results of both studies that were not excluded addressed the topic of pregnancy, the overall average quality ratings for contraception were clearly lagging behind (Figure 5). A direct comparison of the time to pregnancy (TTP) was not possible because only the results of the 2 studies by Wise et al³³ and Stanford et al⁴⁵ were not excluded. Only 4 of 7 excluded studies lined out TTP data. Two of the values were recorded in months, and 2 were recorded in cycles. This makes a comparison of the TTP values impossible because of the natural variations in cycle length. Bradley et al⁴³ reported 3.94 months

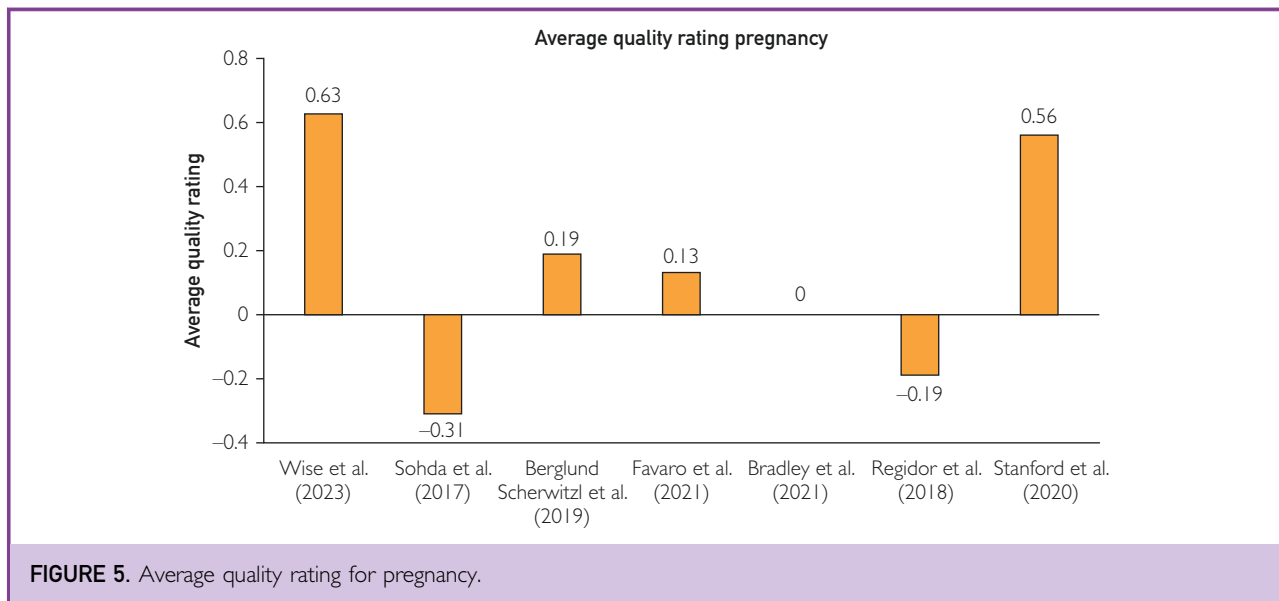


FIGURE 5. Average quality rating for pregnancy.

and Regidor et al⁴⁴ 3.79 months, whereas Favaro et al⁴¹ reported 4.0 cycles and Berglund Scherwitzl et al³⁸ reported 3.7 cycles after previous hormonal contraception and 2.3 cycles after previous contraception with Natural Cycles. The Kaplan-Meier analysis in the study by Berglund Scherwitzl et al³⁸ found a significantly lower probability of becoming pregnant after hormonal contraception, which approached the values without hormonal contraception after around 10 months. Although the study by Berglund Scherwitzl et al³⁸ seem to provide important evidence on the influence of previous hormonal contraception on TTP, the risk of lead-time bias due to previous use of Natural Cycles must be taken into account. There was no comparison group that had not previously used hormonal contraception and was not familiar with the application or natural contraception.

Although Sohda et al³⁴ focused on the topic of pregnancy, they compared different calculation methods in relation to the time of ovulation without specifying a concrete TTP or fertility rate (FR). This study had the lowest average quality rating of the included studies on pregnancy applications, followed by Regidor et al.⁴⁴ Favaro et al⁴¹ reported a lower TTP due to application use in people younger than 35 years of age, low cycle variability, and frequent sexual intercourse. The study by

Bradley et al⁴³ also seems to support the link between a higher probability of pregnancy and lower cycle variability. However, in the study by Bradley et al,⁴³ participants with short cycles of less than 25 to 26 days and with irregular cycles had a higher probability of pregnancy. This could be because shorter cycles have more potential ovulations per year. The fact that Bradley et al⁴³ found a lower TTP with application use for shorter cycles with greater variability than that for shorter cycles with less variability could be because in case of short, regular cycles, an erroneous standard 28-day cycle of the application algorithm resulted in the ovulation time being persistently misrepresented. For participants with short but highly variable cycles, the estimated ovulation time point may have been correct, at least in some cycles that happened to last 28 days.⁴³ Similarly, Bradley et al⁴³ were not conclusive as to why smokers had a shorter TTP. However, this study included in the analysis only participants who became pregnant by the end of the study. So, this complete case study excluded all participants who did not become pregnant.

It was necessary to exclude the results of these studies owing to their poor quality. Consequently, 2 sources remained for this systematic review,^{33,45} which had the best

average quality scores compared with that of the other studies in this systematic review. However, they did not give specific TTP values, but the FR, that is, when a certain sub-cohort (eg, with application use) achieved pregnancy compared with nonuse. Wise et al³³ described a very small advantage of the application although it was unclear whether and which fertility indicators provided an advantage. Not only people with a reduced chance of getting pregnant (older than 25 years of age or already trying to become pregnant for a longer period) but also people without hormonal contraception in the past 3 months and with less than 12 years of education found a higher FR through application use.³³ This influence of school education may indicate an increase in the chances of pregnancy through improved health literacy. Stanford et al⁴⁵ described a 20% higher FR with application use, especially when entering fertility indicators.

Summary of Evaluations on Ovulation Day

Of the 5 studies included on the additional topic of ovulation day, 1 also described the contraceptive effect³⁵ and 2 the effect on participants with the desire to become pregnant.^{34,44} Shilaih et al²⁴ reported the highest average quality score and Sohda et al³⁴ the lowest. The outcomes relating to this topic are not comparable with the topics of pregnancy and contraception but provide interesting background information on the menstrual cycle, which was included in the discussion.

DISCUSSION

Summarized Evaluation Regarding the Research Question

The findings of this study indicate that a high level of e-Health literacy and compliance is essential for the safe use of fertility applications. Otherwise, users can be misled into a false sense of security regarding the application's effectiveness. Despite the exclusion of most results, the average of the average quality ratings of the 10 included studies on contraception was 0.16 (SD, 0.25) and for the 7 studies on pregnancy intention was 0.14 (SD, 0.33). In terms of quality rating (despite the exclusion of results), the studies on

pregnancy intention predominantly reported better values, which was mainly due to 2 nonproducer-funded, controlled studies.^{33,45}

Because the values can only range between +1 and -1 and there are high SDs, the picture is mixed. In addition, no comparability was possible with regard to the outcome, meaning that the question of whether fertility applications are more suitable for people who want to use contraception or who want to get pregnant cannot be answered conclusively in this study.

Effectiveness in Persons With a Current Desire to Have Children

Although reproduction is a natural process, reduced fertility can be caused by the increasing age of first-time mothers in industrialized countries and stress.⁴⁷ The average TTP given by the 4 studies on the topic of the desire to get pregnant was just under 4 months and could not be calculated more precisely because 2 studies stated it in months and 2 in cycles, and these results are not evident owing to poor quality of the studies. In addition, the 2 high-quality studies^{33,45} described inconsistently whether the applications for pregnancy planning generally led to a faster pregnancy. The constant presence of the topic through the application can lead to anxiety and stress, especially among application users who are currently trying to conceive,⁴¹ with negative effects on the probability of pregnancy.⁴⁶

Although controlled studies are qualitatively preferable,²¹ a comparison with other studies is impossible without information on TTP. According to Hong et al,⁴⁸ the TTP described in the literature without supporting methods (such as determining the fertile period) varies and is influenced by biological, sociological and demographic factors as well as the study design (eg, prospective or retrospective). Eisenberg et al⁴⁹ described the average TTP in the United States in 2020 as 5.4 months, compared with 3 months in 2002. However, only people older than 30 years or those who had previously given birth to at least 1 child reported an increase in TTP.⁴⁹ However, owing to the increasing age of first-time mothers in industrialized nations,⁴⁷ the proportion of individuals aged 30 years and older who express a desire to become pregnant is also on the rise.

Effectiveness in Individuals With a Contraceptive Intention

Most contraceptive methods have a lower probability of unwanted pregnancies if used correctly and without interruption.⁵⁰ For example, the Standard Days Method on which the Dot and CycleBeads applications are based, in which unprotected sex is avoided from the 8th to 19th day of the cycle, reported a perfect-use PI of 4.75 for regular cycles.⁵¹ This number is likely to be considered a high probability of pregnancy by many users. However, the typical-use PI was even much higher at a PI of 11.96.⁵¹ According to Trussel and Grummer-Strawn,⁵² only the influence of imperfect use on the contraceptive effect shows the forgiveness of usage errors by the method without unintended pregnancies occurring. The typical-use values of methods with little room for maneuver for the user (eg, copper or hormone intrauterine devices and hormone implants) are significantly closer to the perfect-use values because there is a high inherent effectiveness owing to the reduction of error factors. In contrast, natural contraceptive methods sometimes show large differences in this aspect.⁵²

Even if we disregard the questionable quality of these studies, the difference between the typical-use and perfect-use PI values of these publications ranged between 4.1 and 6.5 and between 9.3 and 5.6 when considering the worst-case typical-use PI. This indicates a lower inherent effectiveness of the methods. This could be because there are more possible sources of error due to the users themselves, for example, when measuring basal body temperature or determining the quality of cervical mucus. Lifestyle also influences cycle variability,¹³ which can affect the accuracy of determining the fertile period.

Moreover, if the quality of the included studies had been high enough to produce results of stronger evidence, the influence of imperfect use on the number of unintended pregnancies would be roughly comparable with the values described by Trussel⁵⁰ for hormonal contraceptive methods that must be used regularly by the users themselves, for example, contraceptive pill or NuvaRing (perfect-use PI of 0.3 vs typical-use PI of 8.0). In contrast, the inherent effectiveness is higher for the usage of hormonal or copper

intrauterine devices because of the reduced likelihood of application errors, that is, less user compliance is required to ensure safe use.⁵⁰ The high rate of dropouts and the low percentages of perfect-use cycles (eg, only 9.6% of the cycles in Berglund Scherwitzl et al³⁷) appear to support this.

Influences on the Generalizability of Study Results

Methodologically, there is often a lack of randomization in the available studies, which reduces the transferability of the study results.⁵⁰ The different study durations also greatly reduce comparability as the probability of pregnancy per cycle decreases: the greatest chance (30%) is present in the first cycle, compared with 1% to 3% per cycle after more than 3 years.⁴⁷

The data on user satisfaction of the contraception applications were not objective, but it is important to state that these data seemed very high at 60% to 99% but were very prone to bias (whether intentional or unintentional). These high values are not consistent with the sometimes very high dropout rates. These facts should be kept in mind for individuals interested in natural contraception and tend to rely in advertisements or studies of poor quality. In addition, the question arises as to whether participants have sufficient digital health literacy (e-Health literacy) and compliance⁵³ to be able to adequately understand opportunities and risks and use the methods correctly in the long term.⁵⁰ In the context of attrition bias, dropouts can lead to the effect of the intervention being overestimated or underestimated because dropouts can be related to the method used itself.⁵⁴

The number of participants varied from 55 to over 1,000,000. Nevertheless, the inclusion of a considerable number of participants does not inherently ensure comprehensive diversification of the overall cohort. It must be acknowledged that the results may not be considered representative of the broader variability observed in the population.

Only 1 study excluded potential participants for whom pregnancy would pose a risk to mother or child.³⁶ Health consequences of unplanned pregnancies should be assessed anamnesticly by medical study staff before

the start of the study and communicated to participants in an understandable way, as their e-Health literacy cannot be assumed across the board.^{53,55} Because the participants in most of the studies (n=11) were selected from already registered application users, the clientele was limited to technologically savvy people with basic knowledge of the application in the sense of self-selection (and a risk of lead-time bias). Only 2 sources used clinics to recruit participants,^{24,35} whereas 1 study used nongovernmental organizations as multipliers.²⁶ In contrast to the potential lead-time bias mentioned earlier, 40% of the participants in the study conducted in Kenya by Shelus et al²⁶ had never used contraception before. Whether this was due to structural, religious, or financial reasons is unclear; a lack of education on contraceptive issues in some emerging and developing countries could also be a reason.⁵⁶

Education, social status, and cycle variability also appear to be correlated. A lower social status is often associated with a lower level of education⁵⁷ and more cycle variability⁵⁸ with negative effects on the contraceptive effect. In addition, Shelus et al²⁶ described higher dropout rates with a higher level of education. The underrepresentation of ethnic groups⁵⁹ is also evident in this systematic review because most of the studies were conducted in industrialized countries. Only 2 studies were conducted in emerging or developing countries.^{25,26} This may limit generalizability because the PI and dropout rates of natural contraceptive methods may differ from country to country.⁶⁰

Physiologically, cycle fluctuations of up to 14 days are possible even in menstruating persons who assume a regular cycle.¹² As many as 9 of the included studies had no participation restrictions regarding cycle length and variability, whereas 7 had restrictions. Although the age of the participants was only slightly restricted in 11 studies (18-45 years) and only 5 prescribed tighter restrictions, both younger users younger than the age of 18 years and older users older than 45 years were not included in the study results. Because irregular cycles are to be expected in these age groups in particular owing to puberty and the onset of menopause, respectively,¹³ age-

adjusted proof of efficacy is particularly important for these user groups. With regard to the desire to have children, the probability of pregnancy is significantly reduced at an older age.⁴⁷ With regard to contraceptive intent, unplanned pregnancy is also associated with increased risks for the pregnant individual and the child, particularly in people younger than 20 years of age⁶¹ and in people aged 35 and older.⁶²

People with a higher BMI can also experience greater cycle fluctuations,^{40,63} which can have an impact on the likelihood of pregnancy. The generalizability of study results may therefore be reduced depending on the BMI. This topic was inadequately mapped in the studies. Moreover, the effect of previous hormonal contraception on effectiveness was not stated in 10 studies, and 3 sources avoided this by excluding users who had used hormonal contraception in the previous months. It is therefore difficult to transfer and compare the results, especially because the duration in which previous hormonal contraception influences the probability of pregnancy is presented differently in the literature and can be up to 42 months.^{38,64}

Methodologically, the measured parameters of the cycle applications are one of the most important indicators of their effectiveness, but only if these parameters are also included in the calculations of the fertile time of the current cycle.⁷ Unfortunately, in 4 studies, the parameters recorded by the applications consisted solely of the start of menstrual bleeding and the (optional provided) time of intercourse. These applications are mostly based on calendar methods such as the Standard Days Method and, according to Frank-Herrmann et al,⁷ are very inaccurate owing to cycle variability. Unfortunately, however, it is unclear to what extent these parameters were used to calculate the fertile period. Only 1 study disclosed the algorithm,³⁰ but this used the Standard Days Method.

Three applications required a special measuring device: these included a temperature measuring arm band,²⁴ an intravaginally worn temperature measuring ring,⁴⁴ and a thermometer, which can also be used independently as a cycle computer.²⁷ The remaining 16 studies did not describe any specific

measuring device required but enabled the input of basal temperature measurements (eg, oral, vaginal, or rectal) or luteinizing hormone measurements. The study situation on peripheral temperature measurement is currently still contradictory.^{5,65}

In terms of study funding, 10 studies were funded by the manufacturers, which can be associated with increased interest in rapid approval and marketing and in study results with high effectiveness.³⁰ Conflicts of interest were visible in 13 studies because at least 1 of the main authors worked for the application manufacturers. Although 5 studies appeared at first glance to have a lower risk of conflict of interest, 3 of them reported how researchers from different studies had professional relationships with (at first glance) different applications.^{25,26,30}

Limitations

Because of different requirements for search strings in PubMed and Google Scholar, the selection of search terms was difficult. The study types were also very inconsistent. Originally, clinical trials (ideally randomized controlled trials) were to be searched for in order to obtain higher quality results. However, it turned out that the sources contained different types of studies, particularly on the topic of contraception, and that there were hardly any randomized controlled trials. The selection of study types was therefore expanded to obtain a sufficient number of results, but this made the sources difficult to compare and, in some cases, brought with it a considerable risk of bias, resulting in only 2 studies of 19 meeting the quality standards.

CONCLUSION

A comparison of the studies was not feasible, precluding a definitive conclusion regarding the applications' capacity to assist users. The inconsistency in application naming further complicates the decision-making process for potential users and practitioners alike. Consequently, the health care system, its stakeholders, and politicians are confronted with the increasing challenge of responding more effectively and equitably to the desires and concerns of individuals seeking hormone-free contraception and support in achieving

pregnancy. This is because younger generations are seeking contraception and child-bearing support options that are straightforward and align with their personal needs. An authoritarian approach and paternalism, as evidenced by legislation and the actions of gynaecologists, coupled with the unclear legal classification of applications and the methodological shortcomings of studies examining many contraceptive methods, will not impede this development.

POTENTIAL COMPETING INTERESTS

The authors report no competing interests.

SUPPLEMENTAL ONLINE MATERIAL

Supplemental material can be found online at <https://www.mcpcdigitalhealth.org/>. Supplemental material attached to journal articles has not been edited, and the authors take responsibility for the accuracy of all data.

Abbreviations and Acronyms: BMI, body mass index; FR, fertility rate; PI, Pearl Index; TTP, time to pregnancy

Correspondence: Address to Jan P. Ehlers, DVM, MA, Didactics and Educational Research in Healthcare, Medicine Department, Faculty of Health, Witten/Herdecke University, Alfred-Henhausen-Str. 50, D-58448 Witten, Germany (jan.ehlers@uni-wh.de).

ORCID

Isabell Rabe:  <https://orcid.org/0009-0003-3268-8620>; Jan P. Ehlers:  <https://orcid.org/0000-0001-6306-4173>

REFERENCES

1. Kondom löst Pille als Verhütungsmittel Nummer eins ab: Neue BZgA-Studiendaten zum Verhütungsverhalten Erwachsener. Bundeszentrale für gesundheitliche Aufklärung. https://www.bzga.de/fileadmin/user_upload/PDF/pressemitteilungen/2023/23_11_16_PM_BZgA-Studiendaten_Verh%C3%BCtungsverhalten_2023.pdf. Accessed April 8, 2024.
2. Le Guen M, Schantz C, Régnier-Loilier A, de La Rochebrochard E. Reasons for rejecting hormonal contraception in Western countries: a systematic review. *Soc Sci Med*. 2021;1284:114247. <https://doi.org/10.1016/j.socscimed.2021.114247>.
3. Dixon SC, Herbert DL, Loxton D, Lucke JC. 'As many options as there are, there are just not enough for me': contraceptive use and barriers to access among Australian women. *Eur J Contracept Reprod Health Care*. 2014;19(5):340-351. <https://doi.org/10.3109/13625187.2014.919380>.
4. Spitzenverband Digitale Gesundheitsversorgung. Stellungnahme des Spitzenverbands Digitale Gesundheitsversorgung e.V. (SVDGV) zum Entwurf eines Gesetzes zur Beschleunigung der Digitalisierung des Gesundheitswesens (Digital-Gesetz, DigiG). 2023: https://www.bundestag.de/resource/blob/975992/ca29e0d18ab43479ec0cc1190988006c20_14_0163-4_-Spitzenverband-Digitale-Gesundheitsversorgung-e-V-DigitalG-dاتا.pdf. Accessed April 8, 2024.

5. Association of the Scientific Medical Societies in Germany; Non-hormonal Contraception. Guideline of the German Society of Gynecology and Obstetrics (S2k-Level, AWMF Registry No. 015/095, Dezember 2023). [Version 1.1]. Published August 7, 2024: <https://register.awmf.org/de/leitlinien/detail/015-095>. Accessed March 11, 2025.
6. Bender Atik R, Christiansen OB, Elson J, et al. ESHRE guideline: recurrent pregnancy loss. *Hum Reprod Open*. 2018;2018(2):hoy004. <https://doi.org/10.1093/hropen/hoy004>.
7. Frank-Herrmann P, Freis A, Freundl-Schütt T, et al. Zyklus-Apps zur Verhütung—sicher oder Gesellschaftsspiel? *Gynäkologe*. 2019;52(2):90-97. <https://doi.org/10.1007/s00129-018-4358-6>.
8. Frank-Herrmann P, Heil J, Gnath C, et al. The effectiveness of a fertility awareness based method to avoid pregnancy in relation to a couple's sexual behaviour during the fertile time: a prospective longitudinal study. *Hum Reprod*. 2007;22(5):1310-1319. <https://doi.org/10.1093/humrep/dem003>.
9. Deutsche Gesellschaft für Gynäkologische Endokrinologie und Fortpflanzungsmedizin e. V. Sektion Natürliche Fertilität. <https://www.dggef.de/sektion-snf/>. Accessed December 28, 2023.
10. Schmidt RF, ed. *Physiologie des Menschen: Mit Pathophysiologie; + IMPP-Fragen Physiologie online*. 30., neu bearb. und aktualisierte Aufl. Springer-Medizin-Verlag; 2007.
11. Malteser Arbeitsgruppe NFP. *Natürlich und sicher—das Praxisbuch: Sichere Empfängnisregelung ohne Nebenwirkungen; Familienplanung mit sensiplan; mit dem richtigen Timing zum Wunschkind*. 18., vollst. überarb. Aufl. TRIAS; 2011.
12. Creinin MD, Keveline S, Meyn LA. How regular is regular? An analysis of menstrual cycle regularity. *Contraception*. 2004;70(4):289-292. <https://doi.org/10.1016/j.contraception.2004.04.012>.
13. Apple Women's Health Study: menstrual cycles today: how menstrual cycles vary by age, weight, race, and ethnicity. Harvard T.H. Chan School of Public Health. 2022: <https://www.hsph.harvard.edu/applewomenshealthstudy/updates/menstrualcyclestoday/>. Accessed April 8, 2024.
14. 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. <https://doi.org/10.21037/mhealth.2018.06.02>.
15. Karasneh RA, Al-Azzam SI, Alzoubi KH, Muflih SM, Hawamdeh SS. Smartphone applications for period tracking: rating and behavioral change among women users. *Obstet Gynecol Int*. 2020;2020:2192387. <https://doi.org/10.1155/2020/2192387>.
16. Duane M, Contreras A, Jensen ET, White A. The performance of fertility awareness-based method apps marketed to avoid pregnancy. *J Am Board Fam Med*. 2016;29(4):508-511. <https://doi.org/10.3122/jabfm.2016.04.160022>.
17. Schulz PJ, Nakamoto K. Health literacy and patient empowerment in health communication: the importance of separating conjoined twins. *Patient Educ Counsel*. 2013;90(1):4-11. <https://doi.org/10.1016/j.pec.2012.09.006>.
18. 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*. 2016;127(6):1153-1160. <https://doi.org/10.1097/AOG.0000000000001444>.
19. Johnson S, Marriott L, Zinaman MJ. Can apps and calendar methods predict ovulation with accuracy? *Curr Med Res Opin*. 2018;34(9):1587-1594. <https://doi.org/10.1080/03007995.2018.1475348>.
20. Setton R, Tierney C, Tsai T. The accuracy of web sites and cellular phone applications in predicting the fertile window. *Obstet Gynecol*. 2016;128(1):58-63. <https://doi.org/10.1097/AOG.0000000000001341>.
21. Braun C, Schmucker C, Nothacker M, et al. *Manual Bewertung des Biasrisikos in Interventionsstudien*. 2nd ed. Cochrane; 2021.
22. Ziegler A, Antes G, König I. Bevorzugte Report Items für systematische Übersichten und Meta-Analysen: Das PRISMA-Statement. *Dtsch med Wochenschr*. 2011;136(08):e9-e15. <https://doi.org/10.1055/s-0031-1272978>.
23. Perez S. Apple stops ignoring Women's Health with iOS 9 HealthKit update, now featuring period tracking. TechCrunch. <https://techcrunch.com/2015/06/09/apple-stops-ignoring-womens-health-with-ios-9-healthkit-update-now-featuring-period-tracking>. Accessed April 8, 2024.
24. Shilain M, Goodale B, Falco L, Kübler F, De Clerck V, Leeners B. Modern fertility awareness methods: wrist wearables capture the changes in temperature associated with the menstrual cycle. *Biosci Rep*. 2018;38(6):BSR20171279. <https://doi.org/10.1042/BSR20171279>.
25. Haile L, Fultz H, 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. <https://doi.org/10.21037/mhealth.2018.06.07>.
26. 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(3):131-141. <https://doi.org/10.1363/43e4617>.
27. Koch MC, Lermann J, van de Roemer N, et al. Improving usability and pregnancy rates of a fertility monitor by an additional mobile application: results of a retrospective efficacy study of Days and DaysView app. *Reprod Health*. 2018;15(1):37. <https://doi.org/10.1186/s12978-018-0479-6>.
28. Koch MC, Lermann J, van de Roemer N, et al. Clarifications concerning the commentary "Published analysis of contraceptive effectiveness of Days and DaysView app is fatally flawed." *Reprod Health*. 2019;16(1):83. <https://doi.org/10.1186/s12978-019-0746-1>.
29. Polis CB. Published analysis of contraceptive effectiveness of Days and DaysView app is fatally flawed. *Reprod Health*. 2018;15(1):113. <https://doi.org/10.1186/s12978-018-0560-1>.
30. Jennings V, Haile L, 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(2):148-153. <https://doi.org/10.1080/13625187.2019.1581164>.
31. Earle S, Marston HR, Hadley R, Banks D. Use of menstruation and fertility app trackers: a scoping review of the evidence (Bd. 47). British Medical Journal Publishing Group; 2021. <https://doi.org/10.1136/bmjrsrh-2019-200488>.
32. Li D, Heyer L, Jennings V, Smith CA, Dunson D. Personalised estimation of a woman's most fertile days. *Eur J Contracept Reprod Health Care*. 2016;21(4):323-328. <https://doi.org/10.1080/13625187.2016.1196485>.
33. Wise LA, Wang TR, Stanford JB, et al. A randomized trial of web-based fertility-tracking software and fecundability. *Fertil Steril*. 2023;119(6):1045-1056. <https://doi.org/10.1016/j.fertnstert.2023.02.005>.
34. 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(11):e391. <https://doi.org/10.2196/jmir.7468>.
35. 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(5):403-408. <https://doi.org/10.3109/13625187.2014.988210>.
36. 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(3):234-241. <https://doi.org/10.3109/13625187.2016.1154143>.
37. 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(6):420-425. <https://doi.org/10.1016/j.contraception.2017.08.014>.

38. Berglund Scherwitzl E, Lundberg O, Kopp Kallner H, et al. Short- and long-term effect of contraceptive methods on fecundity. *Eur J Contracept Reprod Health Care*. 2019;24(4):260-265. <https://doi.org/10.1080/13625187.2019.1621999>.
39. Bull J, Rowland S, Lundberg O, 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(3):1-7. <https://doi.org/10.1136/bmjopen-2018-026474>.
40. Bull JR, Rowland SP, Berglund Scherwitzl E, Scherwitzl R, Danielsson KG, Harper J. Real-world menstrual cycle characteristics of more than 600,000 menstrual cycles. *NPJ Digit Med*. 2019;2:83. <https://doi.org/10.1038/s41746-019-0152-7>.
41. Favaro C, Pearson JT, Rowland SP, et al. Time to pregnancy for women using a fertility awareness based mobile application to plan a pregnancy. *J Womens Health (Larchmt)*. 2021;30(11):1538-1545. <https://doi.org/10.1089/jwh.2021.0026>.
42. Pearson JT, Chelstowska M, Rowland SP, et al. Natural Cycles app: contraceptive outcomes and demographic analysis of UK users. *Eur J Contracept Reprod Health Care*. 2021;26(2):105-110. <https://doi.org/10.1080/13625187.2020.1867844>.
43. Bradley D, Landau E, Jesani N, et al. Time to conception and the menstrual cycle: an observational study of fertility app users who conceived. *Hum Fertil (Camb)*. 2021;24(4):267-275. <https://doi.org/10.1080/14647273.2019.1613680>.
44. Regidor PA, Kaczmarczyk M, Schiweck E, Goetzenjan-Festag M, Alexander H. Identification and prediction of the fertile window with a new web-based medical device using a vaginal biosensor for measuring the circadian and circamensual core body temperature. *Gynecol Endocrinol*. 2018;34(3):256-260. <https://doi.org/10.1080/09513590.2017.1390737>.
45. Stanford JB, Willis SK, Hatch EE, Rothman KJ, Wise LA. Fecundability in relation to use of mobile computing apps to track the menstrual cycle. *Hum Reprod*. 2020;35(10):2245-2252. <https://doi.org/10.1093/humrep/deaa176>.
46. Clancy KBH. Reproductive ecology and the endometrium: physiology, variation, and new directions. *Am J Phys Anthropol*. 2009;(140 suppl 49):137-154. <https://doi.org/10.1002/ajpa.21188>.
47. Taylor A. ABC of subfertility: extent of the problem. *BMJ*. 2003;327(7412):434-436. <https://doi.org/10.1136/bmj.327.7412.434>.
48. Hong X, Yin J, Wang W, Zhao F, Yu H, Wang B. The current situation and future directions for the study on time-to-pregnancy: a scoping review. *Reprod Health*. 2022;19(1):150. <https://doi.org/10.1186/s12978-022-01450-6>.
49. Eisenberg ML, Thoma ME, Li S, McLain AC. Trends in time-to-pregnancy in the USA: 2002 to 2017. *Hum Reprod*. 2021;36(8):2331-2338. <https://doi.org/10.1093/humrep/deab107>.
50. Trussell J. Contraceptive failure in the United States. *Contraception*. 2004;70(2):89-96. <https://doi.org/10.1016/j.contraception.2004.03.009>.
51. Arévalo M, Jennings V, Sinai I. Efficacy of a new method of family planning: the Standard Days Method. *Contraception*. 2002;65(5):333-338. [https://doi.org/10.1016/s0010-7824\(02\)00288-3](https://doi.org/10.1016/s0010-7824(02)00288-3).
52. Trussell J, Grummer-Strawn L. Contraceptive failure of the ovulation method of periodic abstinence. *Fam Plann Perspect*. 1990;22(2):65-75.
53. Aulenkamp J, Mosch L, Schmidt J, Kaufmann M, Wirbelauer L. Digitale bildung. In: Matusiewicz D, Henningsen M, Ehlers JP, eds. *Digitale Medizin*. Medizinisch Wissenschaftliche Verlagsgesellschaft; 2020:75-93.
54. Buchberger B, Elm E von, Gartlehner G, et al. Bewertung des Risikos für Bias in kontrollierten Studien. *Bundesgesundheitsbl*. 2014;57(12):1432-1438. <https://doi.org/10.1007/s00103-014-2065-6>.
55. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. *J Med Internet Res*. 2006;8(2):e9. <https://doi.org/10.2196/jmir.8.2.e9>.
56. Schäffer M. *Start-Ups International: Gründergeschichten Rund Um Den Globus: Erfolgsfaktoren, Motivationen Und Persönliche Hintergründe*. Fachmedien Wiesbaden: Springer; 2015.
57. Rutter S. Soziale Ungleichheit im Bildungssystem. *Sozialanalyse in der pädagogischen Arbeit*. 2021:1-66. https://doi.org/10.1007/978-3-658-32065-2_1.
58. Münster K, Schmidt L, Helm P. Length and variation in the menstrual cycle—a cross-sectional study from a Danish county. *Br J Obstet Gynaecol*. 1992;99(5):422-429. <https://doi.org/10.1111/j.1471-0528.1992.tb13762.x>.
59. Epstein DA, Lee NB, Kang JH, et al. Examining menstrual tracking to inform the design of personal informatics tools. *Proc SIGCHI Conf Hum Factor Comput Syst*. 2017;2017:6876-6888. <https://doi.org/10.1145/3025453.3025635>.
60. World Health Organization. A prospective multicentre trial of the ovulation method of natural family planning. II. The effectiveness phase. *Fertil Steril*. 1981;36(5):591-598. [https://doi.org/10.1016/s0015-0282\(16\)45856-5](https://doi.org/10.1016/s0015-0282(16)45856-5).
61. *World Population Prospects 2022: Summary of Results*. United Nations; 2022.
62. Nugent D, Balen A. The effects of female age on fecundity and pregnancy outcome. *Hum Fertil*. 2001;4(1):43-48. <https://doi.org/10.1080/1464727012000199251>.
63. Apple Women's Health Study: investigators at Harvard T. H. Chan School of Public Health talk about the benefits of tracking your period as a window into health. Harvard T.H. Chan School of Public Health. 2020. <https://www.hsph.harvard.edu/applewomenshealthstudy/learn-more/>. Accessed April 8, 2024.
64. Vessey MP, Wright NH, McPherson K, Wiggins P. Fertility after stopping different methods of contraception. *Br Med J*. 1978;1(6108):265-267. <https://doi.org/10.1136/bmj.1.6108.265>.
65. Wark JD, Henningham L, Gorelik A, Jayasinghe Y, Hartley S, Garland SM. Basal temperature measurement using a multi-sensor amband in Australian young women: a comparative observational study. *JMIR Mhealth Uhealth*. 2015;3(4):e94. <https://doi.org/10.2196/mhealth.4263>.