December 19, 2018: This file was updated as follows:

1. Formatting changes for consistency and ease of reading
2. Table of contents added
3. Appendix 5 was modified in several places:
    a. Berglund 2016: 4 high quality markers (instead of 5)
    b. Burkhart 2000: 6 high quality markers (instead of 8)
    c. Fehring 2013: 6 high quality markers (instead of 7)
    d. Fehring 2017: 4 high quality markers (instead of 5)
    e. Gribble 2008: 5 high quality makers (instead of 4)
4. Key added to describe proposed rank numbers for detailed quality ranking tables (Appendix 6)
5. Bonnar quality ranking form was corrected to indicate that the statistical analysis was moderate and not low quality per a conversation with the author conducted prior to publication (Appendix 6)

Table of Contents: Appendixes

Contents

Urrutia RP, Polis CB, Jensen ET, Greene ME, Kennedy E, Stanford JB. Effectiveness of fertility awareness-based methods for pregnancy prevention: a systematic review. Obstet Gynecol 2018; 132.
The authors provided this information as a supplement to their article.

Appendix 1. Detailed Search Strategy

PubMed

Set	Search	Results
#1	"Natural family planning methods"[MeSH] OR "Natural family planning methods"[tw] OR "Natural family planning method"[tw] OR "Ovulation prediction"[MeSH] OR "ovulation prediction"[tw] OR"Ovulation detection"[MeSH] OR "ovulation detection"[tw] OR "natural family planning"[tw] OR "natural contraception"[tw] OR "natural fertility"[tw] OR "fertility awareness"[tw] OR"Billings Ovulation Method"[tw] OR "Creighton Model"[tw] OR "Symptothermal Method"[tw] OR "Marquette Method"[tw] OR "basal body"[tw] OR "cervical mucus monitoring"[tw] OR "cervical secretions"[tw] OR "cervical secretion"[tw] OR "cervix mucus"[MeSH] OR "cervix mucus"[tw] OR "cervical mucus"[tw] OR "cervical mucus"[tw] OR "modified mucus method"[tw] OR "Standard Days Method"[tw] OR "rhythm method"[tw] OR "calendar method"[tw] OR (charting OR tracking AND (calendar OR fertility)) OR "sensiplan"[tw] OR "procef"[tw] OR "sexual abstinence"[MeSH] OR "sexual abstinence"[tw] OR "periodic abstinence"[tw] OR "Fertility monitoring"[tw] OR "Natural birth control"[tw] OR "Persona"[tw] OR "Cervicovaginal secretion"[tw] OR "Cervicovaginal secretions"[tw] OR "Cervicovaginal fluid"[tw] OR "Cervicovaginal fluids"[tw] OR "Cervicovaginal mucus"[tw] OR "Cervico-vaginal secretion"[tw] OR "Cervico-vaginal secretions"[tw] OR "Cervico-vaginal fluid"[tw] OR "Cervico-vaginal fluids"[tw] OR "Cervico-vaginal mucus"[tw] OR" Vaginal secretion"[tw] OR "Vaginal secretions"[tw] OR "Vaginal fluid"[tw] OR "Vaginal fluids"[tw] OR "Cervical cumin "[tw] OR "Cervical mucins"[tw] OR "TwoDay Method"[tw] OR "TwoDay Algorithm"[tw]	12,373
#2	"Pregnancy"[MeSH] OR "Pregnancy"[tw] OR "Pregnancies"[tw] OR "Pregnant"[tw] OR "Patient compliance"[Mesh] OR "compliance"[tw] OR "Health behavior"[tw] OR "Health Behaviors"[tw] OR "health behaviour"[tw] OR "health behaviours"[tw] OR "Consumer behavior"[tw] OR "consumer behaviour"[tw] OR "consumer behaviours"[tw] OR "Consumer Behavior"[MeSH] OR "contraception	1,986,023

	behavior”[MeSH] OR “contraception behavior”[tw] OR “contraception behaviour”[tw] OR “Patient dropouts”[MeSH] OR “Dropouts”[tw] OR “Effectiveness”[tw] OR “efficacy”[tw] OR “Satisfaction”[tw] OR “Dissatisfaction”[tw] OR “Continuation”[tw] OR “continuance”[tw] OR “Discontinuation”[tw] OR “discontinuance”[tw] OR “Acceptability”[tw] OR “adherence”[tw]	
#3	#1 AND #2	4,516
#4	("Case Reports"[publication type] OR “Editorial“[publication type] OR “Letter”[publication type] OR “Addresses”[publication type] OR “Autobiography”[publication type] OR “Bibliography”[publication type] OR “Biography”[publication type] OR “comment”[publication type] OR “Congresses”[publication type] OR “Consensus Development Conference, NIH”[publication type] OR “Dictionary”[publication type] OR “Directory”[publication type] OR “Festschrift”[publication type] OR “Interactive Tutorial”[publication type] OR “Interview”[publication type] OR “Lectures”[publication type] OR “Legal Cases”[publication type] OR “Legislation”[publication type] OR "News"[publication type] OR “Patient Education Handout”[publication type] OR “Periodical Index”[publication type] OR “Portraits”[publication type] OR “Scientific Integrity Review”[publication type] OR “Video-Audio Media”[publication type] OR “Webcasts”[publication type])	3,390,005
#5	(#1 AND #2) NOT #4	4,333
#6	(#1 AND #2) NOT #4 Filters: Humans	3,303

Total after Internal Duplicates removed: 3,301
Total after External Duplicates removed: 3,289

EMBASE

Set	Search	Results
#1	((fertility) NEAR/3 (awareness) NEAR/3 (method*))	96
#2	(natural NEAR/3 ('family planning' OR fertility OR contraception OR 'birth control'))	1,159
#3	('sexual abstinence'/exp OR ('Creighton model' OR 'periodic abstinence' OR 'sexual abstinence' OR sensiplan OR procef OR persona):ab,ti)	1,188
#4	('ovulation detection'/exp OR 'ovulation prediction'/exp OR ('ovulation detection' OR 'ovulation prediction' OR 'basal body temperature' OR ((cervix* OR cervic* OR vagina*) NEAR/3 (secretion* OR mucus OR mucous OR mucin* OR fluid* OR secretion*))):ab,ti)	8,349
#5	((('calendar' OR 'rhythm' OR 'modified mucus' OR 'standard days' OR 'billings ovulation' OR TwoDay OR symptothermal OR Marquette) NEXT/3 (method* OR algorithm*)):ab,ti)	1,187
#6	(((fertility) NEAR/3 (track* OR chart* OR monitor*)):ab,ti)	178
#7	('pregnancy'/exp OR (pregnancy OR pregnancies OR pregnant):ab,ti)	774,611
#8	('patient compliance'/exp OR (compliance OR adherence OR acceptability OR discontinuation OR continuation OR satisfaction OR dissatisfaction OR efficacy OR effectiveness OR dropout*):ab,ti)	1,444,495
#9	('health behavior'/de OR 'consumer attitude'/exp OR ((consumer OR health OR contracept*) NEAR/3 (behav* OR attitude*)):ab,ti)	72,656
#10	#1 OR #2 OR #3 OR #4 OR #5 OR #6	11,703

#11	#7 OR #8 OR #9	2,236,410
#12	#10 AND #11	3,725
#13	#10 AND #11 AND [humans]/lim	2,879
#14	#10 AND #11 AND [humans]/lim AND ([cochrane review]/lim OR [systematic review]/lim OR [controlled clinical trial]/lim OR [randomized controlled trial]/lim OR [meta analysis]/lim)	233
#15	#10 AND #11 AND [humans]/lim AND ([article]/lim OR [article in press]/lim OR [conference abstract]/lim OR [conference paper]/lim OR [review]/lim)	2,743
#16	#14 OR #15	2,752

Total after Internal duplicates removed: 2,733
Total after External duplicates removed: 1,085

CINAHL

Set	Search	Results
S1	(MH “Family Planning, Natural”) OR (MH “Family Planning”)	4,157
S2	(natural) N3 ("family planning" OR fertility OR contraception OR "birth control"))	228
S3	(MH “Ovulation prediction”) OR (MH “ovulation detection”) OR (MH "Sexual Abstinence") OR (MH “Rhythm Method”) OR (MH “body temperature”)	3,735
S4	persona OR “basal body temperature”	145
S5	(ovulation) N3 (predict* OR detect*)	85
S6	(sexual OR periodic) N3 (abstinence)	749
S7	(Creighton OR Billings OR Symptothermal OR rhythm OR calendar OR TwoDay OR “standard days” OR “modified mucus” OR Marquette) W3 (model* OR method* OR algorithm*)	212
S8	(cervix* OR cervic* OR vagina*) N3 (secretion* OR mucus OR mucous OR mucin* OR fluid* OR secretion*)	315
S9	(fertility OR) N3 (chart* OR track* OR monitor*)	43
S10	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9	8,554
S11	(MH “Pregnancy”) OR (MH “Pregnancy, Unplanned”) OR (MH “Pregnancy, Unwanted”) OR (MH "Attitude to Pregnancy") OR (MH “birth rate”)	126,328
S12	(pregnant OR pregnancy OR pregnancies)	139,067

S13	(MH "Patient satisfaction") OR (MH "behavior") OR (MH "health behavior") OR (MH "health knowledge") OR (MH "Fertility") OR (MH "patient compliance") OR (MH "Effectiveness") OR (MH "consumer satisfaction") OR (MH "patient satisfaction") OR ("patient dropouts")	122,769
S14	(compliance OR adherence OR acceptability OR discontinuation OR continuation OR satisfaction OR dissatisfaction OR efficacy OR effectiveness OR dropout*)	288,174
S15	S11 OR S12 OR S13 OR S14	466,633
S16	S10 AND S15	2,724

Total after Internal duplicates removed: 2,718
Total after External duplicates removed: 2,335

Web of Science

Set	Search	Results
#1	TS=((fertility) NEAR/3 (awareness) NEAR/3 (method*))	63
#2	TS=(natural NEAR/3 ("family planning" OR fertility OR contraception OR "birth control"))	1,199
#3	TS=("Creighton model" OR "periodic abstinence" OR "sexual abstinence" OR sensiplan OR procef OR persona OR "ovulation detection" OR "ovulation prediction" OR "basal body temperature")	3,402
#4	TS=((cervix* OR cervic* OR vagina*) NEAR/3 (secretion* OR mucus OR mucous OR mucin* OR fluid* OR secretion*))	6,019
#5	TS=((calendar OR rhythm OR "modified mucus" OR "standard days" OR "billings ovulation" OR TwoDay OR symptothermal OR Marquette) NEAR/3 (method* OR algorithm*))	1,327
#6	TS=((fertility) NEAR/3 (track* OR chart* OR monitor*))	275
#7	TS=(pregnancy OR pregnancies OR pregnant)	362,876
#8	TS=(compliance OR adherence OR acceptability OR discontinuation OR continuation OR satisfaction OR dissatisfaction OR efficacy OR effectiveness OR dropout*)	1,514,769
#9	TS=((consumer OR health OR contracept*) NEAR/3 (behav* OR attitude*))	62,373
#10	#1 OR #2 OR #3 OR #4 OR #5 OR #6	11,978

#11	#7 OR #8 OR #9	1,902,799
#12	#10 AND #11	2,431
#13	#10 AND #11 **Refined by:DOCUMENT TYPES:** (ARTICLE OR MEETING ABSTRACT OR PROCEEDINGS PAPER OR REVIEW)	2,375

Total after Internal duplicates removed: 2,374
Total after External duplicates removed: 1,074

Clinicaltrials.gov

Searched for key words (fertility OR pregnant OR pregnancy) AND (fertility awareness OR family planning) in any field with no filters regarding study status. Total number of abstracts: 216

Appendix 2. Detailed Inclusion and Exclusion Criteria (PICOTS) Table

Category	Criteria	
	Inclusion	**Exclusion**
Population	Menstruating or recently pregnant women trying to avoid pregnancy	Non-human studies Infertility studies Studies with fewer than 50 women
Geography/ Setting	No limits	
Time period	No limits	
Length of follow up	No restrictions	
Interventions	Specific fertility awareness-based methods including: Standard Days Method, the Rhythm Method, Basal Body Temperature Methods, Billings Ovulation Method, Two Day Method, Modified Mucus Method, Creighton Model Fertility Care System, Symptothermal Methods (e.g., Sensiplan), and Symptohormonal Methods including the Marquette Method, and Persona. Studies that combine pregnancy rates for users of more than one defined (as above) method	Lactational Amenorrhea Method Undefined fertility awareness-based method Unspecified "rhythm" or other method without reference to explicit rules for determining the fertile time
Outcomes	Life table pregnancy probabilities 12 month/13 cycle Pearl pregnancy rates Survival Analysis pregnancy rates (e.g., Kaplan Meier)	Studies with no pregnancy rates or probabilities AND insufficient raw data to calculate a 12 month/13 cycle Pearl pregnancy rate
Publication language	English, German, French, and/or Spanish	All other languages
Admissible evidence (study design and other criteria)	Peer-reviewed, original research; eligible study designs include: Randomized controlled trials Meta-analyses Clinical trials Observational studies with prospectively collected outcomes and pregnancy intentions Correspondence regarding other wise included studies	Case series Case reports Reviews Editorials Letters to the editor (unless pertaining to included studies) Non-peer reviewed studies Studies without prospectively ascertained pregnancies or pregnancy intentions

Appendix 3. Abstraction Form

Abstraction form
Note: Form was created online and cannot easily be exported, but it is available to the public at https://srdr.ahrq.gov. The name of the project is "Efficacy of Fertility Awareness-Based Methods of Family Planning and Modifying Behavioral Factors

1. Publication Information: auto populates with PubMed ID
2. Which key question(s) is/are addressed?

What is the direct evidence of effectiveness of specific FAMs (e.g. Standard Days Method, Rhythm Method, Basal Body Temperature Methods, Billings Ovulation Method, TwoDay Method, Modified Mucus Method, Creighton Model Fertility Care System, Symptothermal Methods including Sensiplan, and Symptohormonal Methods including the Marquette Method, and Persona) to avoid pregnancy as measured by pregnancy rates over time among pregnancy avoiding users of FAMs?

Describe the acceptability of the methods as described in included studies (i.e. return to method after pregnancy event; adherence to the method; reported satisfaction by method users).

What is the evidence that behavioral factors (e.g. intercourse in the fertile window, consistency of observation, adherence to the method, barrier method use in the infertile and fertile times, and withdrawal use during fertile and infertile times) modify typical- and perfect-use pregnancy rates over the first year and longer among pregnancy-avoiding users of FAMs?

What is the evidence that demographic characteristics (age, race, marital status, religion, religiosity, income, education, other markers of SES, geographic location, and reproductive history) affect typical- and perfect-use pregnancy rates the first year and longer among pregnancy-avoiding users of FAMs? What is the evidence that either desire at enrollment to space pregnancies versus desire to avoid pregnancies, different levels of desire to avoid pregnancy (e.g. Likert scale), or baseline family size affect typical- and perfect-use pregnancy rates during the first year of use and longer among pregnancy-avoiding users of FAMs? What is the evidence that changing pregnancy intentions over time affect pregnancy rates?

In FAMs which include women of all reproductive categories, what is the evidence that reproductive categories (postpartum, breastfeeding, perimenopausal, oligo/amenorrhoeic [also subgroups who are more likely to have oligomenorrhea such as women living with HIV and women with eating disorders/athletes, diabetic women, women with PCOS, etc.], post-oral contraceptive or other contraceptive) affect typical- and perfect-use pregnancy rates during the first year of use and longer among pregnancy-avoiding users of FAMs?

What is the evidence that relationship characteristics (concordance of intentions, communication, relationship satisfaction, equality of decision-making, history of intimate partner violence) affect typical- and perfect-use pregnancy rates during the first year of use and longer among pregnancy-avoiding users of FAMs?

Describe the patient education components of each method (format, intensity, duration) and the instructor/teacher characteristics. What is the evidence that specific educational and or teaching components affect the pregnancy rates during the first year of use and longer among pregnancy-avoiding users of FAMs?

3. Is the study peer reviewed (y/n)?
4. Is it possible that the data in this paper is replicated elsewhere (y/n)? If yes, provide details.
5. What is the study design? RCT, Observational, Metanalysis?
6. What are the study funding sources/reported conflicts of interest?
7. List countr(ies) where participants were located.
8. Inclusion criteria?
9. Exclusion criteria with specific categories (breastfeeding, postpartum, subfertility, perimenopausal-age, oligomenorrhea)?
10. Exclusion criteria other.
11. Were women screened for and or treated for STIs during the study (y/n)?
12. Experience with the method (new, experienced, both)?
13. Recruitment or enrollment sites: church/religious organization, religiously affiliated clinic, unaffiliated private clinic, unaffiliated university clinic, unaffiliated community site, public sector clinic, other.
14. Planned duration of follow-up
15. Mean duration of follow-up
16. Dates of data collection (enrollment through final data collection).
17. How is pregnancy avoidance defined by the study including method used and time points assessed.
18. How often were pregnancy intentions assessed
19. What was the method of collecting data: interview, personal diary paper, personal diary computer, other including frequency?
20. Definition of pregnancy: self-report only, routine urine hcg testing, routine blood hcg testing, self-report with targeted pregnancy testing, live birth/miscarriage/abortion event, diary/chart evidence of 17 days or more of elevated basal body temp without menstruation, other.
21. Are pregnancies classified according to the prospectively identified FP intention?
22. Did investigators exclude cycles with no intercourse from the analysis (which ones)?

23. Was barrier and or withdrawal method use assessed in the study including frequency, timing, mode and detail.
24. Were any pregnancies excluded from the analysis (y/n)?
25. Does the study include a measure of typical use pregnancy rates? If yes, describe.
26. Does the study include a measure of perfect use pregnancy rates? If yes, describe.
27. What was the unit of analysis (woman, couple)?
28. Were all cycles in which women prospectively stated that they wished to avoid pregnancy included in the analysis (y/n)?
29. Please describe attrition include reasons for discontinuation.
30. If study is an RCT, was there adequate randomization, adequate allocation concealment, masked outcome assessment, early stoppage (y/n)?
31. For comparative studies was differential loss to follow-up noted (y/n)?
32. If comparative analysis, describe adjustment for potential confounders.
33. Selective outcomes reporting (y/n)?
34. Besides unintended pregnancy, were other adverse events reported (y/n)?
35. Were exploratory analyses performed to try and understand factors associated with unintended pregnancy?
36. Do you have additional concerns for bias not otherwise discussed (y/n)?
37. Summarize strengths of study.
38. Summarize weaknesses of the study.
39. Select study arm (e.g. specific FABM) or arms?
40. For each arm describe, educational setting, type of educator involved, training of teachers, recommended number of encounters, total time needed to learn method, total number of hours recommended, experience of teachers.
41. Was each an FAM method or NFP method? Note: this question was dropped eventually given difficulty answering.
42. Describe specific rules to determine the beginning of the fertile window.
43. Were any devices or Internet applications used as part of the method? If yes, describe.
44. Baseline data (described in means or percents as appropriate and if available): age, race, ethnicity, marital status, religion, educational attainment, socioeconomic status, lactation status, postpartum status, perimenopausal age status, post hormonal contraceptive use (<3 months or <9 months for depo provera), oligomenorrhea or infrequent cycles, parity, baseline pregnancy intentions (spacing v. limiting), changing pregnancy intentions over time, relationship characteristics, coital frequency, barrier method users, withdrawal method users.
45. List outcomes by type of analysis, typical v. perfect use, time point of use.
46. List Effect estimates with standard errors if reported.

Appendix 4. Detailed Quality Criteria Table

Domain	Quality Ranking		
	High (meets all of the following criteria)	**Moderate** (neither high nor low**)**	**Low** (meets any of the following criteria)
Inclusion exclusion criteria	Clear description of inclusion and exclusion criteria, and denominator of screened population defined (e.g. refusal rate)		Inclusion criteria inadequately defined
Exclusion of populations not at meaningful risk of pregnancy	Attempt made to exclude: cycles without sexual activity, women or couples with a reported history of 12+ cycles of unprotected sex and no conception, and women using sterilization, IUDs, or hormonal contraceptive methods		Estimates of outcomes explicitly (or presumably) includes women using hormonal methods, sterilization, IUD
Treatment of women in various reproductive categories that may impact fertility	Either exclude or do appropriate subgroup analyses for pregnancy rates among: perimenopausal aged women (40 and older), oligomenorrheic women as defined by the study, breastfeeding with amenorrhea, less than 3 cycles postpartum, less than 9 months post injectible progestin use, less than 3 months post other hormonal contraceptive use.		
Description of study population	All New users of the method (analyzed separately); Population is characterized with measurement of important sociodemographic characteristics (age, parity)		Unclear proportion of new versus experienced users.
Fertility awareness-based method(s) studied	Clearly stated, including clarity on rules for identifying fertile days		Different fertility awareness-based methods combined for analyses without clear data to support comparability
Teaching of method	Sufficiently described or referenced so as to be replicable including training and certification of teachers, curriculum or materials used, frequency and duration of instruction encounters, where applicable.		
Detection of pregnancy	Routine periodic pregnancy testing (urine, blood or ultrasound) or targeted pregnancy testing for women with signs of pregnancy on charting. Active follow up to diagnose early pregnancies at conclusion of study		Pregnancy detection by self-report only (no evidence of systematic corroboration with charting or testing)
Classification of pregnancy as intended or unintended	Pregnancy intentions assessed prospectively and repeatedly every month or cycle; All cycles (and pregnancies) reported, including intended pregnancies; All cycles (and pregnancies) with prospectively identified intention to avoid pregnancy included in typical use analysis for pregnancy avoidance including “training” or “learning” phase pregnancies.		Any pregnancy excluded without being prospectively classified by the woman as intended. Note: study only ranked as poor if all relevant estimates have this problem.

Concurrent use of coitus-dependent methods	Barrier method use and withdrawal use assessed systematically; Appropriate subgroup analysis conducted to differentiate between effectiveness with and without additional methods		
Study duration	At least 12 months or 13 cycles with 12 month or 13 cycle typical use pregnancy probability reported, except for special populations (e.g. postpartum).		Studies longer than 12 months reporting Pearl Rates that do not include 12 month or 13 cycle rate
Statistical methods	Single-decrement life table analysis; Cycles with no intercourse are excluded from all analyses (for perfect or typical use; an exception can be made for learning cycles)		Insufficient detail or data provided.
Attrition	Loss to follow-up <20%; one year discontinuation rates (and reasons) clearly reported and examination of population differences for those who left the study still at risk for pregnancy (e.g. did not like the method, loss to follow-up)		
Other	No other major concerns about threats to internal validity		Other major concerns

Appendix 5. Study Detail From Moderate-Quality Studies

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
Calendar-based Standard Days Method and variants						
Arevalo (2002),[18,76,78] Standard Days Method	• Bolivia, Peru, Philippines • 478 women, 4,035 cycles • Age: 18–24 y: 24%; 25–29 y: 26%; 30–34 y: 29%; 35–39 y: 21% • Mean no. of births: 2.5 • Cycle length: 25–32 d	12.9 (8.5–15.3)	• 2nd y: 5.2 (1.8–8.5) • 3rd y: 3.4 (0.4–6.3)	4.8 (2.3–7.1)	• Consistent barrier use in fertile window (1st y): 5.7 (3.1–8.2)	7
Burkhart (2000),[26] Standard Days Method variant	• Guatemala • 301 women, cycles NR • Age: 18–24 y: 30%; 25–29 y: 30%; 30–34 y: 22%; 35–39 y: 19% • Mean no. of births: 1:15%; 2: 22%; 3: 23%; 4: 16%; 5+: 25% • Cycle length: 26–32 d	11.2 (7.6–14.9)	NA	NA	NA	6
Gribble (2008),[53] Standard Days Method	• Benin, Ecuador, El Salvador, Honduras, India, Philippines	14.1 (11.8–16.4)	• 2nd y, noncumulative 3.7 (1.9–5.6)	NA	NA	5

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
	• 1,646 women, cycles NR • Mean age: 28–32 y • Mean no. of births: 1.7–3.6 • Cycle length: 26–32 d		• 3rd y, noncumulative 5.9 (3.0–8.8)			
Sinai (2012),[77] postpartum bridge to Standard Days Method	• Peru, Guatemala • 157 women, 746.5 cycles • Mean age: Peru: 25 y; Guatemala: 26 y • No. of births: NR • Breastfeeding; before or within 1st menstrual cycle postpartum	NA	• 6 mo: 11.8 (6.0–17.2)	NA	• 6 mo: 3.7 (0.0–7.4)	7
Mucus-only methods, TwoDay Method						
Arevalo (2004),[19] TwoDay Method	• Guatemala, Peru, Philippines • 450 women, 2,928 cycles • Mean age: 29 y • Mean no. of births: 2.5	13.7 (9.9–17.3)	NA	3.5 (1.4–5.5)	• Consistent condom, withdrawal use during fertile time 6.3, or both (3.6–8.8)	7

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
	• Cycle length: less than 43 d					
Jennings (2011),[58] TwoDay Method standard and quick start	• Peru	NA	• 7-cycle standard: 3.5	NA	NA	4
	• Standard: 40 women, quick start: 120 women		• 7-cycle quick start: 9.9			
	• Cycles NR					
	• Mean age: 32.6 y					
	• Mean no. of births: 1.4–2.0					
	• Cycle length: NR					
Mucus-only methods, Billings Ovulation Method and variants						
Bhargava (1996),[23] Billings Ovulation Method	• India	10.5 (9.1–11.9)	• 18 mo: 14.1 (12.5–15.7)	1.1 (0.5–1.7)	• 18 mo: 1.2 (0.6–1.8)	6
	• 2,059 women enrolled, 21,579 cycles		• 21 mo: 15.9 (14.3–17.5)		• 21 mo: 1.5 (0.9–2.1)	
	• Mean age 26.2 y, range 15–35 y					
	• Mean no. of births: 2.5					
	• Cycle length: “regular cycles” (21–35 + or −5 d)					
	• Colombia	33.6	NA	NA	NA	4

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
Medina (1980),[70] Billings Ovulation Method	• 277 women, 1,967 cycles • Mean age: 26.9 y • Mean no. of births: 2.0 • Cycle length: NR					
Thapa (1990),[79] Billings Ovulation Method	• Indonesia • 453 women, 6,015.5 cycles • Mean age: 29.1 y • Mean no. of children 2.8 • "Regular cycles"	NA	• Learning phase: 2.4 (0.9–3.9)[c] •13 cycles after learning phase: 2.5 (0.9–4.1)[c]	NA	NA	5
Trussell (World Health Organization reanalysis) (1991),[11] Billings Ovulation Method	• New Zealand, India, Ireland, Philippines, El Salvador • 869 women, no. of cycles NR • Mean age: 30.1 y[d] • Mean no. of pregnancies: 3.9 • Cycle length: 25–35 d	22.8	NA	3.4	NA	11

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
Wade (1981),[81] Billings Ovulation Method	• United States • 573 women, no. of cycles NR • Mean age: 26.7 y • Mean no. of children: 1.0 • Cycle length: 24–36 d	22.4[e]	NA	NA	NA	5
Mucus-only methods						
Thapa (1990),[79] Modified Mucus Method and local variant	• Indonesia • Modified mucus: 209 women, 2,663 cycles • Local variant: 188 women, 1,828 cycles • Mean age 28.0 y; 27.7 y (local variant) • Mean no. of children 3.0; 3.1 (local variant) • "Regular cycles"	NA	• 3-cycle learning phase modified mucus: 3.2 (0.9–5.5),[c] local variant: 16.5 (11.2–21.8)[c] •13 cycles after learning phase modified mucus: 10.3 (6.0–14.6),[c] local variant: 11.5 (5.8–17.2)[c]	NA	NA	5

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
Mucus-only methods, Marquette Mucus-only						
Fehring (2013),[39] Marquette Mucus-only	• United States • 160 women, 1,075 cycles • Mean age: 30.4 y • Mean no. of children: 2.1 • Cycle length: 21–42 d	18.5	NA	2.7	NA	6
Fehring (2014),[38] Marquette Mucus-only[j]	• United States • 73 women, cycles NR, all perimenopausal • Mean age: 41.2 y, range 40–55 y • Mean no. of births: 2.6 • Cycle length: NR	4[g]	NA	NA	NA	6
Fehring (2017),[44], Marquette Mucus-only[k]	• United States • 118 women, 481 cycles • Mean age: 30.4 y • Mean no. of children: 2.4 • Cycle length: no restrictions; mean NR	8	• 24 cycles: 19	NA	NA	4

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
Basal body temperature–based methods						
Berglund-Scherwitzl (2016),[22] natural cycles	• Sweden • 4,054 women, 2,085 woman-y • Age: younger than 20 y: 1%; 20–24 y: 32%; 25–29 y: 43%; 30–34 y: 17%; 35–39 y: 5%; 40 y or older: 2% • No. of births: 0: 79%; 1: 11%; 2: 8%; 3: 1%; 4+: 1% • Mean cycle length: 29.9 d	9.8[f,i]	NA	NA	NA	4
Drouin (1994),[35] Bioself	• Canada • 83 women, 745 cycles • Age: younger than 25 y: 23%, 25–29 y: 37%, 30–34 y: 26%, 35 y or older: 15% • Births: 52% 0 children • Cycle length: NR	9.0	NA	NA	NA	3
Symptothermal methods, single check		NA		NA		4

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
Ecochard (1996),[36] French method	• Belgium, France, Switzerland • 626 women, 6,740 cycles • Mean age: 25–35 y • Parity: 3+ greater than 50% • Cycle length: NR		• New (26.7%) and experienced users: 17.6		• Intercourse limited to postovulation (1,211 cycles): 0 (0–3)	
Freundl (1999),[51,88] French method	• France, Great Britain, Spain • 214 women, 1,495 cycles • Mean age: 26.8–37.6 y • No. of children: 0: 8.5%, 1: 10.3%, 2: 23.9%. 3 or more: 57.3% • Cycle length: NR	NA	• New (6.5%) and experienced, 12 cycles: 8.5 (3.6–13.4)[g] • Abstinence-only, new and experienced: 19.2[f,i,l] • Mixed method users, new and experienced: 3.9[f,i,l]	NA	NA	4
Weeks (1982),[83] Billings Ovulation Method plus basal body temperature	• United States	13.2	•Abstinence-only users: 16.8[f]	NA	NA	3

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
	• 148 women, 1,104 cycles • Median age: 26 y, range 17–53 y • Mean no. of children: 1.4, 47% nulliparous • Cycle length: NR		• Barrier method use in any cycle: 16.4[f]			
Symptothermal methods, double-check						
Frank-Herrmann (2007),[47] Sensiplan[m]	• Germany • 900 women, 9,005 cycles • Age: 19–24 y: 24.7%, 25–29 y: 38.6%, 30–35 y: 24.7%, 35–39 y: 8.9%, 40–45 y: 3.2% • No. of births: 0: 51.9%, 1–2: 34.7%, 3 or more: 13.4% • Average cycle length 22–35 d (20% of cycles could deviate)	1.8 (1.0–2.6)	• 24 cycles of use: 2.6 (1.2–4.0) • 13 cycles abstinence only: 1.6 (−0.2 to 3.4) • 13 cycles mixed method users: 2.0 (0.6–3.4)	0.4 (0.1–1.6)	• Intercourse in fertile time always with barrier method: 0.6 (0.1–2.1)	7

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
Freundl (1999),[51,88] Sensiplan	• Austria, Belgium, Czech Republic, France, Germany, Great Britain, Ireland, Italy, Spain, Switzerland • 1,046 women, 16,856 cycles • Mean age: 26.8–37.6 y • No. of children: 0: 36.0%, 1: 13.6%, 2: 24.0%, 3+: 26.5% • Cycle length: NR	NA	• 12 mo new (60%) and experienced: 2.6 (1.4–3.8)[g] • 18 mo new (60%) and experienced: 4.8 (3.0–6.6)[g] • 24 mo new (60%) and experienced: 5.2 (3.2–7.2)[g] • 36 mo new (60%) and experienced: 5.7 (3.5–7.9)[g] • Abstinence-only users, new and experienced: 2.8[f,i,l] • Mixed method users, new and experienced: 2.5[f,i,l]	NA	NA	4
Medina (1980),[70] Thyma	• Colombia • 286 women, 1,882 cycles • Mean age: 27.5 y • Mean parity: 2.4 • Cycle length: NR	33.0	NA	NA	NA	4

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
Wade (1981),[81] Thyma	• United States • 590 women • Mean age: 26.9 y • Mean no. of children: 1.0 • Cycle length: 24–36 d	11.2[e]	NA	NA	NA	5
Urinary hormonal methods, Persona						
Bonnar (1999),[25] Persona	• England, Ireland, Germany • 710 women, 7,209 cycles • Median age: 30 y • Births: 30% 0 births • Cycle length: 23–35 d	25.6 (22.2–29.1)	NR	12.1 (9.3–14.8)	NR	5
Urinary hormonal methods, Marquette Monitor-only						
Fehring (2013),[39] Marquette Monitor-only	• United States • 197 women, 1,546 cycles • Mean age: 29.7 y • Mean no. of children: 1.8 • Cycle length: 21–42 d	6.8	NA	0.0	NA	6
	• United States	3[g]	NA	NA	NA	6

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
Fehring (2014),[38], Marquette Monitor-only[j]	• 35 women, cycles NR, all perimenopausal • Mean age: 41.2 y, range 40–55 y • Mean no. of births: 2.6 • Cycle length: NR					
Fehring (2017),[44], Marquette Monitor-only[k]	• United States • 212 women, 1,681 cycles • Mean age: 30.4 y • Mean no. of children: 2.4 • Cycle length: no restrictions; mean NR	2	24 cycles of use: 6	NA	NA	4
Symptohormonal methods, Marquette Monitor and Mucus						
Fehring (2014),[38] Marquette Monitor and Mucus[j]	• United States • 42 women, cycles NR, all perimenopausal • Mean age: 41.2 y, range 40–55 y	6[g]	NA	NA	NA	6

Author (Year), Method	Country or Countries, No. Enrolled (Primary Analysis), No. of Cycles (in Primary Analysis), Age, Pregnancy History, Cycle Length or Key Characteristics	Typical Use Pregnancy Rate/100 Woman-Years[a] (CI), 1st Year (13 Cycles) of Use	Other Typical Use Pregnancy Rates[a] (eg, Experienced Users, Other Time Points)	Perfect Use Pregnancy Rate[a] (CI), 1st Year (13 Cycles) of Use, New Users	Other Perfect Use Rates[a] (eg, Experienced Users, Other Time Points)	No. of Quality Criteria With High Rating (Out of 13)[b]
	• Mean no. of births: 2.6 • Cycle length: NR					
Fehring (2017),[44] Marquette Modified Mucus Method[k]	• United States • 333 women, 3,086 cycles • Mean age: 30.4 y • Mean no. of children: 2.4 • Cycle length: no restrictions; mean NR	7	24 cycles of use: 18	NA	NA	4

NR, not reported; NA, not applicable.
[a] Rates presented are single decrement life table rates unless otherwise specified by a subnote.
[b] Number of quality criteria ranked high is presented by comparison.
[c] Adjusted for age.
[d] Mean age calculated based on those entering the teaching phase.
[e] Multiple decrement life table.
[f] Pearl rate.
[g] Kaplan-Meier.
[h] Pearl rate includes 61 participants whose pregnancy status could not be determined at the time of dropout.
[i] Pearl rate recalculated from original with 1,300 instead of 1,200 as number of cycles.
[j] Includes some data from other included studies published in 2009 and 2011. Unlike the two source studies, this analysis met our quality criteria and was thus included in this table.
[k] Fehring 2017 includes some women from a study published in 2011 (not included in this table because it did not meet quality criteria) and a small number from Fehring 2014 included in this table.
[l] Subgroup from an earlier preliminary analysis
[m] Freundl 1999 includes incomplete overlap with data from Frank-Herrmann 2007. In the Freundl 1999 study, 339 women with 7362 cycles were included from the Frank-Herrmann site up to 1995. In Frank-Herrman 2007, data from 900 women and 17,368 cycles were reported up to 2005.

Appendix 6. Detailed Quality Ranking Forms

Note: for proposed rank, 1 means high quality, 2 means moderate quality, 2* means moderate quality for a criterion for which there was no low-quality marker, and 3 means low quality

Calendar-Based Methods: Standard Days Method (SDM) and Variants

STUDY: Arevalo 2002
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion criteria clearly described (18-39 years old, married or with stable partner, 26-32 day cycles, willing to avoid intercourse 12 consecutive days every cycle, willing partners, no history of infertility, no high risk of STI, no contraindication for pregnancy), but no refusal rate provided
Exclusion of populations not at meaningful risk of pregnancy	2	Excluded cycles without sexual intercourse (0.35%); also excluded cycles in which another method of family planning was used on days other than the fertile window of days 8 through 19; history of 12+ cycles of unprotected sex and no conception not explicitly excluded; women using sterilization, IUDs or hormonal methods also not explicitly excluded
Treatment of women in various reproductive categories that may impact fertility	2*	Included breastfeeding and post-partum (but not amenorrheic) participants and did not provide subgroup analyses. Other subgroups likely or explicitly excluded
Description of study population	1	All new users and demographic characteristics well described, including age and parity
FABM method(s) studied	1	Standard Days Method with CycleBeads well-referenced and described
Teaching of method	2*	Trained teachers but training details not reported; number and frequency of encounters reported; curriculum and materials reported
Detection of pregnancy	1	All charts reviewed for pregnancy with targeted pregnancy testing for cycles longer than 42 days; also screened for possible pregnancy at every encounter, although most pregnancies were based on self-report
Classification of pregnancy as intended/unintended	2	Pregnancy intentions assessed periodically, but not prior to every cycle. All cycles and pregnancies were included. Teaching phase pregnancies included in estimates

Concurrent use of coitus-dependent methods	1	Subgroup analysis of couples who used barrier or withdrawal during fertile window was presented
Study duration	1	12 month/13 cycle rate reported for typical use
Statistical methods	1	12 month/13 cycle typical use single decrement life table pregnancy rate; cycles with no intercourse included
Attrition	2	Loss to f/u = 7.1% at 13 cycles; reasons for discontinuation reported but not analyzed by differences in demographic variables
Other	1	None
Overall rank	2	
Results included		13 cycle single decrement life table cumulative pregnancy rate of 4.75% for correct use and 11.96% for typical use. 1st year pregnancy rate for condom or withdrawal at every act of intercourse (5.7%)

STUDY: Burkhart 2000
Quality criteria rankings

Domain	Proposed Ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion criteria clearly defined (18-39 years, no sterilization of woman or man, 26-32 day cycles, wanted to avoid pregnancy for next 12 months, not currently using contraceptive method, had not used hormonal contraception in last 3 months, breastfeeding with at least 3 cycles, male partner willing to participate, married or in union, living together for at least one year, willing to abstain from vaginal intercourse for 11 days); no refusal rate provided
Exclusion of populations not at meaningful risk of pregnancy	2	Women with infertility not explicitly excluded; Cycles with no intercourse not explicitly excluded; Women using contraceptive method or had used in last 3 months excluded
Treatment of women in various reproductive categories that may impact fertility	1	Excluded perimenopausal women, breastfeeding amenorrhoeic women, oligomenorrheic women, postpartum and 3 cycles post hormonal contraception (only 7% had used a modern method so likely the rate of depo Provera use <9 m. would be extremely low)
Description of study population	1	All new users; age and parity described
FABM method(s) studied	1	Modification of Standard Days Method; clear rules (Days 9-19)
Teaching of method	2	Teacher training noted but no detail; number of encounters described; materials referenced (cycle beads and coital log)
Detection of pregnancy	1	Targeted pregnancy testing for long cycles; active follow-up described for any open-ended cycles at the end of the study period
Classification of pregnancy as intended/unintended	2	Unclear how often pregnancy intentions were assessed (may have been as frequently as every 2-3 months or as infrequently as only at baseline), but was not as frequent as monthly. No pregnancies or cycles excluded for avoiding women and no learning phase pregnancies excluded
Concurrent use of coitus-dependent methods	2*	Use not assessed , though instructed to avoid barrier methods and withdrawal while using the method
Study duration	1	12m study; 12m pregnancy rates reported
Statistical methods	2	Single decrement life table; cycles with no intercourse were not explicitly excluded
Attrition	2*	Loss to follow-up <10% (2 people); discontinuation reasons reported but no examination of differences by population
Other	1	None
Overall rank	2	
Results included		12 month typical use pregnancy rate

STUDY: Dicker 1989
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	3	Inclusion/exclusion criteria poorly characterized, and no refusal rate documented
Exclusion of populations not at meaningful risk of pregnancy	2	Poorly described exclusion criteria, but no explicit inclusion of women using hormonal contraception
Treatment of women in various reproductive categories that may impact fertility	2*	Not reported if infertile women, cycles with no intercourse and women using other contraceptive methods were excluded; however, there was no explicit inclusion of women using hormonal methods, sterilization or IUD in the analysis
Description of study population	2	New users with age and marital status only described. Parity not described
FABM method(s) studied	1	Clearly described variant of a standard days method; asked to avoid on days 12-17 of the cycle
Teaching of method	2*	Teaching method, number of encounters and curriculum used were not described or referenced
Detection of pregnancy	3	No reporting on methods for detection of pregnancy
Classification of pregnancy as intended/unintended	2	Insufficient information about how pregnancy intentions were assessed. No learning phase described. No explicit exclusion of pregnancies from the analysis
Concurrent use of coitus-dependent methods	2*	Barrier method and withdrawal not assessed though women were asked to choose between barrier methods and rhythm
Study duration	3	24 months of planned for follow-up and pearl rates reported
Statistical methods	2	Pearl rate for 24 months reported making it not comparable to other studies. Cycles with no intercourse not explicitly excluded
Attrition	2	Loss to follow-up likely <20% as total discontinuation was 20.3%. Discontinuation rates and reasons only partially reported. For example, loss to follow-up not included
Other	3	Comparison study between users of other methods but there was no accounting for possible confounders and no statistical assessment of differences
Overall rank	3	
Results included		None

STUDY: Gribble 2008
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion limited but clear as it was a more pragmatic design for a program evaluation (regular cycles (26-32 day), willing and had partner willing to avoid unprotected sex during fertile days. No refusal rate provided
Exclusion of populations not at meaningful risk of pregnancy	2	No explicit or presumed attempt to exclude cycles without sex, women/couples with history of 12+ cycles of unprotected sex and no conception or women using sterilization, IUDs or hormonal methods. But no explicit inclusion of users of other contraceptive methods in the analysis
Treatment of women in various reproductive categories that may impact fertility	2*	Inclusion/exclusion of breastfeeding, post-partum, post-hormonal or oligomenorreic women not reported. Perimenopausal women were included but no subgroup analysis presented
Description of study population	1	All new users and demographic characteristics described as mean age and parity across the 14 countries
FABM method(s) studied	1	Standard Days Method with CycleBeads well referenced and described
Teaching of method	2*	Trained teachers but training details not reported; number and frequency of encounters reported; curriculum and materials reported
Detection of pregnancy	1	Targeted pregnancy detection with women who had cycles longer than 40 days given a pregnancy test and active follow-up at the conclusion of the study. Details confirmed with author Victoria Jennings
Classification of pregnancy as intended/unintended	2	Unclear frequency of pregnancy intention assessment but probably less than monthly. No pregnancies excluded from the analysis
Concurrent use of coitus-dependent methods	2*	Subjects reported using abstinence or condoms on fertile days but this was not described in the analysis and no subgroup analysis was done
Study duration	1	Planned duration of 13 months
Statistical methods	2	12-month pregnancy rates were calculated using life-table analysis but cycles with no intercourse not explicitly excluded
Attrition	2*	Insufficient detail of attrition or reasons for attrition reported
Other	1	None
Overall rank	2	
Results that should be in table		1st year 13-cycle typical use cumulative pregnancy rate of 14.1 per 100 woman-years of use as well as 2nd and 3rd year rates from Sinai 2012

STUDY: Kursun 2014
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	1	Standard Days method (SDM) presented as one method among others at a family planning service delivery setting; up to 250 women could enroll in SDM or until registration period ended; 99 out of 993 women chose SDM; 15 of these were excluded because spouses objected (leaving N=84); oligomenorrheic women excluded
Exclusion of populations not at meaningful risk of pregnancy	2	Cycles with no intercourse not explicitly excluded. No clear attempt to exclude sub-fecund couples, although only 5% of SDM acceptors had no history of pregnancy. Study separately assessed women using different methods, such as IUDs or hormonal methods so these were not included in the effectiveness calculation for SDM
Treatment of women in various reproductive categories that may impact fertility	2*	Excluded oligomenorrheic women; perimenopausal women were included (<5% of total population but unclear exact percent); Unclear if breastfeeding and postpartum women were included. Among SDM users in the study, the last method used in the 3-month period prior to the study was: withdrawal (n=10), condom (n=46), IUD (n=13), nothing (n=11), or "other" (n=4) (in other words, at most, <5% were post-hormonal contraceptives, if any)
Description of study population	1	Demographic characteristics provided with age and parity.Likely only new users or vast majority new users only as SDM had not been offered in regular service setting in Turkey prior to this study (in addition, <5% of women listed "other" as their previous method). See Table 2
FABM method(s) studied	1	SDM (with use of Cycle Beads); with clarity on rules provided
Teaching of method	1	Midwives and nurses providing FP services at participating centers were trained on SDM in an 8-hour course; details of method provided, plus cycle beads as curricular detail. Clients were counseled face-to-face and were given brochures with illustrations
Detection of pregnancy	3	Likely self-report only (follow up interviews were conducted by clinic staff every 3 months by phone), no charts (using beads)
Classification of pregnancy as intended/unintended	2	Baseline assessment of pregnancy intentions only; no evidence that any unintended pregnancies were inappropriately excluded from analysis
Concurrent use of coitus-dependent methods	2*	SDM not assessed with and without barrier method use
Study duration	1	12-month typical pregnancy rate from life tables reported
Statistical methods	3	12-month life table analysis of pregnancy rate; unclear if single or multiple decrement
Attrition	2	As per Table 3, 7 of 84 SDM users were lost to follow up (8.3%); text states "spectrum of contraceptive methods chosen by the women who were followed up to termination or the one-year period did not differ from the methods selected by those lost to follow up (p=0.115)."; no examination of attrition by population differences amongst SDM users
Other	2	Small sample size (n=84)
Overall rank	3	
Results included		None

STUDY: Sinai 2012
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Some inclusion criteria are provided; refusal rate not reported
Exclusion of populations not at meaningful risk of pregnancy	2	Not reported whether exclusions made according to listed characteristics; no explicit inclusion of hormonal method users Included only postpartum women
Treatment of women in various reproductive categories that may impact fertility	1	All women were breastfeeding and amenorrheic. Only women in one reproductive category so did not need to analyze further
Description of study population	1	New users of this method. Age and parity recorded
FABM method(s) studied	1	SDM Bridge with its own clear rules
Teaching of method	1	Training reported (no details); curriculum reported (cycle beads); number of encounters reported
Detection of pregnancy	1	Targeted pregnancy testing of anyone whose cycles lasted more than 42 days; active follow-up
Classification of pregnancy as intended/unintended	2	Pregnancy intentions not collected every cycle. 2 pregnancies were excluded, but before women were eligible for the method
Concurrent use of coitus-dependent methods	2*	Barrier method assessed; no subgroup analysis
Study duration	2	6-month life table pregnancy rate
Statistical methods	2	Used single-decrement analysis, but unclear if cycles with no sexual activity were excluded. Perfect use rate correctly calculated
Attrition	1	Loss to follow-up <20%; reasons for discontinuation reported
Other	1	None
Overall rank	2	
Results included		Typical use 6-month pregnancy rates, and perfect use rates with note that this is for a special population of breastfeeding and amenorrheic women

Calendar-based Method: Rhythm

STUDY: Kambic 1996
Quality ranking criteria

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	3	Inclusion criteria not defined
Exclusion of populations not at meaningful risk of pregnancy	2	Not defined Did not exclude cycles without sexual intercourse
Treatment of women in various reproductive categories that may impact fertility	2*	Not addressed
Description of study population	2	Unknown whether includes established users No description of study population, not even age
FABM method(s) studied	3	Calendar rules not adequately described, all studies combined, including different and unknown rules
Teaching of method	2*	No description
Detection of pregnancy	3	Not addressed
Classification of pregnancy as intended/unintended	3	Not addressed; unknown
Concurrent use of coitus-dependent methods	2*	No assessment of any use of barriers or withdrawal.
Study duration	3	Variable across different studies (attempted to account for this using a linear regression model of duration of use as independent variable and pregnancy rate as outcome)
Statistical methods	2	Estimated 12-month pearl rate
Attrition	2*	Not addressed

Other	3	Concern about the following statement and the validity of the results: "Of the studies in Table 1, only those of Tietze et al. [1951 Fertil Steril] and Dicker et al. [1989 Contraception] are clearly reported trials. The remainder of the reports are subject to interpretation and are contestable."
Overall rank	3	
Results that should be in table		None

STUDY: Guerrero 1970
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	3	Inclusion criteria inadequately defined. Only states that women had shown an interest and lived in the geographic area served by social workers. A refusal rate was provided
Exclusion of populations not at meaningful risk of pregnancy	3	Cycles with no intercourse and sterilized women not explicitly excluded. Women using hormonal contraceptives and treatments (progestins for abnormal bleeding/amenorrhea) were included in the study
Treatment of women in various reproductive categories that may impact fertility	2*	Perimenopausal women were included and not analyzed separately. It is unclear if women with long cycles or postpartum were included. Breastfeeding women may have been included and treated with hormones to induce acycle, and then entered into the study. "As mentioned, the use of the pill was restricted exclusively to the cases of menstrual irregularities [cycles that varied more than 10 days; women treated with "anovulatory drugs" for up to a year] or to the postpartum period." (p. 547, right)
Description of study population	2	Likely all new users but not explicit. Age and parity are characterized and subgroup analyses are performed
FABM method(s) studied	1	Ogino rhythm for preovulation and Knaus temperature with citations
Teaching of method	2*	Teacher training and curriculum details were not reported. Number of visits and duration reported; requirement that husband participate in initial sessions
Detection of pregnancy	3	Unclear pregnancy detection methods
Classification of pregnancy as intended/unintended	2	No pregnancies were explicitly excluded. Pregnancy intentions only measured at baseline
Concurrent use of coitus-dependent methods	2*	Not reported how barrier method use and withdrawal was assessed or how prevalent it was, but women using these methods were reportedly excluded from the effectiveness estimate
Study duration	3	Study was longer than 12 months and reports only Pearl rates for the entire study period
Statistical methods	3	Unclear whether cycles with no intercourse were excluded. Unclear statistical methodology
Attrition	2*	210 women left the study and an additional 276 moved out of the study area. Unclear how many were loss to follow-up
Other	1	None
Overall rank	3	
Results included		None

STUDY: Tietze 1951
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	1	Inclusion criteria adequately described. All women who qualified were enrolled in this clinic database study. Women were referred to the clinic specifically for Rhythm method if experiencing regular cycles (did not vary by more than 8 days) and if sexually active
Exclusion of populations not at meaningful risk of pregnancy	2	Cycles with no intercourse, sterilized women and women using contraceptives were not explicitly excluded; however, women were referred to this clinic specifically for fertility awareness method in the 1940s and 50s. Therefore, very unlikely that any women were concurrently using other methods
Treatment of women in various reproductive categories that may impact fertility	2*	Perimenopausal women included, no subgroup analysis presented. Women with cycle variation >8 days excluded which would likely include amenorrheic women. Unclear whether breastfeeding, post-hormonal contraceptive or postpartum women were included
Description of study population	1	Age, parity well described and all likely new users as they were referred to this clinic to learn the method
FABM method(s) studied	1	Rhythm method with citations and clearly described rules
Teaching of method	2*	Trained healthcare professionals taught the women and followed up with them monthly but details of curriculum and teacher training not reported
Detection of pregnancy	3	Unclear
Classification of pregnancy as intended/unintended	2	Unclear how often pregnancy intentions were assessed, but all women and pregnancies included in the analysis. Authors tried to include the pregnancy rate of those lost to follow-up by randomly finding a subset through home visits and imputing the pregnancy rate to the whole group
Concurrent use of coitus-dependent methods	2*	Not assessed
Study duration	3	No upper limit; not clearly reported and only reported Pearl rates.
Statistical methods	3	Study did not calculate 12 month Pearl rate and did not provide data to allow for calculations; total pregnancy rate calculated with an imputation procedure (see above)
Attrition	2*	Loss to follow-up greater than 20% with no differential treatment
Other	1	No other major concerns
Overall rank	3	
Results included		None

Mucus-Only Method: Billings Ovulation Method (BOM) and variants

STUDY: Ball 1976
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Some inclusion criteria described (age 20-39; carried at least one pregnancy to term; observed at least one ovulatory cycle since last birth); no information on refusal rate
Exclusion of populations not at meaningful risk of pregnancy	2	Unclear if excluded cycles without sexual activity, infertile women or women who were using other contraceptive methods. No explicit suggestion that study included women on other methods, especially since a reason for exit from the study included changed to another method
Treatment of women in various reproductive categories that may impact fertility	2*	Excluded: perimenopausal. Unclear if women with oligomenorrhea, breastfeeding, or post-hormonal contraception were included. Possibly included: fewer than three cycles, postpartum
Description of study population	3	Experience with the method is unclear/not reported. Women were recruited from NFP centers, so may have already been using the method. Age and parity described
FABM method(s) studied	1	Billings Ovulation Method; rules described
Teaching of method	2*	No information provided on any aspect of teaching, curriculum or number of visits
Detection of pregnancy	3	Likely self-report only, with no evidence of systematic corroboration with charting or testing among all participants included in analysis
Classification of pregnancy as intended/unintended	2	Pregnancy intentions not prospectively assessed; no evidence that any pregnancies were inappropriately excluded
Concurrent use of coitus-dependent methods	2*	Neither barrier nor condom use was assessed
Study duration	3	Mean duration of follow up not reported; maximum number of cycles was 22; mean of 13.2 cycles. Duration was longer than 12 months/13 cycles for some participants, and only Pearl rates were provided
Statistical methods	3	Insufficient detail provided on analysis. Errors in the tables, perfect use Pearl rates incorrectly calculated
Attrition	2*	Of the 124 women, 2 were reported as lost to follow-up. No examination of attrition by population differences
Other	2	Errors in tables and calculations, as well as lack of detailed description of methods
Overall rank	3	
Results included		None

STUDY: Bhargava 1996
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Some inclusion criteria were provided (healthy women volunteers aged 15-35, with regular menstrual cycles (26-31 days +/- 5 days), with husband support for participation); refusal rate not provided.
Exclusion of populations not at meaningful risk of pregnancy	2	Not reported whether cycles without sexual activity were excluded. Not reported whether sub-fertile/infertile couples were excluded. Paper states that "no modern methods of FP were used" by these women during the trial period and that women were discontinued if they began using another contraceptive
Treatment of women in various reproductive categories that may impact fertility	2*	Perimenopausal: excluded; Oligomenorrhea: excluded; breastfeeding, amenorrhea post-partum, or post-hormonal contraceptive use: not reported
Description of study population	1	Whether participants were new or experienced users of the method was not explicitly reported, but only 2.4% reported ever using a traditional/natural method in the past. Age and parity were characterized
FABM method(s) studied	1	Billing Ovulation Method with reference to rules
Teaching of method	2*	Instruction delivered to female partners using paper materials; but unclear type or training of educators, experience of teachers, or educational intensity
Detection of pregnancy	2	Author (Saxena) clarified that participants completed a menstrual diary card, which was cross-checked monthly by a trained teacher/social worker for accuracy (participants who missed a visit received a home visit the next day). Participants who reported missing a menstrual period were clinically examined by a doctor and administered a pregnancy test. Thus; self-report with systematic corroboration with charting or testing all participants included in analysis. No active follow-up specified
Classification of pregnancy as intended/unintended	2	Women were discontinued from the study if they were planning for pregnancy (or for other reasons), but it is not reported when or how pregnancy intention was collected; no evidence that any unintended pregnancies were inappropriately excluded
Concurrent use of coitus-dependent methods	1	Authors report "no modern methods of family planning were used by these women during the trial period, either as a backup or combination method"
Study duration	1	Included at least 12 m/13 cycles (cut off period was 21 months) and included a 12 m single decrement typical pregnancy rate report

Statistical methods	2	Used single-decrement analysis, but unclear if cycles with no sexual activity were excluded. Calculated perfect use rates correctly
Attrition	1	Loss to follow up was 16.3% at 21 months (not reported at 12 months, but certainly <20%); discontinuation rates and reasons are clearly reported and examination of attrition by population differences (urban/rural) is reported
Other	1	Women were discontinued due to husband's non-cooperation or to women's lack of comprehension of the method; i.e., incorrect charting of observations for 3 consecutive cycles - the date of last follow up visit. This may limit study generalizability
Overall rank	2	
Results included		12-month rate (perfect and typical) pregnancy probabilities. 95% confidence intervalscalculated from standard errors provided

STUDY: Gomes 1988
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	3	Criteria minimally defined (registered in NFP center 9/1/81-8/31/82, kept chart for 3 months desires to avoid pregnancy) and no refusal rate provided
Exclusion of populations not at meaningful risk of pregnancy	2	Exclusion of cycles with no intercourse and other key populations not recorded, but no explicit inclusion of contraceptive using women
Treatment of women in various reproductive categories that may impact fertility	2*	No report on whether women in any of the categories was excluded and no subgroup analyses
Description of study population	3	Not reported whether new or experienced users. Age and parity described with other demographic characteristics
FABM method(s) studied	2	Described using Billings Ovulation Method, but no reference to rules used
Teaching of method	2*	No detail on teacher training, curricula used, and/or number of encounters provided
Detection of pregnancy	3	Unclear pregnancy detection procedures. Appears to be self-report; with no evidence of systematic corroboration by charting
Classification of pregnancy as intended/unintended	3	Only reported determining pregnancy intentions at baseline; no obvious exclusion of pregnancies from the calculation but given no data about how unintended pregnancies were defined, this is unclear. Also, unclear if pregnancies may have been excluded from the learning phase of 3 cycles
Concurrent use of coitus-dependent methods	2*	Barrier method and/or withdrawal method use not assessed
Study duration	1	12-month typical use pearl rate reported
Statistical methods	2	12-month typical use pearl rate reported; cycles with no intercourse may have been included
Attrition	2*	Loss to follow-up not reported but mean of 5-6 cycles per woman included in analysis so this could be high. No reporting of discontinuation rates and reasons
Other	1	None
Overall rank	3	
Results included		None

STUDY: Johnston 1979
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	3	Attempt to reach all active clients of the clinics; estimated over 65% participation. No exclusion criteria or inclusion criteria reported
Exclusion of populations not at meaningful risk of pregnancy	2	No evidence of inclusion of women using hormonal methods. Unclear how cycles with no intercourse were treated, there was a risk-based adjustment for cycles of low risk. Unclear if women with sub-fertility excluded
Treatment of women in various reproductive categories that may impact fertility	2*	Unclear if breastfeeding, postpartum, post-hormonal contraception women were included in the study and no subgroup analyses reported by these
Description of study population	3	Unknown proportion of new vs. experienced users. Age and parity described and subgroup analyses
FABM method(s) studied	1	Billings ovulation method, Symptothermal method variant A and Symptothermal method variant B. Clear discussion of rules in early methods paper publication
Teaching of method	2*	Teacher training not described; number of encounters not reported; Billings curriculum referenced
Detection of pregnancy	3	Pregnancy confirmed by medical diagnosis or test, but only after self-reported on the questionnaire
Classification of pregnancy as intended/unintended	3	Pregnancy classification characterized partially by retrospective assessment of the behavior of the couple. May have excluded pregnancies that were retrospectively designated as unplanned pregnancies
Concurrent use of coitus-dependent methods	1	Barrier method and withdrawal use were assessed and pregnancy rates were calculated separately for each method and for each method mixed with other method use
Study duration	2	Study was longer than 12 months, but reports life table rates for all users
Statistical methods	3	Unclear if single v. multiple decrement life table. Pearl rate presented is for more than 12 months/1 year. Perfect use appears to include all cycles in denominator and therefore improperly calculated
Attrition	2*	Dropout rates assessed and broken down by cycle length, method, frequency of intercourse and by occurrence of unplanned pregnancies. Unclear reasons for dropout
Other	1	None
Overall rank	3	
Results included		None

STUDY: Klaus 1979

Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	1	Provides a denominator for screened population/refusal rate. Inclusion: women who had charted their cycle for 30 days, were willing to commit themselves to use of this method for fertility control, and to provide information for the study
Exclusion of populations not at meaningful risk of pregnancy	2	Estimates were available that excluded women using "fertility suppression methods" so did not exclude women using contraception. Did exclude cycles without sexual activity, but not unclear whether sub-fertile women were excluded
Treatment of women in various reproductive categories that may impact fertility	2*	Did not exclude or present subgroup analysis for perimenopausal (8% of study pop was 40-44 and 3% was 45+), or breastfeeding or postpartum women (11% were post-partum, post-abortal, or lactational). Not reported if included oligomenorrheic women, post-hormonal contraceptive users ("prior contraceptive histories not available for all subjects"), or STI-positive individuals
Description of study population	3	Included experienced users: "Prior contraceptive histories are not available for all subjects... additional fertility acceptance methods were reported 739 times" (top page 617). Also, as per note in Table, in life table analyses presentation, all segments were considered to be first segments (would bias failure rates downward since the analysis treats all segments as though they are new users)
FABM method(s) studied	1	Billings Ovulation Method with clear statement of the rules
Teaching of method	2*	Educational setting unclear; type of educators unclear; training of teachers including hours of education, length of training, and topics covered unclear; recommended educational intensity unclear; experience of teachers unclear
Detection of pregnancy	3	Not reported, didn't require charts to be turned in monthly. Most participants provided monthly charting of pregnancy status, but 18% of population relied only upon memory
Classification of pregnancy as intended/unintended	2	No evidence that any pregnancies were excluded without being prospectively classified as intended. However, pregnancy intentions were not prospectively assessed every cycle
Concurrent use of coitus-dependent methods	1	Assessed for concurrent method use and conducted subgroup analysis
Study duration	1	Duration was at least 12 months; and a 12-month typical use pregnancy rate was reported

Statistical methods	3	Insufficient detail about methods (i.e., how many women are included in the life table analyses). Unclear if there was a learning phase. All re-entries are considered new entries, which would bias rates downward. Calculations are multiple decrement rather than single decrement. Perfect use analysis incorrect
Attrition	1	At 12 months, 22.1% voluntary withdrawal but only 2.9% loss to follow-up. Discontinuation rates reported and life table analysis should have appropriately accounted for censoring
Other	3	Record keeping was not consistent across centers; nearly 1 in 5 women didn't chart but instead relied on memory
Overall rank	3	
Results included		None

STUDY: Labbok 1988
Quality criteria rankings

Domain	Proposed Ranking	Rationale
Inclusion exclusion criteria	3	Inclusion: secondary analysis of Billings Ovulation Method users from 42 Kenyan sites. Exclusion: sterilized, unable to keep, not trying to avoid. However, details were not provided on the original inclusion criteria. Reported citation is not peer-reviewed or available. No refusal rate
Exclusion of populations not at meaningful risk of pregnancy	2	Excluded sterilized women; Unclear if contraceptive users were included or whether cycles of no intercourse were included
Treatment of women in various reproductive categories that may impact fertility	2*	Included lactating amenorrhea women and did separate subgroup analysis; not reported whether perimenopausal women, oligomenorrheic women, postpartum women and recent contraceptive users were included. No subgroup analyses were provided
Description of study population	3	Data collected from time the couple attended the first training session, unclear if new users; age and parity not characterized though estimates presented for those<30 and >30 years
FABM method(s) studied	2	Billings ovulation method but method rules and/or citation not reported
Teaching of method	2*	Insufficient detail on teacher training, curriculum used and number of educational encounters
Detection of pregnancy	3	Insufficient information; chose not to contact authors as this was the only poor quality indicator
Classification of pregnancy as intended/unintended	2	No pregnancies explicitly excluded but only record of pregnancy avoidance done at baseline
Concurrent use of coitus-dependent methods	2*	Use of barriers and/or withdrawal was not assessed
Study duration	1	12month study; 12 month pregnancy probabilities reported
Statistical methods	3	Unclear whether single or multiple decrement life table "cumulative 12-month" life tables; Unclear whether cycles with no intercourse were included
Attrition	2*	Details about loss to follow-up and discontinuation rates and reasons were not reported
Other	3	Very minimal methodologic data; information was previously presented in non-peer reviewed publication
Overall rank	3	
Results included		None

STUDY: Mascarenhas 1979
Quality criteria rankings

Domain	Proposed Ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion criteria clear (<44 years old, regular menses, if postpartum or lactating must have had 2 regular cycles, willing to participate, cohabitating couples, willing to use the method for 16 cycles, willing to not use other method s of fertility regulation), refusal rate not reported
Exclusion of populations not at meaningful risk of pregnancy	2	Did not exclude cycles without sexual activity. Did not explicitly exclude women/couples with subfertility. Explicitly excluded women using other methods of fertility regulation
Treatment of women in various reproductive categories that may impact fertility	2*	Excluded women with oligomenorrhea (23-35 day cycles) and breastfeeding and postpartum women. Perimenopausal women were included and no subgroup analysis. Unclear if women who were recent users of contraceptives were excluded
Description of study population	2	All new users, but no information regarding distribution or proportion of age, parity, or other sociodemographic characteristics
FABM method(s) studied	2	Billings Ovulation Method, but start of fertile window not explicitly described
Teaching of method	2*	No detail about t teacher training or materials, curricula used and/or number of educational sessions
Detection of pregnancy	3	Unclear methodology with no explicit description of collecting and analyzing charts for signs of pregnancy
Classification of pregnancy as intended/unintended	2	Based on title and low pregnancy, seems likely that the population of users intended to avoid at entry (though this is not clearly stated) and no description of routine monthly assessment of pregnancy intentions
Concurrent use of coitus-dependent methods	2*	Barrier method and/or withdrawal use not assessed, although women were asked at enrollment not to use other methods
Study duration	1	Up to 16 cycles
Statistical methods	2	Life table of unknown type and Pearl Rate (up to 16 cycles) reported
Attrition	2	Loss to follow-up <20%, but no characterization of any reasons other than pregnancy for discontinuation
Other	2	Brevity of report and lack of methodologic details provided limits capacity to evaluate study quality
Overall rank	3	
Number of "1" s	1	
Results included		None

STUDY: Medina 1980

Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Clearly defined inclusion criteria, no refusal rate reported.
Exclusion of populations not at meaningful risk of pregnancy	2	Unclear whether subfertile, sterilized or other contraceptive users were excluded, but no explicit inclusion of contraceptive using women
Treatment of women in various reproductive categories that may impact fertility	2*	Perimenopausal women were excluded, but unclear whether ologomeorrheic, postpartum, lactating and recently post-hormonal contraception users were excluded. No subgroup analyses were reported
Description of study population	2	Likely all new users, but not clearly/explicitly stated; demographic characteristics well-described, including age and parity
FABM method(s) studied	1	Billings Ovulation Method and Thymas double-check symptothermal method both cited references for the rules of fertility
Teaching of method	2*	"Well-trained" teachers; training program lasted 3-5 months, educational setting and type of trainers unclear, materials were referenced
Detection of pregnancy	2	Women had monthly visits in home to review charts and "determine pregnancy status"; no evidence of active follow-up or routine pregnancy testing
Classification of pregnancy as intended/unintended	2	Only described at baseline, although there were monthly visits that could have elucidated this; learning phase pregnancies were included in some of the analyses but estimates including learning phase pregnancies were reported in which all pregnancies and cycles seem to be included
Concurrent use of coitus-dependent methods	2*	Use of barriers and or withdrawal not assessed
Study duration	1	12-month typical use pregnancy rate reported (in life tables [both single and multiple decrement] and Pearl rates)
Statistical methods	2	12-month typical use single decrement life table pregnancy rate; unclear if cycles with no intercourse were included
Attrition	1	Loss to follow-up<20%; reasons for discontinuation reported and demographic differences between total population and dropouts examined in Table 1. Note: very high discontinuation rate which might impact generalizeability
Other	1	No other concerns. Randomized comparison of STM vs Billings provides a comparative analysis of two FABMs
Overall rank	2	
Results included		12-month single decrement life table pregnancy rates for each arm including learning phase (i.e., 33.6 for Billings and 33 for STM)

STUDY: Perez 1983
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	1	Refusal rate provided; conducted physical exam and pap, treated for STIs, but minimal detail on who was included. Some women (3.9%) were subsequently "urged to abandon the method" for having anovulatory cycles, long fertile periods, emotional tension, problems with partner, or inability to learn the method after 4 months of instruction
Exclusion of populations not at meaningful risk of pregnancy	2	Unclear if cycles with no intercourse were excluded or women who were using contraception, sterilized and/or sub-fertile. No obvious inclusion of women using hormonal contraception
Treatment of women in various reproductive categories that may impact fertility	2*	Perimenopausal-age women (40-44) were included and there was no subgroup analysis. Inclusion of breastfeeding, postpartum, recently post-hormonal contraception or oligomenorrheic women were included
Description of study population	2	95% new users of Billings, but their data were not analyzed separately though 23.2% had previously used another FABM. Demographic well-described including age and parity
FABM method(s) studied	1	Billings Ovulation Method with reference to publication providing details of method
Teaching of method	2*	Training of teachers reported; frequency of encounters reported (up to 30 visits); curricula materials not referenced
Detection of pregnancy	3	Self-reporting at regular visits, no systematic reference to charts or pregnancy testing reported
Classification of pregnancy as intended/unintended	2	Assessed at baseline only, not prospective by cycle. No pregnancies explicitly excluded
Concurrent use of coitus-dependent methods	2*	Barrier and or withdrawal use not assessed
Study duration	1	24 months, with life table rates calculated for 12 months
Statistical methods	2	Life tables for 12 months, but cycles with no intercourse not excluded
Attrition	2*	Loss to follow up was documented month-by-month and was less than 20% (86/660). Pregnancy-related reasons for leaving study are of concern as intentions were not assessed prospectively and 66 left "due to planned pregnancies or planning to get pregnant"
Other	1	Exclusion of 1.5% who could not learn to evaluate their fertile and infertile periods over 4 months may limit study generalizability
Overall rank	3	
Results included		None

STUDY: Thapa 1990
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	1	Cohabiting women 20-39 with history of regular menstrual cycles (defined subjectively), no indication of primary or secondary infertility, non-pregnant (menstrual cycle began in last week), willingness to chart and keep records at least during first 3 teaching cycles for BO and MM methods. 850/912 satisfied study eligibility criteria. Exclusion not based on whether subject was able to learn/practice the method
Exclusion of populations not at meaningful risk of pregnancy	2	No intention to use any other method of fertility regulation while using periodic abstinence (three methods of FABM). No exclusion of cycles without sexual intercourse
Treatment of women in various reproductive categories that may impact fertility	2*	Excluded: peri-menopausal, post-hormonal contraceptives for at least 3 months (but not 9 months as the quality chart requires for post-DMPA users); breastfeeding and postpartum included only after 3 cycles resumed
Description of study population	1	New users only – "no prior experience with the practice of BO, MM or LO for family planning."
FABM method(s) studied	1	Three separate methods compared: Billings, Dorairaj modified mucus method, local version of mucus method
Teaching of method	2*	Little detail provided on teaching
Detection of pregnancy	2	At 3-month intervals, a systematic interview was conducted to include menstrual status and pregnancy status. Most women (except not in the LO group) kept charts that were reviewed.
Classification of pregnancy as intended/unintended	2	Pregnancy intentions assessed every 3 months (not every cycle). Learning phase and teaching phase pregnancies reported but analyzed separately
Concurrent use of coitus-dependent methods	2*	Barrier and/or withdrawal use was not assessed
Study duration	1	15 months: 3 month learning phase pregnancy probability and 12 month experienced user probabilities reported separately
Statistical methods	2	Single decrement life table analyses for each method, but cycles with no intercourse included
Attrition	2*	Loss to follow-up was 4.8% for Billings, 6.6% for modified mucus method, and 8.9% for the local variant. Discontinuation rates not examined by population differences
Other	1	No other concerns
Overall rank	2	
Results included		3-month learning phase pregnancy probabilities and 12-month experienced user probabilities for unplanned pregnancy

STUDY: WHO 1981 (Trussell 1991 reanalysis)

Quality criteria rankings

Domain	Proposed Ranking	Rationale
Inclusion/exclusion criteria	2	Clear description of inclusion criteria; but no denominator for screened population is provided
Exclusion of populations not at meaningful risk of pregnancy	1	Original analysis did not exclude cycles without sexual activity, but re-analysis by Trussell did provide this information. Study included only women who had at least one live birth in preceding 5 years in present union, and excluded women who used hormonal contraceptive within the last three cycles prior to admission, and women had to agree not to use any other method of fertility regulation during the effectiveness phase of the study
Treatment of women in various reproductive categories that may impact fertility	1	Excluded: perimenopausal (all were <39), oligomenorreic (all had history of menstrual cycle intervals of 23-35 days), none were lactating, none were post-hormonal contraception (for 3 cycles prior to admission; and prior injectable users were only 0.7% of the overall sample)
Description of study population	1	Inclusion criteria specified that women must not have practiced self-recognition of mucus changes for family planning, thus, all new users, population was well-characterized with measurement of sociodemographic characteristics including age and parity
FABM method(s) studied	1	Billings Ovulation method; stated clearly, including rules for identifying fertile days
Teaching of method	1	In the majority of instances, the teachers were married women who were themselves using the method and had successfully completed the questionnaire; experienced teachers recruited new teachers for the study. All teachers completed an OM questionnaire designed by Billings that tested her knowledge and understanding of the method. Materials provided are described (charts on which data from 3 cycles could be recorded, colored stickers). 3-6 encounters occurred over 3-6 cycles
Detection of pregnancy	2	Not stated that pregnancies were routinely assessed, but seems likely that there was corroboration with charts ("from the chart and a monthly meeting during which the teacher carefully questioned the subject, the teacher transcribed selected cycle details onto a paper form and judged compliance, comprehension, and attitude). Active follow up to determine early pregnancies at the end of the study was not reported
Classification of pregnancy as intended/unintended	1	All pregnancies were reported (no intended pregnancies; 40 women withdrew to become pregnant). In the Trussell reanalysis, all cycles with prospectively identified intention to avoid pregnancy were included in typical use analysis for pregnancy avoidance
Concurrent use of coitus-dependent methods	1	Neither barrier nor withdrawal use was assessed in analysis; but women had to agree not to use any other method of fertility regulation during the effectiveness phase of the study. Author notes (paper 5) that some cycles included use of condoms or withdrawal (0.4% of cycles)

Study duration	1	There were at least 13 cycles, and a 13 cycle typical use pregnancy rate was provided
Statistical methods	1	Single-decrement life table analysis and cycles with no intercourse are excluded
Attrition	1	Among 869 subjects who entered the teaching phase, 45 (~5%) became pregnant in the teaching phase and 99 (11%) withdrew. Among 725 women who entered the effectiveness phase, after 13 cycles (Table 1, paper 2), no women appeared to be lost to follow-up as all had labeled reasons for discontinuation. At most, 99/869 were lost to follow-up (~11%). Differences for other reasons for discontinuation are described in paper 2. Additional detail: They stated in the teaching phase paper "the characteristics of the 725 subjects who entered this (effectiveness) phase of the study did not differ significantly from those of the 869 originally admitted", and noted "there were higher pregnancy and discontinuation rates among subjects for whom teaching had to be extended for one or more cycles beyond the usual time." In the effectiveness phase paper they did examine differences in discontinuation by country and by pregnancy intention, and described consideration of how marital status and agricultural work impacted discontinuation rates for the reason of separation of spouses and departure from the study center in Manila
Other	1	No other major concerns
Overall rank	2	
Results that should be in table		Perfect use including the initial teaching cycle; typical use 22.8% including learning phase cycles and specifically only those initial cycles in which intercourse was reported

STUDY: Wade 1981
Quality Criteria Rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion criteria clearly defined (20-39 years, regular 24-36 day menstrual cycles, not pregnant, couples in a stable relationship, couples who desired to avoid pregnancy for 2 years, couples who agreed to accept the method randomized to and to participate for at least 12 months). No refusal rate provided
Exclusion of populations not at meaningful risk of pregnancy	2	Cycles with no intercourse not excluded, couples with subfertility not explicitly excluded, and couples who were using another method of contraception not explicitly included or excluded
Treatment of women in various reproductive categories that may impact fertility	2*	Perimenopausal-age, oligomenorrheic women excluded. Unclear if recent users of hormonal contraception and or postpartum and lactating women were included
Description of study population	1	All new users. Demographic characteristics well-described including age and parity
FABM method(s) studied	1	Ovulation method and Thyma double-check symptothermal method with rules described and/or citations provided
Teaching of method	2*	Teachers were trained with ad hoc (not standardized) curriculum; periodic supervision from teacher-supervisor; unclear number of sessions and unclear details of curricula/materials provided
Detection of pregnancy	1	Monthly serum or urine pregnancy tests with review of charting, from enrollment on with active follow-up at study conclusion
Classification of pregnancy as intended/unintended	1	Baseline wanted to avoid pregnancy and agreed to participate in study for at least 12 months. Monthly follow-up included willingness to continue in the study (implying continuing want to avoid). Women who wanted to conceive were exited from the study. All cycles an pregnancies with intention to avoid pregnancy were included
Concurrent use of coitus-dependent methods	2*	No assessment of any use of barriers or withdrawal, although method is described as abstinence-based
Study duration	1	Planned to follow participants for at least 1 year
Statistical methods	2	Multiple decrement lifetable analysis and pearl rate analysis at one year
Attrition	2*	Loss to follow-up less than 20%; Reasons for dropout reported but not examined according to population differences
Other	2	Women who could not learn the method were exited from the study which could impact study generalizability. Some gaps and inconsistencies in description; for example, report states that protocol did not allow investigators to withdraw participants during the study phase, but there are involuntary withdrawals listed in the table. Also, unclear what it means that many couples had “failure to follow or apply rules correctly.”
Overall rank	2	
Results included		Multiple decrement pregnancy rates at 1 year starting from the beginning of the training period. Pearl rates also available

STUDY: Weissmann 1972 Quality Criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	3	Inadequate description of inclusion/exclusion criteria (registered in NFP center 9/1/81-8/31/82, kept chart for 3 months, desired to avoid pregnancy) and no refusal rate provided
Exclusion of populations not at meaningful risk of pregnancy	2	Exclusion of cycles with no intercourse, subfertile women and women using other contraceptive methods not reported, but no explicit inclusion of women using hormonal contraception or sterilization
Treatment of women in various reproductive categories that may impact fertility	2*	Perimenopausal women include with no subgroup analysis. Recently post-hormonal contraception users inclusion was not reported. Breastfeeding and recently postpartum women (<6 weeks after deliver) were included without subgroup analysis
Description of study population	3	Not reported whether new or experienced users. Age and parity not described and demographic description incomplete
FABM method(s) studied	1	Ovulation Method; minimal reference to rules used
Teaching of method	2*	Minimal detail reported, except that teachers went and stayed in the communities until people were confident of the method; no info on training of teachers and uncertainty about curricula materials
Detection of pregnancy	3	Not reported how this was done. Appears to be self-report; with no evidence of systematic corroboration by charting
Classification of pregnancy as intended/unintended	3	Only reported determining pregnancy intentions at baseline; apparent exclusion of most pregnancies from the calculation but given no data about how unintended pregnancies were defined, this is unclear. On page 815, 28 couples anxious to have more children abandoned method, but it is unclear if the change in intentions was reported prospectively
Concurrent use of coitus-dependent methods	2*	Barrier and withdrawal use not assessed, although couples had to promise they were not using withdrawal, which was the most widespread method in Tonga
Study duration	3	Length of time and classification of pregnancies confusing. Maximum number of months could have been 20, but the measure does not clearly apply to a first year of use, and we don't know which women used the method for which length of time
Statistical methods	3	Number of pregnancies reported, but rates not calculated, and there is concern about exclusion of pregnancies from failure rates so Pearl rate cannot be confidently calculated

Attrition	2*	Loss to follow-up <20%. Reasons for dropout reported but no rates and no systematic investigation by population differences
Other	1	None
Overall rank	3	
Results included		None

STUDY: Xu 1994
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	3	Poorly characterized study inclusion/exclusion criteria and no information on refusal
Exclusion of populations not at meaningful risk of pregnancy	2	Not reported whether cycles with no intercourse included or whether couples with sub-fertility or who were using other methods of contraception were excluded. Sterilized women were excluded. No women currently using other contraceptive methods were explicitly included
Treatment of women in various reproductive categories that may impact fertility	2*	Perimenopausal-age and oligomenorrheic women, and breastfeeding women were included. Unclear how recent past hormonal contraception use was. Unclear if recently postpartum women were included. No subgroup analyses were reported
Description of study population	2*	Unclear if new users. 24.2% reported prior use of "safe period contraception", possibly a rhythm/calendar method
FABM method(s) studied	1	Billings ovulation method with clear description of rules. Cited Family of the Americas materials
Teaching of method	2*	Description of teacher training not reported, number of educational sessions not reported but time of learning reported to be about 3 months, materials from Family of the Americas cited
Detection of pregnancy	3	No description on how pregnancies were detected
Classification of pregnancy as intended/unintended	2	No evidence of prospective monthly assessment of pregnancy intentions. No evidence that any pregnancies were excluded from the analysis
Concurrent use of coitus-dependent methods	2*	Use of barrier methods and or withdrawal was not assessed
Study duration	2	Up to 24 months but life table analysis provided for 12 and 24 months
Statistical methods	2	Likely multiple decrement life table analysis. Perfect use pregnancy probability not correctly calculated
Attrition	2*	Loss to follow-up <20% due to overall attrition being <20% at 12 months. Discontinuation rates by reason not reported and no discussion of differences by population
Other	1	None
Overall rank	3	
Results included		None

Mucus-only Method: Creighton Model Fertility Care System (CrMS)

STUDY: Doud 1985
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	1	Clients entering NFP department between Oct 1980 and Dec 1982, excluded if had known infertility, pregnant, not genitally active, using barrier methods. 376/584 included in study
Exclusion of populations not at meaningful risk of pregnancy	2	Unclear if cycles without sexual activity were excluded from the study (although non- "genitally active" people were excluded from the study initially); Couples with a history of infertility were excluded but no definition of what this meant. No indication that women using other methods were excluded, except that women using barriers were excluded
Treatment of women in various reproductive categories that may impact fertility	2*	Study included women in various reproductive categories, butdid not analyze them separately
Description of study population	1	Proportion of new versus experienced users provided, but experienced users are not analyzed separately. Population is characterized by age but not parity
FABM method(s) studied	1	Creighton model is assessed with reference to rules used to identify fertile days
Teaching of method	1	Teachers were certified or in process of being certified as NFP practitioners or instructors. Reference to materials used. One-on-one counseling in introductory session, with follow up sessions on a 2, 4, 6, 8, 12, 24, 36, and 52-week schedule
Detection of pregnancy	3	Self-report (sometimes by phone), with no evidence of systematic corroboration with charting or testing among all participants included in analysis
Classification of pregnancy as intended/unintended	3	Assumed that pregnancies stemming from intercourse on a day of fertility were achieving-related. These estimates are not comparable with other methods of contraception and this method should be separately discussed
Concurrent use of coitus-dependent methods	2*	Women using barrier methods were excluded; no apparent attempt to assess systematically for withdrawal
Study duration	2	Study involved 12 months of follow up, but no true typical use pregnancy report is provided due to definition of pregnancy intentionality
Statistical methods	2	Multiple decrement analysis. Cycles with no intercourse not excluded

Attrition	2*	Discontinuation rates provided, LTFU <20%, but no examination of attrition by population differences.
Other	2	No other major concerns; but other minor concerns such as only 65% having a complete pregnancy evaluation
Overall rank	3	
Results		None

STUDY: Fehring 2009
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion clearly defined (18-44 years old with no known infertility, including women who had recently discontinued hormonal contraception. Refusal rate not provided.
Exclusion of populations not at meaningful risk of pregnancy	2	
Treatment of women in various reproductive categories that may impact fertility	2*	Included hormonal contraceptive users recently with no subgroup analysis. Included women with irregular cycles and analyzed separately. Unclear if breastfeeding and/or postpartum women were included.
Description of study population	3	Unclear proportion of experienced users. Age and parity described.
FABM method(s) studied	1	Marquette monitor only method and Marquette Mucus only methods with citations and clearly defined rules.
Teaching of method	2*	Clear information about instruction and interface; no information about teacher training.
Detection of pregnancy	2	Unclear but did imply systematic review of the fertility charts which were submitted to the principle investigator. Unclear whether active follow-up was employed
Classification of pregnancy as intended/unintended	2	Retrospective study design for some of the data but analyzed prospectively collected database data.
Concurrent use of coitus-dependent methods	2*	Barrier and or withdrawal use not assessed
Study duration	2	12 months
Statistical methods	2	Unclear how cycles in the training phase were handled. Does not appear that correct use was not restricted to correct use cycles. Kaplan Meier 12-month probabilities comparable to single decrement life tables are provided.
Attrition	2*	Lost to follow-up reported in 2007 and 1993 studies, however the total sample size does not fit with the samples in this earlier study and there is a lack of details in who was included in this pooled analysis.
Other	2	
Overall rank	3	
Results included		None

STUDY: Hilgers 1998
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Exclusion criteria defined relatively well though not individual differences for each study; refusal rate was defined for those who entered the study but not for those who were initially invited
Exclusion of populations not at meaningful risk of pregnancy	2	Excluded infertile couples; no mention made of exclusion related to other methods thought unlikely given that even couples using barrier methods were excluded; excluded couples who are not sexually active but unclear if cycles with no sexual activity were excluded from analysis
Treatment of women in various reproductive categories that may impact fertility	2*	Included breastfeeding women and perimenopausal-age women as well as post-Pill (within the last yar) and postpartum with no subgroup analyses performed.
Description of study population	2	Appear to be all new users (after attending introductory session); not explicit. Also, do not have parity on all studies and no demographic information on 2 of the cohorts
FABM method(s) studied	1	Creighton Model; clearly referenced and described.
Teaching of method	1	Clearly referenced and described
Detection of pregnancy	3	Not reported how pregnancies were detected. There was some active follow-up but unclear if it was systematic
Classification of pregnancy as intended/unintended	3	Achieving related pregnancies defined as pregnancies where couples knowingly had intercourse on a fertile day; these were removed from use effectiveness estimte
Concurrent use of coitus-dependent methods	2*	No assessment of any use of barriers or withdrawal though users of barrier methods were excluded from the analysis initially
Study duration	1	18 months with 12 month life table rate
Statistical methods	1	Single Decrement Life Table
Attrition	1	Loss to follow-up less than 20% and discontinuation rates and reasons reported
Other	1	None
Overall rank	3	
Results that should be in table		None

STUDY: Howard 1999
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Study provides a retrospective analysis of records of couples who began use of CrMS. All new users of CrMS were eligible to be included in the analysis. The analysis excluded couples with a history of infertility, pregnant women, and women who were not "genitally active". No attempt made to select couples who expressed strong motivation to avoid pregnancy. No denominator for a screened population is available
Exclusion of populations not at meaningful risk of pregnancy	2	No exclusion cycles without sexual activity; investigators did not have this information as this was a review of records and not a prospective study. Investigators excluded couples with a history of infertility (but authors note that it may have included couples with subfertility) and women using hormonal contraceptives
Treatment of women in various reproductive categories that may impact fertility	1	Oligomenorrheic women were included (defined as more than 38 days), older women were included (over 40), discontinued OCs in past year, breastfeeding included, post-partum included, post-abortion included; with separate estimates provided
Description of study population	2	New users of CrMS were considered eligible; population is characterized by age but not parity
FABM method(s) studied	1	Method (Creighton) and rules are clearly stated and referenced
Teaching of method	1	Materials are referenced with detailed description of CrMS instruction. Group introductory section, immediate charting, follow up visit, 8 follow up visits over a year lasting 45-60 minutes
Detection of pregnancy	2	No routine periodic pregnancy testing, but investigators reviewed any evidence of pregnancy, as suggested by prolonged postovulatory phase of a menstrual cycle. Follow up beyond 12 months was accomplished mainly by telephone contact
Classification of pregnancy as intended/unintended	3	Assumed pregnancies stemming from intercourse on a day of fertility were achieving-related. These estimates are not comparable with other methods of contraception
Concurrent use of coitus-dependent methods	1	Couples using barrier, withdrawal, or other methods were excluded (this was uncommon)
Study duration	2	18 months of follow up (or pregnancy or other reason to leave the study), including separate estimates at 12 months, but no actual 12 month "typical use" pregnancy rate is reported due to the definition of "achieving-related"

Statistical methods	2	Single-decrement life table analysis is performed, but unclear if cycles with no intercourse were excluded
Attrition	2	Probability of LFTU at 12 months was 12.41, but this is based on multiple decrement analysis. No examination of attrition by population differences
Other	1	
Overall rank	3	
Results included		None

Mucus-only Method: Two Day Method (TDM)

STUDY: Arevalo 2004
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Criteria clear (18-39, living in union, had previous pregnancy, >3 cycles postpartum, >6 months post-hormonal injection, >3 months post-oral contraception) but no refusal rate provided
Exclusion of populations not at meaningful risk of pregnancy	2	Excluded cycles of women using sterilization, IUDs or hormonal contraception, and those without sexual activity and those in which woman used barrier method or withdrawal on days that were not identified as fertile by the method; no reference to couples with reported history of 12+ cycles of unprotected sex and no conception, but included only women who had had a previous pregnancy
Treatment of women in various reproductive categories that may impact fertility	2	Excluded all categories referenced in quality ranking except only 6 months post-DMPA (not 9 months as had been specified); no subgroup analysis.
Description of study population	1	New users only as the method was the introduction of a new method; demographic profile of women clearly described including age and parity.
FABM method(s) studied	1	Two Day method clearly defined and rules described.
Teaching of method	2*	Detail on instruction of women, but training of providers not described in detail (5-10 trained per site), curricula cited
Detection of pregnancy	1	Pregnancies determined by targeted hormonal testing for all women whose cycle lasted 42 days or longer; active follow-up of women with amenorrhea at conclusion of study.
Classification of pregnancy as intended/unintended	1	Women interviewed every cycle "to assess their use of the method and their pregnancy status," i.e., text does not explicitly refer to intentions and 2.2% of the participants exited the study because they wanted to get pregnant so we infer that women were asked prior to each cycle to assess intentions.
Concurrent use of coitus-dependent methods	1	Assessed barrier used during fertile period, and separate analyses conducted for those using abstinence and those using barrier methods on fertile days
Study duration	1	13 cycles, typical use rate reported
Statistical methods	1	Single decrement life table rates for correct and typical use of method; cycles with no intercourse excluded
Attrition	2*	4.4% lost to follow up, but no analysis of attrition by population differences

Other	2	Removing women from the study for study related reasons may impact study generalizeability.
Overall rank	2	
Results included		Correct use pregnancy rate; correct and incorrect pregnancy rate.

STUDY: Jennings 2011
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	1	Inclusion criteria clearly defined; refusal rate is provided
Exclusion of populations not at meaningful risk of pregnancy	2	Not reported whether exclusions made according to listed characteristics (including cycles with no intercourse); no explicit inclusion of hormonal method users
Treatment of women in various reproductive categories that may impact fertility	2*	Oligomenorrheic women were not included, but perimenopausal-age women definitely included. Breastfeeding and postpartum and recently post-hormonal contraceptive use not explicitly excluded; no subgroup analysis
Description of study population	1	All new users. Age and parity recorded
FABM method(s) studied	1	Two Day Method with rules described and cited
Teaching of method	2*	Training reported but no details; curriculum reported; number of encounters reported
Detection of pregnancy	2	Presumably self-report only and only recollected; with no evidence of systematic corroboration with charting or testing all participants included in analysis.
Classification of pregnancy as intended/unintended	2	Pregnancy intentions not collected every cycle; no pregnancies excluded
Concurrent use of coitus-dependent methods	2*	NR whether barrier method use or withdrawal use happened. No evidence of it being measured
Study duration	2	7-month life table pregnancy rate
Statistical methods	2	Used single-decrement analysis, but unclear if cycles with no sexual activity were excluded.
Attrition	2*	Loss to follow-up <20% (13+17/161) but over 7 cycles; discontinuation reasons reported but not explored by demographic differences.
Other	1	Recollected intercourse during fertile time which may make perfect use probabilities less accurate but perfect use was not reported for this study.

Overall rank	3	This is based on only one three; a lack of clarity on detection of pregnancy. It may be worth following up with the authors to ask more about detection of pregnancy, as this is otherwise a pretty good study, with 6 "1's" in the quality scores.
Results included		Typical use 7-month pregnancy probabilities for each group

Mucus-Only Methods: Modified Mucus Method (MM) and variant

STUDY: Dorairaj 1984
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Clearly described criteria for inclusion (Heterosexual couples recruited with prior proven fertility, not using contraception); Clear description of target population; no description of refusal or participation rate.
Exclusion of populations not at meaningful risk of pregnancy	2	Excluded users of contraception but not reported about sterilization. Cycles with no intercourse included. No subgroup analyses.
Treatment of women in various reproductive categories that may impact fertility	2*	Some women were still amenorrheic from breastfeeding, but some analyses separated these; perimenopausal-age women were included but no subgroup analyses. Postpartum women were included with no subgroup analyses. Unclear whether recent hormonal contraceptive users were included.
Description of study population	1	All new users; included characterization of age, education, income, gravidity, sex preference
FABM method(s) studied	2*	Modified mucus method, abstain during "fertile type" mucus and 2 days later. No description of observational instructions or definition of fertile type mucus or citation
Teaching of method	2*	Teachers were trained with ad hoc 2-month training that included other health topics besides family planning, educational interface with patients and curricula were not well-described.
Detection of pregnancy	2	Systematic interviews at end of cycles 1,2,3,6,9,12. The MMM doesn't use charts.
Classification of pregnancy as intended/unintended	2	Frequent follow-up. All pregnancies reported. A small number of "planned" pregnancies are reported, but it's not stated whether they were designated planned retrospectively or prospectively. Can recalculate the study pregnancy typical use rates with the "planned" pregnancies included.
Concurrent use of coitus-dependent methods	2*	No assessment of any use of barriers or withdrawal, however method is described as abstinence-based
Study duration	1	Participants followed for 12 months/cycles
Statistical methods	2	Multiple decrement lifetable; but raw data are given from which single-decrement life table rate can be calculated. Dropout rate is very low and multiple decrement and single decrement rates should be nearly identical

Attrition	1	Total attrition <10% other than pregnancy; total dropout ~10% including pregnancy; so discussion of rates and reasons of dropout not necessary.
Other	3	Difference between Table 1 and Table 4 implies that all 942 women with breastfeeding amenorrhea at beginning of study were followed completely for 12 months with no drop out and that none of them became pregnant.
Overall rank	2	
Number of "1" s	3	
Results included		None

STUDY: Dorairaj 1991
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	3	Inadequately defined inclusion criteria
Exclusion of populations not at meaningful risk of pregnancy	3	Inadequate detail provided
Treatment of women in various reproductive categories that may impact fertility	2*	No information on how women in various fertility categories were treated
Description of study population	3	Unclear proportion of new versus experienced users
FABM method(s) studied	1	Modified mucus method studied; rules are described
Teaching of method	1	Individual instruction by lay educators (village women leaders supervised by part-time cluster coordinators) in 10 home visits over three menstrual cycles; no written materials used.
Detection of pregnancy	3	Presumably self-report only, with no evidence of systematic corroboration with charting or testing among all participants included in analysis)
Classification of pregnancy as intended/unintended	2	No evidence of inappropriate exclusion of unintended pregnancies
Concurrent use of coitus-dependent methods	2*	No assessment for barrier or withdrawal use
Study duration	3	Timing of follow up is completely unclear
Statistical methods	3	Insufficient detail is provided
Attrition	2*	Attrition not described
Other	2	Insufficient details provided from which study can be assessed
Overall rank	3	
Results included		None

STUDY: Kambic 1994
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion criteria unclear except statement that women had charted for one cycle
Exclusion of populations not at meaningful risk of pregnancy	2	Unclear whether subfertile women/couples, contraceptive using women/couples, or cycles with intercourse were included but no explicit inclusion of women using contraception.
Treatment of women in various reproductive categories that may impact fertility	2*	None of these categories were reported as included or excluded. Important to note that over 50% of women using either method were breastfeeding
Description of study population	2	Unclear proportion of new users, except that the modified mucus method was new in Liberia so likely all new users. Age and parity not described.
FABM method(s) studied	1	Modified Mucus Method (1) and Symptothermal/Ovulation Method (2) – MMM is referenced in detail, ST/OM not. But the analysis was conducted separately – combined ST and OM methods together as reference for MMM, so this rate should be disregarded.
Teaching of method	2*	Insufficient info on teacher training, number of encounters and curricula/materials used
Detection of pregnancy	3	Self-report only, no apparent corroboration with charting
Classification of pregnancy as intended/unintended	2	No apparent exclusion of retrospectively classified pregnancies; intentions assessed at baseline but not thereafter.
Concurrent use of coitus-dependent methods	2*	Barrier method and withdrawal use not assessed.
Study duration	1	12 months of use – multi-decrement life-table discontinuation rates provided; also show data on average length of use of method
Statistical methods	2	Multi-decrement life table; did not exclude cycles without intercourse
Attrition	1	No loss to follow-up reported; discontinuation rates and reasons discussed otherwise.
Other	1	None
Overall rank	3	
Results included		None

STUDY: Thapa 1990 (see above)

Mucus-Only Method: Marquette-Mucus Only

STUDY: Fehring 2013
Quality criteria rankings

Domain	Proposed Ranking	
Inclusion/exclusion criteria	1	A randomized trial with eligibility assessment flowsheet clearly described.
Exclusion of populations not at meaningful risk of pregnancy	2	Excluded men with subfertility problems, not stated for women; cycles without intercourse not excluded; women using contraceptives likely not included
Treatment of women in various reproductive categories that may impact fertility	2	Included perimenopausal-age women. No subgroup analysis. Excluded women with long cycles (longer than 42 days). Excluded users of hormonal contraception and breastfeeding women until they had at least 3 cycles. Unclear about recently postpartum women but likely excluded until 3 normal cycles.
Description of study population	2	Women recruited online, new instruction, received EHFM, so implies new users, but not specifically stated whether this was assessed; Good description of baseline characteristics in both groups
FABM method(s) studied	1	Marquette method, randomized comparison of 2 variations: Cervical Mucus Only and Marquette Monitor only. Rules clearly described and/or cited.
Teaching of method	1	Quick Start instructions online, with quiz to assess understanding; good report of teacher training and educational materials. Could be easily replicated but using Marquette online charting tool and educational system.
Detection of pregnancy	2	Online system actively notifies participant if luteal phase extends 19 days and requests pregnancy test, but no statement about active follow-up at conclusion of study
Classification of pregnancy as intended/unintended	1	System requests this information at beginning of each cycle; Also, they did a robust quantitative informative analysis of the influence of level of motivation measured at the beginning of each cycle on pregnancy rates (separate paper).
Concurrent use of coitus-dependent methods	2*	Barrie and withdrawal not assessed.
Study duration	1	All rates reported at 12 months or 13 cycles
Statistical methods	2	Kaplan-Meier, but cycles without intercourse not excluded
Attrition	2*	Dropout high, <50% continuing in each group at 12 months. Rates and reasons not described.
Other	1	None
Overall rank	2	
Results included		12-cycle unintended pregnancy probabilities by subtype and 12 “month” correct use pregnancy probability by subtype

STUDY: Fehring 2014
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	All clients from two other prospective studies and some additional prospective data who meet age criteria But no denominator given, no idea of who declines
Exclusion of populations not at meaningful risk of pregnancy	2	All comers to the Marquette Method in the specific study groupings.
Treatment of women in various reproductive categories that may impact fertility	1	Entire purpose of study is to look at special "subgroup of women ages 40-55, by design this should include women with longer or shorter or irregular cycles (though some of the extremes of cycle lengths may have been excluded by MM entrance criteria, not entirely clear).
Description of study population	2	New instruction, received EHFM, so implies new users, but not specifically stated whether this was assessed to exclude prior users from entering Good description of baseline characteristics in both groups
FABM method(s) studied	1	MM, 3 variations: EHFM only, CMM only, and EHFM+CMM, analyzed together and separately
Teaching of method	1	Quick Start instructions available online, plus consultation with MM trained teachers, this is replicable
Detection of pregnancy	2	Online system actively notifies participant if luteal phase extends 19 days and requests pregnancy test, but no statement about active follow-up at conclusion of study
Classification of pregnancy as intended/unintended	1	System requests this information at beginning of each cycle
Concurrent use of coitus-dependent methods	2*	Not assessed
Study duration	1	All rates reported at 12 cycles (ideally would have been 13 cycles)
Statistical methods	2	Kaplan-Meier, but cycles without intercourse not excluded
Attrition	2*	dropout high, mean 4.6 cycles per woman (741 cycles over 160 participants)
Other	1	Well-described, all significant concerns captured above
Overall rank	2	
Number of "1" s	6	
Results that should be in table		12-cycle unintended pregnancy rates overall and by subtype of MM, 12 "month" correct use pregnancy rate

STUDY: Fehring 2017
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion criteria clearly described: All participants at online website who charted for at least one cycle and who were not breastfeeding, age 18 or older
Exclusion of populations not at meaningful risk of pregnancy	2	Included cycles without sexual intercourse. Not reported whether subfertile or contraceptive using women were excluded.
Treatment of women in various reproductive categories that may impact fertility	2*	Only 3.4% were "post hormonal contraception" and 3.5% were "pre-menopause" (neither term was defined). Unclear about recently postpartum. No subgroup analyses. Breastfeeding excluded.
Description of study population	2	All new users to website. Adequate characterization of age, prior pregnancies, prior cycle lengths
FABM method(s) studied	1	Three variations of Marquette Model: Mucus Only, Monitor only, and both Mucus and Monitor. Peer-reviewed publications in references clearly describe the method
Teaching of method	2*	Online support from nurses trained in Marquette Model clearly described and materials referenced. Curriculum for training not described.
Detection of pregnancy	2	Prolonged luteal phase in web-based chart, with prompt to user for pregnancy test. Active follow-up not reported.
Classification of pregnancy as intended/unintended	1	Users required to enter whether intention to avoid or achieve in each cycle prospectively. No pregnancies excluded from the analysis.
Concurrent use of coitus-dependent methods	2*	No assessment of any use of barriers or withdrawal use. Method is described as abstinence-based
Study duration	1	Participants followed for 24 months/cycles; 12-month rates also given
Statistical methods	2	Kaplan-Meier survival comparable to single decrement life table. Cycles without intercourse included. Perfect use calculation is done correctly but only provides overall and not for each arm. Also, unintended pregnancies were counted as during typical use when there was missing information about interpreting the fertile window. Therefore the missing data issue would result in a potential bias in the perfect use estimate. A

Attrition	2*	Drop out high and not well-characterized. Mean participation, 7.9, 4.1, 9.2 cycles for EFM, CMM, and EFM + CMM respectively
Other	1	
Overall rank	2	
Results included		Unplanned pregnancy probabilities (Table 1). Correct use probability in text could be reported with limitations noted above.

Basal Body Temperature (BBT) Based Methods

STUDY: Bartzen 1967
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	3	Criteria unclear other than that most women were Bartzen patients ages 19-46, no refusal rate discussed.
Exclusion of populations not at meaningful risk of pregnancy	2	Excluded cycles of women whose husbands were away with the military. Did not identify women not having sexual activity. Unclear whether sub-fertile women included but t included only women who had had a previous pregnancy.
Treatment of women in various reproductive categories that may impact fertility	2*	No detail on lactation, postpartum. Seems to have included women with longer cycles (average longer cycle was 38.7). Included perimenopausal women. NR on contraceptive use recency. No subgroup analyses
Description of study population	3	Not clear whether new or some experienced users. Age and parity described.
FABM method(s) studied	2*	BBT defined in terms of Ogino but vague combination. Beginning of fertility shortest cycle minus 19. End of fertility was the third higher temperature reading above the baseline which was determined after 2 cycles by the researcher. Not sufficient detail to understand how the baseline temperature was calculated this is not reproducible.
Teaching of method	2*	Some detail provided but study clearly documents gaps in instructor's mastery of the detail and we don't know much about how he was trained in BBT. Curriculum was not referenced. This could not be repeated. Frequency of encounters reported and some content of education provided.
Detection of pregnancy	2	Pregnancies self-identified and confirmed with systematic review of charts.
Classification of pregnancy as intended/unintended	3	No information on pregnancy intentions ongoing. Author notes "planned pregnancies were noted as such" but unclear if this was a retrospective designation. Also, no clear definition of pregnancy intentions even at baseline.
Concurrent use of coitus-dependent methods	2*	No systematic assessment of barrier and withdrawal use for all cycles (reported withdrawal retrospectively for studies in which there was a pregnancy only)
Study duration	3	5 years, or 4.75 years with only overall pearl rates that could be calculated.
Statistical methods	3	Pearl rate of gross and method failure; cycles with no intercourse not excluded. Perfect use calculation is incorrect.

Attrition	2*	8.8% lost to follow up; no analysis of attrition by population differences
Other	1	
Overall rank	3	
Results included		None

STUDY: Berglund Scherwitzl 2017
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion criteria clearly defined (pragmatic design but all enrolled women 18 and older who were using the app to avoid pregnancy and who had no medical contraindication to pregnancy. No denominator of invited women. Though this study was conducted electronically and all women who registered for the application during the enrollment period potentially became a participant in the study (i.e., no refusals), those who did not have access to the app for at least 3 months, did not enter data for 18 days, were younger than 18 years or were not using the app to avoid pregnancy were excluded and there are no numbers clarifying this.
Exclusion of populations not at meaningful risk of pregnancy	2	Unclear whether sterilized women were excluded; cycles with no intercourse likely not excluded and this was likely a high proportion of cycles or was underreported; women were not explicitly asked whether they were using hormonal contraception but on the other hand, only 3% reported using an "other method" of contraception on green days (not counting barrier method use)
Treatment of women in various reproductive categories that may impact fertility	2*	Included women with long cycles, women ≥ 40 (2%), and recent hormonal contraception use (unclear how recent). Unclear whether breastfeeding/postpartum women were included. No subgroup analyses
Description of study population	1	Data collected from the time that women first registered for the app so likely new users of this method (and, only 1% of women previously used any other FABM fertility).
FABM method(s) studied	1	Natural Cycles application (based primarily on temperature plus cycle tracking); proprietary algorithm but clear rules: green days for intercourse.
Teaching of method	1	Teaching is not applicable to this electronic application, which is commercially available. The study could be easily repeated by asking women to access the app. However, there is no information provided about how women used the method and how it actually calculated the fertile window.
Detection of pregnancy	2	Pregnancies were primarily detected (123/143) by prospectively collected pregnancy test data. While such testing was not systematically conducted among all participants (raising concern that pregnancies may have been underreported), investigators also used data from the application to identify women with consistently high temperatures or a delayed menstruation (15/143 pregnancies). However, the definitions used to detect these were not provided. The absence of pregnancy is also confirmed if/when a user logs menstruation at the end of study participation. If it was not possible to determine whether a pregnancy had occurred through either of these mechanisms, information from a retrospective survey asking participants whether they had become pregnant using the method was used. The response rate for this survey was 30% overall, but Figure 1 suggests that responses were 100% among women for whom this information was utilized to determine whether a pregnancy had occurred.

		Finally, the investigators provide estimates with the "worst-case scenario" in which it is assumed that all users with unknown pregnancy status became pregnant.
Classification of pregnancy as intended/unintended	2	Pregnancy intention measured only at baseline. Some women noted that they were planning to conceive at the time of the survey interview at the end of the study. It is unclear how any pregnancies were classified in this group (1.8% of post survey group). Learning phase pregnancies may have been excluded from the analysis as women who had access to the app for <3 months were not excluded and women who did not log an ovulation during the first cycle were excluded.
Concurrent use of coitus-dependent methods	2*	Reported only overall use and underreporting likely for per day reporting; no subgroup analyses
Study duration	1	12m study; 12m pregnancy single decrement life table and pearl rates reported
Statistical methods	2	Perfect use pregnancy rates were incorrectly calculated; so will not be abstracted. For typical use pregnancy rates, both a Pearl Index and a life table estimate using a Kaplan-Meier estimator was used. However, only for the Pearl estimate did investigators make the conservative assumption that 61 women for whom pregnancy status could not be determined eventually became pregnant. Since pregnancy testing was not routinely conducted in this study (and since we are nonetheless ranking this study a"2" instead of a "3" on pregnancy detection), we feel it appropriate to use this "worst-case" scenario estimation of the 9.8 Pearl index.
Attrition	2*	Loss to follow-up: 34% discontinued using the application prior to the end of the study; though pregnancy status was still able to be determined for some of them.
Other	2	Under-reporting of intercourse obvious or women were having significantly less intercourse than average. Given that cycles with no intercourse should usually be removed from the analysis, in this study, many cycles should have been removed per the reported intercourse frequency. This means that the pregnancy rate could be greatly underestimated. Also, there are 60 women censored before the first detection of ovulation; unclear if they could have been pregnant (per note on Table 5). Results of this study may be more generalizable than those of other clinical studies because the population was not closely followed.
Overall rank	2	Total pregnancy rate including women lost to f/u and unknown pregnancy status. Perfect use rates of insufficient quality to include

STUDY: Döring 1967 (in German)
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Patients of a private clinic of BBT-only method of family planning; clearly defined inclusion criteria; refusal rate not applicable as all patients included.
Exclusion of populations not at meaningful risk of pregnancy	2	Report implies (does not explicitly state) that only women with proven fertility included; Report implies (does not explicitly state) that cycles without sexual intercourse were excluded; Unclear whether current users of contraception were included.
Treatment of women in various reproductive categories that may impact fertility	2*	Excluded age 40 and older, oligomenorrhea or amenorrhea; unclear about whether postpartum or post contraception women were excluded. No subgroup analyses.
Description of study population	3	Unclear whether new or established users.
FABM method(s) studied	1	Two types of BBT only method described: 1) postovulatory only, starting on third day of temperature rise; 2) postovulatory and pre-ovulatory, which includes as infertile or available the early days of the cycle up until the earliest day of temperature shift in any prior cycle minus 6 days. Temperature rise defined by "3 over 6", reaching at least 0.2 C higher.
Teaching of method	2*	Clinic-based, presumably taught by author himself
Detection of pregnancy	2*	All cycles have temperature charting that was reviewed
Classification of pregnancy as intended/unintended	2*	Emphasizes that all pregnancies are included as unplanned, even if unclear intentions. Provides a breakdown of putative causes or reasons for each pregnancy. However, gives no information on how or when intentions were updated, or how or when participants might exit for the purpose of planning pregnancy
Concurrent use of coitus-dependent methods	2*	No assessment of any use of barriers or withdrawal. (Method is described as abstinence-based)
Study duration	3	Average follow-up is 4 years 10 months, longest 19 years. No one-year rates reported.
Statistical methods	2	1200 Pearl Rate; cycles with no intercourse likely excluded.
Attrition	2*	Incomplete data on dropout. States specifically that he was unable to follow everyone who had received instruction initially in the method, but does not give any numbers for dropouts or LFU.
Other	1	None
Overall rank	3	
Results included		None

STUDY: Drouin 1994
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Adequately described. No refusal rate provided
Exclusion of populations not at meaningful risk of pregnancy	2	Perimenopausal and post-hormonal contraceptive users excluded. No details provided to indicate evaluation of fertility status. No exclusion of cycles without intercourse in presentation of typical use rate
Treatment of women in various reproductive categories that may impact fertility	2*	Insufficient details provided on whether women could have been subfertile or breastfeeding and there was no sub analyses of women that may have represented different reproductive categories
Description of study population	2	Includes 6% of former users of Bioself and none of the pregnancies occurred to these users (3/6 occurred to former users of another FABM, age and parity described
FABM method(s) studied	1	Adequately described
Teaching of method	2*	Trained healthcare professionals provided method instruction. Curriculum, duration, frequency of encounters not reported
Detection of pregnancy	1	Detection of pregnancies not described; correspondence with author suggested that there was systematic assessment of pregnancy (systematic detection of pregnancy testing every 30-60 days)
Classification of pregnancy as intended/unintended	2	Reported that all pregnancies that occurred were a result of intercourse during the fertile window. Thus, can infer that the method related failure rate is 0.
Concurrent use of coitus-dependent methods	2*	Concurrent use assessed and reported that all pregnancies occurred with unprotected intercourse during the fertile window. No assessment was made of the pregnancy rate among cycles with concurrent use of a method.
Study duration	1	12 months
Statistical methods	2	Study duration of 12 months with life table (figure) and pearl estimate reported
Attrition	2*	27/83 women discontinued use of the method. 5 of the 27 discontinued due to change in family planning intentions.
Other	2	None
Overall rank	2	
Results included		Typical and perfect use pearl rate estimates (life table figure does not provide precise estimates obtained)

STUDY: Flynn 1991
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Adequately described. No refusal rate reported
Exclusion of populations not at meaningful risk of pregnancy	2	Authors state only fertile women were included, but it is unclear how fertility status was defined or determined. While hormonal contraceptive users were excluded, it was unclear as to how recent discontinuation of use had occurred among the women who were former users of an oral contraceptive
Treatment of women in various reproductive categories that may impact fertility	2*	No presentation of any sub-analyses among women of different reproductive categories
Description of study population	1	Well-described, including age, parity and experience with FABMs
FABM method(s) studied	1	Adequate characterization of the method
Teaching of method	2*	Method instruction provided in package insert for the Bioself device
Detection of pregnancy	3	No details provided on detection of pregnancy
Classification of pregnancy as intended/unintended	2	One pregnancy excluded as it could not be characterized; examination of "limiters" versus "spacers" conducted
Concurrent use of coitus-dependent methods	1	Assessment of use of abstinence only approach versus barrier method/withdrawal method approach
Study duration	3	24 months
Statistical methods	2	No overall typical use rate and rates were obtained using unconventional method, and while Pearl rates could be calculated by the reviewer, the study duration was 24 months and this rate is not comparable to other Pearl rates of 12 months of follow-up
Attrition	2*	Attrition reported and <20%
Other	1	None
Overall rank	3	
Results included		None

STUDY: Guerrero 1970
Quality criteria rankings (see above)

STUDY: Marshall 1968
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion criteria clear, but denominator of how many assessed or refused not stated.
Exclusion of populations not at meaningful risk of pregnancy	2	Only included couples with a prior live birth together; Unclear whether current contraceptive users were included. Did not exclude cycles without sexual activity
Treatment of women in various reproductive categories that may impact fertility	2*	Unclear whether breastfeeding, postpartum , recent contraceptive users, or oligomenorrheic women were included. No subgroup analysis
Description of study population	3	Unknown proportion of new users; Good characterization of age, duration of marriage, number of prior children, and SES
FABM method(s) studied	1	Two variations of BBT, BBT postovulatory only, BBT with pre-ovulatory calendar calculation. References provided
Teaching of method	2*	Established organization taught method, but no details about training or materials
Detection of pregnancy	2	All temperature records reviewed for signs of prolonged luteal phase. No evidence of active follow-up
Classification of pregnancy as intended/unintended	2	Reported all pregnancies; excluded pregnancies resulting from stopping BBT charting from typical use analysis as "pregnant by design," would be better to call these pregnancies from stopping use of the method ; these pregnancies could be added back into the analyses.
Concurrent use of coitus-dependent methods	2*	No assessment of barrier method use or withdrawal.
Study duration	3	Follow-up is up to 2 years; 1-year rate is not reported separately
Statistical methods	2	Pearl Rate; cycles with intercourse not excluded.; perfect use estimate not correctly calculated (denominator includes all cycles)
Attrition	1	LFU<20%, detailed categories for reasons for discontinuation
Other	1	
Overall rank	3	
Results included		None

Symptothermal Methods: Single check

STUDY: Ecochard 1998
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Some inclusion criteria are provided and clearly defined; refusal rate is not.
Exclusion of populations not at meaningful risk of pregnancy	2	Not reported whether cycles without sexual activity were excluded. Excluded women using contraceptives. Couples with subfertility excluded but not reported how this was defined
Treatment of women in various reproductive categories that may impact fertility	2*	22% joined study in 6 months after delivery. 47% had used other methods in past (unclear time since last use). No subgroup analysis of either though multivariate analysis done for risk of pregnancy by these characteristics. Breastfeeding and perimenopausal-age women were excluded .
Description of study population	2	74.5% had used method for more than one year. Age and parity recorded. No analysis conducted by these characteristics.
FABM method(s) studied	1	Symptothermal method with fertility rules stated (some minor variations), but clear definition of the rules
Teaching of method	2*	Unclear number of encounters, teacher training, curriculum (though example of a chart provided)
Detection of pregnancy	2	Per correspondence with the author Rene Ecochard: all charts reviewed, pregnancy test issued for women with delayed menses/prolonged luteal phase. Not clear about follow up on early pregnancies at end of study.
Classification of pregnancy as intended/unintended	2	Pregnancy intentions reported every cycle; all pregnancies were not reported (including intended pregnancies); all pregnancies were included in typical use analysis; no training phase reported; no evidence of pregnancies excluded from typical use calculation
Concurrent use of coitus-dependent methods	2*	Barrier method use and withdrawal not assessed
Study duration	1	Included at least 12 m/13 cycles (cut-off period was 21 months) and included a 12m single decrement typical pregnancy rate report.
Statistical methods	2	Used single-decrement analysis, but unclear if cycles with no sexual activity were excluded. Perfect use rate incorrectly calculated.
Attrition	1	Loss to follow-up was 0; rates for discontinuation and reasons were reported

Other	1	None. Some pregnancies were excluded from the user plus method failure but this is not a quality issue for the study as a total pregnancy rate was reported
Overall rank	2	
Results included		Total pregnancy rates for 12 months all users (mostly experienced) Pregnancy rate among cycles in which there was only intercourse post temperature shift.

STUDY: Freundl 1999 European
Quality criteria rankings

Domain	Proposed Ranking	Rationale
Inclusion/exclusion criteria	2	Clear description of criteria but no denominator of screened or refusal rate
Exclusion of populations not at meaningful risk of pregnancy	2	Sterilization and IUD excluded; No mention of subfertility; No exclusion of cycles without intercourse
Treatment of women in various reproductive categories that may impact fertility	2*	About 10% over age 40 without subgroup analyses Other reproductive factors like breastfeeding or hormonal contraception all excluded
Description of study population	2	Clear proportion of new versus established users, but not analyzed separately Good characterization of marital status, education, spacer/limiter, # children, religion
FABM method(s) studied	1	Symptothermal double check (Sensiplan) and Symptothermal single check (CLER) with citations for rules and description provided.
Teaching of method	2*	Teachers established, but way method taught or teachers trained not described and implicitly was different between centers (with two different STM methods)
Detection of pregnancy	2	Self-report with chart reviews of all cycles submitted and reviewed for signs of prolonged menses/luteal phase.
Classification of pregnancy as intended/unintended	1	Intention to avoid or conceive recorded at beginning of each cycle
Concurrent use of coitus-dependent methods	1	Systematically looked at subgroup analyses for a subset of the participants (only on the interim report). Barrier and withdrawal use combined together.
Study duration	1	Followed for 1 year
Statistical methods	2	Multiple decrement Kaplan Meier analysis per email communication with Godehardt (statistician)
Attrition	2*	Drop out and loss to follow-up rate low overall (higher in single-check), but no analysis of characteristics associated with dropout.
Other	2	No information on how teachers at each center selected or invited the volunteers
Overall rank	2	
Results include		12-month overall Kaplan Meier pregnancy probabilities; Pearl Rates by STM and STM-mix in interim report.

STUDY: Johnston 1979
Quality criteria rankings (see above)

STUDY: Marshall 1976
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Unclear how recruitment occurred or what participation rate was. Inclusion criteria adequately defined.
Exclusion of populations not at meaningful risk of pregnancy	2	Not reported whether subfertile women or women using contraceptive methods were excluded. Not reported whether cycles with no intercourse were excluded. No explicit inclusion of women using contraceptive agents.
Treatment of women in various reproductive categories that may impact fertility	2*	Women with irregular cycles excluded. For parous women, least one ovulatory cycle since birth. Breastfeeding not addressed. Perimenopausal-age women included in small numbers. No subgroup analyses.
Description of study population	3	Age and parity well described. All users were experienced; no new users.
FABM method(s) studied	1	Clearly defined. Single-check symptom-thermal.
Teaching of method	2*	Instructed within a medical correspondence service; no details given.
Detection of pregnancy	2	Charts returned every 2 cycles, with reminders and assessed for signs of pregnancy; no pregnancies were excluded.
Classification of pregnancy as intended/unintended	2	Unclear how pregnancy intentions assessed.
Concurrent use of coitus-dependent methods	2*	Withdrawal and barrier method use not assessed
Study duration	3	Maximum 2.5 years, reported in one Pearl Rate; one-year rate not reported.
Statistical methods	2	Pearl Rate (see above)
Attrition	1	Low loss to follow up; reasons and rates for discontinuation not reported.
Other	1	No other significant concerns.
Overall rank	3	
Results included		None

STUDY: Rice, Lanctôt, Garcia-Devesa 1981
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion criteria very clear; refusal rate unclear.
Exclusion of populations not at meaningful risk of pregnancy	2	Cycles without intercourse not excluded; no hormonal methods, IUDs, sterilization – as indicated by the fact that people were considered to have dropped out when they transitioned to using these methods. Women with subfertility not explicitly excluded.
Treatment of women in various reproductive categories that may impact fertility	2*	Includes women 40-44 (12.5%) breastfeeding, post-hormonal contraceptive. Unclear about postpartum women. No subgroup analyses
Description of study population	3	Likely (but unclear) that most users were experienced. Members of NFP associations were recruited, and see between Tables II and III: Text suggests more experienced users were biased toward participating in this new study. When couples dropped out to have a baby reentered the study, they were treated as new participants for life-table analysis.
FABM method(s) studied	2	Focus on sympto-thermal method, provides detailed description of method but some aspects of the approach were combined to convey the simplest possible version of ST method in a way that was not consistent across sites.
Teaching of method	2*	Lay educators, but training information and intensity omitted; highlights couple-to-couple approach but doesn't explain on how this works
Detection of pregnancy	2	Self-report, but likely supplemented by monthly chart review.
Classification of pregnancy as intended/unintended	1	Prospective intentions recorded monthly, all cycles reported, analyzed, no cycles excluded; in one site (Colombia), not all questions were asked, but women in this group were considered to be avoiding if it was not specified, so dealt with in most conservative way in analysis.
Concurrent use of coitus-dependent methods	1	Analysis captures use of "contraceptive devices" during any portion of fertile period – see Tables IX and X
Study duration	1	Life table for 24 and 12 months. Also include Pearl rate.
Statistical methods	3	Pearl and life table method used; not clear if single or multiple decrement life table; perfect use pregnancy rates calculated incorrectly

Attrition	1	Detailed analysis of dropouts in terms of reasons for dropout, shift to other methods (p. 224 (only 35 couples LTFU) p. 228 plus Table V)
Other	1	None.
Overall rank	3	
Results included		None

STUDY: Weeks 1982
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion criteria minimal: all clients of Responsible Parenthood of San Diego, Jan 1979 to May 1980. Refusal rate not applicable.
Exclusion of populations not at meaningful risk of pregnancy	2	No exclusions. Did not exclude cycles without sexual activity. No explicit inclusion of women using other contraceptives. It can be inferred that women from this clinic were interested in FABM use.
Treatment of women in various reproductive categories that may impact fertility	1	Subgroup analyses for 11 different user characteristics (but missing number of cycles for denominator, which limits the interpretability somewhat)
Description of study population	1	All new users; Good characterization of characteristics and baseline motivations
FABM method(s) studied	2*	STM, but not described exactly; 9 women (6%) used either mucus only or temperature only
Teaching of method	2*	An organization taught method, but no details about teaching, training or materials
Detection of pregnancy	2	Monthly phone calls reviewing cycle dates (menses, peak day, temp shift) and self-reported pregnancy
Classification of pregnancy as intended/unintended	2	Baseline motivation to prevent or postpone, but no apparent updates after that
Concurrent use of coitus-dependent methods	1	Reported monthly and subgroup analysis performed.
Study duration	2	Follow-up is up to 7.4 cycles on average; method was to follow for "at least 3 months" with upper limit not defined. Planned follow-up one year
Statistical methods	2	Pearl Rate (13-cycle)
Attrition	2*	Loss to follow-up high but not quantified; no assessment of reasons for discontinuation
Other	2	Gaps in description
Overall rank	2	
Results included		Typical use pregnancy probability using the multiple decrement approach. Pearl rates reported for women who were using barrier methods versus abstinence only.

Symptothermal Methods: Double check

STUDY: Frank-Herrmann 2007
Quality criteria rankings

Domain	Proposed Rank	Rationale
Inclusion/exclusion criteria	2	Clearly defined inclusion (19-46 years old; 80% or more of cycles 25-35 days; intended to avoid pregnancy and willing to alert investigator at once with changed intentions; willing to record family planning intentions at start of each cycle; willing to record sexual activity, including sexual intercourse, genital contact, withdrawal, occasional barrier use; new users of the method; willing to participate for 12 cycles; breastfeeding, postpartum, post-contraceptive users only included after 3 cycles with an established 10 day or more luteal phase; no known history of subfertility or infertility; agreement not to use any other forms of contraception); provided denominator of entire dataset but did not specify how many women had been invited to join the data set originally
Exclusion of populations not at meaningful risk of pregnancy	2	Included substantial number of cycles (2625/17638 = 15%) but probably lower with no intercourse recorded in typical use rate; excluded women with known infertility and women using contraceptives;
Treatment of women in various reproductive categories that may impact fertility	2*	Included breastfeeding (but not amenorrhea) participants with an established luteal phase and women 40-45 (other populations excluded) and did not provide subgroup analyses
Description of study population	1	All new users and demographic characteristics well described including age and parity
FABM method(s) studied	1	Sensiplan Symptothermal Method well referenced
Teaching of method	2	Certified and trained teachers but training details not reported; number of encounters and curriculum referenced
Detection of pregnancy	1	All charts reviewed for pregnancy with targeted pregnancy testing for prolonged luteal phase. No active f/u recorded but all cycles included had charts reviewed (including the last one) so I would give this a "1"
Classification of pregnancy as intended/unintended	1	Pregnancy intentions assessed prior to every cycle and all cycles and pregnancies were included where the decision to avoid pregnancy was prospectively stated. Teaching phase pregnancies included in estimates.
Concurrent use of coitus-dependent methods	1	Measured and provided separate estimates for pregnancy for both groups.

Study duration	1	12-month single decrement life table reported for typical use
Statistical methods	2	12-month typical use single decrement life table pregnancy rate; cycles with no intercourse included
Attrition	2	Loss to f/u 6.7% at 13 cycles; reasons for discontinuation reported but not analyzed by differences in demographic variables
Other	1	None. Intensive study design is methodologically strong, but may have resulted in a highly selected group of users, which should be considered with respect to generalizability of the findings. Multiple subgroup analyses are helpful.
Overall rank	2	
Results that should be in table		Definitely included 13 cycle single decrement life table pregnancy probabilities for total (1.79), STM only (1.62) and STM mix (2.02); Included 13 cycle perfect use rate (0.43), genital contact/withdrawal only during fertile time (1.20), barrier use only during fertile time (0.59), unprotected sex during fertile time (7.56), protected and unprotected sex during fertile time (2.18), no documented sexual behavior during fertile time (0.49)

STUDY: Freundl 1999
Quality Criteria Rankings (see above)

STUDY: Medina 1980
Quality Criteria Rankings (See Above)

STUDY: Wade 1991
Quality Criteria Rankings (see above)

Hormonal methods: Persona

STUDY: Bonnar 1999

Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	1	Well-described criteria. Volunteers self-selected through press advertising
Exclusion of populations not at meaningful risk of pregnancy	1	Excluded cycles with no intercourse
Treatment of women in various reproductive categories that may impact fertility	2*	No sub-analyses presented
Description of study population	1	New users clearly characterized
FABM method(s) studied	1	Clearly defined
Teaching of method	2*	No teaching of the method by study design
Detection of pregnancy	2	Self-reported pregnancy with corroboration with data from the monitor
Classification of pregnancy as intended/unintended	2	Unclear on how pregnancy intentions were assessed after initial baseline assessment
Concurrent use of coitus-dependent methods	2*	Not assessed, however at baseline, participants were instructed to abstain during the fertile window and engage in unprotected intercourse during the infertile period
Study duration	2	13 cycles but unclear analysis; per study author single decrement life table but given unclear in manuscript, ranked 2.
Statistical methods	2	Single decrement life table analysis excluding cycles with no intercourse per communication with author G. Freundl. Given not stated clearly in publication, will rank this a 2 and not a 1.
Attrition	2*	Attrition adequately described and 50.4% contributed 13 cycles
Other	1	None
Overall rank	2	
Results included		None

Hormonal Methods: Marquette Monitor

STUDY: Bouchard 2012
Quality criteria rankings

Domain	Proposed Rank	
Inclusion/exclusion criteria	1	Well described (postpartum women able to access on-line Marquette Method charting); excluding post miscarriage. Denominator of excluded number reported.
Exclusion of populations not at meaningful risk of pregnancy	2	Study was based on postpartum women who by definition may not be at risk of pregnancy if not cycling. In addition, cycles after delivery were included even if (especially in breastfeeding women) these cycles are sub-fecund. All women were recently pregnant but did not explicitly exclude subfertile women. Excluded women using hormonal contraception. Cycles with no intercourse not excluded.
Treatment of women in various reproductive categories that may impact fertility	2*	Women were postpartum and breastfeeding only. However, perimenopausal-age women were included and no subgroup analysis was done. Also, unclear if any recent users of contraception though this would be unlikely.
Description of study population	1	All new users of the online charting system. Age and parity described.
FABM method(s) studied	1	Marquette method breastfeeding protocol with citations and clear rules provided.
Teaching of method	2*	Online charting system used with charting, education and health care professional support but no description of teacher training
Detection of pregnancy	2	Luteal phase of greater than 19 days prompted users to take a pregnancy test. No active follow-up reported at conclusion of the study
Classification of pregnancy as intended/unintended	1	Pregnancy intentions were assessed prior to the beginning of every cycle; no cycles excluded
Concurrent use of coitus-dependent methods	2*	Barrier method and withdrawal not assessed. Method promoted as recommending against use of these methods.
Study duration	1	12 months
Statistical methods	3	Life table analysis but the denominator included months since infant DOB, before the method was being used. This will underestimate the true failure rate as the denominator will be inflated with inclusion of months not at risk.

Attrition	2*	Insufficient information but dropout was high (74/198 women left in the study until 12 cycles)
Other	1	None
Overall rank	3	
Results included		None

STUDY Fehring 2009Quality criteria ranking (see above)

STUDY: Fehring 2013
Quality criteria rankings (see above)

STUDY: Fehring 2014
Quality criteria rankings (see above)

STUDY: Fehring 2017
Quality criteria rankings (see above)

Symptohormonal Methods: Marquette Mucus and Monitor

STUDY: Fehring 2008
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	2	Inclusion/exclusion criteria adequately defined (women who sought training in at one of four sites with initial intention to avoid pregnancy; excluded breastfeeding women, infertility and over 42 years). Refusal rate not documented
Exclusion of populations not at meaningful risk of pregnancy	2*	No reporting on whether cycles with no intercourse were included. Women without history of infertility treatment were excluded. No explicit inclusion of women using hormones
Treatment of women in various reproductive categories that may impact fertility	2*	Perimenopausal age women (defined as older than 42) excluded; Oligomenorrheic women included and no subgroup analysis; Breastfeeding excluded; Postpartum rate and post-hormonal contraception not reported
Description of study population	2	New user status unclear. Age and parity characterized
FABM method(s) studied	1	Women could choose either fertility monitor alone or cervical fluid alone or BBT alone or combination. Rules for using each not defined although citation provided for fertility monitor + cervical mucus. Sub-analyses were performed for women using each approach
Teaching of method	1	Training described, number of encounters described, curricula referenced and are available online
Detection of pregnancy	3	Pregnancies were detected by self-report only with no evidence of systematic correlation with charts. Pregnancy evaluations were done for every reported pregnancy. "Pregnancies were included if they were verified by an in-person pregnancy evaluation"
Classification of pregnancy as intended/unintended	2	The charting system "included having the couple verify their intent for using the fertility monitor to either achieve or avoid a pregnancy" but it is unclear how often this was assessed. No evidence that learning phase pregnancies were excluded
Concurrent use of coitus-dependent methods	2*	Barrier method and withdrawal use was not assessed. Women were instructed not to use other methods.
Study duration	2	Up to 12 months of use

Statistical methods	2	Not reported whether cycles with no intercourse were excluded. Kaplan Meier 12 month pregnancy rate comparable to single decrement life table
Attrition	2*	Lost to follow up not reported. Reason and rates for discontinuation not reported
Other	1	None
Overall rank	3	
Results included		None.

STUDY: Fehring 2011
Quality criteria rankings

Domain	Proposed ranking	Rationale
Inclusion/exclusion criteria	1	Study population well-described but all users of the online charting system. Refusal rate not applicable given the online nature of the study.
Exclusion of populations not at meaningful risk of pregnancy	2	Cycles with no intercourse not excluded. Women with subfertility or current contraceptive users not explicitly excluded. Low likelihood that current contraceptive users were included.
Treatment of women in various reproductive categories that may impact fertility	2*	Postpartum, breastfeeding and irregular cycles included with no subgroup analysis. Unclear if recent hormonal contraceptive users were included.
Description of study population	1	Users were all new to using the on-line system; good characterization of the differences between those who did and did not chart (among those who registered on-line). Age and parity described.
FABM method(s) studied	2	Clear method with citations but given that women could use the method in different ways, the groups should have been broken out separately.
Teaching of method	2*	Teacher training not reported, materials and educational process cited and easily available on-line.
Detection of pregnancy	2	If cycle with luteal phase longer than 19 days, participant prompted to take pregnancy test; unclear if active follow-up occurred.
Classification of pregnancy as intended/unintended	1	Prospective, cyclic detection of pregnancy intentions by prompting women with long luteal phase to do a pregnancy test.
Concurrent use of coitus-dependent methods	2*	Barrier and or withdrawal use not assessed
Study duration	2	6 month duration only
Statistical methods	2	Correct use calculation was appropriate (did not include cycles of incorrect use); Kaplan Meier survival analysis comparable to single decrement life table
Attrition	2*	Very high attrition (47/222 remaining at 6 months) in short study with no assessment of the characteristics of those who left versus those who stayed in the study. This is likely due to the fact that all comers to the website were included and some of those were not planning to necessarily avoid pregnancy for a long time.
Other	3	Given the very high attrition in such a short study, the results are difficult to interpret
Overall rank	3	
Results included		None

STUDY: Fehring 2014
Quality criteria rankings (see above)

STUDY: Fehring 2017
Quality criteria rankings (see above)

Appendix 7. Detailed Information About Comparison Studies

Seven moderate quality studies compared users of different fertility awareness-based methods with each other (1-7). Two were RCTs comparing Billings Ovulation Method versus the Thyma symptothermal method users. Neither reported adequate randomization or masked outcome assessment, and both had very high attrition (>70%). In a Colombian study, differences in pregnancy probabilities were not statistically different (33.6 v. 33.0) (4), but in the United States Study, the probability of pregnancy for symptothermal users was significantly lower than for Billings Ovulation Method users (11.2 v. 22.4) (p value <0.01) (7).

Thapa et al. conducted a non-randomized comparative study of the Billings Ovulation Method, the Modified Mucus Method, and a simplified variant of the Modified Mucus Method (2). The Billings Ovulation Method pregnancy probability among experienced users (2.5) was significantly lower than that for Modified Mucus method users (10.3) or the simplified variant users (11.5) (p value <0.001). However, users of the simplified variant of the Modified Mucus Method were less educated and more likely to self-identify as Catholic than users of the other two methods.

Freundl et al., 1999 compared pregnancy probabilities between French symptothermal single-check users and double-check Sensiplan users in a mixed population of new and experienced users (3). Sensiplan users had a lower pregnancy probability (2.6, 95% CI: 1.4-3.8) than single check method users (8.5, 95% CI: 3.6-13.4). However, the Sensiplan users were more likely new to using the method, unmarried, without children, and more highly educated.

Three Marquette Method studies (all of moderate quality), provided effectiveness for the different approaches to use of the method (monitor-only, mucus-only, monitor plus mucus). One reported no significant differences in first-year typical use pregnancy probabilities for users of the three different Marquette methods (5). One moderate quality RCT compared monitor-only versus mucus-only (1) users; about 60% of women dropped out before completing the trial. Randomization was adequately described but there was not masked outcome assessment. The unintended pregnancy probability was significantly lower among monitor-only users (6.8) versus mucus-only users (18.5) ($P < 0.001$).

References

1. Fehring RJ, Schneider M, Raviele K, Rodriguez D, Pruszynski J. Randomized comparison of two Internet-supported fertility-awareness-based methods of family planning. Contraception 2013 Jul;88(1):24-30.
2. Thapa S, Wonga MV, Lampe PG, Pietojo H, Soejoenoes A. Efficacy of three variations of periodic abstinence for family planning in Indonesia. Stud Fam Plann 1990 Nov-Dec;21(6):327-34.
3. Freundl G. European multicenter study of natural family planning (1989-1995): Efficacy and drop-out. Advances in Contraception 1999;15(1):69-82.
4. Medina JE, Cifuentes A, Abernathy JR, Spieler JM, Wade ME. Comparative evaluation of two methods of natural family planning in Columbia. American journal of obstetrics and gynecology 1980 Dec 15;138(8):1142-7.
5. Fehring RJ, Schneider M. Effectiveness of a Natural Family Planning Service Program. MCN The American journal of maternal child nursing 2017 Jan/Feb;42(1):43-9.
6. Fehring RJ, Mu Q. Cohort efficacy study of natural family planning among perimenopause age women. Journal of obstetric, gynecologic, and neonatal nursing : JOGNN / NAACOG 2014 May-Jun;43(3):351-8.
7. Wade ME, McCarthy P, Braunstein GD, Abernathy JR, Suchindran CM, Harris GS, et al. A randomized prospective study of the use-effectiveness of two methods of natural family planning. American journal of obstetrics and gynecology 1981 Oct 15;141(4):368-76.

Urrutia RP, Polis CB, Jensen ET, Greene ME, Kennedy E, Stanford JB. Effectiveness of fertility awareness-based methods for pregnancy prevention: a systematic review. Obstet Gynecol 2018; 132.
The authors provided this information as a supplement to their article.