Fertility awareness-based methods of family planning: A review of effectiveness for avoiding pregnancy using SORT

Michael D. Manhart, PhD,a Marguerite Duane, MD, MHA, FAAFP, b April Lind, MD,c Irit Sinai, PhD,b Jean Golden-Tevald, DOd

From the aCouple to Couple League, Cincinnati, OH; bGeorgetown University, Washington, DC; cPark Nicollet Clinic, Maple Grove, MN; and dMorningstar Family Health Center, Clinton, NJ.

KEYWORDS:
Fertility awareness-based methods;
Natural family planning;
Contraception;
Effectiveness;
Family planning

Each year, over three-fourths of the women of reproductive age in the United States seek family planning services from primary care clinicians. Women and their doctors should be informed of all effective family planning options and their respective effects on a woman’s reproductive health. Family physicians are well-trained to support the behavior choices necessary for the successful adoption of any reversible family planning method. However, many are unfamiliar with fertility awareness-based methods (FABM) of family planning or have misconceptions about their effectiveness, complexity, or suitability for their patients.

FABM teach women to observe the physical signs and symptoms that follow hormonal fluctuations throughout the menstrual cycle to identify a couple’s fertile window, which can be used to avoid or achieve pregnancy. One in 5 women in the United States expressed interest in using FABM when informed about such options. When correctly used to avoid pregnancy, modern FABM have unintended pregnancy rates <5 (per 100 women years). Studies of modern FABM show that their typical unintended pregnancy rates are comparable to those of commonly used contraceptives. This article presents a review of the FABM literature to (1) familiarize the reader with the physiological basis and features of modern FABM, (2) present and utilize a framework to evaluate clinical evidence using the Strength of Recommendation Taxonomy (SORT), which supports the effectiveness of modern FABM for avoiding pregnancy, and (3) serve as a resource for health care professionals offering FABM options to their patients.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Each year, over three-fourths of the women of reproductive age in the United States seek family planning services from primary care clinicians. In the interest of informed consent, women and their physicians should know about all available family planning options and their effect on reproductive health. Fertility awareness-based methods (FABM) of family planning rely on a woman’s understanding and recognition of her fertility. FABM provide couples with the

Fertility Awareness Based Methods of Family Planning Patient Education Handouts available at osteopathicfamilyphysician.org/current.

1877-573X/$ - see front matter © 2013 Elsevier Inc. All rights reserved. http://dx.doi.org/10.1016/j.osfp.2012.09.002
information they need to identify the days in each cycle when the woman is likely to conceive. Couples can use this information to guide their family planning decisions. Natural family planning (NFP) is the subset of FABM where couples refrain from genital contact during the fertile days to avoid pregnancy or engage in sexual intercourse during fertile time to achieve pregnancy. One in 5 women in the United States expresses interest in using FABM when informed about these methods. However, many physicians are unfamiliar with FABM or have misconceptions about their effectiveness, complexity, or suitability for their patients. As a result, FABM use is uncommon in the United States. This article addresses this knowledge gap.

When correctly used to avoid pregnancy, modern FABM have unintended pregnancy rates <5 (per 100 women years). FABM can be broadly classified into 4 types: (1) calendar-based methods depend on cycle length and counting cycle days, (2) cervical mucus-based methods rely on observing and tracking vulvar discharge, (3) sympto-thermal methods (STM) combine cervical mucus and basal body temperature observations, and (4) sympto-hormonal methods combine mucus observation with technology to detect urinary hormonal metabolites associated with ovulation and fertility. These categories exclude an additional FABM, the lactational amenorrhea method, which specifies that the probability of pregnancy is very small (6-month unintended pregnancy rate <2) as long as the woman is less than 6-month postpartum, in postpartum amenorrhea, and breastfeeding her baby fully or nearly fully. As this method is only available to women up to 6-month postpartum and does not require identifying specific fertile days each cycle, we exclude it from this review.

The goal of this review is to (1) familiarize the reader with the features of modern FABM, (2) evaluate clinical evidence supporting the effectiveness of modern FABM for avoiding pregnancy using the Strength of Recommendation Taxonomy (SORT), and (3) serve as a resource for health care professionals offering FABM options to their patients.

Mechanism of action: The physiology underpinning FABM

Fertility is a normal, healthy function of the human body, which allows couples to conceive new life. Healthy men are generally always fertile, beginning at puberty, whereas healthy women are only fertile a few days each menstrual cycle. This fertile window is about 6 days long—the 5 days before ovulation and the day of ovulation.

While only 12% of menstrual cycles are the stereotypical 28 days long, most healthy women have cycles that usually range from 26 to 32 days. A woman’s cycle is managed in the hypothalamus, where pulses of gonadotropin-releasing hormone regulate pituitary output of follicle-stimulating hormone and luteinizing hormone (LH), prompting the ovaries to produce estrogen and progesterone. In the first half of the menstrual cycle (follicular phase), follicle-stimulating hormone stimulates development of the follicle, which contains the ovum, and the growing follicle(s) secretes estrogen. This estrogen has a proliferative effect on the endometrium, and stimulates the glands within the cervix to produce fluid, transparent, and stretchy mucus, which allows for enhanced sperm motility, nourishment, and survival. Estrogen rising to a threshold level provides feedback to the pituitary gland to produce the LH surge, which triggers ovulation and the start of the luteal phase of the cycle. The ruptured follicle (corpus luteum) now predominantly secretes progesterone, which has 3 functions: (1) it matures the thickened endometrium into secretory tissue to nourish an embryo, (2) it changes the thin, watery cervical mucus into a thicker mucus plug, which inhibits sperm penetration, and (3) it has a thermogenic effect causing a rise in basal body temperature.

These hormonal processes have observable signs. Our understanding and recognition of these signs have increased in recent years, resulting in the development of various FABM. FABM are unique among family planning options, in that they can be used to either avoid or achieve pregnancy dependent on a couple’s choices.

Common features of FABM

Modern FABM are based on sound understanding of reproductive biology, follow precise protocols for correct use, and have been tested in well-designed studies to assess efficacy. Table 1 illustrates the common features of the various modern FABM.

The Standard Days Method (SDM) is the only modern calendar-based FABM. It is recommended for women with cycles that usually range from 26 to 32 days and identifies the fertile window as days 8-19 of the cycle, for all users in all cycles. To our knowledge, the various versions of the calendar-rhythm method, introduced around 80 years ago, have not been tested using contemporary clinical designs.

Mucus-based methods rely on observations of cervical mucus to identify the start and end of the fertile window. The Billings Ovulation Method (Billings) and the Creighton Model FertilityCare System (Creighton) instruct users to observe the mucus pattern, where the fertile period starts at the onset of secretions and ends 3 days after the last day of clear, stretchy, or lubricative mucus. In the TwoDay Method (2day), a woman considers herself fertile on any day in which she noted secretions of any kind on that day or the day before.

STM are a group of methods that use mucus observations in combination with daily basal body temperatures to identify the boundaries of the fertile window. Some variations of STM may also include calendar calculations, optional cervical palpation, internal mucus checks, or mid-cycle cues that may indicate proximity to ovulation.

Finally, sympto-hormonal methods employ in-home technology to directly measure urinary hormones, including estrogen metabolites and LH, in combination with standardized mucus observations, to determine the fertile window.
Development of criteria to evaluate evidence

Several family medicine journals have adopted SORT to allow authors to assess the quality of individual studies and overall strength of a body of evidence.\textsuperscript{10} SORT reviews assist physicians in evaluating outcomes that matter to patients and incorporating best evidence practices.

For pragmatic reasons, FABM do not easily lend themselves to randomized controlled trials. The various FABM include different inclusion or exclusion criteria and often attract different people.\textsuperscript{11} When Grimes et al.\textsuperscript{12} conducted a Cochrane systematic review of randomized controlled trials of FABM, they identified only 3 studies, each with methodological problems. Efficacy studies of FABM typically employ single-armed cohort designs, and determining the quality of those studies is not always straightforward. Therefore, we sought to develop criteria to evaluate peer-reviewed published studies of FABM to allow classification under SORT.

Using criteria recommended by Lamprecht and Trussell\textsuperscript{13} and input from several FABM research scientists, we identified Critical, Important, and Useful criteria to evaluate published FABM efficacy studies (Table 2).

Critical criteria included:

- prospective study design;
- only sexually active, fecund women admitted;
- appropriate sample size to address the research question;
- efficacy calculations are from the start of method use; if a learning phase is reported, it is not distinguished from later cycles in efficacy calculations;
- at least a 1-year follow-up;
- analysis of unintended pregnancies employs life table or survival approaches rather than the Pearl index to avoid the documented biases inherent to the Pearl index\textsuperscript{14};
- all pregnancies are recorded;
- the protocol underwent institutional review or was developed by a governmental agency with multidisciplinary input to protect the rights of participants.

The 12 Critical criteria were awarded 4 points for each element present or 0 points if missing. For some criteria, 2 points were awarded when a rationale was evident (eg, a comparative trial which collected complete prospective pregnancy intentions in one arm but not the other). The additional criteria (labeled Important or Useful in Table 2) served to further differentiate between studies, but did not compensate for a deficit in any of the 12 Critical criteria. Using this grading scheme, each study could earn up to 56 points (Table 2).

A Medline search was conducted using the search terms: NFP, fertility awareness, rhythm, calendar method, STM, etc.
ovulation method, effectiveness, and clinical trials. All publications from 1980 onwards were identified, as this year marked the landmark World Health Organization study of Billings, which is considered a seminal work in the study of FABM efficacy. To ensure completeness, the list was compared with a regularly produced bibliography of FABM-related studies [Richard Fehring personal communication].

Two authors independently reviewed each of the selected articles, and their respective scores were compared. Differences in scoring were resolved through group discussions to reach a consensus. Once the studies had assigned consensus scores, they were assigned a SORT evidence level score. Studies earning a positive score in all 12 Critical criteria (Z > 40 points) were considered robust and therefore met the SORT criteria of evidence Level 1.

Evidence for effectiveness in avoiding pregnancy

Our literature search yielded 29 peer-reviewed clinical studies published since 1980 that evaluated the effectiveness of one or more FABM to avoid pregnancy. Despite the paucity of data, at least 1 well-conducted, robust, clinical trial (SORT evidence, Level 1) is documented for each of the major FABM. Two studies of STM tested the method, either with abstinence or with barrier use during the fertile window. Each met the criteria for Level 1 evidence, giving an A rating for the strength of recommendation (SOR), defined as a recommendation based on consistent and good-quality, patient-oriented evidence. Creighton also meets the SOR criteria of A when examining the evidence, but for correct use only. Billings (with 1 Level 1 and 1 Level 2 studies), Marquette, SDM, and 2day (each with 1 Level 1 study) meet the criteria for SOR B, defined as a recommendation based on inconsistent or limited-quality, patient-oriented evidence.

Table 3 illustrates the best evidence of effectiveness in avoiding pregnancy for various FABM and lists the documented correct- and typical-use rates of unintended pregnancy for each. Collectively, these studies reflect exposure of 8200 women in over 107,000 cycles. The major FABM all have well-documented, correct-use effectiveness rates; couples can expect 0.4-5.0 unintended pregnancies per 100 women years when FABM are used.
correctly. Overall, these correct-use rates are comparable to those of many contraceptives.25

Typical-use unintended pregnancy rates for the STM version used in European clinical trials are comparable to rates seen with hormonal contraceptives in a prospective cohort trial,26 with 1-2 unintended pregnancies per 100 women years. In more recent studies, the unintended pregnancy rate for other FABM is 10-14 per 100 women years, similar to typical-use rates for some hormonal and barrier contraceptives. Defining typical-use effectiveness rates for Creighton is challenging due to the underlying difference in approaching pregnancy intentionality. Unlike the studies of other FABM, in which pregnancy intentions were recorded prospectively, in the larger prospective Creighton studies, intentionality was defined by the couple’s behavior during the fertile window.19 When a couple knowingly engaged in sexual intercourse during the identified fertile window, this was designated as achieving-related behavior, regardless of their prospectively stated intention to not conceive. Accordingly, typical-use effectiveness rates for Creighton remain undefined. A current study is examining this issue by simultaneously capturing couple intentionality by several different measures, including those traditionally used in studies of other methods.27

Discussion

A search of literature since 1980 yielded approximately 30 published cohort studies of FABM where pregnancy (either intended or unintended) was an outcome, reflecting the significant lack of research in this area. Yet, all of the major modern FABM have at least 1 well-conducted robust clinical trial documenting effectiveness in postponing pregnancy.

These studies show that FABM have unintended pregnancy rates comparable to those of many other methods.

Adopting any family planning method requires behavior modification; for FABM, the importance of behavioral choices during the fertile days is taught as part of counseling in method use. Some FABM (ie, Billings, STM, and Creighton) provide a teacher-client interaction that includes motivational and structural support for learners in the early days, while habits are being formed. Other FABM (ie, SDM and 2day) can be taught to the clients in a regular office visit. Moreover, resources for several methods can be purchased directly by the user, without need of a teacher, and online instruction is increasingly available.28 Once learned, most FABM can be used throughout a couple’s reproductive life, which renders these methods highly cost-effective.

Our review highlights some limitations of the evidence for FABM that physicians should be aware of. First, there have been about 30 studies of FABM conducted in more than 30 years, and only about one-third of them have been of high quality. Furthermore, some methods have been tested only in developing countries, whereas others have been tested only in Europe or the United States, thus limiting generalization of the findings. Finally, our exploration of electronic charting tools, applications, and devices revealed that many are not clear on the underlying FABM or rules being employed and, in many cases, leave the user unable to determine the level of evidence behind the applications. Physicians therefore should familiarize themselves with qualified teachers or programs in their areas to which they can refer their patients who choose to use FABM, as well as online and other available resources.

A common criticism of FABM is the perceived difficulty in learning to use the methods properly. Nevertheless, some

<table>
<thead>
<tr>
<th>Method</th>
<th>1-y probability unintended pregnancy (%)</th>
<th>SORT evidence level</th>
<th>Score</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct use</td>
<td>Typical use</td>
<td>Score</td>
<td>Citation</td>
</tr>
<tr>
<td>Billings</td>
<td>3.2</td>
<td>22</td>
<td>2*</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>1.1</td>
<td>10.5</td>
<td>1</td>
<td>52</td>
</tr>
<tr>
<td>STM</td>
<td>0.4</td>
<td>1.6</td>
<td>1</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>2.2</td>
<td>1</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>0.6†</td>
<td>2.02†</td>
<td>1</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>0.4†</td>
<td>1.43†</td>
<td>1</td>
<td>55</td>
</tr>
<tr>
<td>Creighton</td>
<td>0.5†</td>
<td>0.9</td>
<td>1</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>0.14†</td>
<td>0.5</td>
<td>1</td>
<td>43</td>
</tr>
<tr>
<td>Marquette</td>
<td>2.1</td>
<td>14.2</td>
<td>1</td>
<td>54</td>
</tr>
<tr>
<td>SDM</td>
<td>4.75</td>
<td>11.96</td>
<td>1</td>
<td>56</td>
</tr>
<tr>
<td>2Day</td>
<td>3.5</td>
<td>13.7</td>
<td>1</td>
<td>56</td>
</tr>
</tbody>
</table>

*As this is a post hoc analysis of the original World Health Organization data using the Critical criteria outlined in Table 2, the evidence level is reduced.
†These 2 studies are the only ones that calculated efficacy rate when FABM are used in conjunction with barrier methods on the fertile days.
‡Study included women with regular cycles, cycle lengths >38 d, >40 y of age, exclusively breastfeeding, and breastfeeding weaning.
§Typical-use effectiveness cannot be defined as in other trials, total pregnancy rate; Howard and Stanford, 17.12%; for Hilgers, not reported.
*SORT study level evidence = 1 for correct use only; see text for further explanation.
FABM have been successfully deployed and demonstrated effectiveness among diverse populations, including a substantial proportion of illiterate individuals, with minimal provider-client interaction.\textsuperscript{9,21,24} In contrast to earlier studies, these studies employed teaching protocols and methods appropriately tailored to the subject populations and demonstrate that education and literacy are not barriers for the effective FABM use.

As men are almost always fertile and conception of a new human life is a couple endeavor, family planning may best be discussed within the context of the couple. FABM encourage both partners to communicate, participate, and cooperate to adapt their behavior to achieve their family planning goals.\textsuperscript{29} Changing behavior may not be easy, but it is worthwhile. As family physicians, we regularly encourage behavioral modification, such as smoking cessation, choosing nutritious foods, developing healthy relationships, and exercising regularly. FABM empower women and couples through an increased understanding of their fertility and carry no risks of medication side effects, and limited data suggest that the quality of the relationship is improved among couples who use NFP (reviewed in Pallone and Bergus\textsuperscript{30}). Furthermore, with several different FABM available, patients have the opportunity to select the method that best fits their particular needs.

Couples interested in FABM are less likely to adopt these methods if their physician provides no information or inaccurate information about effectiveness and use. Commonly available information about FABM reports unintended pregnancy rates of \(~25\) per 100 woman years.\textsuperscript{31,32} These rates are derived from periodic surveys in the United States that ask women of reproductive age who became pregnant unexpectedly in the last year to recall which method they were using at the time of conception.\textsuperscript{33} These surveys pool all FABM data, including women who use their own version of periodic abstinence (as many as 86\% of the respondents), to generate an estimate of the unintended pregnancy rate. The inclusion of extensive data from self-devised methods or the outdated rhythm method does not accurately reflect the actual effectiveness of modern FABM. The author of this widely cited estimate has acknowledged that it masks the differences in the effectiveness of FABM.\textsuperscript{13} Data from high-quality studies show that modern FABM are highly effective, and recurrent reporting of this 1 statistic is misleading to both physicians and patients.

FABM can be used to effectively achieve or avoid pregnancy, because they do not disrupt a woman’s normal physiology as some contraceptives do. As a couple is likely to become pregnant if they have sexual intercourse during the fertile window, some authors label FABM as unforgiving, but this actually reflects the uniqueness of FABM as a true form of family planning. Some FABM programs offer special tips to optimize the probability of conception that are supported by clinical evidence.\textsuperscript{34,35} Because the signs and symptoms of ovulation and fertility are markers of good health, FABM can also aid in the diagnosis and treatment of infertility and other gynecologic problems.\textsuperscript{36,37} Many FABM also provide guidance for use during specific situations such as postpartum, in the premenopause years, or when discontinuing hormonal contraceptives (Table 1). Well-conducted studies among these special populations are rare\textsuperscript{38,39} and represent an important area for further research.

Conclusions

FABM are more than just an effective means to prevent pregnancy. They provide couples with a choice of whether and when to have children, can aid in the diagnosis and treatment of infertility and other gynecologic conditions, and can help couples embrace the emotional and relational aspects of their sexuality.

Development and application of criteria to evaluate the quality of evidence within SORT reveal that contemporary FABM can be as effective as hormonal contraceptives without the inherent health risks. Further research is needed to replicate best evidence for avoiding pregnancy among different populations and define the level of evidence when FABM are used to achieve pregnancy or when used during times of fertility transitions (eg, postpartum, premenopause, or when discontinuing long-acting contraceptives).

Like most contraceptives, FABM require educating patients about appropriate use and rely on the motivation and compliance of the user to be most effective. In contrast to contraceptives, however, most FABM inherently include the education element, and couples learn how their motivation and behavior affect their reproductive health. As family physicians, we are well trained to support such behavior modification. Currently, most physicians are not aware of the availability and effectiveness of FABM, but they should educate themselves and seek out trained providers of FABM in their area, so that they can offer these options and support couples who choose the FABM approach to planning their family.

Acknowledgments

The authors wish to acknowledge the following individuals for their assistance with this project. Ilene Richmond and Geraldine Matus were early members of the team and provided valuable insight and perspective on the use of FABM beyond an NFP context. Drs Richard Fehring and Joseph Stanford provided critical input on the development of the grading scheme to evaluate the cohort trials, in addition to offering useful insight and perspective as this work developed. Dr Amaryllis Sanchez Wohlever provided a critical review of the manuscript.

Appendix. Web resources for further information

FACTS—Fertility Appreciation Collaborative to Teach the Systems:
Billings Ovulation Method:
References