BRIEF REPORT

The Performance of Fertility Awareness-based Method Apps Marketed to Avoid Pregnancy

Marguerite Duane, MD, MHA, FAAFP, Alison Contreras, PhD, FCP, Elizabeth T. Jensen, MPH, PhD, and Amina White, MD, MA

Introduction: In recent years there has been an explosion in the development of medical apps, with more than 40,000 apps now available. Nearly 100 apps allow women to track their fertility and menstrual cycles and can be used to avoid or achieve pregnancy. Apps offer a convenient way to track fertility biomarkers. However, only some use evidence-based fertility awareness-based methods (FABMs), which with ideal use have rates of effectiveness similar to those of commonly used forms of hormonal birth control. Since having a baby or preventing a pregnancy are important responsibilities, it is critical that women and couples have access to reliable, evidence-based apps that allow them to accurately track their fertility.

Methods: We developed a tool to evaluate and rate fertility apps. This tool is specifically designed to help couples avoid pregnancy.

Results: Results showed that the majority of fertility apps are not based on evidence-based FABMs or include a disclaimer discouraging use for avoiding pregnancy. However, at least 1 app in each FABM category (except symptohormonal methods) had a perfect score on accuracy.

Conclusion: Relying solely on an app to use an FABM, without appropriate training in the method, may not be sufficient to prevent pregnancy. (J Am Board Fam Med 2016;29:508–511.)

Keywords: Biomarkers, Birth Rate, Contraception, Female, Fertility, Menstrual Cycle, Natural Family Planning, Pregnancy

The field of women’s health and fertility tracking applications (apps) has recently exploded, with nearly 100 apps available to help women track their cycle.1 The most popular apps have been downloaded over 1 million times each, and up to 60% of women express interest in using natural or fertility awareness-based methods (FABMs) to prevent pregnancy.2 These methods are attractive because they lack medical side effects, are effective, and can empower women with knowledge about their bodies. For each evidence-based method (Billings, Creighton, two-day, symptothermal, symptohormonal, standard days, and lactational amenorrhea methods), there are Strength of Recommendation Taxonomy level 1 studies that demonstrate that these methods, when used correctly, have rates of effectiveness similar to those of commonly used forms of hormonal birth control.3–6

The effectiveness of FABMs depends on women observing and recording fertility biomarkers and following evidence-based guidelines. Apps offer a convenient way to track fertility biomarkers, but only some use evidence-based FABMs.2 Until now
there have been no objective assessments of the apps designed for use to avoid pregnancy. In this study we developed a rating tool with specific criteria to quantify an app’s response to real cycle data based on the clinical guidelines evaluated in level 1 studies.

Methods

We identified 95 apps for study via iTunes, Google, and Google Play searches. Of those, we excluded 55 apps because they either had a disclaimer prohibiting use for avoiding pregnancy or did not claim to use an evidence-based FABM as described in Manhart et al.3

The rating system was developed based on criteria used by Family Practice Management to evaluate medical apps.7 We rated each app for 10 clearly defined criteria (each on a 5-point scale), which were weighted based on their level of importance for avoiding pregnancy (Table 1).

A standardized data set of 7 cycles of daily fertility observations, derived from real cycle data, was used to determine the apps’ accuracy in identifying potential days of fertility. For each cycle, evidence-based fertile days (FDs) were determined by applying specific guidelines for each FABM, as evaluated in peer-reviewed studies.3 The accuracy of each app was determined by comparing evidence-based FDs to the fertile days of each cycle as identified by the app, called the app-defined FDs (Figure 1).

Apps that did not predict fertile days scored high on accuracy only if they recommended prior FABM training apart from the app.

Results

Of those reviewed, 30 apps predicted days of fertility for the user and 10 did not. Table 2 ranks the apps based on the mean accuracy and authority.

Table 1. Criteria for Rating Apps

<table>
<thead>
<tr>
<th>Very important criteria (weighted ×3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authority</td>
</tr>
<tr>
<td>Source of rules app uses to determine fertile days</td>
</tr>
<tr>
<td>Accuracy (method)</td>
</tr>
<tr>
<td>Effectiveness of the method on which app is based</td>
</tr>
<tr>
<td>Accuracy (observations)</td>
</tr>
<tr>
<td>Accuracy of app in predicting fertile days</td>
</tr>
<tr>
<td>Support</td>
</tr>
<tr>
<td>Ways to have questions answered</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Important criteria (weighted ×2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
</tr>
<tr>
<td>Ability to enter additional data, use with irregular cycles</td>
</tr>
<tr>
<td>Cost/pricing</td>
</tr>
<tr>
<td>Cost of app, transparency in pricing</td>
</tr>
<tr>
<td>Ease of use</td>
</tr>
<tr>
<td>Ease to learn and use the app and share data</td>
</tr>
<tr>
<td>Confidentiality</td>
</tr>
<tr>
<td>Presence of a user agreement detailing confidentiality</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Helpful criteria (weighted ×1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developer/sponsor</td>
</tr>
<tr>
<td>Associated with recognized FABM provider</td>
</tr>
<tr>
<td>Platforms available</td>
</tr>
<tr>
<td>Availability on multiple platforms</td>
</tr>
</tbody>
</table>

FABM, fertility awareness-based method.

Figure 1. A comparison of evidence-based fertile days (FDs) and app-defined FDs.
Table 2. Ranking of Fertility Apps Based on Mean Accuracy and Authority Scores*

<table>
<thead>
<tr>
<th>Ranking</th>
<th>Name of App</th>
<th>FABM</th>
<th>Platforms Available</th>
<th>Accuracy and Authority Score (Mean)</th>
<th>Total Score (Mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Apps that predict fertile days (n = 30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Ovulation Mentor†</td>
<td>Ovulation</td>
<td>Web</td>
<td>4.7</td>
<td>4.4</td>
</tr>
<tr>
<td>2</td>
<td>Sympto.org†</td>
<td>Symptothermal method</td>
<td>iOS/Android/Web</td>
<td>4.5</td>
<td>4.1</td>
</tr>
<tr>
<td>3</td>
<td>iCycleBeads†</td>
<td>Standard days method</td>
<td>iOS/Android</td>
<td>4.3</td>
<td>3.9</td>
</tr>
<tr>
<td>4</td>
<td>LilyPro†</td>
<td>Symptothermal method</td>
<td>iOS</td>
<td>4.3</td>
<td>3.8</td>
</tr>
<tr>
<td>5</td>
<td>Lady Cycle†</td>
<td>Symptothermal method</td>
<td>Android</td>
<td>4.3</td>
<td>4.1</td>
</tr>
<tr>
<td>6</td>
<td>mfNFP.net†</td>
<td>Symptothermal method</td>
<td>iOS/Android/Web</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>7</td>
<td>MyFertilityCharts.com</td>
<td>Symptothermal method</td>
<td>iOS/Android/Web</td>
<td>3.5</td>
<td>3.7</td>
</tr>
<tr>
<td>8</td>
<td>CycleProGo</td>
<td>Symptothermal method</td>
<td>iOS/Android/Web</td>
<td>3.4</td>
<td>3.9</td>
</tr>
<tr>
<td>9</td>
<td>2Day Method</td>
<td>Two day</td>
<td>iOS</td>
<td>3.3</td>
<td>3.2</td>
</tr>
<tr>
<td>10</td>
<td>Ova</td>
<td>Symptothermal method</td>
<td>Web</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td>11</td>
<td>OvaGraph</td>
<td>Symptothermal method</td>
<td>iOS/Android/Web</td>
<td>3.1</td>
<td>3.6</td>
</tr>
<tr>
<td>12</td>
<td>OvuView</td>
<td>Symptothermal method</td>
<td>Android</td>
<td>2.7</td>
<td>2.8</td>
</tr>
<tr>
<td>13</td>
<td>FemCal</td>
<td>Symptothermal method</td>
<td>iOS</td>
<td>2.5</td>
<td>2.2</td>
</tr>
<tr>
<td>14</td>
<td>Ovatemp</td>
<td>Symptothermal method</td>
<td>iOS</td>
<td>2.3</td>
<td>3.2</td>
</tr>
<tr>
<td>15</td>
<td>Natural Cycles</td>
<td>Proprietary</td>
<td>iOS/Android</td>
<td>2.3</td>
<td>3.1</td>
</tr>
<tr>
<td>16</td>
<td>Cylendar</td>
<td>Symptothermal method</td>
<td>Web</td>
<td>2.1</td>
<td>2.2</td>
</tr>
<tr>
<td>17</td>
<td>My Fertility MD</td>
<td>Ovulation</td>
<td>iOS</td>
<td>2.1</td>
<td>3.2</td>
</tr>
<tr>
<td>18</td>
<td>Menstrual Cycle Woman Log</td>
<td>Rhythm†</td>
<td>Android</td>
<td>2.1</td>
<td>1.9</td>
</tr>
<tr>
<td>19</td>
<td>Menstruation &amp; Ovulation</td>
<td>Rhythm</td>
<td>iOS</td>
<td>2.1</td>
<td>1.2</td>
</tr>
<tr>
<td>20</td>
<td>Cycles</td>
<td>Standard days method</td>
<td>iOS</td>
<td>1.7</td>
<td>2.2</td>
</tr>
<tr>
<td>21</td>
<td>iCycly</td>
<td>Rhythm</td>
<td>iOS</td>
<td>1.3</td>
<td>1.4</td>
</tr>
<tr>
<td>22</td>
<td>Period Log</td>
<td>Rhythm</td>
<td>iOS</td>
<td>1.1</td>
<td>1.8</td>
</tr>
<tr>
<td>23</td>
<td>Period Pace</td>
<td>Rhythm</td>
<td>iOS</td>
<td>1.0</td>
<td>1.7</td>
</tr>
<tr>
<td>24</td>
<td>Period &amp; Ovulation Calendar</td>
<td>Rhythm</td>
<td>iOS</td>
<td>1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>25</td>
<td>Pink Pad Pro</td>
<td>Rhythm</td>
<td>iOS/Android</td>
<td>1.0</td>
<td>2.3</td>
</tr>
<tr>
<td>26</td>
<td>Fertility Calendar</td>
<td>Rhythm</td>
<td>Android</td>
<td>1.0</td>
<td>1.4</td>
</tr>
<tr>
<td>27</td>
<td>Fertility Clock</td>
<td>Rhythm</td>
<td>iOS</td>
<td>0.7</td>
<td>1.4</td>
</tr>
<tr>
<td>28</td>
<td>Woman Calendar</td>
<td>Rhythm</td>
<td>iOS</td>
<td>0.7</td>
<td>1.2</td>
</tr>
<tr>
<td>29</td>
<td>Fertility &amp; Ovulation</td>
<td>Rhythm</td>
<td>iOS</td>
<td>0.5</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>Apps that did not predict fertile days (n = 10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>NFP Charting</td>
<td>Ovulation</td>
<td>iOS/Android/Web</td>
<td>4.3</td>
<td>4.5</td>
</tr>
<tr>
<td>2</td>
<td>Symptopro</td>
<td>Ovulation</td>
<td>Web</td>
<td>4.3</td>
<td>3.7</td>
</tr>
<tr>
<td>3</td>
<td>Fertility Pinpoint</td>
<td>Symptothermal method</td>
<td>iOS/Android</td>
<td>4.2</td>
<td>4.2</td>
</tr>
<tr>
<td>4</td>
<td>Kindara</td>
<td>Symptothermal method</td>
<td>Web</td>
<td>4.0</td>
<td>4.1</td>
</tr>
<tr>
<td>5</td>
<td>Groove Fertility Pro</td>
<td>Symptothermal method</td>
<td>iOS</td>
<td>2.7</td>
<td>3.2</td>
</tr>
<tr>
<td>6</td>
<td>FEMM</td>
<td>Symptohormonal method</td>
<td>iOS</td>
<td>2.7</td>
<td>3.2</td>
</tr>
<tr>
<td>7</td>
<td>NFP Project Caruso</td>
<td>Rhythm</td>
<td>Android</td>
<td>2.7</td>
<td>2.1</td>
</tr>
<tr>
<td>8</td>
<td>Charting App</td>
<td>Ovulation</td>
<td>Android</td>
<td>2.0</td>
<td>2.3</td>
</tr>
<tr>
<td>9</td>
<td>Lady Timer</td>
<td>Proprietary</td>
<td>iOS</td>
<td>1.3</td>
<td>2.5</td>
</tr>
<tr>
<td>10</td>
<td>Knowhen</td>
<td>Ovulation</td>
<td>Android</td>
<td>1.0</td>
<td>2.2</td>
</tr>
</tbody>
</table>

*The data is representative of the apps at the time they were reviewed (Spring of 2015). Current versions of the app may be different.

†This app had either a perfect score on accuracy (app-defined fertile days = evidence-based fertile days) or no false negatives (days of fertility classified as infertile).

‡Rhythm is not an evidence-based FABM.

Glow was originally included in the study because it did not appear to meet the exclusion criteria. However, the developers clarified that the app is not a substitute for an FABM. Therefore, it should not be used by women attempting to avoid pregnancy through the use of an FABM. FABM, fertility awareness-based method.
scores, since the total scores include some reviewer subjectivity (such as ease of use) and users may be more concerned with accuracy. Only 6 apps (marked with * in Table 2) had either a perfect score on accuracy (app-defined FDs = evidence-based FDs) or no false negatives (days of fertility classified as infertile).

Discussion
The majority of fertility apps are neither designed for avoiding pregnancy nor founded on evidence-based FABMs. Several popular apps use their own algorithms, which are difficult to assess because they have not been evaluated in peer-reviewed literature. Attractive apps are not necessarily effective and vice versa. At least 1 app had a perfect score on accuracy in each FABM category except sympto-hormonal methods. Apps that do not predict days of fertility may be still useful for experienced FABM users to electronically record their data. Success using FABMs depends on many factors, including the ability to accurately make and classify daily observations. Relying solely on an FABM app may not be sufficient to prevent pregnancy.

For a list of the apps excluded and additional SORT Level 1 studies, please visit: www.FACTSaboutFertility.org.

The authors acknowledge the physicians and researchers who contributed to the development of the rating tool and the standardized cycle data: Dr. Megan Janni, Dr. Gavin Puthoff, Dr. Mary Desi, Dr. Laura Covert, Dr. Catherine Ferguson, Brittany Kudrna, Dr. Richard Fehring, Dr. Joe Stanford, Dr. Mike Manhart, and Dr. Hanna Klaus. The authors also thank the individuals who assisted with the reviews of the apps, including Erin Adams, Teresa Bippus, Anna Churchill, Ana-Maria Dumitru, Tracie Drayer, Chloe Emmanuelle, Dr. Luis Garcia, Jeannette Garcia, Dr. Maritana Giron, Tracee Linder, Dr. Karen Poehalios, Cristina de Rosa, Shawna van Uden, Harri Wettstein, and Brian Young. Finally, the authors thank Christina Verni for editorial assistance.

References

doi: 10.3122/jabfm.2016.04.160022