# Fertility Awareness-based Method Apps for the Billings Method

Audrey D Smith and John L Smith PhD

*A version of this paper was submitted to the Journal of the American Board of Family Medicine in February 2017 but was not accepted for publication.*

## ABSTRACT

Recent research has shown that some fertility awareness-based method apps for family planning are highly accurate in predicting the fertile days in a cycle. The advances in the technology are demonstrated.

The competencies that are necessary for using the Billings Method of family planning are discussed in the light of a comparison of two apps designed for using the method. One app (NFP Charting) is only effective if the user has been fully trained in the method, the other (billingsMentor) may be used effectively without previous training in the method. The limitations that can be experienced by users are identified.

Being able to observe and record the fertility symptom and signs correctly is essential for all users. The significance of data visualisation techniques for fertility charts is demonstrated. A classification of users is defined. The implications for training in a method for app users vary with the type of user.

## INTRODUCTION

Fertility awareness-based methods (FABMs) of family planning are methods that use *fertility indicators* and associated *symptoms* or *signs* that change with hormone fluctuations throughout a woman's ovulatory cycle to predict a woman's fertility (Pallone and Bergus 2009).

Duane et al (2016) identified a total of 95 fertility tracking apps via iTunes, Google and Google Play. They compared 40 of these that were FABM apps, using a rating system based on criteria previously used to evaluate medical apps. They divided these apps into two groups, depending on whether the app predicted days of potential fertility or not. A heavily weighted criterion for the first group was the accuracy with which the app predicted potentially fertile days. The highest rated app in this group was ovulationMentor (the web service that is now also available as the IOS/Android app called billingsMentor (2014)). The highest rated app in the second group was NFP Charting (2007). Both these apps are designed for use of the Billings Method (Billings and Westmore, 2011), but with very different requirements.

Duane et al (2016) stated "The effectiveness of FABMs depends on women observing and recording fertility biomarkers and following evidence-based guidelines". "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".

NFP Charting addresses these reservations by recommending training in the Billings Method before or in conjunction with using the app. The user is responsible for applying every aspect of the method. In contrast, billingsMentor supports independent use of the method without prior training by automating most of the tasks involved in using the method. This paper examines the scope of automation and the changing role for training by reference to the user tasks in billingsMentor and NFP Charting.

# METHOD

Permission to use the data in Figures 3, 4 and 5 was obtained from each of the three billingsMentor users by on-line consent form.

The Billings Method uses fertility indicators generated in response to the ovarian hormones, oestrogen and progesterone. A rise in the level of oestrogen stimulates the cervix to generate several different types of mucus, and these are observable at the vulva. Progesterone causes a thick impenetrable mucus that blocks the entrance of the cervix (Odeblad, 1997).

There are two fertility indicators used in the Billings Method: (1) the sensation at the vulva and (2) the appearance of any discharge observed at the vulva. We speak of the *fertility symptom* and *signs* meaning the description or value of any observation of a fertility indicator. Sensation at the vulva maps to one symptom but there are four signs that may be observed for the appearance of a discharge: quantity, blood, colour and fluidity.

There are three tasks involved in using the Billings Method.

1. observing and describing the fertility symptom and signs;
2. classifying the group of fertility symptom and signs each day and generating a fertility chart;
3. decision-making based on rules (evidence-based guidelines) governing the timing of intercourse for avoiding and achieving pregnancy.

These tasks were analyzed in the context of the two apps identified above.

# RESULTS

## Observing and Describing Symptom and Signs

Observing the fertility symptom and signs is always a personal task for the woman. There is an optional “Comments” field in the NFP Charting data entry form (Figure 1) for describing fertility symptoms using free-form text. The user is only required to enter a classification of the group of symptom and signs (see next section).

There are two decision trees in billingsMentor to help a beginner observe her symptom and signs. Table 1 is an expansion of the simpler tree that is for sensation at the vulva. There are also links to short tutorials on the form for entering descriptions, as well as an on-line manual.

The development of the interactive form for entering descriptions in billingsMentor has been described previously (Smith and Smith, 2014). Table 2 shows the partially ordered lists of descriptors for the sensation symptom and the signs. The user is required to enter a complete description of her symptom and signs at the end of the day by choosing the highest ordered descriptor from each list that applied at any time during the day. There are about 150 realistic combinations of descriptors possible. The app ensures that the user enters a complete and realistic combination.

Table 1. Decision tree in billingsMentor for beginners observing sensation at the vulva (a drop-down list is the response of the user interface to tapping or clicking on the sensation data entry field).

| |
|---|
| Were you aware of a sensation at the vulva at any time during the day?<br>YES - Did the sensation change throughout the day?<br>YES - Choose the highest option in the drop-down list that describes one of the sensations.<br>NO - Choose a descriptor from the drop-down list.<br>DON'T KNOW - Use the 'More detailed help' link.<br>NO - Choose "none (dry)" from the drop-down list.<br>DON'T KNOW - Have you ever felt a sensation at the vulva?<br>YES / DON'T KNOW - Defer entering observations for at least a day or two and pay attention to any sensation as you go about your normal activities. You may need to wait until you next begin to menstruate. Remember that it is normal for there to be days when no sensation is felt. Use the 'More detailed help' link below.<br>NO - There may be days when nothing is felt at the vulva. Just be aware of any sensation as you go about your normal activities. Use the 'More detailed help' link below. You may want to wait a day or two before beginning to enter observations. |

Table 2. Descriptors used in billingsMentor data entry.

| Symptom/Signs | Descriptors |
|---|---|
| Sensation | dry (none), sticky, damp, moist, wet, slippery, wet and slippery |
| Quantity | none, only seminal fluid, only spotting, trace, some, lots |
| Blood | none, blood-color, spotting, clots, blood |
| Color | opaque, blood-color, clear |
| Fluidity | thick, creamy, thin, strings (stringy) |

## Classifying Symptom and Signs

The classification task is somewhat involved (Billings and Westmore, 2011; Smith and Smith, 2014). The order in which a beginner encounters the different classifications is determined by the training process (in the case of NFP Charting) and also depends on how soon a “Peak” day (a time close to ovulation) occurs. The NFP Charting user records her classification of the daily symptom and signs using one of the six stickers labelled Red, Green, Yellow, White Baby, Green Baby or Yellow Baby. The user must learn the semantics of these stickers in the context of classification. The stickers can be embellished with adornments shown to the right of the sticker menu in Figure 1 to indicate more detailed classification. The classification stickers for an example ovulatory cycle that could occur for any healthy fertile women are shown in Figure 2a.

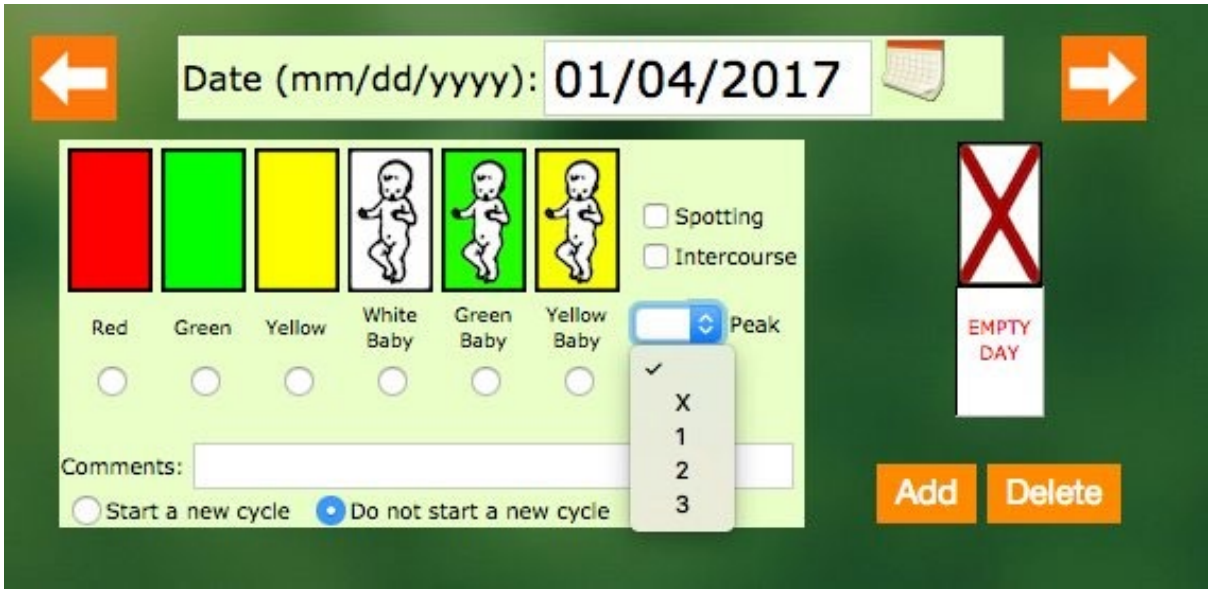



Figure 1 The form for describing and classifying daily observations in NFP Charting.

BillingsMentor does all classification automatically in response to data entry on each day (see Figure 2b). It informs the user about its decisions and the implications. The user may override a classification subject to validation against the underlying fertility model of the Billings Method.

The primary classification types are defined by the labels (infertile, wait, possibly-fertile, peak of fertility) on the four levels of the vertical axis of Figure 2b. The semantics of each level are intuitive. This provides a useful simple representation of fertility. The "wait" level is a fertility level that accounts for uncertainty. This can occur in the pre-ovulatory phase when seminal fluid prevents an accurate observation or when a change occurs from a day with possibly-fertile classification to a day which in other circumstances would be classified as infertile (see Billings and Westmore, 2011, 37-38).

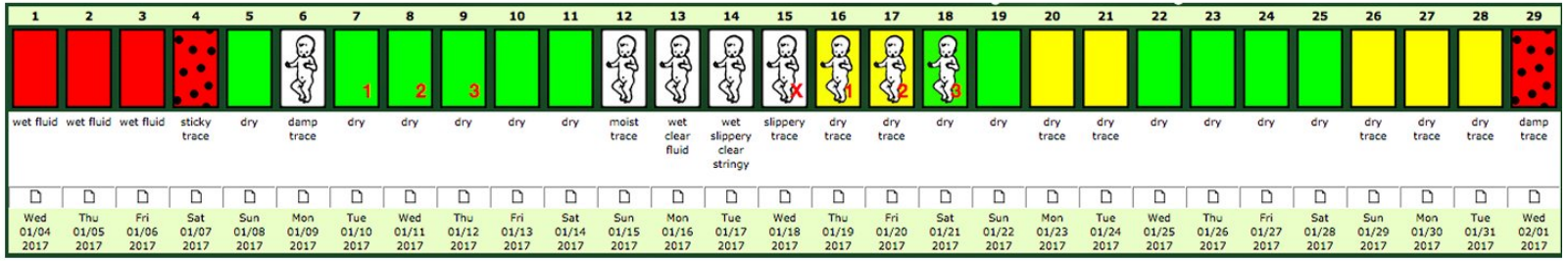
There is a corresponding set of sub-classifications in billingsMentor to those conveyed by the adornments in Figure 2a.

## Data Visualisation

Data visualization is the graphical display of abstract information for two purposes: sense-making (also called data analysis) and communication. Data visualization is only successful to the degree that it encodes information in a manner that our eyes can discern and our brains can easily understand (Few, 2013). This is the role of the fertility chart by allowing the comparison of fertility classifications over time in order to convey fertility awareness. Figure 2b is the equivalent billingsMentor chart to Figure 2a. There are three important differences.

Figure 2b uses the recommended (Few, 2013) and widely understood two-dimensional column chart to visualise levels of fertility over time. Figure 2a uses special graphic symbols (stickers) in a one-dimensional (time) chart.

(a)




(b)

| 2017 | 01-04 | 01-05 | 01-06 | 01-07 | 01-08 | 01-09 | 01-10 | 01-11 | 01-12 | 01-13 | 01-14 | 01-15 | 01-16 | 01-17 | 01-18 | 01-19 | 01-20 | 01-21 | 01-22 | 01-23 | 01-24 | 01-25 | 01-26 | 01-27 | 01-28 | 01-29 | 01-30 | 01-31 | 02-01 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 |
| peak | | | | | | | | | | | | | | | X | | | | | | | | | | | | | | |
| possibly_fertile | | | | | | | | | | | | | | | | 1 | 2 | 3 | | | | | | | | | | | |
| wait | | | | | | | 1 | 2 | 3 | | | | | | | | | | | | | | | | | | | | |
| infertile | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| *Sensation* | wet | wet | wet | sticky | dry | damp | dry | dry | dry | dry | dry | moist | wet | wet slippery | slippery | dry | dry | dry | dry | dry | dry | dry | dry | dry | dry | dry | dry | dry | damp |
| *Appearance* | fluid | fluid | fluid | trace | | trace | | | | | | trace | clear fluid | clear stringy | trace | trace | trace | | | trace | trace | | | | | trace | trace | trace | trace |

Figure 2 Fertility charts generated by (a) NFP Charting and (b) billingsMentor

Figure 2b uses its vertical dimension to denote the four levels of fertility: infertile, wait, possibly-fertile, Peak of fertility. The sticker adornment indices 1, 2 and 3 encode "wait" in Figure 2a.

Figure 2b uses a sequential color palette that encodes the Billings fertility model characteristics of the group of symptom and signs. This is recommended practice when data ranges from relatively low to relatively high values (Waskporn, 2012). It is still effective when printed in grey scale. Figure 2a uses highly contrasting colors contrary to recommended practice.

## Decision Making

There are four rules for avoiding a pregnancy (Billings and Westmore, 2011) that the NFP Charting user must interpret against the classifications of the cycle shown in the example of Figure 2a. The accuracy of the classifications is critical to decision making for both apps.

BillingsMentor generates the interpretation of the rules for both avoiding and achieving a pregnancy on each day. They are accessible for the current day, and retrospectively, by tapping/clicking on the interactive fertility chart. These interpretations refer to the sub-classification of counting indices when appropriate.

Duane et al (2016) measured the number of false negative and false positive days pertaining to potential fertility (the developing pattern) for predictive apps. BillingsMentor received a perfect score in the "accuracy and authority" criterion. Because sperm can survive for up to five days in hospitable cervical mucus, false negatives are seriously misleading for those wanting to avoid pregnancy.

False negative and false positive days can also occur in the classification of "wait" and infertile days. These classifications are significant for users wishing to avoid a pregnancy because the rules recommend that the user refrain from intercourse on "wait" days. Taking the human experts as the benchmark, a test of billingsMentor (Smith and Smith, 2014) showed it produced seven percent fewer days available for intercourse.

What should happen when users do not follow the evidence-based guidelines tends to be swept under the carpet in most FABM literature. If the users want to achieve a pregnancy they may lessen their chance by not following the guidelines. If they want to avoid a pregnancy they may become pregnant or the presence of seminal fluid may just thwart the app in its task to classify the underlying symptom and signs. Similar comments apply when users do not record their symptoms for every day.

When users do not follow the rules billingsMentor always recovers the situation using a conservative approach to classification. The outcome for the user wanting to avoid a pregnancy is that there will be usually be a day or more following the day when a rule was violated that the classification and interpretation of the rules is to wait (refrain from intercourse). Rule violation is always marked with a black triangle. Examples are shown in Figures 3 and 4.

## Learning

Both NFP Charting and billingsMentor accommodate conventional training procedures by enabling a teacher and a student to share her fertility data and communicate.

The current billingsMentor environment provides learning experiences through normal daily use in the following ways.

1. Online help and tutorials;
2. Validation of the group of symptom and signs;
3. Recognition of a Peak day;
4. Messages pertaining to the recognition of significant events and associated guidance;
5. Numerous advice and warning messages pertaining to proper use of the method;
6. "What if" exercises can be conducted by the user entering and deleting test sequences of symptom and signs. They can be entered retrospectively, so experiments can be carried out in a few minutes. Thus, examples of the context of a Peak and the evidence for a Basic Infertile Pattern as described by Billings and Westmore (2011) can be constructed.
7. By retrospective inspection of the interpretation of the rules.

## User Types

Table 3 gives a breakdown of user types for billingsMentor over the past few years. These percentages were calculated after eliminating the large proportion of registrants who did not enter a minimum number of daily symptom and signs.

Only about 25 per cent of clients were having regular fertile cycles when they began using billingsMentor. The research (Duane et al 2016, Smith and Smith, 2014) shows that billingsMentor is a highly reliable resource for these women if they observe and record their symptom and signs correctly.

The largest group of users (44 per cent) had ceased hormonal medication and most were in a period of infertility, sometimes prolonged. Figure 3 shows a section of a chart for a woman who had ceased hormonal medication a month earlier. After the first bleeding episode there were signs of raised oestrogen but no Peak before the next bleeding episode, and then she reached a Peak seventeen days later. The subsequent luteal phase was only eleven days which is just within the range expected after a healthy ovulation (Brown, 2011). The following cycles became more regular with good luteal phases. Some women take several months before fertility returns and we have data where fertility has not returned after a year.

Table 3. Types of billingsMentor User grouped by Purpose and Fertility Status.

| Purpose | Most recent Fertility Event before using billingsMentor | | | | Total |
|---|---|---|---|---|---|
| | Birth or Fertility Returned | Nursing/ Breastfeeding | Ceased hormonal medication | None | |
| Avoid Pregnancy | 16.5% | 11.4% | 23.7% | 9.1% | 60.9% |
| Achieve Pregnancy | 5.8% | 0.6% | 17.6% | 6.6% | 30.5% |
| Not sexually active | 0.2% | 0.1% | 2.4% | 5.8% | 8.6% |
| Total | 22.5% | 12.1% | 43.8% | 21.5% | 100% |

BillingsMentor maintains two sequences of day numbers. For example, in Figure 3 on March 26 the indices are 1 and 33. The start of a bleeding episode corresponds to the start of a new row of the chart and the first day is always numbered 1. The second index indicates the number of days since the last menstruation or else the beginning of charting. This index highlights that not all bleeding is menstruation as discussed below.

Women who are nursing constitute 12% of users and these usually experience a period of infertility. Figure 4 is part of the chart of a woman who had a long period of infertility while nursing. There is a short interval during which there were signs of blood (days 123-127) preceded by a noticeable change on two or three days, indicative of a slight rise in the oestrogen level and then a fall without any Peak occurring. This is a typical experience prior to a return of fertility. A basic infertile pattern consisting of the three groups of symptom and signs “dry; no discharge”, “damp; no discharge” and “damp; trace” had been recognised prior to the days shown on the chart.

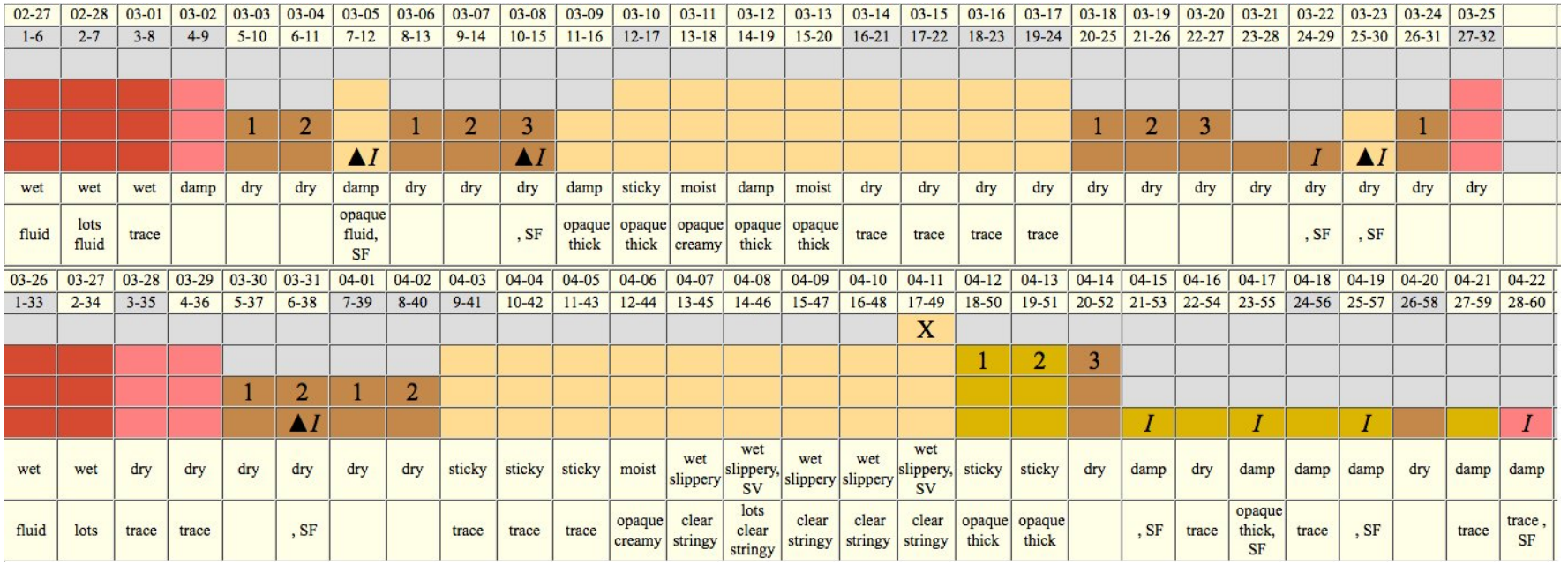



Figure 3 billingsMentor chart for a woman shortly after ceasing hormonal medication.

| 03-04 | 03-05 | 03-06 | 03-07 | 03-08 | 03-09 | 03-10 | 03-11 | 03-12 | 03-13 | 03-14 | 03-15 | 03-16 | 03-17 | 03-18 | 03-19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 97 | 98 | 99 | 100 | 101 | 102 | 103 | 104 | 105 | 106 | 107 | 108 | 109 | 110 | 111 | 112 |
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
| | | *I* | | *I* | | *I* | | | | *I* | | | *I* | | |
| dry | dry | damp | damp | dry | damp | damp | damp | damp | damp | damp | moist | damp | dry | damp | dry |
| | | trace | , SF | | , SF | | , SF | | | trace | trace , SF | | | , SF | |

| 03-20 | 03-21 | 03-22 | 03-23 | 03-24 | 03-25 | 03-26 | 03-27 | 03-28 | 03-29 | 03-30 | 03-31 | 04-01 | 04-02 | 04-03 | 04-04 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 113 | 114 | 115 | 116 | 117 | 118 | 119 | 120 | 121 | 122 | 123 | 124 | 125 | 126 | 127 | 128 |
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
| | | | | 1 | 2 | 3 | | | 1 | | | | | | 1 |
| *I* | | | | | | | | | | ▲*I* | | | | | |
| dry | moist | damp | moist | damp | damp | damp | moist | moist | damp | damp | moist | damp | damp | damp | damp |
| | opaque fluid, SF | trace | opaque fluid | trace | trace | trace | opaque fluid | opaque fluid | trace | trace | trace | trace | trace | trace | trace |

| 04-05 | 04-06 | 04-07 | 04-08 | 04-09 | 04-10 | 04-11 | 04-12 | 04-13 | 04-14 | 04-15 | 04-16 | 04-17 | 04-18 | 04-19 | 04-20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 129 | 130 | 131 | 132 | 133 | 134 | 135 | 136 | 137 | 138 | 139 | 140 | 141 | 142 | 143 | 144 |
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
| 2 | | | | 1 | | 1 | 2 | 3 | | | | | | | |
| ▲*I* | | | | | | | | | | | | | | | |
| damp | moist | damp | damp | damp | moist | damp | damp | damp | damp | dry | dry | dry | dry | damp | damp |
| | opaque fluid, SF | opaque fluid | opaque fluid | trace | opaque fluid | trace | trace | trace | trace | | | | | trace | trace |

Figure 4. billingsMentor chart for a woman during a period of nursing.

Women who have never given birth constitute the remaining group of users. Some of these were having difficulty in conceiving while others were just trying to understand their fertility. Figure 5 is a chart for a woman exhibiting a pattern of continuous discharge and a short interval between bleeding episodes. The woman received treatment from a medical specialist. The chart was useful for monitoring purposes.

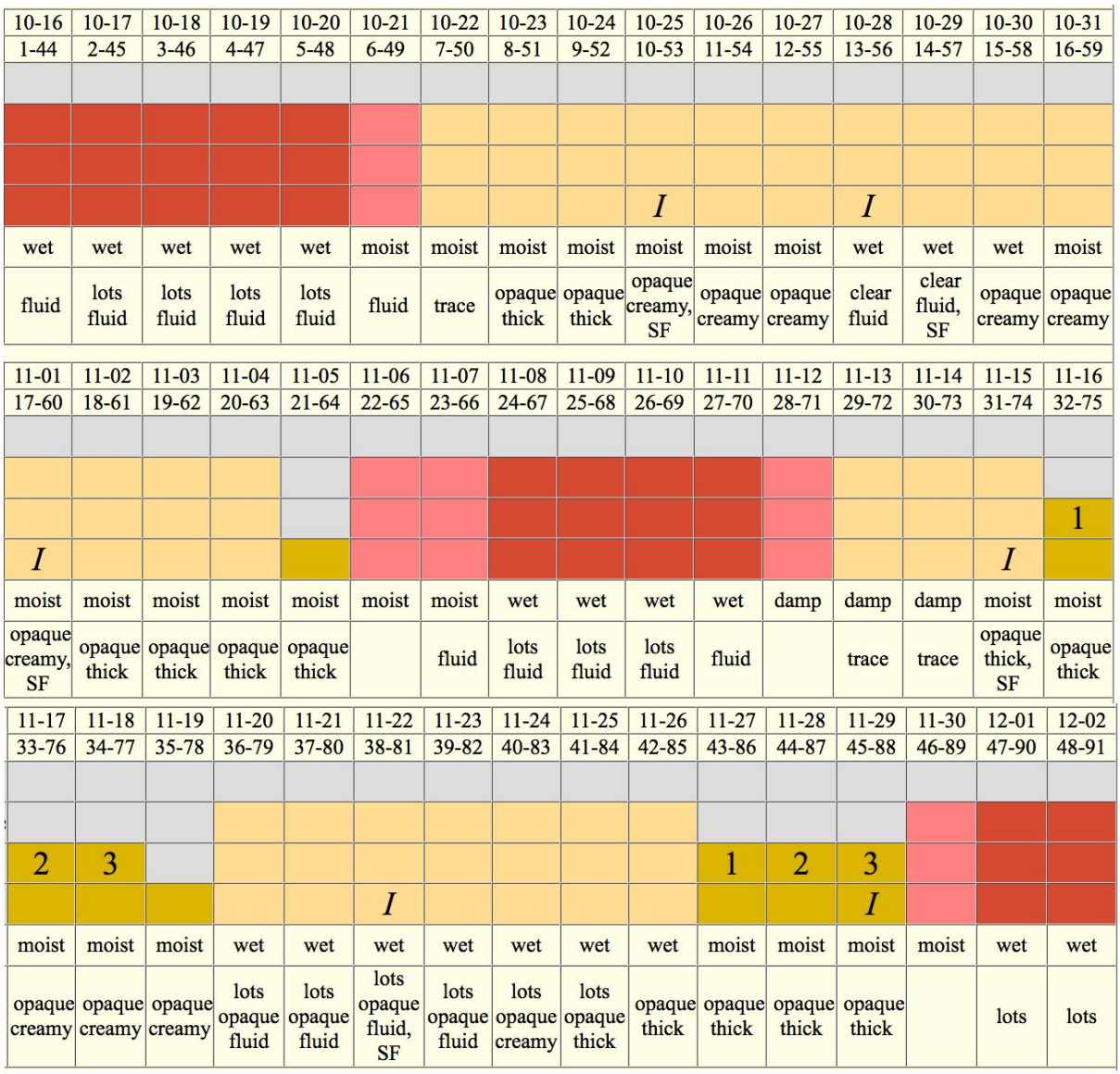


Figure 5. billingsMentor chart for an infertile woman experiencing continuous discharge.

billingsMentor analyses charts looking for anovulatory cycles or long periods of apparently fertile symptom and signs. The process is triggered when there is no Peak day or regularly occurring infertile pattern recognised within an appropriate time interval after the woman begins to enter symptom and signs. This may lead to an advisory message recommending that the user seek advice from a fertility expert.

# DISCUSSION

More research is required to determine the level of accuracy of observation and data entry, both for women who have had training by attending lectures and traditional consultation, and for those relying completely on the guided data entry provided by an app.

The use of stickers and symbols in NFP Charting combines classification with these particular forms of visualisation. They are a hangover from the hard copy record and fertility chart in an era before computers and digital displays became household items. They have influenced the visualisation used in most apps in every FABM category today.

BillingsMentor has an independent classification stage that allows the user to select any one of three forms of visualisation. Figure 2b demonstrates that a two-dimensional chart is best suited to visualising fertility over time in the context of the Billings Method. Communicating the main fertility classification levels by vertical axis increments and using color for the sub-classifications offers clarity that is helpful to anyone wanting to understand the method by using the app. The sequential color palette is even effective in grey scale as shown by Figures 3, 4 and 5. The ubiquity of presentation systems such as Microsoft PowerPoint and the display of simple statistics in the daily media means that this form of visualisation is completely intuitive to women today.

Because predictive apps with good user interfaces simplify tasks or completely relieve the user in some cases, training takes on a new meaning. By obviating the need for education in the more functional aspects of the Billings Method, billingsMentor sharpens the focus on fertility awareness.

The popularity of apps for all sorts of health-related services is creating new expectations for training. The real-time response of the app means that waiting for a response by email, message system or by appointment is no longer satisfactory. People may not want to become attuned to a method in every life circumstance. Training needs are about using the method for the current fertility situation in the context of the app.

The user statistics in Table 3 show a higher proportion of billingsMentor users wishing to achieve a pregnancy than would be the case in the wider community. They also demonstrate the importance of methods based on cervical mucus. Some other fertility indicators do not exhibit useful signs when women are experiencing long periods of infertility.

Orienting training to the needs of the different types of user in Table 3 is a way forward. When ovulation is delayed, a key to user satisfaction is for the app to identify the pattern of infertility. It was in these cases that a human expert performed better than billingsMentor (Smith and Smith, 2014). Until the algorithms equate to a human expert it is important that the app notify the user when expert assistance is necessary, possibly resulting in the use of the override functions in the app to declare a basic infertile pattern. Similarly, the user should be alerted when infertility patterns associated with conditions that require medical intervention are recognized.

The obvious event of menstruation dominates the thinking of most women, but it is ovulation that determines fertility and when menstruation will occur. Table 3 shows that more than half the women who started using billingsMentor had either stopped hormonal medication or they were nursing. In these cases, it is common for there to be a rise in oestrogen levels followed

by a bleeding episode recurring several times before the first ovulation takes place (Brown, 2011). Unless these normal occurrences are explained to the user, confusion and distrust can arise.

The reference document on the BillingsMethod (Billings and Westmore, 2011) does not address rule breaking in decisions. Consistent with always providing a classification and interpretation of the rules, billingsMentor generates a fall-back result in every case where the rules are broken. This is important in real life situations such as couples who are risk taking.

# REFERENCES


Billings, E. and Westmore, A. 2011. The Billings Method: Using the body's natural signal of fertility to achieve or avoid pregnancy. 9th ed. Australia: Penguin Books 2011.

Brown J. 2011. Types of ovarian activity in women and their significance: the continuum (a reinterpretation of early findings). Human Reproduction Update. 17:141–158.

BillingsMentor. 2014. billingsMentor. http://www.billingsmentor.org. Accessed February 13, 2018.

Duane, M. Contreras, A. Jensen, E.T. White, A. 2016. The performance of fertility awareness-based method apps marketed to avoid pregnancy. J Am Board Fam Med 29:508–11.

Few, S. 2013. Data Visualization for Human Perception: The Encyclopedia of Human-Computer Interaction, 2nd Ed. Interaction Design Foundation. Available at: https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/data-visualization-for-human-perception. Accessed February 13, 2018.

NFP Charting. 2007. NFP Charting. https://www.nfpcharting.com. Accessed February 13, 2018.

Odeblad, E. 1997. Cervical mucus and their functions. The Irish Colleges of Physicians and Surgeons 26:27-32.

Pallone, S.R. and Bergus, G.R. 2009. Fertility awareness-based methods: another option for family planning. J Am Board Fam 22:147–57.

Smith, A.D. and Smith, J.L. 2014. billingsMentor: Adapting natural family planning to information technology and relieving the user of unnecessary tasks. Linacre Quarterly 81:219-238.

Waskporn M. 2012. Choosing color palettes. Seaborn: statistical data visualization. Available at: http://seaborn.pydata.org/tutorial/color_palettes.html. Accessed February 13, 2018.