BillingsMentor is an automated Web-based service for the Billings Method of natural family planning in which the guidance and interpretation previously communicated from teacher to student is provided by programmed algorithms. There are two functions: (1) to instruct the client to generate proper descriptions of her fertility symptoms; and (2) to interpret the symptoms efficiently according to the Billings Method and to communicate the results to the client. The efficiency of BillingsMentor was tested by using the historical records of students who were under the guidance of a teacher to emulate their experience under the guidance of BillingsMentor. The results indicate that BillingsMentor performs as well as the teacher/student in recognizing the peak of fertility but it is slightly less efficient than the teacher/student in establishing the basic infertile pattern. Advantages that arise from adapting natural family planning to information technology are discussed.

Keywords: Natural family planning, Ovulation detection, Cervical mucus method, Billings Method, Informatics, EHealth

INTRODUCTION

Natural family planning (NFP) is implicit in obedience to the teaching Magisterium of the Catholic Church as it pertains to conjugal life (Norris 2013). Subject to this moral imperative every effort should be made to adapt NFP to changing cultural norms and to exploit technology for the benefit of any NFP user. The authors have applied their respective long-term experience in the teaching of NFP and in research and development in information technology to develop an online Web-based guidance and decision support system (BillingsMentor 2014) for the Billings Method (Billings and Westmore 2011) of NFP. This is the first publicly available NFP system to automatically perform all of the method-related tasks other than personal observation on behalf of the client. The purpose of this paper is to report an experiment that compares the efficiency of this system to that of a particular online Web-based teaching system in which an expert clinician guided the student (WOOMB 2010). The design of the decision support system and the results that are presented on the efficiency of the system should be of interest to all NFP stakeholders.

The fertility model that underlies the Billings Method is based on two fertility indicators that a woman can easily observe without requiring any devices. These reflect the cervical mucus and vaginal responses (Odeblad 1994, 1997) to the ovarian hormones, oestrogen and progesterone (Brown 1997, 2011). The indicators are the sensation felt at the...
Figure 1. Example bar chart for a cycle with the equivalent symbols that are specified in the Billings Method literature. Row 1 is the day number (days 1–29), rows 2–5 are the four levels of the bar chart, rows 6 and 7 show the fertility symptoms, and row 8 shows the equivalent symbols.
vulva (sensation) and the appearance of any discharge presenting at the vulva (appearance). Changes in these indicators can be interpreted according to the Billings Method to determine the approximate time of ovulation and days when conception might result from intercourse.

The Billings Method

The processes of the Billings Method are aimed at recognizing patterns in the sequence of daily observations of the symptoms exhibited by the fertility indicators (see figure 1 for an example; the day numbers now quoted only pertain to the example). A changing developing pattern (from a dry or sticky sensation to wet/slippy, days 10–15) followed by a definite change (day 16) determines the Peak of fertility (the approximate time of ovulation—day 15) in each ovulatory cycle. The sequence of daily observations is divided into two phases that occur cyclically. The luteal phase (Brown 1997) when the uterus is primed to accept a fertilized ovum is from a Peak to the beginning of the following menstruation (period) (days 16–29). The pre-ovulatory (pre-Peak) phase is from menstruation to the following Peak (days 1–15). A significantly unchanging sequence of symptoms in the pre-ovulatory phase determines a basic infertile pattern (BIP) (days 6–9) on which the couple rely in order to avoid pregnancy.

The Rules of the Billings Method

The rules of the Billings Method in terms of avoiding a pregnancy are simply stated as follows (Billings and Westmore 2011). There is other advice available about scheduling intercourse for users who want to become pregnant.

1. Avoid intercourse on days of heavy menstrual bleeding.
2. You can have intercourse on alternate evenings of your BIP.
3. When there is a change from your BIP, wait and see.
4. You may have intercourse at any time from the beginning of the fourth day after the Peak until your next period (menstruation).

The first three rules only apply prior to the Peak. The meaning of “wait and see” in rule 3 is that the user must wait for three days of BIP to occur before having intercourse again. Often the change covered by rule 3 is the beginning of the changing developing pattern culminating in a Peak (the fertile phase), so it is only in the case of delayed ovulation that there will be three days of BIP following the change. The restriction to alternative evenings in rule 2 is so that any change from the BIP will be observable and not obscured by the effects of seminal fluid. The delay until the fourth day specified in rule 4 accommodates possible differences between the time of ovulation and the time of the Peak and the survival time of the ovum (egg).

All methods of NFP rely on effective user education and conformance to the rules of the method. One critic (Betts 1984) of the Billings Method reported that many users in a trial claimed the method was “too complicated” or “too difficult to apply,” had “too many rules to follow” or “too many qualifications to remember.” The authors empathize with this viewpoint. What Betts was undoubtedly referring to was not just the complication of “wait and see” in rule 3, but also what can be termed the meta-rules (i.e., rules governing the rules) of the Billings Method. The four rules above only specify when intercourse may occur. Meta-rules comprise the knowledge that must be acquired in order to
apply the four rules. This relates to the underlying model of fertility that is based on the functions of the hormones, the cervix and sperm survival. Some meta-rules concern the empirical and qualitative guidelines that are associated with recognizing the Peak and the BIP. Others are for special circumstances such as when a Peak has been recognized but a period (menstrual bleeding) does not follow within the usual time limit. There are also meta-rules that specify when a user’s BIP should be re-evaluated. Beyond the complication of the meta-rules there is also the unanswered question: what quality and breadth of education in the method does the end user receive in real life, even with a teacher?

Some meta-rules that arise in the discussion of the results are now itemized (for a full exposition, see Billings and Westmore 2011).

A. The length of the fertile phase (beginning with the change described in rule 3) depends on mucus being present, typically for two or three days but it may be longer. The last day that the woman senses slipperiness or lubrication at the vulva is the Peak.

B. Normal luteal phases have a maximum length of 16 days followed by menstrual bleeding. Rule 4 above applies during the luteal phase. If bleeding does not start by day 17 past the Peak the application of rule 4 must cease and rules 2 and 3 should be applied until another Peak is identified.

C. If a wet or slippery sensation occurs during the first three days after a Peak has been identified then the counting of days as specified in rule 4 should cease and rules 2 and 3 should be applied until another Peak is identified.

D. If a woman is having regular cycles of normal length (maximum of 35 days) and she experiences a day or days following menstruation when her symptoms are of no sensation and no discharge (nothing felt/nothing seen) then this is her BIP (called a dry BIP).

E. If a woman is having regular cycles of normal length and she does not have a dry BIP, she may have a non-dry BIP. There is some sensation and/or some discharge in an unchanging pattern over the days following menstruation. This pattern should be verified over three usually consecutive cycles before it is applied in rules 2 and 3.

F. If a woman is in a time of infertility following the birth of a baby or following cessation of hormonal medication, or she is having long cycles, there may be a combination of symptoms that constitute a BIP instead of a dry or simple non-dry BIP. The combined BIP is identified as an unchanging pattern which is a combination of symptoms not leading to a wet or slippery sensation.

G. When fertility returns (a Peak is observed) after a time of infertility the current BIP is no longer applicable and a new BIP must be evaluated according to meta-rules D and E.

The internet together with personal computing and communications has now afforded various forms of automated support for many daily activities, so the authors developed billingsMentor to simplify and improve use of the Billings Method by removing the shortcomings mentioned above. The service is available at no charge (billingsMentor 2014). The programmed algorithms that underlie the service have two main functions: (1) to instruct the client to generate proper descriptions of her fertility symptoms; and
(2) to interpret the symptoms efficiently according to the Billings Method and to communicate the results to the client.

**Features of billingsMentor**

The instruction functions of billingsMentor are delivered through an on-line help manual, and by interactive guidance as the client enters the description of her daily fertility symptoms, and by using an interactive bar chart for visualizing fertility (figure 1) that is instantly available and displays more information than conventional charts in an intuitive way. Use of the bar chart eliminates the need for the client to learn the encodings for the stickers and symbols used in conventional fertility charts.

The automated guidance offered when the client enters her fertility symptoms is centered on a data entry form (figure 2)

**Table 1. The vocabulary for billingsMentor**

| Indicator characteristic | Key discriminatory words                                      |
|--------------------------|-------------------------------------------------------------|
| Sensation                | Dry (none), sticky, damp, moist, wet, 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)                       |
that was developed after analysis of the descriptive and semantic issues involved in traditional user charting and teacher/student interactions. The authors previously developed the first online Web-based teaching system for the Billings Method (WOOMB 2010). An analysis of more than 52,000 daily observation descriptions submitted by 213 long-term clients of that system over a period from 2006 to 2010 revealed a strong preference for a relatively small number of key words. This was to be expected as the user vocabulary in any discipline is primarily influenced by the educational material available and all the published literature on the Billings Method is based on a small vocabulary (Billings and Billings, 1997). A preferred vocabulary was identified and divided into the five categories shown in table 1, one category corresponding to the single characteristic of sensation, the others corresponding to four characteristics that can be observed in the appearance of any discharge. A small number of other words were identified, but their semantic mapping to the words in table 1 was obvious. The other feature of the chosen vocabulary is that it covers the range of intensities of each characteristic with a granularity aimed at being sufficiently discriminating to address the requirements of the Billings Method for recognizing change.

There are two versions of billingsMentor currently available. They differ in only one respect. Version 1 supports a wider vocabulary than that defined in table 1. The data entry form is different from figure 2 and a valid description of the fertility symptoms may consist of one or more English language phrases. It is version 2 that is the focus of the experiment described below. The client selects a descriptor from the five drop down lists apparent in figure 2, the activation of the lists being controlled so that a rational and complete selection is generated by the client.

The programmed controller for the data entry generates interactive reminders about the semantics of particular words and enforces a reporting discipline. Analysis of both the teaching service records and those of more than 270 clients of billingsMentor, version 1, and 250 clients of billingsMentor, version 2, showed that when clients are left to their own inclinations the descriptions that they record are often incomplete. Clients of the teaching service and of billingsMentor, version 1 are allowed to use a wide vocabulary of key words and qualifiers but they are not required to describe all of the four characteristics of a discharge. Clients of billingsMentor, version 2 must record a complete description, including both color and fluidity when more than a trace of discharge is present.

In their teaching document (Billings and Billings, 1997) Drs. John and Evelyn Billings stated that “Each woman is an individual and will describe her own mucus patterns in her own way.” This should not be taken as proscribing a limited vocabulary for describing the fertility symptoms in a discipline. The experiment reported here sheds light on whether the vocabulary shown in table 1, when used in conjunction with the reporting discipline of billingsMentor, version 2, is sufficient. There is another significant motivation for determining this approach. It is a straightforward task to port version 2 to any written language for which the character set is supported by the Unicode Transformation Formats of the internet. Thus, version 2 is already offered in a number of other languages including Spanish, Portuguese, French, Chinese, Russian and Arabic.

The main purpose here is to test the efficiency of billingsMentor in applying the rules of the Billings Method by using the available historical data from the teaching service in a laboratory experiment.
**METHOD**

Exempt status for this research was granted by the Marquette University Institutional Review Board. Consent to use data in this research was obtained from each participant by on-line consent form.

The authors introduced the English language teaching service in 2005 (WOOMB 2010) for the Billings Method that supports many teachers, each teacher having a teaching stream (New students are no longer accepted because of newer services such as billingsMentor). Students found the service through the normal channels associated with internet browsing. Teacher and student can communicate using an asynchronous message system in which messages are stored in the database along with the student’s descriptions of her daily fertility symptoms. Teacher and student can view a derived fertility chart display, together with the messages. The student is not restricted in the vocabulary that she can use to describe her fertility symptoms, but concise phrases are required according to the established practice for manual charting of fertility symptoms. The teacher can log changes to the student’s description in conjunction with their message exchanges.

The historical records for those 54 students of the teaching service who had received guidance from a teacher for at least 300 days (the fee was $100) were recovered, and client registrations corresponding to each of the 54 students were established in the internet service for billingsMentor (version 2).

The first 300 daily observations from each student were entered through the internet interface of billingsMentor (figure 2), emulating the actions that a client would have taken over a period of 300 days. In most cases, the observations were for 300 consecutive days, although occasionally a student would leave a gap of several days in her observations. In seven cases there was a long gap in the sequence because the student became pregnant during her tuition and subsequently returned for tuition while breastfeeding her baby.

The teacher/student records were generated between 2006 and 2009. In 2012, there was a change to the guidelines for recognizing a Peak that resulted in meta-rule A being stated as above (Ovulation Method Research & Reference Centre of Australia 2012). A simple modification to the programmed algorithms in billingsMentor for recognition of a Peak was made for the purpose of the experiment so that the algorithms conformed to the teaching at the time the records were made.

**Descriptions in the Teacher/Student Records**

The natural language descriptions of observations taken from the teaching service records were mapped to the format required in the data entry form shown in figure 2. The overall mapping to the billingsMentor vocabulary (table 1) was straightforward. Adjectives and adverbs that qualified a key descriptor in a teaching service description were subsumed in the mapping. Where there was more than one key descriptor per characteristic, the one higher on the drop list was chosen, as instructed on-line.

There were 6,225 observations (38.4% of all observations) described in the teacher/student records in which it would have been possible and relevant for the student to have used both color and fluidity to describe the appearance of a discharge. This is exclusive of observations in which bleeding or signs of blood occurred. There were 3,326 observations in which color was recorded but fluidity was not recorded, 1,033 observations in which fluidity was recorded but color was not recorded and
1,866 observations in which both color and fluidity were recorded.

Three general guidelines were followed in mapping the teacher/student descriptions. They were only relevant when there was no indication of the presence of blood and the quantity of discharge was more than a trace. The colors of “white,” “cloudy” and “yellow” were mapped to the color “opaque.” If the description of

| 01-18 | 01-19 | 01-20 | 01-21 | 01-22 | 01-23 | 01-24 | 01-25 | 01-26 | 01-27 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|       |       |       |       |       |       |       |       |       |       |
| wet   | wet   | wet   | dry   | dry   | wet   | dry   | damp  | damp  | damp  |
| fluid (bled) | fluid (bled) | fluid | trace (little bits of brown ...) | (dry) | SF (wet seminal fluid) | (dry) | opaque thick (thick cloudy once) | opaque thick (cloudy sticky) | opaque thick (cloudy sticky) |
|       |       |       |       |       |       |       |       |       |       |
| 01-28 | 01-29 | 01-30 | 01-31 | 02-01 | 02-02 | 02-03 | 02-04 | 02-05 | 02-06 |
| damp  | damp  | damp  | damp  | dry   | dry   | damp  | damp  | damp  | damp  |
| opaque thick (cloudy sticky) | opaque thick (cloudy sticky) | opaque thick (yellow sticky) | opaque thick (thick cloudy) | opaque thick (thick cloudy ...) | opaque thick (thick cloudy) | opaque thick (thick cloudy) | opaque thick (thick cloudy) | trace (shiny clear not very much ...) |
|       |       |       |       | (dry) | (dry) | (dry) | (dry) | (dry) | (dry) |
| 02-07 | 02-08 | 02-09 | 02-10 | 02-11 | 02-12 | 02-13 | 02-14 | 02-15 | 02-16 |
| wet   | wet   | wet   | dry   | dry   | damp  | dry   | dry   | dry   | dry   |
| trace (shiny clear not very much ...) | opaque thick (thick cloudy shiny) | opaque thick (thick cloudy) | opaque stringy (thick cloudy stringy) | opaque thick (cloudy thick) | opaque thick (cloudy thick sticky) | opaque thick (cloudy thick sticky) | opaque stringy (clear thick stickiness) | clear thick (clearstringy sticky) | (dry) |
|       |       |       |       |       |       |       | (dry) | (dry) | (dry) |
| 02-17 | 02-18 | 02-19 | 02-20 | 02-21 | 02-22 | 02-23 | 02-24 | 02-25 | 02-26 |
| 3 l   | i     | 0     | 0     | 1=    | 2=    | 3=    | 0     | 11    | 2 l   |
| dry   | dry   | dry   | dry   | dry   | dry   | wet   | wet   | wet   | dry   |
| opaque thick (cloudy sticky) | clear fluid (wet ... shiny clearstringy) | opaque stringy (thick cloudy stringy) | trace (tiny smear) | fluid (wet bledd) | fluid (wet bledd) | fluid (wet not as much) | trace (a little red ...) |
|       |       |       |       |       |       |       |       |       |       |
| 02-27 | 02-28 | 03-01 | 03-02 | 03-03 | 03-04 | 03-05 | 03-06 | 03-07 | 03-08 |
| dry   | damp  | dry   | dry   | dry   | dry   | wet   | wet   | wet   | wet   |
| trace (a little red ...) | opaque stringy (thick eggwhite cloudy) | opaque stringy (thick eggwhite shyly stickly) | opaque stringy (shiny little thick eggwhite) | trace (tiny smear) | opaque stringy | clear stringy (clearstringy stickiness) | clear stringy (clearstringy stickiness) | clear stringy (clearstringy stickiness) | clear fluid (shiny clearstringy) |
|       |       |       |       |       |       |       |       |       |       |
| 03-09 | 03-10 | 03-11 | 03-12 | 03-13 | 03-14 | 03-15 | 03-16 | 03-17 | 03-18 |
| O     | O     | O     | O     | O     | O     | O     | O     | O     | O     |
| wet   | wet   | wet   | wet   | wet   | wet   | wet   | wet   | wet   | wet   |
| clear stringy (... clearstringy stickiness) | clear stringy (... clearstringy stickiness) | clear stringy (... clearstringy stickiness) | trace (tiny smear) | trace (tiny smear) | trace (tiny smear) | trace (tiny smear) | opaque thick (cloudy sticky) |
|       |       |       |       |       |       |       |       |       |       |
| 0 O   | O     | O     | O     | O     | O     | O     | O     | O     | O     |
| clear stringy (... clearstringy stickiness) | clear stringy (... clearstringy stickiness) | clear stringy (... clearstringy stickiness) | trace (tiny smear) | trace (tiny smear) | trace (tiny smear) | trace (tiny smear) | opaque thick (cloudy sticky) |

Figure 3. Reconstructed comparative chart for 60 daily records from student A3.
fluidity was omitted then it was described as “thick” if the color was “opaque,” and “thin” (fluid) if the color was “clear,” and vice versa if the description of color was omitted.

Comparative Charts

Figure 3 is a reconstructed chart showing the outcomes of the teaching service for a particular student (ID = A3) juxtaposed with the analysis made by billingsMentor. There are six sets of five rows and ten columns in figure 3 that describe the analyses for 60 consecutive days from the student using standard Billings Method chart symbols. The first row contains an abbreviated date (mm/dd) and the third and fourth rows contain the description of sensation and appearance, respectively. If the description given in the teaching service differed textually from the mapped description in billingsMentor, it is given in parenthesis. The second row represents the fertility analysis made by billingsMentor, encoded using symbols. The fifth row is the analysis made by the student under the guidance of her teacher, also encoded in symbols. Differences in the two analyses will be discussed in the next section.

Evaluation

Two criteria were used to measure the performance of billingsMentor against the outcomes from the teaching service. The first is the number of Peaks recognized correctly. The second is the number of days classified correctly as belonging to the BIP or belonging to a luteal phase. The results are reported for three groups of students. Group A consisted of 19 students who had no recent fertility history events (i.e., not breastfeeding or no recent use of hormonal contraception). Group B was 21 students who had recently given birth and were breastfeeding (nursing) their babies. Group C was 14 students who had recently ceased hormonal medication.

Brown (2010, 2011) noted the correlation between ovulatory and anovulatory events at the Peak and associated subsequent bleeding. If the changing developing pattern that culminates in a Peak day is followed by menstrual bleeding within 16 days (the luteal phase) then

| Table 2. Aggregate number of Peaks and associated luteal phases recognized for students in each group* |
|------------------|------------------|------------------|------------------|
|                  | Peaks recognized | Teacher/student only | billingsMentor only |
|                  | Peaks | Luteal phases | Peaks | Luteal phases | Peaks | Luteal phases |
|                  | L1† | L2‡ | L3§ | L4¶ | L1† | L2‡ | L3§ | L4¶ | L1† | L2‡ | L3§ | L4¶ |
| A                | 64  | 45  | 0   | 11  | 8   | 13  | 11  | 2   | 0   | 0   | 8   | 4   | 0   | 1   |
| B                | 65  | 43  | 4   | 15  | 3   | 12  | 4   | 0   | 6   | 2   | 16  | 8   | 0   | 3   |
| C                | 42  | 34  | 1   | 6   | 1   | 6   | 5   | 0   | 1   | 0   | 14  | 9   | 0   | 3   |
| Total            | 171 | 122 | 5   | 32  | 12  | 31  | 20  | 2   | 7   | 2   | 38  | 21  | 0   | 7   |

*The luteal phase data are classified by the type of luteal phase.
†Normal length (11–16 days).
‡Pregnancy occurred.
§Short length (less than 11 days).
¶Length exceeded 16 days – no ovulation.
this is further evidence that ovulation has occurred near the time of the Peak, but only intrusive monitoring such as ultrasound around the time of the Peak can be totally conclusive. If the bleeding occurs within less than 11 days (short luteal phase) then some function in the ovulatory mechanism was faulty but the Peak day is still significant for the Billings Method user. If bleeding does not occur within 16 days then either there is a pregnancy or there was no ovulation at the time of the suggested Peak (only a rise in the level of the oestrogen hormone) or else there was a fault in the maturation of the follicle. Thus, it is informative to classify the Peaks according to length of the subsequent luteal phase (see table 2).

The prompt and reliable recognition of the BIP (see meta-rules D, E and F) is an important function in the use of the Billings Method. The BIP days may be either infertile or possibly infertile (wait days) e.g., those subject to a count of three days after a change (rule 3, “wait and see”). The BIP is only relevant to days in the pre-ovulatory phase. The number of days classified as BIP can be inversely

Figure 4. Number of Peaks recognized for each student in each group classified as (i) recognized in common, (ii) recognized by teacher/student only, and (iii) recognized by billingsMentor only.
dependent on the number of Peaks that have occurred. Therefore, when evaluating the performance in recognizing infertile days the number of days in the luteal phases has to be taken into account as well, because it is the overall number of days available for intercourse when wanting to avoid a pregnancy that is important, and the luteal phase after the third day is a time of infertility (rule 4).

RESULTS

Recognizing the Peak Day

Figure 4 displays the results for the recognition of Peak days for each student in each group.

There were no Peaks recognized (possible infertility) for five of the students in group A (member IDs 1, 2, 7, 15 and 17) and also for five students in groups B or C (no return of fertility). The aggregate statistics for Peaks recognized and the lengths of the luteal phases are shown in table 2. The Peaks are divided into three classes: those that were recognized in common by the teacher/student and by billingsMentor (171), those that were recognized by the teacher/student only (31), and those that were recognized by billingsMentor only (38). For example, figure 4(A) member ID 9 shows three Peaks recognized in common, one by the teacher/student only and two by billingsMentor only.

Table 2 shows that the majority of luteal phases are of normal length, and most of the remaining cases are short-length luteal phases. The bleeding that ends these types of luteal phase (types L1 and L3 in table 2) lends weight to the proposition that the preceding Peak was identified correctly and that there was a completed ovulation in the case of the normal length luteal phases.

There were seven pregnancies in the data recorded (the purpose of the student at the time of the pregnancy is not clear from the records and not relevant here) and these are identified in table 2 (luteal phase type L2). When a woman becomes pregnant the fertility model of the Billings Method is no longer applicable. However, it is helpful to continue to observe symptoms for evidence of implantation, such as appearance of blood, but it is not relevant to identify Peaks even if there are symptoms similar to a developing pattern.

Apart from the Peaks near the time when a student became pregnant there were a larger number of Peaks identified that were not followed by menstrual bleeding within the prescribed time following a completed ovulation (luteal phase type L4 in table 2). Bleeding occurred within a few days past the limit for a normal luteal phase (day 18 onwards) in nine of the twelve cases where the preceding Peaks were identified in common. In the other three cases in common, there were fertile symptoms before any bleeding event but the fertile symptoms occurred after the limit day. This is significant because the circumstances are addressed by meta-rule B when the BIP becomes relevant again.

There were two Peaks among those identified only by the teacher/student in which the duration of a normal luteal phase was exceeded. In one case bleeding began on day 18 past the identified Peak. In the other case fertile symptoms arose on day 14 and another Peak was identified on day 21 and followed by a short luteal phase. There was an error on the part of the teacher/student in this case. On day 3, past the first Peak there was a return to a wet sensation. This circumstance is addressed in meta-rule C but it was not applied by the teacher/student in this case.

There were ten Peaks among those identified only by billingsMentor in which the duration of a normal luteal phase was
exceeded. The action in the case of five Peaks was clear cut; there was no ovulation and it was merely a matter of the system applying meta-rule B. Two of the five were consecutive Peaks 20 days apart. In this case the possibility of a Peak was acknowledged at the time by the teacher but neither Peak was marked by the teacher/student. This situation is sometimes described as a “doubtful Peak.” The other five Peaks arose in the records of two students. The first student was diagnosed with polycystic ovary syndrome (PCOS) and the teacher elected not to identify any Peaks because there were recurring Peak symptoms. These conflicting symptoms caused billingsMentor to identify a Peak and later advise that the Peak should be rejected (by using the manual system override) and then a subsequent Peak was identified followed by a normal luteal phase. The message dialogue for the second student indicated that she was not able to develop a discipline for

![Diagram](attachment:image.png)

**Figure 5.** Cases of students where disputed Peak(s) occurred, i.e., one or more Peaks were recognized by teacher/student only and/or by billingsMentor only. The chart indicates the identity of the first occurrence of a disputed Peak and the number of Peaks recognized in common that preceded the disputed Peak. There were no disputed peaks for some students.
observing her symptoms, nor was she able to discriminate between a wet sensation and a slippery sensation. This resulted in consecutive Peaks similar to those described for the previous student. It is likely there was a faulty ovulation mechanism in the case of the first student, but in the second case it is possible that the billingsMentor interpretations were misleading because of the incorrect descriptions by the student.

In the case of another student, the teacher expressed surprise and suggested that the student had PCOS when a Peak that was recognized in common was not followed by bleeding. Subsequently, billingsMentor generated two consecutive Peaks (not recognized by the teacher/student) in a similar management pattern to the PCOS student above. The second Peak was followed by a short luteal phase.

There were three reasons that billingsMentor did not recognize 31 Peaks that the teacher/student recorded (the records of eleven students were involved). There were 15 Peaks that were not followed by a sustained definite change as specified in meta-rule C and as implemented in billingsMentor (i.e., meta-rule C was not applied by the teacher/student or there was a stringy discharge in the symptoms recorded on the second and/or third day after the Peak for which the action in meta-rule C is also applied by billingsMentor). For 10 Peaks there was no slippery/lubricative sensation in the developing pattern leading to the Peak, in violation of meta-rule A. For six Peaks the developing pattern did not satisfy the billingsMentor criteria that determine if there has been a sufficient developing pattern preceding the Peak.

BillingsMentor recognized 38 Peaks in addition to the 171 Peaks recognized in common with the teacher/student. Apart from the ten already discussed, the remainder were all followed by bleeding consistent with a normal luteal phase (21) or a short luteal phase (7). The apparent errors by the teacher/student in these cases may have been due to incorrect student records that were not revised by the teacher, or to some “doubtful Peaks,” or that the students were not confident enough to make a decision in the early stages of their education (see figure 5).

There were disputed Peaks (i.e., the 31 Peaks recognized only by the teacher/student and the 38 Peaks recognized only by billingsMentor) in the case of 32 students. Figure 5 shows whether the first occurrence of a disputed Peak for a student was one recognized by the teacher/student or by billingsMentor, and how many Peaks were recognized in common before the occurrence of the first disputed Peak. The teacher/student recognized the first disputed Peak in 11 cases and billingsMentor in 21 cases. The average number of common Peaks preceding the teacher/student cases was 2.3 and the average preceding the billingsMentor cases was 0.6. These data suggest that the Peaks recognized by teacher/student only, tend to occur after a number of Peaks have been recognized in common and/or by billingsMentor only, whereas the Peaks recognized by billingsMentor only, tend to occur early on. This will be discussed in the final section.

Other Classified Days (BIP and Luteal Phase)

The classification of days in the cycles as belonging to a luteal phase or being part of a BIP is summarized for each student in figure 6 and the aggregates for each group are given in table 3. Data are given for the teacher/student and for the interpretations done by billingsMentor when applying the meta-rules D, E, and F. The teacher/student records often have
a greater number of non-dry BIP days recognized when compared to the billingsMentor emulation. There were exceptional circumstances with no Peaks for three students for whom billingsMentor failed to recognize a BIP, but the teacher/student did:

- Client A7 was aged 47 years old and had many periods of bleeding;
- Client A15 had recently given birth but was not breastfeeding, English was her second language and she used very long descriptions of her discharges with possibly unintended variations; she was

Figure 6. Number of classified days (luteal phase, dry BIP and non-dry BIP) for each student charted as stacked columns. There are two columns for each student. The left column is the classification done by the teacher/student, the right column is the interpretation of the daily sequences done by billingsMentor.
unconvinced about the efficacy of the Billings Method;
• Client A17 was acting as surrogate and taking fertility medication; her sequence of daily symptoms was interspersed with many sequences of stringy discharge.

Humans have amazing pattern recognition abilities, so one can expect that billingsMentor will always fall short of human expert performance in recognizing a non-dry BIP. For example, in figure 3 the teacher/student had classified components of a combined BIP by day 02–11, whereas billingsMentor was later. However, there are some significant reasons for the apparent improved performance by the teacher/student in the case of 16 other students. In six cases a non-dry BIP, or a combined BIP, was declared at the beginning of charting. This may have been based on previous information, or there was a questionable student decision, but we did not copy this declaration in the billingsMentor emulation. In these cases the lesser performance of billingsMentor was not just due to the normal delay in recognizing a BIP, but in some cases the immediate sequence of symptoms was not sufficient to support the recognition of the declared BIP. There were five cases in which meta-rule G was ignored by the teacher/student after a return of fertility and a non-dry BIP was declared immediately. BillingsMentor would take three cycles, or a two-week interval after a long cycle had been identified, before identifying a BIP. There were two cases where a student began charting again after a pregnancy and the teacher/student declared a non-dry BIP immediately while billingsMentor followed the procedures specified in meta-rule F. There were three cases in which the non-dry BIP was changed in two consecutive cycles by the teacher/student, contravening meta-rule E. In a few cases where a combined BIP was apparent in the teacher/student chart, it was not clear that the teacher and student were in agreement with the different symptoms that were marked as belonging to the BIP. Other reasons for the lesser performance by billingsMentor were inconsistent use of language by the student, and gaps in the teacher/student records that caused billingsMentor to reassess the BIP.

The aggregate number of BIP days recognized by the teacher/student (5,520) exceeds that for billingsMentor (4,507) by 953 (table 3). The circumstances of the 16 students described above accounts for well over 300 days of non-dry BIP in the difference. Taking this into account billingsMentor classified 7,058 days or 93 percent of a modified total count of 7,602 days for the teacher/student (table 3).

There were 173 daily observations involving records from 19 students that

| Group | Classified days |
|-------|-----------------|
|       | Teacher/student | billingsMentor |
|       | Luteal | BIP | Total | Luteal | BIP | Total |
| A     | 954    | 1,781 | 2,735 | 884    | 1,472 | 2,356 |
| B     | 876    | 2,161 | 3,037 | 951    | 1,736 | 2,687 |
| C     | 578    | 1,578 | 2,156 | 656    | 1,359 | 2,015 |
| Total | 2,408  | 5,520 | 7,928 | 2,491  | 4,567 | 7,058 |

Table 3. Aggregate lengths of luteal phases and number of days classified as BIP for each group
billingsMentor classified as belonging to a non-dry BIP when the teacher/student records did not. The main reason is that billingsMentor has a limited vocabulary for describing symptoms (see table 1) so that some descriptions that were equated under billingsMentor led to a change being detected by the teacher/student. The majority of these cases arose because the vocabulary has only three colours for a discharge, “opaque,” “blood-coloured” and “clear.” Thus “white,” “yellow” and “cloudy” are equated to “opaque,” so that when these colors occur sequentially a change is only detected by billingsMentor if there is a change in one of the other characteristics (quantity, blood, fluidity or sensation). Similarly, there were a number of fluidity descriptions that equated to “thick.” The other cases arose when there was only a small quantity of discharge (trace, smear, etc.); then billingsMentor does not rely on any color (other than blood) or fluidity that might be recorded. These differences are illustrated on days 02–21, 03–02 and 03–04 in figure 3. There was always good contextual justification for the classification of these daily symptoms by billingsMentor.

**DISCUSSION**

The Billings Method chart is the primary response of either system to student/client input. The main difference is that in the traditional teacher/student environment, the student must first interpret and encode her fertility by choosing a sticker or symbol for each day. There is a range of 19 different stickers and 13 different symbols (Billings and Billings, 1997). The chart is then generated and displayed (see figures 1 and 3). BillingsMentor interprets and encodes fertility on behalf of the client by analyzing the sequence of symptoms and then generating a chart according to the rules and meta-rules. The option most commonly chosen for chart visualization by users of billingsMentor is the bar chart, which is unique to billingsMentor and as noted above it is intuitive and has more information content. For example, the same Billings Method sticker/symbol is used to denote a wait day after intercourse (rule 2) and a possibly fertile day (rule 3), whereas there is a clear distinction in the height of the bar chart.

There were numerous errors in sticker selection by students over 300 days, varying from no errors to a maximum of 43 by a single student, with a median of 5 errors. Any apparent errors in the records for the first four weeks were ignored in these statistics because introductory teaching uses a restricted set of stickers. The subsequent errors mostly arose in the encoding of days that match the BIP and particularly any “wait days” (see rule 3) that must include the day count (1, 2 and 3). There was also a lot of confusion among students in the selection of stickers that denote spots of blood or blood color, not included above. The billingsMentor functions that generate a chart have been thoroughly tested and always generate a chart that is consistent with its analysis. This degree of quality control is an important advance.

Depending on her level of understanding, the student may rely on the teacher to correct the chart by changing the selected stickers. She may also communicate (usually via messages) with the teacher for advice and explanation. The degree of satisfaction with this type of service is dependent on the teacher’s availability and preparedness to respond to messages promptly. Future systems could use real-time video channels such as Skype, but there will always be a trade-off between minimizing cost and maximizing service.

There were no service standards for the teaching system reported here. The
teachers were highly qualified but, except for one, all their experience was in clinical mode, dealing with students face to face and by appointment. There was considerable variation in the frequency of interaction with students, from fixed intervals similar to clinical appointments to monitoring daily and interacting as required. There was also considerable variation in the way teachers dealt with the educational task over the internet media. Some teachers would concentrate on tutoring and leave the student to make any corrections to the chart while others insisted on a correct chart to the point of correcting the chart for the student.

The data raise a question about meta-rule A for recognizing a Peak day and the revision announced in 2012 (Ovulation Method Research & Reference Centre of Australia 2012). Referring to the day following the last day of slippery sensation, the Ovulation Method Research and Reference Centre of Australia (OMR&RCA) stated: “If she says … it was indeed wet, then we would be questioning the Peak” and “One of the problems we found when people were teaching ‘no longer wet or slippery’ was that women were waiting until dry before the count. This often meant that there was a delay of a few days after actual Peak before the Peak Rule was applied, leaving very little time for intercourse.” Of the 171 Peaks recognized in common, 33 Peaks (19%) involved a wet sensation being recorded after the last day of slippery sensation. In view of the reason expressed by OMR&RCA for the revision to meta-rule A, it seems preferable in the future for billingsMentor to implement the old meta-rule with a warning of the slight risk involved for the client who is wanting to avoid a pregnancy, and only if fertile symptoms arise subsequently, or bleeding does not occur within 16 days, should there be an automatic revoking of the Peak and a reversion to rules 2 and 3. (Currently billingsMentor follows the direction given by OMR&RCA, queries the user about the wet sensation and does not recognize a Peak unless the user changes her description.)

The events from more than 16,000 daily records from 54 students of the Billings Method have been emulated. It is our experience that most women who seek help with NFP are not experiencing regular normal length cycles with the predictability that comes in that type of circumstance. The students of the teaching service in this experiment consisted of 39 percent who were nursing a baby, 26 percent who had recently ceased taking contraceptive medication, and of the remainder many had irregular fertility. Therefore, while the input data may not be representative of a wider population of healthy fertile women, it did provide a more rigorous test of billingsMentor than would have been the case with women experiencing regular normal length cycles. The experiment confirmed the efficiency of billingsMentor in interpreting fertility symptoms from clients according to the Billings Method, and did not reveal any shortcomings from the use of the restricted vocabulary and reporting discipline. We have been able to explain all the cases in which there was not full agreement between billingsMentor and the outcomes for the teacher/student, so we did not assess the results with a statistical method.

BillingsMentor performed as well, if not better, than teacher/students in recognizing Peak days. The teacher/student combination classified more days as being of non-dry BIP than billingsMentor. The difference shown in table 3 is not a fair indication of the relative efficiency of billingsMentor as explained in the previous section, because billingsMentor applies the rules of the Billings Method with equal
precision in every circumstance. In this respect the indications in figure 5 for disputed Peaks need further investigation. One explanation for some of the disputed Peaks recognized by billingsMentor is that in recognizing a Peak the teacher/student started out with caution but there was learning over time both for the student and the teacher, the latter seeing the patterns of the particular student in the light of experience, be that appropriate or not. BillingsMentor does not learn about an individual's fertility patterns in the sense of an artificial intelligence system so it always make the same decision in the same immediate circumstances, thus providing a consistent explanation for its recognition of a disputed Peak in the first weeks that were emulated. An explanation for disputed Peaks recognized by the teacher/student is that the student was acting somewhat independently and some errors were not corrected by the teacher.

There is no record of when a decision was made by the teacher/student to mark a Peak. The convention is that a Peak should be recognized at the end of the following day, when the definite change from a slippery sensation can be confirmed. In the early learning stages it is clear that teachers were coaching students to make decisions and that there were delays in making decisions. This may account to some degree for the trend shown in figure 5 for billingsMentor to be more likely to recognize a Peak in the first weeks of tuition.

There are two noticeable shortcomings in the descriptions recorded in the other systems compared with those in billingsMentor, version 2: (1) clients are less inclined to report traces of discharge; (2) teacher/students described the color of a discharge more often than the possibly more important characteristic of fluidity. BillingsMentor offers better guidance than the teaching environment in the following circumstances:

• The discipline enforced in the observation input module ensures that the client always enters a complete description of her symptoms, in contrast to the somewhat lax student recording noted above;

• The client gets immediate guidance about the implications of her current symptoms and her overall fertility status. When using a teaching service, the student may wait days or weeks for the teacher to re Exxhn her chart;

• BillingsMentor generates a chart that is consistent with its interpretations. Even in the best of circumstances the student who is not using billingsMentor can be left with a chart that contains errors.

Other important features of billingsMentor that were not tested in the experiment reported here are:

• the client can override billingsMentor to declare a new BIP or declare/reject a Peak;

• BillingsMentor inserts a warning symbol on the fertility chart for any day when a client breaks a rule of the Billings Method. It then adjusts the count after a change or the “wait day” after intercourse appropriately.

The results indicate that billingsMentor is a reliable tool for using the Billings Method without the guidance of a teacher. BillingsMentor can also be recommended for use in a teaching clinic that services students on-line. In this mode, the clinician can undertake a less demanding teaching role with only occasional monitoring of the student’s chart to address those cases where further guidance or interpretation would be appropriate. Similarly, there is value for the medical clinician who is treating infertility.

BillingsMentor always gives the same interpretation in the same circumstances whereas the empirical and qualitative
aspects of the Billings Method mean that some meta-rules cannot be precisely specified and there will always be a range of teacher/student interpretations in the marginal circumstances. On the other hand, it is apparent that there is some flexibility among teachers in the application of the well-defined aspects of the meta-rules, as discussed in the results for the non-dry BIP.

**Future**

There is potential for a modest technical project to improve the algorithms underlying billingsMentor and to host them in any written language on personal devices (e.g., iPhone app). The comparable task to make high-quality and comprehensive teaching services widely available is an enormous organisational and economic challenge. An extension to billingsMentor could allow the user to make declarations about her existing fertility (current BIP, whether fertility has returned, and whether the initial day is part of a pre-Peak phase or a luteal phase). BillingsMentor could also be adapted to operate primarily as an interactive learning tool. A more challenging task would be to introduce an artificial intelligence component capable of learning and applying inference and deduction using an individual’s fertility history.

It remains to demonstrate that clients can learn to observe their fertility indicators correctly under the automated guidance from billingsMentor in order to prove that it is complete and sufficient in its functionality as an unguided service. Then, by exploiting information technology the practice of NFP can be reduced to the task of reporting fertility symptoms and acting according to the interpretations given.

In view of the many difficulties that arise from controlling fertility by artificial means, and the harm to fertility that is often experienced as a side effect, and the difficulties with fertility that women experience in stressful lives, it is hoped that systems such as billingsMentor will lead to a cultural change. In particular, by adapting the technology to the personal communications devices of the future it would become a convenient means for women of any language to gain fertility awareness.

**ACKNOWLEDGEMENTS**

The authors acknowledge the contribution of the senior teachers from the World Organisation of the Ovulation Method Billings and the Ovulation Method & Research Centre of Australia who helped make the teaching service possible. The authors acknowledge Dr. Richard Fehring, Marquette University, for his assistance and encouragement to publish this research. The authors acknowledge Dr. Philip Batterham for his advice on statistics, and Dr. Frank Long, Dr Adrian Thomas, and Dr. Paul Burt for their comments on draft versions of the paper.

**SOURCE OF FUNDING**

The authors have had no source of funding.

**REFERENCES**

Betts, K. 1984. The Billings Method of family planning: An assessment. *Studies in Family Planning* 15: 253–66.

Billings, E., and A. Westmore. 2011. *The Billings Method: Using the body’s natural signal of fertility to achieve or avoid pregnancy*. Australia: Penguin Books; first published 1980, new edition 2011.
Billings, E., and J. Billings. 1997. Teaching the Billings Ovulation Method part 2; variations of the cycle and reproductive health. Ovulation Method Research and Reference Centre of Australia.
billingsMentor. 2014. billingsMentor. http://www.billingsmentor.org.
Brown, J. 1997. Ovarian activity and fertility and the Billings Ovulation Method. Ovulation Method Research and Reference Centre of Australia.
Brown, J.B. 2000. Billings Method science. http://www.billingsmethod.org/bom/science/variants.html.
Brown, J.B. 2011. Types of ovarian activity in women and their significance: The continuum (a reinterpretation of early findings). Human Reproduction Update 17: 141–58.
Norris, C. 2013. Why NFP. Linacre Quarterly 80: 218–21.
Odeblad, E. 1994. The discovery of the different types of cervical mucus and the Billings Ovulation Method. Bulletin of the Ovulation Method Research and Reference Centre of Australia 21(3): 3–35.
Odeblad, E. 1997. Cervical mucus and their functions. The Irish Colleges of Physicians and Surgeons 26: 27–32.

Ovulation Method Research & Reference Centre of Australia. Bulletin of the Ovulation Method Research and Reference Centre of Australia 39(3): 12–14.
WOOMB. 2010. Learn the Billings Ovulation Method over the Internet. https://web.archive.org/web/20101222182458/http://woomb.org/bom/teaching/learningBOM.html.

BIOPGRAPHICAL NOTE

Audrey Smith is a natural family planning consultant with continuous teaching experience in the Billings Method spanning forty years. Her email address is audrey.smith@apex.net.au

John L. Smith Ph.D., is a retired engineer with a career in research and development in information systems spanning forty years. His email address is jlsmith@apex.net.au