European Journal of Obstetrics & Gynecology and Reproductive Biology, 47 (1992) 121-127 992 Elsevier Science Publishers B.V. All rights reserved 0028-2243/92/$05.00

The reliability, acceptability and applications of basal body temperature (BBT) records in the diagnosis and treatment of infertility

Antonio R. Martinez, Marcel H.A van Hooff, Erik Schoute, Maartje van der Meer, Frank J.M. Broekmans and Peter G.A. Hompes

Division of Reproduclire Endocrinology and Fertility, Department of Obsterics and Gynecology, Free University Hospital, Amsterdam, The Netherlands

Accepted for publication 13 August 1992

Summary
The possibilities and limitations of basal body temperature (BBT) records as an adjunct in the management of infertility were re-evaluated. To assess its accuracy as an index of ovulation, 172 charts were analyzed by three different physicians. While the average true positive rate was 90%, the false negative rate was only 2%. The remaining graphs (8%) were classified as non-interpretable, probably reflecting measurement problems. Retrospective assessment of 210 biphasic records showed the thermal nadir to occur within I day of the urinary luteinizing hormone (LH) surge in 75% of the cases, and in 90% when 2 days where considered. This confirms BBT as a relatively accurate guide for retrospective identification of the pericivulatory period. Moreover, results of a study conducted to investigate how patients experienced daily recording of BBT graphs suggest that the method is well accepted by a high proportion of women. From all these it appears that there are many indications where BBT graphs can still be applied. Development. of new electronic devices may further improve the reliability, acceptability and applications of the BBT records in the fertility investigation.

Basal body temperature; LH test; Infertility; Ovulation

Introduction

The observation that the basal body temperature (BBT) has a biphasic pattern during the menstrual cycle was first reported by William S.

Correspondence to. Marcel H.A. van Hooff, M.D., Division Reproductive Endocrinology and Fertility, Department Obsterics and Gynecology, Free University Hospital, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands.

Squire in 1868 in the transactions of the London Obstetrical Society [1]. Mary Putnam Jacobi, one of the first American woman physicians, also noticed the biphasic curve of the basal temperature in women (21, but they both failed to connect this finding to ovulation. A possible relationship between the timing of evulation and the shift in the BBT was first suggested in 1904 by the Dutch gynecologist Theodoor Hendrik van de Velde [3] Later in 1926, he clarified that the corpus luteum was responsible for higher temperatures in the second half of the cycle and that the end of corpus luteum activity caused menstruation. He also clearly stated that the temperature shift was related to ovulation [4). Since that time there has been good evidence for the clinical use of the BBT graph [5-8] and it has been widely applied as a simple, inexpensive method to determine whether ovulation has occurred and to estimate its timing in the menstrual cycle. However, often more has been expected from the BBT graph, both in concurrent and retrospective interpretation, than it can actually provide. As a consequence, its value has been questioned and unccrtainty has emerged with regard to the concrete possibilities and applications of the BBT chart. Moreover, the recent development of new methods and techniques allowing a more accurate prediction and detection of ovulation, has raised questions on the current place of the BBT record in the diagnosis and treatment of infertility problems.

The present study is intended to re-assess the reliability of the method as well as its possibilities and limitations in fertility investigation. In addition, the acceptability by the patients and their personal experience with the performance of BBT charts were evaluated. The applicability of the method and the potential role that novel computerized measuring devices may play in this area is further discussed.

Reliability of BBT records

Detection of ovulation

Basically, a BBT record provides information on the existence of either a biphasic or a monophasic pattern. A biphasic BBT graph is assumed to be indicative of ovulation and reflects the central effect of increasing plasma progesterone levels secreted by a functional corpus luteurn [7]. However, ovulation has been reported to occur in 3% to 20% of monophasic BBTs when based on hormonal or sonographic criteria 18-121. Another study [13] found that when using the occurrence of pregnancy as ovulation criteria, all charts studied (n = 110) showed a biphasic BBT pattern. On the other hand, the chance that biphasic BBTs could be observed during anovulatory cycles is almost nihil [8,121. These data suggest that the method has in fact a high specificity, defined as the ability to identify the true-negatives, whereas the sensitivity, the ability to identify true-positives, has been found variable according to the different studies.

The accuracy in identification of a BBT pattern is mainly affected by factors related to the method itself, and by the interpretation of the observer. Factors related to the method include either improper recording, illness, diet, medication and sleeping pattern alterations, or inability of an adequate level of progesterone to exert its thermogenic effect. Correct interpretation of a record has been commonly associated to the knowledge and experience of the observer.

Most of the published studies evaluating the reliability of BBT graphs as a method of ovulation detection have expressed their results in terms of monophasic or biphasic patterns. However, such a differentiation is in fact not always very clear, particularly when based on subjective criteria 141.

To investigate to which extent the occurrence of poorly defined or erratic patterns could affect the accuracy of BBT readings as an index of ovulation, 172 charts from 54 normally cycling women (mean 3.2, SD ± 1.7, range 1-6 charts per woman) were independently analyzed by three different physicians. They were unaware of the patients medical history and were asked to quailify the records as either biphasic, monophasic or non-interpretable, according to their own interpretation criteria. One hundred and twelve of the cycle's (65%) were spontaneous, whereas 60 cycles (35%) were stimulated with 100 mg clomiphene citrate (CC) (Serophene, Scrono, Geneva, Switzerland) daily from days 3 to 8 of the cycle. All cycles were ovulatory as confirmed by daily vaginal ultrasound monitoring [15]. The three physicians participating in the study worked in our Division of Reproductive Endocrinology and Fertility and had different grades of experience with the interpretation of BBT charts. Observer A was a recently initiated fellow, observer B was an experienced gynecologist, and observer C was an advanced resident in obstetrics and gynecology.

Observers
*** Biphasic * Non-interp. * Monophasic**
Fig. 1. Percentage of BBT graphs (n - 172) qualified as biphasic, non-interpretable or monophasic by three different ob- servers.

The obtained results are shown in Fig. 1. The number of records identified as biphasic were 170 (98.8%) for observer A, 157 (91.3%) for observer B, and 140 (81.4%) for observer C, giving an average true positive rate of 90.5%. There were 2 charts marked as monophasic by observer B, and 5 by observer C, which represents a false negative rate of 1.3% and 2.9%, respectively. The remaining 2 (1.3%), 13 (7.5%) and 27 (15.7%) charts were qualified as non interpretable by the respective observers. There were no substantial differences observed in the interpretation of records from either spontaneous or CC stimulated cycles.

The rate of true-positive findings observed in this study coincides with data in the literature [8,10,12] and reconfirms the discrepancy between investigators concerning the interpretation of BBT charts. However, in this study such discrepancies were found to be mainly based on the number of charts considered as non-interpretable. The observation of these irregular patterns, which may probably reflect methodologic and/or technical problems rather than the occurrence of definite monophasic records, seems to be the major cause of erroneous interpretations and could explain the discrepancies between studies. The strength of the criteria used to define a biphasic pattern appeared to be associated with the experience of the observer. The low percentage (2%) of clear false-negative patterns found in this study contributes to validate the accuracy of BBT graphs as an index of ovulation. Moreover, these findings suggest that the incidence of noninterpretable records could be substantially reduced not only by improving aspects related to the interpretation criteria but following a refinement of the measurement techniqe [14].

Ovulation prediction
Fig. 2. Time relationship between the day of urinary LH surge detection (day 0) and the day of the BBT nadir in sponta- neous (n = 94) and clormiphene citrale (n = 116) stimulated cycles.

The anticipation of ovulation using a concurrent BBT graph has long been used as an aid to conception and for timing of insemination. The two most common suggestions for the BBT graph as a concurrent method of predicting ovulation are a dip (nadir) in the curve, that signals the approach of ovulation, and a subsequent sharp rise between two successive days indicating ovulation. Although it is now well accepted that the BBT graph is an unreliable marker for the prediction of ovulation [16], it still could be useful as a simple method for retrospective identification of the presumptive day of ovulation.

Among the numerous parameters used to detect the day of ovulation, the identification of the luteinizing hormone (LH) surge appears to be the most reliable indicator of impending ovulation [171. The relationship between the thermal nadir and the LH surge have been stressed by different studies in the past [18-201. In order to asses the accuracy of BBTs in retrospectively predicting the day of the Urinary LH surge, we analyzed 210 records of 88 patients (mean 1.8, SD ± 0.7, range 1-4 cycles per patient) who underwent intrauterine insemination with their husbands' semen. Ninety-four (45%) cycles were spontaneous and 116 cycles (55%) were stimulated with 100 mg CC (Serophene, Serono) daily from days 3 to 8 of the cycle. Monitoring of an LH surge was performed in morning and evening urine samples by means of the LH Color' test (Organon, Oss, The Netherlands), which detects the LH peak on average 16 to 28 hours before ovulation. The characteristics of the test as well as its clinical Teliability for the prediction of the time of ovulation have been previously reported [21]. All cycles included in the study had presented a positive test, presumably reflecting the occurrence of a urinary LH surge above 50 U/l, and were ovulatory as confirmed by daily vaginal ultrasound [15). A reproductive endocrinologist, experienced in the reading of BBTs and having no prior knowledge of patients' results, was asked to review the charts and to predict retrospectively the day of the LB surge by determining the thermal nadir (subjective interpretation). The day of urinary LH Surge detection was considered as day 0, and the number of cycles where the nadir differed around the LH surge were expressed as cumulative percentages for each subsequent day.

The results obtained from this study are shown in Fig. 2. There were no differences in the relationship of thermal nadir to urinary LH surge between spontaneous and CC stimulated cycles. From all cycles, in 66 cases (31.5%) the LH surge occurred on the day of the nadir, in 157 cases (74.8%) when ± I day was considered, and in 193 cases (91.8%) the nadir was found within ±2 days of the LH surge. For the remaining cases the surge fell more than 2 days from the nadir.

These findings closely coincide with other published studies relating the BBT nadir to the day of the LH surge [18-201 and reconfirms the inaccuracy of BBT records for the prediction of ovulation. However, when retrospectively analyzed, in 75% of the cases the thermal nadir was noted within I day of the LB surge and in more than 90% of the cases when 2 days where considered. This indicates that BBT charts still can provide a reasonably accurate guide for retrospective assessment of events related to the periovulatory period.

Acceptability

Despite the extensive use of BBT graphs in clinical practice, and the active participation it demands from the patients, not much attention has been payed to the customers concern with the method in most of the reported studies. The following study was conducted to investigate how women experienced daily recording of HBT graphs and the factors which could affect their qualification of the procedure.

The first 100 patients who, from the start of the study, attended our Division of Reproductive Endocrinology and Fertility, were given a standard questionaire. The answers were returned anonimously and always on a voluntary basis. The first questions were related to demographic data, education, kind of work and reason of consult. Patients were asked for the number of cycles they had kept a BBT chart (=< 3,4-6, => 7) and whether it was related to: (a) the fertility investigation, (b) monitoring of a treatment or, (c) voluntary use as a method of self-observation. To qualify the burden associated to the performance of the method four possible alternatives were offered: (a) very burdensome, (b) burdensome, W little burden, or W no burden. Finally, patients were asked whether they considered to keep BBT records as a part of their treatment and/or investigation as either: (a) a positive experience, (b) a negative experience, or W indifferent.

Eighty-one out of the 100 distributed questionaires were returned. All patients but one had previous experience with keeping BBTs, and these were included in further analysis. Mean age of the patients was 31.4 years (SD ± 4) ranging between 25 and 43 years. Regarding education 3 (4%) had only completed primary school, 40 (50%) had completed high school, and 37 (46%) had received higher education. AS related to the work situation in 2 cases the answers were equivocal, 37 patients (47%) worked more than 20 hours a week, 21 patients (27%) worked 20 hours or less, and the remaining 20 (26%) had no active job.

The reason for consulting was infertility in 74 cases (92.5%), other reasons in 3, and not known in another 3 cases. From the total group 69 patients (96%) were currently keeping a BBT chart or had done so within the last year, whereas 11 (14%) stopped recording more than I year ago. The number of cycles that patients had kept a BBT chart according to the indications is given in Fig. 3. In a majority of the cases (54.5%) the indication was related to the fertility investigation, in 35 cases (29%) it was used for monitoring of a treatment, and in 20 cases (16.5%) it was done just as a method of self-observation. In almost half of the cases BBTs had been kept for more than 6 cycles. Most of the patients (95%) recorded rectal temperature, 2 patients did it vaginally, one axilar and one orally.

Regarding the burden associated with the procedure, it was qualified as very burdensome by 10 patients (12%), burdensome by 23 (29%), as a little burden by 35 (44%) and as no burden by 12 (15%). When asked about their personal experience with the method, 6 patients (7.5%) expressed it was for them a negative experience, 37 patients (46%) considered as positive the experience of keeping the BBTs during their investigation, and for 37 patients (46%) it was indifferent. Factors such as age of the patients, education or the category of infertility did not substantially affect either the degree of burden associated with the procedure or the way the patients experienced the method. However, when considering the patients' occupational situation, it was found that a remarkably higher proportion (73%) of patients working more than 20 hours a week had expressed small or no burden associated with the performance of the method as compared to the groups working 20 hours or less (47.5%) or without an active job (40%), (XI test: $P < 0.05$). When associating the burden of the procedure with the experience of the method (Fig. 4), it was found that while most patients with a negative experience considered the method burdensome (5 out of 6), still 10 (30%) of the patients who regarded the procedure as burdensome expressed to have a positive experience with the method.

These findings suggest that, despite the intensive participation required, the use of BBT records is a well accepted method by a high proportion of our patients population. The active involvement in the investigation and/or treatment of their fertility problems as well as the awareness of their own biological changes are speculative reasons that could account for this. Whether user acceptability might further be improved with the advent of novel and more practical measuring devices still remains to be elucidated.

Applications

Fig. 4. Percentage of patients expressing a negative, indiffer- ent or positive experience with the method as associated to the related burden (B + VB: burdensome to very burdensome; L+NB: little or no burden).

The BBT offers the advantage of being a simple, cost-effective and reliable method for retrospective timing and identification of the occurrence of mutation. Its possibilities and limitations well understood, the BBT may still be used as a valuable auxilliary too] in the diagnosis and treatment of infertility problems. The BBT serves Several functions: (1) gives a rapid and objective impression about the length of the cycle, possible occurrence of ovulation, and the characteristics of the menstrual pattern; (2) indicates frequency and timing of sexual intercourse; (3) provides information for scheduling of post-coital tests and cervical factor evaluation. By this means a great proportion of normal cases can be easily evaluated, while considering the use of more sophisticated techniques only for cases with doubtful or negative results; (4) evidences the occurrence of a short luteal phase (< 11 days) and helps in planning of endometrial biopsies; (5) aids in timing and interpretation of hormonal determinations; and (6) monitors the effect of ovulation -inducing agents.

Besides its applications in the diagnosis and treatment of infertility, BBT records have been extensively used as an important element of family planning practice. For this purpose, new sophisticated digital-electronic BBT recording devices have been developed in an effort to simplify and improve these measurements [22]. Most of them are oral thermometers that utilize microcomputcr-assisted repeated calculations giving information on a liquid crystal display in only a few seconds. These devices document BBT measurements in a built-in computer program- Algorithms are incorporated that indicate the fertile and non-fertile phase of the cycle, whether ovulation has been occurring, and the expected more fertile phase of the current cycle. Some devices also allow the possibility of printing out previously recorded readings. One of the most sophisticated devices available is the Baby Comp (Valley Electronics, Eschenlohe, Germany), which possesses an 8-bit micro-processor, with 128 computational registers, and the possibility to store information of up to 48 cycles (4 years). It is also programmed to recognize operational errors. In case of unusually marked temperature variations or skipped measurements, the device automatically calculates the missing readings and these will not be used in the evaluation. By pushing a button information is incorporated on whether or not the patient is menstruating. The possibility that a conception has occurred is displayed after 10-18 days, and calculation of the delivery date is based on the probable day of conception.

The possibility to modify these devices to process information related to parameters that can be used by the clinicians in the diagnosis and treatment of infertility would still further improve the reliability of BBT measurements, its acceptability by the patients as well as their clinical applications.

Conclusions

The BBT graph has largely been used as an important adjunct in the fertility investigation. The main advantages associated to this method have been its simplicity, noninvasiveness and low cost. In the last years, with the advent of more sophisticated methodologies, its value has been questioned and doubts on its current place have arisen. In the present study the reliability and applicability of BBT records was reevaluated in. the actual context of new fertility management techniques.

The method showed a high sensitivity when used as an index of ovulation. This was in agreement with previous studies. From our investigation, however, it appears to be the occurrence of irregular records, here classified as non-interpretable, the main contribution to erroneous results. Future refinements in the recording and measurement techniques might further improve the accuracy of the BBT graph for the detection of ovulation. The unreliability of BBT records for the prediction of impending ovulation, as compared to rapid urinary LH surge detection, was reconfirmed. In addition, it was demonstrated that it still can provide a reliable guide when used for retrospective identification of the presumptive day of ovulation.

The way patients experienced daily keeping BBT records and their reliance on the method has, to our knowledge, not been previously reported. This pilot study, which was representative of our patient population, clearly suggests that the method is well accepted by a majority of the women. Although 41% found the procedure burdensome, only 7.5% of the patients reported a negative experience with the method.

These findings indicate that there are many indications, in the study and management of fertility problems, where the BBT graph can still be applied. Moreover, it is to be expected that the development of new electronic devices allowing the incorporation of relevant fertility data will further improve the reliability, acceptability and applications of BBT records in the fertility investigation.

References

1. Squire WS. Puerperal temperatures. Trans Obstet Soc, (London) 1868;9:129-144.

2. Jacobi MP. The question of rest for women during men struation, New York: Putnam's Sons, 1877.

3. Van de VeIde TH. Uber den Zusammenhang zwischen Ovarialfunction, Wellenbewegung und Menstrualblutung und uber die Enstenhung des sogenannten Mittelschmerzcs. Haarlem, De Erven F. Bohn, 1904.

4. Van de Velde TH. Die Vollkommene Elie: Eine Studie uber Are Physiologic und Technik. Leipzig, Sttutgart: Benno Konegan,1926.

5. Tompkins P. The use of basal temperature graphs in determining the day of ovulation. JAMA 1944;124-.698704.

6. Marshall J. Thermal changes in the normal menstrual cycle. Br Med J 1963; 1: 102-107.

7. Moghissi KS. Prediction and detection of ovulation. Fertil Steril 1980,34:89-98.

8. Corson SL Ovulation prediction in the treatment of infer tility. I Reprod Med 1986;31:739-741.

9. Jobansson EDB, Larsson-Cohn U, Gernzell CA. Monophasic basal body temperature in ovulalory menstiual cycles. Am 3 Obstet Gynecol 1972; 113:993-999.

10. Moghissi KS. Accuracy of basal body temperature for ovulation detection. Fertil Steril 1976:27:1415-1421.

11. Lenton &0, Weston GA, Cooke ID. Problems in using has I body temperature recordings in an infertility clinic. Br Zed J 1987;1:803-807.

12. Bauman JE: Basal body temperature. Unreliable method of ovulation prediction. Fend Steril 1981;36:729-733.

13. Newill RGD, Katz M. The basal body temperature chart in artificial insemination by donor pregnancy cycles. Fertil Steril 19M;38:431-438.

14. McCarthy JJ Jr, Rockette HE. A comparison of methods to interpret the basal body temperature graph. Fertil Steril 1983;39:640-646.

15. Ritchie WGM. Ultrasound in the evaluation of normal and induced ovulation. Fertil Steril 1985;43:167-175.

16. McCarthy JJ Jr., Rockette HE. Prediction of ovulation with basal body temperature. I Reprod Med 1986;31:742749.

17. World Health Organization Task Force Investigators. Temporal relationships between ovulation and defined changes in the concentration of plasma estradiol l7B-lluteinizing hormone, follicle-stimulating, hormone and progesterone. Am J Obstet Gynecol 1980;138:383-390.

18. Morris NM. Underwood LE, Easterling W Jr. Temporal relationship between body temperature nadir and lulcinizing hormone surge in normal women. Fertil Steril 1976 27.780-783.

19. De Mouzon J, Testart J, Lefevre B, Pouly J-1, Frydman R. Time relationships between basal body temperature and ovulation or plasm@ progestins. Fertil Steril 1984;41:254259.

20. Quagliarello J, Amy M, Inaccuracy of basal body temperature charts in predicting urinary luteinizing hormone surges. Fertil Steril 1986,45:334-337.

21. Martinez AR, Bernardus RE, Kucharska D, Schoemaker J. Urinary Luteinizing hormone testing and prediction of ovulation in spontaneous, clomiphene citrate and human menopausal gonadotropin-stimulated cycles: a clinical evaluation. Acta Endocrinol (Copenh) 1991;124:357-363.

22. Albertson BD, Zinaman MJ. Me prediction of ovulation and monitoring of the fertile period. Adv Contracep 1987;3:263-289.