How Objective is Visual Evaluation of Antepartum and Intrapartum Cardiotocograms?

Herman P. Van Geijn, *Dick K. Donker, and *Arie Hasman

Department of Obstetrics and Gynecology, Free University Hospital, De Boelelaan 1117, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands; and *Department of Medical Informatics and Statistics, University of Limburg, P.B. Box 616, 6200 MD Maastricht, The Netherlands

Decisions to intervene during pregnancy or labor because of fetal distress can be determined by many factors such as maternal hypertension, fetal growth retardation, a decrease in fetal movements, a diminished amount of amniotic fluid, an abnormal blood flow in fetal or umbilical vessels, and an abnormal fetal heart rate pattern. The final decision to actually perform a cesarean section (in pregnancy or labor) or a forceps or ventouse delivery (in labor) is still in the majority of cases based upon the heart rate pattern recorded from the fetus.

Fetal heart rate (FHR) as a sign of fetal well-being was at first recognized around 1820. Until the years 1965 to 1970, intermittent auscultation using a stethoscope was the only method available. Electronic monitoring of FHR and maternal uterine activity (cardiotocography) was then introduced by Caldeyro-Barcia et al. (1), Hon and Quilligan (2), and Hammacher (3) at approximately the same time in South America, the United States, and Europe, respectively.

Continuous monitoring of FHR, in comparison with intermittent auscultation, is now easily performed, provides accurate and continuous information, and results in a substantial document. It is relatively cheap compared with other surveillance techniques. Cardiotocographic monitoring is used in at-risk pregnancies, for example 30 to 60 min several times per week or on a daily basis. During labor it is applied for a longer period and often continuously during the major part of the first and second stage of labor.

Signs of fetal distress include tachycardia (baseline FHR more than 160 beats/min), bradycardia (baseline FHR less than 100 beats/min), diminished or absent variability, absence of accelerations for more than 40 min, and presence of decelerations either spontaneously, during movements of the fetus, or in relation to uterine contractions (4). Particularly variable and late decelerations are associated with a lowering of fetal pH. Increasing duration and depth of decelerations and shortening of pauses between decelerations are among the factors that lead the obstetrician to believe that fetal condition is deteriorating (5).
VALIDITY OF CARDIOTOCOGRAPHY

Although techniques of electronic fetal heart rate monitoring (EFM) have improved in the last decades, the validity of cardiotocography remains a controversial issue. Interobserver agreement in the reading of fetal heart rate tracings never exceeds 64% (6,7). Division of fetal heart rate tracings into segments of baseline, accelerations, and decelerations is a major problem, even for experienced obstetricians (8). Antepartum FHR testing has low sensitivity and a high rate of false positives. Results from EFM with regard to prediction of fetal outcome are comparable to fetal movement counting (9).

The rapid introduction and wide scale application of cardiotocography in obstetric practice was primarily based upon empirical evidence. Randomized trials followed decades later. After a period of overestimated appraisal, the technique is currently under attack. A committee of the Institute of Medicine of the National Academy of Sciences in the United States in 1989 even concluded that “Americans have adopted cardiotocography as standard practice, at considerable added expense to routine obstetrical care, despite the failure of scientific evidence to support its use” (10). Nevertheless, daily obstetric care relies substantially on the application of cardiotocography. Improvement of fetal surveillance became the primary goal of the European Community–sponsored concerted action project “New Methods for Perinatal Surveillance.” One of the subprojects was defined to examine reasons for the low validity of cardiotocography (sensitivity 30–50%; specificity 90–92%) (11). Twenty-two well-known obstetricians from ten European Community countries were invited to participate in a study on classification and interpretation of cardiotocograms in relation to obstetric decision making.

Panel of Referees

The 21 obstetricians were brought into a classroom environment. Actual obstetric cases were presented to this panel of referees. The cases originated from a database of high-risk obstetric patient files collected within the European Community. The referees received extensive data on medical and obstetric history, course of pregnancy, the complete cardiotocographic recordings, and information from additional monitoring of maternal and fetal condition. The data were presented up to a certain point in time, when the referees had to read and interpret the cardiotocographic tracings, assess fetal condition, and decide upon obstetric management.

In total, 13 cases were submitted. Decision points included antepartum situations ($n = 3$) and intrapartum events during the first ($n = 5$) or second stage ($n = 5$) of labor. At each decision point the referees were asked to fill in a questionnaire and to segment and characterize the cardiotocographic tracing preceding the decision point. Classification of FHR phenomena was according to Hon (Table 1). In addition the referees were asked to assess subjectively the cardiotocographic tracing as normal, possibly abnormal, abnormal, or highly abnormal (terminal).
TABLE 1. Classification scheme for fetal heart rhythm features

<table>
<thead>
<tr>
<th>Baseline variability</th>
<th>BI</th>
<th>BN</th>
<th>BR</th>
<th>BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acceleration</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deceleration</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early</td>
<td>DE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>DV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late</td>
<td>DL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>DO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undefined</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The data derived from the segmented and classified cardiotocographic tracings were stored per referee in a MUMPS database, which was converted into so-called classification profiles for each type of fetal heart rate segment (12). The amplitude of a profile signal is equal to the number of referees classifying that particular characteristic of the fetal heart rate at each moment.

CLASSIFICATION AND INTERPRETATION OF CARDIOTOCOGRAPHIC TRACINGS

A summary of the referees' assessment of fetal condition compared with assessment of FHR tracings is presented in Table 2. A "normal" tracing was considered to be associated with a normal fetal condition in most instances. The same but opposite phenomenon was apparent for the traces assessed as "highly abnormal." However, unequivocal opinions appeared to be present in a majority of the cases, although the cases selected for presentation were not collected in a random fashion. In about two-thirds of the decision points, FHR traces were assessed as "possibly abnormal" or "abnormal." In these cases the interpretation of the traces was associated with an ambiguous estimation of fetal condition.

TABLE 2. Assessment of fetal heart rate: cardiotocography (CTG) traces versus assessment of fetal condition

<table>
<thead>
<tr>
<th>CTG</th>
<th>Fetal condition</th>
<th>Good</th>
<th>Possibly endangered</th>
<th>Endangered</th>
<th>Severely endangered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td></td>
<td>53</td>
<td>6</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Possibly abnormal</td>
<td></td>
<td>22</td>
<td>69</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Abnormal</td>
<td></td>
<td></td>
<td>35</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>Highly abnormal</td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>
The major reason for the variation in assessment of fetal condition is a considerable difference in subjective assessment of the cardiotocographic tracings among the referees. Major differences were observed among the various referees in reading and describing exactly the same FHR traces. Examples are given in Figs. 1 and 2, representing the classification of deceleration types and baseline variability by two of
the referees, both having looked at exactly the same traces. Although referee 9 considered about half the baseline segments to have increased variability, referee 10 recognized increased variability in only a small proportion of baseline segments. Whereas referee 9 classified FHR decelerations as either early or variable in nearly half the instances, referee 10 regarded decelerations as variable almost exclusively. This latter referee classified only a very few decelerations as early and none as late.

An overview of each referee’s assessment of baseline variability and deceleration types is presented in Figs. 3 and 4. Obstetricians apparently differ greatly regarding their frame of reference.

FACTORS CAUSING DIFFICULTIES IN READING AND INTERPRETING FHR TRACES

A number of factors may explain the difficulties encountered in reading and interpreting FHR traces. One or more of the following factors may contribute:

1. The FHR pattern is too indirect a measure of fetal condition.
2. The FHR pattern primarily represents fetal central nervous system (CNS) functioning. Fetal circulatory and cardiac functioning is apparent only in cases of a deteriorating fetal condition.
3. Cardiotocography should be considered merely a screening method and should not be applied as a diagnostic tool.
ANTE- AND INTRAPARTUM CARDIOTOCOGRAMS

4. Visual analysis of FHR traces creates, among obstetricians, large differences in assessment of FHR features such as accelerations, decelerations, and variability. Exact criteria are not commonly accepted.

5. Even if criteria have been accepted they are very loosely applied. Visual analysis of FHR traces often lacks discipline.

6. Many factors involve the final FHR pattern (Table 3). The obstetrician reading a FHR tracing should be aware of the complexity of the end result of all these factors.

Interpretation of fetal heart rate patterns requires full knowledge of all clinical factors.

| FIG. 4. Classification of decelerations by the 21 referees.

<table>
<thead>
<tr>
<th>TABLE 3. Influences on fetal heart rate and variability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mother</strong></td>
</tr>
<tr>
<td>Drugs</td>
</tr>
<tr>
<td>Fever, infection</td>
</tr>
<tr>
<td>Hyperthyroidism</td>
</tr>
<tr>
<td>Position</td>
</tr>
<tr>
<td>Convulsions, shock</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
factors such as duration of pregnancy and maternal medication. The obstetrician should have in mind the concept of fetal behavioral states and should be able to conceptualize the actual situation with regard to fetal CNS and circulatory functioning.

IMPROVEMENT OF FHR MONITORING

Classification of FHR Patterns

First of all, adequate baseline recognition is essential. This requires a look at the full length of a tracing, lasting for a minimum of 30 min. “Jogging” of the fetus, hiccups (13), and maternal medication such as beta-mimetics can cause an increase in the baseline heart rate.

Visual analysis of full-length recording prevents variable decelerations from being classified incorrectly as early or late. The flow sheet published by Hon in 1968 (14) in his guidelines, *An Atlas of Fetal Heart Rate Patterns*, stresses the necessity first to observe whether “the FHR pattern reflects the waveform of uterine contraction.” If this is not the case, decelerations are automatically classified as “variable decelerations.” To quote Hon again: “The FHR pattern of variable deceleration is of prime importance in clinical obstetrics. It is the offending deceleration pattern in about 90 percent of the patients who have been clinically diagnosed as having fetal distress.”

The Concept of Fetal Behavioral States

With improving resolution capacity of ultrasound equipment it became more and more attractive to study fetal behavioral states. Since respiration is not continuously present in fetal life and the existence of open or closed eyes is hardly recognizable in the fetus, other criteria for fetal behavioral states were required than are applied to the newborn infant (15). In 1982, Nijhuis and coworkers (16) proposed criteria for fetal behavioral states. Presence or absence of fetal eye and body movements and the heart rate pattern (HRP) were included in their criteria. Four fetal behavioral states were described, IF to 4F, in analogy to the neonatal states 1 to 4. Their definitions of behavioral state are as follows:

*State IF*: Eye and body movements absent, HRP A.
*State 2F*: Eye and (periodic) body movements present, HRP B.
*State 3F*: Eye movements present, body movements absent, HRP C.
*State 4F*: Eye and (continuous) body movements present, HRP D.

The fetal heart rate patterns A to D are defined as follows:

*HRP A (state IF)*: Stable heart rate with a narrow oscillation bandwidth.
HPR B (state 2F): Wider oscillation bandwidth than pattern A, with frequent accelerations in association with movements.

HPR C (state 3F): Stable heart rate with a wider oscillation bandwidth than pattern A and a more regular oscillation frequency than pattern B. There are no accelerations.

HPR D (state 4F): Unstable heart rate pattern, showing large and long-lasting accelerations, often fused into sustained tachycardia.

This behavioral-state concept has been widely accepted. Originally it was claimed that behavioral states could only be identified from 36 weeks of gestation onward. Before 36 weeks of gestation, however, coordination among the cycling patterns of heart rate variation, eye movements, and body movements is already present, suggesting a more or less gradual development of behavioral states (17,18). Fetal behavioral states are also recognizable during active labor (19). Within states, there is a close relationship between movements and heart rate phenomena (20–22). In the healthy fetus, the heart rate pattern is strongly influenced by type, duration, and timing of movements.

Automated Analysis of FHR

From the foregoing it is clear that uniform and disciplined classification of baseline and accelerations, decelerations, and variability appears virtually impossible in daily obstetric practice. It gives rise to a wide variety of opinions on the assessment of fetal condition and related obstetric management.

In recent years, stimulated especially by Professor Dawes in Oxford, successful attempts have been made to produce a system for automated analysis of FHR (23–25). This system has been commercialized and released by Oxford Sonicaid Medical Systems. Currently it is being tested in a randomized trial on conventional monitoring versus computerized FHR monitoring within the European Community concerted action project “New Methods for Perinatal Surveillance” (26). Automated analysis of FHR allows further investigations on the relationships between fetal (patho)physiologic variables (27,28).

CONCLUSIONS

Reading and interpreting FHR patterns requires knowledge about all factors that can influence sympathetic and parasympathetic regulation of fetal heart rhythm such as gestational age, maternal medication and type, and duration and timing of fetal movements. It is necessary to be thoroughly aware of the concept of fetal behavioral states. Further studies are required to understand fully the physiologic and pathophysiologic mechanisms at the basis of fetal heart rhythms. Future prospects include quantitative automated analysis of FHR patterns. Finally, it is necessary to integrate the conclusions on individual FHR patterns into a general concept in which clinical, biochemical, and biophysical data referring to fetal condition should be in harmony.
REFERENCES

DISCUSSION

Dr. Dawes: How accurate are these new methods? How does the phono compare with the direct ECG, for example.

Dr. Van Geijn: We are examining this at the moment in a collaborative investigation between our department of medical physics and the obstetric department, using the phono transducer primarily to detect respiratory movements in the fetus. When we compare ultrasound observations with the phono data it is clear that in many instances breathing movements can be readily detected with the phono system. It must therefore be able to detect fetal body movements. We are currently testing how the phono system handles fetal heart rate, uterine contractions, and fetal body movements.

Dr. Hobel: Fetal assessment is especially important when there is meconium in the amniotic fluid. In the cases you presented to your panel, did any of them have meconium? How did the physicians handle this situation?

Dr. Van Geijn: We gave the panel all the information available, so if there was meconium we informed them. The primary aim was to see how they handled the heart rate patterns and I do not know how they dealt with the meconium problem as an individual variable. However, I fully agree with you that meconium is a clinical feature that should be considered very carefully. The recent study by Rossi, et al. (1) again demonstrated the strong relationship between fresh meconium and certain heart rate patterns, in particular the absence of accelerations, a slight increase in baseline heart rate, and the association with postmaturity.

Dr. Guesry: I am struck by the fact that, since your study was retrospective, you knew the actual outcome. You knew what should have been a good answer to the cardiotocographic and other findings. Have you compared the panel’s answer to the ideal answer? Was the response of the clinicians in charge of these particular cases better than average? Maybe the actual decision was better than it could have been from a panel of people who were divorced from the clinical reality.

Dr. Van Geijn: You are correct that the panel did not know about the outcome. Having studied the cases I am convinced that the clinicians in charge of the cases did no better than the panel would have done. All the panel members agreed that the cases were presented in an appropriate way and close to the real clinical situation. It was after all not much different from the usual day-to-day situation when a resident or colleague comes with a case and presents you, as a staff member, with the clinical data.

What the study proved was that even obstetricians who are experts in fetal heart monitoring and surveillance have the same difficulties as average clinicians in a district hospital. Many other studies over the past 5 or 10 years have shown the same results. What we were especially interested in was why cardiotocography does not fulfill clinical expectations. I suspect there are many reasons for this. One is that heart rate is too indirect a parameter, relating primarily to the function of the central nervous system and having little bearing on cardiac metabolic
function. Only when the fetal condition deteriorates does it reflect cardiac function to some extent. Second, we are badly educated on the meaning and definition of heart rate patterns. There are at least 20 definitions and nobody sticks to any of them. Obstetricians are very undisciplined in looking at heart rate patterns. Nobody really does it in the way I presented: reading, description, interpretation. And I think that many obstetricians are not aware of the wide variation in normal fetal heart rate patterns, many of which are erroneously considered abnormal.

**Dr. Jouppila:** Our problem with cardiotocography (CTG) is that we have so many false-positive results. Using pH determinations we can markedly decrease the number of false-positive CTG interpretations, and I believe these should be done on a regular basis, particularly in university departments where there are many young doctors in training.

**Dr. Van Geijn:** In the study by Jongsma and Eskes (2) it was shown that the sensitivity of CTG was about 30%. When fetal blood sampling was used in addition, the sensitivity increased to nearly 50%. I agree with you. We should encourage fetal blood sampling in labor.

**Dr. Hope:** I think the results of your study were determined by your selection of cases rather than by the accuracy of the technique.

**Dr. Van Geijn:** I agree in general that selection can be an important factor, but I am strongly convinced that the cases used were typical of those that occur daily in normal clinical practice.

**Dr. Dawes:** A recent study by John Patrick and his colleagues in London, Ontario, is relevant. John Patrick died earlier this year and I don’t know when the studies will be published. They were covering 100 records circulated on three occasions, mixed-up, to the same five observers, all of whom were expert in CTG analysis. The research team used the kappa statistic to assess inter- and intraobserver reliability. They came to the same general conclusion as Prof. van Geijn, that evaluation of heart rate variability was statistically not different from random, and assessment of baseline heart rate was hardly any better. When there was more than one deceleration on a record, there was a 40% variation in interpretation between individuals or by the same individual a month or two later. These are astonishing inconsistencies in visual interpretation of CTG records.

**REFERENCES**
