REVIEW

Standardization of fetal heart rate pattern management: Is international consensus possible?

Julian T. Parer

Division of Maternal Fetal Medicine, Department of Obstetrics, Gynecology and Reproductive Sciences, University of California, USA

Agreement about the terminology and descriptions of fetal heart rate (FHR) patterns (nomenclature) is now well established, largely based on the report of the National Institute of Child Health and Human Development (NICHD) workshop of 1997, but consensus on FHR interpretation and management has been extraordinarily difficult to achieve in US obstetrics. Interpretation deals with the significance for the fetus in terms of risk of potentially damaging metabolic acidemia. It is also now understood that part of this interpretation is recognizing or projecting the probability of a pattern of lower risk of acidemia evolving into one with a higher risk so that timely intervention can occur. Management means how the obstetrical team actually responds to a FHR pattern to minimize fetal metabolic acidemia without excessive operative or other interventions.

Many professional bodies and individuals, particularly overseas, have classified FHR patterns and recommended management approaches. For various reasons none of these guidelines has achieved widespread international adoption.

Evidence is accumulating that a 5-tier system does relate to degrees of acidemia and fetal damage and, if appropriately rule based, can improve consistency in interpretation among providers. There is also emerging evidence that if taught and accepted hospital-wide, such an approach can reduce newborn metabolic acidemia without increased obstetrical intervention.

An obvious solution is for professional associations to set up a framework that conforms to the currently available data (admittedly limited), which can be tested for effectiveness. The Japan Society of Obstetrics and Gynecology has done this with 5 tiers on a national level and is expecting validation (or the opposite) to emerge from subsequent studies.

A. Previous attempts to standardize fetal heart rate pattern management

Continuous electronic fetal heart rate (FHR) monitoring was developed in the 1960s, and by the 1970s had spawned a number of companies who marketed devices which were easily adapted to clinical usage, at least in North America. Pattern descriptions (i.e. nomenclature) were rapidly accepted in the USA, due to the publications of Hon and colleagues, although other nomenclature was in use in Uruguay with the work of Caldeyro Barcia and his co-workers, and in Europe.

The significance of the various patterns from the point of view of fetal health was relatively desultory with much emphasis on the reputation of various patterns rather than their stricter relationship to fetal hypoxia, or its measured surrogate, fetal or umbilical arterial acidemia. Thus, the development of suggested management approaches to avoid fetal acidemia during the actual process of labor was substantially delayed.

Attempts to standardize such approaches to management on an international level began with the International Federation of Societies of Gynecology and Obstetrics (FIGO). Though this approach was widely accepted, it lacked specific recommendations for action during labor to avoid newborn acidemia, without paying too high a price of obstetrical intervention. It had little impact on clinical practice in the USA.

In the mid-1990s, the National Institute of Child Health and Human Development (NICHD) sponsored a group of
FHR Pattern Management

18 clinicians who convened with the hope of standardizing FHR nomenclature, and recommending management practices.\textsuperscript{4} The nomenclature was defined quite rigidly, with a view to accuracy in visual interpretation, but also with the expectation that we were on the brink of using computers to more accurately define and read the FHR patterns. Thus the definitions were quantitative and generally unambiguous for this hoped for application, which did not occur. A further disappointment of the FHR group was that it could not achieve consensus on management approaches, except at extremes, those being the “normal” tracings (normal variability [amplitude range 6–25 bpm]), normal baseline [110–160 bpm] and absence of decelerations), and “abnormal” tracings, with absent variability, and decelerations or a bradycardia. The normal trace signified a non-hypoxic, non-acidemic baby, and the abnormal trace as defined above signified a fetus with a high risk of being hypoxic and acidic, to the point of death or subsequent neurologic damage. These extremes of FHR patterns would be called Category I and Category III in the subsequent NICHD conference of 2008.\textsuperscript{5}

There was more activity internationally as the Royal College of Obstetricians & Gynaecologists (RCOG) presented a thorough review of the literature and a 3-tier system of management.\textsuperscript{6} This was followed by the publications of the Royal Canadian College of Obstetricians and Gynaecologists, and the Royal Australian and New Zealand College of Obstetricians and Gynaecologists. These bodies generally espoused a 3-tier system of categorization of FHR patterns, with the first generally being called normal, and a 3rd tier with some semantic variation on the term “abnormal”. Then there is described an intermediate pattern, once again with various nomenclatures, such as “suspicious”.

In 2007, a 5-tier system was published.\textsuperscript{7} This in broad terms included the normal and abnormal tracings, but described 3 gradations within the intermediate category of risk of fetal acidemia, or risk of evolution to a more serious pattern.

The NICHD reconvened in 2008, and reiterated most of the findings from 1997,\textsuperscript{5} and added a 3-category classification system of FHR patterns, similar in many ways to previous publications. It was rapidly noted by numerous investigators that Category II, termed the “indeterminate” category, generally included more than 80% of all babies during the birth process. It was described by some critics as too heterogeneous for use in labor management\textsuperscript{8} and poor at discriminating degrees of risk of newborn acidemia.\textsuperscript{9} Nevertheless, because of the perceived authority of the NICHD, it was widely accepted despite its limited utility.

The American College of Obstetricians & Gynecologists put out a Practice Bulletin in 2010\textsuperscript{10} (in fact they had 3 publications on FHR since 2005) accepting the 3 categories, and appeared to rationalize Category II by describing 6 groups within it, primarily based on different types of FHR patterns.

Recognizing the lack of utility of the NICHD Category II for actual labor management, Clark et al.\textsuperscript{11} suggested a further algorithm based on presence or absence of normal FHR variability, labor progress and stage of labor, and used certain duration-of-pattern recommendations for when to intervene with delivery. This algorithm was accompanied by numerous qualifiers and is currently undergoing study. An attempt by us to reconcile it with the 5-tier system failed.

### B. Development of the 5-tier framework

A framework for standardizing management of FHR patterns was developed based on the risk of acidemia, and a further novel approach, a projected risk of a recurrent FHR pattern evolving into one of higher risk of acidemia.\textsuperscript{7} Five gradations of risk were recognized (Table 1).

<table>
<thead>
<tr>
<th>Five Gradations of Acidemia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green No acidemia.</td>
</tr>
<tr>
<td>Blue No central fetal acidemia (adequate oxygen).</td>
</tr>
<tr>
<td>Yellow No central fetal acidemia, but FHR pattern suggests intermittent reductions in O\textsubscript{2} which may result in fetal O\textsubscript{2} debt.</td>
</tr>
<tr>
<td>Orange Fetus potentially on verge of decompensation.</td>
</tr>
<tr>
<td>Red Evidence of actual or impending damaging fetal asphyxia.</td>
</tr>
</tbody>
</table>

The FHR framework categorizes 134 different possible combinations of baseline variability, baseline heart rate, and graded decelerations into 5 color-coded levels. A summary of the combinations of features defining the various colors is shown in Figure 1. Findings within normal limits for all features were coded “green”. Progressively abnormal combinations of features result in classifications of “blue”, “yellow”, “orange”, and “red” for each 10-minute segment of the FHR tracing, with the most severe pattern encountered used for classification. Each FHR level has corresponding management approaches, directed towards an operative intervention if required (Figure 2). For example, no intervention is recommended for “green”. For “blue”, nurses inform the clinician of the FHR patterns and attempt
amelioration through such conservative therapeutic maneuvers as position change, hyperoxia, correction of hypotension, assuring adequate intravascular volume and avoidance of excessive contractions (Table 2). For “yellow”, the clinician is required to be at the bedside for evaluation, confirmation of amelioration techniques, and increased surveillance. For “orange”, preparation for urgent delivery is started by moving the patient to the operating room, checking the availability of personnel for anesthesia and newborn resuscitation, and continuing amelioration techniques. For persistent “red” patterns, urgent delivery is performed by cesarean section or operative vaginal delivery.

**Figure 1.** Risk categories for fetal acidemia related to FHR variability, baseline rate and presence of recurrent decelerations.

*Mod, Moderate; Sev, Severe; VD, Variable decelerations; LD, Late decelerations; PD, Prolonged decelerations; Brd, Bradycardia; Tachy, Tachycardia; G, Green; B, Blue; Y, Yellow; O, Orange; R, Red.*

**Figure 2.** Proposed management of the color coded categories.
C. Evidence for effectiveness of the 5-tier system

The 5-tier system has been examined in a number of publications with supportive evidence for its ability to predict fetal acidemia. Elliot and coworkers\(^{12}\) used specialized software to categorize nearly 2,500 FHR tracings with diverse outcomes, using the 5-tier system. They found that metabolic acidemia and neurologic morbidity were related to both the degree and duration of tracing abnormality.

The ability of clinicians to agree on interpretation was further demonstrated when five experts in examining over 700 tracings were able to agree nearly 90% of the time within one color code.\(^{13}\) A further bonus of this study demonstrated extremely close agreement of the experts with a computer-read analysis of the tracings.

In a retrospective case control study of twenty-four cases with umbilical cord arterial pH \(< 7.0\), compared with twenty-four cases with a pH \(> 7.2\), the 5-tier system had a superior sensitivity to the 3-tier system.\(^{14}\) In another study, five different management systems were compared (several of them used only in Europe) and the 5-tier system once again was found to be the most discriminatory.\(^{15}\)

In a “before-and-after” study in a small hospital in Japan, the occurrence of metabolic acidemia (pH \(< 7.15\) and base excess \(< -12\) mmol/l) decreased sevenfold following education of the staff in the 5-tier system\(^{16}\) (Figure 3). Further analysis of the FHR patterns showed that there were fewer FHR patterns signifying severe acidemia, so we believed that the clinicians learned to better select those fetuses at risk following training and did not increase their intervention rate of Caesarean or vacuum-assisted deliveries.\(^{17}\)

D. Adoption of 5-tier systems

The most extensive adoption of a 5-tier system of FHR

<table>
<thead>
<tr>
<th>Table 2. Conservative therapeutic interventions to ameliorate FHR patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conservative Therapeutic Interventions to Ameliorate FHR Patterns</td>
</tr>
<tr>
<td>Position Change</td>
</tr>
<tr>
<td>Hyperoxia</td>
</tr>
<tr>
<td>Correct hypotension</td>
</tr>
<tr>
<td>Adequate intravascular volume</td>
</tr>
<tr>
<td>Correct excessive contractions (oxytocin?)</td>
</tr>
<tr>
<td>Avoid constant pushing</td>
</tr>
<tr>
<td>Tocolysis</td>
</tr>
<tr>
<td>Amnioinfusion to correct amniotic fluid deficit</td>
</tr>
</tbody>
</table>

Figure 3. Rate of metabolic acidemia, defined as umbilical arterial cord blood pH and Base Excess (BE) below specific thresholds.

The figure shows the highly significant and sustained decrease following a 6 month period of training the obstetrical staff in the 5-tier management system (reference 16).
management is in Japan, where it is supported nationwide by the Japan Society of Obstetrics and Gynecology (JSOG). This system is a modification of that published in 2007 with less possible categories available, because of the combining of decelerations initially defined as “moderate” into a single category of severe. The authority of the JSOG is such that it is recommended for all labor and delivery units in Japan. Its extensive use by obstetricians, midwives, and nurses in Japan is greatly facilitated by the ready availability of a pocket card summarizing the system. A start has been made on demonstrating the effectiveness of the Japanese system. It was shown in a small patient series that the patterns correlated with umbilical cord acid base state.

The original 5-tier system has been adopted by some specific medical groups within the USA. For example, the Providence Health System in Portland, Oregon uses the algorithm in its hospitals albeit with some modifications. They have an educational package with clarifications of some concepts (Figure 4).

Mainstream Health California, based in the San Francisco Bay Area, has an iPad-based program incorporating the 5-tier color-coded system. Modifications have been made to the original system, to facilitate usage. One such modification is incorporation of clinical information into the classification. For example, in laboring women with preeclampsia, intrauterine growth restriction and certain medical illnesses, the color classification is advanced by 1 step.

Another attempt to simplify the 5-tier color coded algorithm and to reconcile it with the NICHD categories has been presented by King. She has presented a simplified flow chart of the 5 colors and called them Category I, IIa, IIb, IIc, and III. This algorithm does not incorporate the baseline, which is done as a separate interpretative step.

It is obvious that modifications made to the original framework are designed to simplify its apparent complexity, which was an early criticism. Substantial progress towards achieving simplicity came with the development of an app by an Oregon obstetrician, Dr. Melanie Plaut (FHR 5-Tier, iTunes). This cheap and readily available app incorporates all of the NICHD definitions, and an easy to navigate input of FHR characteristics, with automatic presentation of the color code, notes on the risk of fetal acidemia, and clinical management recommendations.

A Spanish-language app has also been developed by Dr. Francisco Guerra, an obstetrician in Southern Chile. This is readily available and incorporates much physiologic educational background (personal communication, 2014).

In the United Kingdom, in response to a belief that the current RCOG recommendations lack incorporation of many details of FHR patterns which are important for management, Ugwumadu presented an expanded classification. This is readily identifiable as similar in many respects to the 5-tier system, as can be seen from the modification of his figure (Figure 5).

In considering the above geographically widespread adoptions of the 5-tier system, it is clear that numerous clinicians find the 3-tier system inadequate for day-to-day management. I believe the reason for this is

<table>
<thead>
<tr>
<th>Category</th>
<th>Definition</th>
<th>Risk of Acidemia at time of evaluation</th>
<th>Risk of Progression of Acidemia</th>
<th>Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green (Primarily Category I)</td>
<td>No acidemia</td>
<td>None</td>
<td>Very low</td>
<td>None</td>
</tr>
<tr>
<td>Blue (Category II)</td>
<td>No central acidemia</td>
<td>None</td>
<td>Low</td>
<td>Conservative interventions</td>
</tr>
<tr>
<td>Yellow (Category II)</td>
<td>Intermittant hypoxia that may lead to acidemia</td>
<td>None</td>
<td>Moderate</td>
<td>Conservative interventions with surveillance</td>
</tr>
<tr>
<td>Orange (Category II)</td>
<td>Potential for fetal decompensation</td>
<td>Low</td>
<td>High</td>
<td>Conservative interventions &amp; preparation for delivery</td>
</tr>
<tr>
<td>Red (Category III)</td>
<td>Possible actual or impending fetal asphyxia</td>
<td>High</td>
<td>Present</td>
<td>Deliver unless FHR abnormality can be immediately resolved</td>
</tr>
</tbody>
</table>

“Risk” is a probability statement about acidemia, not an absolute measure.

Figure 4. A modified educational chart used by Providence Health of Portland, Oregon (courtesy of Dr. Mark Tomlinson).
that FHR pattern interpretation is complex, due to the dynamic changes during labor of the 3 variables of FHR variability, baseline and decelerations, each of which has subclassifications. Denial of this complexity will result in loss of detail, and reduces interpretative accuracy.

**E. Summary of the international status of standardization of FHR pattern management**

It is clear that at present there is no international
consensus in FHR pattern management, although there are certain nations where there is internal agreement.

Generally speaking, all systems agree on the first category ("normal") and the final category ("III", "pathologic", or "red"). This is an improvement on a previous 2-tier system describing "reassuring" vs. "nonreassuring", where there was virtually no agreement on the latter, it being left up to the individual practitioner. The major disagreement is in the middle category of the 3-tier system, and whether it should be expanded.

There is currently an international effort under the auspices of FIGO, which includes scores of experts appointed by FIGO, various national organizations, and individuals appointed because of their contribution to the literature. As of this writing, consensus has been reached on approximately half of the agenda. It is unclear whether this document will achieve international standardization as there appear to be certain impediments.

Firstly, although standard nomenclature and definitions of FHR patterns have been widely accepted in USA since 1997, those in other countries often have different views. The FIGO committee’s definitions are compatible with those of the NICHD, but it is unclear whether other countries will follow. Part of the reason for this appears to be the territoriality associated with the subject. That is various national organizations feel that they “own” the subject of FHR interpretation and, when there are national differences that their own approach is the “best” approach. This disagreement may even be the case in different organizations within a nation.

Another reason for these discrepancies is that there is such widespread usage of FHR monitoring in developed countries, it is now virtually impossible to do a randomized controlled trial (RCT). Although lesser quality evidence supports the effectiveness of the technique in avoiding some fetal acidemia, no one approach to interpretation and management has been shown to be superior to any other.

A second major impediment to international standardization is in the different scaling used in FHR monitors. In most cases the vertical scaling adopted is 30 bpm per cm, but major differences occur in horizontal scaling. In the USA and Japan this is 3 cm/min, but in Europe it is mostly 1 cm/min. This results in a bunching up of the signal which exaggerates the degree of variability. As will be seen from the algorithm presented, FHR is the most powerful prognosticator of absence of acidemia. At present, there is no evidence that one paper speed is superior to the other or that a clinician familiar with one paper speed interprets similarly to the practitioner familiar with the other paper speed.

The USA has not yet adopted a standardized system of management incorporating specific clinical actions, although Japan has adopted a 5-tier color coded approach. Although it is unlikely that RCT will be done to demonstrate effectiveness, other study designs, such as case-control, or “before -an-after” trials are being done. It is likely that lack of consensus internationally will continue until one system shows superior effectiveness in reducing newborn acidemia without excessive obstetrical intervention, by sophisticated trials short of an RCT.

Conflict of interest
None.

References


