Continuous electronic fetal heart monitoring versus intermittent auscultation during labor: Would the literature outcomes have the same results if they were interpreted following the NICHHD guidelines?

Mariarosaria Di Tommaso1,A–F, Serena Pinzauti1,C–F, Silvia Bandinelli1,B,C, Chiara Poli1,B,C, Antonio Ragusa2,E,F

1 Department of Health Sciences, University of Florence, Italy
2 San Giovanni Calibita Fatebenefratelli Hospital, Rome, Italy

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Abstract

Background. All guidelines regarding electronic fetal heart monitoring (EFM) before 2008 were designed to avoid more hypoxia than acidosis. In addition, the results of the Cochrane meta-analysis of 2013 do not show a significant improvement in neonatal outcomes using EFM or intermittent auscultation (IA).

Objectives. We retrospectively evaluated the results on delivery outcomes arising from a comparison between EFM and IA during labor of 2 specific and high-quality trials. We hypothesized that revisiting the delivery outcomes through the adoption of the recent National Institute of Child Health and Human Development (NICHHD) guidelines, the reported delivery outcomes would be different.

Material and methods. The study retrospectively evaluated the results on delivery outcomes arising from the comparison between EFM and IA during labor of the “Dublin trial” and “Vintzileos trial” published, respectively, in 1985 and 1993. A translational model was constructed to recalculate these results, applying a correction factor to estimate the number of pathological patterns using the NICHHD guidelines for EFM.

Results. After the reevaluation of the 2 trials using the proposed correction factor, the comparison of the re-calculated cesarean section and operative delivery rates for fetal distress between EFM and IA group were no longer statistically significant, both in the Dublin trial and Vintzileos trial. Even the comparison of the re-calculated incidence of the rate of non-reassuring fetal heart rate (FHR) patterns in the EFM and IA groups has not given any indication of significance for the Vintzileos trial.

Conclusions. Our results lead to reconsidering the results of the Dublin trial and Vintzileos trial in terms of operational rates of births, hypothesizing that these results would have been significantly lower if FHR traces were interpreted using the current NICHHD guidelines, which aim to identify potential acidotic fetuses rather than hypoxic ones.

Key words: fetal heart rate, intermittent auscultation, fetal hypoxia, fetal acidosis, electronic fetal heart monitoring
Introduction

Electronic fetal monitoring (EFM) was introduced into clinical practice in the seventies and since then it has become the most commonly used method for the surveillance of fetal well-being during labor. Since the first EFM classification of Boylan in 1987, the major international scientific societies have developed specific guidelines that, based on objective parameters, classify the EFM into risk categories with the goal of addressing obstetricians in order to act appropriately to reduce neonatal morbidity, and to avoid inappropriate cesarean section and operative delivery.

Most of the recommendations and guidelines focused on avoiding fetal hypoxia, but in 2008 a consensus of the National Institute of Child Health and Human Development (NICHHD), the American College of Obstetricians and Gynecologists (ACOG) and the Society for Maternal Fetal Medicine clearly stated that the aim of EFM was to avoid the birth of acidotic fetuses and that the concept of fetal acidosis should replace the concept of fetal hypoxia. The consensus reviewed and updated the definitions of fetal heart rate (FHR) tracings with the development of new guidelines that were subsequently revisited in 2009 by the ACOG with the definition of specific clinical management for each FHR category. This interpretation of EFM tracings clearly highlighted that FHR deceleration represents a reversible sign of fetal hypoxia in utero in terms of a defense mechanism of the fetus towards the hypoxic stimulus, while the alteration of the FHR variability represents the highest risk factor for fetal acidosis.

From the introduction of EFM, several randomized trials have also compared its efficacy to intermittent auscultation (IA) of FHR during labor, suggesting that EFM reduces perinatal mortality, but it may increase the cesarean delivery rate without performing an assessment of fetal acid-base status. The synthesis of all these reports was summed up in 2013 by a Cochrane review, which reported that EFM during labor is associated with a reduction in neonatal seizures, but with no significant differences in cerebral palsy, infant mortality or other standard measures of neonatal well-being; this review also confirmed that EFM was associated with an increase in cesarean and operative delivery. However, this Cochrane review also showed that many of the epidemiological studies included had limitations and only a few studies were randomized and based on a larger population. In particular, only 2 of the 13 clinical trials included in the review were considered to be of high quality: the Dublin trial published in 1985 and the trial published in 1993 by Vintzileos et al.

Based on these premises, the present study aims to retrospectively evaluate the 2 specific and high-quality trials that have contributed to the Cochrane meta-analysis – the Dublin and Vintzileos trials. Because it was not possible for us to re-evaluate all the single tracings, we hypothesized that, after revisiting the operative delivery and cesarean section rates, through the adoption of NICHHD guidelines throughout the construction of a translational model, these outcomes will be different in the light of recent guidelines.

Table 1. Characteristics of the Dublin trial and Vintzileos trial

| Characteristics                  | Dublin trial | Vintzileos trial |
|----------------------------------|--------------|------------------|
| Study population, n              | 12,964       | 1,428            |
| EFM                              | 6,530        | 746              |
| IA                               | 6,554        | 682              |
| Inclusion criteria               | singleton    | singleton        |
| gestation age ≥28 weeks          | gestation age ≥26 weeks |
| active labor                     | spontaneous labor |
| No meconium-stained amniotic fluid |               |
| Exclusion criteria               | fetal malformation | congenital or fetal chromosomal defects |
| Setting                          | National Maternity Hospital, Dublin, Ireland | Alexandra Hospital, Athens, Greece |
| Marika Iliadi Hospital, Athens, Greece |          |
| Type of midwifery assistance     | one to one   | one to one       |
| Type of EFM interpretation       | Boylan Classification (1987) | ACOG Classification (1989) |
| Type of IA                       | every 15 min during the I stage of labor | every 15 min during the I stage of labor |
| every interval between contraction during the II stage of labor | every 5 min during the II stage of labor |
| 1 min auscultation               | 1 min auscultation |
| Support techniques to EFM        | fetal Ph from scalp blood sampling | none |
| Measures outcome                 | intrapartum death | neonatal complications |
|                                  | neonatal death | intrapartum fetal death |
|                                  | neurological sequelae | neonatal and perinatal death |
|                                  | other neurological problems | |

EFM – electronic fetal monitoring; IA – intermittent auscultation.
Material and methods

The study retrospectively evaluated the results on delivery outcomes arising from the comparison between EFM and IA during labor of the “Dublin trial” and “Vintzileos trial” published, respectively, in 1985 and 1993.11,12 The characteristics of the trials are summarized in Table 1.

The Dublin trial used the specific criteria for EFM classification that was analogous to the Boylan classification for EFM further published in 1987.1 Fetal heart rate patterns were considered “suspicious or ominous” in the case of at least one of the following criteria: marked tachycardia or bradycardia, moderate tachycardia or bradycardia with reduced variability, a pattern of late decelerations, moderate and severe variable decelerations, and other confusing patterns which could not be interpreted (Table 2). In the Dublin trial, if one of these patterns lasted for at least 10 min and conservative measures failed, the fetal pH was assessed with scalp blood sampling during the 1st stage of labor, or if delivery had been completed, during the 2nd stage of labor. Childbirth was performed as soon as possible in the case of pH < 7.20. If the pH was between 7.20 and 7.25 and the FHR pattern was not reassuring, delivery was completed as soon as possible; if the pH was greater than 7.25 and the pattern was not reassuring, the pH of the scalp was repeated after 0.5–1 h. Similarly, in the IA group, FHR auscultation was considered not reassuring when FHR >160 bpm or <100 bpm during 3 contractions without response to conservative measures.

The Vintzileos trial applied the ACOG guidelines of 198913 and EFM tracings were considered “non-reassuring” in the presence of at least one of the following criteria: late decelerations, persistent prolonged decelerations, severe variable decelerations, variable decelerations with loss of variability, persistent tachycardia with decreased variability, persistent decreased variability and sinusoidal pattern, early decelerations, late decelerations, and severe variable decelerations with loss of variability (Table 2). The IA was considered non-reassuring when the FHR was persistently lower than 100 bpm during and after uterine contraction or persistently higher than 160 bpm. Both groups were managed similarly in the case of non-reassuring FHR, with the application of conservative measures for at least 20 min and after that a cesarean section was performed.

In order to compare the delivery outcomes of the 2 studies, we structured a summary table for each trial in which the specific results were reported regarding fetal distress (Tables 3,4). In particular, the Dublin trial included, as an indication for operative or cesarean delivery for fetal distress, all the “suspicious FHR patterns” for fetal distress, while the Vintzileos trial referred to “suspected fetal distress” as an indication for operative or cesarean delivery in the presence of “not-reassuring FHR patterns”. Both trials referred to intrapartum EFM anomalies as suspected fetal distress, since the resulting neonatal outcomes were analyzed as distinct parts and were not included in our revision.

In a second instance, we construed a translational model by recalculating these results, applying a correction factor to estimate the number of pathological patterns using the NICHD guidelines for EFM. The NICHD guidelines defined as “category III” the presence of sinusoidal patterns or a low FHR variability associated with at least one of the following criteria: recurrent late decelerations, recurrent variable decelerations and bradycardia (Table 2).2

The correction factor was calculated using as a reference a large randomized prospective study evaluating

---

Table 2. Pathological EFM patterns according to the different classifications

| Boylan Classification 1987 | ACOG classification 1989 | NICHD classification 2008 |
|---------------------------|-------------------------|--------------------------|
| marked tachycardia or bradycardia, moderate tachycardia or bradycardia with reduced variability, pattern of late decelerations, moderate and severe variable decelerations, other confusing patterns which could not be interpreted. | late decelerations, persistent prolonged decelerations, severe variable decelerations – variable decelerations with loss of variability, persistent tachycardia with decreased variability, persistent decreased variability, sinusoidal pattern, early decelerations, late decelerations and severe variable decelerations with loss of variability. | sinusoidal pattern, low FHR variability associated to at least 1 of the following criteria: recurrent late decelerations, recurrent variable decelerations, bradycardia. |

EFM – electronic fetal monitoring. FHR – fetal heart rate.

---

Table 3. Outcomes of the Dublin trial for cesarean section and spontaneous operative delivery for fetal distress (modified from MacDonald et al.)11

| Type of delivery | Dublin trial | p-value |
|------------------|--------------|---------|
| Cesarean section | | |
| Fetal distress  | EFM (n = 6,474) | IA (n = 6,490) |
| | 158 (2.4) | 144 (2.2) |
| | 25 (0.4) | 10 (0.2) |
| Operative spontaneous delivery | | |
| Fetal distress | EFM (n = 6,474) | IA (n = 6,490) |
| | 528 (8.2) | 407 (6.3) |
| | 190 (2.9) | 75 (1.2) |

*p < 0.05; EFM – electronic fetal monitoring; IA – intermittent auscultation.
the relationships among abnormal cardiotocography and umbilical cord blood pH, Apgar score, and meconium-stained amniotic fluid. Steer et al.\textsuperscript{14} analyzed 698 FHR patterns, defining as “abnormal” 289 patterns that were comparable to category III of NICHD. In particular, they excluded from abnormality traces with a baseline between 120 bpm and 160 bpm, a baseline variability of 15 bpm, the presence of accelerations, synchronous decelerations with uterine contraction and an amplitude of less than 40 bpm in the 1\textsuperscript{st} stage of labor. Decelerations during the 2\textsuperscript{nd} stage of labor were always considered normal, as well as a baseline between 100 bpm and 120 bpm with normal variability. Considering the reported rate of abnormal FHR patterns and relating that to the total patterns, from this proportion we calculated that 41.4% (289/698) of the traces were homologated to category III of NICHD.

After this estimation, the number of pathological EFM tracings of the 2 trials that hesitate in a cesarean section or instrumental delivery for fetal distress were recalculated by subtracting 58.6% and approximated with excess.

All data was analyzed using SPSS software (IBM SPSS v. 23; IBM Corp, Armonk, USA). Delivery characteristics were compared in univariate statistical analyses using Pearson’s $\chi^2$ test and Fisher’s exact test. The results were reported as numbers and percentages, as appropriate. P-value <0.05 was considered statistically significant.

### Results

Evaluating the results of the Dublin trial, the number of cases of fetal distress that constituted the indication for cesarean section in the EFM group decreases from 25 (0.4%) to 10.3 (0.1%); similarly, the number of operative deliveries for fetal distress in the EFM group is reduced from 190 (2.9%) to 78.6 (1.2%) (Table 5).

The reevaluation of the Vintzileos trial showed that the rate of obstetrical intervention for fetal distress, both cesarean section and operative delivery, in the EFM group decreases from 84 (11.2%) to 34.8 (4.6%). In particular, in the EFM group, the rate of operative deliveries changes from 44 (5.8%) to 18.2 (2.4%) and the rate of cesarean sections decreases from 40 (5.4%) to 16.6 (2.2%) (Table 6). Similarly, the number of non-reassuring FHR tracings among patients who underwent EFM evaluation goes down from 175 (23.4%) to 72.4 (9.7%) (Table 6).
After the reevaluation of the 2 trials using the proposed correction factor, the comparison of the recalculated cesarean section and operative delivery rates for fetal distress between the EFM and IA groups was no longer statistically significant, both in the Dublin trial and Vintzileos trial (Tables 5,6). Even the comparison of the recalculated incidence of the rate of non-reassuring FHR patterns in the EFM and IA groups does not give any indication of significance for the Vintzileos trial (Table 6).

Discussion

Our study aimed to reevaluate the results of the 2 most representative trials reported in scientific literature and by the Cochrane review on the subject of EFM vs IA during labor using the EFM tracings indications in the recent NICHD guidelines. Our retrospective reevaluation noted that the previous differences in terms of obstetrical intervention for delivery in the case of fetal distress (cesarean section or operative spontaneous delivery) disappear after reading the EFM tracings using the NICHD classification.

We suppose that the interpretative key of these results is represented by the increasing importance of the role of fetal acidosis as a “unique” indication for cesarean delivery or instrumental delivery. As stated in the NICHD guidelines, the concept of fetal hypoxia should be replaced by the concept of fetal acidosis since, in the past, fetal hypoxia, often represented by fetal deceleration, was considered the indication for delivery; nowadays hypoxia is no longer an indication. Specifically, decelerations, if not associated with absent variability, should be considered an aspect of fetal reactivity, instead of fetal distress. Fetal deceleration represents a compensatory reaction that may subsequently worsen in a decompensatory phase represented by reduction of variability. So, if fetal hypoxia is embodied by FHR deceleration, fetal acidosis is mostly represented by the reduction of FHR variability that should therefore assume the most relevant value in the readings of EFM patterns. Category III of the NICHD guidelines is, in fact, clearly defined by the presence of sinusoidal patterns or absent baseline FHR variability in association with at least one among bradycardia and recurrent late or variable decelerations. Category III embodies an increased risk of fetal acidosis at the time of observation and, if unresolved, these traces should be treated with a prompt delivery.4

Reviewing the 2 trials, it appears clear that suspicious or non-reassuring FHR characteristics have been revisited with the introduction of the NICHD guidelines and some criteria that were previously used to define pathological patterns have since been reconsidered with the NICHD classification. In both classifications used in the trials, the Boylan classification from 1987 and the ACOG classification from 1989, a definition of normal FHR baseline between 120 bpm and 160 bpm was used, while nowadays the lower limit for this parameter is fixed at 110 bpm. Furthermore, applying the Boylan classification, the Dublin trial considered “dangerous” and an indication for cesarean section at least 1 criterion among severe tachycardia with low variability, severe bradycardia, and late and variable decelerations irrespective of FHR variability, which is an essential criterion to allocate the tracings in category III of the NICHD classification. Similarly, in the Vintzileos trial, the reduction of FHR variability is an isolated parameter sufficient to expedite delivery, without any association of other tracing characteristics, like in category III of the NICHD.

Recently, Clark et al.15 revisited the current NICHD approach to the management of category II FHR patterns, indicating operative or cesarean delivery in cases of FHR patterns less severe than those of category III. Considering that algorithm for category II, the results of our reevaluation of the Dublin and Vintzileos trials could possibly be less impressive, but we chose not to include the Clark algorithm in our analysis since we preferred to focus only on consolidated national guidelines on intrapartum monitoring. Future research will probably elucidate this interesting approach to category II and its impact on clinical practice and previous trials.

From our comparison, it is evident that the indications for obstetrical intervention in the 2 randomized trials fail when EFM incorporates the concept of acidosis. This difference was less pronounced in the Dublin trial where the fetal pH was assessed with scalp blood sampling.

In the clinical practice, the biggest hazard for the obstetrician is represented by the misinterpretation of the EFM tracings, since EFM has an extreme intra-observer and inter-observer variability. Blackwell et al. recently demonstrated that while the intra-observer agreement is substantial for the reading of EFM tracings with the NICHD guidelines, the inter-observer agreement is moderate for category I and II, but poor for category III.16

The misjudgment of physiological FHR variations as signs of hypoxia may inevitably expose inappropriate obstetrical interventions. The ACOG recently reported that the second most common indication for primary cesarean section is the presence of non-reassuring or indeterminate EFM tracings, accounting for about 23% cases of cesarean section.17,18 Given that, a standardized approach to EFM tracing management could reduce the rate of inappropriate cesarean sections. Surgical intervention is frequently determined by the presence of abnormal EFM that does not represent a clear situation of hypoxia or acidosis, since FHR variations may occur before the onset of acidosis. Electronic fetal heart monitoring may identify fetuses that are not exposed to hypoxic stimuli (category I), fetuses that are able to overcome the hypoxic stimulus to which they are exposed by compensatory responses (category II) and fetuses that are incapable of overcoming the hypoxic stimulus and are confronted with decompensation (category III).
The strength of our study is represented by the critical revision of the well-established association between EFM and the increment of operative delivery in comparison with IA in low-risk labor, demonstrating that a proper reading of the FHR tracings in view of the recent guidelines does not increase the incidence of surgical intervention. Moreover, our study points out the concept of acidosis in reading the FHR traces that may alter the outcome of previous trials; similarly, the use of fetal ECG ST-segment analysis has been questioned as an adjunct to conventional intrapartum electronic FHR monitoring, since a recent multicenter trial demonstrated that it did not improve the perinatal outcomes or decrease operative-delivery rates.10

A limitation of our study is that our results were obtained by recalculating previous data with a correction factor rather than through a direct assessment of individual cases. Other limitations are represented by the intrinsic bias of the 2 trials examined. Concerning the Dublin trial, the major confounding factors are the use of support techniques such as pH measurement of the fetal scalp and the transition from one group to another (EFM and IA), which may have masked some benefits arising from one or the other of the monitoring techniques, and may have excluded high-risk patients for meconium-stained amniotic fluid or prolonged labor.

Different aspects may alter the results of the Vintzileos trial. As before, no additional methods of fetal surveillance were used (i.e., pH of the fetal scalp). Moreover, both low-risk and high-risk pregnancies were included in the study as well as preterm labors, and the study was conducted based on a previous statistical assessment carried out to identify the sample number needed to demonstrate the real difference in perinatal mortality rates between the 2 groups.

In conclusion, all guidelines regarding EFM before 2008 were designed more to avoid hypoxia than acidosis. In addition, the results of the Cochrane meta-analysis of 2013 do not show a significant improvement in neonatal outcomes using EFM or IA. These considerations lead us to reconsider the results of the Dublin trial and Vintzileos trial in terms of operational rates of births (cesarean sections and operative spontaneous delivery), hypothesizing that these results would have been significantly lower if FHR traces had been interpreted using the current NICHD guidelines which aim to identify potential acidotic fetuses rather than hypoxic ones. Future clinical prospective trials will be needed to resolve any doubts and confirm our results.

References
1. Boylan P. Intrapartum fetal monitoring. Baillieres Clin Obstet Gynaecol. 1997;1(1):73–95.
2. Macones GA, Hankins GDV, Spong CY, Hauth J, Moore T. The 2008 National Institute of Child Health and Human Development Workshop Report on Electronic Fetal Monitoring. Obstet Gynecol. 2008;112(3):661–666.
3. American College of Obstetricians and Gynecologists. ACOG Practice Bulletin. Clinical Management Guidelines for Obstetrician-Gynecologists, Number 70, December 2005 (Replaces Practice Bulletin Number 62, May 2005). Intrapartum fetal heart rate monitoring. Obstet Gynecol. 2005;106(6):1463–1466.
4. American College of Obstetricians and Gynecologists. ACOG Practice Bulletin No. 106: Intrapartum fetal heart rate monitoring: Nomenclature, interpretation, and general management principles. Obstet Gynecol. 2009;114(1):192–202.
5. Madaan M, Trivedi SS. Intrapartum electronic fetal monitoring vs intermittent auscultation in postcesarean pregnancies. Int J Gynaecol Obstet. 2006;94(2):123–125.
6. Haverkamp AD, Orleans M, Langendoerfer S, McFee J, Murphy J, Thompson HE. A controlled trial of the differential effects of intrapartum fetal monitoring. Am J Obstet Gynecol. 1979;134(4):399–412.
7. Wood C, Renou P, Oats J, Farrell E, Beischer N, Anderson I. A controlled trial of fetal heart rate monitoring in a low-risk obstetric population. Am J Obstet Gynecol. 1981;141(5):527–534.
8. Di Tommaso M, Seravalli V, Cordisco A, Consorti G, Meacci F, Rizzello F. Comparison of five classification systems for interpreting electronic fetal monitoring in predicting neonatal status at birth. J Matern Fetal Neonatal Med. 2013;26(5):487–490.
9. Alfivric Z, Devane D, Gyte GM. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. In: Alfivric Z, ed. Cochrane Database of Systematic Reviews. Chichester, UK: John Wiley & Sons, Ltd; 2013.
10. Belfort MA, Saade GR, Thom E, et al. A randomized trial of intrapartum fetal ECG ST-segment analysis. N Engl J Med. 2015;373(7):632–641.
11. MacDonald D, Grant A, Sheridan-Pereira M, Boylan P, Chalmers I. Comparison of five classification systems for interpreting electronic fetal monitoring in predicting neonatal status at birth. J Matern Fetal Neonatal Med. 2013;26(5):487–490.
12. Vintzileos AM, Nochimson DJ, Guzman ER, Knuppel RA, Lake M, Schifrin BS. Intrapartum electronic fetal heart rate monitoring versus intermittent auscultation: A meta-analysis. Obstet Gynecol. 1995;85(1):149–155.
13. ACOG technical bulletin. Fetal heart rate patterns: Monitoring, interpretation, and management. Number 207 – July 1995 (replaces No. 132, September 1989). Int J Gynaecol Obstet. 1995;51(1):65–74.
14. Steer PJ, Eigbe F, Lissauer TJ, Beard RW. Interruptions of labor: abnormal cardiotocograms in labor, meconium staining of the amniotic fluid, arterial cord blood pH, and Apgar scores. Obstet Gynecol. 1998;92(3):715–721.
15. Clark SL, Nageotte MP, Gartie TJ, et al. Intrapartum management of category II fetal heart rate tracings: Towards standardization of care. Am J Obstet Gynecol. 2013;209(2):89–97.
16. Blackwell SC, Grobman WA, Antoniewicz L, Hutchinson M, Gyamfi-Bannerman C. Interobserver and intraobserver reliability of the NICHD 3-Tier Fetal Heart Rate Interpretation System. Am J Obstet Gynecol. 2011;205(4):378.e1–378.e5.
17. American College of Obstetricians and Gynecologists, Society for Maternal-Fetal Medicine. Obstetric care consensus No. 1: Safe prevention of the primary cesarean delivery. Obstet Gynecol. 2014;123(3):693–711.
18. Barber EL, Lundsberg LS, Belanger K, Pettker CM, Funai EF, Illuzzi JL. Indications contributing to the increasing cesarean delivery rate. Obstet Gynecol. 2011;118(1):29–38.