Validation of functional fetal autonomic brain age score fABAS in 5 min short recordings

To cite this article: Dirk Hoyer et al 2015 Physiol. Meas. 36 2369

View the article online for updates and enhancements.

Related content
- Monitoring fetal maturation—objectives, techniques and indices of autonomic function
  Dirk Höyer, Jan ebrowski, Dirk Cysarz et al.
- Linear and nonlinear measures of fetal heart rate patterns evaluated on very short fetal magnetocardiograms
  Eder Rezende Moraes, Luiz Otavio Murta Jr, Oswaldo Baffa et al.
- Development of the fetal autonomic nervous system
  U Schneider, E Schleussner, A Fiedler et al.

Recent citations
- Martin Black award for the best paper published in 2015
  Randall Moorman and Maggie Simmons
Validation of functional fetal autonomic brain age score fABAS in 5 min short recordings

Dirk Hoyer¹, Uwe Schneider², Eva-Maria Kowalski¹,³, Alexander Schmidt¹, Otto W Witte¹, Ekkehard Schleußner², Wolfgang Hatzmann¹, Dietrich HW Gröneneyer² and Peter van Leeuwen⁵

¹ Biomagnetic Center, Hans Berger Department of Neurology, Jena University Hospital, Friedrich Schiller University, Jena, 07743, Germany
² Department of Obstetrics, Jena University Hospital, Friedrich Schiller University, Jena, 07743, Germany
³ Institute of Biomedical Engineering and Informatics, Ilmenau University of Technology, Ilmenau, 98693, Germany
⁴ Faculty of Health, University Witten/Herdecke, Bochum, 58448, Germany
⁵ Gröneneyer Institute for Microtherapy, University Witten/Herdecke, Bochum, 58448, Germany

E-mail: dirk.hoyer@med.uni-jena.de

Received 15 July 2015, revised 1 September 2015
Accepted for publication 14 September 2015
Published 21 October 2015

Abstract
With the objective of evaluating the functional maturation age and developmental disturbances we have previously introduced the fetal autonomic brain age score (fABAS) using 30 min fetal magnetocardiographic recordings (fMCG, Jena). The score is based on heart rate pattern indices that are related to universal principles of developmental biology. The present work aims at the validation of the fABAS methodology on 5 min recordings from an independent database (fMCG, Bochum).

We found high agreement of fABAS obtained from Jena normal fetuses (5 min subsets, n = 364) and Bochum recordings (n = 322, normal fetuses). fABAS of 48 recordings from fetuses with intra-uterine growth restriction (IUGR, Bochum) was reduced in most of the cases, a result consistent with IUGR fetuses from Jena previously reported. fABAS calculated from 5 min snapshots only partly covers the accuracy when compared to fABAS from 30 min recordings. More precise diagnosis requires longer recordings.

fABAS obtained from fMCG recordings is a strong candidate for standardized assessment of functional maturation age and developmental
disturbances. Even 5 min recordings seem to be valuable for screening for maturation problems.

Keywords: fetal autonomic brain age score fABAS, fetal development, IUGR, fetal magnetocardiography, validation study, prenatal diagnosis, CTG

(Some figures may appear in colour only in the online journal)

Introduction

Fetal brain development and autonomic control includes growth, maturation and adaption of the hypothalamic-pituitary-adrenal (HPA) axis and autonomic nervous system (ANS) structure and function. Its evaluation is important with respect to the early identification of fetal developmental disorders since these have implications for health problems in later life which cannot completely be compensated for by later postnatal therapies (Barker 1998, Van den Bergh et al 2005).

In general, the association between ANS activity and heart rate variability (HRV) patterns provides a widely used diagnostic approach based on the electrocardiogram which, in postnatal life, is easily acquired and forms the basis for heart rate variability analysis. Since the fetus is extensively shielded and protected inside the maternal uterus, noninvasive fetal heart beat detection provides a unique window for respective prenatal HRV analysis. The various methods for fetal heart beat detection suffer from different disadvantages. While clinically established ultrasound based cardiotocography (CTG) disregards the single heart beats and, hence, fast vagally modulated rhythms, fetal electrocardiography (fECG) suffers from high rates of failed detection (van Leeuwen et al 2014). Magnetocardiography (fMCG) precisely detects each heart beat but the duration of data acquisition is limited since the women are required to remain in a stable position without movements. So far, none of these methods is without disadvantages and it remains an open question which technology and investigation modalities are best suited for the assessment of the functional maturational age and developmental disturbances. In that context, several aspects such as data length, state classification, sampling precision, HRV indices, and comparable standards are current subjects of research.

Various linear and nonlinear/complexity HRV indices have been previously reported as depending on fetal age and developmental disturbances such as intrauterine growth retardation (Van Leeuwen et al 1999, 2003, Schneider et al 2006, Ferrario et al 2009, Hoyer et al 2009, Huhn et al 2011, Wallwitz et al 2012) but their precision was low with regard to individual diagnosis. Furthermore, their comparison and validation across study centers is pending. In our view, the selective assignment towards vagal and sympathetically mediated rhythm addresses only a part of autonomic control and may not be appropriate to reflect the complex operation of the ANS. The good performance of complexity measures as well as overall variability give respective indication. Furthermore, the consideration of fetal behavioural states and particular heart rate patterns such as heart rate accelerations plays an important role. But both can only partly be identified in general, and, in particular if considering their premature forms before 32 weeks of gestation. With respect to those limitations, we recently developed a functional fetal autonomic brain age score (fABAS) (Hoyer et al 2013b). It is based on HRV indices that reflect universal developmental characteristics, namely increasing fluctuation amplitude, increasing complexity and pattern formation. fABAS precisely predicted fetal maturation age in data sets of quiet and active sleep as well as, with slightly lower precision, in 30 min recordings without state selection (Hoyer et al 2013b, 2014). Those results seem
to support the conceptional advantage of using universal developmental characteristics for
describing developing fetal behavior and complex ANS function. It may also help to deal with
methodological limitations in state classification. The limited available data length is a further
important issue with regard to practical application in clinical routine diagnosis.

The present work is focused at the meaningfulness of short recordings with respect to the
fABAS (Hoyer et al. 2013b, 2014) and its validation between different study centers. Using
previous results obtained from fetal magnetocardiographic (fMCG, Jena) recordings of 30 min
duration, in this work we address the question to what extent the fABAS score can be vali-
dated in independent data (fMCG, Bochum) in which the recording duration is short (5 min)
and a sleep classification is lacking.

Methods

Subjects and data acquisition

Fetal magnetocardiographic nonstress recordings were performed in two centers using com-
parable standard obstetric observation methods and similar acquisition modalities (e.g. 1024
Hz sampling rate) (Jena: (Hoyer et al 2009, 2013a), Bochum: (Van Leeuwen et al 1999, 2003,
2013)).

All Jena measurements were taken using the 195 channel vector-magnetograph ARGOS
200 (ATB Chieti, Italy) in a magnetically shielded room (two-layer shielded room, ATB
Chieti, Italy) at the Biomagnetic Center of the Jena University Hospital. In Bochum the data
were collected using one of two biomagnetometer systems: (a) a 37-channel system (Siemens
Krenikon, Erlangen, Germany), consisting of first-order gradiometers with a baseline of 7 cm
with a probe diameter of 19 cm or (b) a 67-channel system (BTI 1300C, San Diego, USA),
consisting of magnetometer sensing coils arranged concentrically on a slightly curved sur-
face with a diameter of 33 cm and including a configuration of 11 reference coils used to
detect ambient noise. Both systems were installed in a standard magnetically shielded room
(Vakuumschmelze AK3b, Hanau, Germany).

The respective preprocessing and heart beat detection are comparable between the centers
and do not differ concerning the heart beat interval detection.

The learn set from the Biomagnetic Center of the Jena University Hospital consisted of 364
fMCG recordings of normal fetuses aged 22–38 weeks of gestation (WGA). The entire 30 min
and subsections of the first 5 min were analysed. The validation set from the Grö
nemeyer
Institute for Microtherapy, University Witten/Herdecke, Bochum consisted of 5 min record-
ings of 322 of normal fetuses aged 19–40 WGA and 48 IUGR fetuses aged 27–39 WGA. The
WGA distribution of the normal data (figure 1) is more Gaussian leading to an over-weighted
contribution of recordings around about 25–37 WGA of the Jena database in contrast to the
relatively uniform distribution between 19–40 WGA of the Bochum database.

Exclusion criteria in the normal group were as follows. Maternal: known heart diseases,
diabetes mellitus, medication affecting cardiac function/rhythm, abuse of nicotine, alcohol or
drugs, previous administration of synthetic glucocorticoids, uterine contractions during the
recording. Fetal: known chromosomal abnormalities, sonographically identified malforma-
tions, fetal cardiac arrhythmias.

IUGR was diagnosed on the basis of ultrasound examination (estimated weight < 10th
Percentile with respect to gestational age). Distinction from constitutionally small fetuses was
based on either Doppler ultrasound (pathological uteroplacental perfusion > 24 WGA, Jena)
or paediatric assessment post natum (Bochum).
In both centers the investigations were approved by the Local Ethics Committee and all the subjects gave their written informed consent.

**fABAS and statistical analysis**

fABAS was previously introduced in detail (Hoyer et al 2013a). It was designed as a multivariate linear regression model under consideration of 5 heart rate variability indices that were selected with respect to universal developmental principles as shown in table 1.

fABAS was previously fitted to 10 min sections of active and quiet sleep, respectively (Hoyer et al 2013b), as well as to 30 min recordings without state selection (Hoyer et al 2014) by multivariate regression under consideration of those five HRV parameters. All HRV parameters have been calculated according to the previous work. Only, with respect to the present short data sets, VLF/LF was calculated based on 3 min sections (Welch algorithm) in contrast to the previously used 5 min sections. In patients with repeated recordings the intra-individual correlation was ignored since the score is intended to be appropriate for single recordings. All models were significant ($p < 0.05$).

fABAS cross validation to the 5 min subsections was first based on 30 min recordings (learn set) and in a second step refitted to a learn set drawn from 5 min sections/recordings. Goodness of fit is provided as standard error SE and corrected determination coefficients $R^2$. The fABAS values of the IUGR fetuses were plotted versus mean and 95% CI of the distribution of the normal group fABAS values.

This work contains (i) the comparison of fABAS values from 30 min to those from 5 min subsections of normal fetuses from the Jena database, (ii) fABAS transition (cross validation) between 5 min subsections of Jena database and the recordings of the Bochum data base, and (iii) the comparison of fABAS values between normal and IUGR fetuses of the Bochum database.

**Results**

**Normal group Jena**

While roughly similar (see figure 2) the precision of fABAS with coefficients fitted to the 30 min recordings was reduced in their application to the first 5 min sections of the identical
recordings. The standard error SE increased from 2.73 to 3.7 weeks (see table 2). Also, fABAS fitted to the 5 min data did not reach the precision obtained from the 30 min recordings (SE = 3.39).

Cross validation with normal group Bochum

The 5 min data sets of both databases provide almost comparable fABAS results. It should be taken into consideration that there are two confounding effects. Beside the different study centers, the Jena data set ranges from about 22 to 38 WGA with a clustering around 25 to 37 WGA) while the Bochum data set is distributed relatively uniformly between 19 to 40 WGA. Therefore, the coefficients of determination $R^2$ in the Bochum database are systematically increased due to their larger range of gestational ages (table 2). In this circumstance, the standard errors are more representative. When applying the fABAS coefficients obtained from the Jena learn set to the Bochum validation set, the increase of SE by 1.18 week is partly caused by the incomplete WGA range in the learn set. In contrast, when applying the coefficients obtained from the Bochum learn set to the Jena validation set, SE is reduced by 0.93 week and approaches the values of the Jena learn set. Those results indicate the high agreement in the investigated overlapping WGA range. Furthermore, they show that fABAS may also appropriate for the evaluation of the maturation age over the larger WGA range from about 19 to 40 WGA.

Comparison of normal and IUGR fetuses

The individual fABAS values, mean and 95% confidence interval of the Bochum normal group model fit are shown in figure 3, left panel. The fABAS values of the IUGR group appear between fitted mean and lower 95% confidence interval of the normal group. Higher fABAS values, but still below the 95% CI of the normal group were found only in four recordings at >35 WGA.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Calculation</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMP</td>
<td>20–95 Inter-quantile distance of detrended NN interval series</td>
<td>Fluctuation range of heart beat intervals</td>
</tr>
<tr>
<td>gMSE(3)</td>
<td>Generalized mutual information at coarse graining level 3 of NN interval series</td>
<td>Complexity of sympatho-vagal modulations</td>
</tr>
<tr>
<td>skewness</td>
<td>Skewness of NN interval series</td>
<td>Change from predominant decelerations towards predominant acceleration patterns</td>
</tr>
<tr>
<td>pNN5</td>
<td>Percentage of differences between adjacent NN intervals that are &gt;5 ms.</td>
<td>Formation of vagal rhythms</td>
</tr>
<tr>
<td>VLF/LF</td>
<td>Ratio between very low (0.02–0.08 Hz) and low (0.08–0.2 Hz) frequency band power</td>
<td>Baseline fluctuation in relation to sympatho-vagal modulations</td>
</tr>
</tbody>
</table>

Table 1. HRV parameters constituting fABAS and their interpretation (adapted from (Hoyer et al 2013b).
Established FIGO-CTG monitoring of fetal well-being requires recordings of up to 60 min in order to determine an active behavioral section according to criteria defined from 34 weeks of gestation upwards (Nijhuis et al 1982, Pardey et al 2002, FIGO 2011). The maturation stage of the fetal autonomic brain age was recently assessed in two ways: (1) by separate fABAS settings during active and quiet sleep related behavioral states of 10 min sections (Hoyer et al 2013b) and (2) by a fABAS setting that integrated changing states and heart rate pattern over an unselected period of 30 min (Hoyer et al 2014). From this work, several questions arose. On the one hand, it remained unclear whether short recordings, which do not allow a reliable distinction between fetal behavioral states, were applicable to reflect autonomic brain age. On the other hand, little is known about the reproducibility of the fABAS score determined in the data obtained by investigators in other centers using comparable equipment and protocols. Since we compared data sets of different age ranges the gold standard of Bland-Altman-plot was not applicable and we discuss the results based on coefficient of determination $R^2$ and standard error $SE$ of the regression models. It should be taken into consideration that $R^2$ is influenced by the range of investigated age (variance to be explained by the score) and, hence, only partly appropriate to compare data sets of different age ranges. In contrast, the standard error of the estimated age provides a direct measure of the precision of the estimated age. For example, considering the learn sets: $SE = 2.73$ obtained from the 30 min Jena recordings indicates an error range of estimated age of 2.73 weeks maturation age, that error is increased towards $SE = 3.69$ in the 5 min sections of that recordings. That means the error of maturation age prediction increases by 0.88 week, which equals about 6 d gestational age. $R^2$ decreases

Table 2. fABAS cross validation results, standard error $SE$ and coefficient of determination $R^2$.

<table>
<thead>
<tr>
<th>fABAS</th>
<th>Learn set</th>
<th>Validation set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$SE$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>30 min $\rightarrow$ 5 min (both Jena)</td>
<td>2.73</td>
<td>0.61</td>
</tr>
<tr>
<td>5 min (Jena) $\rightarrow$ 5 min (Bochum)</td>
<td>3.39</td>
<td>0.39</td>
</tr>
<tr>
<td>5 min (Bochum) $\rightarrow$ 5 min (Jena)</td>
<td>4.37</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Discussion

Established FIGO-CTG monitoring of fetal well-being requires recordings of up to 60 min in order to determine an active behavioral section according to criteria defined from 34 weeks of gestation upwards (Nijhuis et al 1982, Pardey et al 2002, FIGO 2011). The maturation stage of the fetal autonomic brain age was recently assessed in two ways: (1) by separate fABAS settings during active and quiet sleep related behavioral states of 10 min sections (Hoyer et al 2013b) and (2) by a fABAS setting that integrated changing states and heart rate pattern over an unselected period of 30 min (Hoyer et al 2014). From this work, several questions arose. On the one hand, it remained unclear whether short recordings, which do not allow a reliable distinction between fetal behavioral states, were applicable to reflect autonomic brain age. On the other hand, little is known about the reproducibility of the fABAS score determined in the data obtained by investigators in other centers using comparable equipment and protocols.

Since we compared data sets of different age ranges the gold standard of Bland-Altman-plot was not applicable and we discuss the results based on coefficient of determination $R^2$ and standard error $SE$ of the regression models. It should be taken into consideration that $R^2$ is influenced by the range of investigated age (variance to be explained by the score) and, hence, only partly appropriate to compare data sets of different age ranges. In contrast, the standard error of the estimated age provides a direct measure of the precision of the estimated age. For example, considering the learn sets: $SE = 2.73$ obtained from the 30 min Jena recordings indicates an error range of estimated age of 2.73 weeks maturation age, that error is increased towards $SE = 3.69$ in the 5 min sections of that recordings. That means the error of maturation age prediction increases by 0.88 week, which equals about 6 d gestational age. $R^2$ decreases
from 0.61 to 0.27 in this case of identical age distribution. In contrast, the increase of $SE$ from the Jena 5 min to the Bochum 5 min data sets ($3.39 \rightarrow 4.57$) appears in connection with increasing $R^2$ ($0.39 \rightarrow 0.52$): This relationship is mainly caused by a clearly enlarged age distribution range in connection with the only slightly increased $SE$. Hence, in the present validation analysis the effects of data length, range of gestational age, and study center interfere.

The main results of the present validation study are (i) fABAS obtained from 5 min recordings, in spite of being less precise, reflect essential characteristics of functional brain maturation age and IUGR related changes, (ii) there is a high agreement between the results of both study centers, (iii) fABAS methodology seems to be appropriate for the extended range of gestational age from about 19 WGA upwards. Finally, the concept of estimating the functional fetal brain maturation based on universal developmental indices, namely heart rate fluctuation amplitude, complexity and pattern formation that constitute fABAS, was confirmed.

The slight loss of precision of fABAS between 30 min and 5 min in the Jena data set disregarding activity characteristics is not surprising. Previously, we were able to fit separate models more precisely for pre-classified sections of quiet and active sleep state as well as when time segments according to particular signal patterns were considered (Hoyer et al. 2013b, 2014). There, we used 10 min sections for a reliable state classification. The randomness of 5 min snapshots is obvious. Concerning their state classification, in addition to the short data length, fundamental physiological limitations are due to the immense evolving state dynamics in the investigated period from 15 WGA upwards. The present results suggest that one single linear model may completely describe the extended investigated age range, though with high variability. Previously, we found small quadratic parts in models of both active and quiet sleep and 30 min state-independent recordings of fetuses aged 21/25–40 WGA (Hoyer et al. 2013b, 2014), all models were clearly more precise than the present analysis of 5 min recordings. We can only speculate on the linearity across the extended range of analysed gestational ages. It is known that different parts of the ANS predominantly mature during different age periods with remarkable stages around 28 and 32 WGA (Nijhuis et al. 1999, Schneider et al. 2009). The resulting question, if and to which extent the developmental process according to the overall (universal) developmental characteristics reflects those particular periods, cannot be explored in more detail from the highly variable 5 min data studied here. Longer data sets are required. Nevertheless, the present results suggest that fABAS provides an estimate of the functional maturation age over that extended period as a mainly linear continuous process.

![Graph showing fABAS values of the normal group of the Bochum fetal monitoring database, mean ± 95%CI (left) and its overlay with the IUGR group’s values (dark dots, right).](image)
The 5 min data results from Bochum and Jena nicely validate each other. fABAS previously developed using the Jena data set that range mainly from about 25 to 37 WGA could be extended to the range of at least 19 to 40 WGA. This result confirms the concept of designing a maturation age score based on universal developmental characteristics that are included in fABAS. The previously introduced factors of fABAS that reflect developmental characteristics such as increasing fluctuation amplitude, increasing complexity, and pattern formation (Hoyer et al. 2013b) seem to apply very well over the extended WGA range studied here. Our results are highly suggestive of consistency between the study centers taking into account that only short data sets were used and fetal activity not accounted for. Since even 30 min recording do not necessarily contain both, a quiet and an active behavioral states segment, longer recordings would be preferable. Recordings over 24h would allow the assessment of the entire circadian sleep-wakefulness cycle. But those recordings are rarely performed, due to practical reasons. According to the longest duration of 40 min of a quiet state in normal fetuses, in the FIGO cardiotocography guidelines (FIGO 2011) recordings of up to 60 min are recommended in order to acquire one active period for analysis. A corresponding investigation period of fMCG recordings according to our previous analyses, namely separately within each state as well as across different states, seems reasonable. An extension to 90 min might bring some reserve for a certain appearance of both states and better evaluation of developmental disturbances.

In a previous analysis of 20 IUGR fetuses (26–39 WGA) from the Jena database, we found a relative state distribution of quiet sleep/active sleep/active wakefulness of 21%/58%/21% compared to 26.4%/66.8%/6.8% in normal fetuses (Hoyer et al. 2013b). Hence, fHR patterns in IUGR seem to suggest more active wakefulness. However, this assumption is inconsistent with the usual reduction in fetal activity due to limited nutritional supply and, as information on fetal movements is lacking, must be interpreted with caution (Baschat 2010). In contrast, HRV was reduced in the sections of quiet sleep (traditional indices, Schneider et al. 2006) and active sleep (Hoyer et al. 2013b). fABAS values of the IUGR recordings in active sleep were situated around the lower 95% CI of the normal group in active sleep and mainly separable from the normal group (82.3% correctly classified by means of logistic regression).

In the present analysis of IUGR fetuses from the Bochum database, most of the fABAS values are below the mean but within the 95% CI of the normal group. Only in 4 cases aged above 35 WGA fABAS was clearly larger, but still below the upper 95% CI. This result is consistent with the previous results, if we take the following methodological differences into account. The lower discriminatory power can be explained by the missing state classification. By visual inspection of the heart rate patterns the 4 cases with larger fABAS values can clearly be assigned to active awakeness. The lower values reflect a mixture of active and quiet sleep.

Those results reconfirm the importance of a state selective scoring. Nevertheless, despite of the lower discriminatory power, the systematic IUGR related decrease of fABAS as a function of the gestational age previously found could be confirmed. Reduced short and long term variability were found in IUGR fetuses using standard FIGO-CTG monitoring (Nijhuis et al. 2000). This result is consistent with our own findings, however without the additional information on maturation age provided by fABAS.

Conclusion

The cross validation of results for the data from the two study centers showed that fABAS provides a general score of functional autonomic brain age. Recordings over 5 min lack a certain amount of information that can be obtained from 30 min recordings. Nevertheless, even 5 min
recordings seem to be valuable for the screening of the autonomic function in the human fetus. Hence, fABAS obtained from fMCG recordings was confirmed as being a strong candidate for standardized assessment of functional brain maturation age and developmental disturbances.

Acknowledgments

Dirk Hoyer and Uwe Schneider acknowledge the support by Deutsche Forschungsgemeinschaft: ‘Biomagnetic investigations of fetal autonomic and central nervous maturation and its disturbances due to intrauterine growth restriction and glucocorticoid administration’ (DFG: Ho 1634/12-2, Schn 775/2–3). Franziska Jaenicke, Esther Heinicke, Anja Rudolph, Ulrike Wallwitz, Isabel Kynass, Franziska Bode, Janine Tegtmeyer, Kathrin Kumm, and Samuel Nowack contributed to fMCG recordings and/or data management.

References

Barker D J 1998 In utero programming of chronic disease Clin. Sci. 95 115–28
FIGO 2011 (Federation Internationale De Gynecologie Et D’Obstetrics) Maternity Guidelines, Federal Gazette (Germany: Institute for Quality and Efficiency in Health Care)
Hoyer D et al 2014 Fetal autonomic brain age scores, segmented heart rate variability analysis, and traditional short term variability Front. Hum. Neurosci. 8 948
Van Leeuwen P, Geue D, Lange S, Hatzmann W and Gronemeyer D 2003 Changes in the frequency power spectrum of fetal heart rate in the course of pregnancy *Prenat. Diagn.* **23** 909–16
Van Leeuwen P, Lange S, Bettermann H, Gronemeyer D and Hatzmann W 1999 Fetal heart rate variability and complexity in the course of pregnancy *Early Hum. Dev.* **54** 259–69