Investigation of quantitative imaging biomarkers for assessing perinatal outcomes

Elisenda Bonet Carné

PhD THESIS

Programa de Doctoral en Biomedicina

Neurociències

Universitat de Barcelona

Investigation of quantitative imaging biomarkers for assessing perinatal outcomes

Submitted by

Elisenda Bonet Carné

To obtain the degree of "Doctor in Biomedicine"

and the International Doctor Mention

Director:

Professor Eduard Gratacós Solsona

Director:

Professor Ferran Marqués Acosta

Universitat de Barcelona

Facultat de Medicina

Programa de Doctorat de Biomedicina

Neurociències

A Thesis submitted by Elisenda Bonet Carné for the PhD degree (Doctor in Biomedicine, University of Barcelona) including the mention of "European Doctor" under the direction of Eduard Gratacós Solsona, Professor of Obstetrics and Gynecology at Barcelona University and Ferran Marqués Acosta, Professor of Signal Theory and Communications at the Universitat Politècnica de Catalunya, BarcelonaTECH.

Elisenda Bonet Carné

Barcelona, September 2014.

Professor Eduard Gratacós Solsona

Director, Barcelona Center for Maternal-Fetal and Neonatal Medicine and Fetal I+D Research Center of Fetal and Perinatal Medicine,

Hospital Clinic i Hospital Sant Joan de Deu

Professor of Obstetrics and Gynecology, Universitat de Barcelona

Professor Ferran Marqués Acosta

Professor in Signal Theory and Communications

Dean of ETSETB – TelecomBCN

Universitat Politècnica de Catalunya, BarcelonaTECH

We confirm that **Elisenda Bonet Carné** has conducted under our supervision the studies presented in the Thesis "**Investigation of quantitative imaging biomarkers for assessing perinatal outcomes**". The present Thesis has been structured following the normative for PhD Thesis as a compendium of publications, to obtain the degree of **International Doctor in Biomedicine** and the mentioned studies are ready to be presented to a Tribunal.

In addition, the co-directors also confirm that none of the co-authors has used, or is going to use, any of the articles here presented in another PhD Thesis.

Codirector **Codirector**

Eduard Gratacós Solsona **Ferran Marqués Acosta**

Barcelona, September 2014.

Prof Eduard Gratacós and Prof Ferran Marqués as codirectos of this Thesis, we declare that **Elisenda Bonet Carné** has performed under our supervision the studies presented in the Thesis entitled "**Investigation of quantitative imaging biomarkers for assessing perinatal outcomes**". This Thesis has been structured following the normative for PhD Thesis as a compendium of publications and the PhD candidate specific role performed in each publication is explained below:

Study 1. T. Cobo, E. Bonet-Carne, M. Martinez-Terron, A. Perez-Moreno, N. Elias, J. Luque, I. Amat-Roldan, M. Palacio. Feasibility and Reproducibility of Fetal Lung Texture Analysis by Automatic Quantitative Ultrasound Analysis and Correlation with Gestational Age.Fetal Diagn Ther. 2012 Apr; 31(4):230-6. The authors Teresa Cobo and **Elisenda Bonet-Carne** have equally contributed to this study. **Elisenda Bonet-Carne** has performed the image feature extraction, the statistical learning algorithm processing, and contributed to writing and revising the manuscript.

Study 2.M. Palacio, T. Cobo, M. Martinez-Terron, G.A. Ratta, E. Bonet-Carne, I. Amat-Roldan, E. Gratacos. Performance of an automatic quantitative ultrasound analysis of the fetal lung to predict fetal lung maturity. Am J Obstet Gynecol. 2012 Dec;207(6): 504.e1-5. **Elisenda Bonet-Carne** has contributed to the study by performing the image processing, the machine learning prediction, writing and revising the manuscript.

Study 3. E. Bonet-Carne, M. Palacio, T. Cobo, A. Perez-Moreno, M. Lopez, JP. Piraquive, JC. Ramirez, F. Marques, E. Gratacos. Quantitative ultrasound texture analysis of fetal lungs to predict neonatal respiratory morbidity.Ultrasound in Obstetrics & Gynecology. **Elisenda Bonet-Carne** has substantially contributed to the study by performing the development, image quality control, data analysis revision, writing and revision of the manuscript.

Study 4. M. Sanz-Cortes, GA. Ratta, F. Figueras, E. Bonet-Carne, N. Padilla, A. Arranz, N. Bargallo, E. Gratacos. Automatic Quantitative MRI Texture Analysis in Small-for-Gestational-Age Fetuses Discriminates Abnormal Neonatal Neurobehavior. PLoS ONE 2013 8(7): e69595. **Elisenda Bonet-Carne** has contributed to the study, by performing the image analysis, writing and revision of the manuscript.

Codirector **Codirector**

Eduard Gratacós Solsona **Ferran Marqués Acosta**

Barcelona, September 2014.

PRESENTATION

This Thesis has been structured following the normative for PhD Thesis, as a compendium of publications, to obtain the degree of International Doctor in Biomedicine. It was approved by the "Comissió del programa de Doctorat en Biomedicina" on 20th June 2011. The studies included in the Thesis belong to the same research line, leading to four papers already published or submitted for publication in international journals:

Study 1. T. Cobo, E. Bonet-Carne, M. Martinez-Terron, A. Perez-Moreno, N. Elias, J. Luque, I. Amat-Roldan, M. Palacio. Feasibility and Reproducibility of Fetal Lung Texture Analysis by Automatic Quantitative Ultrasound Analysis and Correlation with Gestational Age.Fetal Diagn Ther. 2012 Apr; 31(4):230-6.

* T.C. and E.B.-C. contributed equally to this paper.

Status: published

Journal Impact factor: 1.902

Quartile: 2nd, area Obstetrics & Gynecology.

Study 2.M. Palacio, T. Cobo, M. Martinez-Terron, G.A. Ratta, E. Bonet-Carne, I. Amat-Roldan, E. Gratacos. Performance of an automatic quantitative ultrasound analysis of the fetal lung to predict fetal lung maturity. Am J Obstet Gynecol. 2012 Dec;207(6):504.e1-5.

Status: published

Journal Impact factor: 3.877

Quartile: 1st, area Obstetrics & Gynecology.

Study 3. E. Bonet-Carne, M. Palacio, T. Cobo, A. Perez-Moreno, M. Lopez, JP. Piraquive, JC. Ramirez, F. Marques, E. Gratacos. Quantitative ultrasound texture analysis of fetal lungs to predict neonatal respiratory morbidity.Ultrasound in Obstetrics & Gynecology.

Status: accepted,5th June 2014. Ref: UOG-2014-0288.R1

Journal Impact factor: 3.557

Quartile: 1st, area Obstetrics & Gynecology.

Study 4. M. Sanz-Cortes, GA. Ratta, F. Figueras, E. Bonet-Carne, N. Padilla, A. Arranz, N. Bargallo, E. Gratacos. Automatic Quantitative MRI Texture Analysis in Small-for-Gestational-Age Fetuses Discriminates Abnormal Neonatal Neurobehavior. PLoS ONE 2013 8(7): e69595.

Status: published

Journal Impact factor: 3.730

Quartile: 1st, area Multidisciplinary sciences.

TABLE OF CONTENT

TABLE OF CONTENT

1. INTRODUCTION

1. INTRODUCTION

1.1. Quantitative Medical Imaging

Imaging techniques are commonly used in medicine to create images of the human body parts that are hidden by the skin and bones. Medical imaging can be used to diagnose as well as to treat disease; this Thesis is focused in quantitative medical imaging used as a diagnostic technique.

The main objective of diagnostic studies based on images is the characterization of tissues: images are acquired in order to determine whether the tissues in the selected area for study show normal (healthy tissue) or pathological characteristics. The process to classify an image as pathological or not requires a complex evaluation performed by an experienced professional. However, there are some diagnostic problems for which simple visual analysis of the image is insufficient for the specific characterization of the tissue. Subtle changes in tissue brightness or texture are difficult or impossible to identify by subjective interpretation [1, 2].

Quantitative analysis of medical images may increase repeatability and assist in solving ambiguities in the interpretation of different images. As a starting point, researchers naturally considered the image characteristics that radiologists use explicitly or implicitly in their evaluation of tissue appearance. Intensity, morphology and texture are usually quoted as important characteristics. Image texture analysis is known to be a particularly sensitive characteristic in the evaluation of pathologies. The human observer has a limited sensitivity to textural properties, whereas mathematical techniques for texture analysis give quantitative and therefore objective elements [3] which are invisible to the human eye [4].

The texture of image refers to the appearance, structure and arrangement of the parts of an object within the image. Texture analysis is a technique that extracts patterns from images based on the characterization of the microstructural information that may not be assessed visually [5, 6]. Texture analysis may be performed on any medical

image as ultrasound images [7], however most applications or studies have been performed on MRI because of the great amount of detail provided by this technique [8].

Some disciplines, as Maternal-Fetal medicine, are perfect candidates for quantitative imaging as diagnostic tools due to the lack of accessibility to the fetal tissues. Although there are plenty of imaging techniques, ultrasound is the one extensively used among obstetricians and is a central diagnostic technology in Maternal-Fetal Medicine because ultrasound does not use radiation, does not require special facilities to perform an analysis, there are some pocket size ultrasound machines and its use is economic. Additionally, Magnetic Resonance Imaging (MRI) is the most common technique for structural neuroimaging since it is a non-invasive technique that provides a great contrast between grey and white matter which could be useful to evaluate the central nervous system [9].

1.1.1. Ultrasound imaging

Ultrasound imaging is a non-invasive way to acquire images from body parts (for example, from soft tissues) based on acoustic waves. To generate ultrasound images a transducer is required to convert electricity into sound using piezoelectric crystals. The ultrasound waves are sent from the transducer and propagate through different tissues and they return reflected as echoes to the transducer. Those echoes are converted back into electrical impulses and are processed in order to form the ultrasound image. The waves are reflected at the surfaces between the tissues of different acoustic density. Ultrasound images can detect tissue variations because each type of tissue presents different acoustic properties. Different ultrasound machines and proves are currently available which makes ultrasound a bedside technique for tissue examination.

Ultrasound image interpretation can be difficult, especially when the target is to detect subtle tissue pathologies or changes. Several studies demonstrate that theoretically

spectral ultrasound parameters are related to tissue microstructure and that different tissue architectures can be interpreted (e.g. Lizzi et al. 1983 [10]). From early 80's the idea that quantitative ultrasound analysis may increase repeatability and assists in solving ambiguities in the interpretation of ultrasound examinations was widespread. Moreover, Insana et al. (1988) used quantitative techniques to estimate tissue characteristics and imaging techniques to extract diagnostic information of either the features contained in the ultrasound images or the reflected acoustic waves; thus demonstrated that specific image features can be measured consistently [7]. Furthermore, recent experiments, performed on cell samples exposed to a chemotherapeutic drug, illustrate that it is possible to detect and measure cell level changes in ultrasonic images [11, 12].

Typical ultrasound imaging frequencies range from 1 to 15MHz and lateral resolution corresponds to 3mm to 0.3mm. This resolution is best at the focal length distance and widens away from this distance in a non-uniform way because of diffraction effects [13]. Due to the resolution, much information can be extracted from ultrasound images. This information, or image features, characterizes the biological tissue that has been analysed and can be understood as several markers which can be used to monitor or detect specific biological processes. Features can be used as an input to computer science, artificial intelligence or machine learning algorithms to generate computational techniques to predict or monitor a specific biological process. Several studies were performed to investigate the use of quantitative ultrasound for different medical diagnostic applications, including breast cancer [14, 15]and liver disease [16-18]. Over years powerful quantitative techniques for ultrasound image analysis have been developed thanks to improvements in computer capacity and image resolution [8]. Up to the authors' knowledge, nowadays elastography, which has been used successfully in liver diseases (Fibroscan®), is the only image analysis diagnostic technique used in clinical practice, besides of those based on morphometric measures [19].

1.1.2. Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is a non-invasive technique based on tissue response to radio-frequency pulses. In MRI, the patient is placed in the magnetic field, and a pulse of radio waves is generated by antennas (coils) positioned around the patient. The protons in the patient absorb the radio waves, and subsequently reemit this radio wave energy after a period of time that depends upon the spatially dependent magnetic properties of the tissue. The radio waves emitted by the protons in the patient are detected by the antennas that surround the patient. MRI produces a set of images that depict slices through the patient. Each point in an image depends on the micromagnetic properties of the tissue corresponding to that point in the body. Due to each type of tissue presents different local magnetic properties, MRI has a high sensitivity to anatomical variations.

In MRI, frequency is proportional to magnetic field strength. There is no optimum field strength for MR imaging although certain MR-based techniques require high fields nevertheless clinical MR imaging does not. The diverse nature of applications requires different systems operating at an appropriate field [20].

Several studies were performed to investigate the use of quantitative MRI for different medical diagnostic applications [5, 21-23],also including breast cancer [24] and liver disease [25]. As far as the author knowledge, a computer software for calculation of texture parameters in MR images, MaZda, that extracts information from MR images is the only final software that is currently available for this purpose [26].

1.2. Quantitative imaging in Fetal Lung Maturity

The most common cause of mortality and neonatal morbidity in preterm and early term fetuses is lung immaturity that causes neonatal respiratory morbidity, defined as respiratory distress syndrome or transient tachypnea of the newborn [27, 28]. Neonatal respiratory morbidity is not restricted to very preterm births and remains high among late-preterm and early-term infants born before 39 weeks of gestation

[29-31]. Fetal Lung Maturity (FLM) is mainly determined by production of pulmonary surfactant by type II pneumocytes which increases during the gestation. Thus, the most accurate non-invasive predictor of lung maturity is gestational age. At present, assessment of FLM is performed by means of laboratory tests in amniotic fluid [32-36] which require the performance of an invasive procedure. Besides being risky, it cannot be performed in all clinical settings due to the required facilities. The need of amniocentesis has resulted in a decline in the use of this information clinically. It is clear that for some indications, delivery should occur regardless of FLM results. However, there is an open debate about the value of FLM testing in the decision-making process for those clinical situations in which late preterm or early-term delivery may seem a reasonable option but delivery could be postponed if fetal lung immaturity assessed [28]. Determining the risk of FLM without the need for an invasive technique might have a tremendous impact in the clinical management of such cases. Aside from economic implications, avoiding the need of amniocentesis would be associated with less patients' discomfort and related complications, and controversies about indications for fetal lung maturity assessment could be approached from a different perspective.

Fetal lung maturity assessment by non-invasive methods is an unsolved problem despite 20 years of extensive research focused on the development of quantitative imaging solutions based on ultrasound imaging to test fetal lung maturity, reported approaches included gray level measurements, lung tissue motion and relative features of lung to placental or liver images [37-42]. These studies suggest the potential of quantitative analysis of ultrasound images to predict fetal lung maturity but the diagnostic accuracy was insufficient for clinical use.

Therefore, despite advancing in the knowledge of risk factors and the introduction of many public health strategies designed to reduce the risk of neonatal respiratory morbidity, its prenatal prediction still remains a main challenge.

1.3. Quantitative imaging in Fetal Brain

Smallness for gestational age affects 10% of all pregnancies [43]. In clinical practice when an estimated fetal weight is below the tenth percentile and Doppler assessment of the umbilical artery is normal, the diagnosis of a small-for-gestational-age (SGA) is reached [44-46]. Although some fetuses with this diagnosis are constitutionally small, in a substantial proportion of cases, the diagnosis of SGA identifies mild forms of fetal growth restriction due to placental insufficiency that are not expressed by umbilical artery Doppler. Therefore, fetal development occurs in suboptimal conditions, with a deprived delivery of oxygen and nutrients to the fetal brain [47]. Under these conditions brain reorganization may take place, among other changes of the so-called fetal programming [48]. Different authors have shown how despite the fact that most SGA fetuses reach term without signs of deterioration, there is a proportion of them that presents an increased risk for an adverse perinatal outcome [49-51] with an abnormal neonatal neurobehavior [52, 53] and impaired neurodevelopment in early childhood [54].

The most common technique to acquire a brain image is MRI. However in Maternal-Fetal field, cranial ultrasound is the preferred imaging modality to assess fetal or neonatal brain status since it allows visualization of brain structures in a non-invasive approach [55]. Cranial Ultrasound is bedside feasible, non-invasive and non-expensive. On the other hand, MRI is difficult to perform in infants, especially in those preterm neonates who need special cares. Previous works from our research group were focused on extracting textural information from neonatal brain ultrasound images [56, 57]. When MRI and Cranial Ultrasound are compared as predictors of neurodevelopmental outcomes, they present a high variability in the studies but a close concordance between both techniques [58]. A previous study from our research group demonstrated that the evaluation of fetal brain using MRI texture analysis is feasible. Texture analysis on fetal brain MRI showed discrimination based on brain textural features between SGA and adequate-for-gestational-age (AGA) fetuses [59].

Considering its prevalence, SGA constitutes a challenge and an opportunity for public health to improve the impact of prenatal conditions in quality of life. However, at present the detection of SGAs at risk of abnormal neurodevelopment is limited since standard clinical examinations fail to identify significant differences.

1.4. Relevance and justification of the research

Imaging techniques as ultrasound, MRI (Magnetic Resonance Imaging) or CT (Computed Tomography) are extensively used as diagnostics technologies in medicine. In fact, ultrasound imaging is a central diagnostic tool used in Maternal-Fetal Medicine. Although it has reached its full potential for the diagnosis of macro-structural changes, ultrasound contains much more information, which the subjective human inspection is not able to distinguish [60, 61]. The proposed PhD Thesis has the main objective of finding potential imaging biomarkers, using quantitative image analysis, to improve the subjective inspection of images in order to predict altered outcomes in fetal population.

In order to accomplish the objectives of this PhD Thesis and fully investigate possible imaging biomarkers for Maternal-Fetal medicine, a clear outcome and huge databases are required. Images of the fetal brain could not be used as an appropriate model to demonstrate the main hypothesis due to the uncertainty associated to the neurodevelopmental outcomes. Furthermore, huge databases are required to develop an image biomarker and MR images are more difficult to acquire than ultrasound ones because of the acquisition discomfort and prize. For these reasons, neonatal respiratory morbidity will be used as the main pathological model for this Thesis (**STUDY 1**, **2** and **3**) and brain images will be used to test the transversely of the quantitative texture analysis in other pathological model (**STUDY 4**).

This PhD Thesis is part of two larger projects. On one hand, fetal thorax ultrasound images were used for the fetal lung maturation studies (**STUDY 1**, **2** and **3**) which are part of a Transmural Biotech and Fetal and Perinatal Medicine Research Group of

Hospital Clínic of Barcelona larger project. On the other hand, **STUDY 4** is part of a larger prospective research program on Small for Gestational Age (SGA) fetuses involving image fetal acquisition and short- and long-term postnatal follow-up at Fetal and Perinatal Medicine Research Group of Hospital Clínic of Barcelona.

In summary, this Thesis consists on different studies to develop a non-invasive imaging biomarker to predict neonatal respiratory morbidity as a clinical outcome. In this case, the development of the imaging biomarker may help to plan delivery in some cases, and might have an impact in obstetric management. In general, imaging biomarkers could contribute to clinical diagnosis in a non-invasive manner. In **STUDY 1**, we evaluated the feasibility and reproducibility of a texture feature extractor software for the estimation of quantitative features in fetal lung. Specifically, we explored the ability to correlate quantitative image information with gestational age as a preliminary evidence to justify further research on non-invasive assessment of fetal lung maturity. **STUDY 2** correlated quantitative image features with fetal lung maturity assessed by an amniotic fluid test (TDx-FLM II [62, 63]). Then, quantusFLM^{TM} (Transmural Biotech, SL, Barcelona, Spain) was specifically designed to predict neonatal respiratory morbidity through the analysis of textural image features from fetal lung images. In **STUDY 3** we described the basic principles of a novel method to predict neonatal respiratory risk in a non-invasive manner, quantusFLM^{TM}; and we tested its performance. Transversely of texture analysis methods using another pathological model and a different acquisition image technique was tested in **STUDY 4**.

The main goal of this Thesis is to use quantitative image features to predict the risk of abnormal clinical outcomes. In order to achieve this objective this Thesis is divided in four specific studies.

2. HYPOTHESES

2. HYPOTHESES

2.2. Main hypothesis

Image texture analysis methods could be developed for the analysis of medical images (i.e. ultrasound or magnetic resonance imaging) in the field of fetal medicine applications, to characterize microstructural information that may not be assessed by standard clinical evaluation, in a reproducible and reliable manner allowing its use as imaging biomarker to assess or predict specific clinical outcomes. Fetal ultrasound of the lung and the assessment of fetal lung maturity is an optimal candidate model to test a first approach in order to develop such methods.

2.3. Specific hypotheses

1. Quantitative image analysis of ultrasound images of fetal lung tissue allows extracting reproducible features and patterns that correlate with gestational age.

2. Quantitative image analysis of ultrasound images of fetal lung tissue allows extracting reproducible features and patterns that correlate with the results of lung maturity tests in amniotic fluid.

3. Quantitative image analysis of ultrasound images of the fetal lung can be used to develop reproducible imaging biomarkers to predict the risk of neonatal respiratory morbidity.

4. Quantitative image analysis methods are potentially useful to assess other fetal areas and other techniques (i.e. magnetic resonance imaging) and to identify patterns associated with changes in fetal brain development induced by prenatal conditions, such as intrauterine growth restriction.

3. OBJECTIVES

3. OBJECTIVES

3.1. Main objective

To explore the development of a series of new methods based on image texture analysis allowing the analysis of medical images (i.e. ultrasound or magnetic resonance imaging) in the field of fetal medicine applications -mainly fetal lung maturity and fetal brain assessment-, to test their reproducibility and to select the best performing approach to develop an imaging biomarker predicting a clinical outcome of interest.

3.2. Specific objectives

1. To assess the relationship between ultrasound texture analysis of the fetal lungs and the stages of lung maturation occurring across gestational age.

2. To evaluate the correlation between quantitative ultrasound analysis of the fetal lung and the results of fetal lung maturity assessed by clinical standard methods in amniotic fluid.

3. To develop and evaluate the performance of a novel method to predict neonatal respiratory morbidity based on quantitative analysis of fetal lung by ultrasound.

4. To test whether texture analysis in magnetic resonance images could identify patterns associated with an abnormal neurobehavior in small for gestational age neonates.

4. MATERIALS AND RESULTS

4.1. STUDY 1

Feasibility and Reproducibility of Fetal Lung Texture Analysis by Automatic Quantitative Ultrasound Analysis and Correlation with Gestational Age

T. Cobo, E. Bonet-Carne, M. Martinez-Terron, A. Perez-Moreno, N. Elias, J. Luque, I. Amat-Roldan, M. Palacio. Fetal Diagn Ther. 2012 Apr; 31(4):230-6.

* T.C. and E.B.-C. contributed equally to this paper.

Status: published

Journal Impact factor: 1.902

Quartile: 2nd

Original Paper

Fetal Diagnosis and Therapy

Fetal Diagn Ther 2012;31:230–236
DOI: 10.1159/000335349

Received: November 2, 2011
Accepted after revision: November 23, 2011
Published online: April 24, 2012

Feasibility and Reproducibility of Fetal Lung Texture Analysis by Automatic Quantitative Ultrasound Analysis and Correlation with Gestational Age

Teresa Cobo[a] Elisenda Bonet-Carne[b] Mónica Martínez-Terrón[a]
Alvaro Perez-Moreno[b] Núria Elías[a] Jordi Luque[b] Ivan Amat-Roldan[b]
Montse Palacio[a, c]

[a]Department of Maternal-Fetal Medicine, Institute Clinic of Gynecology, Obstetrics and Neonatology, Hospital Clinic, Institut de Investigacions Biomèdiques August Pi-Sunyer (IDIBAPS), Universitat de Barcelona, [b]Transmural Biotech SL, and [c]CIBER de Enfermedades Raras (CIBERER), Barcelona, Spain

Key Words
Fetal lung maturity · Quantitative ultrasound

Abstract
Objective: To evaluate the feasibility and reproducibility of fetal lung texture analysis using a novel automatic quantitative ultrasound analysis and to assess its correlation with gestational age. ***Methods:*** Prospective cross-sectional observational study. To evaluate texture features, 957 left and right lung images in a 2D four-cardiac-chamber view plane were previously delineated from fetuses between 20 and 41 weeks of gestation. Quantification of lung texture was performed by the Automatic Quantitative Ultrasound Analysis (AQUA) software to extract image features. A standard learning approach composed of feature transformation and a regression model was used to evaluate the association between texture features and gestational age. ***Results:*** The association between weeks of gestation and fetal lung texture quantified by the AQUA software presented a Pearson correlation of 0.97. The association was not influenced by delineation parameters such as region of interest (ROI) localization, ROI size, right/left lung selected or sonographic parameters such as ultrasound equipment or transducer used. ***Conclusions:*** Fetal lung texture analysis measured by the AQUA software demonstrated a strong correlation with gestational age. This supports further research to explore the use of this technology to the noninvasive prediction of fetal lung maturity.

Introduction

The most common cause of neonatal morbidity and mortality in preterm and term fetuses is lung immaturity. The strongest predictor of lung maturity is gestational age. Thus, infants who are born at <39 weeks have significantly higher rates of neonatal morbidities, including respiratory distress syndrome (RDS) when compared with infants born at ≥39 weeks of gestation [1, 2].

T.C. and E.B.-C. contributed equally to this paper.

KARGER
Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1015–3837/12/0314–0230$38.00/0

Accessible online at:
www.karger.com/fdt

Dr. Montse Palacio
Department of Maternal-Fetal Medicine
Hospital Clinic, Sabino de Arana 1
ES–08028 Barcelona (Spain)
Tel. +34 93 227 9946, E-Mail mpalacio@clinic.ub.es

The surfactant to albumin ratio [3] is the most common method to assess lung maturity but requires an amniocentesis. In addition, the accuracy of this test is not optimal assuming that the optimal cutoff to predict RDS results in a sensitivity of 89% with a specificity of 83%, respectively [4].

However, these studies have shown weak correlations hindering their incorporation to clinical application. Over the last 30 years, the prediction of lung maturity by non-invasive methods has been extensively explored [5–12] with promising results. Thus, changes in placental maturity observed by ultrasound have been proposed as markers of fetal lung maturity with contradictory results [5–7]. Similar discordant results have been reported when lung echogenicity changes, compared with liver, were suggested as markers of lung maturity [8–11].

Quantitative ultrasound analysis has been proposed to improve robustness and to solve the current ambiguities in extracting additional information from ultrasound images. This approach is based on an image processing method to the already acquired ultrasound picture, which estimates tissue characteristics by means of a quantitative analysis. The use of quantitative texture analysis in ultrasound has previously been investigated for medical diagnostic applications, including breast cancer [13–16] and liver disease [17–19]. There are few references in the literature on the assessment of fetal lung texture by quantitative ultrasound tissue characterization. Although different methodologies have been proposed, they all report differences in lung texture features along gestation when compared to the reflection pattern of liver [20–24].

The Automatic Quantitative Ultrasound Analysis (AQUA) is a custom-developed software, which remains unmodified under illumination changes as it does not use direct grey level from the image or tissue references. This method estimates texture features based on conditional random fields (CRF) so that image texture features converge robustly to different tissue characteristics independently of the overall acquisition context (i.e. scanner settings or operator skills). In a previous study, the AQUA software was tested in preterm neonates and achieved a high accuracy in the early identification of brain changes in subclinical stages [25]. However, the feasibility of conducting reproducible measurements in fetal tissue with this software has not been explored.

In this study, we evaluated the feasibility and reproducibility of the AQUA software for the estimation of quantitative features in fetal lungs. We explored the ability of the features extracted by AQUA to provide information related to gestational age as preliminary evidence to justify further research on noninvasive assessment of lung maturity. In addition, the reproducibility and robustness of the method under different examination settings were explored.

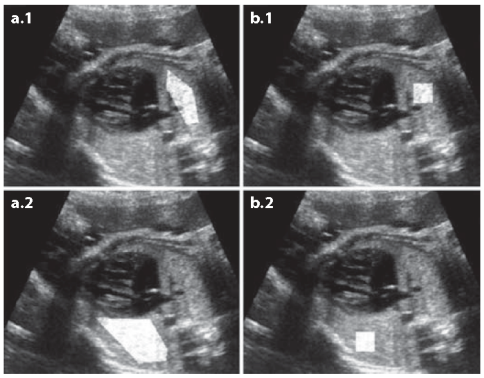


Fig. 1. Manual (**a**) and square (**b**) delineation of each lung. 1 = Right lung; 2 = left lung. Gestational age = 33.4 weeks.

Methods

The study population included singleton pregnancies attending the Maternal-Fetal Medicine Department at Hospital Clinic in Barcelona for routine pregnancy ultrasound scans from June 2010 to December 2010. Multiple pregnancy and structural/chromosomal anomalies were excluded from the study. The study protocol was approved by the local Ethics Committee (ID 3823–2007) and pregnant women provided written informed consent. In all pregnancies gestational age was calculated based on the crown-rump length at first trimester ultrasound. The number of cases included per each gestational age was approximately ≥30.

Image Acquisition and Regions of Interest

A semilateral transverse four-cardiac-chamber view plane for lung image acquisition was performed using a Siemens Sonoline Antares (Siemens Medical Systems, Malvern, Pa., USA), Voluson 730 Pro, Voluson 780 Pro (GE Medical Systems, Milwaukee, Wisc., USA) and ALOKA Prosound Alpha-7 (Hitachi Aloka Medical, Ltd.) ultrasound equipment, all equipped with curved linear transducer with a frequency range from 3 to 7.5 MHz. Equipment settings were adjusted at the discretion of the clinician performing the ultrasound in order to obtain the optimal image quality according to clinical criteria.

Data was digitally collected in the original Digital Imaging and Communication in Medicine (DICOM) format and stored for the off-line analysis in a custom made program with a Graphic User

Table 1. Demographic characteristics and pregnancy outcome of the study population

Parameter	Total (n = 957)	GA < 22.0 (n = 137)	GA ≥ 38.0 (n = 112)	GA 22.0–37.6 (n = 708)
Maternal age, years	31.1 (5.7)	30.5 (6.2)	30.8 (5.7)	32.6 (5.3)
Nulliparity, n (%)	288 (30)	42	194	52
GA at delivery, weeks	39.5 (1.6)	39.1 (2.6)	39.4 (1.3)	40.4 (1.2)
Birthweight	3,320.1 (499.1)	3,280.1 (576.7)	3,274.4 (474.7)	3,484.5 (465.8)
5-Apgar	9.9 (0.5)	9.8 (1.1)	9.9 (0.4)	9.9 (0.1)
UA pH	7.21 (0.35)	7.23 (0.07)	7.22 (0.06)	7.16 (0.73)
Preeclampsia, n (%)	6 (0.6)	2 (1.4)	3 (2.6)	1 (0.1)
IUGR, n (%)	11 (1.1)	2 (1.4)	6 (5.3)	3 (0.4)
Preterm birth < 34, n (%)	1 (0.1)	1 (0.7)	0	1 (0.1)
NICU admission, n (%)	15 (1.5)	3 (2.1)	11 (9.8)	1 (0.1)
RSD, n (%)	6 (0.6)	0	6 (5.3)	0
Exitus, n (%)	0	0	0	0

Data given as means ± SD or %. GA = Gestational age; UA = umbilical artery; IUGR = intrauterine growth restriction; NICU = neonatal intensive care unit; RSD = respiratory distress syndrome.

Interface (GUI) using MATLAB R2007b (version 7.5.0.342; MATLAB; The MathWorks Inc., Natick, Mass., USA). Manual (freehand) and square (60x60 pixels) delineation were performed in each lung by one physician ensuring that the delineation only included lung tissue as illustrated in figure 1. Maternal and neonatal outcomes were recorded. Clinical statistical analyses were performed with the SPSS 18.0 statistical software. Statistical learning analyses were performed using R statistical software [26].

AQUA and Statistical Learning Algorithm

A computing model of gestational age based on textural features from transverse fetal lung ultrasound images was developed following two basic steps: (1) feature extraction, and (2) statistical learning algorithm. Feature extraction was performed by means of an AQUA algorithm, mainly based on conditional random fields (CRF) and previously reported by Tenorio et al. [25], based on texture analysis. Details about the CRF methodology are available in the appendix.

Image features were calculated from each scan, lung and delineation independently. Neither other regions nor scans were used. The analysis was performed using different delineations to ensure that feature extraction was lung side independent since in some acquisitions the image quality differs in the left and right lung. The feature extraction method obtained 15,300 features per delineation. A statistical learning algorithm was then applied to select only the most relevant features, those not influenced by noise and clinical acquisition variability.

Regarding the statistical learning algorithm the data was split into two bags. Bag 1 contained a cohort of fetuses with gestational ages at acquisition below 22 weeks (Bag 1a) and ≥38 weeks of gestation (Bag 1b) and Bag 2 contained the cohort of fetuses between 22 and 38 weeks of gestational age. These sets were used in different parts of the learning algorithm; Bag 1 was used for feature transformation and Bag 2 to compute the regression model.

The feature transformation algorithm was based on (1) a feature selection step by means of Mahalanobis distance between ultrasound lung features from all Bag 1 fetuses, and (2) a dimension reduction step by means of principal component analysis. Therefore, the relevant features providing the best information related to gestational age were selected.

A regression model was computed using Bag 2 fetuses to test the robustness of the previously selected features extracted by the AQUA software and their association with gestational age. The regression model was based on the random forest model to estimate weeks of gestation using 30 relevant transformed features extracted by AQUA as input. Since the same sample bag was used to train and test, the regression was the same, and to ensure that the model was not over-fitted, the number of features introduced into the model was less than one third of the sample size [27].

The R^2, Pearson correlation and root mean square error (RMSE) were used as regression fitting indicators. Additional statistics were computed to evaluate the robustness of the algorithm with regard to scanner settings and acquisition equipment.

Results

From June 2010 to December 2010 a total of 957 singleton pregnancies were included in the study. Maternal characteristics are summarized in table 1. No significant changes in characteristics disaggregated by group were found.

The feature transformation algorithm was first trained with the ultrasound lung features of 249 fetuses, 112 with less than 22 weeks of gestation (Bag 1a) and 137 with more than 38 weeks (Bag 1b), obtaining 30 features. A regression model to correlate these 30 features with gestational

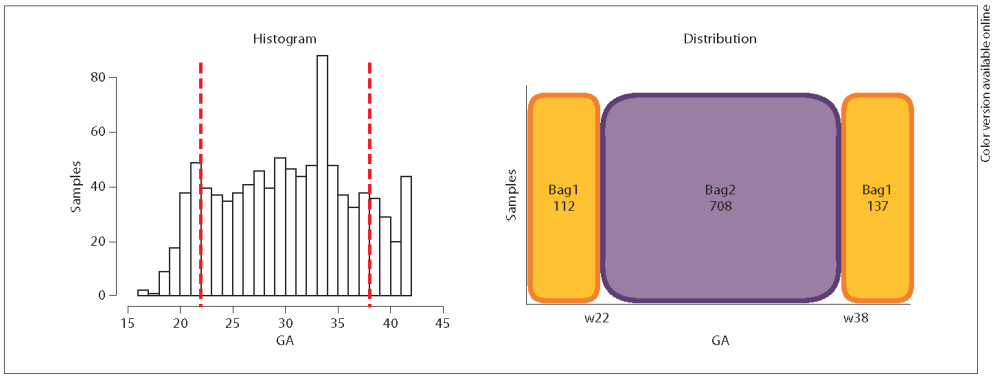


Fig. 2. Histogram and distribution of total sample (n = 957). Thresholds in red (in the online version only) and the final sample size of distributions.

age was then evaluated in a large cohort of 708 fetuses between 22 and 38 weeks of gestational age (Bag 2).

Sample distribution by weeks of gestation is shown in figure 2. The analysis showed a strong correlation of 97% between the AQUA lung extracted and transformed features (region of interest (ROI) size and side independent) and gestational age (table 2).

The robustness of features extracted by AQUA and transformed with regard to different scanner equipment is illustrated in figure 3. GE equipment does not provide transducers frequency information and therefore the images acquired with GE were not included in this analysis. Regarding the frequency setting (fig. 4), of 491 acquisitions evaluated, no significant differences were found.

Finally, there was a strong correlation between the estimated gestational age and the observed gestational age (R = 0.98) (fig. 5).

Table 2. Correlation between AQUA analysis and gestational age according to lung side and ROI size

	LLM	RLM	LLS	RLS
Pearson correlation	0.98	0.97	0.98	0.97
R^2	0.86	0.86	0.86	0.86
RMSE	1.61	1.58	1.62	1.62

LLM = Left lung manual delineation; RLM = right lung manual delineation; LLS = left lung square delineation; RLS = right lung square delineation.

Discussion

This study provides preliminary evidence that the AQUA software, which is based on texture models of CRF, can extract reproducible quantitative image features from lung ultrasound images related to gestational age. Since gestational age is strongly associated with fetal lung maturity, the findings of this study open a pathway for future research evaluating the relationship between texture analysis and lung maturity. The analyses performed with AQUA were not influenced by delineation parameters such as ROI localization, ROI size, right/left lung selected or sonographic parameters such as ultrasound equipment or transducers frequency used.

Quantification of lung structure using a noninvasive methodology remains a challenge to predict lung maturity in several obstetric situations. The use of quantitative ultrasound tissue characterization of normal fetal lung development has been investigated in recent years, showing better accuracy than lung echogenicity to detect histological changes [20–24].

Thus, Sohn et al. [20] reported a methodology to determine the maturity of fetal lung by comparing the frequency characteristics of lung echoes to those from the fetal liver as a reference organ. They concluded that the borderline between fetal lung immaturity and maturity is a relation called Q_{mean} ($Frequency_{mean}$ liver/$Frequency_{mean}$ lung) of 1.1 in which a higher ratio indicates lung

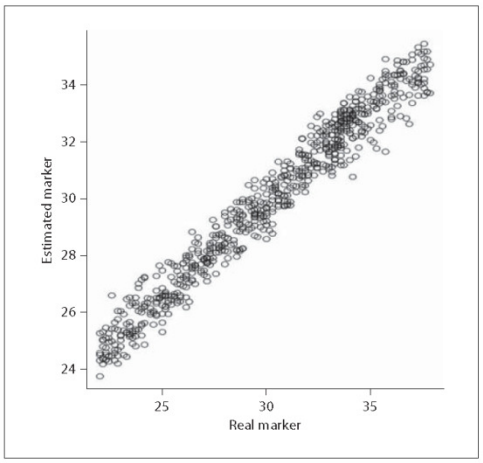


Fig. 3. Relation between the estimated and the observed gestational age. R Pearson correlation = 0.98. Due to the similarity between the results only those obtained with left lung manual delineation are presented.

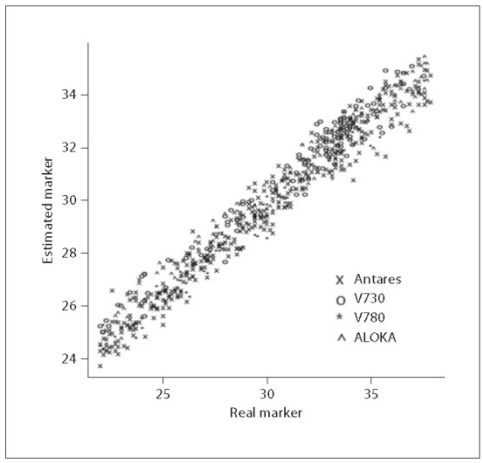


Fig. 4. Robustness of GA estimation using features extracted by AQUA, related to scanner equipment.

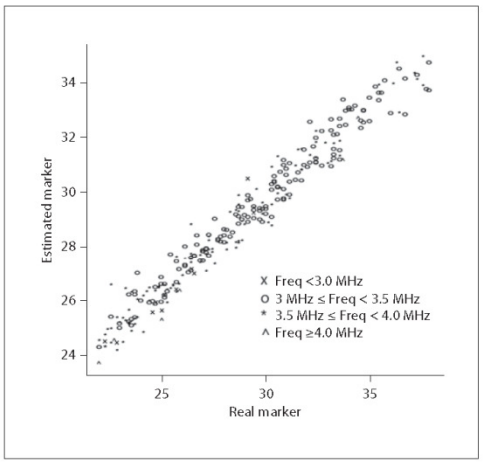


Fig. 5. Independence of GA estimation using features extracted by AQUA, regarding frequency of the probe.

immaturity and a lower ratio lung maturity. Maeda and coworkers evaluated grey-level histogram width (GLHW) of fetal lung and liver showing that an increase in GLHW lung-to-liver ratio at ≥30 weeks predicted RDS with an accuracy of about 80–90% [21, 22]. Prakash et al. [23] proposed three different training and test sets using a complex algorithm of lung-to-liver feature values with an accuracy from 73 to 96% to correctly classify their high pulmonary risk group. Finally, Tekesin et al. [24] examined the tissue-specific grey scale evaluating lung-to-liver grey level distribution. Although no significant differences in the mean grey value of fetal lung were observed along gestation, the ratio between lung and liver increased significantly from 24 to 31 weeks.

In line with previous data [20–24], the AQUA software provides excellent information about lung structural changes in transverse ultrasound images of fetal lung without the use of any other region or scan (such as liver) as a tissue reference to compute relative echodensities. Thus, AQUA is not influenced by ROI localization, providing facilities for image acquisition to sonographer. The robustness of the gestational age estimation using features extracted by AQUA with regard to scanner equipment as well as their independency of the variable

of the different scanner settings, simplify its inclusion in clinical practice.

The strength of this study is the large cohort of patients included to test the features extracted by the AQUA software. Moreover, AQUA might overcome some limitations of previous feature extraction methods; since no other region is needed as a reference. Thus, the algorithm proved to be robust when using different settings and equipment during acquisition.

This study presents some limitations. Firstly, since not all equipments provide frequency we could not ensure that scanner frequency does not affect the validation in all instances. However, no differences were found when different settings were used. Secondly, this study provides evidence that texture information extracted by the AQUA software could be used to estimate gestational age. However, no blind samples were used to test the predictive capacity of the model. Additional research is needed to evaluate its ability as an image biomarker predictor of lung maturity.

In summary, this study demonstrates that evaluation of lung texture analysis by the AQUA software is feasible and reproducible. In addition, texture quantitative features were highly correlated with gestational age. These results support further research to establish the potential use of the AQUA software as a noninvasive predictive method of fetal lung maturity.

Acknowledgements

This work was supported by a Rio Hortega grant from Instituto de Salud Carlos III in Spain (CM09/00213). E.B.-C. and Transmural Biotech SL thank the Comissionat per a Universitat i Recerca del Departament d'Innovació, Universitats i Empreses de la Generalitat de Catalunya, for financial support.

Appendix

In general terms, a CRF is a probabilistic graph model that encodes the conditional probability distribution between two sets of variables, the target variables and the observed variables [28]. In pattern recognition problems, the two sets of variables are defined according to specific scheme of neighboring pixels (nodes of the graph model) and pixel-to-pixel interactions (edges of the graph model) are quantified by the mentioned conditional probability distributions. These sorted conditional probability distributions can then be used to define a rich set of image features that characterize a particular ROI of a medical image. One of the main strengths of the CRF representation is that we do not impose any constraint on the distribution of the observed variables, and this enables to incorporate data which might be poorly understood (like ultrasound waves propagating through an unknown and heterogeneous media like mother's womb).

References

1 Spong CY, Mercer BM, D'Alton M, Kilpatrick S, Blackwell S, Saade G: Timing of indicated late-preterm and early-term birth. Obstet Gynecol 2011;118:323–333.
2 Teune MJ, Bakhuizen S, Bannerman CG, Opmeer BC, van Kaam AH, van Wassenaer AG, Morris JM, Mol BW: A systematic review of severe morbidity in infants born late preterm. Am J Obstet Gynecol 2011;205:374e1–9.
3 Grenache DG, Gronowski AM: Fetal lung maturity. Clin Biochem 2006;39:1–10.
4 Bennasar M, Figueras F, Palacio M, Bellart J, Casals E, Figueras J, Coll O, Gratacos E: Gestational age-specific cutoff levels of TDX-FLM II for the prediction of neonatal respiratory distress syndrome. Fetal Diagn Ther 2009;25:392–396.
5 Harman CR, Manning FA, Stearns E, Morrison I: The correlation of ultrasonic placental grading and fetal pulmonary maturation in five hundred sixty-three pregnancies. Am J Obstet Gynecol 1982;143:941–943.
6 Golde SH, Platt LD: The use of ultrasound in the diagnosis of fetal lung maturity. Clin Obstet Gynecol 1984;27:391–401.
7 Grannum PA, Berkowitz RL, Hobbins JC: The ultrasonic changes in the maturing placenta and their relation to fetal pulmonic maturity. Am J Obstet Gynecol 1979;133:915–922.
8 Fried AM, Loh FK, Umer MA, Dillon KP, Kryscio R: Echogenicity of fetal lung: Relation to fetal age and maturity. AJR Am J Roentgenol 1985;145:591–594.
9 Cayea PD, Grant DC, Doubilet PM, Jones TB: Prediction of fetal lung maturity: Inaccuracy of study using conventional ultrasound instruments. Radiology 1985;155:473–475.
10 Bowerman RA DK, Carson PL, et al: Ultrasonic prediction of pulmonary maturity: Correlation with l/s ratio. Annual Meeting American Institute of Ultrasound in Medicine, Kansas City, 1984.
11 Feingold M, Scollins J, Cetrulo CL, Koza D: Fetal lung to liver reflectivity ratio and lung maturity. J Clin Ultrasound 1987;15:384–387.
12 Zilianti M, Fernandez S: Correlation of ultrasonic images of fetal intestine with gestational age and fetal maturity. Obstet Gynecol 1983;62:569–573.
13 Chen DR, Chang RF, Kuo WJ, Chen MC, Huang YL: Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks. Ultrasound Med Biol 2002;28:1301–1310.
14 Wan C, Du J, Fang H, Li F, Wang L: Evaluation of breast lesions by contrast enhanced ultrasound: Qualitative and quantitative analysis. Eur J Radiol.
15 Feleppa EJ, Mamou J, Porter CR, Machi J: Quantitative ultrasound in cancer imaging. Semin Oncol 2011;38:136–150.
16 Caproni N, Marchisio F, Pecchi A, Canossi B, Battista R, D'Alimonte P, Torricelli P: Contrast-enhanced ultrasound in the characterisation of breast masses: utility of quantitative analysis in comparison with MRI. Eur Radiol 2011;20:1384–1395.

17 Hartman PC, Oosterveld BJ, Thijssen JM, Rosenbusch GJ: Variability of quantitative echographic parameters of the liver: intra- and interindividual spread, temporal- and age-related effects. Ultrasound Med Biol 1991;17:857–867.
18 Kadah YM, Farag AA, Zurada JM, Badawi AM, Youssef AM: Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images. IEEE Trans Med Imaging 1996;15:466–478.
19 Icer S, Coskun A, Ikizceli T: Quantitative grading using grey relational analysis on ultrasonographic images of a fatty liver. J Med Syst 2011 Apr 28. [Epub ahead of print].
20 Sohn C, Stolz W, Bastert G: Diagnosis of fetal lung maturity by ultrasound: a new method and first results. Ultrasound Obstet Gynecol 1991;1:345–348.
21 Maeda K, Utsu M, Yamamoto N, Serizawa M: Echogenicity of fetal lung and liver quantified by the grey-level histogram width. Ultrasound Med Biol 1999;25:201–208.
22 Serizawa M, Maeda K: Noninvasive fetal lung maturity prediction based on ultrasonic gray level histogram width. Ultrasound Med Biol 2010;36:1998–2003.
23 Prakash KN, Ramakrishnan AG, Suresh S, Chow TW: Fetal lung maturity analysis using ultrasound image features. IEEE Trans Inf Technol Biomed 2002;6:38–45.
24 Tekesin I, Anderer G, Hellmeyer L, Stein W, Kuhnert M, Schmidt S: Assessment of fetal lung development by quantitative ultrasonic tissue characterization: a methodical study. Prenat Diagn 2004;24:671–676.
25 Tenorio V, Bonet-Carne E, Botet F, Marques F, Amat-Roldan I, Gratacos E: Correlation between a semiautomated method based on ultrasound texture analysis and standard ultrasound diagnosis using white matter damage in preterm neonates as a model. J Ultrasound Med 2011;30:1365–1377.
26 Team. RDC: R: A language and environment for statistical computing. R foundation for statistical computing, ed http://www.R-project.org., 2008.
27 Hastie T, Tibshirani R, J Friedman: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, ed 2. New York, Springer, 2009.
28 Lafferty J, McCallum A, Pereira F: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. Int Conf Machine Learning, 2001.

4.2. STUDY 2

Performance of an automatic quantitative ultrasound analysis of the fetal lung to predict fetal lung maturity

M. Palacio, T. Cobo, M. Martinez-Terron, G.A. Ratta, E. Bonet-Carne, I. Amat-Roldan, E. Gratacos. Am J Obstet Gynecol. 2012 Dec;207(6): 504.e1-5.

Status: published

Journal Impact factor: 3.877

Quartile: 1st

Results from this study have been presented in the following congresses:

1. Ratta GA, Palacio M, Cobo T, Martinez-Terron M, Elias N, Bonet E, Amat I, Gratacos E. Non-invasive fetal lung maturity prediction through automatic quantitative ultrasound analysis texture extractor. ISUOG.September 9-13, 2012. Copenhagen, Denmark.

2. Bonet-Carne E, Cobo T, Luque J, Martinez-Terron M, Perez-Moreno A, Palacio M, Gratacos E, Amat-Roldan I. Consistent association between image features of fetal lungs from different ultrasound equipments and fetal lung maturity from amniocentesis. IEEE ISBI. May 2-5, 2012. BCN, Spain.

3. Palacio M, Cobo T, Martinez-Terron M, Ratta GA, Elias N, Bonet E, Amat-Roldan I, Gratacos E. Performance of an automatic quantitative ultrasound analysis (AQUA) texture extractor to predict fetal lung maturity assessed by TDx-FLM in amniotic fluid. SMFM.February 6-11, 2012. Dallas, USA.

RESEARCH www.AJOG.org

OBSTETRICS

Performance of an automatic quantitative ultrasound analysis of the fetal lung to predict fetal lung maturity

Montse Palacio, MD, PhD; Teresa Cobo, MD, PhD; Mònica Martínez-Terrón, MD; Giuseppe A. Rattá, MSc, PhD; Elisenda Bonet-Carné, MSc; Ivan Amat-Roldán, MSc; Eduard Gratacós, MD, PhD

OBJECTIVE: The objective of the study was to evaluate the performance of automatic quantitative ultrasound analysis (AQUA) texture extractor to predict fetal lung maturity tests in amniotic fluid.

STUDY DESIGN: Singleton pregnancies (24.0-41.0 weeks) undergoing amniocentesis to assess fetal lung maturity (TDx fetal lung maturity assay [FLM]) were included. A manual-delineated box was placed in the lung area of a 4-chamber view of the fetal thorax. AQUA transformed the information into a set of descriptors. Genetic algorithms extracted the most relevant descriptors and then created and validated a model that could distinguish between mature or immature fetal lungs using TDx-FLM as a reference.

RESULTS: Gestational age at enrollment was (mean [SD]) 32.2 (4.5) weeks. According to the TDx-FLM results, 41 samples were mature and 62 were not. The imaging biomarker based on AQUA presented a sensitivity 95.1%, specificity 85.7%, and an accuracy 90.3% to predict a mature or immature lung.

CONCLUSION: Fetal lung ultrasound textures extracted by AQUA provided robust features to predict TDx-FLM results.

Key words: fetal lung maturity, pregnancy, quantitative ultrasound

Cite this article as: Palacio M, Cobo T, Martínez-Terrón M, et al. Performance of an automatic quantitative ultrasound analysis of the fetal lung to predict fetal lung maturity. Am J Obstet Gynecol 2012;207:504.e1-5.

The most common cause of mortality and neonatal morbidity in preterm and early term fetuses is lung immaturity. The strongest predictor of lung maturity is gestational age, although a quantifiable risk of pulmonary morbidity caused by lung immaturity may be present at any gestational age. Thus, infants who are born at less than 39 weeks have significantly higher rates of neonatal morbidity, including respiratory distress syndrome (RDS) when compared with infants born at a gestation of 39 weeks or longer.[1,2]

The current methods used to test fetal lung maturity (FLM), including lamellar body count, lecithin-sphingomyelin ratio, or TDx fetal lung maturity assay II test (TDx-FLM II; Abbott Laboratories, Abbott Park, IL),[3,4] are performed in amniotic fluid and, consequently, require an invasive procedure. Over the last 30 years, the prediction of lung maturity by noninvasive ultrasound methods has been extensively explored. Earlier studies comparing fetal lung echogenicity with the placenta,[5-7] fetal gut,[8] or liver[9,10] demonstrated ultrasonographic changes associated with fetal lung maturation. Later studies have used approaches based on quantitative ultrasound analysis to explore the potential of ultrasound to predict fetal lung maturation.

Quantitative ultrasound is based on applying processing methods to ultrasound images. This allows extracting quantitative features and potentially identifying subclinical tissue differences that escape subjective inspection. The use of quantitative ultrasound analysis has previously been investigated for medical diagnostic applications, including breast cancer[11,12] and liver disease.[13-15] A few studies have explored the assessment of fetal lung maturity by quantitative ultrasound tissue characterization with different methodologies,[16-20] with all reporting differences in lung texture features along gestation, individually or compared with the reflection pattern of other organs such as the liver. However, controversial results, limitations in the sample size, and difficulties in recording the parameters proposed with standard means have prevented further

From the Department of Maternal-Fetal Medicine, Institut Clínic de Ginecologia, Obstetrícia i Neonatologia, Hospital Clinic, Universitat de Barcelona (Drs Palacio, Cobo, Martínez-Terrón, and Gratacós); the Fetal and Perinatal Medicine Research Group, Institut d'Investigacions Biomediques August Pi i Sunyer (Drs Palacio, Cobo, Rattá, and Gratacós); Centro de Investigación Biomédica en Red de Enfermedades Raras (Drs Palacio and Gratacós); and the Transmural Biotech, S.L. (Ms Bonet-Carné and Mr Amat-Roldán), Barcelona, Spain.

Received April 6, 2012; revised August 18, 2012; accepted September 26, 2012.

T.C. was supported by a Río Hortega Grant from the Instituto de Salud Carlos III in Spain (CM09/00213). Financial support to E.B.-C. and Transmural Biotech SL was provided by the Comissionat per a Universitat i Recerca del Departament d'Innovació, Universitats i Empreses de la Generalitat de Catalunya. This project was also supported by The People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7 under REA Grant Agreement no. 217911 (u-Volumes).

Transmural Biotech has collaborative research agreements with the Hospital Clinic for the joint development and potential exploitation of imaging biomarkers. However, Transmural Biotech did not provide funding for this study.

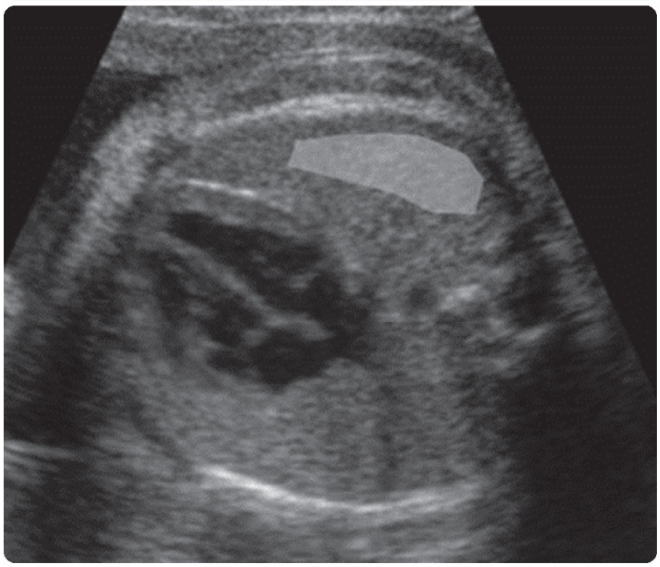
The authors report no conflict of interest.

Partial results of this study were presented orally at the 32nd annual meeting of the Society for Maternal-Fetal Medicine, Dallas, TX, Feb. 6-11, 2012.

Reprints not available from the authors.

0002-9378/$36.00 • © 2012 Published by Mosby, Inc. • http://dx.doi.org/10.1016/j.ajog.2012.09.027

FIGURE 1
View of fetal lung acquisition



Cross-sectional view of the fetal thorax with the manually delineated region of interest in the proximal lung.

Palacio. Fetal lung maturity using AQUA texture descriptors. Am J Obstet Gynecol 2012.

development into clinically applicable solutions.

Among various quantitative imaging methods, texture analysis is proposed as a powerful approach to extract quantitative features directly from medical images.[19,21] We have previously developed an automatic quantitative ultrasound analysis (AQUA) algorithm, which is invariant under illumination changes and does not use direct gray level from the image or tissue references. The method estimates texture features based on conditional random fields so that image texture features converge robustly to identify different tissue characteristics independently of the overall acquisition context, including scanner settings.[22] In a previous study, we demonstrated that AQUA could extract features from fetal lung ultrasound images, showing a strong correlation with gestational age.[23]

In this study, we evaluated the performance of the AQUA texture extractor to predict fetal lung maturity, as assessed by the TDx-FLM II test in amniotic fluid.

Materials and Methods

The study was carried out at the Maternal-Fetal Medicine Department at Hospital Clinic in Barcelona from October 2010 to March 2011. The study population included singleton pregnancies with gestational ages between 24.0 and 41.0 weeks.

Patients were selected from among pregnant women undergoing amniocentesis to assess fetal lung maturity for medical indications or to exclude infection in cases of preterm labor or preterm rupture of membranes. In addition, women scheduled for elective term cesarean section for obstetrical indications were also included. In this latter group, the amniotic fluid was obtained intraoperatively after hysterotomy. Multiple pregnancies and structural/chromosomal anomalies were considered noneligible for this study.

The study protocol was approved by the local Ethics and Institutional Review Board (ID 3823-2007), and all patients provided written informed consent. In all pregnancies gestational age was calculated based on the crown-rump length at first-trimester ultrasound. Maternal and neonatal outcomes were recorded. Descriptive statistics were performed with the SPSS 18.0 statistical software (SPSS Inc, Chicago, IL).

Fetal lung maturity test

A fetal lung maturity was tested in amniotic fluid using the surfactant-to-albumin ratio by fluorescence polarization (TDx-FLM II).[4] Gestational age–specific cutoff values were used to classify the results as mature or immature. This adjustment for gestational age results in a significant improvement in the capacity of the test to predict the respiratory distress syndrome and may simplify clinical decisions.[24]

Image acquisition and delineation of lung tissue

Ultrasound images were obtained the same day of amniotic fluid collection in all cases. A semilateral transverse 4 cardiac-chamber view plane for lung image acquisition was performed using Siemens Sonoline Antares (Siemens Medical Systems, Malvern, PA) and Voluson 730/780 Pro (GE Medical Systems, Milwaukee, WI) ultrasound equipment, all equipped with a curved linear transducer with a frequency range from 3 to 7.5 MHz. Equipment settings were adjusted at the discretion of the clinician performing the ultrasound to obtain the optimal image quality according to clinical criteria.

Images were digitally collected in the original Digital Imaging and Communication in Medicine format and stored for off-line analysis. Manual delineation was performed in the proximal lung (according to the distance to the ultrasound transducer) with a custom-made program with a Graphic User Interface tool developed with MATLAB R2007b (version 7.5.0.342; MATLAB; The MathWorks Inc, Natick, MA). Care was taken

to delineate only lung tissue as illustrated in Figure 1.

AQUA and machine learning prediction

AQUA software was applied to extract the texture features from the manually delineated lungs. Once AQUA was executed, each image became a set of descriptors (15,300 per delineated lung), which include a high amount of textural-related information per delineation. Dimensionality reduction and machine learning-based models were then applied to select a set of descriptors that correlated robustly with the results of the FLM test in amniotic fluid. For this purpose, genetic algorithms[25] and support vector machines[26] were applied.

Genetic algorithms are computational techniques used to resolve the problem of time that would need an exhaustive algorithm to select every possible combination of the 15,300 descriptors to create and validate a predictor. Support vector machines (SVMs) are supervised machine learning algorithms that can be used to predict or classify sets of labeled objects. These SVMs work in 2 phases. Phase 1 is called training, and at this stage, a subset of objects (represented by their feature vectors and the class or label they belong to) are given as inputs to the system. Using the feature vectors, the SVM builds a mathematical model. In phase 2, which is called testing or validation, the model automatically determines the class or label of the new objects. The success rate in this stage is measured, and therefore, the accuracy of the model created is expressed.

In the present study, genetic algorithms identified a set of 31 critical descriptors, which were used in subsequent steps. The support vector machines were then used to develop a mathematical model that combined the set of 31 descriptors to evaluate their theoretical predictive value of the results of lung maturity in amniotic fluid. The association and predictive value were tested with a strategy consisting in the definition of 2 randomly created datasets, containing the same number of women, "A" and "B" with a similar number of mature and immature fetuses. The set of 31 descriptors was used to train a machine learning model.

The accuracy of the model proposed was double tested: the first model was trained with data set B and validated in A. Subsequently, the groups were interchanged and a second model was trained with dataset A and validated in B using the same 31 descriptors. The mean accuracy resulting from the 2 validation tests was calculated and stored as the fitness.

TABLE
Baseline characteristics of the population studied (n = 103)

Variable	Mean (SD) or n (%)
Maternal age, y	31.4 (5.6)
Nuliparity	60 (58.3)
Gestational age at amniocentesis, wks	32.2 (4.5)
24 to <28	19/103 (18.5)
28 to <32	26/103 (25.2)
32 to <34	21/103 (20.4)
34 to <37	14/103 (13.6)
≥37	23/103 (22.3)
Indication for amniocentesis	
Preterm labor	41/103 (39.8)
Preeclampsia/IUGR	8/103 (7.8)
PPROM	29/103 (28.2)
Methrorrhagia	4/103 (3.9)
Elective	21/103 (20.4)
Gestational age at delivery, wks	34.3 (4.1)
Delivery <37	63/101 (62.4)
Delivery <34	43/101 (42.6)
Delivery <28	6/101 (5.9)
Birthweight, g	2322 (976)
pH AU <7.10	3/73 (4.1)[a]
Apgar 5 minutes <7	4/101 (4.0)
NICU admission	56/101 (54.5)
Respiratory distress syndrome	17/101 (16.8)
Days at NICU, d	19.8 (21.0)[b]
Neonatal death	5/101 (5.0)

Values were expressed in mean (SD) or n (percentage) when appropriate. Two women and their babies were lost to follow-up. *IUGR*, intrauterine growth retardation; *NICU*, neonatal intensive care unit; *PPROM*, premature preterm rupture of membranes.

[a] pH AU was not available for all babies; [b] Days at NICU of those admitted (n = 56).

Palncio. Fetal lung maturity using AQUA texture descriptors. Am J Obstet Gynecol 2012.

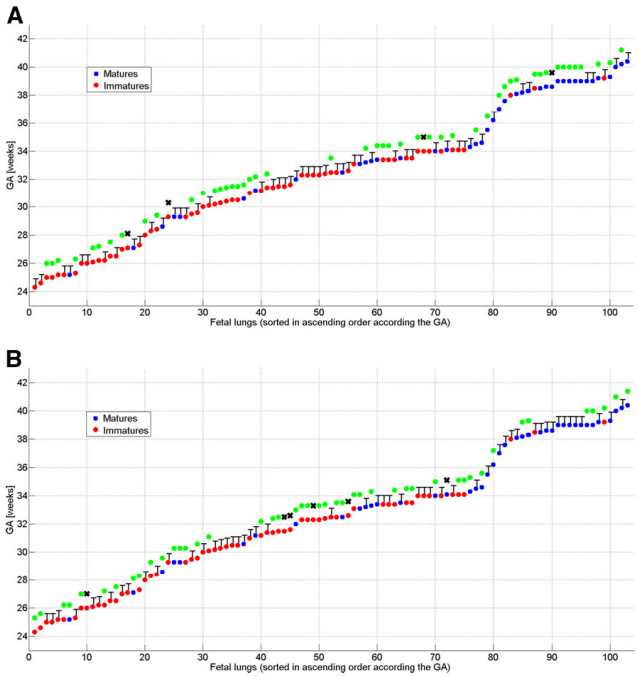
RESULTS

A total of 103 women from 24.0 to 41.0 weeks of gestation were included in the study. The baseline characteristics of the study population are described in the Table. Considering gestational age cutoff values of the TDx-FLM test,[24] 62 and 41 cases were classified as immature or mature, respectively.

The database was split randomly into 2 similar datasets or subgroups. Subgroup A included 31/21 immature/mature lungs and subgroup B included 31/20 immature/matures lungs. These data sets were used for training and testing the algorithm as described in the *Material and Methods* section and Figure 2.

Figures 2, A and B, show the prediction rates attained with the combination of genetic algorithm and support vector

FIGURE 2
Accuracy of the prediction of fetal lung maturity



Blue and *red dots* in both charts represent the 103 women studied (X axis). *Blue* indicates that the TDx-FLM result was mature, whereas *red* indicates immature. Women were randomly split into two subgroups **(A and B)**, containing a similar proportion of immature/mature results. **A,** Subgroup B was used for training the texture algorithm (indicated with a T above each dot), which was then tested to predict the outcome of TDx-FLM in amniotic fluid in cases belonging to subgroup A. A *green circle* and a *black cross* indicate a correct or a wrong prediction, respectively. The accuracy for a correct prediction was 92.31%. **B,** Subgroup A was used for training and subgroup B for validation of the predictive value. Again, correct (*green circles*) and wrong predictions (*black crosses*) are plotted. The accuracy for a correct prediction was 88.24%.

TDx fetal lung maturity assay (TDx-FLM II); Abbott Laboratories, Abbott Park, IL.

Palacio. Fetal lung maturity using AQUA texture descriptors. Am J Obstet Gynecol 2012.

machine approaches in the 103 women, using the 31 previously selected AQUA descriptors. The predictive accuracy ranged from 88.24% to 92.31%, for an average accuracy of 90.27%. The sensitivity was 95.1% and the specificity was 85.7% to predict fetal lung maturity, as assessed by TDx-FLM II.

COMMENT

This study provides evidence that AQUA can extract robust quantitative image features from lung ultrasound images, which strongly correlate with fetal lung maturity as assessed by a standard fetal maturity test as TDx-FLM II. These findings open the possibility to explore the introduction of non-invasive techniques into clinical practice to test fetal lung maturity. In view of the increasingly recognized importance of respiratory morbidity[1,2] and the growing numbers of late preterm pregnancies undergoing elective delivery, avoidance of the need for invasive techniques may have a tremendous impact on the clinical management of these cases.

These results are in line with previous studies showing that ultrasound images contain nonvisible information that can be extracted for clinical purposes. This notion has already been demonstrated for breast and liver disease,[11-15] but the results in fetal lung ultrasound analysis have remained nonconclusive.

Among recent studies, some failed to identify differences in the patterns of features assessed with quantitative ultrasound.[16,17] Prakash et al[18] evaluated the ability to predict lung maturity, using gestational age as a surrogate, by means of ultrasound parameters and reported a classification accuracy ranging from 73% to 96%. In a study by Tekesin et al,[19] the mean gray value of the fetal lung showed a changing pattern with fetal lung development. However, no significant differences were observed above 32 weeks' gestation, precluding its use in clinical practice. Finally, Serizawa and Maeda[20] published a form of tissue characterization named ultrasonic gray level histogram width. This method, combined with gestational age, predicted the occurrence of RDS with a sensitivity of 0.96 and a specificity of 0.72, which was comparable with invasive amniotic fluid tests. The method required examination of the lung and liver, and it was designed in a relatively reduced sample of 22 and 25 fetuses with and without RDS, respectively, but the results were promising.

The method tested here did not use direct gray level or other tissue references such as the liver. This may represent a substantial advantage for clinical use as compared with the other quantitative analyses discussed in the previous text. In addition, the AQUA texture extractor has been shown to be unaffected by moderate changes in the acquisition settings as well as by the use of different ultrasound transducers, thereby facilitating its eventual clinical application.[23] These particular properties make appropriate viewing of the fetal lung easy to obtain and further simplify its inclusion in clinical practice.

In view of recent evidence stressing the impact of respiratory morbidity in even mild degrees of prematurity, determining

the status of lung maturity has become a need in a substantial number of late preterm pregnancies before the induction of labor or elective cesarean section is indicated for nonurgent reasons.[1,2] This entails performing invasive procedures in thousands of pregnancies yearly, whereas a noninvasive method is not available in the clinical practice and only gestational age is a reliable approach to fetal lung maturity status. In addition, aside from the obvious implications for management costs, and despite the relatively low rate of complications after third-trimester amniocentesis, 0.7-3% of events in otherwise uncomplicated pregnancies must be taken into account.[27,28]

This study has some limitations. Although the study provides proof of principle that assessment of fetal lung maturity could be achieved by means of a noninvasive method, the outcome of clinical interest is, obviously, the risk of respiratory morbidity. To provide quick assessment of the potential performance of the method we decided to use amniotic fluid results, and therefore, we used a surrogate of fetal lung maturity. However, we acknowledge that these results require confirmation with the clinical outcomes of interest (ie, respiratory morbidity and RDS). Taking into account that TDx-FLM II also has a limited sensitivity and specificity (89% and 83%, respectively)[24] for the prediction of RDS, it would be of great interest to assess whether noninvasive evaluation of the fetal lung texture might provide more precise information. Ongoing clinical studies are currently evaluating the ability of the texture extractor reported here to predict neonatal respiratory outcomes. Because the prevalence of these outcomes in advanced gestational ages may be lower than 5-10%,[1,2] large sample sizes are required. Such large studies will allow defining more accurate predictive algorithms adjusted by gestational age intervals.

Another limitation is that to ensure the reliability of the models created, the quantity of descriptors selected by the algorithm should have been lower, given the reduced number of subjects under study. Nonetheless, we plan to overcome this issue in future research by enlarging the databases with a higher number of subjects.

In summary, this study provides evidence that AQUA descriptors contain information that robustly correlates with the results of fetal lung maturity tests in amniotic fluid. These findings support further research on the value of the AQUA texture extractor as a noninvasive quantitative image biomarker to predict a relevant clinical outcome such as RDS. Should these results be confirmed in larger studies, the need for amniocentesis to assess fetal lung maturity might be avoided in clinical practice. ■

REFERENCES

1. Spong CY, Mercer BM, D'Alton M, Kilpatrick S, Blackwell S, Saade G. Timing of indicated late-preterm and early-term birth. Obstet Gynecol 2011;118:323-33.
2. Teune MJ, Bakhuizen S, Bannerman CG, et al. A systematic review of severe morbidity in infants born late preterm. Am J Obstet Gynecol 2011;205:374.e1-9.
3. Wijnberger LD, de Kleine M, Voorbij HA, et al. Prediction of fetal lung immaturity using gestational age, patient characteristics and fetal lung maturity tests: a probabilistic approach. Arch Gynecol Obstet 2010;281:15-21.
4. Grenache DG, Gronowski AM. Fetal lung maturity. Clin Biochem 2006;39:1-10.
5. Grannum PA, Berkowitz RL, Hobbins JC. The ultrasonic changes in the maturing placenta and their relation to fetal pulmonic maturity. Am J Obstet Gynecol 1979;133:915-22.
6. Harman CR, Manning FA, Stearns E, Morrison I. The correlation of ultrasonic placental grading and fetal pulmonary maturation in five hundred sixty-three pregnancies. Am J Obstet Gynecol 1982;143:941-3.
7. Golde SH, Tahilramaney MP, Platt LD. Use of ultrasound to predict fetal lung maturity in 247 consecutive elective cesarean deliveries. J Reprod Med 1984;29:9-11.
8. Zilianti M, Fernandez S. Correlation of ultrasonic images of fetal intestine with gestational age and fetal maturity. Obstet Gynecol 1983; 62:569-73.
9. Fried AM, Loh FK, Umer MA, Dillon KP, Kryscio R. Echogenicity of fetal lung: relation to fetal age and maturity. AJR Am J Roentgenol 1985;145:591-4.
10. Feingold M, Scollins J, Cetrulo CL, Koza D. Fetal lung to liver reflectivity ratio and lung maturity. J Clin Ultrasound 1987;15:384-7.
11. Chen DR, Chang RF, Kuo WJ, Chen MC, Huang YL. Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks. Ultrasound Med Biol 2002;28:1301-10.
12. Wan C, Du J, Fang H, Li F, Wang L. Evaluation of breast lesions by contrast enhanced ultrasound: qualitative and quantitative analysis. Eur J Radiol 2012;81:e444-50.
13. Hartman PC, Oosterveld BJ, Thijssen JM, Rosenbusch GJ. Variability of quantitative echographic parameters of the liver: intra- and interindividual spread, temporal- and age-related effects. Ultrasound Med Biol 1991;17:857-67.
14. Kadah YM, Farag AA, Zurada JM, Badawi AM, Youssef AM. Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images. IEEE Trans Med Imaging 1996;15:466-78.
15. Icer S, Coskun A, Ikizceli T. Quantitative grading using grey relational analysis on ultrasonographic images of a fatty liver. J Med Syst 2012;36:2521-8.
16. Sohn C, Stolz W, Bastert G. Diagnosis of fetal lung maturity by ultrasound: a new method and first results. Ultrasound Obstet Gynecol 1991;1:345-8.
17. Maeda K, Utsu M, Yamamoto N, Serizawa M. Echogenicity of fetal lung and liver quantified by the grey-level histogram width. Ultrasound Med Biol 1999;25:201-8.
18. Prakash KN, Ramakrishnan AG, Suresh S, Chow TW. Fetal lung maturity analysis using ultrasound image features. IEEE Trans Inf Technol Biomed 2002;6:38-45.
19. Tekesin I, Anderer G, Hellmeyer L, Stein W, Kuhnert M, Schmidt S. Assessment of fetal lung development by quantitative ultrasonic tissue characterization: a methodical study. Prenat Diagn 2004;24:671-6.
20. Serizawa M, Maeda K. Noninvasive fetal lung maturity prediction based on ultrasonic gray level histogram width. Ultrasound Med Biol 2010;36:1998-2003.
21. Castellano G, Bonilha L, Li LM, Cendes F. Texture analysis of medical images. Clin Radiol 2004;59:1061-9.
22. Tenorio V, Bonet-Carne E, Botet F, Marques F, Amat-Roldan I, Gratacos E. Correlation between a semiautomated method based on ultrasound texture analysis and standard ultrasound diagnosis using white matter damage in preterm neonates as a model. J Ultrasound Med;30:1365-77.
23. Cobo T, Bonet-Carne E, Martinez-Terrón M, et al. Feasibility and reproducibility of fetal lung texture analysis by automatic quantitative ultrasound analysis and correlation with gestational age. Fetal Diagn Ther 2012;31:230-6.
24. Bennasar M, Figueras F, Palacio M, et al. Gestational age-specific cutoff levels of TDx-FLM II for the prediction of neonatal respiratory distress syndrome. Fetal Diagn Ther 2009;25:392-6.
25. Holland J. Adaptation in natural and artificial systems. Ann Arbor, MI: University of Michigan Press; 1975.
26. Cortes C, Vapnik V. Support-vector networks. Machine Learning 1995;20:273-97.
27. Zalud I, Janas S. Risks of third-trimester amniocentesis. J Reprod Med 2008;53:45-8.
28. Gordon MC, Narula K, O'Shaughnessy R, Barth WH Jr. Complications of third-trimester amniocentesis using continuous ultrasound guidance. Obstet Gynecol 2002;99:255-9.

4.3. STUDY 3

Quantitative ultrasound texture analysis of fetal lungs to predict neonatal respiratory morbidity

E. Bonet-Carne, M. Palacio, T. Cobo, A. Perez-Moreno, M. Lopez, JP. Piraquive, JC. Ramirez, F. Marques, E. Gratacos. Ultrasound in Obstetrics & Gynecology.

Status: accepted, 5th June 2014. Ref: UOG-2014-0288.R1

Journal Impact factor: 3.557

Quartile: 1st

Results from this study have been presented in the following congresses:

1. Bonet-Carne E, Palacio M, Cobo T, Perez-Moreno A, Lopez M, Piraquive JP, Ramirez JC, Marques F, Gratacos E. Validation of a quantitative ultrasound texture analysis of fetal lungs to predict neonatal respiratory morbidity. ECPM. June 4-7, 2014. Firenze, Italy.

2. Bonet-Carne E, Palacio M, Cobo T, Perez-Moreno A, Lopez M, Piraquive JP, Ramirez JC, Marques F, Gratacos E. Quantitative ultrasound texture analysis of fetal lungs to predict neonatal respiratory morbidity. ISUOG. September 14-17, 2014. Barcelona, Spain.

QUANTITATIVE ULTRASOUND TEXTURE ANALYSIS OF FETAL LUNGS TO PREDICT NEONATAL RESPIRATORY MORBIDITY

Elisenda Bonet-Carne[1], MSc; Montse Palacio[2,3], MD, PhD; Teresa Cobo[2,3], MD, PhD; Alvaro Perez-Moreno[1], MSc; Marta Lopez[2], MD; Juan Pablo Piraquive[2], MD; Juan Carlos Ramirez[2], MD; Ferran Marques[4], PhD; Eduard Gratacos[2,3], MD, PhD

[1] Transmural Biotech SL., Barcelona, Spain.

[2] Department of Maternal-Fetal Medicine, ICGON, Hospital Clinic of Barcelona, University of Barcelona, Spain.

[3] Fetal and Perinatal Medicine Research Group (IDIBAPS and CIBERER), Barcelona, Spain.

[4] Universitat Politecnica de Catalunya BarcelonaTECH, Barcelona, Spain.

Corresponding author

Elisenda Bonet-Carne
Transmural Biotech S.L.
Sabino de Arana 38, 1 1
08028 Barcelona
Spain
+34 934 487 496
e.bonetcarne@transmuralbiotech.com

Key words:

Fetal Lung Maturity, Neonatal Respiratory Morbidity, Image quantitative analysis, Texture analysis

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/uog.13441

ABSTRACT

OBJECTIVE: To develop and evaluate the performance of a novel method to predict neonatal respiratory morbidity based on quantitative analysis of fetal lung by ultrasound.

METHODS: A huge number of non-clinical and fetal lung ultrasound images were used to develop a computerized method based on texture analysis and machine learning algorithms, trained to predict neonatal respiratory morbidity risk on fetal lung ultrasound images. The method, termed quantitative ultrasound fetal lung maturity analysis (quantusFLM™), was then validated blindly in 144 fetuses delivering at 28.0-39.0 weeks' gestational age. Lung ultrasound images in DICOM format were obtained within 48 hours of delivery and the ability of the software to predict neonatal respiratory morbidity, defined as either respiratory distress syndrome or transient tachypnea of the newborn, was determined.

RESULTS: Mean gestational age at delivery was 36.0 (SD 3.3) weeks. There were 29/144 (20.1%) events of neonatal respiratory morbidity. Quantitative texture analysis predicted neonatal respiratory morbidity with a sensitivity, specificity, positive predictive value and negative predictive value of 86.2%, 86.9%, 62.5%, and 96.2% respectively.

CONCLUSIONS: Quantitative ultrasound fetal lung maturity analysis predicted neonatal respiratory morbidity with an accuracy comparable to current tests using amniotic fluid.

INTRODUCTION

Neonatal respiratory morbidity, defined as respiratory distress syndrome or transient tachypnea of the newborn, is the leading cause of mortality and morbidity associated with prematurity[1,2]. Neonatal respiratory morbidity is not restricted to very preterm births and remains high among late-preterm and early-term infants born before 39 weeks of gestation[3-5]. Fetal lung maturity (FLM) is mainly determined by pulmonary surfactant and it can only be assessed with laboratory tests in amniotic fluid[6-10]. The need of amniocentesis has resulted in a decline in the use of this information clinically. Non-invasive prediction of FLM on fetal lung ultrasound images has been attempted for 25 years by means of gray level measurements, lung tissue motion and relative features of lung to placental or liver images, among others[11-16]. These studies showed a good correlation with respiratory morbidity, but the diagnostic accuracy was insufficient for clinical use.

Over years powerful quantitative techniques for ultrasound image analysis have been developed thanks to improvements in computer capacity and image resolution[17]. Specifically, texture analysis approaches are computerized methods that can analyze medical images and identify subtle changes in the aspect, or texture, which are invisible to the human eye[18]. These textural patterns can then be used to train algorithms to predict clinical information. Recent studies have demonstrated that texture analysis of fetal lung ultrasound images are able to identify patterns of features that strongly correlated with gestational age[19], or with the results of FLM tests in amniotic fluid[20]. These studies provided a proof of concept of the potential of texture based methods, but common problems of other quantitative imaging methods remained, such as the lack of robustness to blind testing due to variable acquisition conditions. Up to date the ability of texture analysis of fetal lung ultrasound images to blindly predict the risk of neonatal respiratory morbidity has not been demonstrated.

To address these limitations, we have developed a new method termed quantitative ultrasound fetal lung maturity analysis (quantusFLM™), which combines various image texture

extractors and machine learning algorithms. In this study we describe the basic principles of this novel method, and the results of a validation study to blindly predict the risk of neonatal respiratory morbidity.

METHODS

A novel quantitative ultrasound fetal lung maturity analysis was planned to offer automatic assessment of neonatal respiratory morbidity risk using an ultrasound image of the lateral axial transverse section of the fetal thorax at the level of the 4-chamber section of the fetal heart. quantusFLM™ was specifically designed to be composed of two modules, a texture feature extractor and a classifier. The latter uses information from the extracted features to assess the risk of respiratory morbidity (Figure 1). Finally, the algorithm performance was validated using blind samples in order to evaluate its potential use in clinical practice.

Development of a textural feature extractor module

The first aim was to develop a textural feature extractor module which showed the highest robustness when tested using images acquired under different conditions. This module is used to compact all the information contained in an image (or a region of an image) to a few features which contain relevant information. For a given application, a feature extractor module obtains the relevant information of an image. This way, it allows to represent the information conveyed by the pixels of an image by a much more compact, application-dependent set of values. Features should be invariant to clinical acquisition conditions such as changes in lighting, shadows, rotation or resolution due to the lack of accessibility and control over the fetus position during the image acquisition procedure. The region of interest for the analysis is the region corresponding to fetal lung tissue, which is manually delimited by the operator. Hence, the feature extractor module must be invariant to region of interest shape and size.

Different existing methods for image texture extraction based on wavelets, co-occurrence matrix, histogram gradients, binary patterns, scale and rotation invariant features methods and standard first and second order statistic measures were evaluated[21-26]. Our previously reported method for feature extraction, AQUA[27], was also tested in these series of

experiments, but was discarded after the first rounds of experiments due to poor robustness with respect to image variability. For each different extractor method, the process may be considered to be robust when the robustness is demonstrated within a certain range in the acquisition conditions that are not critical since certain acquisition parameters can be controlled to some degree (for instance, minimum ROI size or shadows). Moreover, to ensure the feasibility of quantusFLM™, the method should not require an area bigger than 400 pixels to extract robust features.

OUTEX[28] and PHOTEX[29] databases composed by texture images acquired using different controlled parameters as illumination, spatial resolution and rotation angles were used for this purpose. A total of 13.171 images were used for illumination and rotation experiments and 11.178 were used for the resolution ones. Moreover, each of these images was automatically divided in 25 non-overlapping equal size images and in 30 overlapping different size regions to test the robustness versus the region of interest. Affinity and measures such as correlation, city block, Chebyshev or Euclidean distance were used to evaluate the performance of the tested methods.

A newly developed feature extractor was then further refined and tuned up using real fetal lung ultrasound scans. The images used were a set of 957 samples obtained in a previous study[19]. Using real lung images was critical to determine the final combination of textural features in a manner that maximizes the robustness of the extracted features under those variations occurring in fetal lung ultrasound acquisition in clinical practice. Over 7 billion computerized experiments were performed to construct the quantusFLM™ feature extractor module which combines the features that demonstrated to be invariant to geometric and photometric transformations (in acquisition condition ranges).

Development and training of a classification algorithm

The classification algorithm must combine the features obtained with the first module to predict the occurrence of neonatal respiratory morbidity. Different supervised machine learning methods including regression models, classification trees and neural networks were trained[30-34] to combine the appropriate subset of features to identify those that contain relevant information to classify fetuses with high or low respiratory morbidity risk. The final quantusFLM™ classification algorithm is a sequence of various machine learning steps that combines the textural features selected (previously obtained with the extractor module) and clinical data (gestational age) to estimate the prediction. The parameters of this model were estimated and tuned up using 390 samples recruited prospectively among singleton pregnancies between 24.6 and 41.6 weeks gestation. The cases used for training the algorithm were not used in any previous study. Lung ultrasound scans were obtained within 48 hours of delivery, and the occurrence of neonatal respiratory morbidity was recorded. Ethical board review approval was obtained (ID 2013/8892) and informed consent was obtained in all cases. The final algorithm combined hundreds of textural features associated with the occurrence of neonatal morbidity. The theoretical diagnostic performance obtained was as following: 87% of accuracy [95% Confidence interval 82-90%], Sensitivity of 91% [77-98%], Specificity of 86% [82-90%], Positive Predictive Value of 47% [35-59%] and Negative Predictive Value of 98% [96-99%].

Validation

The validation was carried out at the Maternal-Fetal Medicine Department at Hospital Clinic Barcelona. The study protocol was approved by the local Ethics and Institutional Review Board (ID 2013/8892), and all patients provided written informed consent. The validation group consisted in new cases recruited prospectively for the purposes of this study until the sample size fixed for the validation according to the study design (n=150) was reached. In this validation study, the population included singleton pregnancies with gestational ages between

28.0 and 39.0 [weeks.days]. Validation samples were not used for the generation of the algorithm or for other previous study.

Pregnant women who were scheduled for or at risk of delivery within 48 hours were considered eligible. This included patients scheduled for elective delivery or cesarean section, or patients with pregnancy complications including threatened preterm labor, rupture of membranes or various pregnancy complications, mainly preeclampsia and intrauterine growth restriction. Multiple pregnancies and those with fetal structural or chromosomal anomalies were considered non-eligible. In all pregnancies gestational, age was calculated based on the crown-rump length at first-trimester ultrasound. At enrolment, an ultrasound scan of the fetal lungs was obtained but only patients actually delivering within 48 hours were finally included in the study. Maternal baseline features and neonatal outcomes were recorded prospectively. Neonatal respiratory morbidity was defined as respiratory distress syndrome or transient tachypnea of the newborn. Respiratory distress syndrome was defined as respiratory symptoms (eg, grunting, flaring, tachypnea, retractions) or need for supplemental oxygen requirement, together with compatible chest radiography findings and NICU admission for respiratory support and transient tachypnea of the newborn was determined by chest radiography impression and clinical diagnosis established by the clinician in charge[3].

Ultrasound images were obtained using a pre-established acquisition protocol by clinicians at the Day Assessment Unit of the Department. A lateral axial transverse section of the fetal thorax at the level of the 4-chamber section of the fetal heart was acquired (Figure 2). Images were acquired either with Siemens Sonoline Antares (Siemens Medical Systems, Malvern, PA) or Voluson 730/830 Pro (GE Medical Systems, Milwaukee, WI) ultrasound equipments, with curved linear transducers of 3-7.5 MHz. For the purposes of the validation, the preset used to perform the ultrasound did not contain any type of post-processing options, such as image smoothing options, nor any Doppler measurements, calipers or pointers. The use of tissue harmonic imaging, and adjustment of image settings such as gain, frequency and time-gain

compensation, were at the discretion of the physician performing the ultrasound. Images were digitally collected in the original Digital Imaging and Communication in Medicine (DICOM) format and stored for off-line analysis.

All DICOM images were inspected for quality control by two members of the research team (A.P-M., E.B-C.). Images were discarded if the lung area contained calipers or obvious acoustic shadows created by bony structures. Images passing quality criteria were processed by clinicians (T.C., M.P.) using a specific software that allowed delineating the regions of interest (ROI) for analysis. Delineation aimed to include the largest possible area of the fetal lung proximal to the transducer (Figure 3). Care was taken to delineate only lung tissue, avoiding the heart, great vessels and surrounding areas. In order to ensure the robustness of the analysis, the software did not accept a ROI containing less than 400 pixels. The same software was used to perform the automatic analysis. Neonatal outcomes were collected (J.P.) and data analysis was performed by one member of the research team (J.C.) who did not have intervened in the other steps of the study.

The sample size for the study was established arbitrarily at 150 subjects, since it was estimated that this would allow including 25-35 cases of neonatal respiratory morbidity. Descriptive statistics were performed with R language[35].

RESULTS

A total of 150 images from different pregnancies were analyzed in this study. After analysis for quality criteria, six (3.7%) images were excluded, which left a total of 144 images for analysis. Baseline characteristics of the study population are described in Table 1 and perinatal outcomes are displayed in Table 2. Mean gestational age at delivery was 36.0 (SD 3.3) weeks. There was a total of 29/144 (20.1%) events of neonatal respiratory morbidity, of which 15/29 (51.7%) corresponded to respiratory distress syndrome and 14 to transient tachypnea of the newborn.

Quantitative ultrasound fetal lung maturity analysis predicted neonatal respiratory morbidity with a sensitivity, specificity, positive predictive value and negative predictive value of 86.2%, 86.9%, 62.5%, and 96.2% respectively. Performance stratified per gestational age showed similar values for cases delivering at 28.0 to 33.6 as compared to those delivering at 34.0-39.0 weeks' gestational age (Table 3).

DISCUSSION

This study provides evidence that quantitative texture analysis of lung ultrasound images predicts neonatal respiratory morbidity with similar accuracy to current tests in amniotic fluid.

Previous studies exploring quantitative assessment of fetal lung ultrasound to predict FLM used a variety of techniques. Prakash et al.[12] compared ratios of fetal lung to liver image feature values, with reported accuracies ranging from 73% to 96%. La Torre et al.[15] correlated accurately several patterns of fetal breathing movements with fetal lung maturity tests. Tekesin et al.[13] evaluated the mean gray value of fetal lungs, showing a changing pattern with fetal lung development. However, no significant differences were observed above 32 weeks' GA. Later, Serizawa and Maeda[14] tried to predict fetal lung immaturity by comparison of the ultrasonic gray level histogram width (GLHW) of the fetal lung and liver, in 22 fetuses with respiratory distress syndrome and 25 controls. In that study, GLHW combined with gestational age identified differences associated with the occurrence of respiratory distress syndrome, with a theoretical sensitivity of 0.96 and specificity of 0.72. In general, previous studies demonstrated a correlation between quantitative image analysis and FLM, but either diagnostic accuracy was low or results were not blindly validated, preventing its implementation into clinical practice.

Up to the authors' knowledge, this is the first study reporting blind validation of quantitative imaging analysis software specifically designed to predict neonatal respiratory morbidity. A difference with previous studies is that the software was constructed using a large amount of theoretical and real images in order to address the main difficulties of quantitative analysis in real practice. The first difficulty is the variability in image acquisition, which may create false associations in small sample size studies. The texture extractor here used works with textural features selected after millions of computer tests, which are highly correlated with the ultrasound texture of the fetal lung, and highly robust to changes in the angle of insonation and to the adjustments of image settings occurring in real practice of fetal ultrasound. In these

respects, this approach represented a huge step forward and a completely new approach in relation with our previously reported methodology AQUA[27]. While AQUA was a useful method to demonstrate proof of principle, as a texture extractor it had important limitations in terms of robustness, which were addressed by the newly developed feature extractors of quantusFLM™. The second challenge of quantitative analysis is automation. Thus, once robust textures have been selected, automated classifiers are required, and this entails very large databases of real cases, to avoid spurious associations in the selected textural features. Again, quantusFLM™ represents a completely new approach, since it incorporates highly robust machine learning classification algorithms that have been trained extensively, which allowed a fully automated performance and immediate clinical use.

The proposed algorithm performance is comparable to that reported with currently used tests in amniotic fluid[36-40]. Thus, average reported sensitivity and specificity of Lecithin/Sphingomyelin Ratio was 74% (range 48-96%) and 98% (range 81-100%), respectively. Lamellar Body Count has a reported sensitivity and specificity of 86% (71-100%) and 86% (60-100%), the Phosphatidyglycerol test 91% (86-94%) and 72% (67-79%) and surfactant/Albumin Ratio 90% (83-96%) and 76% (95-88%), respectively. The average sample size used in these studies was 167 (ranging from 28 to 301), which is similar to this study.

The results of this study open the possibility of using non-invasive approaches for prenatal prediction of FLM. Despite improvements in clinical practice such as prenatal corticoids and postnatal surfactant, respiratory morbidity remains a leading cause of neonatal morbidity and mortality in late preterm (34 0/7–36 6/7 weeks of gestation) and even in early-term (37 0/7–38 6/7 weeks of gestation)[3,41]. It is clear that for some indications, delivery should occur regardless of FLM results. However, there is an open debate about the value of FLM testing in the decision-making process for relative indications or borderline clinical situations in which late preterm or early-term delivery may seem a reasonable option but delivery could be postponed if fetal lung immaturity assessed[2,42-44]. In fact, recent data shows that

approximately 1 in 15 neonates were delivered late preterm for "soft" indications[45,46]. Determining the risk of FLM without the need for an invasive technique might have a tremendous impact in the clinical management of such cases. Aside from economic implications, avoiding the need of amniocentesis would be associated with less patients' discomfort and related complications, and controversies about indications for FLM assessment could be approached from a different perspective.

This study has some limitations that are worthwhile mentioning. The study was single-center and image acquisition and delineation were performed by highly trained personnel in a clinical research setting. In this study, only 3.7% images were excluded, but we acknowledge that this number could be higher with a larger number of operators and settings. However, we believe that the rate of failed acquisitions should be low with minimal training, since the axial sections used are standard of practice for any healthcare provider trained in fetal ultrasonography. The sample size in this study was similar to those of clinical studies to validate amniotic fluid tests, but it prevented evaluating the performance within narrow gestational age ranges. We acknowledge that a larger sample size should be used to obtain this information. A multicenter international study to validate the results here reported is now underway.

In summary, this study provides evidence that a purpose-developed software based on quantitative texture analysis of fetal lung ultrasound images predicts neonatal respiratory morbidity. The performance obtained was similar to those reported for commercial FLM tests in amniotic fluid. These results should be confirmed in larger multicenter studies.

REFERENCES

1. Teune MJ, Bakhuizen S, Gyamfi Bannerman C, Opmeer BC, van Kaam AH, van Wassenaer AG, Morris JM, Mol BW. A systematic review of severe morbidity in infants born late preterm. American journal of obstetrics and gynecology. 2011;205(4):374.e1-9.

2. Spong CY, Mercer BM, D'Alton M, Kilpatrick S, Blackwell S, Saade G. Timing of indicated late-preterm and early-term birth. Obstetrics and gynecology. 2011;118(2 Pt 1):323-33.

3. Consortium on Safe Labor, Hibbard JU, Wilkins I, Sun L, Gregory K, Haberman S. Respiratory morbidity in late preterm births. JAMA : the journal of the American Medical Association. 2010;304(4):419-25.

4. Tita AT, Landon MB, Spong CY, Lai Y, Leveno KJ, Varner MW, Moawad AH, Caritis SN, Meis PJ, Wapner RJ, Sorokin Y, Miodovnik M, Carpenter M, Peaceman AM, O'Sullivan MJ, Sibai BM, Langer O, Thorp JM, Ramin SM, Mercer BM; Eunice Kennedy Shriver NICHD Maternal-Fetal Medicine Units Network. Timing of elective repeat cesarean delivery at term and neonatal outcomes. The New England journal of medicine. 2009;360(2):111-20.

5. Clark SL, Miller DD, Belfort MA, Dildy GA, Frye DK, Meyers JA. Neonatal and maternal outcomes associated with elective term delivery. American journal of obstetrics and gynecology. 2009;200(2):156.e1-4.

6. Gluck L, Kulovich MV, Borer RC, Jr., Brenner PH, Anderson GG, Spellacy WN. Diagnosis of the respiratory distress syndrome by amniocentesis. American journal of obstetrics and gynecology. 1971;109(3):440-5.

7. Gluck L, Kulovich MV. Lecithin-sphingomyelin ratios in amniotic fluid in normal and abnormal pregnancy. American journal of obstetrics and gynecology. 1973;115(4):539-46.

8. Neerhof MG, Dohnal JC, Ashwood ER, Lee IS, Anceschi MM. Lamellar body counts: a consensus on protocol. Obstetrics and gynecology. 2001;97(2):318-20.

9. Besnard AE, Wirjosoekarto SA, Broeze KA, Opmeer BC, Mol BW. Lecithin/sphingomyelin ratio and lamellar body count for fetal lung maturity: a meta-analysis. European journal of obstetrics, gynecology, and reproductive biology. 2013;169(2):177-83.

10. ACOG Practice Bulletin No. 97: Fetal lung maturity. Obstetrics and gynecology. 2008;112(3):717-26.

11. Maeda K, Utsu M, Yamamoto N, Serizawa M. Echogenicity of fetal lung and liver quantified by the grey-level histogram width. Ultrasound in medicine & biology. 1999;25(2):201-8.

12. Prakash KN, Ramakrishnan AG, Suresh S, Chow TW. Fetal lung maturity analysis using ultrasound image features. IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society. 2002;6(1):38-45.

13. Tekesin I, Anderer G, Hellmeyer L, Stein W, Kuhnert M, Schmidt S. Assessment of fetal lung development by quantitative ultrasonic tissue characterization: a methodical study. Prenatal diagnosis. 2004;24(9):671-6.

14. Serizawa M, Maeda K. Noninvasive fetal lung maturity prediction based on ultrasonic gray level histogram width. Ultrasound in medicine & biology. 2010;36(12):1998-2003.

15. La Torre R, Cosmi E, Anceschi MH, Piazze JJ, Piga MD, Cosmi EV. Preliminary report on a new and noninvasive method for the assessment of fetal lung maturity. Journal of Perinatal Medicine. 2003; 31(5):431-4.

16. Cosmi EV, Anceschi MM, Cosmi E, Piazze JJ, La Torre R. Ultrasonographic patterns of fetal breathing movements in normal pregnancy. International Journal of Gynaecology & Obstetrics. 2003; 80(3):285-90.

17. Insana MF, Garra BS, Rosenthal SJ, Hall TJ. Quantitative ultrasonography. Medical progress through technology. 1989;15(3-4):141-53.

18. Bergen JR, Adelson E. Theories of visual texture perception. Spatial vision. 1991;10:114-34.

19. Cobo T, Bonet-Carne E, Martínez-Terrón M, Perez-Moreno A, Elías N, Luque J, Amat-Roldan I, Palacio M. Feasibility and reproducibility of fetal lung texture analysis by Automatic Quantitative Ultrasound Analysis and correlation with gestational age. *Fetal diagnosis and therapy*. 2012;31(4):230-6.

20. Palacio M, Cobo T, Martínez-Terrón M, Rattá GA, Bonet-Carné E, Amat-Roldán I, Gratacós E. Performance of an automatic quantitative ultrasound analysis of the fetal lung to predict fetal lung maturity. American journal of obstetrics and gynecology. 2012; 207(6):504 e1-5.

21. Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 1989; 11(7):674-693.

22. Daubechies I. Orthonormal bases of compactly supported wavelets. Communications on pure and applied mathematics. 1988; 41(7):909-996.

23. Dalal N, Triggs B. Histograms of oriented gradients for human detection. Computer Vision and Pattern Recognition, Computer Society Conference on. 2005; 1:886-893.

24. Haralick RM, Shanmugan K, Dinstein I. Textural Features for Image Classification. Systems, IEEE Transactions on . 1973; SMC-3:6:610-621.

25. Turk M, Pentland A. Eigenfaces for Recognition. Journal of Cognitive Neuroscience. 1991; 3(1):71-86.

26. Lowe DG. Object recognition from local scale-invariant features. Computer vision, The proceedings of the IEEE International Conference on. 1999; 2:1150-1157.

27. Tenorio V, Bonet-Carne E, Botet F, Marques F, Amat-Roldan I, Gratacos E. Correlation Between a Semiautomated Method Based on Ultrasound Texture Analysis and Standard Ultrasound Diagnosis Using White Matter Damage in Preterm Neonates as a Model. Journal of ultrasound in Medicine. 2011; 30(10):1365-1377.

28. Ojala T, Maenpaa T, Pietikainen M, Viertola J. OUTEX - new framework for empirical evaluation of texture analysis algorithms. Pattern Recognition Proceedings, International Conference on. 2002; 1:701-706.

29. PhoTex database. Texture lab, Heriot-Watt University, Edinburgh, UK. Available on-line at http:// http://www.macs.hw.ac.uk/texturelab/resources/databases/photex/.

30. Holland PW, Welsch, RE. Robust Regression Using Iteratively Reweighted Least-Squares. Communications in Statistics: Theory and Methods. 1977; A6:813–827.

31. Hastie T, Tibshirani R. Discriminant adaptive nearest neighbor classification. Pattern Analysis and Machine Intelligende, IEEE Transactions on.1996; 18(6):607-616.

32. Kung SY, Taur JS. Decision-based neural networks with signal/image classification applications. .Neural Networks, IEEE Transactions on. 1995; 6(1):170-181.

33. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: data mining, inference, and Prediction. Second Edition. Springer series in statistics. Springer, 2009.

34. Gelfand SB, Ravishankar CS, Delp EJ. An iterative growing and pruning algorithm for classification tree design. Systems, Man and Cybernetics, IEEE International Conference on. 1989; 2:818-823.

35. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Viena, 2013.

36. Wijnberger LD, Huisjes AJ, Voorbij HA, Franx A, Bruinse HW, Mol BW. The accuracy of lamellar body count and lecithin/sphingomyelin ratio in the prediction of neonatal respiratory distress syndrome: a meta-analysis. BJOG : an international journal of obstetrics and gynaecology. 2001;108(6):583-8.

37. Karcher R, Sykes E, Batton D, Uddin Z, Ross G, Hockman E, Shade GH Jr. Gestational age-specific predicted risk of neonatal respiratory distress syndrome using lamellar body count and surfactant-to-albumin ratio in amniotic fluid. American journal of obstetrics and gynecology. 2005;193(5):1680-4.

38. Hagen E, Link JC, Arias F. A comparison of the accuracy of the TDx-FLM assay, lecithin-sphingomyelin ratio, and phosphatidylglycerol in the prediction of neonatal respiratory distress syndrome. Obstetrics and gynecology. 1993;82(6):1004-8.

39. Russell JC, Cooper CM, Ketchum CH, Torday JS, Richardson DK, Holt JA, Kaplan LA, Swanson JR, Ivie WM. Multicenter evaluation of TDx test for assessing fetal lung maturity. Clinical chemistry. 1989;35(6):1005-10.

40. Haymond S, Luzzi VI, Parvin CA, Gronowski AM. A direct comparison between lamellar body counts and fluorescent polarization methods for predicting respiratory distress syndrome. American journal of clinical pathology. 2006;126(6):894-9.

41. Sengupta S, Carrion V, Shelton J, Wynn RJ, Ryan RM, Singhal K, Lakshminrusimha S. Adverse neonatal outcomes associated with early-term birth. JAMA pediatrics. 2013;167(11):1053-9.

42. ACOG committee opinion no. 560: Medically indicated late-preterm and early-term deliveries. Obstetrics and gynecology. 2013;121(4):908-10.

43. Towers CV, Freeman RK, Nageotte MP, Garite TJ, Lewis DF, Quilligan EJ. The case for amniocentesis for fetal lung maturity in late-preterm and early-term gestations. American journal of obstetrics and gynecology. 2014;210(2):95-6.

44. Simon AE, Uddin SG. National trends in primary cesarean delivery, labor attempts, and labor success, 1990-2010. American journal of obstetrics and gynecology. 2013;209(6):554 e1-8.

45. Laughon SK, Reddy UM, Liping S, Jun Z. Precursors for Late Preterm Birth in Singleton Gestations. Obstet Gynecol. 2010;116(5):1047-1055.

46. Reddy UM, Ko CW, Raju TN, Willinger M. Delivery indications at late-preterm gestations and infant mortality rates in the United States. Pediatrics. 2009; 124(1):234-240.

Table 1. Baseline characteristics of the study group.

	n = 144
Maternal age, y	32.8 (5.7)
Nuliparity	86 (59.7)
Antenatal Corticosteriods	52/144 (36.1)
Gestational age at scan, weeks.days	36.0 (3.3)
28.0 to < 33.6	38/144 (26.4)
≥ 34.0	106/144 (73.6)

Values are expressed as mean (SD) or n (%) as appropriate.

Table 2. Perinatal outcomes of the study population.

	n = 144
Gestational age at delivery (weeks.days)	36.1 (3.3)
28.0 to 33.6	36 (25.0)
34.0 to 39.0	108 (75.0)
Reason for delivery (28.0-33.6)	
Spontaneous preterm labor	12/144 (8.3)
Maternal / Fetal conditions	24/144 (16.7)
Reason for delivery (34.0-39.0)	
Spontaneous preterm labor (34.0 to 36.6)	3/144 (2.1)
Maternal / Fetal conditions (34.0 to 36.6)	19/144 (13.2)
Early-term elective delivery (37.0 to 39.0)	86/144 (59.7)
Cesarean Delivery	97/144 (67.4)
Birth weight (g)	2644 (888)
Umbilical artery pH < 7.10 [a]	6/132 (4.5)
Apgar 5 minutes < 7	0/144 (0)
Admission to NICU	53/144 (36.8)
Respiratory morbidity	29/144 (20.1)
Days at NICU [b]	25.6 (33.1)
Neonatal death	0/144 (0%)

NICU: neonatal intensive care unit

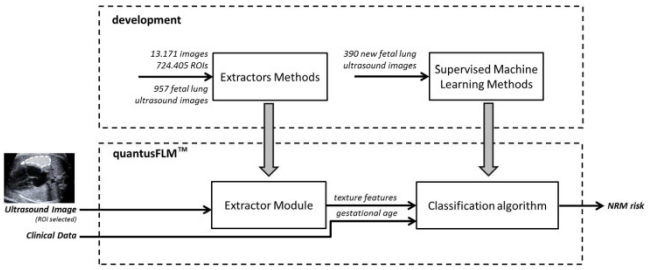
Values are expressed in mean (SD) or n (percentage) when appropriate.

[a]pH AU was not available for all newborns. [b]Of those admitted.

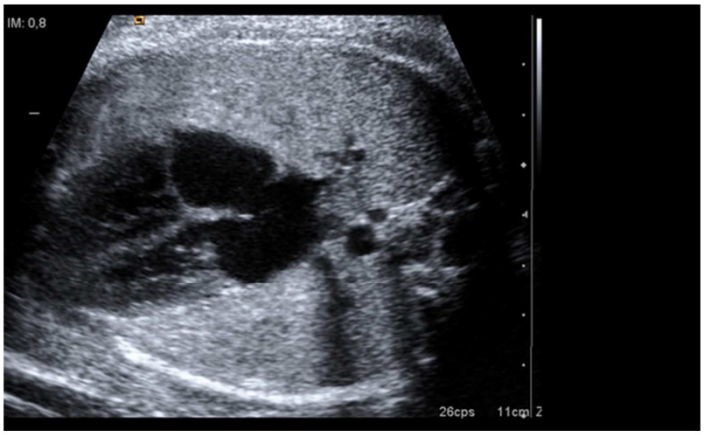
Table 3. Algorithm performance stratified by gestational age subgroups.

	Gestational age at scan		
	[28.0-39.0]	**[28.0-33.6]**	**[34.0-39.0]**
Total Samples	144	38	106
NRM (%)	29 (20.1%)	21 (55.3%)	8 (7.5%)
TP	25	19	6
TN	100	16	84
FP	15	1	14
FN	4	2	2
Accuracy	86.8%	92.1%	84.9%
Sensitivity	86.2%	90.5%	75.0%
Specificity	86.9%	94.1%	85.7%
PPV	62.5%	95.0%	30.0%
NPV	96.2%	88.9%	97.7%

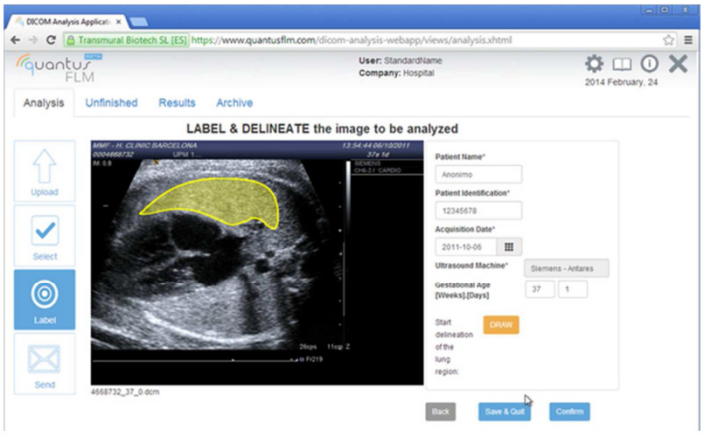
NRM: Neonatal Respiratory Morbidity. TN: True Positive, TN: True Negative, FN: False Negative, FP: False positive, PPV: Positive Predictive Value and NPV: Negative Predictive Value.



Scheme of quantusFLM development and final trained algorithm.



Standard lateral axial section of the fetal thorax in a 4-chamber view as used for image analysis in this study



Platform delineation screen showing a manual delineation of the region of interest in the lung proximal to the transducer.

4.4. STUDY 4

Automatic Quantitative MRI Texture Analysis in Small-for-Gestational-Age Fetuses Discriminates Abnormal Neonatal Neurobehavior

M. Sanz-Cortes, GA. Ratta, F. Figueras, E. Bonet-Carne, N. Padilla, A. Arranz, N. Bargallo, E. Gratacos. PLoS ONE 2013 8(7): e69595.

Status: published

Journal Impact factor: 3.730

Quartile: 1st

Results from this study have been presented in the following congress:

1. Sanz-Cortes M, Figueras F, Padilla N, Bargallo N, Bonet-Carne E, Luque J, Amat-Roldan I, Gratacos E. Evaluation of a computer-based analysis of brain textures on fetal MRI to detect changes in small for gestational age (SGA) fetuses and to predict neurodevelopmental outcome. ISUOG.September 18-22, 2011. Los Angeles, USA.

OPEN ACCESS Freely available online PLOS ONE

Automatic Quantitative MRI Texture Analysis in Small-for-Gestational-Age Fetuses Discriminates Abnormal Neonatal Neurobehavior

Magdalena Sanz-Cortes[1,2,3]*, Giuseppe A. Ratta[1,2,3], Francesc Figueras[1,2,3], Elisenda Bonet-Carne[1,2,3], Nelly Padilla[1,2,3], Angela Arranz[1,2,3], Nuria Bargallo[4,5], Eduard Gratacos[1,2,3]

1 Maternal-Fetal Medicine Department, ICGON, Hospital Clinic, Universitat de Barcelona, Barcelona, Spain, **2** Fetal and Perinatal Medicine Research Group, Institut d'Investigacions Biomediques August Pi i Sunyer (IDIBAPS), Barcelona, Spain, **3** Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Barcelona, Spain, **4** Department of Radiology Hospital Clinic, Centre de Diagnostic per la Imatge, Hospital Clínic, Barcelona, Spain, **5** Image platform IDIBAPS, Barcelona, Spain

Abstract

Background: We tested the hypothesis whether texture analysis (TA) from MR images could identify patterns associated with an abnormal neurobehavior in small for gestational age (SGA) neonates.

Methods: Ultrasound and MRI were performed on 91 SGA fetuses at 37 weeks of GA. Frontal lobe, basal ganglia, mesencephalon and cerebellum were delineated from fetal MRIs. SGA neonates underwent NBAS test and were classified as abnormal if $\geq$1 area was $<5^{th}$ centile and as normal if all areas were $>5^{th}$ centile. Textural features associated with neurodevelopment were selected and machine learning was used to model a predictive algorithm.

Results: Of the 91 SGA neonates, 49 were classified as normal and 42 as abnormal. The accuracies to predict an abnormal neurobehavior based on TA were 95.12% for frontal lobe, 95.56% for basal ganglia, 93.18% for mesencephalon and 83.33% for cerebellum.

Conclusions: Fetal brain MRI textural patterns were associated with neonatal neurodevelopment. Brain MRI TA could be a useful tool to predict abnormal neurodevelopment in SGA.

Citation: Sanz-Cortes M, Ratta GA, Figueras F, Bonet-Carne E, Padilla N, et al. (2013) Automatic Quantitative MRI Texture Analysis in Small-for-Gestational-Age Fetuses Discriminates Abnormal Neonatal Neurobehavior. PLoS ONE 8(7): e69595. doi:10.1371/journal.pone.0069595

Editor: Grainne M. McAlonan, King's College London, United Kingdom

Received February 9, 2013; **Accepted** June 10, 2013; **Published** July 26, 2013

Funding: This work was supported by grants from The Cerebra Foundation for the Brain Injured Child (Carmarthen, Wales, UK) http://www.cerebra.org.uk/English/Pages/home.aspx, the Thrasher Research Fund (Salt Lake City, USA) http://www.thrasherresearch.org/, the Fundacion Dexeus (Barcelona, Spain) http://www.fundaciondexeus.org. Dr. Sanz Cortes was supported by a Rio Ortega research fellowship, Spain. Dr. Padilla was supported by Instituto de Salud Carlos III Sara Borrell, Spain. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Dr. Eduard Gratacos has been an Editor of PLOS ONE in the past. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

* E-mail: msanz1@clinic.ub.es

Introduction

Smallness for gestational age affects 10% of all pregnancies [1]. In clinical practice when an estimated fetal weight is below the tenth centile and Doppler assessment of the umbilical artery is normal, the diagnosis of a small-for-gestational-age (SGA) is reached [2,3,4]. Although some fetuses with this diagnosis are constitutionally small, in a substantial proportion of cases, the diagnosis of SGA identifies mild forms of fetal growth restriction due to placental insufficiency that are not expressed by umbilical artery Doppler. Therefore, fetal development occurs in suboptimal conditions, with a deprived delivery of oxygen and nutrients to the fetal brain [5]. Under these conditions brain reorganization may take place, among other changes of the so-called fetal programming [6]. Different authors have shown how despite the fact that most SGA fetuses reach term without signs of deterioration, there is a proportion of them that present an increased risk for an adverse perinatal outcome [7,8,9] with an abnormal neonatal neurobehavior [10,11] and impaired neurodevelopment in early childhood [12]. Considering its prevalence, SGA constitutes a challenge and an opportunity for public health to improve the impact of prenatal conditions in quality of life. However, at present the detection of SGAs at risk of abnormal neurodevelopment is limited since standard clinical examinations fail to identify significant differences. For this purpose, it is crucial to develop new biomarkers based on the characterization of distinctive brain patterns associated with abnormal neurodevelopment. Quantitative imaging based on texture analysis might offer an opportunity for the development of such biomarkers.

Quantitative imaging techniques are based on the application of imaging physics for the development of algorithms improving the information obtained from medical images. These techniques attempt to improve the performance of subjective inspection by extracting quantitative information that may detect non-visible changes and be used in a more objective fashion for prediction, diagnosis and monitoring. Among various approaches, Texture Analysis (TA) is a technique that extracts patterns from images based on the characterization of the microstructural information

that may not be assessed visually [13]. It has been widely used in different pathologies [13,14,15,16,17], being able to classify pathological from healthy tissues in liver [18], breast [17] and tumors [14].We have previously tested a TA software whose efficacy has been shown by different studies on preterm transcranial ultrasound imaging demonstrating a high accuracy in the early identification of preterm white matter damage in subclinical stages [19], on fetal MRIs showing a discrimination based on brain textural features between SGA and AGA fetuses [20] and also when applied on fetal lung ultrasound images, showing a high correlation with gestational age [21]. In a previous study we provided evidence that fetuses with SGA presented statistical differences in their brain MRI textural patterns with respect to controls [20]. In this study we explored whether these patterns showed a correlation with neonatal neurobehavior.

The aim of the study was to test the hypothesis that SGA fetuses show abnormalities in different brain areas reflected by changes in TA, which can be associated to an abnormal neonatal neurobehavior.

Materials and Methods

2.1 Subjects

This study is part of a larger prospective research program on IUGR involving fetal assessment and short and long term postnatal follow-up at the Hospital Clinic (Barcelona, Spain). A prospective cohort of 91 SGA singleton fetuses, defined as an estimated and confirmed birthweight below the 10th centile according to local standards [22] with normal UA pulsatility index (PI) (below the 95th centile) [23], was included for this study. Exclusion criteria were non-cephalic presentation, the presence of congenital malformations, chromosomal abnormalities, perinatal infections and chronic maternal pathology.

Prenatal and neonatal data were prospectively recorded. The protocol was approved by the institutional ethics committee of the Hospital Clinic of Barcelona and all participants gave written informed consent for exams performed on themselves on the basis of this trial and on their neonates as their legal guardians (Institutional Review Board 2008/4422).

2.2 Data acquisition

2.2.1 Ultrasound data. Gestational age was corrected from fetal crown-rump length in the first trimester [24]. Prenatal Doppler ultrasound examinations were performed within one week from MRI scan. Weight estimation, placental and amniotic fluid evaluation were performed using a Siemens Sonoline Antares ultrasound machine equipped with a 6–2 MHz linear-curved-array transducer. Umbilical artery Doppler spectral parameters were obtained automatically from three or more consecutive waveforms with the angle of isonation as close to zero as possible from a free floating cord loop.

2.2.2 Fetal MRI. All cases were scanned at 37 weeks of gestation in a TIM TRIO 3.0 T scanner (Siemens, Germany) without sedation. A body coil with 8 elements was wrapped around the mother's abdomen. Routine fetal imaging took from 15 to 30 min. Fetal neuroimaging consisted on single-shot, fast spin echo T2 weighted sequences (TR 990 ms, TE 137 ms, slice thickness 3.5 mm, FoV 260 mm, voxel size 1.4×1.4×3.5 mm, in plane resolution 192, flip angle 180°, acquisition time 24 seconds) acquired in the three orthogonal planes. If the quality of the images was distorted due to fetal movements, consecutive repetitions were acquired until an acceptable quality image was obtained.

Structural MRI images were reviewed for the presence of anatomical abnormalities by an experienced neuroradiologist, blinded to group membership.

2.2.3 Neurobehavioral performance. Postnatal follow up was offered to all patients. Neonatal Behavioral Assessment Scale (NBAS) test is a standard method for evaluating newborns' capacity to respond to the environment, which reflects brain maturation [25]. It was performed in all 91 patients prospectively at 42±1 weeks by 1 of 2 observers accredited by The Brazelton Institute (Harvard Medical School, Boston, MA) that were blinded to the SGA diagnosis of this group and their perinatal outcomes. This test evaluates 35 items that are rated on a 1 to 9 scale, where 9 is the best performance for some areas and for others this is represented by the central score of 5 [26]. Items are grouped into 6 clusters, including habituation (habituation to light, rattle, bell and tactile stimulation of the foot), motor (general tone, elicited activity, spontaneous activity and motor maturity), social-interactive (responses to visual, animate and inanimate auditory stimuli and alertness), organization of state (irritability, state lability, maximal excitation and reaction time) and regulation of state (self-quieting and hand-to-mouth responses). The social-interactive cluster was subscored for visual and auditory stimuli. In addition, as reported recently by the authors of the NBAS test [27], an aggregation of individual items (alertness, quality of the alert responsiveness and cost of attention) was used to evaluate the capacity of the newborn's attention. Neonates were assessed in the afternoon, between feedings in a small, semidark quiet room with a temperature between 22° and 27°C in the presence of a parent.

In order to categorize the scores from the studied clusters of the NABS test to determine cases and controls, fifth centile was calculated for each cluster determining in each subject if their performance on that cluster was above or below this centile cutoff.

2.3 Classification of the study groups

All 91 SGA neonates that composed our sample were divided into two groups: Cases and controls based on abnormal or normal NBAS test results. SGAs were classified as cases if any of the studied cluster's score (habituation, motor, social-interactive, organization of state, regulation of state and attention) was below the 5th centile, and they were classified as controls if all the scores were above the 5th centile.

2.4 Delineation of Regions of Interest (ROIs)

A custom-made Graphical User Interface (GUI) tool on MATLAB R2007b (version 7.5.0.342; MATLAB; the MathWorks Inc., Natick, Massachusetts, USA) was used to manually delineate all nine regions of interest (ROIs). Before delineation, all images were checked for artifacts. If the anatomic area to be delineated showed a suboptimal quality, it was discarded. Delineation was performed by two experienced operators in neuroanatomy blinded to group membership. Right and left supra- and infraventricular frontal lobe, right and left basal ganglia, mesencephalon and cerebellum were selected as clinically relevant ROIs in the studied condition (Figure 1), following the criteria for delineation and image reorientation steps as explained elsewhere [20].

2.5 Image analysis

2.5.1 TA and Statistical learning algorithm. The TA method used in the software that was applied is based on wavelet decomposition [28] using Daubechies orthogonal wavelet basis [29]. Wavelets were used to decompose the images in a pyramidal scheme as described by Quellec *et al.* [30] and modified for the use in medical images as previously described [19]. Texture descrip-

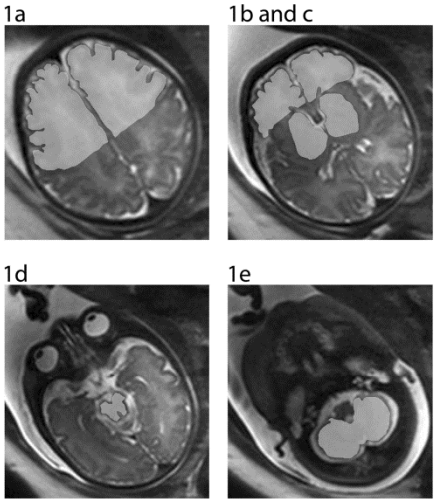


Figure 1. Image selection and ROI delineation: 1a.1b- Right and Left supraventricular frontal lobe; 2a.2b.- Right and Left infraventricular frontal lobe; 3a. 3b- Right and left basal ganglia;4a. 4b.-Right and left Mesencephalon; 5- Cerebellum.
doi:10.1371/journal.pone.0069595.g001

tors of an image were described as the concatenation of the marginal distributions of each equalized sub-band image.

This method was applied to the delineated ROIs obtaining a set of 15,300 descriptors per ROI. Based on their anatomic functionality and clinical relevance, descriptors from all 9 delineated ROIs were grouped into 4 main areas:

1) Frontal lobe: left and right infra- and supra ventricular frontal lobe.
2) Basal ganglia: left and right basal ganglia.
3) Mesencephalon: left and right mesencephalon.
4) Cerebellum.

Both infra and supraventricular frontal lobe regions were grouped into one single vector to represent the complexity of the frontal lobe at two different levels. In some cases, one of the ROIs to be merged was not delineated due to an insufficient image quality leading inevitably to a decrease in the number of delineated areas in the frontal lobe. Due to this limitation, we performed our discriminative analysis based on 81 subjects for frontal lobe area, 88 for basal ganglia and mesencephalon and 83 subjects for cerebellum (Table 1).

2.5.2 Selection of descriptors and identification algorithm. Computational models were applied in order to select an appropriate subset of descriptors to identify differences between SGAs with normal or abnormal NBAS test results. To this end, a combination of two artificial intelligence methods were applied: Support Vector Machines (SVM) and Genetic Algorithms (GALs) [31]. As a result, a compact subset of descriptors (between 28 and 77 depending on the area) was automatically selected.

Table 1. Study areas obtained from delineated ROIs.

	Cases (N = 42)	Controls (N = 49)	Total
Frontal lobe	38	43	81
Basal ganglia	41	47	88
Mesencephalon	41	47	88
Cerebellum	38	45	83

doi:10.1371/journal.pone.0069595.t001

The procedure initially splitted the total sample into two subsets of equal size (subsets "A" and "B"). Firstly, a model was created with subsample "A" and validated with "B". The accuracy was calculated as the percentage of correctly identified SGAs with normal or abnormal NBAS test results in the validation subset "B". Subsequently, groups were permuted: a model was created with subsample "B" (using the same subset of descriptors) and validated with "A", obtaining a second identification percentage. The mean accuracy resulting from the two tests to identify group membership and therefore the possibility of obtaining a normal or abnormal NBAS test was finally measured.

Each model validation result provided a score per subject that was useful for further group comparisons. In order to obtain these scores, the algorithm was designed in a way that the cut-off was assigned to "0", which is the standard value for SVM. Output values above "0" indicated a high risk for abnormal neurobehavior and below "0" indicated low risk.

2.6 Statistical analysis

2.6.1 Demographic and clinical data. Student's t test for independent samples and Pearson's X^2 or Fisher's exact tests were used to compare quantitative and qualitative data, respectively. Multivariate analysis of covariance was conducted to analyze the results of the NBAS test and the prediction scores for each area between the two groups. For the first analysis a model was carried out for each different set of skills (habituation, motor, organization of state, regulation of state, social interactive and attention) with the study group included as a factor and smoking during pregnancy, gender, Apgar score below 7, days of adaptation from birth to the test and gestational age at the moment of NBAS test as covariates. For the second analysis, a multivariate analysis of covariance was conducted to analyze the results from the prediction scores adjusting for the same covariates.

Results were considered to be significant at $p<0.05$. All statistical calculations were done using the software package SPSS statistical software, version 17.0(SPSS for Windows, SPSS Inc, Chicago, IL).

Results

3.1 Study groups characteristics

Anthropometric, ultrasound and MRI data were obtained from all patients included in the study. All fetal MR images were considered as normal, not finding signs of intracranial pathology.

As shown in Table 2, when we compared clinical characteristics between SGAs with normal and abnormal NBAS test results, no differences that could explain their different neurobehavioral outcome were found: Both populations were similar in terms of gender, birthweight and biometries. Also, no differences were found concerning signs of perinatal distress or in the rate of breast feeding at discharge or in length of NICU admission (Table 3).

Table 2. Maternal characteristics of the population.

	Cases (N = 42)	Controls (N = 49)	P
Maternal age (y)	31.14+5.9	32.14+5.4	0.41
BMI (kg/m^2)	21.53+3.9	22.0+3.2	0.51
Primiparity	83.3	69.4	0.12
Non-white ethnicity	26.2	20.4	0.51
Smoker	26.2	20.4	0.51
Superior studies	43.9	51.1	0.5
Low -Socio-economic status [a]	26.5	18.4	0.41
GA at US (w)	37.08+1.06	37.2+0.86	0.43
GA at MRI (w)	37.25+1.01	37.32+1.01	0.75

Results are expressed as mean + and standard deviation or percentage determined by Students t-test for independent samples, Pearson's X^2 or Fisher's exact test as appropriate. Y: years. BMI: Body mass index. GA: Gestational age.US: Ultrasound. MRI: Magnetic Resonance Imaging; w = weeks. [a] Routine occupations, long-term unemployment, or never worked (United Kingdom National Statistics Socio-economic Classification).
doi:10.1371/journal.pone.0069595.t002

Table 3. Perinatal outcomes of the population.

	Cases (N = 42)	Controls (N = 49)	P*
GA at delivery (w)	38.79+1.03	38.88+1.26	0.72
Labor induction	76.2	77.1	0.92
Emergency Cesarean section	23.8	30.6	0.46
Cesarean section	26.2	42.9	0.09
Birthweight (g)	2446+289	2422+324	0.70
Birth weight centile	2.45 +2.74	2.93+2.79	0.41
Male	52.4	65.3	0.21
Head circumference (cm)	32.6+1.14	32.5+2.13	0.74
Length (cm)	46.04+1.79	46.06+2.2	0.95
5 minute Apgar score of <7	2.4	4.1	0.65
Neonatal acidosis [a]	10	17.4	0.32
Breast feeding during neonatal period	92.7	93.9	0.82
NICU admission	4.8	4.1	0.87
NICU stay length (d)	0.38+1.78	0.12+0.72	0.35

Results are expressed as mean + and standard deviation or percentage determined by * Student's t-test for independent samples, Pearson's X^2 or Fisher's exact test as appropriate. GA: Gestational age. NICU: Neonatal Intensive care unit. g = grams; d = days; w = weeks. [a] Umbilical artery pH <7.15 and base excess >12 mEq/L.
doi:10.1371/journal.pone.0069595.t003

3.2 NBAS test results

As shown in Table 4, both populations showed similar adaptation times from birth until the performance of the NBAS test and age at the moment of the test. Concerning the scores of the NBAS test, overall worse results were found in the abnormal NBAS test results group and were more pronounced in the habituation and regulation of state clusters (Table 4).

3.3 Automatic identification of study group membership based on fetal brain MRI TA

The mean accuracy obtained after the application of the procedure previously explained for each area was 95.56% in basal ganglia area, 95.12% in frontal lobe, 93.18% in mesencephalon and 83.33% in cerebellum.

There was a significant difference between the scores representing TA for each algorithm output between SGAs with normal and abnormal NBAS test results, in all studied areas. This comparison was adjusted for smoking status, gender, Apgar score below 7, days of adaptation and age at NBAS test. The distribution of the scores obtained with the TA-based algorithms for each area is displayed in Figure 2.

Discussion

This study provides evidence that fetal brain MRI textural patterns are associated with neonatal neurobehavior and sets the basis for further research on in utero imaging biomarkers based on quantitative assessment of brain microstructure.

The correlation between TA and functional outcome has previously been demonstrated in adults with neurological conditions and apparently normal MRI scans, such as in mild traumatic brain injury or mild cognitive impairment [13,32]. In these conditions, TA was able to identify differences in relation with the progression of the disease and indicate the most affected areas. To our best knowledge, this is the first time in which brain quantitative imaging in fetuses has been used to establish associations with post-natal neurobehavior. The results are in line with the existence of brain reorganization in IUGR. Different lines of evidence have shown that fetuses and infants affected with early and severe IUGR have significant differences in brain metabolism, sulcation, composition, and microstructure [33,34,35,36]. Furthermore, correlations between these brain disturbances and neurological performance have been reported [33,35].The majority of earlier studies were conducted in early-onset IUGR. However recent evidence supports that late-onset IUGR have changes in the same direction, including differences in brain metabolism and microstructure [37] and signs of increased axonal loss at 5 years of age [38]. Along the same lines, in a previous study we reported differences in textural patterns on fetal brain MRI between term SGA and AGA fetuses [20]. In the present study we provide evidence that these patterns are correlated with post-natal neurobehavior.

From a pathophysiological point of view, textural patterns could reflect brain microstructural alterations in late-onset IUGR fetuses. Brain reorganization is thought to underlie developmental deficits of SGA infants, which show cognitive disadvantages from the neonatal period until adolescence [10] [39,40]. It is increasingly accepted that subtle changes in brain morphology may be present years before the clinical onset of neuropsychiatric and neurodegenerative diseases [41,42].These changes could be identified by quantitative imaging in order to define "early endophenotypes" as markers of future functional outcome [33]. Therefore, results obtained from this study encourage further research aiming at the identification of such "imaging endophenotypes" in IUGR, and possibly other neurocognitive disorders of fetal and perinatal origin.

In this study we chose several brain areas that might potentially be involved in brain reorganization affecting neurodevelopment. Attention skills are generally attributed to the frontal lobe, due to its importance for cognitive tasks and the results of MRI studies of attention deficit and hyperactivity disorder [43]. On the other hand, potential cerebellar microstructure alterations could be preferentially involved in lower scores of the motor cluster, including motor learning, memory and cognition and in behavior

Table 4. Clinical information and results regarding NBAS test.

	Total sample (N = 91)	Cases (N = 42)	Controls (N = 49)	P*
GA at NBAS test (w)	42.68 ± 2.82	43.11 ± 2.81	42.73 ± 2.5	0.49
Adaptation time (d)	28.44 ± 17.24	30.21 ± 18.06	26.92 ± 16.54	0.36
Scores from the clusters in the BAS test				
Habituation	6.33 ± 1.34	6.07 ± 1.64	6.57 ± 0.91	0.03
Social-interactive	6.17 ± 1.14	5.97 ± 1.23	6.33 ± 1.04	0.32
Motor	5.43 ± 0.69	5.33 ± 0.91	5.52 ± 0.43	0.82
Organization of state	3.90 ± 0.95	3.73 ± 1.16	4.05 ± 0.71	0.82
Regulation of state	4.26 ± 1.38	3.49 ± 1.25	4.93 ± 1.12	< 0.01
Attention	6.24 ± 1.58	5.94 ± 1.71	6.5 ± 1.42	0.48

Results are expressed as mean ± and standard deviation. *MANCOVA statistical analysis was used to compare scores in cases vs controls from each area of the NBAS test adjusting for smoking status, gender, Apgar score below 7, gestational age at NBAS test and days of adaptation. GA: Gestational age. w = weeks; d = days.
doi:10.1371/journal.pone.0069595.t004

[44]. However, brain neurostructure and organization undergoes substantial changes during the two first years of age, and in general extrapolation of observations from older children or adults to fetal and perinatal life is not feasible. In this study we did not find definite correlations between specific areas and behavioral domains. Actually, basal ganglia, frontal lobe and mesencephalon obtained similarly high accuracies in predicting their neurobehavioral outcome. As mentioned, this was somewhat expected. The contribution from each brain area to the NBAS test is unknown, probably existing direct or indirect influences from all areas in various NBAS clusters at this primitive stage of neurodevelopment.

From a clinical perspective, the study provides further evidence to support the existence of changes in brain development, which could be used for diagnosis of true forms of fetal growth restriction in utero. Identifying at-risk patients lays the basis for timely interventions in utero to decrease the rate of adverse perinatal results [8] and for selection of newborns for targeted interventions. Evidence from randomized trials indicates how preterm-born IUGR neonates that received the Newborn Individualized Developmental Care and Assessment program (NIDCAP) showed better neurobehavior, electrophysiology and brain structure than those receiving standard care [45]. Other interventions with demonstrated impact include breast feeding, with a positive effect on brain white matter growth [46] and a worse adherence in IUGR newborns due to a poorer regulation and organization of state during the neonatal period [47]. The potential clinical value of TA in the identification of risk requires a great deal of further research. At this point, most quantitative imaging-based methods are still far from clinical applications. Specifically, TA-based applications require developing robust algorithms based on large databases, software user interface platforms and feasibility studies

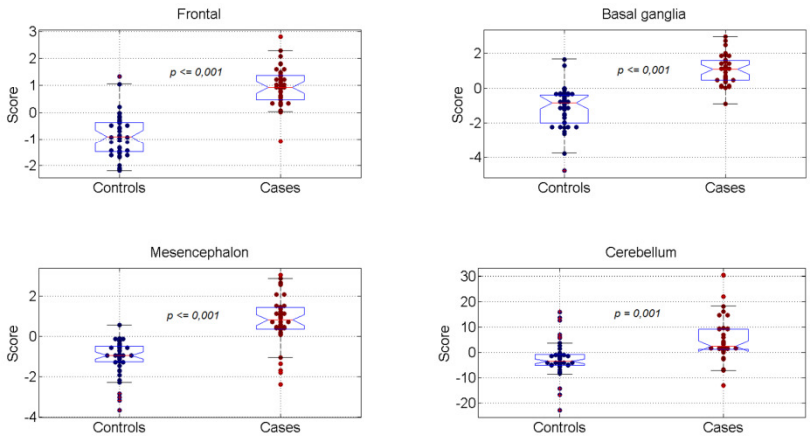


Figure 2. Score distribution of MRI TA-based algorithms for each area under study. Significant differences are present in all areas between cases and controls.*MANCOVA statistical analysis was used to compare scores from each brain area adjusting for smoking status, gender, Apgar score below 7, age at NBAS test and days of adaptation to NBAS test.
doi:10.1371/journal.pone.0069595.g002

demonstrating its value in clinical practice, and it is likely to be years before these studies are completed.

One strength of this study is that it evaluates brain MRI TA from a homogeneous cohort of term SGA fetuses selected in utero and prospectively followed up until the neonatal period. The correlation with neurobehavioral scores weeks after birth supports the importance of prenatal factors as a strong independent contributor to neurodevelopment, irrespective of postnatal events. It was remarkable that there was not any a priori potential bias on NBAS examiners since SGA newborns with normal and abnormal NBAS were homogeneous with respect of weight and length. In addition, study groups were similar in terms of perinatal outcomes and other potential confounding factors, such as days of adaptation, breast feeding or educational level from the mother.

However, we grant some limitations and technical considerations in this study. We acknowledge that this study based its functional outcome on neonatal neurobehavioral scores and not in long term cognitive evaluation. However, increasing evidence supports a neurobiological basis for infant or neonatal behavior [48], linking neonatal neurobehavioral skills with later neurocognitive development [25,36,47,49,50] and showing how scores on neurobehavioral tests predict IQ at 6 years of age [25].

Concerning the statistical learning algorithm that was developed for this study, it should be stated that the limited sample size prevented the use of an external validation by an independent group. However, the two-fold cross validation used can minimize the impact of this limitation. We acknowledge that these results are preliminary and require confirmation in larger sample sizes allowing external validation.

In summary, this study provides evidence that fetal brain quantitative imaging based on MRI TA has a potential in predicting an abnormal neurobehavioral outcome. This study supports further research on quantitative imaging techniques to develop imaging biomarkers of abnormal neurodevelopment in late-onset IUGR fetuses.

Author Contributions

Conceived and designed the experiments: MSC GAR EBC AA NP FF NB EG. Performed the experiments: MSC GAR EBC AA. Analyzed the data: MSC GAR EBC AA NP FF NB EG. Contributed reagents/materials/analysis tools: MSC GAR EBC AA NP FF NB EG. Wrote the paper: MSC GAR EBC AA NP FF NB EG.

References

1. Bernstein IM, Horbar JD, Badger GJ, Ohlsson A, Golan A (2000) Morbidity and mortality among very-low-birth-weight neonates with intrauterine growth restriction. The Vermont Oxford Network. Am J Obstet Gynecol 182: 198–206.
2. Gagnon R, Van den Hof M (2003) Society of Obstetricians and Gynaecologists of Canada, Diagnostic Imaging Comitee and Council.The use of fetal Doppler in Obstetrics. J Obstet Gynaecol Can 25: 601–614.
3. Royal College of Obstetrics and Gynaecology. Green-top guideline. (2002) The Investigation and Management of the Small-for-Gestational-Age Fetus. London, England.
4. American College of Obstetrics and Gyneacologists (1997).Committee on Obstetric Practice. Utility on antepartum umbilical artery Doppler velocimetry in intrauterine growth restriction:number 188, October 1997 (replaces no.116, November 1992). Int J Gynaecol Obstet 59: 269–270.
5. Rees S, Mallard C, Breen S, Stringer M, Gock M, et al. (1998) Fetal brain injury following prolonged hypoxemia and placental insufficiency: a review. Comp Biochem Physiol A Mol Integr Physiol 119: 653–660.
6. Godfrey KM, Barker DJ (2001) Fetal programming and adult health. Public Health Nutr 4: 611–624.
7. Doctor BA, O'Riordan MA, Kirchner HL, Shah D, Hack M (2001) Perinatal correlates and neonatal outcomes of small for gestational age infants born at term gestation. Am J Obstet Gynecol 185: 652–659.
8. Cruz-Martinez R, Figueras F, Hernandez-Andrade E, Oros D, Gratacos E (2011) Fetal brain Doppler to predict cesarean delivery for nonreassuring fetal status in term small-for-gestational-age fetuses. Obstet Gynecol 117: 618–626.
9. Figueras F, Eixarch E, Gratacos E, Gardosi J (2008) Predictiveness of antenatal umbilical artery Doppler for adverse pregnancy outcome in small-for-gestational-age babies according to customised birthweight centiles: population-based study. BJOG 115: 590–594.
10. Figueras F, Oros D, Cruz-Martinez R, Padilla N, Hernandez-Andrade E, et al. (2009) Neurobehavior in term, small-for-gestational age infants with normal placental function. Pediatrics 124: e934–941.
11. Padidela RN, Bhat V (2003) Neurobehavioral assessment of appropriate for gestational and small for gestational age babies. Indian Pediatr 40: 1063–1068.
12. Figueras F, Eixarch E, Meler E, Iraola A, Figueras J, et al. (2008) Small-for-gestational-age fetuses with normal umbilical artery Doppler have suboptimal perinatal and neurodevelopmental outcome. Eur J Obstet Gynecol Reprod Biol 136: 34–38.
13. Holli KK, Harrison L, Dastidar P, Waljas M, Lähdesmäki S, et al. (2010) Texture analysis of MR images of patients with mild traumatic brain injury. BMC Med Imaging 10: 8.
14. Kjaer L, Ring P, Thomsen C, Henriksen O (1995) Texture analysis in quantitative MR imaging. Tissue characterisation of normal brain and intracranial tumours at 1.5 T. Acta Radiol 36: 127–135.
15. Freeborough PA, Fox NC (1998) MR image texture analysis applied to the diagnosis and tracking of Alzheimer's disease. IEEE Trans Med Imaging 17: 475–479.
16. Bonilha L, Kobayashi E, Castellano G, Coelho G, Tinois E, et al. (2003) Texture analysis of hippocampal sclerosis. Epilepsia 44: 1546–1550.
17. Holli K, Laaperi AL, Harrison L, Luukkaala T, Toivonen T, et al. (2010) Characterization of breast cancer types by texture analysis of magnetic resonance images. Acad Radiol 17: 135–141.
18. Jirak D, Dezortova M, Taimr P, Hajek M (2002) Texture analysis of human liver. J Magn Reson Imaging 15: 68–74.
19. Tenorio V, Bonet-Carne E, Botet F, Marques F, Amat-Roldan I, et al. (2011) Correlation between a semiautomated method based on ultrasound texture analysis and standard ultrasound diagnosis using white matter damage in preterm neonates as a model. J Ultrasound Med 30: 1365–1377.
20. Sanz-Cortes M, Figueras F, Bonet-Carné E, Padilla N, Bargallo N, et al. (2013) Fetal brain MRI texture analysis identifies different microstructural patterns in adequate- and small-for-gestational-age fetuses at term. In press. Fetal Diagn Ther.
21. Cobo T, Bonet-Carne E, Martinez-Terron M, Perez-Moreno A, Elias N, et al. (2012) Feasibility and reproducibility of fetal lung texture analysis by automatic quantitative ultrasound analysis and correlation with gestational age. Fetal Diagn Ther 31: 230–236.
22. Figueras F, Meler E, Iraola A, Eixarch E, Coll O, et al. (2008) Customized birthweight standards for a Spanish population. Eur J Obstet Gynecol Reprod Biol 136: 20–24.
23. Arduini D, Rizzo G (1990) Normal values of Pulsatility Index from fetal vessels: a cross-sectional study on 1556 healthy fetuses. J Perinat Med 18: 165–172.
24. Robinson HP, Fleming JE (1975) A critical evaluation of sonar "crown-rump length" measurements. Br J Obstet Gynaecol 82: 702–710.
25. Canals J, Hernandez-Martinez C, Esparo G, Fernandez-Ballart J (2011) Neonatal Behavioral Assessment Scale as a predictor of cognitive development and IQ in full-term infants: a 6-year longitudinal study. Acta Paediatr 100: 1331–1337.
26. Brazelton TB, Nugent JK (1995) Neonatal Behavioral Assessment Scale. London: McKeith Press.
27. Sagiv SK, Nugent JK, Brazelton TB, Choi AL, Tolbert PE, et al. (2008) Prenatal organochlorine exposure and measures of behavior in infancy using the Neonatal Behavioral Assessment Scale (NBAS). Environ Health Perspect 116: 666–673.
28. Mallat SG (1989) A theory for multiresolution signal decomposition: The wavelet representation.. IEEE Transactions on pattern analysis and machine intelligence 11(7); 674–693.
29. Daubechies I (1988) Orthonormal bases of compactly supported wavelets. Communications on Pure and Applied Mathematics 41(7);909–996.
30. Quellec G, Lamard M, Cazuguel G, Cochener B, Roux C (2010) Wavelet optimization for content-based image retrieval in medical databases. Med Image Anal 14; 227–241.
31. Kernytsky A, Rost B (2009) Using genetic algorithms to select most predictive protein features. Proteins 75: 75–88.
32. de Oliveira MS, Balthazar ML, D'Abreu A, Yasuda CL, Damasceno BP, et al. MR imaging texture analysis of the corpus callosum and thalamus in amnestic mild cognitive impairment and mild Alzheimer disease. AJNR Am J Neuroradiol 32: 60–66.
33. Dubois J, Benders M, Borradori-Tolsa C, Cachia A, Lazeyras F, et al. (2008) Primary cortical folding in the human newborn: an early marker of later functional development. Brain 131: 2028–2041.
34. Story L, Damodaram MS, Allsop JM, McGuinness A, Patel A, et al. (2011) Brain metabolism in fetal intrauterine growth restriction: a proton magnetic resonance spectroscopy study. Am J Obstet Gynecol 205: 483.e481–488.

35. Padilla N, Falcon C, Sanz-Cortes M, Figueras F, Bargallo N, et al. (2011) Differential effects of intrauterine growth restriction on brain structure and development in preterm infants: a magnetic resonance imaging study. Brain Res 1382: 98–108.
36. Lodygensky GA, Seghier ML, Warfield SK, Tolsa CB, Sizonenko S, et al. (2008) Intrauterine growth restriction affects the preterm infant's hippocampus. Pediatr Res 63: 438–443.
37. Sanz-Cortes M, Figueras F, Bargallo N, Padilla N, Amat-Roldan I, et al. (2010) Abnormal brain microstructure and metabolism in small-for-gestational-age term fetuses with normal umbilical artery Doppler. Ultrasound Obstet Gynecol 36: 159–165.
38. Pueyo V, Oros D, Valle S, Tuquet H, Guerri N, et al. (2012) Axonal loss and cognitive deficits in term infants born small for gestational age with normal umbilical artery Doppler. Ultrasound Obstet Gynecol 40:297–303.
39. O'Keeffe MJ, O'Callaghan M, Williams GM, Najman JM, Bor W (2003) Learning, cognitive, and attentional problems in adolescents born small for gestational age. Pediatrics 112: 301–307.
40. Larroque B, Bertrais S, Czernichow P, Leger J (2001) School difficulties in 20-year-olds who were born small for gestational age at term in a regional cohort study. Pediatrics 108: 111–115.
41. Galaburda AM, Bellugi U (2000) V. Multi-level analysis of cortical neuroanatomy in Williams syndrome. J Cogn Neurosci 12 Suppl 1: 74–88.
42. Nordahl CW, Dierker D, Mostafavi I, Schumann CM, Rivera SM, et al. (2007) Cortical folding abnormalities in autism revealed by surface-based morphometry. J Neurosci 27: 11725–11735.
43. Depue BE, Burgess GC, Bidwell LC, Willcutt EG, Banich MT (2010) Behavioral performance predicts grey matter reductions in the right inferior frontal gyrus in young adults with combined type ADHD. Psychiatry Res 182: 231–237.
44. Baillieux H, De Smet HJ, Paquier PF, De Deyn PP, Marien P (2008) Cerebellar neurocognition: insights into the bottom of the brain. Clin Neurol Neurosurg 110: 763–773.
45. Als H, Duffy FH, McAnulty G, Butler SC, Lightbody L, et al. (2012) NIDCAP improves brain function and structure in preterm infants with severe intrauterine growth restriction. J Perinatol.
46. Isaacs EB, Fischl BR, Quinn BT, Chong WK, Gadian DG, et al. (2010) Impact of breast milk on intelligence quotient, brain size, and white matter development. Pediatr Res 67: 357–362.
47. Lundqvist-Persson C (2001) Correlation between level of self-regulation in the newborn infant and developmental status at two years of age. Acta Paediatr 90: 345–350.
48. Herschkowitz N, Kagan J, Zilles K (1997) Neurobiological bases of behavioral development in the first year. Neuropediatrics 28: 296–306.
49. Tolsa CB, Zimine S, Warfield SK, Freschi M, Sancho Rossignol A, et al. (2004) Early alteration of structural and functional brain development in premature infants born with intrauterine growth restriction. Pediatr Res 56: 132–138.
50. Olson SL, Bates JE, Sandy JM, Schilling EM (2002) Early developmental precursors of impulsive and inattentive behavior: from infancy to middle childhood. J Child Psychol Psychiatry 43: 435–447.

5. DISCUSSION

5. DISCUSSION

5.1. General overview

This Thesis consists of different studies to test the use of quantitative image features to predict the risk of abnormal clinical outcomes. The Thesis is particularly focused on the use of fetal thorax ultrasound images to predict neonatal respiratory morbidity as a clinical outcome. The work provides evidence supporting the concept that purpose-developed quantitative imaging techniques may help in clinical diagnosis. Specifically, the performance obtained with the software developed to predict neonatal respiratory morbidity was similar to that reported for commercial fetal lung maturity tests in amniotic fluid.

Two different populations have been used in this Thesis to test the transversal usefulness of quantitative image features as a source of clinical information using different image modalities and different outcomes. Firstly, ultrasound images of fetal thorax were used to test quantitative images to predict neonatal respiratory morbidity (Studies 1, 2 and 3). Secondly, fetal brain MR images were used to test quantitative image features using abnormal neurobehavior as clinical outcome (Study 4).

5.2. Quantitative imaging in Fetal Lung Maturity

Our **first study** demonstrates that quantitative image features extracted from fetal thorax ultrasound images correlate with gestational age. Since gestational age is strongly associated with fetal lung maturity, the findings of this study open a pathway for future research evaluating the relationship between texture analysis and lung maturity.

The use of quantitative ultrasound tissue characterization of normal fetal lung development has been investigated in recent years, showing a better accuracy than lung echogenicity to detect histological changes [37-41, 64]. Thus, Sohn et al.[64]

reported a methodology to determine the maturity of fetal lung by comparing the frequency characteristics of lung echoes to those from the fetal liver as a reference organ. Maeda and co-workers evaluated the grey-level histogram width (GLHW) of fetal lung and liver combined with gestational age showing that an increase in GLHW lung-to-liver ration at > 30 weeks of gestational age predicted respiratory distress syndrome with an accuracy of about 80-90%, a sensitivity of 96% and a specificity of 72%, which was comparable with invasive amniotic fluid tests. It was designed in a relatively reduced sample of 22 and 25 fetuses with and without RDS, respectively, but the results were promising [37, 40]. La Torre et al. [41] correlated accurately several patterns of fetal breathing movements with fetal lung maturity tests. Finally, Serizawa and Maeda [40] published a form of tissue characterization named ultrasonic gray level histogram width. Prakash et al. [38] also used lung-to-liver features obtaining an accuracy from 73% to 96% to correctly classify their high pulmonary risk group. Tehesin et al. [39] also evaluated lung-to-liver grey level distribution. The results obtained in our **first study** are in line with these previous studies. The extractor software used provided relevant information about lung structural changes in transverse ultrasound images of fetal lung without the use of any other region or scan (such as liver) as a tissue reference to compare relative echodensities. The study also showed that the features extracted correlated robustly with gestational age with regard to scanner equipment as well as different scanner settings. Furthermore, since no other region was needed as a reference, the method presented might overcome some limitations of previous studies. It is worth mentioning that no blind samples were used to test the predictive capacity of the model. We acknowledge that additional research is necessary to evaluate the ability of the software to be used as an image biomarker predictor of fetal lung maturity.

After testing the texture analysis of fetal lung ultrasound images and its correlation with gestational age, in our **second study** we tested the correlation between texture analyses and the existing fetal lung maturity test[62, 63].Thus the **second study** provided evidence that the image features from lung ultrasound images correlate with

fetal lung maturity test assessed by a standard test as TDx-FLM II. These findings opened the possibility to explore the introduction of non-invasive techniques into clinical practice to test fetal lung maturity. In view of the increasingly recognized importance of respiratory morbidity [27, 28] and the growing numbers of late preterm pregnancies undergoing elective delivery, avoidance of the need for invasive techniques may have a tremendous impact on the clinical management of these cases. These results are in line with previous studies showing that ultrasound images contain non-visible information that can be extracted for clinical purposes. This notion has already been demonstrated for breast and liver disease [14, 15], but the results in fetal lung ultrasound analysis have remained non-conclusive. Among recent studies, some failed to identify differences in the patterns of features assessed with quantitative ultrasound [37, 64]. However, no significant differences were observed above 32 weeks' gestation, precluding its use in clinical practice. In a similar manner than in our first study, the method tested here did not use direct gray level or other tissue references such as the liver. This may represent a substantial advantage for clinical use as compared with the previous quantitative analyses studies. The particular properties of the evaluated software (it seems to be unaffected by moderate changes in the acquisition settings) might facilitate its inclusion in clinical practice. It is important to note that in this study we used a surrogate of fetal lung maturity, such as amniotic fluid analysis. Taking into account that TDx-FLM II has also suboptimal sensitivity and specificity (89% and 83%, respectively) [65] for the prediction of RDS, it would be of great interest to assess whether non-invasive evaluation of the fetal lung texture might provide more precise information. Thus, the results of this study supported further research on ultrasound texture analysis as a non-invasive quantitative imaging biomarker to predict the clinical outcome, neonatal respiratory morbidity, which was to be tested in our **third study**.

At this point, the ability of texture analysis of fetal lung ultrasound images to blindly predict the risk of neonatal respiratory morbidity was yet not demonstrated. In addition, the software used in study 1 and 2 presented some problems of robustness

when it was tested blindly in new samples. To address this limitation, we decided to develop a completely new method from the beginning, by abandoning all previous approaches and testing systematically all available quantitative feature extractors and machine learning approaches, in order to obtain a combination that offered the best possible robustness. We termed this new methodology quantitative ultrasound fetal lung maturity analysis (quantusFLM™), which combines in a unique and innovative manner several steps of different image texture extractor and machine learning algorithms. This, in the **third study** we described the basic principles of this novel method, and the results of a validation study to assess the ability of the method to blindly predict the risk of neonatal respiratory morbidity. Remarkably, this study demonstrated that the performance of quantusFLM™ was comparable to that reported with the use of current tests in amniotic fluid [66-70]. Of note, the average sample size used in these studies was 167 (ranging from 28 to 301), which is similar to the one used in this study (144). The clinical implications of this study are important, as it opens the possibility of using non-invasive approaches for the prenatal prediction of fetal lung maturity. We acknowledge that a larger sample size should be used to test the performance within narrower gestational age ranges. Currently, a multicentre international study to validate the results is underway. In summary, this study provides evidence that a purpose-developed software based on quantitative texture analysis of fetal lung ultrasound images predicts neonatal respiratory morbidity with a similar performance to that reported for commercial fetal lung maturity tests in amniotic fluid.

5.3. Quantitative imaging in Fetal Brain

In our **fourth study**, we evaluated the ability of image texture analysis to detect abnormalities in different fetal brain areas, and tested their association with abnormal neonatal neurobehavior. This study demonstrated the potential of quantitative imaging texture analysis for other image acquisition techniques and clinical outcomes. The study provided evidence that fetal brain MRI textural patterns were associated

with neonatal neurobehavior and set the basis for further research on in utero imaging biomarkers based on quantitative image analysis.

The correlation between quantitative texture analysis and functional outcome had previously been demonstrated in adults with neurological conditions and apparently normal MRI scans, such as in mild traumatic brain injury or mild cognitive impairment [5, 71]. In these conditions, texture analysis was able to identify differences in relation with the progression of the disease and indicate the most affected areas. To our best knowledge, this is the first time in which brain quantitative imaging in fetuses has been used to establish associations with post-natal neurobehavior. The results are in line with the existence of brain reorganization in IUGR (intrauterine growth restriction). Different lines of evidence have shown that fetuses and infants affected with early and severe intrauterine growth restriction have significant differences in brain metabolism, sulcation, composition, and microstructure [72-75]. Furthermore, correlations between these brain disturbances and neurological performance have been reported [72, 74]. Along the same lines, in a previous study we reported differences in textural patterns on fetal brain MRI between term SGA and AGA fetuses [59]. In the present study we provide evidence that these patterns are correlated with post-natal neurobehavior. These changes could be identified by quantitative imaging in order to define ‘‘early endophenotypes’’ as markers of future functional outcome [72]. From a clinical perspective, this provides further evidence to support the existence of changes in brain development, which could be used for diagnosis of true forms of fetal growth restriction in utero. Identifying at-risk patients lays the basis for timely interventions in utero to decrease the rate of adverse perinatal results [50] and for selection of newborns for targeted interventions. The strength of this study is that it evaluates brain MRI texture analysis from a homogeneous cohort of term SGA fetuses selected in utero and prospectively followed up until the neonatal period. We acknowledge that these results are preliminary and require confirmation in larger sample sizes allowing external validation.

In summary, this study provided evidence that fetal brain quantitative imaging based on MRI quantitative texture analysis has a potential in predicting an abnormal neurobehavioral outcome. However, the potential clinical value of quantitative texture analysis in the identification of risk requires a great deal of further research. At this point, most quantitative imaging-based methods are still far from clinical applications.

5.4. Limitations and technical considerations

Several study limitations and technical considerations in relation with the work here presented should be mentioned. In the first study, the frequency of acquisition was not always available. This, we could not rule out that scanner frequency was not a critical issue for blind validation using new samples. To support this notion, when the extractor method [56] used in Study 1 and 2 was tested using databases with images acquired with controlled parameters (such as the spatial resolution of the images) it showed a very poor robustness to slight variations in image characteristics.

Each study provided a further step towards the goal of developing a clinically relevant proof of principle. In the first study the outcome used was gestational age, a surrogate measure of fetal lung maturity. In the second study, the outcome also was a surrogated of fetal lung maturity but, in this case, we used the amniotic test result. TDx-FLM II [65] is a better surrogated value than gestational age but it also has a limited sensitivity and specificity (89% and 83%, respectively) for the prediction of respiratory distress syndrome. However we acknowledge that these results require confirmation with clinical outcome. Study 3 overcomes this limitation.

Another limitation, present in Study 1, 2 and 4 was the lack of a blind validation of the method, mainly because of the reduced number of the available samples. Texture analysis-based applications require developing robust algorithms based on large databases, software user interface platforms and feasibility studies demonstrating its value in clinical practice. In contract, Study 3 used more than 1,000 real ultrasound images, and more than 13,000 from other databases [76, 77], to develop

quantusFLM™. However, we acknowledge that the sample used for blind validation should be expanded. Thus, only 144 completely new samples could be used to test the performance of the algorithm. The algorithm was evaluated in only one center and the image acquisition and delineation were performed by highly trained personnel in a clinical research setting. To overcome these limitations a multicenter international study is now underway.

5.5. Concluding remarks and future work

This Thesis consists of different studies focused in advancing towards the development of non-invasive imaging biomarkers to predict perinatal clinical outcomes. The majority of the work was focused on developing a quantitative imaging biomarker for neonatal respiratory morbidity.

To our knowledge, the third study of this Thesis is the first study reporting blindly validation of quantitative imaging analysis software specifically designed to predict neonatal respiratory morbidity.

Future work related with neonatal respiratory morbidity will be focused on the undergoing multicentre international study and on continuous improvement, for example, performing automatic image lung segmentation. Other lines will be explored to test the transversal value of the methodology in other clinical conditions where quantitative imaging analysis could improve the current clinical practice, essentially by reducing the need of invasive diagnostic procedures.

6. CONCLUSIONS

6. CONCLUSIONS

1. Quantitative image analysis can extract information from subtle differences in the texture of the fetal lung ultrasound images that are related to gestational age tissue microstructure changes and consequently, to the pathophysiological process of fetal lung maturation.

2. Fetal lung ultrasound texture analysis correlates with fetal lung maturity assessed by laboratory standard methods in amniotic fluid.

3. Quantitative image analysis of fetal lungs can predict neonatal respiratory morbidity with similar performance to currently used laboratory methods in amniotic fluid.

4. Non-invasive texture analysis could be used in magnetic resonance to identify patterns associated with abnormal neurobehavior in small-for-gestational-age babies.

7. REFERENCES

7. REFERENCES

1. Hope, T., et al. *Ultrasonic tissue characterization as a predictor of white matter damage: results of a preliminary study*. in *Ultrasonics Symposium, 2004 IEEE*. 2004: IEEE.
2. Hope, T., et al. *Texture-based tissue characterization: a novel predictor for brain injury*. in *Proceedings of the Second IASTED International Conference on Biomedical Engineering. Calgary, Alberta, Canada: International Association of Science and Technology for Development*. 2004.
3. Bruno, A., et al., *Texture analysis in medical imaging.* STUDIES IN HEALTH TECHNOLOGY AND INFORMATICS, 1997: p. 133-164.
4. Bergen, J.R. and E.H. Adelson, *Early vision and texture perception.* Nature, 1988. **333**(6171): p. 363-364.
5. Holli, K.K., et al., *Texture analysis of MR images of patients with mild traumatic brain injury.* BMC medical imaging, 2010. **10**(1): p. 8.
6. Allison, J.W., et al., *Understanding the process of quantitative ultrasonic tissue characterization.* Radiographics, 1994. **14**(5): p. 1099-1108.
7. Insana, M.F., et al., *Quantitative ultrasonography.* Medical progress through technology, 1988. **15**(3-4): p. 141-153.
8. Castellano, G., et al., *Texture analysis of medical images.* Clinical radiology, 2004. **59**(12): p. 1061-1069.
9. Neuroradiology, A.S.o., *ACR-ASNR Practice Guideline for the Performance and Interpretation of Magnetic Resonance Imaging (MRI) of the Brain.* 2013.
10. Lizzi, F.L., et al., *Theoretical framework for spectrum analysis in ultrasonic tissue characterization.* The Journal of the Acoustical Society of America, 1983. **73**(4): p. 1366-1373.
11. Kolios, M.C., et al. *An investigation of backscatter power spectra from cells, cell pellets and microspheres*. in *Ultrasonics, 2003 IEEE Symposium on*. 2003: IEEE.
12. Tunis, A.S., et al., *Monitoring structural changes in cells with high-frequency ultrasound signal statistics.* Ultrasound in medicine & biology, 2005. **31**(8): p. 1041-1049.
13. Szabo, T.L., *Diagnostic ultrasound imaging: inside out*. 2004: Academic Press.
14. Chen, D.-R., et al., *Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks.* Ultrasound in medicine & biology, 2002. **28**(10): p. 1301-1310.
15. Wan, C., et al., *Evaluation of breast lesions by contrast enhanced ultrasound: Qualitative and quantitative analysis.* European journal of radiology, 2012. **81**(4): p. e444-e450.
16. Hartman, P.C., et al., *Variability of quantitative echographic parameters of the liver: intra-and interindividual spread, temporal-and age-related effects.* Ultrasound in medicine & biology, 1991. **17**(9): p. 857-867.
17. Kadah, Y.M., et al., *Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images.* Medical Imaging, IEEE Transactions on, 1996. **15**(4): p. 466-478.

18. Icer, S., A. Coskun, and T. Ikizceli, *Quantitative grading using grey relational analysis on ultrasonographic images of a fatty liver.* Journal of medical systems, 2012. **36**(4): p. 2521-2528.
19. Sandrin, L., et al., *Shear elasticity probe for soft tissues with 1-D transient elastography.* Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on, 2002. **49**(4): p. 436-446.
20. P(ed.)., R., *Magnetic Resonance in Medicine. The Basic Textbook of the European Magnetic Resonance Forum.6th edition.* 2012.
21. Kjaer, L., et al., *Texture analysis in quantitative MR imaging: tissue characterisation of normal brain and intracranial tumours at 1.5 T.* Acta Radiologica, 1995. **36**(2): p. 127-135.
22. Freeborough, P.A. and N.C. Fox, *MR image texture analysis applied to the diagnosis and tracking of Alzheimer's disease.* Medical Imaging, IEEE Transactions on, 1998. **17**(3): p. 475-478.
23. Bonilha, L., et al., *Texture analysis of hippocampal sclerosis.* Epilepsia, 2003. **44**(12): p. 1546-1550.
24. Holli, K., et al., *Characterization of breast cancer types by texture analysis of magnetic resonance images.* Academic radiology, 2010. **17**(2): p. 135-141.
25. Jirak, D., et al., *Texture analysis of human liver.* Journal of Magnetic Resonance Imaging, 2002. **15**(1): p. 68-74.
26. Szczypinski, P.M., M. Strzelecki, and A. Materka. *Mazda-a software for texture analysis*. in *Information Technology Convergence, 2007. ISITC 2007. International Symposium on*. 2007: IEEE.
27. Teune, M.J., et al., *A systematic review of severe morbidity in infants born late preterm.* American journal of obstetrics and gynecology, 2011. **205**(4): p. 374. e1-374. e9.
28. Spong, C.Y., et al., *Timing of indicated late-preterm and early-term birth.* Obstetrics and gynecology, 2011. **118**(2 Pt 1): p. 323.
29. Hibbard, J.U., et al., *Respiratory morbidity in late preterm births.* JAMA, 2010. **304**(4): p. 419-425.
30. Tita, A.T.N., et al., *Timing of elective repeat cesarean delivery at term and neonatal outcomes.* New England Journal of Medicine, 2009. **360**(2): p. 111-120.
31. Clark, S.L., et al., *Neonatal and maternal outcomes associated with elective term delivery.* American journal of obstetrics and gynecology, 2009. **200**(2): p. 156. e1-156. e4.
32. Gluck, L., et al., *Diagnosis of the respiratory distress syndrome by amniocentesis.* Obstetrical & Gynecological Survey, 1971. **26**(10): p. 708-710.
33. Gluck L, K.M., *Lecithin-sphingomyelin ratios in amniotic fluid in normal and abnormal pregnancy.* Am J Obstet Gynecol., 1973. **115**(4): p. 539-46.
34. Neerhof, M.G., et al., *Lamellar body counts: a consensus on protocol.* Obstetrics & Gynecology, 2001. **97**(2): p. 318-320.

35. Besnard, A.E., et al., *Lecithin/sphingomyelin ratio and lamellar body count for fetal lung maturity: a meta-analysis.* European Journal of Obstetrics & Gynecology and Reproductive Biology, 2013. **169**(2): p. 177-183.
36. ACOG, *Practice Bulletin No. 97: Fetal lung maturity.* Obstetrics and gynecology, 2008. **112**(3): p. 717-26.
37. Maeda, K., et al., *Echogenicity of fetal lung and liver quantified by the grey-level histogram width.* Ultrasound in medicine & biology, 1999. **25**(2): p. 201-208.
38. Bhanu Prakash, K.N., et al., *Fetal lung maturity analysis using ultrasound image features.* Information Technology in Biomedicine, IEEE Transactions on, 2002. **6**(1): p. 38-45.
39. Tekesin, I., et al., *Assessment of fetal lung development by quantitative ultrasonic tissue characterization: a methodical study.* Prenatal diagnosis, 2004. **24**(9): p. 671-676.
40. Serizawa, M. and K. Maeda, *Noninvasive fetal lung maturity prediction based on ultrasonic gray level histogram width.* Ultrasound in medicine & biology, 2010. **36**(12): p. 1998-2003.
41. La Torre, R., et al., *Preliminary report on a new and noninvasive method for the assessment of fetal lung maturity.* Journal of perinatal medicine, 2003. **31**(5): p. 431-434.
42. Cosmi, E.V., et al., *Ultrasonographic patterns of fetal breathing movements in normal pregnancy.* International Journal of Gynecology & Obstetrics, 2003. **80**(3): p. 285-290.
43. Bernstein, I.M., et al., *Morbidity and mortality among very-low-birth-weight neonates with intrauterine growth restriction.* American journal of obstetrics and gynecology, 2000. **182**(1): p. 198-206.
44. Gagnon, R. and M. Van den Hof, *The use of fetal Doppler in obstetrics.* Journal of obstetrics and gynaecology Canada: JOGC= Journal d'obstetrique et gynecologie du Canada: JOGC, 2003. **25**(7): p. 601-14; quiz 615-6.
45. Royal College of Obstetrics and Gynaecology, G.-t.G., *The Investigation and Management of the Small-for-Gestational-Age Fetus.* 2002.
46. opinion, A.c., *Utility of antepartum umbilical artery Doppler velocimetry in intrauterine growth restriction. Number 188.* Committee on Obstetric Practice. American College of Obstetricians and Gynecologists., 1997.
47. Rees, S., et al., *Fetal brain injury following prolonged hypoxemia and placental insufficiency: a review.* Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 1998. **119**(3): p. 653-660.
48. Godfrey, K.M. and D.J.P. Barker, *Fetal programming and adult health.* Public health nutrition, 2001. **4**(2b): p. 611-624.
49. Doctor, B.A., et al., *Perinatal correlates and neonatal outcomes of small for gestational age infants born at term gestation.* American journal of obstetrics and gynecology, 2001. **185**(3): p. 652-659.
50. Cruz-Martinez, R., et al., *Fetal brain Doppler to predict cesarean delivery for nonreassuring fetal status in term small-for-gestational-age fetuses.* Obstetrics & Gynecology, 2011. **117**(3): p. 618-626.

51. Figueras, F., et al., *Predictiveness of antenatal umbilical artery Doppler for adverse pregnancy outcome in smallâ€ forâ€ gest□tion□lâ€ □ge b□bies □ccording to customised birthweight centiles: popul□tionâ€ b□sed study.* BJOG: An International Journal of Obstetrics & Gynaecology, 2008. **115**(5): p. 590-594.
52. Figueras, F., et al., *Neurobehavior in term, small-for-gestational age infants with normal placental function.* Pediatrics, 2009. **124**(5): p. e934-e941.
53. Padidela, R.N. and V. Bhat, *Neurobehavioral assessment of appropriate for gestational and small for gestational age babies.* Indian pediatrics, 2003. **40**(11): p. 1063-1068.
54. Figueras, F., et al., *Small-for-gestational-age fetuses with normal umbilical artery Doppler have suboptimal perinatal and neurodevelopmental outcome.* European Journal of Obstetrics & Gynecology and Reproductive Biology, 2008. **136**(1): p. 34-38.
55. De Vries, L.S., *Neurological assessment of the preterm infant.* Acta Paediatrica, 1996. **85**(7): p. 765-771.
56. Tenorio, V., et al., *Correlation between a semiautomated method based on ultrasound texture analysis and standard ultrasound diagnosis using white matter damage in preterm neonates as a model.* Journal of Ultrasound in Medicine, 2011. **30**(10): p. 1365-1377.
57. Bonet-Carne, E., et al. *Evaluation of semiautomated quantification of cranial ultrasound images in newborns as a predictor of Neonatal Behavioral Assessment Scale*. in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*. 2011: IEEE.
58. Nongena, P., et al., *Confidence in the prediction of neurodevelopmental outcome by cranial ultrasound and MRI in preterm infants.* Archives of Disease in Childhood-Fetal and Neonatal Edition, 2010. **95**(6): p. F388-F390.
59. Sanz-Cortes, M., et al., *Fetal brain MRI texture analysis identifies different microstructural patterns in adequate and small for gestational age fetuses at term.* Fetal diagnosis and therapy, 2013. **33**(2): p. 122-129.
60. Hope, T.A. and S.n.E. Iles, *Technology review: the use of electrical impedance scanning in the detection of breast cancer.* Breast Cancer Research, 2004. **6**(2): p. 69-74.
61. Hope, T.A., et al., *Selecting and assessing quantitative early ultrasound texture measures for their association with cerebral palsy.* Medical Imaging, IEEE Transactions on, 2008. **27**(2): p. 228-236.
62. Wijnberger, L.D.E., et al., *Prediction of fetal lung immaturity using gestational age, patient characteristics and fetal lung maturity tests: a probabilistic approach.* Archives of gynecology and obstetrics, 2010. **281**(1): p. 15-21.
63. Grenache, D.G. and A.M. Gronowski, *Fetal lung maturity.* Clinical biochemistry, 2006. **39**(1): p. 1-10.
64. Sohn, C.H., W. Stolz, and G. Bastert, *Diagnosis of fetal lung maturity by ultrasound: a new method and first results.* Ultrasound in Obstetrics & Gynecology, 1991. **1**(5): p. 345-348.

65. Bennasar, M., et al., *Gestational age-specific cutoff levels of TDX-FLM II for the prediction of neonatal respiratory distress syndrome.* Fetal diagnosis and therapy, 2009. **25**(4): p. 392-396.
66. Wijnberger, L.D.E., et al., *The accuracy of lamellar body count and lecithin/sphingomyelin ratio in the prediction of neonatal respiratory distress syndrome: a metaâ€ analysis.* BJOG: An International Journal of Obstetrics & Gynaecology, 2001. **108**(6): p. 583-588.
67. Karcher, R., et al., *Gestational age-specific predicted risk of neonatal respiratory distress syndrome using lamellar body count and surfactant-to-albumin ratio in amniotic fluid.* American journal of obstetrics and gynecology, 2005. **193**(5): p. 1680-1684.
68. Hagen, E., J.C. Link, and F. Arias, *A comparison of the accuracy of the TDx-FLM assay, lecithin-sphingomyelin ratio, and phosphatidylglycerol in the prediction of neonatal respiratory distress syndrome.* Obstetrics & Gynecology, 1993. **82**(6): p. 1004-1008.
69. Russell, J.C., et al., *Multicenter evaluation of TDx test for assessing fetal lung maturity.* Clinical chemistry, 1989. **35**(6): p. 1005-1010.
70. Haymond, S., et al., *A direct comparison between lamellar body counts and fluorescent polarization methods for predicting respiratory distress syndrome.* American journal of clinical pathology, 2006. **126**(6): p. 894-899.
71. De Oliveira, M.S., et al., *MR imaging texture analysis of the corpus callosum and thalamus in amnestic mild cognitive impairment and mild Alzheimer disease.* American Journal of Neuroradiology, 2011. **32**(1): p. 60-66.
72. Dubois, J.r.m., et al., *Primary cortical folding in the human newborn: an early marker of later functional development.* Brain, 2008. **131**(8): p. 2028-2041.
73. Story, L., et al., *Brain metabolism in fetal intrauterine growth restriction: a proton magnetic resonance spectroscopy study.* American journal of obstetrics and gynecology, 2011. **205**(5): p. 483. e1-483. e8.
74. Padilla, N., et al., *Differential effects of intrauterine growth restriction on brain structure and development in preterm infants: a magnetic resonance imaging study.* Brain research, 2011. **1382**: p. 98-108.
75. Lodygensky, G.A., et al., *Intrauterine growth restriction affects the preterm infant's hippocampus.* Pediatric research, 2008. **63**(4): p. 438-443.
76. Texture lab, H.-W.U., Edinburgh, UK, *PhoTex database*. 2014.
77. Ojala, T., et al. *Outex-new framework for empirical evaluation of texture analysis algorithms*. in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*. 2002: IEEE.

8. ACKNOWLEDGMENTS

8. ACKNOWLEDGEMENTS

Financial support

We (Elisenda Bonet Carné and Transmural Biotech S.L.) thank the Comissionat per a Universitat i Recerca del Departament d'Innovació, Universitats i Empreses de la Generalitat de Catalunya, for its financial support.

Study 1 and 2 of this Thesis was supported by a Rio Hortega grant (Dr. Cobo) from Instituto de Salud Carlos III in Spain (CM09/00213). This project (STUDY 2) was also supported by The People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7 under REA Gran Agreement no. 217911 (u-Volumes). Study 4 was supported by grants from The Cerebra Foundation for the Brain Injured Child (Carmarthen, Wales, UK), the Thrasher Research Fund (Salt Lake City, USA), the Fundacio Dexeus (Barcelona, Spain). Dr. Sanz-Cortes was supported by a Rio Hortega research fellowship and Dr. Padilla was supported by Instituto de Salud Carlos III Sara Borrell.

I would like to thank to all the people who have made this Thesis possible, to those who have help me during all this period in so many ways. My sincere thanks.

I would also like to thank my directors, Eduard Gratacós and Ferran Marqués, as they both provided me with many great points to include and gave me advice whenever it was required. To Eduard, for being the best example to follow, for encouraging my research and for allowing me to grow as a researcher and as a person during these years. To Ferran, for sharing your precious time, for your excellent comments and suggestions and for your encouragement during these years.

To all co-authors of the papers whose work has been essential in order to achieve this Thesis and from whom I could learnt. And of course, to all the technicians that performed the acquisitions.

Voldria agrair a tothom que m'ha ajudat durant aquests anys, després de tot aquest temps es fa difícil escollir com agrair-vos als que d'alguna manera o d'una altra m'heu acompanyat en aquest camí, és massa complicat posar-ho en paraules. Tot i això, m'agradaria nombrar sense deixar-me a ningú, a aquelles persones que han tingut un paper especial durant aquest temps.

A la Montse, per ensenyar-me i explicar-me sense desesperar-se, per resoldre tots els dubtes rellevants o no sobre la prematuritat i la vida, per contagiar-me el seu entusiasme en tot, i per guiar-me en els moments més desesperants.

Al Luis, per la seva paciència, per donar-me suport en el necessari i per animar-me quan ho he necessitat.

A totes les persones del grup d'investigació de medicina fetal i perinatal i de Transmural Biotech, perquè després de tants anys, hi ha hagut moments millors o pitjors, però és on he après la major del què sé i on he anat evolucionant gràcies a tots vosaltres. En especial a Elisenda i la Míriam, per despertar-me la curiositat en el món de la patologia fetal, per les anècdotes, pel rigor científic que m'han ensenyat i per moltes coses més. A la Tere i la Magda, per la seva paciència i les seves ensenyances. Al Francesc, pel seu rigor estadístic. I a la Violeta amb qui vaig començar l'aventura. També, als enginyers amb els que vam començar, el Joan, la Patri, l'Àlvaro i al Dafnis, per les estones freaks, els sopars "enginyerils" i les birretes. Sobretot al Daf, per compartir tot aquest camí, m'hagués desesperat sense el teu suport. A la Pat, la Mercè, la Mara i la Clara, pels riures, pels cafès, per alegrar-me els dies. Al Roc, per la seva visió particular. A l'Àlvaro, per les cerveses, cafès, entrepans a les tantes, els moments d'alegria i de tristor, per compartir els últims anys. A la Maite, Gemma, Balma, Pere... Uf, sou tants...

Als amics, per no deixar-me tirar la tovallola i per fer-me riure.

A l'Ari i la Lluïsa, pels vespres, per dir-nos-ho tot o no dir-nos res, per estar al meu costat en tot moment, per ser les Teres. Al trio més divertit, a la Marteta, l'amiga de telecos, a l'Arnau i a la nova incorporació, al Manel, perquè sou genials. Al Valencia, perquè tot i veure'ns poc procurem fer-ho el suficient com perquè sembli que l'últim cop era ahir. A la Queralt i a la Mireia, perquè passen els anys però seguim com sempre.

A totes aquelles persones que, tot i no compartir moments de feina, m'han fet somriure i m'han donat l'energia suficient per tirar endavant, a l'Ari Recasens i la Silvia del Màster, als del grup de la Uni, a les demés nenes de Manresa, a les de Badalona, del futbol... moltes gràcies a totes.

Al tota la meva família, en especial als meus pares, pel seu suport incondicional i pels seus ànims. Per guiar-me quan ho he necessitat però també per deixar-me equivocar. A la meva germana petita, per ser tan alegre, per compartir tots els moments, per deixar el que estàs fent per escoltar les meves tonteries.

Per aguantar-me quan no m'hagués aguantat ni jo, el meu més sincer agraïment per l'Albert. Per ser com ets i, sobretot, per deixar-me ser com sóc.

9. APPENDIX: RESUM EN CATALÀ

(SUMMARY IN CATALAN LANGUAGE)

9. APPENDIX: RESUM EN CATALÀ

(SUMMARY IN CATALAN LANGUAGE)

9.1. Introducció

Les tècniques d'imatge s'utilitzen en medicina per crear imatges de les parts del cos humà que estan ocultes per la pell i els ossos. Les imatges mèdiques es poden utilitzar tant per diagnosticar com per a tractar malalties; aquesta Tesi es centra en el tractament quantitatiu d'imatges mèdiques com a tècnica de diagnòstic.

L'objectiu principal dels estudis de diagnòstic basats en imatges és la caracterització dels teixits: les imatges s'adquireixen per tal de determinar si els teixits en l'àrea d'estudi son normals (teixit sa) o patològics. L'anàlisi quantitatiu d'imatges mèdiques pot augmentar la repetibilitat i ajudar a resoldre les ambigüitats en la interpretació de diferents imatges. Algunes disciplines, com la medicina Materno-Fetal, són candidates perfectes per l'anàlisi quantitatiu d'imatge com a eina de diagnòstic a causa de la manca d'accessibilitat als teixits fetals.

Encara que hi ha moltes tècniques d'imatge, l'ecografia s'utilitza preferentment entre els obstetres i esdevé una tecnologia principal de diagnòstic en la medicina Materno-Fetal, sobretot perquè és no-invasiva, no requereix d'instal·lacions especials, es transportable i el seu ús és econòmic. Diversos estudis demostren que les característiques de la senyal d'ultrasò estan relacionades amb la microestructura del teixit i que diferents arquitectures tissulars poden es poden reconèixer i interpretar en les imatges. Des de principis dels anys 80 la idea de que l'anàlisi qualitatiu d'ultrasò pot augmentar la repetibilitat i ajudar en la resolució d'ambigüitats s'ha generalitzat. Per altra banda, la ressonància magnètica és la tècnica més comú per a la neuroimatge estructural. També és una tècnica no invasiva que ofereix un gran contrast entre la matèria blanca i la matèria gris, per aquesta raó esdevé molt útil per avaluar el sistema nerviós central. També s'han realitzat diversos estudis per investigar l'ús de la ressonància magnètica quantitativa per a diferents aplicacions mèdiques de diagnòstic.

Aquesta tesi doctoral forma part de dos projectes més amplis i s'utilitzen dues patologies diferents. D'una banda, es van utilitzar imatges d'ultrasò del tòrax fetal per als estudis de maduració pulmonar fetal (Estudis 1, 2 i 3). D'altra banda, l'Estudi 4 forma part d'un estudi sobre els fetus petits-per-edat-gestacional (PEG). A continuació hi ha una breu explicació dels problemes clínics associats als dos projectes.

9.1.1. Pulmó Fetal

La causa més comuna de mortalitat i morbiditat en prematurs i en fetus nascuts a precoçment a terme és la immaduresa pulmonar que causa morbiditat respiratòria neonatal, definida com el distrés respiratori o la taquipnea transitòria del nounat. Actualment, l'avaluació de la maduresa pulmonar fetal es realitza mitjançant proves de laboratori del líquid amniòtic, que requereixen la realització d'un procediment invasiu per obtenir les mostres, l'amniocentesi. Aquesta, a part de ser arriscada, no es pot realitzar en tots els entorns clínics a causa de les instal·lacions requerides. Per això, la necessitat de l'amniocentesi ha resultat en una disminució en l'ús d'aquesta informació clínicament. L'avaluació de la maduresa pulmonar fetal amb mètodes no invasius és un problema no resolt malgrat hi ha hagut més de 20 anys d'investigació en el desenvolupament de solucions basades en imatges d'ultrasò. Els estudis realitzats anteriorment suggereixen que l'anàlisi quantitatiu d'imatge d'ultrasò per a predir la maduresa pulmonar fetal té potencial, però la precisió en el diagnòstic no és suficient per a l'ús clínic.

Per tant, tot i avançar en els factors de risc i la introducció d'estratègies de salut pública destinades a reduir el risc de morbiditat respiratòria neonatal la seva predicció segueix sent un repte.

9.1.2. Cervell Fetal

Un 10% dels embarassos resulten en un nounat petit-per-edat-gestacional (PEG). Encara que alguns fetus amb aquest diagnòstic són constitucionalment petits, en una proporció substancial dels casos, el diagnòstic de PEG identifica les formes lleus de la restricció de creixement intrauterí. Per tant, en aquests casos, el desenvolupament

fetal es produeix en condicions subòptimes, amb un subministrament que priva de l'oxigen i nutrients necessaris al cervell fetal. Diferents autors han demostrat que tot i que la majoria de PEG no presenten signes de deteriorament, hi ha una proporció que presenta un augment en el risc de patir un resultat perinatal advers amb un neurocomportament anormal neonatal i un desenvolupament neurològic deteriorat durant la primera infància.

Tenint en compte la seva prevalença, el diagnòstic dels PEG constitueix un repte i una oportunitat per millorar l'impacte de les condicions prenatals en la qualitat de vida. No obstant, en l'actualitat la detecció dels PEG que tenen risc de desenvolupament neurològic anormal és limitada ja que els exàmens clínics estàndards fallen en la identificació de diferencies significatives en aquest subgrup.

9.2. Objectius

Les tècniques d'imatge, com l'ultrasò o la imatge per ressonància magnètica, s'utilitzen extensament com a eines de diagnòstic mèdic. De fet, l'ultrasò és la tècnica de diagnòstic per imatge més utilitzada en el camp de la medicina Materno-Fetal ja que permet realitzar imatges de forma no invasiva, la màquina per a l'adquisició és portable i el seu ús econòmic. Encara que l'ultrasò ha assolit el seu màxim potencial per al diagnòstic de canvis macroestructurals, les imatges d'ultrasò tenen molta més informació que l'ull humà no és capaç de distingir. Aquesta proposta de Tesi té com a principal objectiu investigar l'ús potencial de biomarcadors d'imatge, utilitzant l'anàlisi quantitatiu de textures, per millorar la interpretació subjectiva de les imatges per tal de predir diferents alteracions en la població fetal.

La hipòtesi principal d'aquest projecte és que els mètodes d'anàlisi de la textura de la imatge es podrien utilitzar per l'anàlisi d'imatges mèdiques (per exemple, en ecografies o imatges de ressonància magnètica) en el camp de la medicina fetal, per caracteritzar informació de la microestructura que no es pot detectar mitjançant una avaluació clínica estàndard. L'ecografia del pulmó fetal per l'avaluació de la maduresa pulmonar fetal és un candidat òptim per a desenvolupar aquests mètodes.

Les hipòtesis específiques són:

- L'anàlisi quantitatiu de la textura de les imatges d'ultrasò del pulmó fetal permet l'extracció de forma reproduïble de característiques del teixit que es correlacionen amb l'edat gestacional del fetus.
- L'anàlisi quantitatiu de la textura de les imatges d'ultrasò del pulmó fetal permet l'extracció de característiques que es correlacionen amb els resultats de maduració pulmonar obtinguts amb les proves realitzades en líquid amniòtic.
- L'anàlisi quantitatiu de la textura de les imatges d'ultrasò del pulmó fetal es pot utilitzar per desenvolupar biomarcadors d'imatge per predir el risc de morbiditat respiratòria neonatal.
- Els mètodes d'anàlisi quantitatiu de la textura de les imatges es poden utilitzar per avaluar altres regions del fetus i, altres tècniques d'adquisició d'imatge (com per exemple, imatges de ressonància magnètica), poden ser utilitzats per identificar patrons associats amb els canvis que es produeixen en el desenvolupament del cervell fetal deguts a la restricció de creixement intrauterí.

L'objectiu principal d'aquest projecte és explorar i testar el desenvolupament de nous mètodes basats en l'anàlisi quantitatiu de la textura de les imatges mèdiques (per exemple, imatges d'ultrasò i de ressonància magnètica) en el camp de la medicina fetal, principalment per l'avaluació de maduresa pulmonar fetal i del cervell fetal.

Els objectius específics són:

- Avaluar la relació entre la textura d'imatges d'ultrasò del pulmó fetal i un procés fisiològic ben definit, les etapes de la maduració pulmonar durant la gestació.
- Avaluar la possibilitat d'utilitzar mètodes quantitatius d'anàlisis d'imatges d'ultrasò del pulmó fetal i la seva correlació amb els resultats de la maduració

pulmonar fetal obtinguts amb els mètodes estàndard que requereixen de l'anàlisi de líquid amniòtic.

- Desenvolupar i avaluar el rendiment d'un nou mètode per predir el risc de morbiditat respiratòria neonatal basat en l'anàlisi quantitatiu de textures d'imatges d'ultrasò pulmonars fetals.
- Provar si l'anàlisi quantitatiu de la textura en imatges de ressonància magnètica podria identificar patrons associats amb un neurocomportament anormal en nadons petits l'edat gestacional, per avaluar el comportament transversal de l'anàlisi quantitatiu d'imatges.

Aquesta tesi s'ha dividit en quatre estudis i s'ha estructurat seguint la normativa per a la tesi doctoral, com a compendi de publicacions, per obtenir el grau internacional de Doctor en Biomedicina. Els estudis inclosos en la Tesi pertanyen a la mateixa línia d'investigació i consten de quatre articles ja publicats o presentats per a la seva publicació en revistes científiques internacionals.

9.3. Mètodes i Resultats

Per tal d'assolir els objectius d'aquesta tesi i investigar possibles biomarcadors d'imatges per a la medicina Materno-Fetal és necessari disposar d'un model patològic clar i de grans bases de dades d'imatges. A causa de la manca de mostres i de la incertesa associada als resultats de neurodesenvolupament, l'estudi del cervell fetal no seria un candidat adequat per a ser utilitzat com a model per la hipòtesi principal. Els problemes respiratoris neonatals, que es presenten poques hores després del naixement, poden proporcionar grans bases de dades sense haver d'esperar per al resultat clínic. Per aquest motiu, la morbiditat respiratòria neonatal s'utilitzarà com el principal model patològic d'aquest projecte (Estudis 1, 2 i 3) i les imatges de ressonància magnètica del cervell fetal s'utilitzaran per provar la transversalitat de l'anàlisi quantitatiu de la textura de la imatge en un altre model patològic (Estudi 4).

9.3.1. Estudi 1. Viabilitat i reproductibitat de l'anàlisi automàtic de textures del pulmó fetal i la seva correlació amb l'edat gestacional.

Corresponent a l'article: T. Cobo, E. Bonet-Carne, M. Martinez-Terron, A. Perez-Moreno, N. Elias, J. Luque, I. Amat-Roldan, M. Palacio. Feasibility and Reproducibility of Fetal Lung Texture Analysis by Automatic Quantitative Ultrasound Analysis and Correlation with Gestational Age. Fetal Diagn Ther. 2012 Apr; 31(4):230-6.

* T.C. and E.B.-C. han contribuït equitativament en aquest treball.

Estat: publicat

Factor d'impacte de la revista: 1.902

Quartil: 2n, àrea Obstetricia i Ginecologia.

Objectiu: Avaluar la factibilitat i reproductibilitat de l'anàlisi de la textura del pulmó fetal utilitzant un nou mètode quantitatiu d'anàlisi d'ultrasò i, avaluar la seva correlació amb l'edat gestacional.

Mètodes: Estudi observacional transversal prospectiu. Per avaluar les característiques de la textura, 957 imatges 2D del pulmó fetal dret i esquerra corresponents al pla quatre càmeres cardíaques foren delineades. Les imatges es corresponien a fetus que es trobaven entre les 20 i les 41 setmanes de gestació. La quantificació de la textura pulmonar es va realitzar amb el software AQUA (Automatic Quantitative Ultrasound Analysis), que s'utilitza per extreure les característiques de les imatges. Posteriorment, un procediment d'aprenentatge que consistia en una transformació de les característiques i un model de regressió es va utilitzar per avaluar l'associació entre les característiques de la textura i l'edat gestacional.

Resultats: L'associació entre les setmanes de gestació i la textura del pulmó fetal quantificada mitjançant AQUA van presentar una correlació de Pearson de 0,97. L'associació no es va veure influenciada pels paràmetres de la delineació, com la ubicació de la regió d'interès, la mida d'aquesta o si el pulmó seleccionat era l'esquerra o dret. Tampoc es va veure influenciada pels paràmetres de l'adquisició de les imatges d'ultrasò, ni per l'equip d'ultrasò o el transductor utilitzat.

Conclusions: L'anàlisi de textura del pulmó fetal mitjançant AQUA demostra una forta correlació amb l'edat gestacional. Això dóna suport a explorar l'ús d'aquesta tecnologia per a la predicció no invasiva de la maduresa pulmonar fetal.

9.3.2. Estudi 2. Realització d'un anàlisi automàtic quantitatiu d'ultrasò del pulmó fetal per predir la maduresa pulmonar fetal.

Corresponent a l'article: M. Palacio, T. Cobo, M. Martinez-Terron, G.A. Ratta, E. Bonet-Carne, I. Amat-Roldan, E. Gratacos. Performance of an automatic quantitative ultrasound analysis of the fetal lung to predict fetal lung maturity. Am J Obstet Gynecol. 2012 Dec;207(6): 504.e1-5.

Estat: publicat

Factor d'impacte de la revista: 3.877

Quartil: 1r, àrea Obstetricia i Ginecologia.

Objectiu: L'objectiu de l'estudi va ser avaluar el funcionament de l'anàlisi automàtic quantitatiu de la textura de l'ultrasò (AQUA) per predir la maduresa pulmonar fetal determinada pel líquid amniòtic.

Mètodes: S'analitzaren embarassos únics (24.0 a 41.0 setmanes) als que s'havia realitzat una amniocentesis per determinar la maduresa pulmonar fetal (utilitzant el test TDx-FLM II). Manualment es delineà un rectangle sobre el pulmó fetal en el pla de la imatge d'ultrasò corresponent a les 4-càmeres cardíaques del tòrax fetal. AQUA s'utilitzà per a transformar la informació de la delineació en un conjunt de característiques. S'utilitzaren algoritmes genètics per extreure les característiques més rellevants. Després es va crear i validar un model que podria distingir els pulmons fetals madurs dels immadurs utilitzant el test TDx-FLM II de referència.

Resultats: L'edat gestacional de la mostra va ser (mitjana [desviació estàndard]) de 32.2 [4.5] setmanes. D'acord amb els resultats del TDx-FLM II 41 mostres es corresponien a pulmons madurs i 62 a immadurs. El biomarcador d'imatge basat en AQUA va presentar una sensibilitat del 95.1%, una especificitat del 85.7% i una precisió del 90.3% per predir un pulmó madur o immadur.

Conclusions: Les textures de les imatges d'ultrasò dels pulmons fetals extretes utilitzant AQUA proporcionen unes característiques robustes per predir els resultats del test TDx-FLM II.

9.3.3. Estudi 3. Anàlisis quantitatiu de textura d'imatges d'ultrasò per predir morbiditat respiratòria neonatal.

Corresponent a l'article: E. Bonet-Carne, M. Palacio, T. Cobo, A. Perez-Moreno, JP Piraquive, M. Lopez, E. Gratacos. Quantitative ultrasound texture analysis of fetal lungs to predict neonatal respiratory morbidity. Ultrasound in Obstetrics & Gynecology.

Estat: acceptat, 5 de Juny 2014. Ref: UOG-2014-0288.R1

Factor d'impacte de la revista: 3.557

Quartil: 1r, àrea Obstetricia i Ginecologia.

Objectiu: Desenvolupar i avaluar el funcionament d'un nou mètode per predir la morbiditat respiratòria neonatal. El mètode està basat en l'anàlisi quantitatiu dels pulmons del fetus mitjançant una ecografia.

Mètodes: Es van utilitzar un gran nombre d'imatges no clíniques i d'imatges pulmonars fetal d'ultrasò per desenvolupar un mètode computacional basat en l'anàlisi de textura i en algorismes d'aprenentatge automàtic, entrenat per a predir el risc de morbiditat neonatal en imatges d'ultrasò de pulmons fetals. El mètode, anomenat anàlisi quantitatiu d'ultrasò de la maduresa pulmonar fetal (quantusFLM), després va ser validat a cegues en 144 fetus que naixeren entre les 28.0 i les 39.0 setmanes d'edat gestacional. Les imatges dels pulmons, guardades en format DICOM, es van adquirir dins de les 48 hores prèvies al naixement i s'utilitzaren per determinar la capacitat del mètode per predir la morbiditat respiratòria neonatal, definida com a síndrome de distrés respiratori o taquipnea transitòria del nounat.

Resultats: L'edat gestacional mitjana al part va ser de 36.0 setmanes (3.3 DE). Hi va haver 29/144 (20.1%) casos de morbiditat respiratòria amb una sensibilitat, especificitat, valor predictiu positiu i valor predictiu negatiu de 86.2%, 86.9%, 62.5% i 96.2% respectivament.

Conclusions: El mètode d'anàlisi quantitatiu d'ultrasò de la maduresa pulmonar fetal (quantusFLM) va predir la morbiditat respiratòria neonatal amb una precisió comparable a les proves actuals que utilitzen líquid amniòtic.

9.3.4. Estudi 4. Ànalisi de textura automàtic d'imatges de ressonància magnètica en fetus petits-per-edat-gestacional i el seu ús per discriminar neurocomportament anormal neonatal.

Corresponent a l'article: M. Sanz-Cortes, GA. Ratta, F. Figueras, E. Bonet-Carne, N. Padilla, A. Arranz, N. Bargallo, E. Gratacos. Automatic Quantitative MRI Texture Analysis in Small-for-Gestational-Age Fetuses Discriminates Abnormal Neonatal Neurobehavior. PLoS ONE 2013 8(7): e69595.

Estat: publicat

Factor d'impacte de la revista: 3.730

Quartil: *:* 1r, àrea de ciències multidisciplinars.

Antecedents: Hem provat la hipòtesi de si l'anàlisi de textura d'imatges de ressonància magnètica pot identificar patrons associats a un neurocomportament anormal en els nounats petits-per-edat-gestacional (PEG).

Mètodes: Es van adquirir imatges d'ultrasò i de ressonància magnètica en 91 fetus PEG a les 37 setmanes d'edat gestacional. El lòbul frontal, els ganglis basals, el mesencèfal i el cerebel es van delinear en les imatges de ressonància magnètica fetal. Els nadons PEG es van sotmetre a la prova NBAS i van ser classificats com anormals si una o més àrees estaven per sota del percentil 5 i, com a normals si totes en totes les àrees van obtenir un resultat superior al percentil 5. Les característiques de textures associades amb el neurodesenvolupament es van seleccionar i es van aplicar tècniques d'aprenentatge automàtic per modelar un algorisme predictiu.

Resultats: Dels 91 nounats PEG, 49 es van classificar com a normals i 52 com a anormals. LA precisió per predir un neurocomportament anormals basant en anàlisi de textures va ser del 95.12% pel lòbul frontal, 95.56% pels ganglis basals, 93.18% pel mesencèfal i el 83.33% pel cerebel.

Conclusions: Els patrons en la textura de la ressonància magnètica cerebral es van associar amb el desenvolupament neurològic neonatal. L'anàlisi de textures en la ressonància magnètica cerebral podria ser una eina útil per predir el desenvolupament neurològic anormal en els PEG.

9.4. Discussió

Aquesta Tesi consisteix en diferents estudis per provar que l'ús de l'anàlisi quantitatiu de la textura d'imatges mèdiques pot predir el risc de resultats clínics anormals. La Tesi es centra principalment en l'ús d'imatges d'ultrasò del tòrax fetal per predir la morbiditat respiratòria neonatal. Aquest treball proporciona evidència que dóna suport al concepte prèviament estès de que les tècniques quantitatives d'imatge extreuen informació del teixit estudiat que pot ser d'ajuda en el diagnòstic clínic. Concretament, els resultats obtinguts amb el mètode desenvolupat per predir la morbiditat respiratòria neonatal foren similars als reportats per les proves comercials que utilitzen líquid amniòtic per l'anàlisi.

Els estudis de maduresa pulmonar (Estudi 1, 2 i 3) estan enfocats a utilitzar imatges del tòrax fetal per a predir la morbiditat respiratòria neonatal, cada treball intenta superar les limitacions que presentava l'estudi anterior. En el primer estudi, es correlacionaven les característiques del pulmó amb l'edat gestacional, demostrant que es pot extreure informació de forma no invasiva que es correlaciona amb un procés fisiològic normal, la maduració pulmonar, l'edat gestacional es va utilitzar com a mesura subrogada de la maduresa pulmonar fetal. En el segon estudi, el resultat utilitzat també va ser una mesura subrogada però, en aquest cas, es va utilitzar el resultat del test TDx-FLM II que utilitzava una mostra de líquid amniòtic. D'aquesta manera es va demostrar que les textures contenien informació relacionada amb la maduració pulmonar. Tot i això, el test TDx-FLM II també presenta una sensibilitat i especificitat limitada (89% i 83% respectivament). En l'estudi 3, per fer front a les limitacions anteriors, es va desenvolupar un nou mètode d'anàlisi d'imatges (quantusFLM®) i s'utilitzà com a mesura de la maduresa pulmonar fetal el resultat clínic real, l'aparició o no de morbiditat respiratòria neonatal. A part de descriure els principis bàsics d'aquest nou mètode, en el tercer estudi s'avalua la metodologia de forma cega, utilitzant mostres noves. El resultats de la validació al predir cegament el risc de morbiditat respiratòria neonatal utilitzant quantusFLM® són comparables als reportats per les proves que s'utilitzen actualment amb mostres de líquid amniòtic. Tot i això l'avaluació s'ha

realitzat en un únic centre, per això s'està realitzant un estudi multicèntric per avaluar el funcionament de quantusFLM® en més centres i utilitzant més mostra.

En el quart estudi es va avaluar la capacitat de l'anàlisi d'imatges de ressonància magnètica per detectar anomalies en diferents àrees del cervell del fetus que podrien estar associades amb un neurocomportament neonatal anormal. Aquest estudi es va utilitzar per provar la transversalitat de les tècniques d'anàlisi quantitatiu d'imatge utilitzant diferents tipus d'imatges i de resultats clínics. Aquest estudi proporciona evidència de que la ressonància magnètica del cervell dels fetus conté informació textural que està associada amb el neurocomportament neonatal. Addicionalment, l'estudi posa les bases per a realitzar una major investigació en el tema dels biomarcadors d'imatge de cervell fetal basats en anàlisi quantitatiu.

La limitació principal, present ens els estudis 1, 2 i 4, és que no hi ha una validació cega dels models, sobretot degut al nombre reduït de mostres de les que es disposava. Com a limitació general de l'anàlisi quantitatiu de textures és que es requereixen grans bases de dades per a crear biomarcadors d'imatges. En aquells casos en que la prevalença de la patologia sigui petita, serà un repte desenvolupar un biomarcador quantitatiu d'imatge basat en textures, si més no, utilitzant la metodologia aquí presentada.

Com a punt fort de l'estudi 4 cal destacar que determinar el risc de morbiditat respiratòria neonatal sense la necessitat d'una tècnica invasiva podria tenir un impacte important en el maneig clínic d'aquests casos. Segons el coneixement dels autors, el quart estudi presentat en aquesta Tesi és el primer estudi que avalua cegament un mètode basat en anàlisi quantitatiu d'imatges per a predir la morbiditat respiratòria neonatal.

9.5. Conclusions

1. L'anàlisi quantitatiu d'imatges pot extreure informació de les imatges d'ultrasò del pulmó fetal que està relacionada amb els canvis microestructurals i, en conseqüència, amb el procés fisiopatotològic de la maduració pulmonar fetal durant la gestació.

2. L'anàlisi de textures de les imatges d'ultrasò del pulmó fetal es correlen amb la maduració pulmonar fetal, quan aquesta es determina mitjançant les tècniques de laboratori estàndard basades en mostres del líquid amniòtic.

3. L'anàlisi quantitatiu d'imatge d'imatges d'ultrasò de pulmons fetals pot avaluar la morbiditat respiratòria neonatal de forma no invasiva amb una precisió similar als tests utilitzats actualment basats en mostres de líquid amniòtic.

4. L'anàlisi quantitatiu de la textura de la imatge es podria utilitzar en imatges de ressonància magnètics de cervells fetals per identificar patrons associats amb el neurocomportament anormal en nadons petits per edat gestacional.