CONSISTENT ASSOCIATION BETWEEN IMAGE FEATURES OF FETAL LUNGS FROM DIFFERENT ULTRASOUND EQUIPMENTS AND FETAL LUNG MATURITY FROM AMNIOCENTESIS

E. Bonet-Carne[1], T. Cobo[2], J. Luque[1], M. Martínez-Terrón[2], A. Perez-Moreno[1], M. Palacio[2,3], E. Gratacos[2,3], I. Amat-Roldan[1]

[1]TransMural Biotech, S.L., Barcelona, Spain
[2]Department of Maternal-Fetal Medicine, Institut Clinic of Gynecology, Obstetrics and Neonatology, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, Barcelona, Spain
[3]Centro de investigación Biomédica en Red de Enfermedades Raras, CIBER-ER, Instituto Salud Carlos III, Barcelona, Spain

ABSTRACT

Respiratory disorders in premature are outlined as one of the major causes of newborn morbidity. Therefore, estimation in utero of the risk that a particular newborn will present respiratory distress syndrome is critical in perinatal medicine. Fetal lung maturity assessment by non-invasive methods is an unsolved problem since 80's. There has been a significant effort reporting that ultrasound images carry information associated with changes of fetal lung microstructure but it required the lung-to-liver measurements to be calibrated, which limits clinical feasibility. In this work, we explored the association between image features only from a fetal lung computed by a previously reported probabilistic model and fetal lung maturity. Two studies validated image features obtained from different US equipments: 957 scans were correlated to gestational age ($r \geq 0.97$); and 83 scans were modeled to estimate a normalized lung maturity index in 7 blind samples. Bootstrapping provided median $r \geq 0.82$.

Index Terms— Image texture analysis, Statistical Learning, fetal lung, ultrasound, lung maturity

1. INTRODUCTION

Neonatal morbidity derived from respiratory disorders in late preterm infants has been underlined by a number of large studies. Therefore, estimation in utero (prior to birth) of the risk that a particular newborn will present respiratory distress syndrome is one of the most important aspects related to patient management in perinatal medicine. Currently, this risk is estimated through analysis of amniotic fluid obtained by amniocentesis. Amniocentesis is a highly invasive procedure that requires insertion of a needle to withdraw the fluid from the amniotic sac. According to statistical studies, about 1-2% of women experience cramping, spotting/bleeding, or leakage of amniotic fluid after an amniocentesis, and the total related loss rate after this invasive procedure is estimated to be around 0.5-1% [1]. Therefore, a non-invasive quantification of lung maturation will avoid risks associated to amniocentesis. In this line, the idea of predicting lung maturity by ultrasound (US) has been explored by various studies as it has the potential to provide a non-invasive mean to assess the risk of a fetus of suffering respiratory distress after delivery. A wide range of different US approaches to quantify fetal lung maturity (FLM) have been investigated: placenta grading, physical measurements of fetal size, Doppler flow assessment in trachea, lung tissue motion and relative image features from lung in combination to placenta or liver as tissue reference [2,3]. Among all, the most successful approach has relied on using relative image features based on the texture properties of lung and liver. This holds on three principles: (1) a known biological mechanism, fetal lungs experience important architectural and cytological changes throughout gestation and these changes do not occur in the liver for the last trimester of gestation, (2) a fundamental physical principle, the existing relationship between the complex US backscattering signal caused by a particular type of tissue microstructure that has been repeated in many studies [4] and (3) an integrative technological solution, that quantifies US signal by computing relative image features between lung and liver as using the liver as tissue reference helps to correct for the clinical variability of the acquired image. In these studies, a number of different measurements were simultaneously computed in lung and liver, like relative echogenicity, relative attenuation of frequencies, relative shift in the RF frequency during the process of maturation, and relative histogram-based and relative fractal-based image features [3]. These studies reinforce the idea that texture analysis is a correct method to extract image features related to characteristics of tissue microstructure. In spite of

the literature related to assessment of fetal lung maturity by US, quantification of fetal lung structure using a non-invasive methodology is an unmet challenge of perinatal medicine.

In this work we propose the utilization of a new set of image features based on a previously reported probabilistic model [5-7] extracted only within a single ROI at the one of lungs to evaluate its correlation with gestational age and fetal lung maturity. From the clinical prospective, the ability of assessing lung maturity by imaging only the lung offers an important advantage as standard US planes can be used. Particularly, we propose the 4-chamber view of the heart (Fig.1A) that is routinely performed in pregnant women after 20 weeks of gestation.

2. MATERIALS AND METHODS

2.1. Patients

Study population was singleton pregnancies attending the Maternal-Fetal Medicine Department at Hospital Clinic in Barcelona for routine pregnancy US scans from June to December 2010. 957 pregnant women accepted to participate in this study. Women were recruited until 30 per gestational age (GA) from 20 to 40 weeks were reached. GA of cohort at scan acquisition ranged from 16 to 42 weeks and was estimated by standard procedure, accounting for last menstrual period and biometries. This study protocol was approved by the local Ethics Committee. Pregnant women were informed and provided a written consent.

2.2. Image acquisition and Regions of Interest

A semi-lateral transverse 4 chamber view of the heart was acquired routinely by different physicians. The scans were performed using Siemens Sonoline Antares (Siemens Medical Systems, Malvern, PA, USA), Voluson 730 Pro, Voluson 830 Pro (General Electric Medical Systems, Milwaukee, Wisc., USA) and ALOKA Prosound Alpha-7 (Hitachi Aloka Medical, Ltd) ultrasound equipments, all of them equipped with curved linear transducer with a frequency range from 2 to 9MHz. Equipment settings, gain, frequency and time-gain compensation were not fixed and could be adjusted at the discretion of the clinician performing the ultrasound in order to obtain the optimal image quality according to clinical criteria. Ultrasound scans were digitally collected and stored for off-line analysis in a custom made program for delineation. Each scan was evaluated offline for the presence of abnormalities and all maternal and neonatal outcomes were recorded. Manual (free-hand) delineation was done in all 957 images and square delineation (60x60pixels) was done in each lung by a physician in a total of 90 images ensuring that delineation only included lung tissue as illustrated in Figure 1.B.

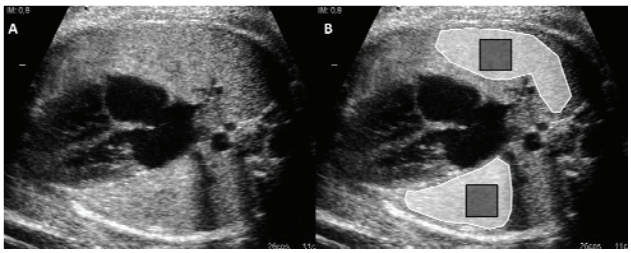


Figure 1. A. semi-lateral transverse US sacan of 4 chamber view of heart. B. Manual delineation of proximal and distal lung in white and square delineation in black.

2.3. Feature extraction

Image features were computed by AQUA [5], a software that estimates texture features based on probabilistic model as described in ref. [6-7]. Image features were calculated from each US scan within a single ROI at the lung. Fundamental steps of AQUA algorithm are shown in Figure 2.

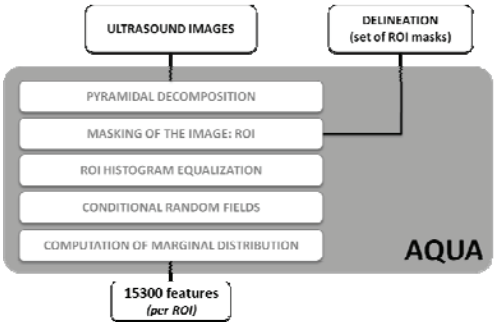


Figure 2. Main steps of AQUA algorithm

2.4. Association of AQUA features with fetal lung maturity

Association of AQUA features with fetal lung maturity was computed in two experiments: (1) The manually delineated lungs of 957 fetuses in which data was split in two bags: one for feature transformation and another for regression model and test; and (2) The square delineation of 90 fetuses in which data was split in three bags: one for feature transformation, another for regression model and a third one for testing. Bootstrapped calculations were also computed to evaluate the robustness of the method.

2.4.1.- Feature Validation 1

Data and Bags 1. 957 images were analyzed. Data was split into two bags. Bag 1 contained those fetuses with gestational ages at acquisition below 23 weeks (Bag 1a) and above or equal 39 weeks of gestation (Bag 1b) which corresponds to 10 and 90 centile respectively. Bag 2 contained those images between 23 and 39 weeks of gestation as shown in Figure 3A. The learning sequence was

divided into two parts, (1) feature transformation and (2) regression model used to test the hypothesis. Bag 1 (249 images) was used to find an optimal feature transformation and Bag 2 (708 images) to compute a regression model.

Feature Transformation 1. The 15,300 features were first selected by thesholding them at 0.5 of Mahalanobis distance between extreme groups (Bag 1a and 1b). Secondly, principal component analysis reduces data dimensionality.

Regression model 1. A random forest is computed within Bag 2 to estimate weeks of gestation. As input were used only 30 principal components to ensure no over-fitting.

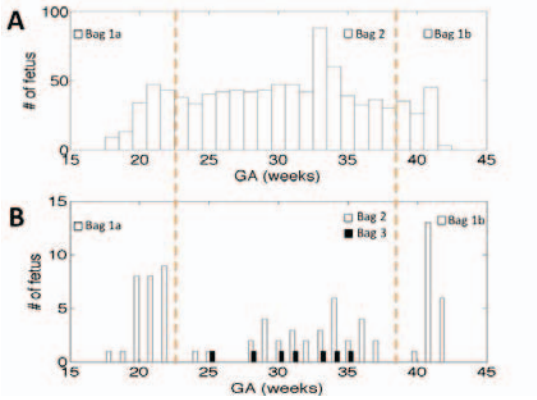


Figure 3. Histogram of total sample for (A) Feature Validation 1 (N=957) and (B) Feature Validation 2 (N=90). (A) Two bags: Bag1 with extreme fetuses (GA<23 and GA>39), Bag 2, elsewhere. (B) Three bags: extreme fetuses, immature (GA<23) and mature (GA>39); bag 2 and 3 with the rest of fetuses. Bag 2, includes only those fetuses in which FLM was measured by amniocentesis (gold standard). Bag 2, includes only those fetuses in which FLM was measured by amniocentesis (gold standard).

2.4.2.- Feature Validation 2

Data and Bags 2. A sub-set of 90 images was analyzed (70% from Siemens and 30% from GE, US machines). Data was split into three bags as shown in Figure 3B. Bag 1 with fetuses below 23 (immature) or above 39 (mature) weeks of GA. The other bags within 22 and 39 weeks of GA: Bag2 with measured FLM by amniocentesis one day earlier or after US scan was acquired and Bag 3 with no FLM measurement by amniocentesis. The learning sequence was divided into three parts, (1) feature transformation, (2) regression model and (3) blind test. The learning sequence was bootstrapped to compute an interval of confidence of this validation.

Feature Transformation 2. Selection of features with SNR>10 within immature fetuses and mature fetuses, and Mahalanobis distance was larger than 0.5. Dimension reduction was done by keeping only the first component analysis. All computations were performed using only 15 immature and 15 mature samples of Bag 1 chosen randomly.

Regression model 2. Data of Bag 3 and the rest of Bag 1 not used for feature transformation is used to compute a linear model based on the 1st principal component and the GA at date of US scan. The estimated GA is then divided by 40 for normalization purposes, as 40 weeks is the average gestation period.

Blind Test 2. Bag 2 is bootstrapped 100 times to avoid effect of outliers and a correlation distribution is obtained between the *regression model 2* and FLM of Bag 2, which was not used in the whole process.

2.4.3.- Measurements of fitting

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 equipments.

3. RESULTS

3.1.- Results: Feature Validation 1

No significant differences between clinical characteristics were found in disaggregated sample bags (Bag1a, Bag1b and Bag2). The gestational age prediction was evaluated in a large cohort of 708 fetuses ranging 22 to 39 weeks of gestational age (Bag 2). The analysis showed a strong Pearson correlation above of 0.97 between AQUA lung features and gestational age (Table 1).

	RIGHT lung free-hand	LEFT lung free-hand
Pearson Correlation	0.97	0.98
R^2	0.86	0.86
RMSE	1.58	1.16

Table 1. Fitting results for feature validation 1.

Additional results (Figure 4) illustrate the robustness of AQUA features with regard to different US equipments and different frequency settings. Frequency information could not be extracted from US scans of General Electric equipments and were not included in the frequency analysis. The sample was then reduced to 491 acquisitions only for this analysis and no significant differences were found.

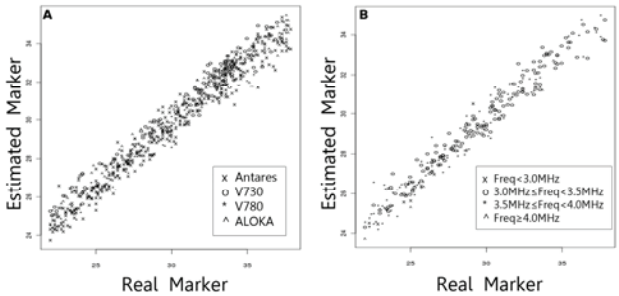


Figure 4. Additional experiment results to illustrate independency of AQUA features to different acquisition machines and settings. A. Scatter plot showing predictions for different US machines scans (N=957). B. Scatter plot illustrating different acquisition frequencies (N=491).

3.2.- Results: Feature Validation 2

Delineation of a Right Lung square resulted in a median of Pearson correlation with FLM measurement by amniocentesis (gold standard) of 0.94, with an confidence interval at 95% (CI95) of [0.87,0.97]. Left Lung resulted in a Pearson's correlation of median 0.82 and CI95 of [0.75, 0.9]. Comparatively, correlation to a random process resulted in a median Pearson correlation of 0.06 and CI95 of [-0.87, 0.87]. These results are shown in Fig. 5 and Table 2. This methodology was repeated using manual delineation and it underperforms (data not shown).

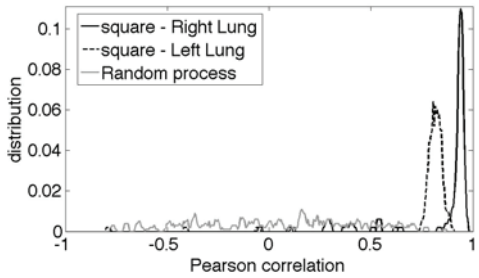


Figure 5. Distribution of Pearson correlation from 100 bootstraps of features computed from Right and Left square ROI at the lungs and a random process with FLM results obtained by amniocentesis (gold standard).

Results from fetuses that belonged to the blind set can be read in Table 2, which contains the individual measurements (GA at US scan and FLM value provided by amniocentesis) and estimated lung maturity (as the median of the bootstrap) for each individual fetus and lung side. Correlation of GA at US and FLM-amniocentesis is 0.64.

	GA at US (weeks)	FLM amniocentesis (a.u.)	RIGHT normalized lung maturity	LEFT normalized lung maturity
fetus 1	25.00	4.66	0.68	0.69
fetus 2	32.71	10.80	0.76	0.83
fetus 3	33.57	10.80	0.76	0.83
fetus 4	29.57	14.56	0.80	0.90
fetus 5	34.71	28.00	0.92	0.91
fetus 6	30.71	14.00	0.85	0.81
fetus 7	28.00	14.00	0.79	0.85

Table 2. GA at US scan, FLM by amniocentesis (gold standard), median of the 100 bootstraps as estimation of lung maturation from Right and Left lungs only from the image features on a blind sample of 7 fetuses (Bag 2, Fig.3B).

4. CONCLUSIONS

Quantification of lung structure using a non-invasive methodology remains a challenge to predict lung maturity in several obstetric situations. Main contributions of this study to overcome this situation are: (1) utilization of a new set of image features based on a previously reported probabilistic model to correlate with fetal lung maturity and (2) demonstration that this set of features does not require tissue reference and therefore, a standard US plane (4 chamber view of heart) routinely performed in pregnant women after 20 weeks of gestation can be effectively used. The proposed set of features shows good performance on three critical aspects: (1) it strongly correlates with GA as shown in feature validation; (2) features are consistent and reproducible with different types of US equipments, which would simplify its inclusion in clinical practice; and (3) it is informative of the underlying biological process related to the appearance of surfactant (FLM by amniocentesis is an indirect measurement of the tissular and cytological changes occurring in fetal lungs) as shown in feature validation 2 that included a blind test of 7 samples.

Although results of this study are promising, evaluation of the predictive value within a larger blind set and correlation to effects of therapies (i.e. corticoids) on fetuses at risk needs to be thoroughly studied in order to develop and fully understand the implications of the first US-imaging biomarker to assess fetal lung maturity in clinics. These future studies will provide evidence to completely characterize the meaning of the computed image features in relation to this important issue of perinatal medicine.

5. REFERENCES

[1] B.F. Crandall, J. Howard, T.B. Lebherz, L. Rubinstein, W.F. Sample, and D. Sarti, "Follow-up of 2000 second-trimester amniocenteses," *Obstet Gynecol.*, 56(5), pp. 625–628, 1980.

[2] P.A. Grannum, R.L. Berkowitz, and J.C. Hobbins, "The ultrasonic changes in the maturing placenta and their relation to fetal pulmonic maturity," *Am J Obstet Gynecol.*, 133(8), pp. 915-922, 1979.

[3] B.K.N. Prakash, A.G. Ramakrishnan, S. Suresh, and T.W.P. Chow, "Fetal lung maturity analysis using ultrasound image features," *IEEE Trans Inf Technol Biomed.*, 6(1), pp. 38-45, 2002.

[4] A.S. Tunis, G.J. Czarnota, A. Giles, M.D. Sherar, J.W. Hunt, and M.C. Holios, "Monitoring structural changes in cells with high-frequency ultrasound signals statistics," *Ultrasound in Med and Biol.*, 31(8), pp. 1041-1049. 2005.

[5] Software available upon registration or email at www.transmuralbiotech.com.

[6] V. Tenorio, E. Bonet-Carne, F. Botet, F. Marques, I. Amat-Roldan, and E. Gratacos, "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(10), pp. 1365-77, 2011.

[7] E. Bonet-Carne, V. Tenorio, F. Figueras, E. Gratacos, and I. Amat-Roldan, "Evaluation of semiautomated quantification of cranial ultrasound images in newborns as a predictor of Neonatal Behavioral Assessment Scale," *IEEE International Symposium on Biomedical Imaging,* Chicago, IL, 2011 pp. 41-44.

6. ACKNOWLEDGEMENTS

Elisenda Bonet-Carne and Transmural Biotech acknowledge "Comissionat per a Universitat i Recerca del Departament d'Innovació, Universitats i Empreses de la Generalitat de Catalunya" by its financial support. Teresa Cobo was supported by a Rio Hortega grant from Instituto de Salud Carlos III in Spain (CM09/00213). Mónica Martínez-Terrón was supported for a grant from Hospital Clínic de Barcelona (2011-2012)