Creating dynamic and customized fetal growth curves using cloud computing

Mario A. Bochicchio, Antonella Longo, Lucia Vaira, Antonio Malvasi, and Andrea Tinelli

Abstract—**The modern cloud-based solutions are gaining the upper hand in the most common industrial areas. In the healthcare sector, cloud computing is starting to set foot on, and the major applications are about the possibility to collect and share medical data. One of the hot topics argued in the obstetrics and gynecologist community, is related to the fetal growth and, in particular, to the necessity to have updated standards, since the current references lack information like ethnicity and maternal biometric parameters (essential to evaluate the correct growing parameters). We believe that cloud computing could help the development of such a kind of dynamic and customized fetal growth curves, which in turn could improve the possible detection of anomalies and pathological states during the whole pregnancy period. The paper presents a proposal for resolving the problem of the fetal biometric data obsolescence and the inability to use them in a custom or adapted fashion giving the opportunity to manage clinical data which are dispensed "as a service" on a global scale. The objective of the study is to create and validate a database collecting several fetal growth curves (obtained by means of the available results beginning from 1963 until 2013).**

I. Introduction

Recent developments in healthcare technology enable the collection, storage, management, and sharing of massive amounts of medical data. The urgent need for pervasive and ubiquitous real-time access to patients' data from anywhere and from any digital device has become essential for proper diagnosis and treatment procedure so that it leads to achieve high quality of medical services.

The obstetrics and gynecologist sector stands as emblematic: technologies are underutilized and are not treated as a means to improve patient care or, if utilized, they embrace the image acquisition devices (3D or 4D ultrasound scan) and not consider the data management and analysis that could improve the fetal growth assessment.

Throughout the pregnancy, in order to correctly interpret the data collected during the different growth stage, the doctor assesses a set of parameters, such as Biparietal Diameter (BPD), Head Circumference (HC), Abdominal Circumference (AC), Femur Length (FL), and compares them with particular reference charts, representing the fetal growth curves, which show the average trend fetal growth over time allowing to detect potential form of abnormal growth. The problem lies in the fact that the commonly used growth curves are at least five decades old and may not be suitable for the current population. These curves are those developed by Lubchenco [1] Usher and McLean [2] and Babson and Brenda [3] all suffering from the relatively indigent population included, so they not address the issue of variation in intrauterine growth due to ethnicity, which is an important factor to take into account.

Beyond ethnicity, many other factors affect fetal growth including physiological and pathological variables, such as maternal height and weight, drug or tobacco exposure, genetic syndromes, congenital anomalies and placental failure. Also the fetus gender plays a role in the birth size [4]. Some authors addressed this problem trying to adjust for most of the influential factors in their own environment. In the early 1992, Gardosi [5] introduced the idea of individualized fetal growth charts according to specific maternal and fetal characteristics. Later, other authors provided several fetal growth charts considering different subgroups of population (UK, Canadian, Kuwait, New Zealand, Saudi Arabia, Thailand, USA, Tunisian, Pakistani, Turkey, India and so on) but all these charts cannot be applied universally to different population and suffer from many limitations: they include small, and therefore insufficient, samples at low gestational age; they have hospital-based samples and they use inadequate statistical modeling techniques and adopt inaccurate methods to calculate gestational age.

Since there is considerable methodological heterogeneity in studies of fetal biometry, a standardization of methodologies is deeply necessary in order to make correct diagnosis avoiding unnecessary obstetric interventions at the time of delivery.

In this scenario, we feel that porting this kind of analysis to the Cloud could make possible approaches on a global scale and the development of quite significant statistical samples and furthermore could provide a greater sustainability since it's possible to address to a large target

Manuscript received October 9, 2001. (Write the date on which you submitted your paper for review.) This work was supported in part by the U.S. Department of Commerce under Grant BS123456 (sponsor and financial support acknowledgment goes here). Paper titles should be written in uppercase and lowercase letters, not all uppercase. Avoid writing long formulas with subscripts in the title; short formulas that identify the elements are fine (e.g., "Nd–Fe–B"). Do not write "(Invited)" in the title. Full names of authors are preferred in the author field, but are not required. Put a space between authors' initials.

F. A. Author is with the National Institute of Standards and Technology, Boulder, CO 80305 USA (corresponding author to provide phone: 303-555-5555; fax: 303-555-5555; e-mail: author@ boulder.nist.gov).

S. B. Author, Jr., was with Rice University, Houston, TX 77005 USA. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar. colostate.edu).

T. C. Author is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA, on leave from the National Research Institute for Metals, Tsukuba, Japan (e-mail: author@nrim.go.jp).

978-1-4799-3163-7/13/$31.00 ©2013 IEEE

of people by means of advertisement and the distribution of services directly to the final user.

In this paper we propose a cloud approach, based on the SaaS (Software as a Service) paradigm and on data anonymization techniques in charge of collect, process and share fetal growth data on a global scale in order to create dynamic and customized fetal growth curves that will take into account the above-mentioned factors influencing fetal growth. The user can select and consult the preferred curves using a contemporary, large, racially diverse sample and compare with his own values for example.

Clinicians can also upload the data collected during the visits in order to contribute to the development of a new and updated tool for growth assessment.

The paper is organized as follows: Section 2 introduces related works. Section 3 presents our approach. Section 4 discusses the experimental results and the prototype system developed. Finally, we draw conclusions and discuss our future works in Section 5.

II. Related Works

One of the key benefits deriving from the cloud computing usage could be the ability to exchange data between disparate systems, and this is very important for healthcare IT. Despite widespread adoption of cloud-based solutions by most industries, cloud computing has been slowly embraced by healthcare and the biggest impeding factors are concerns about security and performance of cloud services causing the poor usage by the citizen (this is the official reason for the closure of Google Health and Revolution Health for example). In our case, instead, there is a strong desire on the part of pregnant women to know and assess the health status of their own child.

Despite all opposition, many institutions have moved to the cloud to lower their storage costs and facilitate the exchange of images and medical record, such as Microsoft Health Vault [6] or the Merge Healthcare's Project Honeycomb [7], the Accenture Medical Imaging Solution [8] or the Individual Clinical Portability of the ULSS number 8 of Asolo [9].

Considering the obstetrics and gynecologist field, for more than half a century, several clinicians and investigators have proposed different reference data for assessing birth weight for gestational age [10], [11], [12] but none of these works dynamically develops the curves using the clinical routine practice data achieved throughout the ambulatory visits and no one proposed a cloud-based solution to enlarge the number of analyzed patients and to have a more significant statistical sample for perform the analysis.

One of the major developments in this field has been the introduction and usage of software tools that allow you to get customized charts. The most known is surely GROW by J. Gardosi [13] which provides customized growth charts based on a regression model for birth weight, fitted to a very large group of neonates. It addresses the issue related to the data updating to use as the reference standard for the development of growth curves, but data are not subject to regular updating, they descend from "*static*" archives, they don't differ in ethnic origin and other important factors and so they are not suitable to assess the biometric parameters influencing the fetal growth. Furthermore it obligates to purchase and maintain software.

III. Our Approach

In this paper we propose an alternative approach, based on cloud, on the SaaS paradigm and on data anonymization techniques to collect, process and share growth data on a global scale in order to resolve the problem of the fetal biometric data obsolescence and the inability to use them in a custom or adapted fashion.

Analyzing the SaaS paradigm from the healthcare point of view, the biggest advantage of SaaS is the simplified management of all health sensible data. In fact, SaaS healthcare services can enable scenarios where all medical assistants have access to patient data anytime and anywhere with any device equipped with Internet connection.

In particular, to ground our proposal, we focus on the well-known and widely diffused Amazon public Cloud provider, namely the Amazon Web Services (AWS), which delivers a set of services that together form a reliable, scalable, and inexpensive computing platform.

The issue of clinical data handling is well known in literature. Personal Health Records (PHRs) contain extremely sensitive personal information and thus pose strong privacy concerns. This means that special care has to be adopted to preserve the privacy. The privacy and security aspects are out of the scope of this paper, but in summary they have been addressed by separating the personal data from the clinical data using hash keys for reconciliation purposes. Data protection laws in several countries, require knowledge of where data is stored. For this reason, storage of patient data in the cloud will be very difficult to implement in countries like Spain, France or Italy. However, several cloud providers allow obligatory data storage in a specific geographic location. Thus, the problem addressed can be minimized and even countries with higher restriction laws might accept the solution [14].

A. Scenario definition

The scenario presumes that throughout the pregnancy, the mother is subjected to routine checks to ensure proper fetal growth. The patients and the family are interested in the stages definition of pregnant and family biometric parameters in order to provide all relevant information for the development of dynamic and customized fetal growth curves. The medical and paramedical staff is involved in the stages of acceptance, preparation and training for pregnancy, execution of routine exams related to the fetal and maternal health status, and in the management of any form of abnormalities, pathologies, diagnosis and prognosis up to the psychological support.

B. Requirement analysis

The main identified requirements are:

1) R1: development of a web and mobile system for ubiquitous access;
2) R2: respect the constraints of safety, reliability, privacy and anonymity;
3) R3: collection of significant amounts of data for statistical evaluation of the examined parameters;
4) R4: guarantee service levels for the medical staff usage;
5) R5: flexibility of the information schema in order to support analysis types in evolution;
6) R6: ability to segment the analyzed sample according to various dimensions of analysis, especially in relation to biometric parameters, maternal ethnicity, environmental factors and lifestyle.

The multidimensional requirement (R6) is new if compared to the requirements of the other authors who had proposed customized growth curves. It has been adopted because it allows the detection and the study of correlations that are still not known among the listed parameters and a dynamic point of view: it's possible to make dynamic analysis without following a fixed pattern.

C. Data modeling

For the modeling of the clinical data, we started from the Silverston Universal Data Model [15] which provides an outstanding starting point for delivering flexible data models, in order to map and model the main components of our data model and from the HL7 Reference Information Model [16] which is an abstract information model for healthcare provided by the HL7 Organization. Each doctor should have available a local database, on which are recorded only his own patients and related reports. Because of the lacking of a database model that is unique for all doctors of this sector, we have developed a flexible model that could allow the modeling of different kind of exams and hence it can subsequently be specialized according to the needs. On account of the limited space available, we can't show the ER diagram, but in summary, we have focused on the main person and report attributes firstly in a separate way and then, from the interviews conducted with the doctors, we understand the need to move some patient attributes to the medical report entity. For instance, attributes like weight can't be static patient's attributes, since it grows during the pregnancy. A little fragment of the whole ER diagram is shown in Fig. 1.

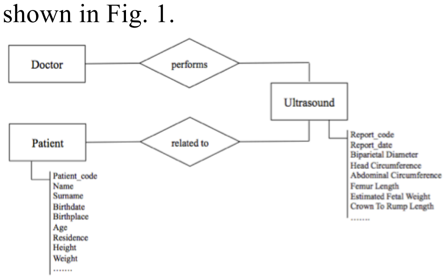

Fig. 1. A fragment of the ER report-oriented diagram

D. Architecture modeling

The whole application is split in two parts as can be seen from the Fig. 2: the first one, including the Dynamic Fetal Growth Curves Builder (DFGC Builder), the doctor's interfaces and all the services (encryption, anonymization etc.) with high privacy and security requirements is located at the University of Salento on a dedicated server. The second one, including the patients' interfaces and the public statistical data about fetal growth is located in a public cloud on the web. The DFGC Builder includes all the algorithms used to build the personalized growth curves starting from the patient data and the clinical measured features.

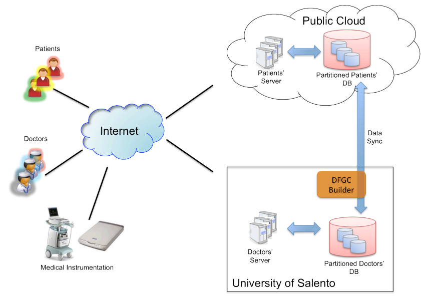

Figure 2. Architecture diagram

E. Implementation

For a multidimensional analysis of the fetal biometric data and the development of customized fetal growth curves, we used a data warehouse system, which is in charge of collecting and integrating data coming from heterogeneous sources in order to dynamically define the personalized growth curves matching the right fetus-maternal parameters for each specific case. To accomplish these goals and satisfy these needs, each user must be able to access certain resources and certain data to perform only specific tasks, avoiding to compromise the data security. We built an ad hoc virtual machine, which contains our application, the data warehouse and all the tools adopted for the development: the suite of open source Business Intelligence products (Pentaho Business Analytics with all its service components) and the MySQL DBMS.

IV. Results and Discussion

The development of personalized fetal growth curves is a problem characterized by relevant size if we consider that there are about 140 ML of newborns per year.

The dimensionality is not the only aspect to take into account. Even the diversity of ethnic groups in each country is quite relevant. For instance, in Italy (for which we know very well the entire statistical parameters) there are 21 different ethnic groups and for all groups are used the same fetal growth curves. But each ethnic group has distinct biometric parameters and so it requires distinct growth curves. Our approach has been validated by a small number of doctors, which have evaluated and verified: the functional correctness of the proposed model and approach, the simplicity and usability of the entire system and finally the correctness and the completeness of the informative model.

In order to test the system, in this first stage and in the absence of data that are distinguished according to different ethnic groups, we have produced a biometric parameters generator, which produces a significant number of samples starting from the percentile used for the different populations, all respecting the mean and the standard deviation that characterize the Gaussian distribution for a specific gestational week. The generated data allowed us to prove the correctness of the elaboration of the fetal growth curves distinguished by ethnicity and the validation of the elaboration time. The performed analysis include the standard biometric parameters evaluated on several fetuses, of different gestational age, with the possibility to explore the data from different points of views. For instance, it's possible to obtain the 50th percentile related to the Head Circumference in relation to the mother ethnic group as shown in Fig. 3.

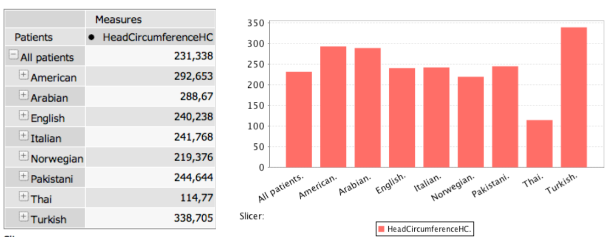

Figure 3. Head Circumference Average Value according to the mother's ethnicity

V. Conclusion and Future works

The fetal growth assessment is a relevant problem since it concerns about 140 ML of newborns per year. Due to the population and "*ethnicity reshufflement*", it's by nature a global phenomena and cloud computing is the most suitable approach to use in order to obtain personalized growth curves' development and an adequate response time (< 3 sec).

In order to prove the feasibility of the proposed approach we have developed a prototype based on a relational database, but being our problem global by nature, in the future we plan to provide a more scalable solution that will make use of NoSQL database which have usually the goal of offering horizontal scalability, on the basis of a distributed implementation and deployment. According to [17] more than fifty NoSQL systems have been already implemented, each with different characteristics (e.g., different data model and different API to access the data, as well as different consistency and durability guarantees). As pointed up in [17] this lack of standard is problematic to application developers.

References

[1] Lubchenco L.O., Hansman C., Boyd E. Intrauterine growth in length and head circumference as estimated from live births at gestational ages from 26 to 42 weeks. *Pediatrics*. 37, 403–408. 1966

[2] Usher R., McLean F. Intrauterine growth of live-born Caucasian infants at sea level: standards obtained from measurements in 7 dimensions of infants born between 25 and 44 weeks of gestation". *J Pediatr*. 74, 901–910. 1969

[3] Babson S.G., Benda G.I. Growth graphs for the clinical assessment of infants of varying gestational age. *J Pediatr*. 89, 814–820. 1976

[4] Groveman S.A. New Preterm Infant Growth Curves Influence of Gender and Race on Birth Size. Masters thesis. Philadelphia, PA: *Biotechnology and Bioscience, Drexel University*. 2008

[5] Gardosi J., Chang A., Kalyan B., Sahota D., Symonds E.M. Customised antenatal growth charts. *Lancet*. 339, 283-287. 1992

[6] Microsoft. Microsoft HealthVault. Retrieved from http://www.microsoft.com/en-us/healthvault/organize/medical-records.aspx 2012

[7] Merge Honeycomb, available at http://www.merge.com/Solutions/Cloud/Merge-Honeycomb-Archive.aspx

[8] Violino, B. Accenture, AT&T Offer Cloud-Based Medical Imaging. Retrieved from http://www.information-management.com/news/Accenture-ATT-cloud-computing-medical-imaging-solution-10021574-1.html 2011

[9] PIC ULSS 8 Asolo available at https://www.ulssasolo.ven.it

[10] Gardosi J., Mongelli M., Wilcox M., Chang A. An adjustable fetal weight standard. *Ultrasound Obstet Gynecol* 1995; 6:168-174. 1995

[11] Bonellie S., Chalmers J., Gray R., Greer I., Jarvis S., Williams C. Centile charts for birth-weight for gestational age for Scottish singleton births. *BMC Pregnancy Childbirth*. 8,5–14. 2008

[12] Fenton T.R.: A new growth chart for preterm babies: Babson and Bendas chart updated with recent data and a new format. *BMC Pediatr*. 3, 13–22. 2003

[13] GROW software http://www.gestation.net

[14] Lus A. Bastio Silva, Carlos Costa and Jos Luis Oliveira. Strengths and Weaknesses of Using Cloud Computing Data Ownership & Protection Issues, *Imaging Management*, vol. 11 (5), p 14-15. 2012

[15] Silverston L., The Data Model Resource Book, Vol. 2: A Library of Universal Data Models by Industry Types, Revised Edition, *John Wiley and Sons*, New York (April 2001)

[16] HL7 Reference Information Model, Health Level Seven, Inc., available at http://www.hl7.org/v3ballot/html/infrastructure/rim/rim.htm

[17] Stonebraker M. Stonebraker on NoSQL and enterprises. *Commun. ACM*, 54 (8): 10–11, 2011