Percentile Growth Curves for Placenta Measures: A Dynamic Shiny Application

Samuel Alves[1(✉)], Ana Cristina Braga[2], and Rosete Nogueira[3,4]

[1] University of Minho, 4710-057 Braga, Portugal
samuelcnalves@gmail.com
[2] ALGORITMI Centre, University of Minho, 4800-058 Guimarães, Portugal
acb@dps.uminho.pt
[3] School of Medicine and Life and Health Sciences Research Institute (ICVS) and ICVS/3B's - PT Government Associate Laboratory, University of Minho, Braga, Portugal
[4] Embryo & Fetal and Placental Pathology Laboratory, CGC, Unilabs, Porto, Portugal

Abstract. For decades, researchers and health professionals have been using fetal and newborns measurements to evaluate its development. In recent years, there have been new studies suggesting that the placenta's measurements and its evolutions are capable of reflecting changes in the fetus's development and even newborn and adult diseases.

Most of these analyses are done using growth curves that use linear regression methodologies such as previous studies done. To account for errors associated with this regression and use a more robust method, quantile regression is used to create the placenta's growth curves. The dataset used for this study was collected on Portuguese CGC Genetics and involves the Portuguese parturient population from different regions.

It is also an objective of this study to create a dynamic application that allows the researcher or health professional to enter placental growth values and compare them to the created growth curves to evaluate the evolution of the placenta. This application uses a CSV file with the information gathered from the placenta and is uploaded to the application which then plots the values on the created growth curves. The application also allows the user to edit the values. This application was created on Shiny and can be accessed at https://samuelalves.shinyapps.io/APP2/.

Keywords: Placenta · Growth curves · Quantile regression · Shiny

1 Introduction

The placenta is one of the largest and first organs to develop during pregnancy, it is considered by some experts as one of the most important organs due to its crucial role in the normal development of the baby. As such the placenta is considered highly regulated and coordinated to allow highly efficient exchange of different products and nutrients between the mother and the fetus.

This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the R&D Units Project Scope: UIDB/00319/2020

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022
O. Gervasi et al. (Eds.): ICCSA 2022 Workshops, LNCS 13377, pp. 543–554, 2022.
https://doi.org/10.1007/978-3-031-10536-4_36

The xenobiotic protector effect of the placenta is also highly important just as is the effect on hormone regulation and its ability to react to the different changes in the circulatory and nutritional environment to allow different proportions of nutrients to be supplied to the fetus.

There have been different studies relating the measurements of the placenta to different problems during the pregnancy, the newborn and the health later in life (Asgharnia et al. 2008; Barker et al. 1990; Hindmarsh et al. 2001).

In the most recent years, there has been an increase in interest in studying this organ although there are some gaps in the knowledge about it and its biological significance due to the lack of interest from researchers to study it compared to the fetus.

Additionally, there is a problem associated with the fact that it is difficult to get the placenta's measurements and standardise them. Moreover, there is still a lack of the placenta measurements data and the problem is that the values change according to the different races and the available data is considered old.

To improve the knowledge about the placenta, a work is previously done by Doctor Rosete Nogueira (Nogueira et al. 2019), which made possible the creation of percentile curves of the Portuguese population's placenta.

Considering this, the goal of this study was to create growth charts with placenta dimensions of the Portuguese population and create an application that allows the user (either a health professional or researcher) to enter the placental values and compare them with the previously created on this study growth charts. This will allow the clinician to assess the growth of the placenta in different stages of the pregnancy process.

This paper is structured as follows. Section 2 presents a state of the art, where some works related to the models and statistical methods used, as well as the software are briefly reported. The theoretical methodology of quantile regression is presented in Sect. 3. The exploratory results of the dataset used as well as the growth curves for the measures and the development of an application, are in Sect. 4 and Sect. 5, respectively. Finally, the main conclusions and the proposed future work are described in Sect. 6.

2 State of the Art

2.1 Models and Statistical Methods

Since the beginning of the creation of growth charts, there have been several different ways to do it. It started with fitting smoothing curves on sample quantiles of the segmented age groups. The problem associated with these methods comes from the fact that these are not robust (robustness is considered the resilience of statistical procedures to deviations from the assumptions of hypothetical models (Koenker and Bassett 1978)) to outliers and a large sample is needed to estimate the percentiles with good precision (Chen 2005).

Another problem associated with the usual growth charts created is that they are built on the assumption that the measurements are normally distributed which is not true since for example, if adult heights from a relatively homogenous population are considered close to normal, children's heights are usually non-normally distributed and this problem can be even bigger in other measurements such as weight.

To solve these problems, the majority of the studies about growth charts use regression models.

Regression is a statistical method that allows the user to study the relationships between variables. Of the many regression methods available, quantile regression is a method that is used to estimate functional relations between variables for all portions of a probability distribution (Cade and Noon 2003).

Quantile regression is a nonparametric method that due to the increase in computing power is becoming more frequently used. It is independent of any distribution or transformation to normality since it estimates distribution directly and as such, it is more robust against outliers (Kiserud et al. 2018).

2.2 Shiny Applications

Shiny is a package used in R, that allows the user to easily create dynamic web applications using R code.

With Shiny we can make the Back End and Front End of an application, interact with languages external to R, such as CSS, HTML and JavaScript and also host an app on a cloud.

A great disadvantage of shiny is that it requires timely updates, i.e., as the functions used in the app get outdated sometimes with newer package versions, it becomes necessary to update the shiny app from time to time.

Several authors have already shown its usefulness and how it may revolutionize the way models' simulations are shared. It has also been shown by authors different ways that these applications can be shared such as ShinyApps.io, Shiny Server and Shiny Server Pro which allows the sharing of these applications to a wider audience due to being hosted on a web server instead of locally (Wojciechowski et al. 2015).

3 Quantile Regression

Quantile regression was proposed by Koenker and Bassat and is a method used to estimate functional relations between variables for all portions of a probability distribution which allows the studying of the conditional distribution of Y on X at a different location, offering a global view of the interrelations of X and Y (Cade and Noon, 2003).

An important note has to be made that this work is done on a non-normal distributed sample. The distribution can also be different about the mean or a specific quantile. For example, modelling using only the mean (which happens when using linear regression) can lead to missing important aspects of the association between the outcome and its predictors (Beyerlein 2014).

As such the quantile regression method was chosen to allow to fill gaps created by other methods and allow the creation of a better solution for the researchers.

According to Buhai 2004, considering the ordinary quantile, which has a real value random variable Y, with the following distribution:

$$F(y) = \Pr(Y \leq y) \tag{1}$$

Thus for any $\tau \in (0, 1)$, the τ-*th* quantile of Y is defined as

$$Q(\tau) = inf\{y : F(y) \geq \tau\} \tag{2}$$

The median is $Q(1/2)$, the first quartile $Q(1/4)$ and the first decile $Q(1/10)$. The quantile function allows the user to create a complete characterization of Y, like the distribution of function F. The quantiles can be written as solutions to this optimization:

$$\rho_\tau(u) = u(\tau - I(i < 0)), \quad \tau \in (0, 1) \tag{3}$$

where $I(.)$ is the usual indicator function. Thus, the solution to the minimization is

$$\widehat{\alpha}(\tau) = \arg min_{\xi \in \mathbb{R}} E[\rho_\tau(Y - \xi)] \tag{4}$$

The sampled analogue of $Q(\tau)$ is based on a random sample $\{y_1, \ldots, y_n\}$ of Y. The τ-*th* quantile can be identified as any solution to:

$$\widehat{a}_\tau = \arg\min_{\xi \in \mathbb{R}} \sum_{i=1}^{n} \rho_\tau(Y - \xi) \tag{5}$$

Let $\mathbf{x}_i$, $i = \overline{1 \ldots n}$, a $K \times 1$ vector of regressors. It can be written the equivalent to the first expression as

$$F_{u_\tau}\left(\tau - \mathbf{x}_i^T \boldsymbol{\beta}_\tau | \mathbf{x_i}\right) = \Pr(y_i \leq \tau | \mathbf{x_i}) \tag{6}$$

Which is a different form of:

$$y_i = \mathbf{x}_i^T \boldsymbol{\beta}_\tau + u_{\tau_i} \tag{7}$$

considering the distribution of the error term, u_{τ_i} is left unspecified, the only constraint being the quantile restriction $Q_\tau\left(u_{\tau_i} | \mathbf{x_i}\right) = 0$.

Using an analogy, the estimation of conditional mean functions as in

$$\widehat{\beta} = \arg min_{\boldsymbol{\beta} \in \mathbb{R}^K} \sum_{i=1}^{n} \left(y_i - \mathbf{x}_i^T \boldsymbol{\beta}\right)^2 \tag{8}$$

The linear conditional quantile function,

$$Q_Y(\tau | \mathbf{X} = \mathbf{x}) = \mathbf{x}_i^T \boldsymbol{\beta}_\tau \tag{9}$$

Can be estimated by solving the equivalent of expression 8:

$$\widehat{\beta}_\tau = \arg min_{\beta \in \mathbb{R}^K} \sum_{i=1}^{n} \rho_\tau \left(y_i - \mathbf{x}_i^T \boldsymbol{\beta}\right)^2 \tag{10}$$

According to Yu *et al.* (2003) to discuss model estimation we need to take in mind the principles of the least-squares estimation.

Considering the regression model that is the simplest, the β is estimated through the quadratic loss function $r(u) = u^2$ and given a data set with observations $\{x_i, y_i\}_{i=1}^{n}$, its estimation is done by minimizing:

$$\sum_{i=1}^{n} r\left(y_i - {x_i}^T \beta\right) = \sum_{i=1}^{n} r\left(y_i - {x_i}^T \beta\right)^2 \tag{11}$$

As such, modelling using only the mean which is what happens in linear regression can lead to missing important aspects of the association between the outcome and its predictors (Beyerlein 2014).

Thus, the quantile regression allows for the examination of changes in distribution and provides more specific information than linear regression, allowing for a better understanding of how pertinent variables affect the entire distribution of an outcome variable rather than just the mean of the distribution (Hajovsky et al. 2020).

4 Case of Study

4.1 Dataset

The dataset used for this study was previously collected on CGC Genetics and is composed of placentas collected from 12 to 41 gestational ages of singleton pregnancies. Of the 2248 entries only 1951 had the available information for this study and as such only these were used. To not allow the direct identification of the patient, the name and information of the patient were changed to a process number.

Table 1 presents a brief description of the variables available in the dataset for this study. As a note, since the placenta is not a perfect circle, two diameters were measured: Diameter1 which is the biggest diameter and Diameter2 which is the smallest.

Table 1. Description of variables available on the dataset.

Variable name	Description	Scale of measurement	Values
MaternalAge	Maternal age	Proportional	Years
GA	Gestational age	Proportional	From 12 to 41
Fetusgender	Fetus gender	Nominal	Levels: M - Male, F - Female, A - Ambigous, 9 - Missing Information
Fetalweight	Birthweight	Proportional	Grams, g
Placentalshape	Placental shape	Nominal	Levels: Normal, Bilobate, Circunvalate, Membranous
Placentalweight	Placental weight	Proportional	Centimetre, cm

(*continued*)

Table 1. (*continued*)

Variable name	Description	Scale of measurement	Values
Diameter1	Diameter 1 - the smallest value of placental diameter	Proportional	Centimetre, cm
Diametre2	Diameter 2 - the highest value of placental diameter	Proportional	Centimetre, cm
Placentalthickness	Placental thickness	Proportional	Centimetre, cm
Placental_vol	Placental volume	Proportional	Cubic Centimetre, cm^3

Table 2 presents some of the summary statistics for the quantitative metrics. We can see that the gestational age varies between 12 and 41 weeks, and the fetal weight has a minimum of 5.4 g and a maximum of 4880 g. The mean was 1248.7 g and 50% of the fetus presents a weight less than 766 g. Concerning the placenta weight, it was observed a minimum of 6 g and a maximum of 995.0 g. The mean placenta weight observed in this sample was 233. 2 g, and the median was 195 g.

Table 2. Descriptive statistics for the quantitative measures.

	GA	Fetalweight	Placentalweight	Diameter1	Diameter2
Min.	12.0	5.4	6.0	1.7	1.50
1st Q	19.0	235.5	101.0	10.0	8.0
Median	26.0	766.0	195.0	13.0	11.0
Mean	26.71	1248.7	233.2	14.48	12.8
3rd Q	35.0	2220.0	345.5	17.0	14.6
Max.	41.0	4880.0	995.0	32.0	30.0

4.2 Growth Curves

Using the previously described dataset the growth charts were made. These charts were created using the package *quantregGrowth (*which uses the package *quantreg*) and *ggplot2*, both are R's packages.

The plots created are illustrated on graphs in Figs. 1, 2, 3 and 4. The values used are not differentiating the genders since splitting the data would lead to a less robust chart and there were no observable differences between the genders that would justify the separation. The choice of percentiles is in line with those that are most common in the medical community.

The 97th percentile in both Fig. 1 and Fig. 2 show a decrease in the slope between the intervals of 35 to 40 weeks of gestational age. Even though we are still not sure about what is happening to cause this, we have two hypotheses: from a statistical point of view, there could be the need for more data to allow to create of a better curve. From a biological point of view, this could show alterations in the placenta in the late stages of the pregnancy.

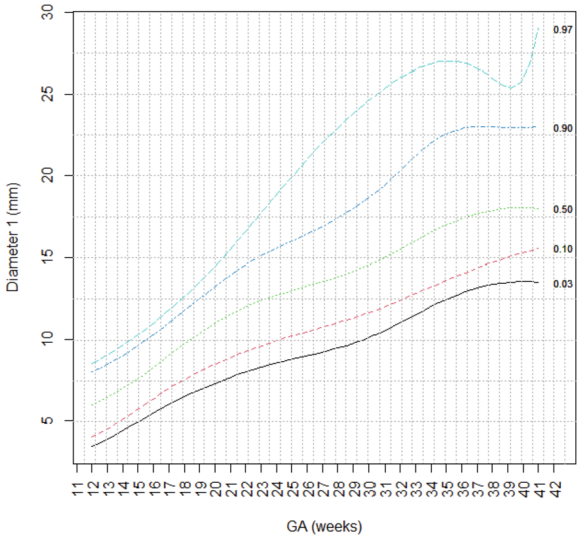

Fig. 1. Growth curve for diameter1 vs GA.

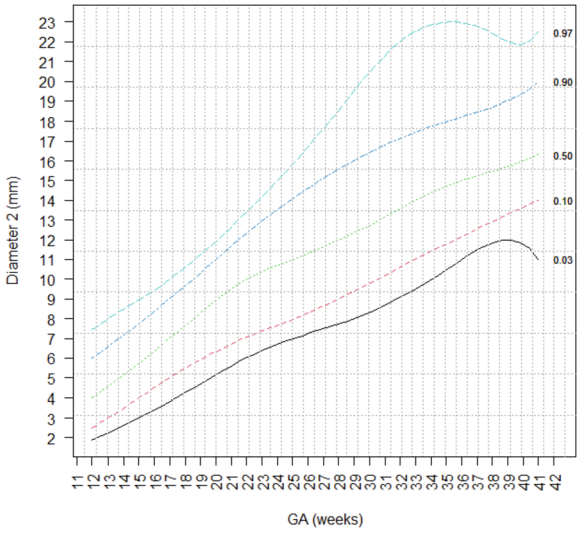

Fig. 2. Growth curve for diameter2 vs GA.

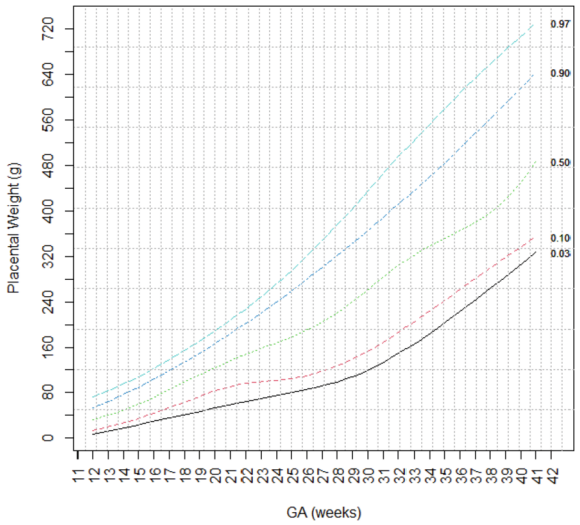

Fig. 3. Growth curve for placental weight vs GA.

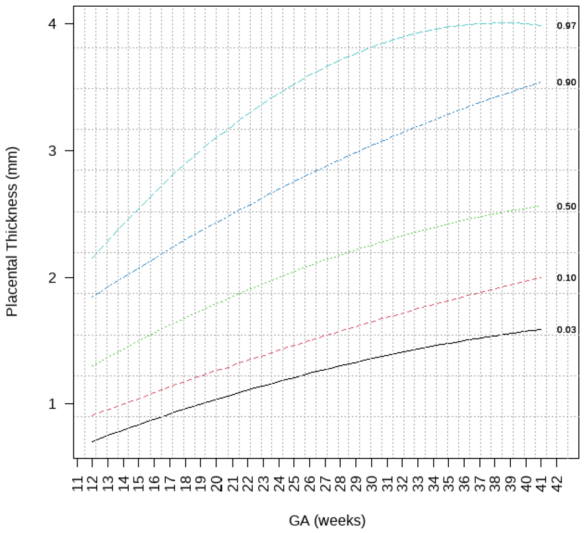

Fig. 4. Growth curve for placenta thickness vs GA.

5 Application Development

5.1 System Architecture

Packages Used. The main package used to create the application is the RStudio's with package Shiny. Another package used is *ggplot2* to show the charts and *rhandsontable* for the tables.

Application. To illustrate how the application works we design the workflow presented in Fig. 5.

The user starts by uploading the files that will create the charts. The User Interface is composed of 4 tabs: 1 tab for the plot and the other for the table for both diameters. It has also a sidebar composed of 2 forms to allow the user to upload the d1 and d2.csv files containing their values. It also has a button that allows the user to save the changes made to the files.

The files uploaded by the user are uploaded on a reactive function to allow the data to be dynamically changeable while the application is executed. After that, two reactive objects are created to store the values that the user changes. Another function is to when the button is pressed, save the CSV file with the changes made by the user by allowing the user to download it.

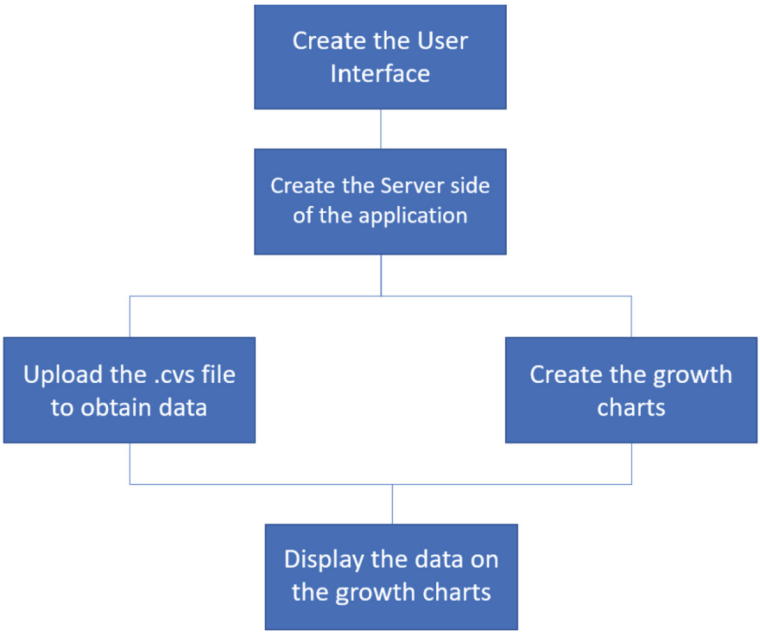

Fig. 5. Workflow for the application.

5.2 Instruction's Manual to Use the Application

The way the application works is by the user uploading a.csv with the different values gathered from the analysis of the placenta. The file has the value 0 for all the gestational ages at the beginning. The user will then change them with the values gathered. After all the changes are made the user can save the CSV file. All the values will also be plotted on the chart allowing the user to see the progression of those values.

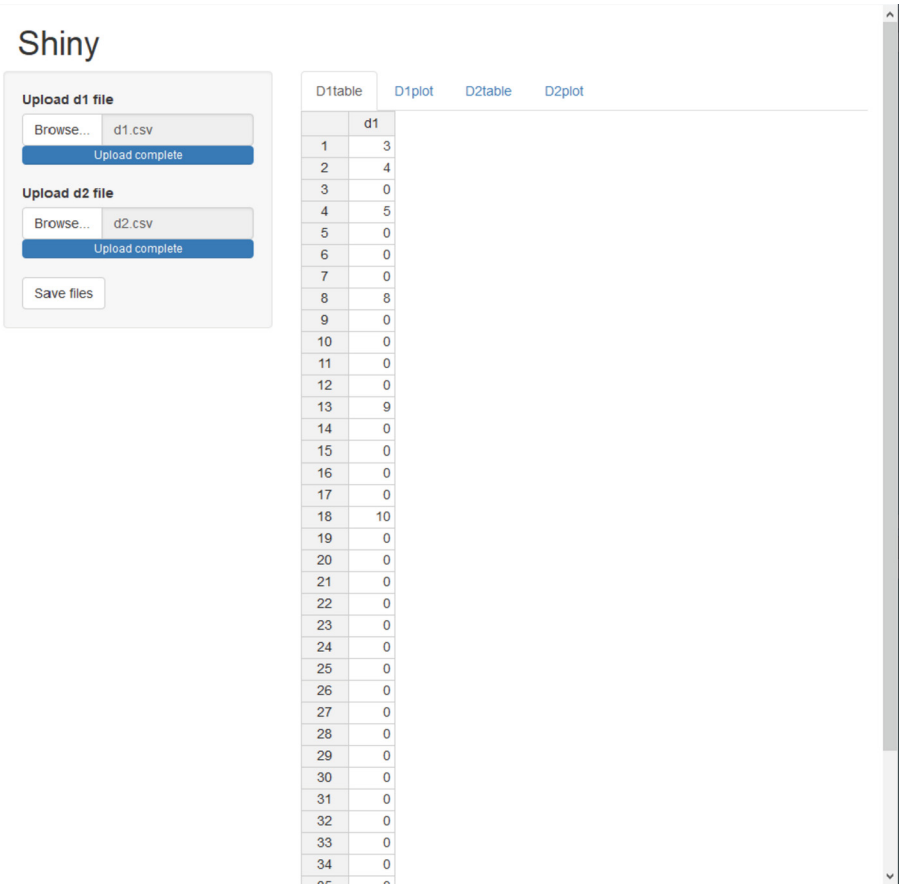

Fig. 6. Application's display of the table.

As illustrated in Fig. 6. the files after being uploaded are shown on the table and updated on the chart like in Fig. 7.

Figure 7 shows the chart of the uploaded file. It is plotted all the values as points (all values higher than 11 gestational weeks since the study was done only after it) and the growth curves each with a different colour.

The user can then edit the values on the app. At the moment, to see the changes on the chart the user has to save the file and then upload it again.

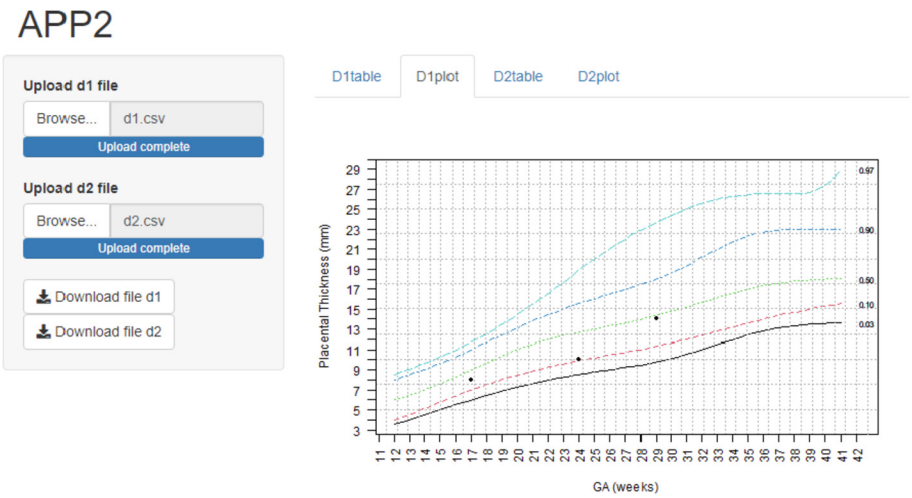

Fig. 7. Application's display of the chart.

6 Conclusion and Future Work

The newly created placenta's growth curves were generated using a more robust regression method consolidating the Portuguese population's placenta growth curves allowing the researchers and doctors to start using these in their studies and patient appointments to better evaluate the fetus's current and future health.

The created application allows the user (a researcher or a clinician) a way to see the evolution of the placenta's dimensions using a simple and robust graph.

Regarding future works, the UI could be improved, allowing the user to change the values and see the changes immediately on the plots without needing to re-upload the file. An interesting idea would be to set up the applications in a real-world situation and be used by many researchers and medical professionals and hear their opinions about the applications and start improving based on their insights.

References

Asgharnia, M., Esmailpour, N., Poorghorban, M., Atrkar-Roshan, Z.: Placental weight and its association with maternal and neonatal characteristics. Acta Medica Iranica **46**(6), 467–472 (2008)

Barker, D.J.P., Bull, A.R., Osmond, C., Simmonds, S.J.: Fetal and placental size and risk of hypertension in adult life. BMJ **301**(6746), 259–262 (1990). https://doi.org/10.1136/bmj.301.6746.259

Beyerlein, A.: Quantile regression - Opportunities and challenges from a user's perspective. Am. J. Epidemiol. **180**(3), 330–331 (2014)

Buhai, S.: Quantile regression: overview and selected applications. Ad-Astra-The Young Rom. Sci. J. 1–20 (2004)

Cade, B.S., Noon, B.R.: A gentle introduction to quantile regression for ecologists. Front. Ecol. Environ. **1**(8), 412–420 (2003)

Chen, C.: Growth charts of body mass index (BMI) with quantile regression. In: Proceedings of the 2005 International Conference on Algorithmic Mathematics and Computer Science, AMCS 2005 **1**, 114–120 (2005)

Hajovsky, D.B., Villeneuve, E.F., Schneider, W.J., Caemmerer, J.M.: An alternative approach to cognitive and achievement relations research: an introduction to quantile regression. J. Pediatric Neuropsychol. **6**(2), 83–95 (2020). https://doi.org/10.1007/s40817-020-00086-3

Hindmarsh, P.C., Geary, M.P.P., Rodeck, C.H., Jackson, M.R., Kingdom, J.C.P.: Effect of early maternal iron stores on placental weight and structure. Obstetric Gynecol. Survey **56**(2), 66–67 (2001)

Kiserud, T., et al.: The World Health Organization fetal growth charts: concept, findings, interpretation, and application. Am. J. Obstet. Gynecol. **218**(2), S619–S629 (2018)

Koenker, R., Bassett, G.: Regression quantiles. Econometrica **46**, 33–50 (1978)

Nogueira, R., et al.: Placental biometric parameters: the usefulness of placental weight ratio and birth/placental weight ratio percentile curves for singleton gestations as a function of gestational age. Jcap **4**, 1–3 (2019)

Wojciechowski, J., Hopkins, A.M., Upton, R.N.: Interactive pharmacometric applications using R and the Shiny package. CPT: Pharmacomet. Syst. Pharmacol. **4**(3), 146–159 (2015)

Yu, K., Lu, Z., Stander, J.: Quantile regression: applications and current research areas. J. Royal Stat. Soc. Ser. D: Stat. **52**(3), 331–350 (2003)