BIOCYBERNETICS AND BIOMEDICAL ENGINEERING XXX (2018) XXX–XXX

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bbe

Biocybernetics and Biomedical Engineering

Original Research Article

Statistical methods for constructing gestational age-related charts for fetal size and pregnancy dating using longitudinal data

Martin Hynek [a,b,*], Jan Kalina [c], Jana Zvárová [c,d,1], Jeffrey D. Long [e,f]

[a] Gennet, Centre for Fetal Medicine and Reproductive Genetics, Prague, Czech Republic
[b] Department of Gynecology and Obstetrics, Thomayer Hospital, Prague, Czech Republic
[c] Institute of Computer Science, The Czech Academy of Science, Prague, Czech Republic
[d] Institute of Hygiene and Epidemiology, First Faculty of Medicine, Charles University, Prague, Czech Republic
[e] Department of Psychiatry, Carver College of Medicine, University of Iowa, Iowa City, IA, United States
[f] Department of Biostatistics, College of Public Health, University of Iowa, Iowa City, IA, United States

ARTICLE INFO

Article history:
Received 2 February 2018
Received in revised form
2 August 2018
Accepted 6 September 2018
Available online xxx

Keywords:
Fetal size
Fetal reference charts
Longitudinal data
Mixed effects models
Regression quantiles
Multimodel inference

ABSTRACT

The assessment of fetal size and the accurate estimation of gestational age are of crucial importance for proper pregnancy management. The information is almost exclusively based on ultrasound measurements of fetal biometric parameters and the means for evaluating these measurements are age-related reference charts (centile charts) allowing interpretation of obtained fetal measurement in comparison with the expected average measurement in the reference population. The construction of such reference charts requires an appropriate statistical methodology. The most frequent method for the construction of fetal reference charts from cross-sectional data is the parametric approach with fractional polynomials regression functions for the mean and standard deviation of each fetal measurement. This article suggests how this method can be extended to longitudinal data using fractional polynomials in linear mixed effect regression. The presented approach includes maximum likelihood estimation for fitting first- and second-order fractional polynomial models, and multimodel inference using Akaike's information criterion and related tools as a suitable strategy for model selection. Finally, an example of the suggested approach is presented.

1. Introduction

The assessment of fetal size and the accurate estimation of gestational age (GA) are of crucial importance for proper pregnancy management. Early detection of fetal growth restriction or macrosomia may decrease associated morbidity and mortality [1,2]. Precise information on GA may prevent unnecessary obstetric interventions at the time of delivery [3].

* *Corresponding authorat*: Gennet, Centre for Fetal Medicine and Reproductive Genetics, Kostelní 9, 170 00 Praha 7, Czech Republic.
E-mail address: martin.hynek@gmail.com (M. Hynek).
[1] Passed away in July 2017.
https://doi.org/10.1016/j.bbe.2018.09.001

The information is almost exclusively based on ultrasound measurements of fetal biometric parameters (e.g. crown-rump length, head circumference, femur length, etc.), which almost invariably increase with GA. The means for evaluating these measurements are age-related reference charts (centile charts) allowing interpretation of obtained fetal measurement in comparison with the expected average measurement in the reference population.

A variety of statistical methods for constructing reference charts has been suggested. To allow for elaboration, let y be the fetal measurement of interest, and let t be gestational age. The statistical objective is to estimate the centile, $c_\alpha(y|t)$, of the conditional distribution of y (given t) for specified values of α (e.g. $\alpha = 0.95$ for the $(100 \cdot \alpha)$th = 95th centile). Moreover, $c_\alpha(y|t)$ is required to be a smooth function of t [4]. There are two main issues in the estimation of $c_\alpha(y|t)$: approximating the distribution of $y|t$, and smoothing estimated centiles on t. Several approaches have been suggested to tackle these issues, including parametric, semiparametric and nonparametric techniques. Detailed overviews and comparisons of different approaches can be found in the literature [5–8].

Interest is usually in the construction of two types of reference charts: charts of fetal size and dating charts. The former is used to assess fetal size, whereas the latter to predict GA. It is incorrect to use the size charts to estimate GA; proper dating charts should be constructed [9]. If the charts of fetal size model the fetal size as a function of GA, the dating charts are produced in similar fashion to size charts, except that they are based on modelling GA as a function of fetal dimension using the conditional distribution, $t|y$.

An appropriate method for constructing reference charts fulfills certain requirements. Altman and Chitty (1993) state that reference centiles should change smoothly with gestation, and they should provide a good fit to the raw data. It is desirable for the statistical model to be as simple as is compatible with these requirements [9]. The WHO Multicentre Growth Reference Study Group (MGRS) [8] agreed on primary criteria for method selection, which include the ability to:

- estimate precisely outer centiles,
- estimate centiles simultaneously in such a way that they are constrained to be ordered (not to cross),
- estimate Z-score and centiles using direct formulae,
- apply continuous age smoothing, and
- account for both skewness and kurtosis when necessary.

In addition, the secondary criteria were:

- the ability to assess fit to the data,
- ease of explanation and clear documentation,
- useful for application to different anthropometric measures, so that WHO growth curves would rely on a single approach.

The vast majority of literature dealing with fetal reference charts describes statistical methods suitable for cross-sectional data, i.e. each fetus contributes only one observation to a reference sample. Because single-visit data is easiest to collect and methodology for computing centiles is straightforward, most published papers in past decades have used cross-sectional datasets to produce reference charts. However, longitudinal data are becoming more common and different techniques must be used for data in which some or all subjects are measured more than once [10]. Among the challenges, serial measurements of individual fetuses induce correlated data that needs to be accounted for in the statistical modeling. There is also the question of how to compute degrees of freedom (*df*), especially when the number of repeated measurements varies among the fetuses [9].

The aim of this article is to compile and propose statistical methods for constructing age-related fetal reference charts using longitudinal data. The method will be used to produce reference charts of fetal size and charts for pregnancy dating for the Czech population in a future analysis. The study will be based on a large dataset collected during routine ultrasound scans in the Fetal Medicine Centre Gennet in Prague. Because a considerable proportion of our fetuses are measured more than once during the course of pregnancy, we are focusing on accommodating currently available methods for construction of fetal reference charts to longitudinal data.

This article focuses solely on statistical methods. The equally important issues concerning sample size, sample selection, inclusion and exclusion criteria, and data collection, are not discussed here and can be found in the relevant literature [9,11–13].

2. Statistical methods

When searching for a suitable statistical method that fulfils the requirements mentioned above, we have to realize two specifics of fetal biometry. First, it is well established that mean fetal dimensions increase monotonically during pregnancy. The between-subject variability of fetal dimensions also tends to increase (spread out over time), which is summarized by the standard deviation (SD) over age. It is crucially important to consider not only the relation between the mean and GA, but also the relation between the SD and GA [9]. Second, we know from many previous studies that the distribution of fetal dimensions is close to normal for any GA [13]. In addition, the mean is a non-linear function of GA for most fetal dimensions, and so too the SD. Therefore, the most used method for construction of fetal reference charts is linear regression using fractional polynomials (FPs) [14], fitted separately for the mean and the SD, assuming normality at each GA, but non-linearity as a function of GA [4,14–16].

First, we will describe the mean and SD method using conventional polynomials (CPs) for the cross-sectional data, then introduce its modification using FPs, and discuss methods for checking the goodness-of-fit. Finally, we will propose an extension of the method to longitudinal data.

2.1. *Cross-sectional data*

2.1.1. *Mean and SD model*

The original mean and SD approach represents parametric modeling proposed by Altman [15] and Royston and Wright [16]. The method assumes normally distributed fetal measurements at each GA and uses CP regression to model the mean and the SD as a function of GA. A desired centile curve is then constructed with the centile defined as

Please cite this article in press as: Hynek M, et al. Statistical methods for constructing gestational age-related charts for fetal size and pregnancy dating using longitudinal data. Biocybern Biomed Eng (2018), https://doi.org/10.1016/j.bbe.2018.09.001