

Reference Growth Charts for Saudi Arabian Children and Adolescents

P. J. Foster & T. Kecojević

First version: 3 September 2009

Research Report No. 12, 2009, Probability and Statistics Group
School of Mathematics, The University of Manchester

Reference growth charts for Saudi Arabian children and adolescents

P. J. Foster and T. Kecojević

Abstract The purpose of this study is to provide Saudi Arabian population reference growth standards for height, weight, body mass index (BMI), head circumference and weight for length/stature. The estimated distribution centiles are obtained by splitting the population into two separate age groups: infants, birth to 36 months and children and adolescents, age 2 to 19 years. The reference values were derived from cross-sectional data applying the LMS method of Cole and Green (*Stat. in Medicine* 1992; **11**:1305-1319) using the `lmsqreg` package in R (public domain language for data analysis, 2009). The report provides an overview of how the method has been applied, more specifically how the relevant issues concerning the construction of the growth charts have been addressed and is illustrated by just using the girls weight data (birth to three years old). These issues include identifying the outliers, diagnosing the appropriate amounts of smoothing and averaging the reference standards for the overlapping 2 to 3 year age range. The use of ANCOVA has been introduced and illustrated as a tool for making growth standard comparisons between different geographical regions and between genders.

KEY WORDS: growth curves; quantile regression; LMS; cross-sectional data

P. J. Foster

School of Mathematics, University of Manchester, Manchester, UK

e-mail: Peter.Foster@manchester.ac.uk

T. Kecojević

Lancashire Business School, University of Central Lancashire, Preston, UK

e-mail: TKecojevic@uclan.ac.uk

Prepared for presentation at: *Conference on Quantitative Social Science Research Using R*, Fordham University, New York, June 18-19, 2009.

1 Introduction

The growth standards are derived from a cross sectional sample of healthy children and adolescents aged from birth to 19 years. The sample was randomly selected by a stratified multistage probability sampling procedure from each of the 13 administrative regions of the Kingdom of Saudi Arabia, ensuring both national and urban/rural representation. The anthropometric data comprises 51,485 observations of which 25,987 are made on boys and 25,498 on girls. Those measurements include: length, for the children 2 years of age and below, height, for children above 2 years of age, weight and head circumference. All possible efforts have been made to ensure reliability and the accuracy of the measurements.

The reference growth charts we have constructed describe the dependance of height, weight, boy mass index (BMI) and head circumference on age, and weight on length/stature for two age ranges, birth to 36 months and 2 to 19 years. They were constructed using the LMS (Lamda-Mu-Sigma) method of Cole and Green [8] in \mathbb{R} , a public domain language for data analysis (R Development Core Team (2009)). The LMS method provides a way of obtaining growth standards for healthy individuals and is based on normalizing the conditional distribution of a measure using the power transformation of Box and Cox [1]. The package `lmsqreg` developed by Carey [2] implements the LMS method in \mathbb{R} . Use of the LMS method was a requirement of the study.

In this paper we discuss various issues involved in using the LMS methodology, all of which are specifically illustrated using the Saudi girls weight data for those from birth to three years old. Section 2 focusses on identifying and removing extreme outliers prior to estimating the centile curves. In Section 3 we describe the model and how it is fitted to the data. Goodness-of-fit is considered in Section 4 while in Section 5 we describe a simple solution to obtaining a common set of centiles in the overlapping 2-3 years age-range. In Section 6 we propose using ANCOVA to investigate differences in growth patterns in different geographical regions and also between the sexes. This proves to be a much more informative approach than that described in the literature which does not take age into account. Finally, we add some discussion and further suggestions.

2 Outliers

An *Outlier* is a sample value that lies outside the main pattern or distribution of the data and in the context of quantile regression, which was first introduced by Koenker and Basset [14], it will be one which has a much larger or smaller response value at a given age when compared with other responses at a similar age. Quantile regression measures the effect of covariates not only in the center of the distribution but also in the upper and lower tails. Extremely low and extremely upper quantiles are of interest regarding growth charts and therefore it is important to deal with the issue of removing the potential outliers with a cautiousness.

The outlier should not be regarded as a pejorative term; outliers may be correct, but they should be checked for transcription error [18]. The quantile regression model is a natural extension of the linear regression model. If an outlier is included in the data which is used to estimate the quantiles then it may be highly influential on the fitted regression line in that the line may be pulled in a disproportionate manner towards the outlying value or it may cause a failure in the algorithm used to estimate the quantiles [14]. This latter point is particularly true with respect to the LMS procedure, as according to Carroll [4] the choice of the transformation $L(x)$ is highly sensitive to outliers in the data. We have also found that if the outliers are not removed it can result in the numerical failure of the model fitting algorithm in the function `lmsqreg`.

The lack of a methodology to assess the direct effect of an individual observation on the LMS methodology has prompted us to approximate the LMS model using a cubic regression line to model the relationship between a response and covariate (such as weight and age). Approximating the LMS model in this way enables us to identify the outliers in that space with respect to this mode, that hopefully are also the outliers with respect to the LMS model. To fit this cubic regression line we have used a robust regression procedure.

Robust regression deals with cases that have very high leverage, and cases that are outliers. Robust regression represents a compromise between the efficiency of the ordinary least squares (OLS) estimators and the resistance of the least absolute value (LAV) estimators, both of which can be seen as special cases of M -estimation [13]. It is a form of weighted least squares regression, which is similar to least square in that it uses the same minimization of the sum of the squared residuals, but it is done iteratively. Based on the residuals a new set of weights are determined at each step. In general, the larger the residuals, the smaller the weights. So the weights depend on the residuals. At the same time, the residuals depend on the model and the model depends on the weights. This generates an iterative process and it goes on until the change in the parameter estimates are below a preset threshold. At the end, instead of all points being weighted equally, the weights vary and those with the largest weights contribute more to the fit.

There are a few types of weighting schemes, M -estimators, that can be implemented [18]. In Huber's [13] weighting, observations with small residuals get a weight of 1, the larger the residual, the smaller the weight. M -estimation, introduced by Huber (1964) can be regarded as a generalisation of maximum-likelihood estimation (MLE), hence the term ' M '-estimation [10].

Consider the linear model

$$y_i = x_i' \beta + \varepsilon_i \quad i = 1, \dots, n \quad (1)$$

where the $V(\varepsilon_i) = \sigma^2$ and $Cov(\varepsilon_i, \varepsilon_j) = 0, i \neq j$. If ε_i has density f , we can define $\rho = -\log f$, where the function ρ gives the contribution of each residual to the objective function. Then the MLE $\hat{\beta} = b$ solves

$$\min_{\beta} \sum_i -\log f(y_i - \mu_i) = \min_{\beta} \sum_i \rho(y_i - \mu_i) \quad (2)$$

where $\mu_i = x_i' \beta$ and so $\hat{\mu}_i = x_i' b$.

Let $\psi = \rho'$ be the derivative of ρ . Then we will have $\sum_i \psi(y_i - \hat{\mu}_i) x_i' = 0$ or $\sum_i w_i (y_i - \hat{\mu}_i) x_i' = 0$ where the weight $w_i = \psi(y_i - \hat{\mu}_i) / (y_i - \hat{\mu}_i)$. This suggests an iterative method of solution, updating the weights at each iteration [18].

If $\rho(x) = x^2$, the solution is the conditional mean and the median is $\rho(x) = |x|$. The function

$$\psi(x) = \begin{cases} -c & x < -c \\ x & |x| < c \\ c & x > c \end{cases} \quad (3)$$

is known as *Winsorizing* and brings in extreme observations to $\mu \pm c$. The corresponding function $\rho = -\log f$ is

$$\rho(x) = \begin{cases} x^2 & \text{if } |x| < c \\ c(2|x| - c) & \text{otherwise} \end{cases} \quad (4)$$

and equivalent to a density with a Gaussian centre and double-exponential tails. This estimator is due to Huber. Note that its limit as $c \rightarrow 0$ is the median, and as $c \rightarrow \infty$ the limit is the mean. The value $c = 1.345$ gives 95% efficiency at the normal [18].

Venables and Ripley's MASS package [17] introduces the `r1m` function for fitting a linear model by iterated re-weighted least squares (IWLS) regression using Huber's M -estimator with tuning parameter $c = 1.345$ and also incorporating a robust estimate of the scale parameter σ , where $\hat{\sigma} = s$. The details are; if we assume a scaled pdf $f(e/\sigma)/\sigma$ for ε and set $\rho = -\log f$, in this case the MLE minimizes

$$\min_{\beta} \left[\sum_i \rho \left(\frac{y_i - \mu_i}{\sigma} \right) + n \log \sigma \right] \quad (5)$$

Assuming that σ is known and if $\psi = \rho'$, the the MLE b of β solves

$$\min_{\beta} \sum_i x_i \psi \left(\frac{y_i - \mu_i}{\sigma} \right) = 0 \quad (6)$$

A common way to solve the above equation is by IWLS, with weights

$$w_i = \psi \left(\frac{y_i - \hat{\mu}_i}{\sigma} \right) / \left(\frac{y_i - \hat{\mu}_i}{\sigma} \right) \quad (7)$$

Of course, in practice the scale σ is not known. However, as mentioned above σ is estimated by a robust MLE-type estimate denoted by s .

A cubic polynomial using the `r1m` function in R has been fitted to the *log*-transformed data (in a bid to stabilize the variance over age) using *MM*-estimation that combines the resistance and robustness, whilst gaining the efficiency of *M*-estimation.

```

> library(MASS)
> mp<-rlm(log(weight)~1+agey+I(agey^2)+I(agey^3), method="MM")
> summary(mp)

Call: rlm(formula = log(weight) ~ 1 + agey + I(agey^2) +
          I(agey^3), method = "MM")

Residuals:
    Min       1Q   Median       3Q      Max
-0.784423 -0.098632 -0.001707  0.096245  0.708137

Coefficients:
            Value      Std. Error t value
(Intercept)  1.1731     0.0034   342.0763
agey         2.0866     0.0148   140.8422
I(agey^2)    -1.1875     0.0145  -81.6688
I(agey^3)     0.2223     0.0036   61.0781

Residual standard error: 0.144 on 6123 degrees of freedom

```

After fitting this cubic line we have used the weights produced in a robust regression procedure to identify the most extreme values. The observations with the big residuals are down weighted, which reflects that they are atypical from the rest of the observations when it comes to fitting such a model. Observations with 0 weight ($w_i = 0$) are deemed to be extreme and so are then removed from the data before running the LMS model fitting algorithm (Figures 1 and 2). Please note that weight referred to in the Figure 1 corresponds to girls actual body weight.

row.names	weight	agey	w
811	15.7	0.6	0
1235	2.5	0.2617	0
2240	4.8	1.21	0
5963	30.2	2.96	0
1149	11	0.33	0.002548397
1993	6.8	2.7819	0.00273806
3250	20	1.86	0.02485549
4170	8	0.15	0.02486067
1201	9.8	0.3	0.06277497
1305	8.4	0.2	0.06360398
3454	14	0.65	0.1121188
86	18.5	1.97	0.1242005
4131	3	0.259	0.1309751
3595	7	0.1206	0.1326621
5740	6	1.0648	0.1556946
5392	10	0.36	0.1615647

Fig. 1 Identifying the outliers, Girls Weight, age birth to 36 months

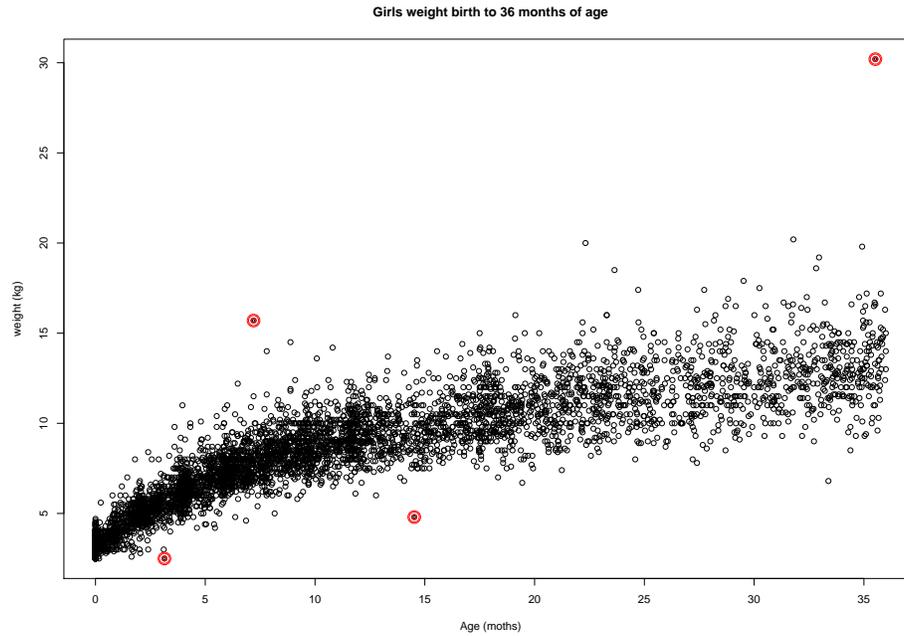


Fig. 2 Identified outliers, Girls Weight, age birth to 36 months

The World Health Organisation (WHO) has defined limits for acceptable data based on 1977NCHS/WHO growth charts and recommends that the exclusion range for weight-for-age should be $|z| > 5$ [5]. After the final LMS model for girls weight (age birth to three) was fitted, we used the `zscores` function from the `lmsqreg` package in R to calculate z -scores for the four identified outliers and these are given in Table 1. Each omitted case has an $|z|$ greater than 5 tying in with the WHO guideline.

<code>z-scores {lmsqreg}</code>			
<code>row.names</code>	<code>weight</code>	<code>age</code>	<code>z-score</code>
811	15.7	0.6	5.22488
1235	2.5	0.2617	-6.70038
2240	4.8	1.21	-5.29738
5963	30.2	2.96	6.49793

Table 1 z -scores of the four identified outliers for girls weight, age birth to 36 months

3 LMS

Under the assumption of normality, growth curves can be constructed by estimating the age specific mean and standard deviation, say $\mu(t)$ and $\sigma(t)$, so that chosen quantile curve for $\alpha \in [0, 1]$ can then be obtained as

$$\hat{Q}(\alpha | t) = \hat{\mu}(t) + \hat{\sigma}(t)\Phi^{-1}(\alpha) \quad (8)$$

where $\Phi^{-1}(\alpha)$ denotes the inverse of the standard normal distribution function. Providing that assumption of normality holds at each age, such a curve should split the population into two parts with the proportion α lying below the curve, and the proportion of $1 - \alpha$ above the obtained curve [19].

Although adult heights in a reasonably homogeneous population are known to be quite close to normal, in general anthropometric data are known to be not normally distributed [19]. Anthropometry tends to be right skew rather than left skew, which is why a log transformation which treats the two tails of the distribution differently is often suggested as a means of obtaining a symmetric distribution [7]. A log transformation can be viewed as a particular power transformation of the data but there is a whole family of such powers. Cole [6] suggested that in principle, there is no reason why a general power transformation should not be applied to the data. The maximum likelihood estimate (MLE) for the power, which both minimises the skewness and optimises the fit to normality, is ideally suited to the problem of skew data. However, it only operates on individual groups and does not allow for the skewness to change in a smooth manner over the range of the covariate.

The LMS, or $\lambda\mu\sigma$, approach of Cole [6] provides a way of obtaining normalised growth centiles that deals quite generally with skewness as well as non-constant variance. The method enables us to fit the growth standards to all forms of anthropometry by making the simple assumption that the data can be normalised by using a smoothly varying Box-Cox transformation, so that after the transformation of the measurements $Y(t)$ to their standardised values $Z(t)$ they will be normally distributed.

$$Z(t) = \frac{[Y(t)/\mu(t)]^{\lambda(t)} - 1}{\lambda(t)\sigma(t)} \quad (9)$$

With these normalised measurements, the desired quantile curve for $\alpha \in [0, 1]$ can then be obtained using the following model

$$Q(\alpha | t) = \mu(t)[1 + \lambda(t)\sigma(t)\Phi^{-1}(\alpha)]^{1/\lambda(t)} \quad (10)$$

which summarises the construction of the centiles by three smooth curves, ie. functions, representing the skewness, the median and the coefficient of variation. The LMS method works with power transformed measurements, but converts the mean back to original units and uses coefficient of variation (CV) rather than standard de-

viation of the data. In this way the results for different power transformations can be compared, and the best (Box-Cox) power can be identified as the one which gives the smallest CV [7]. This method provides a coherent set of smoothed centiles and the shape of the power curves provide information about the changing skewness, median and coefficient of variation of the distribution.

The three parameters λ , μ and σ were assumed to change smoothly with age. Green [11] has proposed to estimate the three curves by maximizing the penalised likelihood,

$$\ell(\lambda, \mu, \sigma) - v_\lambda \int (\lambda''(t))^2 dt - v_\mu \int (\mu''(t))^2 dt - v_\sigma \int (\sigma''(t))^2 dt \quad (11)$$

where $\ell(\lambda, \mu, \sigma)$ is the Box-Cox log-likelihood function derived from (9),

$$\ell(\lambda, \mu, \sigma) = \sum_{i=0}^n [\lambda(t_i) \log \frac{Y(t_i)}{\mu(t_i)} - \log \sigma(t_i) - \frac{1}{2} Z^2(t_i)] \quad (12)$$

and $Z(t_i)$ are the SD scores corresponding to $Y(t_i)$. In this way, the three curves are constrained to change smoothly as the covariate changes and, like the centiles, they can be plotted against the covariate (Figures 3 and 4). The curves are fitted using cubic splines to give a non-linear regression, and the extent of the smoothing required can be expressed in the terms of smoothing parameters $(v_\lambda, v_\mu, v_\sigma)$. These quantities are defined to be the traces of the relevant smoothing matrices and are referred to as the "equivalent degrees of freedom" (edf) [19]. Cole and Green [8] argued that the distributions of $(v_\lambda, v_\mu, v_\sigma)$ in the LMS model are largely independent of each other, implying that one edf can be optimised while fixing the other two.

```
> mw3<-lmsqreg.fit(weight, age, edf=c(7,13,9),
  pvec = c(0.03, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.97))
> plot(mw3)
> points(age, weight, pch=".",col="red")
```

Carey [2] has developed the `lmsqreg` package that implements the LMS method in R. Smoothed centiles curves have been fitted to the reference data using `lmsqreg.fit` function with suggested starting edf values setting of 3, 5 and 3 for λ , μ and σ , respectively [3]. The strategy is then to optimise the μ curve edf, by increasing/decreasing the edf by 1 until the change in penalised likelihood is small i.e. less than 2. Once the μ curve is fitted, the process is repeated on the σ curve avoiding the value for edf of 2 which would force a linear trend on the μ curve. Finally, the λ curve was fitted similarly to the σ curve (Figure 3). However, in cases of fitting the centiles curves for weight measurement age 2 to 19 years for both sexes λ had to be set to the value of zero, which constrains the entire curve to be a constant value and forces a log transformation (Figure 4). The same had to be applied for the fitting of male head circumference age 2 to 19 years.

```
> mw19<-lmsqreg.fit(weight, age, edf=c(0,14,8), pvec = c(0.03,
  0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.97), lam.fixed=0)
> plot(mw19)
```

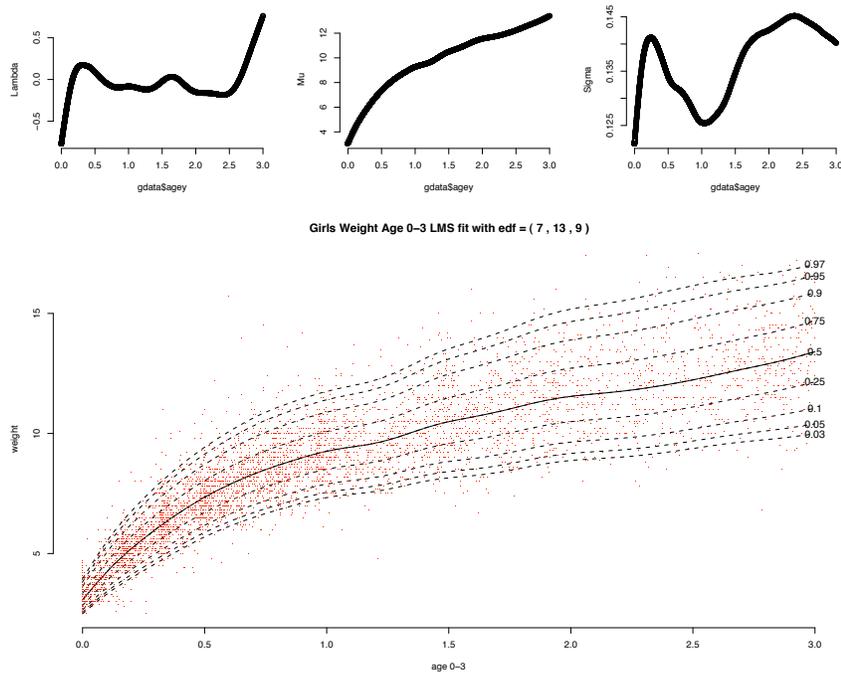


Fig. 3 Centile curves for girls weight birth to 36 months of age

```
> points(age, weight, pch=".", col="cyan")
```

Following the suggested strategy, the data was over-fitted and the curves were clearly undersmoothed. As Cole [8] implies the case for making the centile curves smooth is to some extent cosmetic - the centiles are more pleasing to the eye when smoothed appropriately but it is also in the belief that the true population centiles will themselves change smoothly. Any non-parametric curve estimation method requires some means of controlling the smoothness of the fitted functions. For the LMS method this control is provided by the edf parameters $(\nu_\lambda, \nu_\mu, \nu_\sigma)$.

As indicated by Carey [2] the value in which to increase/decrease edf and the change in penalised likelihood depends on the sample size. For large samples the change of less than 2 units is not significant therefore the large change is needed and the final decision should depend on the appearance of the curve. In order to overcome the over-fitting of the curves the edf values had to be relaxed.

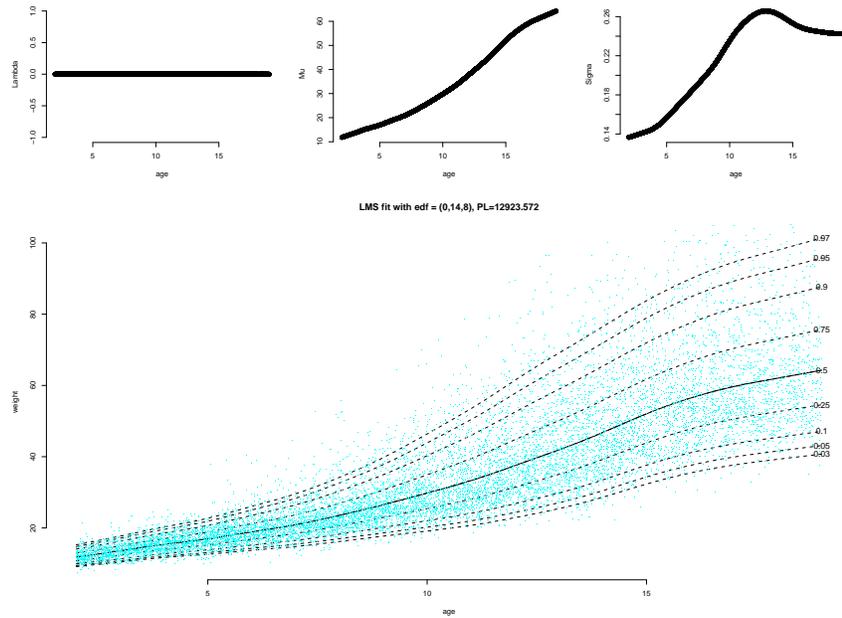


Fig. 4 Centile curves for boys weight 2 to 19 years of age

4 Smoothing and evaluation

The number of effective degrees of freedom is a convenient parameter that expresses the amount of adjustment necessary for smoothing a set of data. Adjustment of edf values was done following Carey's [2] algorithm, this time decreasing the value for ν_μ by 1 until the curve appeared to be smooth. The same procedure was followed for ν_σ and lastly for ν_λ (Figure 5). Finally, the adequacy of the chosen model is evaluated using the original data.

As discussed by Green [11], the distribution theory for model evaluation statistics formed on the bases of changes in penalised likelihood is currently still undeveloped. We have adopted a local-test based approach to formal model evaluation. Carey's *lmsqreg* package [2] provides as a part of the output for a fitted model a collection of model-based z -scores derived from the given quantile regression model. They are stratified based on the covariate t , and within this strata, z -scores are tested for marginal Gaussianity (Kolmogorov-Smirnov test), zero mean (Student's t -test) and unit variance (χ^2 test) [3].

```
> mw3
Dependent variable: gdata$weight , independent variable: gdata$agey
The fit converged with EDF=( 4,6,3 ), PL= 9198.316

KS tests: (intervals in gdata$agey //p-values)
```

```

(-0.001,0]      (0,0.348] (0.348,0.802] (0.802,1.54]      (1.54,3] Overall
      0.000          0.000          0.271          0.324          0.676          0.001

t tests: (intervals in gdata$agey //p-values)
(-0.001,0]      (0,0.348] (0.348,0.802] (0.802,1.54]      (1.54,3] Overall
      0.006          0.000          0.562          0.369          0.568          0.810

X2 tests (unit variance): (intervals in gdata$agey //p-values)
(-0.001,0]      (0,0.348] (0.348,0.802] (0.802,1.54]      (1.54,3] Overall
      0.000          0.000          0.717          0.050          0.462          0.979

```

The above output from the final fitted model shows that the hypotheses of a zero mean, unit variance normal distribution in the intervals close to birth are rejected. The original data is strongly skewed and the edf parameters finally selected are not able to transform the data sufficiently well, with the final empirical distribution being slightly skewed. If the smoothing parameters are increased, in particular v_λ , the normality of the transformed data can be successfully achieved. However, as discussed earlier in section 3, we reduced the values of the optimal smoothing parameters in order to obtain smoother estimated centiles curves.

Table 2 reports on the accuracy of the quantile regression fit in terms of the discrepancy between the nominal and empirical proportions of data lying beneath selected quantile function for age group birth to 3 years. By and large these results show that the quantiles of the fitted models do fit the data well.

sex	variable	N	<i>p</i>								
			0.03	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.97
female	weight	6,123	0.025	0.052	0.090	0.240	0.506	0.755	0.905	0.950	0.972

Table 2 Table entries are quantile coverage probability estimates. Measurement: Age: birth to 36 months.

5 Averaging

We were required to produce reference standards for two age groups: birth to 36 months of age and 2 to 19 years of age. The overlap for the two sets of charts occurs for ages between 2 and 3 years. The values for both sets of standards in the overlapping age range is a product of the model fitted to the whole data set for each specific age group. This means that the centile curves for a particular measurement in this overlapping period will not be the same for the two sets of charts as they are based on using different data outside the range 2 to 3 years (Figure 6).

One of the arguments of `lmsqreg.fit` function is `targlen` which defines the number of points at which smooth estimates of λ , μ , and σ should be extracted for quantile plotting. For both sets of charts we have adopted the default value of 50 for the `targlen` argument. For the overlapping period 2 to 3 years this produces

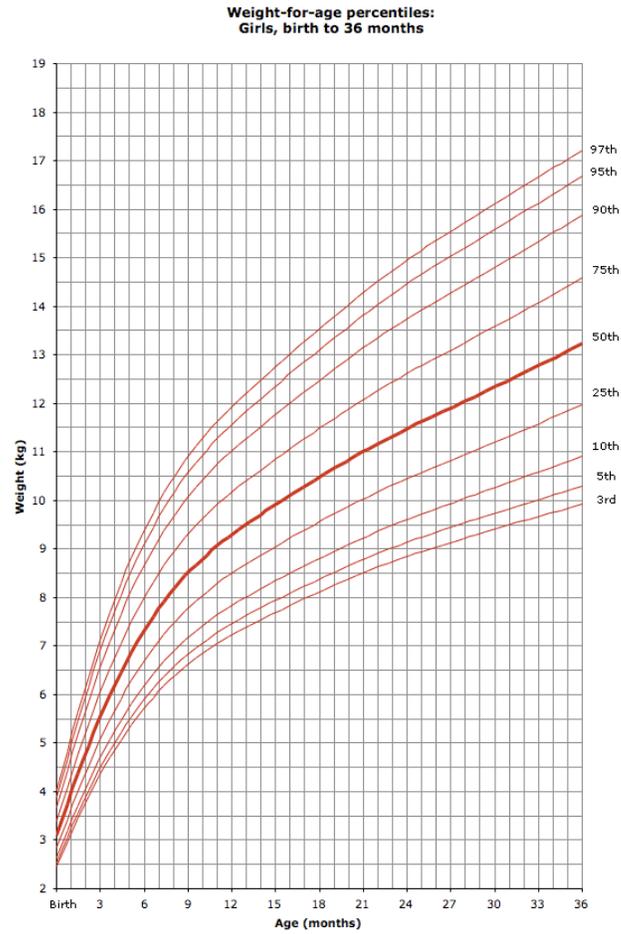


Fig. 5 Final smooth centile curves for girls weight birth to 36 months of age

17 points in the birth to 36 months chart and 3 points in the chart for age 2 to 19 years (Figure 7).

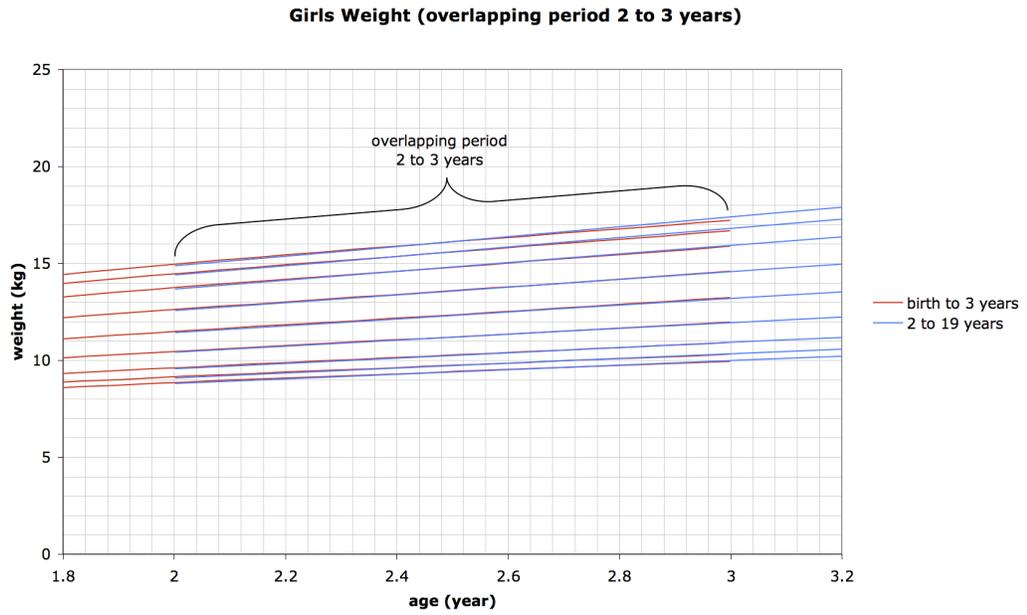


Fig. 6 Overlapping charts: girls weight

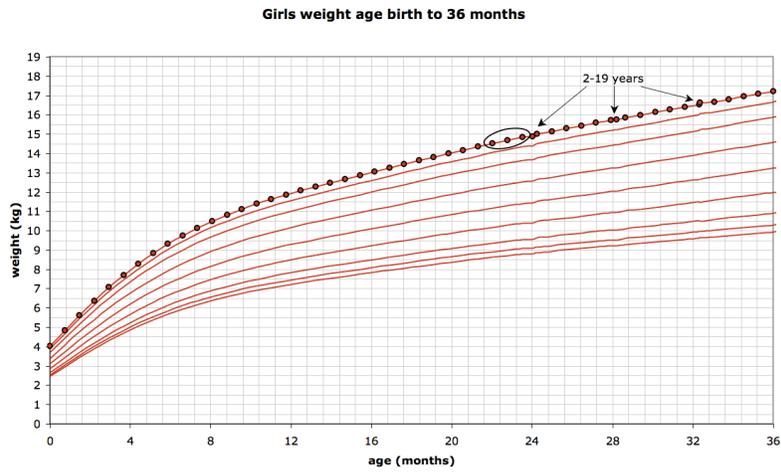


Fig. 7 Centile curves for girls weight birth to 36 months of age

In order to make the centile curves for a particular measurement for this overlapping period the same for the two sets of charts we have re-estimated the curves using the following cubic polynomial:

$$\hat{y}_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 \quad (13)$$

To estimate this cubic polynomial for each of the centiles at the lower and upper boundaries of the overlapping period we have used three adjacent points from each of the charts (Figures 7 and 8), using the least squares estimator given by (14).

$$\hat{Y} = X \left[X'X \right]^{-1} X'Y \quad (14)$$

For the overlapping period new estimates were calculated using the newly found polynomial resulting in a smooth overlap (Figure 9). This means that the centiles for a particular measurement will be the same in the birth to 36 months chart as in the 2 to 19 years age chart.



Fig. 8 Centile curves for girls weight age 2 to 19 years

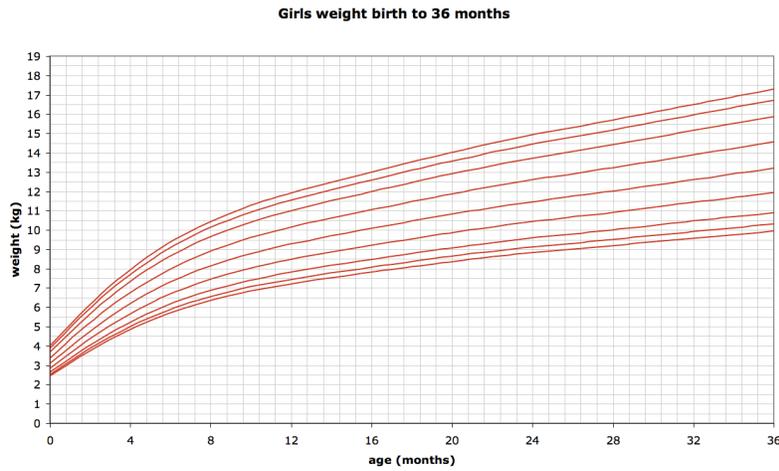


Fig. 9 Final smooth centile curves for girls weight birth to 36 months of age

6 Comparisons using ANCOVA

6.1 Comparing Geographical Regions

In the following analysis the aim is, for a particular measurement, sex and age group, to compare the growth trends over age in different geographical regions. These are:

- i. North,
- ii. Southwest,
- iii. Central.

This means that we are looking at a large proportion of the original data used to fit the LMS models but not all of it as some individuals live in regions other than those listed above.

One approach, for a particular measurement and sex, would be to fit a different LMS model to the data in each region and then to compare the fitted models. However, we are not aware of any existing methodology to make such direct LMS model comparisons. In our proposed approach, we have taken the final LMS model fitted to all the data and used it to transform all the individual measurements into standard deviation scores.

Then, in step 1, a separate cubic regression curve was fitted, where the response ("y-variable") is the SDS score and the covariate ("x-variable") is age, to the data in each of the three regional groups. These regression lines describe how the mean SDS score of a given measurement changes with age in each region. The fit of the three cubic regression curves were then compared with the fit of three quadratic regression curves. If the difference in fits was not statistically significant then the quadratic

models were accepted and they were then compared with three linear regression curves and so on until the simplest model that might be fitted is three different constant horizontal lines. The three final regression lines can be plotted to provide a graphical description of the differences (Figure 10).



Fig. 10 SDS score regression models in the three geographical regions for weight vs. age; age: birth to 36 months; sex: female

If there are no differences in the three regions in how a particular measurement for a given age group and sex changes with age then a single common regression line would be an appropriate model for all the data in the three regions. Therefore, in step 2, such a model was fitted to the data. It would be expected that it would be fairly close to the zero line but not identically zero because we have not used all the original data in this analysis as explained above. The degree of this line (cubic, quadratic, etc.) was chosen to be the same as that of the best fitting three separate ones.

The next stage is to statistically test the fit of the model involving three separate regression lines with the fit of the model based on a single common regression line. We would expect the total residual sum of the squares for the model involving three regression lines to be less than that which just involves one but we need to test whether the difference is statistically significant. The method we have used is a standard "*F-test*" in this context, which is appropriate because the standardised data is Normally distributed. If the *p*-value of this test is small (less than 0.05) the conclusion would be that the single regression line is inadequate and there are significant

differences between the regions in how mean SDS score of a given measurement of a given age group of a given sex changes with age (Table 3).

age: birth to 36 months		
sex	variable	<i>p</i>
female	weight	$< 10^{-6}$

Table 3 Resulting *p*-values when testing a common regression model vs. different regression models for the three regions.

Finally, after finding a significant result we can then go on to use the same methodology as above but just use pairs of regions in turn to see which are significantly different from each other.

This procedure can be summarized for a given sex and measurement by the following steps:

- [i] STEP 1: Find the best fitting polynomials having the lowest possible common degree for each of the three regions.
- [ii] STEP 2: We want to answer the question "Is a common polynomial of the same degree as found in STEP 1 appropriate for all three regions or do the polynomials vary with region?"
 ie. for a particular measurements, sex and age group we want to test:
 $H_0 : E[z | age] = \beta_0 + \dots + \beta_q age^q$ for each region, where $q \leq 3$ is the degree of the common best fitting polynomial.
 vs. H_1 : The polynomial for at least two regions differ.
- [iii] STEP 3: After finding a significant result in STEP 2 carry out pairwise comparisons between the regions.

Note that the notation $E[z|age]$ denotes the mean value of z at the given age. The hypothesis H_0 says that a common polynomial of degree q describes the trend in z -scores over age in each region. On the other hand, the alternative hypothesis, H_1 , says that the trend in z -scores is described by different polynomials of the same degree q in the regions.

Nb. in the following tables p denotes the "p-value" found when testing as above H_0 vs. H_1 . Its' value corresponds to the probability of observing a test statistic value at least as large as we have done and is calculated under the assumption that the null hypothesis, H_0 , is true. For the stepwise and overall tests in STEP 1 and STEP 2 above it is common practice to reject H_0 in favor of H_1 if $p < 0.05$ or if the sample size is large than we may use $p < 0.01$. To account for carrying out multiple comparisons (or multiple hypothesis tests) between pairs of regions for a particular measure as in STEP 3, we would suggest that H_0 is rejected in favor of H_1 if $p < 0.003$ (ie. $0.01/3$) using the Bonferroni method which divides the total significance level into 3 equal proportions corresponding to the number of pairwise comparisons we are carrying out.

Note that the results of the analyses carried out in STEP 2 for age birth to 3 years are given in Table 4. The coefficients of the polynomials in the three separate regions, as well as for all three regions together, are in Tables 4. Those polynomials

for girls weight age birth to 3 years are plotted in Figure 10. Table 5 details the p -values for all the pairwise comparisons between regions.

sex: female					
variable	region	β_0	β_1	β_2	β_3
weight	central	0.04201	0.38775	-0.34224	0.09184
	north	0.4412	-1.2076	0.91695	-0.18133
	southwest	-0.36297	-0.20639	0.10564	-0.02648
	all	0.039995	-0.005139	-0.020645	0.007602

Table 4 Estimates of the model parameters for individual regions and all three regions together for female weight, age range birth to 36 months.

sex:female, age birth to 36 moths	
weight	
	P
north-central	$< 10^{-6}$
southwest-central	$< 10^{-6}$
southwest-north	$< 10^{-6}$

Table 5 p values for the pairwise comparisons between the different regions using ANCOVA

There are clearly significant differences between the regions for each of the measurements for both sexes in each of the age ranges.

6.2 Comparing Males and Females

Standard deviation scores were used to compare the growth patterns between boys and girls using very similar methodology as to that described above when comparing the geographical regions. In order to make comparisons for a given measure between genders, we have used the relevant fitted girls' LMS model to standardise both girls and boys measures using the `zscores` function from Carey's [2] `lmsqreg` package.

We can then plot these standardised measures against age and construct separate regression lines for boys and girls. Considering that the data were standardised by the girls model, it is evident that the appropriate regression model for girls would be zero. However, the z scores of the boys could be explained by an appropriate polynomial regression model (up to cubic polynomial), describing the existing differences between boys and girls. If there are differences then this will be indicated by a non-zero regression line and we can test whether the two lines are significantly different from each other using ANCOVA. We have also superimposed girls and boys

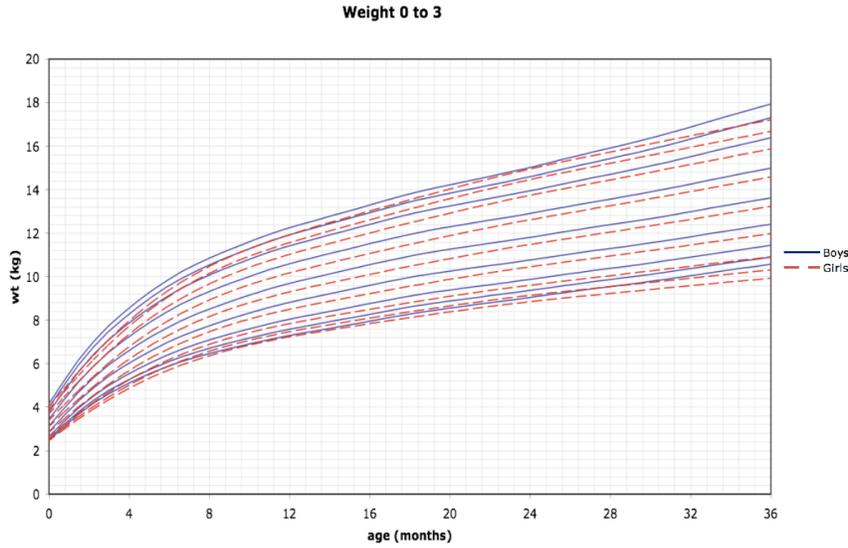


Fig. 11 Comparisons of the growth charts for weight measurement between male and female birth to 3 years of age.

centiles for a given measure on the same plot to give another graphical impression of any differences (Figure 11). For children aged 0-3 we found significant differences for each measure and the fitted regression lines (Table 6) describe how the differences (measured in girls standard deviation scores) change with age (Figure 12).

Age birth to 36 months				
variable	β_0	β_1	β_2	β_3
length	0.21719	0.12799	-0.06722	-
head circumference	0.22312	0.75785	-0.56355	0.12686
weight	0.16774	0.70106	-0.65468	0.14854
body mass index	-	0.52468	-0.50530	0.12154

Table 6 Estimates of the model's parameters.

7 Discussion

This study was set up by the Saudi medical authorities who required growth charts based entirely on data collected from Saudi children and adolescents rather than

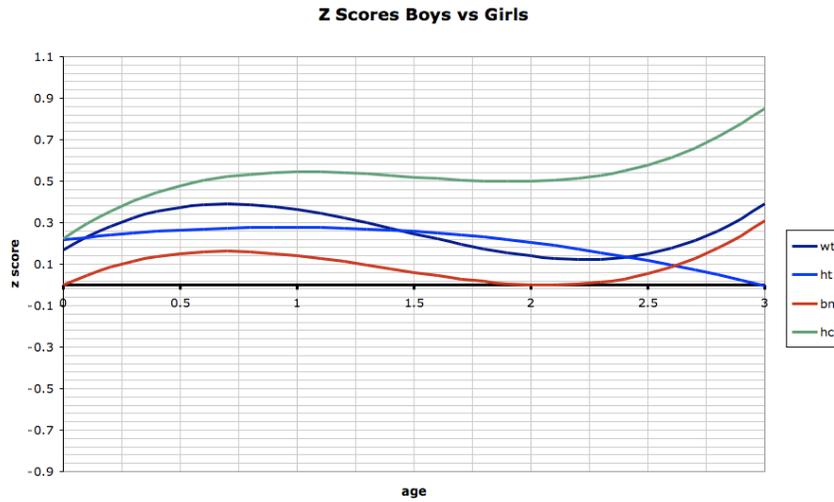


Fig. 12 Comparisons of growth patterns between boys and girls birth to 3 years of age

using a more general alternative, such as those provided by the WHO. We have seen that for girls weight (birth to 36 months) the age-specific conditional quantile estimates we have constructed using the LMS method by and large successfully capture the main features of the data and this also proved to be true for the other growth parameters. In further work we have compared the new Saudi charts with the 2006 WHO standards and found that there are marked differences in corresponding centiles. Use of the WHO standards in Saudi Arabia would, for example, increase the prevalence of undernutrition, stunting and wasting [9].

An essential part of our procedure was to try to identify outliers to be removed from the data prior to estimating the LMS model. We used the robust regression `rlm` function to do this, basing our assessment on the weight attached to each observation by the procedure. We should stress that we were not using this model to make any formal inferences about the form of the conditional mean function. As seen in Section 2, this worked well with four cases being removed. If these cases were included then there is a numerical failure in the LMS model estimation algorithm. All four deleted cases had z-scores greater than 5 in absolute value. The only other case which had an absolute z-score bigger than 5 is case 4131 with a z-score of -5.078, who can be seen listed in Figure 1. This corresponds to a girl aged 0.259 years (3.11 months) who had a weight of only 3.0 kg which is a little higher than case 1235 whose weight was only 2.5 kg at a similar age and who was deleted from the data.

A number of authors have reported that there can be a problem with significant kurtosis in the residuals from the LMS method [15]. Stasinopoulos and Rigby have developed the more flexible Box-Cox power exponential model to overcome this

where they add an extra parameter to model kurtosis [16]. The LMS model we have used is a special case of this. In this study we did not consider this alternative.

8 Acknowledgment

We are very grateful to Professor Mohammad I El Mouzan, MD; King Saud University, Riyadh, for providing the data.

References

1. Box GEP, Cox DR. *An analysis of transformations*. Journal of the Royal Statistical Society, Series B, 1964; **26**:211–252.
2. Carey VJ. *LMSqreg: An R package for ColeGreen reference centile curves*, 2002. <http://www.biostat.harvard.edu/carey>.
3. Carey VJ, Youg FH, Frenkel LM, McKinney RM. *Growth velocity assessment in pediatric AIDS: smoothing, penalized quantile regression and the definition of growth failure*. Statistics in Medicine, 2004; **23**:509–526
4. Carroll RJ. *Two examples of transformations when there are possible outliers* Applied Statistics, 1982; **31**:149–152
5. Centers for Disease Control and Prevention (CDC). *Cut-offs to define outliers in the 2000 CDC Growth Charts*, 2009. <http://www.cdc.gov/nccdphp/dnpa/growthcharts/resources/BIV-cutoffs.pdf>
6. Cole TJ. *Fitting smoothing centile curves to reference data*. Journal of the Royal Statistical Society, Series A—General 1988; **151**:385–418.
7. Cole TJ. *The LMS method for constructing normalized growth standards*. European Journal of Clinical Nutrition, 1990; **44**:45–60.
8. Cole TJ, Green PJ. *Smoothing reference centile curves: the LMS method and penalized likelihood*. Statistics in Medicine, 1992; **11**:1305–1319.
9. El Mouzan, MI, Foster PJ, Al Herbish AS, Al Salloum AA, Al Omar AA, Qurachi MM, Kecojevic T. *The Implications of using the world health organization child growth standards in Saudi Arabia*. Nutrition Today, 2009; **44**(2):62–70.
10. Fox J. 2002. *An R and S-Plus companion to applied regression*. SAGE Publications
11. Green PJ. *Penalized likelihood for general semi-parametric regression models*. International Statistical Review, 1987; **55**:245–259
12. Huber PJ. *Robust estimation of a location Parameter*. Annals of Mathematical Statistics, 1964; **35**:73–101
13. Huber PJ, 1981. *Robust Statistics*. New York: John Wiley and Sons
14. Koenker R, Basset G. *Regression Quantiles*. Econometrica, 1978; **46**:33–50
15. Stasinopoulos M, Rigby B, *Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution*. Statistics in Medicine, 2004; **23**:3053–3076
16. Stasinopoulos M, Rigby B, Akantziliotou C. *GAMLASS: An R package for generalised additive models for location, scale and shape*, 2009. <http://studweb.north.londonmet.ac.uk/stasi-nom/gamlss.html>
17. Venables WN, Ripley BD. *MASS: An R package in the standard library of Venables and Ripley*, 2007. <http://cran.r-project.org/src/contrib/Descriptions/VR.html>
18. Venables WN, Ripley BD. (2002) *Modern applied statistics with S*. 4th Edition. New York: Springer science+business media, Inc.
19. Wei Y, Pere A, Koenker R, He X. *Quantile regression methods for reference growth charts*. Statistics in Medicine, 2006; **25**:1396–1382