

Comparing Reference Charts for Cross-Section and Longitudinal Data

Thomas H. Scheiwe, Ph.D.
Mei-Jie Zhang, Ph.D.
Anders Juul, M.D.

Technical Report 25

March 1997

Division of Biostatistics
Medical College of Wisconsin
8701 Watertown Plain Road
Milwaukee WI 53226
Phone: (414)456-8280

Comparing Reference Charts for Cross-sectional and Longitudinal Data

Thomas H. Scheike
Department of Biostatistics
University of Copenhagen
Blegdamsvej 3, 2200 KNH. N., Denmark
email : ts@ubism.u.d

Mei-Jie Jiang¹
Division of Biostatistics
Medical College of Wisconsin
email : meijie@hp01.biostat.mcw.edu
and

Anders Juul
Department of Growth and Reproduction
The National University Hospital, Copenhagen
Blegdamsvej 5, 2200 KNH. N., Denmark

¹Research partially supported by National Cancer Institute Grant 1 R01 CA54706-03 and Grant PO1-CA-40053, and by Institutional Research Grant 170H from American Cancer Society.

Comparing Reference Charts for Cross-sectional and Longitudinal Data

Abstract

Reference charts are valuable tools for clinicians in their daily work on pediatric clinics. Reference charts are often constructed by smoothing techniques, and in this paper we present a newly dev

1 Introduction

reference chart is a graph showing the distribution of some measurement of interest and age. This is usually done by displaying the median and various percentiles over the range of ages. When the measurements are approximately normally distributed, perhaps after an appropriate transformation, the median is equivalent to the mean and this is usually used for estimation purposes. Further, when measurements are approximately normal the percentiles can all be expressed as a simple function of the mean and the standard deviation. Therefore it often suffices to estimate a mean function that relates the expected value of the measurement to age, and the variance function that relates the variation of measurements to age. The mean curve and the variance function may be estimated when some assumptions are made, usually one assumes that they are smooth curves, and then uses a smoothing technique to estimate the curves. Quite often smoothness is a reasonable assumption that can be justified based on biological reasoning. Even with smoothness, however, many difficulties are still present, and these are not the issue of this paper. Cole & Green¹ reviews methodological issues of construction for reference charts.

It is important to distinguish between reference charts used for cross-sectional purposes, i.e., deciding whether or not a given measurement at a given age is normal compared to the distribution in a comparable population, or longitudinal purposes, i.e., deciding whether or not the growth of a child is normal based on repeated measurements. When evaluating the development of the measurement of interest the techniques used should reflect the longitudinal aspect. The typical clinical situation is the following: a child returns for measurements at the pediatric clinic, or shows up with a record of earlier measurements, now, based on the current and the

2 A Log-Rank Test for Comparing Regression Functions

In this section we present a longitudinal regression model, for independent identically distributed subjects, that models the current measurement given the time it was measured, the previous measurements and the times of these, see Scheike⁴ or Scheike Zhang³ for further details. This is expressed through the conditional regression model

$$Y_{i,j} = m(V_{\tau_{i,j}}^i) + \epsilon_{i,j}, \quad \text{for } j = 1, \dots, N_i, \quad i = 1, \dots, n. \quad 1)$$

Think of $Y_{i,j}$ as the j^{th} measurement of the i^{th} subject at time $\tau_{i,j}$. We assume that $m(\cdot)$ is a smooth function and that V_s^i is an observable process that only depends on past observations. Note that the regression

where $\hat{H}(y)$ is introduced to avoid edge effects of the kernel estimators. Using local-linear smoothers, or smoothers without edge-problems, this issue can be ignored for applications.

Scheike & Zhang³ studied the asymptotic distribution of $T(z)$ and showed that if $n_j/(n_1 + n_2) \rightarrow p_j$, for $j = 1, 2$, and under sufficient smoothness and other weak regularity conditions, it follows that $\sqrt{n_1 + n_2}T(z)$ converge towards a Gaussian martingale with mean zero (under the hypothesis) and variance function

$$H(y) = p_1^{-1} \int_a^z \frac{\sigma_1^2(y)}{\alpha_1(y)} + p_2^{-1} \int_a^z \frac{\sigma_2^2(y)}{\alpha_2(y)},$$

that can be estimated consistently by

$$\widehat{H}(y) = \left(\frac{n_1}{n_1 + n_2}\right)^{-1} \int_a^z \frac{\widehat{\sigma}_1^2(y)}{\widehat{\alpha}_1(y)} dy + \left(\frac{n_2}{n_1 + n_2}\right)^{-1} \int_a^z \frac{\widehat{\sigma}_2^2(y)}{\widehat{\alpha}_2(y)} dy$$

One consequence of the Proposition is that

$$\sqrt{n_1 + n_2}T(z) \approx N(0, H(y)),$$

i.e., $T(z)$ is approximately normally distributed with a variance we can estimate.

We now define the log-rank (two-sample) test-statistic of the hypothesis $H_0: m_1(\cdot) = m_2(\cdot)$ on the interval $[a, S - a]$ as

$$LR = \sqrt{n_1 + n_2}T(S - a) / \sqrt{\widehat{H}(S - a)}$$

where S is the upper limit of comparison. The two sample log-rank test, LR , have an asymptotically standard normal distribution under the null hypothesis of $m_1(z) = m_2(z)$ on $[a, S - a]$. The test-statistic works best if $m_1(\cdot) \leq m_2(\cdot)$ or $m_2(\cdot) \leq m_1(\cdot)$.

If this is not the case one may instead consider the maximal deviation test-statistic

$$M \stackrel{\text{def}}{=} \sup_{z \in [a, S - a]} |T(z)|. \tag{4}$$

To work out the log-rank test-statistic we thus need to have estimates of $m_k(\cdot)$, $\sigma_k^2(\cdot)$ and $\alpha_k(\cdot)$, and we therefore propose that these are given when reference charts are presented. The next two sections consider the implementation of the log-rank test-statistic in two practical situations. Section 3 contains an application to cross-sectional data, and Section 4 discusses a longitudinal situation.

3 Comparing Cross-Sectional Growth Data

The average height in the population has been increasing with time - the so called "secular trend". Consequently, construction of reference charts for height must be renewed regularly. The secular change in mean height in a population is the result of a general

and an estimate of the variance of this quantit

$$\hat{H} = \sum_i (1_{i+1} - 1_i) \frac{s_{1,i}}{f_{1,i}} + \sum_i (2_{i+1} - 2_i) \frac{s_{2,i}}{f_{2,i}}.$$

Then the the log-rank test-statistic is computed as $LR = T/\sqrt{\hat{H}}$. Using the summations provided in the table we get $T = 1985.0 - 1952.2 = 32.8$ and $\hat{H} = 9.2 + 0.8 = 10.0$ which results in a log-rank test-statistic of LR=10.3, that is approximatel standard normal under the null-h pothesis, and therefore is equivalent to a p-value less than 0.0001.

The average height of the recent stud is 3.5 cm greater than the Danish standard from ndersen et al.⁵. This equals approximatel 0.5 standard deviation (depending on age) and implies that, with the use of the reference charts based on heights obtained more that 20 ears ago, a smaller fraction of short children in 1992 will be classified as pathologicall short, i.e., with a height that is more that 2 standard deviations smaller than the average height. The difference between the 2 studies must be ascribed to socio-economic differences and secular changes.

4 Co parin Lon itudinal rowth Data

In this section we wish to compare the growth of patients with two t pes of skeletal d splasia, namel h pochondroplasia (H po) and achondroplasia (cho). Our data were provided b the Department of Growth and Reproduction at the Universit hospital in Copenhagen and consists of longitudinal measurements of height and weight for 36 patients with h pochondroplasia and 42 patients with achondroplasia.

Skeletal d splasias represent more than 200 different clinical t pes of short limbed dwarfism of which chondroplasia and H pochondroplasia are the most common t pes of skeletal d splasias. The sewere dwarfism and d sproportion of the bod in patients with achondroplasia is caused b a point mutation on chromosome 4 in the fibroblast growth factor receptor (FGFR3) gene which can be demonstrated in all patients with achondroplasia. B contrast, patients with h pochondroplasia represent a more heterogenous group; some patients have the same clinical appearance as patients with achondroplasia and similar degree of growth retardation, whereas others have an almost normal clinical phenot pe and growth. The mean standing height was approximatel 3 SD's below the mean for age-matched helath children. Point mutations in the FGFR3 gene have been demonstrated in 50-60 % of patients with h pochondroplasia onl . Whereas, several studies have reported on actual heights in patients with skeletal d splasia, little is known on possible differences in the linear growth pattern in different t pes of skeletal d splasias. We refer to Hertel⁶ for further details on skeletal d splasia.

The focus of this section is on the longitudinal aspect of the data, with the specific aim of deciding whether or not the two t pes of skeletal d splasia results ina3lasia resultfl in)30TDfl[p)-30 2p;er

models

It appears that patients with hypochondroplasia grow faster than patients with achondroplasia, and if we apply our regression log-rank test to the 2-dimensional regression model with the following region of previous height and time since previous measurement : $[50, 120] \times [0.2, 1.9]$ our test statistic can be calculated for a choice of the two dimensional band-widths. Figure 2 shows the difference in the cumulative regression functions (T, z) for $b_1 = 5.0$ and $b_2 = 0.2$. For this choice of bandwidths we get a test-statistic evaluated in the endpoint $(T = 120, 1.9)$ on 94.1 with variance 586, and this results in a $LR = 3.9$ test-statistic which is approximately standard normal thus resulting in a p-value of approximately 0.0001. Further smoothing of the regression functions results in the same conclusion although the test-statistic decreases some. Note that one would expect the test-statistic to have good power in this application since the hypochondroplasia diagnosis appears to result in a consistently better growth than the achondroplasia diagnosis. A similar comparison of the increase

ere

fl in)Tjfl1126

and the methodology is therefore limited to low dimensions. In an example we compared the longitudinal growth for two different diagnosis of skeletal dysplasia where an ideal 3-dimensional regression function was reduced to a 2-dimensional regression function that provided an adequate description of the data.

6 Appendix: Formulas for estimators

In this appendix we provide formulas for estimators of the quantities that are used in the log-rank test statistic.

We provide a Nadaraya-Watson (ND) type estimator of the regression functions and the variance function as well as an estimator of $\alpha(y)$.

Let $K(\cdot)$ be a kernel function with support on $[-1; 1]$, $\int_{-1}^1 K(u) du = 1$ and $\int_{-1}^1 u K(u) du = 0$, and let $b = (b_1, \dots, b_d)$ be a d -dimensional bandwidth, $|b| = b_1 \cdot \dots \cdot b_d$, $b_i \in]0; \infty[^d$. Define further $C_K \stackrel{\text{de}}{=} \int_{-1}^1 K^2(u) du$, $d_K \stackrel{\text{de}}{=} \int_{-1}^1 u^2 K(u) du$ and $e_K \stackrel{\text{de}}{=} \int_{-1}^1 K(u) du$. We assume that e_K is 0 to obtain an asymptotically unbiased result for our estimator. We abuse notation by letting K denote a d -dimensional kernel as well as a one-dimensional through the product kernel, i.e., $K(y, b) \stackrel{\text{de}}{=} K\left(\frac{y_1}{b_1}, \dots, \frac{y_d}{b_d}\right) \stackrel{\text{de}}{=} \prod_{i=1}^d K\left(\frac{y_i}{b_i}\right)$.

Now, the ND estimator $\hat{m}_{ND}(y) = \frac{1}{n} \sum_{i=1}^n K\left(\frac{y - y_i}{b}\right) y_i$ is an estimator of $m(y)$.

References

- [1] T. J. Cole and P. J. Green. Smoothing reference centile curves: The L - S method and penalized likelihood. *StatMed*

Figure 1. Estimated mean (thin line) and data points from Copenhagen boys 1988-1992 (dots), and estimated mean and 95 % reference area for Danish standard reference from 1971 (thick lines).

Table 1. Estimate of density multiplied by sample size, estimate of mean, estimate of variance, and log-rank test for difference between mean curves of Copenhagen boys and Danish standard reference.

Figure 2. Normalised difference in cumulative regression functions for H₀ (rho = 0) for bandwidth h = 5, 0.2). The log-rank test results in p-value at 0.0001.

