Comparing Reference Charts for Cross-Section and Longitudinal Data

Thomas H. Scheike, 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 Plank 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@kubism.ku.dk

Mei-Jie Zhang [1]
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

[1]Research 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 developed non-parametric test for comparing these charts.

We illustrate the method by two examples. The first example compares cross-sectional data on height in children from two Danish studies from 1970 and 1990, respectively. A second example shows how longitudinal data on growth for two types of skeletal dysplasia may be compared. In the cross-sectional setting the test compares the average height over the range of ages for the two groups, and in the longitudinal setting the test compares, similarly, the average height given a particular history of development.

1 Introduction

A 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 & Green[1] 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 new measurement and the history of earlier measurements the pediatrician wish to assess if the child is growing normally. It may be very misleading to consider the change of percentiles on the cross-sectional growth-chart, see Cole[2]. Therefore, rather than considering the cross-sectional growth chart other methods should be used. One such approach is to use the previous measurements to predict the current new measurement, and then construct a reference chart based on the values of the previous measurements. We return to this issue in Section 4 where we compare the growth of two types of dwarfism based on longitudinal data.

The main objective of this work is to present a new non-parametric test for comparing reference charts for both cross-sectional and longitudinal data, see Scheike & Zhang[3].

The paper is organised as follows : Section 2 presents some methodological issues and the test-statistic used for the comparison; Section 3 contains an adaptation of the methodology to the cross sectional situation, and also presents an application to comparing standard reference charts of height versus age constructed from a study from 1970 and a recent 1990 study, respectively; Section 4 presents an example of longitudinal growth data on two types of skeletal dysplasia.

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 Scheike[4] or Scheike & Zhang[3] for further details. This is expressed through the conditional regression model

$$Y_{i,j} = m(V^i_{\tau_{i,j}}) + \epsilon_{i,j}, \qquad \text{for } j = 1, ..., N_i, \; i = 1, ..., n. \tag{1}$$

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^i_s is an observable process that only depends on past observations. Note that the regression function is equal to the conditional mean of the current measurement given the time of the measurement and all past information, i.e.,

$$E(Y_{i,j}|\tau_{i,j}, \; (Y_{i,k}, \tau_{i,k}), \; k = 1, ..., j-1) = m(V^i_{\tau_{i,j}}), \tag{2}$$

Assuming that the conditional variance of the noise terms, $\epsilon_{i,j}$, is also a deterministic, continuous and bounded function of the observable process V^i_s , we have that

$$\begin{aligned} E(\epsilon_{i,j}|\tau_{i,j} = s, \; (Y_{i,k}, \tau_{i,k}), \; k = 1, ..., j-1) &= 0, \\ E({\epsilon_{i,j}}^2|\tau_{i,j} = s, \; (Y_{i,k}, \tau_{i,k}), \; k = 1, ..., j-1) &= \sigma^2(V^i_s). \end{aligned}$$

Apart from the regression relationship we also assume that the measurement times occur continuously in time, such that a non-parametric estimation of the regression function is possible. Section 3 contains an example of a cross-sectional regression model, where there is only one measurement per subject and V^i_s=s. Section 4 deals with a situation where we believe (approximately of course) that the current measurement can be predicted from the history of that subject based only on the current age, the previous measurement and the age at which it was taken.

The regression function may be estimated by standard smoothing techniques applied to the measurements and the predictors (V^i_s), and the variance function is obtained similarly by smoothing of the squared residuals.

The aim of this work is to establish a non-parametric test to compare the regression functions for two independent groups of subjects. Below, we provide a description of the asymptotic distribution of a log-rank test to evaluate this hypothesis.

We denote the number of subjects in the two groups as n_1 and n_2, the two regression functions as $m_1(\cdot)$ and $m_2(\cdot)$, the two variance functions as $\sigma^2_1(\cdot)$ and $\sigma^2_2(\cdot)$, and the density functions for the distribution of the regressors ($V^i_{\tau_{i,j}}$) as $\alpha_1(y)$ and $\alpha_2(y)$. All these quantities are estimated by standard smoothing techniques, and the estimators are denoted by a $\widehat{\;}$, such that, e.g., $\widehat{m(\cdot)}$ is a non-parametric estimator of $m(\cdot)$. Computational details of the estimators are given in the appendix.

The non-parametric log-rank test is test based on comparing the cumulative regression functions, we therefore consider the the process, $T(z)$, defined as follows

$$T(z) \stackrel{\text{def}}{=} \int_a^z (\widehat{m}_1(y) - \widehat{m}_2(y))dy, \tag{3}$$

where a 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 & Zhang[3] 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

$$\widehat{H(y)} = (\frac{n_1}{n_1+n_2})^{-1} \int_a^z \frac{\widehat{\sigma_1^2(y)}}{\widehat{\alpha_1(y)}} dy + (\frac{n_2}{n_1+n_2})^{-1} \int_a^z \frac{\widehat{\sigma_2^2(y)}}{\widehat{\alpha_2(y)}} dy$$

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 $H_o : m_1(\cdot) = m_2(\cdot)$ 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 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 section consider the implementation of the log-rank test-statistic in two practical situations. Section 3 contains and 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 increase in height at all ages after birth as well as an advancement of the pubertal growth spurt. However, the latter only affects height during adolescence but may have no or little importance for the secular trend for final adult height. Secular changes in height and growth can be considered as a marker of changing health status (hygienic, nutritional) of a population. Consequently, secular changes may be more pronounced in groups of low social class, malnourished individuals who experience more marked improvements in their general health status compared to the overall population. This implies that one can expect a greater secular increase in the lower limits of the reference chart compared the the mean curve and the upper limit. This emphasizes the existence of several pitfalls when comparing reference charts to describe secular changes if these charts are based on populations that vary in socio-economic and general health aspects.

Eight-hundred and twentyfive male children and adolescents from 8 different grammar and highschools in the Copenhagen area agreed to participate in a study of growth during 1988-1992. The schools are located in an area with a superior socio-economic status compared to the rest of the country. The children were 6 to 20 years of age, and those with chronic diseases and/or on medication were excluded from the analyses. Height was measured with a portable Harpenden Stadiometer by trained pediatricians, from the Department of Growth and Reproduction at the University hospital in Copenhagen, to nearest 0.1 cm. In this work we compare the mean height in the 1988-1992 study with a reference chart on height from 1970-1971, see Andersen et al.[5] . The Andersen study is considered as the Danish standard.

In this section we focus attention on cross-sectional growth data, and provide explicit formulas for the log-rank test introduced in the previous Section. A typical presentation of cross-sectional data, that exhibit approximately normally distributed noise, will be a table listing the average height and variance for ages at, say, yearly intervals. The average heights and standard deviations are estimated by smoothing techniques, see Figure 1 and Table 1. In Figure 1 the thick lines represent the Danish standard reference curves, i.e., median, 2.5 % and 97.5 % percentiles, based on a large study performed in 1970, the points are data from the recent 1990 study of boys from Copenhagen and the thin line is the median (mean) curve based on these data. It appears that there has been an increase in height during the last 20 years.

To compute the LR test statistic for the hypothesis of equal regression functions, we need some definitions for the entries of a table of reference values. Let for subtable j, $a_{j,i}, x_{j,i}, s_{j,i}$ and $f_{j,i}$ denote the age and corresponding average height, variance and number of individuals in age group (density of age multiplied by the size of the group). Note that the density multiplied by the size of the sample is an estimate of the number of observations used for the estimation of the average height for a particular age group. Now, we can compute an approximate log-rank test-statistic for the hypothesis that the two groups have the same mean-regression functions by computing the difference in the cumulative regression functions for the two groups

$$T = \sum_i (a_{1,i+1} - a_{1,i})x_{1,i} - \sum_i (a_{2,i+1} - a_{2,i})x_{2,i}$$

and an estimate of the variance of this quantity

$$\widehat{H} = \sum_i (a_{1,i+1} - a_{1,i})\frac{s_{1,i}}{f_{1,i}} + \sum_i (a_{2,i+1} - a_{2,i})\frac{s_{2,i}}{f_{2,i}}.$$

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

The average height of the recent study is 3.5 cm greater than the Danish standard from Andersen et al.[5]. This equals approximately 0.5 standard deviation (depending on age) and implies that, with the use of the reference charts based on heights obtained more that 20 years ago, a smaller fraction of short children in 1992 will be classified as pathologically 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 Comparing Longitudinal Growth Data

In this section we wish to compare the growth of patients with two types of skeletal dysplasia, namely hypochondroplasia (Hypo) and achondroplasia (Acho). Our data were provided by the Department of Growth and Reproduction at the University hospital in Copenhagen and consists of longitudinal measurements of height and weight for 36 patients with hypochondroplasia and 42 patients with achondroplasia.

Skeletal dysplasias represent more than 200 different clinical types of short limbed dwarfism of which Achondroplasia and Hypochondroplasia are the most common types of skeletal dysplasias. The sewere dwarfism and dysproportion of the body in patients with achondroplasia is caused by a point mutation on chromosome 4 in the fibroblast growth factor receptor (FGFR3) gene which can be demonstrated in all patients with achondroplasia. By contrast, patients with hypochondroplasia 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 phenotype and growth. The mean standing height was approximately 3 SD's below the mean for age-matched helathy children. Point mutations in the FGFR3 gene have been demonstrated in 50-60 % of patients with hypochondroplasia only. Whereas, several studies have reported on actual heights in patients with skeletal dysplasia, little is known on possible differences in the linear growth pattern in different types of skeletal dysplasias. We refer to Hertel[6] for further details on skeletal dysplasia.

The focus of this section is on the longitudinal aspect of the data, with the specific aim of deciding whether or not the two types of skeletal dysplasia results in different growth patterns. In Cole[2] there is a thorough discussion of how to construct conditional reference charts that can be used for monotoring longitudinal growth. Cole focuses on parametric models which may be viewed as a special case of the models presented in the previous section and here.

A conditional reference chart is constructed from a conditional mean function and a conditional variance function, which may be estimated by standard smoothing techniques. We consider a longitudinal regression model where the regression function is a function of the height at the previous measurement, current age and increment of age since the previous measurement, as follows

$$Y_{k,i,j} = m_k(Y_{k,i,j-1}, \tau_{k,i,j}, \tau_{k,i,j} - \tau_{k,i,j-1}) + \epsilon_{k,i,j}, \quad \text{for } j = 2, ..., N_i^k, \ \ i = 1, ..., n_k, k = 1, 2, \tag{5}$$

and

$$E({\epsilon_{i,j}}^2 | \tau_{i,j}, \ (Y_{i,k}, \tau_{i,k}) \ k = 1, ..., j-1) = \sigma^2(Y_{k,i,j-1}, \tau_{k,i,j}, \tau_{k,i,j} - \tau_{k,i,j-1}). \tag{6}$$

This regression and variance function may be estimated for some values of the regressors. A conditional reference chart would display curves for the upper and lower 2.5 % percentile and the median. Thus, if normality of the residuals is approximately true the curves are estimates of the following three curves

$$m(h_0, a_0, s) \pm 2\sigma(h_0, a_0, s) \text{ and } m(h_0, a_0, s),$$

for given initial age (a_0) and initial height (h_0) and for for s varying over some appropriate range. Pursuing the issue non-parametrically this is a formidable task that will need a lot of data. Therefore one would probably settle for a parametric version of the conditional reference charts as in Cole[2] for our data. In principle, however, this may be done and then a log-rank test-statistic may be computed by integrating the estimated regression function over some relevant area of the 3-dimensional regressor space. When the issue is only comparison of the growth patterns things are somewhat easier due to the improved rate of convergence of the cumulative regression functions which are the objects of comparison.

Analysing the data it turned out that a very good description of the data was obtained from considering the simpler (incorrect) sub-model

$$Y_{k,i,j} = m_k(Y_{k,i,j-1}, \tau_{k,i,j} - \tau_{k,i,j-1}) + \epsilon_{k,i,j}, \quad \text{for } j = 2, ..., N_i^k, \ \ i = 1, ..., n_k, k = 1, 2, \tag{7}$$

and

$$E({\epsilon_{i,j}}^2 | \tau_{i,j}, \ (Y_{i,k}, \tau_{i,k}) \ k = 1, ..., j-1) = \sigma^2(Y_{k,i,j-1}, \tau_{k,i,j} - \tau_{k,i,j-1}). \tag{8}$$

The residuals in the model represent further biological variation and measurement error and were expected to be right-skewed and biased for small values of time-increments. Residuals plots, however, revealed that this was not a serious problem for our study. The second regressor of the ideal model (age) did not contribute much additional information in terms of predicting the growth of hypo or acho patients. Therefore, when the objective is solely to compare the growth patterns the above simpler sub-model may be used. When a difference is found it can be concluded that the longitudinal growth are different for the two diagnosis. In contrast, however, equivalent behaviour for two groups based on the limited model will only make this conclusion valid for the observed age-span.

It appears that patients with hypochondroplasia grow faster than patients with achondroplasia, and if we apply our regression log-rank test to the 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 Hypo diagnosis appears to result in a consistently better growth than the Acho diagnosis. A similar comparison of the increase of weight did not result in a significant difference between the two groups, with a p-value at 0.48 for the bandwidths chosen as for height.

Despite the recents capability of locating mutations in specific genes resulting in specific phenotypes, accurate diagnosis in children with skeletal dysplasia is often difficult. In this context differences in the linear growth pattern as well as in body proportions (Hertel[6]) between these two types of skeletal dysplasia may improve correct diagnosis in these growth-retarded children and may enhance our knowledge on the different pathogenetic background for these types of dwarfism.

Due to the rather limited data available we have been forced to ask rather general questions, but in principle, when the conditional regression and variance functions are estimated one may proceed to compare these for various areas of the regressor space and thereby obtain a more detailed analysis of where the differences originate from.

5 Discussion

In this work we have presented a new non-parametric test for comparison of the mean-curve of two reference charts. After an appropriate transformation the mean curves may often be interpreted as median-curves. The methodology is applicable for both standard reference charts of, say, height versus age, and conditional reference charts. Reference charts are often constructed from an assumption of smoothness and the present non-parametric test ties naturally into this framework where parametric assumption are unwanted. The test focuses on the mean curve of the reference chart although the variance function may also differ and contain important information about the distribution of the measurement of interest. This may be particularly relevant when studying secular changes that results in a more pronounced change for individuals in either tail of the distribution. The secular change of body-mass index, e.g., is believed to result in a distribution with the same median but with increasing skewness.

The proposed test is good at detecting differences where the same pattern is consistent across ages. When this is not expected to be the case alternative tests should be used.

The proposed methodology breaks down when the smoothing procedure breaks down 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 $[-1;1]$, $\int K(u)du = 1$ and $\int uK(u)du = 0$, and let $b = (b_1, ..., b_d)$ be a d-dimensional bandwidth, $|b| = b_1 \cdot ... \cdot b_d$, $b \in]0;\infty[^d$. Define further $C_K \stackrel{\text{def}}{=} \int K^2(u)du$, $d_K \stackrel{\text{def}}{=} \int u^2 K(u)du$ and $e_K \stackrel{\text{def}}{=} \int uK(u)du$. 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., $K(y,b) \stackrel{\text{def}}{=} K(\frac{y_1}{b_1}, ..., \frac{y_d}{b_d}) \stackrel{\text{def}}{=} \prod_{i=1}^d K(\frac{y_i}{b_i})$.

Now, the ND estimator, $\widehat{m(y)}$, of the regression function $m(y)$ is defined by

$$\widehat{m}(y) \quad \stackrel{\text{def}}{=} \quad \frac{\widehat{r}(y)}{\widehat{\alpha}(y)}$$

where

$$\widehat{r}(y) \quad \stackrel{\text{def}}{=} \quad \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i} Y_{i,j}\,\tfrac{1}{|b|}K(y - V^i_{\tau_{i,j}}, b),$$

and

$$\widehat{\alpha}(y) \quad \stackrel{\text{def}}{=} \quad \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i} \tfrac{1}{|b|}K(y - V^i_{\tau_{i,j}}, b).$$

$\widehat{\alpha}(y)$ is an estimator of $\alpha(y)$.

The variance functions, $\sigma^2(\cdot)$, can be estimated by the following squared-residual kernel estimator

$$\widehat{V}(y) = \frac{V(y)}{\widehat{\alpha}(y)} - (\widehat{m}(y))^2 \tag{9}$$

where

$$V(y) \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_t^i}(Y_{i,j})^2\,\tfrac{1}{|b|}K(y - V^i_{\tau_{k,i,j}}, b) \tag{10}$$

References

[1] T. J. Cole and P. J. Green. Smoothing reference centile curves: The LMS method and penalized likelihood. *StatMed*, 11:1305–1319, 1992.

[2] T. J. Cole. Growth charts for both cross-sectional and longitudinal data. *StatMed*, 13:2477–2492, 1994.

[3] T. H. Scheike and M. Zhang. Log-rank and maximal deviation tests for regression models for longitudinal data. *Submitted*, 1996.

[4] T. H. Scheike. Parametric regression for longitudinal data with counting process measurement times. *Scandinavian Journal of Statistics*, 21:245–263, 1994.

[5] E. Andersen, B. Hutchings, J. Jansen, and M. Nyholm. Højde og vægt hos danske børn. *Ugesrift for Læger, In Danish*, 1974.

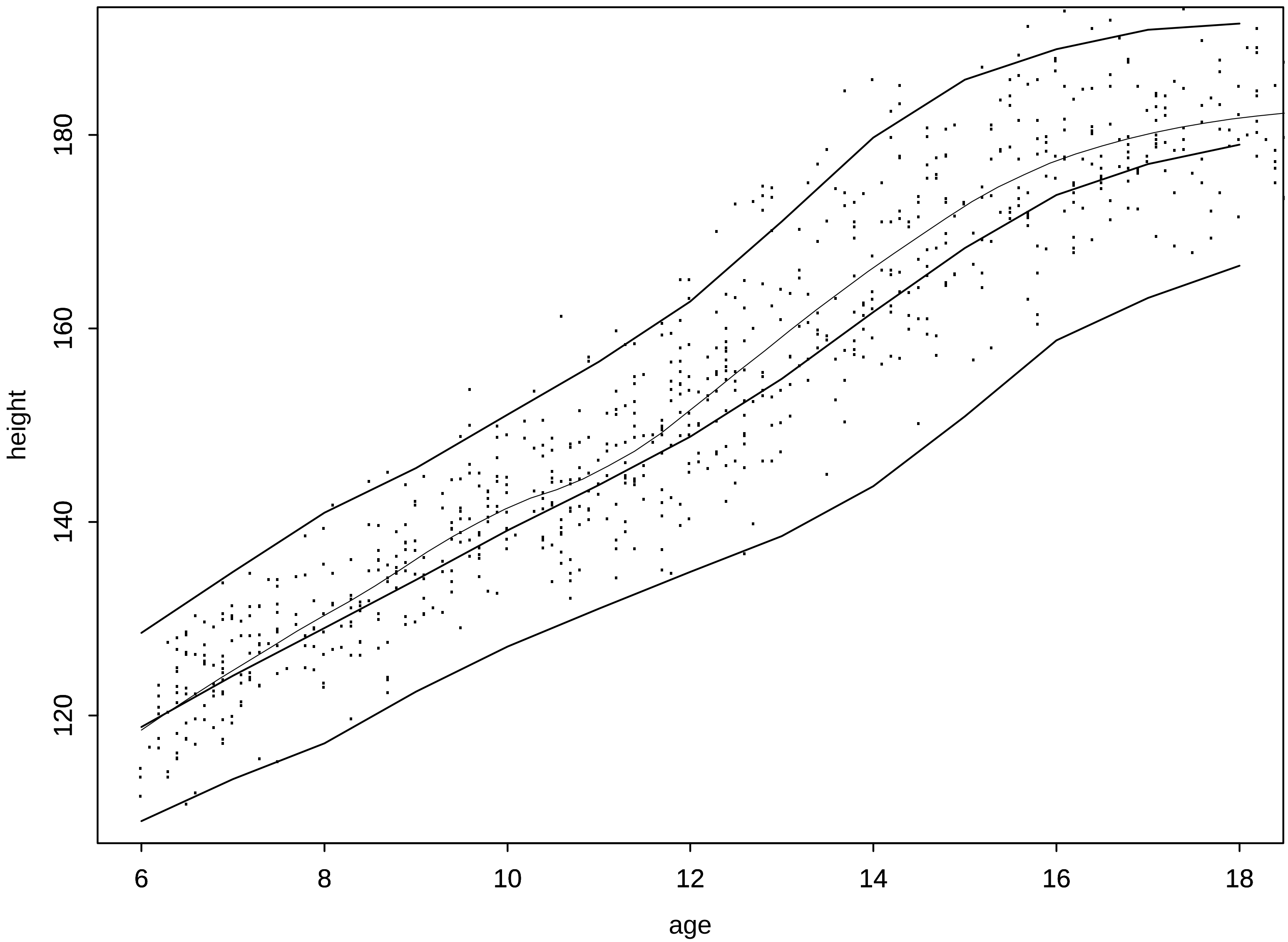
[6] N. T. Hertel. *Linear Growth, Body Proportions and Biochemical Markers of Growth in Healthy Children and Individuals with Three Types of Skeletal Dysplasia.* PhD thesis, Department of Growth and Reproduction, University of Copenhagen, 1996.

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).

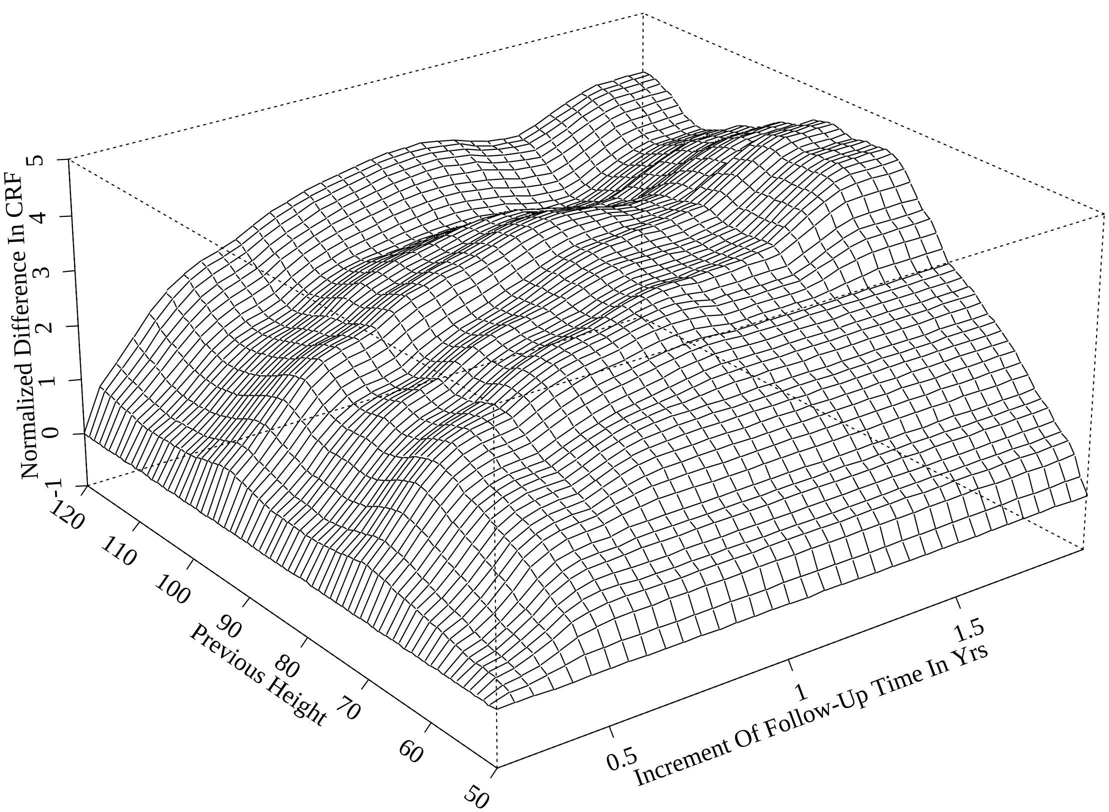
Table 1. Estimate of density multiplied by sample size, estimate of mean, estimate of variance, and log-rank test for difference between mean curves of Copenhagen boys and Danish standard reference.

Figure 2. Normalised difference in cumulative regression functions for Hypo-Acho for bandwidth (5,0.2). The log-rank test results in p-value at 0.0001.

	Copenhagen Boys 1990 (825)				Danish Reference 1971 (10925)			
age	# boys	Mean	Variance (σ^2)	σ^2 / # boys	# boys	Mean	Variance (σ^2)	σ^2 / # boys
6	37.02	118.49	21.18	0.57	269	118.8	23.72	0.09
7	61.08	124.63	21.59	0.35	402	124.1	28.73	0.07
8	64.35	130.31	22.48	0.35	932	129	35.52	0.04
9	67.64	136.14	24.08	0.36	1175	134	33.29	0.03
10	72.53	141.44	27.74	0.38	1124	139.1	36	0.03
11	79.77	145.27	35.47	0.44	1152	143.8	40.83	0.04
12	83.37	151.58	53.98	0.65	967	148.8	48.72	0.05
13	74.98	159.12	67.93	0.91	1103	154.8	66.1	0.06
14	68.56	166.25	71.85	1.05	1127	161.7	81.18	0.07
15	66.79	172.65	63.72	0.95	1017	168.3	75.69	0.07
16	63.43	177.32	50.21	0.79	786	173.8	56.7	0.07
17	49.67	180.1	42.72	0.86	536	177	47.89	0.09
18	23.89	181.73	37.55	1.57	345	179	39.19	0.11
sum	813.07	1985.05	-	9.23	10395	1952.20	-	0.82
	Log-rank test-statistic from Table LR=-10.3,p $<$ 0.000							



height
180
160
140
120
6
8
10
12
14
16
18
age



Normalized Difference In CRF
-1
0
1
2
3
4
5
Previous Height
120
110
100
90
80
70
60
50
Increment Of Follow-Up Time In Yrs
0.5
1
1.5

Figure 2