International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 4, 2005, pp 79-98

Growth Comparison among Philippine Children: A Regression Approach

Md. Ayub Ali, Kumi Ashizawa, Chiyoko Kumakura and Ayano Kusumoto

Laboratory of Growth and Ergonomics,
Otsuma Women's University, Sanban-cho 12, Chiyodaku, Tokyo 102-8357, Japan

[Received February 24, 2004; Revised March 16, 2005; Accepted July 17, 2005]

Abstract

Stature, body weight, and the body mass index (BMI) were quantified in cross-sectional sample of 3154 Philippine children aged 7-17 years. This sample was consisted of 1564 boys and 1590 girls in which 396 boys and 350 girls from three villages of Isabela province, 475 boys and 483 girls from Makati City, 151 boys and 235 girls from Neguilian of Isabela province, and 542 boys and 520 girls from Quezon City. Mahalonobis Distance and Cook's Distance were used to remove the outliers and influential data points. Methods of regression analysis, MANOVA and LSD test under post-hoc comparison of means were used to quantify the growth of the children. Several predicted equations for different regions were proposed to quantify the growth of the Philippine children.

Keywords and Phrases: Quantification, Growth, Philippine children, Regression equation, Regional and Sex differences.

AMS Classification: 62F10, 62H12, 62H15, 62J05, 62J07, 62J10, 62J15, 62P10, 62P25.

1 Introduction

The meaning of growth refers to the increase in the size of the body as a whole or in the size of specific parts or segments of the body from the conception of maturity. The mean differences in size among individuals within the same population appear due to the effects of the environmental factors, e.g., diet, family size, housing, fatigue, privation, or psychosocial stress (Eveleth and Tanner, 1990). The growth of children and adult physical status is influenced by these factors, however it is impossible to isolate these factors one from the other. Children who live in urban areas are usually taller and heavier than those who live in rural areas. This seems to be the result of better living conditions, e.g., better sanitation, food availability, access to adequate health care, education, etc (Eveleth and Tanner, 1976). Conversely, some researchers have found small urban-rural differences in growth compared to differences among social strata (Martorell and Habicht, 1985; Singh et al., 1987; Lopez Blanco et al., 1987, 1992). Also, sexual differences are not similar in different populations (Bogin et al., 1990).

The growth of the children is the mirror of the society in which they live. Many studies on this issue have been reported for different populations (Ashizawa et al., 1987, 1993, 1997, 1998; Cameron et al., 1992; Dater, 1993; Eveleth and Tanner, 1976, 1990; Lopez Blanco et al., 1992; Malina et al., 1985; Martorell and Habicht, 1985; Mueller and Titcomb, 1977; Singh et al., 1987). Within-population variability in quantitative traits has largely been ignored in human population biology. Recently, Ashizawa et al. (1998) studied the growth of Philippine children, however there is scope to quantify that growth statistically. Thus, the purpose of this present study is to quantify the growth, and the growth differences with respect to different regions and sex of the Philippine children.

2 Data and Methods

Data. A cross-sectional raw data of stature and weight of 3154 Philippine children aged 7-17 years was considered from the data set of Ashizawa et al. (1998). The sample size was not consistent with that of Ashizawa et al. (1998) due to omitting the outliers and influential data points from the whole sample. This sample was consisted of 1564 boys and 1590 girls in which 396 boys and 350 girls from three villages of Isabela Province (Isabela group), 475 boys and 483 girls from Makati, 151 boys and 235 girls from Neguilian of Isabela province (Neguilian group), and 542 boys and 520 girls from Quezon City (QC). The father's occupations were variable with different regions (Fig. 1). In Isabela, 87% of the parents were tenant farmers cultivating paddy, peanuts, maize, and so on. The number of children in a family was sometimes more than 8, and 50% of them do not survive beyond age 5. About 55% parents are farmer in Neguilian of Isabela province whereas only 2% parents are farmer in Makati city and almost no parent is farmer in QC. In QC, about 53% of the parents are officer, 27% are professionals, and 7% are labor.

In Makati, about 38% parents are labor and 16% are officer. Comparing among these four study areas, QC is rich area in Metro Manila, Makati is below rich but not poor, Neguilian is almost same or little poor than Makati, and the village Isabela is almost poor. Everyday, the subjects in QC were taking breakfast in the morning, tiffin (with snacks or sweets, and cocacola) at about 10:00 am, lunch at about 12:00 am, tiffin again at about 3:00 pm, and dinner at night. This pattern was not similar in other regions.

Linguistically, in Isabela, 97% subject were Ilocano speakers, while in Naguilian, 63% were Ilocano and 34% were Ibanag speakers. In Metro Manila, Tagalog-speaking subjects were 76% in Makati and 91% in QC. This implies that the migration rate from outside the capital was higher into Makati than into QC. For more details, see Ashizawa et el., 1998.

Method. Stature and weight were, measured by Ashizawa and her colleagues (1998), used in the present study. The Body Mass Index (BMI) was calculated by the equation

$$BMI = weight\ in\ kg/(stature\ in\ meter)^2$$

Methods of regression analysis (Draper and Smith, 1966; Montgomery and Peck, 1982), was used to quantify the growth of the children. Sometimes it is necessary to specify a regression equation with an intercept or without an intercept (intercept forced to zero, regression through the origin). Regression without the intercept is often used in analyses of economic data in cases when, by definition, the regression line describing the relationship between some variables would be predicted to have a zero intercept.

For example, if one were to correlate tax revenues with gross national product (GNP) then it is obvious that, if there is zero GNP, there is zero tax revenue. However, in the majority of applications (in particular in the social and natural sciences) variables of interest are measured on more or less arbitrary scales where the zero points have no special meaning. For example, if we consider BMI is a function of body weight, or stature, the intercept term is meaningless there because one can not think BMI when any of those measurements is zero. In such a situation, inclusion of intercept term result low value of R^2. On the other hand, if we consider stature is a function of age for example, the intercept term is then essential because the stature at age zero (at the time of birth) is not zero. If one is interested to calculate his height, weight and BMI provided that he has only the information of age, nothing else, he may use the following regression equations.

Set 1:
$$Model1 : Ht = \alpha + \beta Age + \varepsilon;$$
$$Model2 : Wt = \alpha + \beta Age + \varepsilon;$$

Set 2:
$$Model3 : BMI = \beta Wt + \varepsilon;$$
$$Model4 : BMI = \beta Ht + \varepsilon;$$

where ε is random error assumed that $\varepsilon \sim NID(0, \sigma^2)$.

MANOVA test is considered to test the overall differences among the variables whether they are significantly different or not, and LSD test under post-hoc comparison of means is considered to test the significance of their pairwise differences. The test of normality, Shapiro-Wilk's W test extended by Royston (1982), was employed to test the normality of the variables (using the software STATISTICA) which allows it to be applied to large samples (with up to 2000 observations). If the W statistic is significant, then the hypothesis that the respective distribution is normal should be rejected. The Shapiro-Wilk's W test is the preferred test of normality because of its good power properties as compared to a wide range of alternative tests (Shapiro, Wilk, & Chen, 1968).

To know how well the regression equations will predict on independent samples of the population individuals, cross-validated correlation, a model validation technique, is considered (Stevens, 1996; p. 96). The cross validity predictive power, denoted by ρ^2_{cv} , is defined as:

$$\rho^2_{cv} = 1 - \frac{(n-1)(n-2)(n+1)}{n(n-k-1)(n-k-2)}(1-R^2);$$

where n is the sample size, k is the number of predictors in the regression equation and the cross-validated R is the correlation between observed and predicted values of the dependent variable. Using the above statistic, it can be concluded that if the prediction equation is applied to many other samples from the same population, then $(\rho^2_{cv} \times 100)\%$ of the variance on the predicted variable would be explained by the regression equation (Stevens, 1996; p. 100).

3 Results

Age-group-wise means, standard deviations were almost the same as those of Ashizawa et al. (1998). The test of non-normality, using Shapiro-Wilk's W test extended by Royston (1982), of BMI, height and weight were shown in Table 1. This table indicates that the BMI, height and weight are almost normally distributed except the BMI for boys in QC due to their non-normality of the weight. The blank spaces in the table imply the non-availability of the measurement data.

MANOVA test indicates that overall differences of height, weight and BMI among different age groups, regions and sexes. Least significance difference (LSD) test was adopted to test significance of the pairwise regional differences in height, weight, and BMI at different age groups. These results were shown respectively in Table 2, Table 3, and Table 4. These tables divulged that the people of QC were different in growth of height, weight, and as well as BMI from that of other 3 regions up to age 17 years. This difference in BMI was continued after age 17.

Age-group-wise sex differences in height, weight and BMI of Philippine children for different regions were investigated and shown in Table 5. This table shows that the sex difference for BMI in QC was insignificant except the age 8. In Neguilian, this difference in height was significant from age 14. But, an opposite picture was took place in BMI. In Isabela, the sex differences were insignificant except age 12 for height and weight, and age 12 & 13 for BMI.

As the growth differences due to region and sex were observed, different regression equations with different regions and sex were necessary to quantify the growth.

Table 6 exhibits the regression coefficients for the Philippine boys. These coefficients were highly significant ($p < 0.0001$) with a smaller amount of standard errors of the estimate. This table also exhibits that R^2 was sufficiently large for the models of BMI on height, BMI on weight for all the regions.

Table 7 exhibits that the regression coefficients for the Philippine girls. These coefficients were also highly significant ($p < 0.0001$) with a smaller amount of standard errors of the estimate. The sufficiently large R^2 were found for the models of BMI on height, BMI on weight for all the regions.

Analyses of residual for predicted equations were considered to understand the preciseness of the prediction. Mean, standard deviation, standard error (SE), and 95% confidence limits of the estimated residuals were shown in Table 8. This table shows that the mean residuals in all models of both boys and girls for different regions are very small, and small amount of standard error of the estimates imply the preciseness of the predicted equation, however, a little large standard deviations for Model 1 and Model 2 of the both sexes indicate scatteredness of the observed height and weight with respect to their chronological ages.

4 Discussions

Ashizawa et al. (1998) applied the test of non-normality (on BMI only) using the methods Kolmogorov-Smirnov and Shapiro-Wilk's, whereas in the present study, Shapiro-Wilk's W test extended by Royston (1982) was considered. It is note that this test is the preferred test of non-normality because of its good power properties. For this, the present study sometimes showed different results with that of Ashizawa et al. (1998) due to the appropriateness of the method. Non-normality test of the height and weight were incorporated and found that the results were not consistent with that of BMI.

Significant regional differences among the growth of BMI, height and weight implies that the growth of the children differ with the deviation of socio-economic conditions of the family. Age-wise regional differences in height growth were found among the four population groups except the population pair Makati and Neguilian for boys and girls (except age 13), and Makati and Isabela for age groups 7 to 10 years for girls and

7 years for boys (Table 2). In weight, QC is significantly larger from the other regions up to age 17 for boys and age 16 for girls. This is due to taking heavy intake and in accord with their social status.

For Philippine boys, Table 6 implies that the growth intensities of height and weight were higher in QC followed by Makati and then Isabela. The Neguilian group was not comparable due to the unavailability of the data up to age 12. This is also similar with their social status. For Philippine girls, the growth intensities of weight were higher in QC followed by Makati and then Isabela (Table 7). The Neguilian group also was not comparable due to the unavailability of the data up to age 12.

By **ANCOVA** (Analysis of Covariance, the results are not shown), it is found that the pairwise regional differences on BMI appears due to the significant differences in height as well as weight. But, opposite picture is found only for Neguilian and QC. In this two Provinces, the BMI difference is significant only due to significant different in weight, not in height. This may partially due to taking heavy intact.

The mean residuals for all 4 models of different regions were very small for both boys and girls implying that the estimated lines passed the observed data through such points by which the error sum of squares were minimized, however, a little larger standard deviations were observed in Model 3 and Model 4 those in accord with the results of Ashizawa et al. (1998). The standard error of the estimated mean residuals were also smaller for both the sexes with different regions and models. The present study exhibits that the 95% confidence bounds for residuals (Table 8) show good estimations of the regression models. The lengths of the 95% confidence intervals for mean residual were also comparatively more condensed than those of Ali and Ohtsuki (2001).

Using the information of age only, one can calculate his height, weight and BMI from the proposed regression equations.

These estimated equations were

$$
\begin{array}{rcl}
\textbf{First Set} & : & \\
Boy(Isabela) & : & Ht = 79.95409 + 4.42095Age \\
(Makati) & : & Ht = 78.64460 + 4.94855Age \\
(Negulian) & : & Ht = 100.6982 + 3.56450Age \\
(QC) & : & Ht = 84.18365 + 5.12612Age \\
Girl(Isabela) & : & Ht = 74.77612 + 5.05225Age \\
(Makati) & : & Ht = 88.32894 + 3.92114Age \\
(Negulian) & : & Ht = 140.4214 + 0.60350Age \\
(QC) & : & Ht = 95.29539 + 3.91937Age \\
Boy(Isabela) & : & Wt = 1.691118 + 2.128151Age \\
(Makati) & : & Wt = -4.15351 + 2.95793Age \\
(Negulian) & : & Wt = -8.41417 + 3.39012Age \\
(QC) & : & Wt = -1.32015 + 3.51909Age \\
Girl(Isabela) & : & Wt = -2.60921 + 2.62416Age \\
(Makati) & : & Wt = -1.79882 + 2.72632Age \\
(Negulian) & : & Wt = 21.50038 + 1.31074Age \\
(QC) & : & Wt = 3.409541 + 2.881255Age
\end{array}
$$

$$
\begin{array}{rcl}
\textbf{Second Set} & : & \\
Boy(Isabela) & : & BMI = 0.612214Wt \\
(Makati) & : & BMI = 0.474361Wt \\
(Negulian) & : & BMI = 0.403653Wt \\
(QC) & : & BMI = 0.428830Wt \\
Girl(Isabela) & : & BMI = 0.608278Wt \\
(Makati) & : & BMI = 0.500959Wt \\
(Negulian) & : & BMI = 0.443115Wt \\
(QC) & : & BMI = 0.462948Wt \\
Boy(Isabela) & : & BMI = 0.117737Ht \\
(Makati) & : & BMI = 0.114681Ht \\
(Negulian) & : & BMI = 0.114785Ht \\
(QC) & : & BMI = 0.127641Ht \\
Girl(Isabela) & : & BMI = 0.118007Ht \\
(Makati) & : & BMI = 0.120245Ht \\
(Negulian) & : & BMI = 0.123479Ht \\
(QC) & : & BMI = 0.128460Ht
\end{array}
$$

From the first set, we can predict height or weight for any age. Using the predicted values of these height or weight we can further predict the BMI from the Second set, need not both height and weight. The data of QC boys was affected with little non-normality. But, one can use the result of the QC boys with caution as the sample size was large. After age 10 in boys and 13 in girls, height of Neguilian children overtake the height of QC children. This may in accord with their social and economical conditions. Comparing the slopes of the predicted equations, it is found that the growth rate in height and weight of boys is higher in the region QC followed by Makati and Isabela. These rates for girls are consistent with those in weight but not in height. In height, it is higher in the region Isabela followed by QC and then Makati. This may perhaps due to not participating in physical exercise/work of the female of the Isabela region, or to take food like rice, sweet, etc. The growth rate in BMI for both sexes in the Isabela region is higher comparing with other regions. This may due to the lack of awareness of the physical fitness, because most of the people of the Isabela region are poor farmer or labor. To the best of our knowledge, there is no other reference regarding this approach. So the further discussion is kept limited.

The predicted equations are cross validated by the cross validity predictive power as described in method section. Estimated cross validity predictive power, ρ^2_{cv}, of the predicted equations for different regression lines of Philippine boys and girls are shown in Table 9. This table indicates that for any independent sample of the Philippine population (studied here) more than 99% of the variance on the predicted variable would be explained by the predicted equations. In other words, the expected amounts of shrinkage of R^2 are very small for all regression equations of boys and girls, implying a highly cross validated.

Acknowledgements

We are grateful to Prof. M.A. Basher Mian, Executive Editor, and the anonymous reviewers for their cooperation, comments and suggestions on an earlier version of the manuscript. We are also grateful to the Japan Society for the Promotion of Science (JSPS) for their financial support during this study.

References

[1] Ali, M.A. (2000). Secular trends in growth of the Japanese and prediction of their adult stature. D.Sc. dissertation. Tokyo Metropolitan University, Tokyo.

[2] Ali, M.A. and Ohtsuki, F. (2001). Prediction of adult stature for Japanese population: A stepwise regression approach. Am. J. Hum. Biol. 13: 316-322.

[3] Ashizawa, K., Kusumoto, A. and Kawabe, T. (1987). Height, weight, and chest circumference in a group of Philippine school children compared with the Japanese standard. Journal of Human Ergology, 16: 43-53.

[4] Ashizawa, K., Kusumoto, A., Sugane, A., Dobashi, K., Kawabata, M. and Inomata, E. (1993). Anthrometrical data of urban and rural Japanese children. Anthropological Science, 101: 129-156.

[5] Ashizawa, K., Kumakura, C., Kusumoto, A. and Narasaki, S. (1997). Relative foot size and shape to general body size in Javanese, Filipinas and Japanese with special reference to habitual footwear types. Ann. Hum. Biol., 24: 117-129.

[6] Ashizawa, K., Kumakura, C. and Kusumoto, A. (1998). Growth of Philippine children in reference to socioeconomic environment. Anthropological Science. 106: 77-94.

[7] Bock, R.D., du Toit, S.H.C. and Thissen, D. (1994). AUXAL: Auxological analysis of longitudinal measurements of human stature. Chicago: SSI.

[8] Bogin, B., Wall, M. and MacVean, R.B. (1990). Longitudinal growth of high socioeconomic status Guatemalan children analized by the Preece-Baines function: An international comparison. Am. J. Hum. Biol. 2: 271-281.

[9] Bogin, B., Wall, M. and MacVean, R.B. (1992). Longitudinal analysis of adolescent growth of Ladino and Mayan school children in Guatemala: Effects of environment and sex. Am. J. Phys. Anthropol. 89: 447-457.

[10] Cameron, N., Kgamphe, J.S., Leschner, K.F. and Farrant, P.J. (1992). Urban-rural differences in the growth of South African black children. Ann. Hum. Biol. 19: 23-33.

[11] Count, E.W. (1943). Growth patterns of human physique: an approach to kinetic anthropometry. Human Biology, 15:1-32.

[12] Datar, F.A. (1993). Seasonality: Its impact on the growth of Filipino pre-school children in a fishing community. Ph.D. thesis (Faculty of Graduate School of State University of New York at Buffalo).

[13] Draper, N.R. and Smith, H. (1966). Applied Regression Analysis. New York: Wiley & Sons.

[14] Eveleth, P.B. and Tanner, J.M. (1976). Worldwide Variation in Human Growth. Cambridge University Press, Cambridge.

[15] Eveleth, P.B. and Tanner, J.M. (1990). Worldwide Variation in Human Growth. 2nd edition. Cambridge University Press, Cambridge.

[16] Lopez Blanco, M., Landaeta Jimenez, M. and Mendez Castellano, H. (1992). Urban-rural differences in the growth status of Venezuelan children. Am. J. Hum. Biol. 4: 105-113.

[17] Malina, R.M., Little, B.B., Buschang, P.H., DeMoss, J. and Selby, H.A. (1985). Socioeconomic variation in the growth status of children in a subsistence agricultural community. Am. J. Physical Anthropl. 68: 385-391.

[18] Martorell, R. and Habicht, J.P. (1986). Growth in early child-hood in developing countries. In F. Falkner amd J.M. Tanner (eds.): Human Growth, a Comprehensive Treatise, Vol. 3. New York: Plenum Press, pp. 241-262.

[19] Montgomery, D.C. and Peck, E.A. (1982). Introduction to Linear Regression. Analysis. John Wiley & Sons.

[20] Mueller, W.H. and Titcomb, M. (1977). Genetic and environmental determinants of growth of school-aged children in a rural Colombian population. Ann. Hum. Biol. 4: 1-15.

[21] Royston, P. (1982). An extension of Shapiro and Wilk‘s W test for normality to large samples. Applied Statistics, 3: 115-124.

[22] Shapiro, S.S., Wilk, M.B. and Chen, H.J. (1968). A comparative study of various test for normality. Journal of American Statistical Association, 63: 1343-1372.

[23] Singh, S.P., Sidhu, L.S., and Malhotra, P. (1987). Growth performance of Punjabi children aged 6-12 years. Ann. Hum. Biol., 14: 169-179.