Imputing Longitudinal Growth Data in International Pediatric Studies: Does CDC Reference Suffice?

Zhiguo Li, PhD[1], Jorma Toppari, MD, PhD[2], Markus Lundgren, MD, PhD[3], Brigitte I. Frohnert, MD, PhD[4], Peter Achenbach, MD[5], Riitta Veijola, MD, PhD[6], and Vibha Anand, PhD[7] for the T1DI study group

[1]Center for Computational Health IBM Research, NY, NY, [2]Institute of Biomedicine and Population Health Research Centre, University of Turku and Department of Pediatrics, Turku University Hospital, Turku, Finland, [3]Department of Clinical Sciences Malmö, Lund University/CRC, Skåne University Hospital, Malmö, Sweden, [4]Barbara Davis Center for Diabetes, University of Colorado, Denver, CO, USA, [5]Institute of Diabetes Research, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich-Neuherberg, Germany, [6]Department of Pediatrics, PEDEGO Research Unit, University of Oulu and Oulu University Hospital, Oulu, Finland, [7]Center for Computational Health IBM Research, Cambridge, MA

Abstract

This study investigates a missing value imputation approach for longitudinal growth data in pediatric studies from multiple countries. We analyzed a combined cohort from five natural history studies of type 1 diabetes (T1D) in the US and EU with longitudinal growth measurements for 23,201 subjects. We developed a multiple imputation methodology using LMS parameters of CDC reference data. We measured imputation errors on both combined and individual cohorts using mean absolute percentage error (MAPE) and normalized root-mean-square error (NRMSE). Our results show low imputation errors using CDC reference. Overall height imputation errors were lower than for weight. The largest MAPE for weight and height among all age groups was 4.8% and 1.7%, respectively. When comparing performance between CDC reference and country-specific growth charts, we found no significant differences for height (CDC vs. German: p =0.993, CDC vs. Swedish: p=0.368) and for weight (CDC vs. Swedish: p=0.513) for all ages.

Introduction

The problem of missing data commonly exists in observational studies.[1] We explored imputation methodology to address missing values in childhood growth data from prospective natural history studies of type 1 diabetes (T1D), including two in the US (DAISY[2], DEW-IT[3]) and three in the EU (BABYDIAB[4], DiPiS[5], DIPP[6]). The primary aim of these studies was to study development of islet autoimmunity (IA) and progression to type 1 diabetes (T1D) in infants and young children with high genetic or familial risk. Study subjects were followed for development of IA for a period of up to 15 years or until diagnosis of T1D, whichever came first. Data collection efforts mainly focused on immunological (islet autoantibodies) and metabolic markers (blood glucose levels) that are implicated in development of T1D.

In a collaboration with JDRF and their academic partners, the type 1 data intelligence (T1DI) study group combined and harmonized data from these large observational studies into the T1DI cohort. The cohort has over 24,000 subjects and the study group is currently evaluating various outcomes of interest from this large dataset, including the potential effect of childhood growth on risk of IA and development of T1D. In practice, childhood growth is often assessed via measurement of height and weight at periodic intervals (e.g. 6m, 1y, 2y visit) during pediatric office visits. During study follow, a subject may not have growth assessment at every visit for blood sampling for autoantibody or blood glucose measurement. In order to test interesting hypotheses, such as the effect of childhood growth on IA development, we need to link available growth data from patient charts with research data measured at more frequent intervals (i.e. every 3 to 9 months in these studies). However, this linking, which may be fuzzy based on nearest age visit or similar methods) often creates holes (sparsity) in the longitudinal research data making them difficult or impossible to analyze unless meaningful imputations are performed for growth variables.

In the T1DI cohort, missing data exist in weight and height measurements of children participating in follow-up. For meaningful inclusion of these variables in association analyses or predictive models (such as for IA development and T1D onset), data need to be imputed to coincide with timing of research visits. Furthermore, on the

basis of the observed/imputed body weight and height, other growth measurements such as the Body Mass Index (BMI) or other derived features such as percentiles and velocities of weight, height, and BMI may be computed which can then be also included in downstream models. Of note imputation of growth data for the T1DI cohort is hindered by a lack of a common reference for childhood growth data. The T1DI cohort growth data are from disparate geographic areas with variable demographic compositions and applicable population growth references may vary. For example, in the US, CDC growth charts[7] are used widely and are parameterized for childhood and adolescent growth from age 2 to 20 years in intervals of 6 months[8]. In many EU countries, country-specific growth charts or references exist.[9,10] However, these references are either not easily available or parameterized for similar use. Thus, in this study we hypothesized that the childhood growth in the T1DI cohort could be imputed using the CDC LMS parameters. We further hypothesized that using the CDC reference for imputation would produce overall low error rates. This approach could provide several other advantages in large international studies such as use of one common reference, model and fit. In this work we test our hypotheses using a novel multiple imputation (MI) methodology as described below.

Regardless of the reference data used, missing data should be handled and analyzed with caution depending on variable types and their mechanism of missingness. The missingness mechanism can be broadly classified into three types according to the characteristics with regards to randomness[11]: missing not at random (MNAR), missing at random (MAR), and missing completely at random (MCAR). Significant interactions between observed variables are an indicator of MNAR data. In general, missing value imputation for MNAR is not recommended, and hence few algorithms exist for this type of data. In contrast, a wide variety of methods are available for handling missing values in MAR and MCAR types. In practice, selecting the most appropriate method depends on the problem to be addressed and the available data. Single imputation methods simply replace missing data points with a single fixed value (e.g., zero, the mean, median, or most frequent value) by assuming that the data are MCAR. These methods could reduce variability of data and generate biased estimates of error variances.[12] Advanced methods such as model-based methods impute the missing values using a predictive distribution that models the underlying data missingness mechanism. The most commonly used model-based imputation methods include multiple imputation (MI), full information maximum likelihood (FIML) and maximum likelihood expectation-maximization (EM) imputation. In particular, MI assumes MAR and has been widely applied in analysis of clinical data.[13] Some popular algorithms in this category include nonparametric missing value imputation using random forest[14], K-nearest neighbors (KNN), multiple imputation by chained equations (MICE)[15], and Amelia.[16] More recently, more sophisticated models have been used for missing value imputation including multilayer perceptron, self-organizing maps as the predictive models. These have been applied to estimation of missing data for applications such as breast cancer diagnosis and cardiovascular data for clinical trials.[17,18] In our study, we use a general MI method and apply it to missing values of growth data in pediatric studies.

Methods

To describe characteristics of data missingness, we calculated missing value ratio. CDC growth charts and LMS parameters as well as country-specific growth charts are used for missing value imputations. We describe these sources below followed by our particular implementation of the MI algorithm, which considers uncertainty of the missing data and combines multiple sets of plausible imputed values.

Data Missingness: Missingness for height and weight were separately assessed for each subject in every dataset. We define missing value ratio as the number of rows (visits) where measurement (for height or weight separately) was missing divided by the total number of rows for that subject in the longitudinal dataset. For example, if the missing value ratio of height is 0.1, then 10% of longitudinal measurements of height were missing for that subject.

CDC Growth Charts and LMS Parameters: In the United States, the CDC growth charts[19] are widely used to monitor growth patterns of children and adolescents during clinical care. CDC charts describe values for height or weight at different percentiles by gender and age in months (from birth to 20-years in increment of 1 month). The smoothed percentile curves can be reproduced using LMS parameters, i.e. median (M), generalized coefficient of variation (S), and the power in the Box-Cox transformation (L). These parameters were calculated based on the Box-Cox transformations to normality. From these parameters, a measurement value can be calculated if the percentile is known, and vice versa. CDC growth charts span the age range of 0 – 36 months and 2 – 20 years. The LMS parameters, for 3rd, 5th, 10th, 25th, 50th, 75th, 90th, 95th, and 97th percentiles for weight, height and BMI are provided. For finer age intervals, interpolation could be used to obtain the LMS values. In this study, we use the

CDC reference to impute missing values in the body weight and height using the LMS parameters. We evaluate these values against the country-specific growth charts.

Country-specific growth charts: These were obtained from our collaborators for Swedish (DiPiS) and German (BABYDIAB) cohorts. There was no similar reference chart available for the Finnish cohort to our team. We used country-specific charts to impute missing values and compared these with CDC growth charts. The Swedish growth charts[10] cover the age range of 0 – 2191 days (72 months) and 61-228 months. It provides the values of standard deviations (-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5) for weight, height, and BMI. These growth charts were used for imputing missing values in DiPiS. Note that the LMS parameters are not provided in these charts. The German growth charts[9] cover the age range of 0 – 216 months. They list the LMS parameters and the percentiles (3rd, 10th, 25th, 50th, 75th, 90th, and 97th) for weight and height. They also provide the LMS parameters and the percentiles (3rd, 10th, 50th, 90th, and 97th percentile) for BMI. These charts were used for imputing missing values in BABYDIAB.

Of note, CDC growth charts cover measurements for up to 20-years of age, the German growth charts up to 18 years, and the Swedish charts up to 19 years. Thus, to address measurements until age of 20, extrapolation was used for the latter two cohorts. In this study, we imputed the missing values in the body height and weight using CDC growth charts for all five (US and EU) data sets.

Missing Value Imputation method: We use the growth charts with LMS parameters for missing value imputations and in that respect, it is a model-based method. Additionally, we apply multiple imputations to derive the missing value, i.e. "multiply impute" missing data in a single cross-section from an available time series of data. The flow chart is illustrated in Figure 1. The work was implemented using R (v3.6.3). The inputs of this algorithm include sex and age of a subject; the outputs are the imputed values and the corresponding percentiles for height/weight. This missing value imputation procedure consists of the following steps: For a growth measure (height or weight),

- **Step I**: For an individual subject, locate the ages at which the measurement values were missing and find the sex. Then identify the most proximal previous and subsequent ages with observed values. For each of the ages with missing values, repeat the following steps II-IV:

- **Step II**: For the previous age with known measurement value, use growth chart to compute the corresponding LMS values for that age through interpolation (if needed). Using the observed measurement and interpolated LMS values, compute its percentile (the percentile of the previous measurement) using the LMS formula. Do the same for the subsequent visit with a measured value to compute its percentile.

- **Step III**: if both the previous and the subsequent percentiles are obtained, take the average of these two as the percentile for the age at which we want to impute the missing value; otherwise, just use the available percentile (either the previous percentile or subsequent percentile).

- **Step IV**: for the age with a missing value, use the growth charts to obtain the LMS values through interpolation of the interval covering that age. Using the percentile value calculated in Step III and the interpolated LMS values to determine the imputed value, which is used to replace the missing measurement at that age.

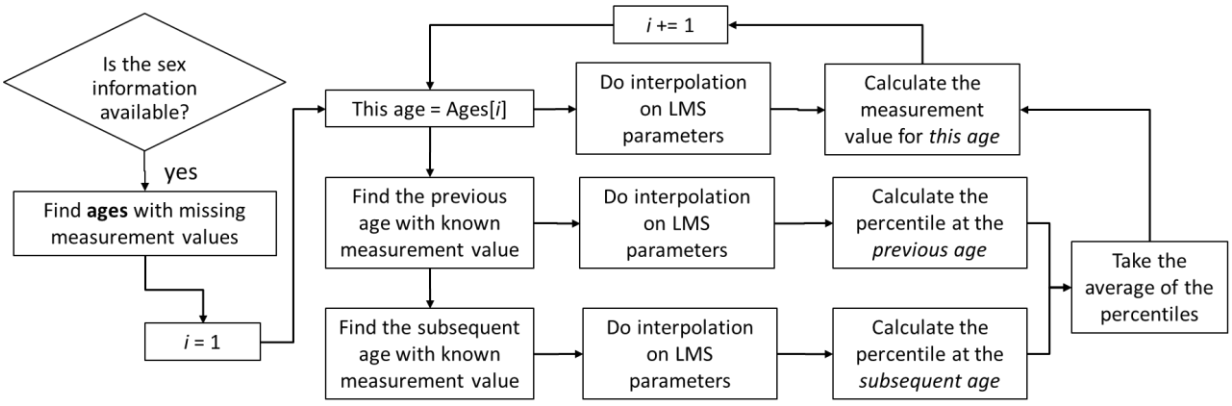

Figure 1. Flow chart for missing value imputation using the growth charts

Note that in Step III, in case that only a single percentile is available, extrapolation is used to impute the missing values and could cause larger imputation errors since it assumes that the growth percentile remains unchanged regardless of the length of time interval.

Performance measures: To assess the validity of missing value imputation method and compare the performance of the missing value imputation using country specific growth charts, we calculated imputation errors as follows. First, we randomly selected 20% of observed measurements and then removed their values creating a knock-out dataset. We then applied the missing value imputation method to the knock-out dataset and imputed the "missing" values. Finally, the imputation errors were computed. The same procedure was implemented using CDC or country-specific references. We used the following metrics for the imputation errors– mean absolute percentage error (MAPE)[20] and normalized root-mean-square error (NRMSE)[21]. MAPE is defined as

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i^{imp} - y_i^{act}}{y_i^{act}}\right|$$

where y_i^{imp} and y_i^{act} are imputed and actual (observed) measurement values, respectively, for *i*= 1, 2, …, *n* and *n* is the total number of imputed values. There is some variation in NRMSE definitions; however, the most common one was used in this study as follows,

$$NRMSE = \frac{\left[\frac{1}{n}\sum_{i=1}^{n}\left(y_i^{imp} - y_i^{act}\right)^2\right]^{1/2}}{\frac{1}{n}\sum_{i=1}^{n} y_i^{act}}$$

where the denominator is the average of actual measurements. From the above definitions, we can see that MAPE is more sensitive than NRMSE to the errors when the actual values are small, while NRMSE puts more weight on larger deviations than MAPE.

We calculated mean and standard deviation of MAPE or NRMSE values using 100 repetitions of knock-out and imputations and computed p-values for comparison with country specific references. All significance was tested at $P < 0.05$ using the Student's t test.

Results

In total there are 23,201 subjects for whom growth data were available (Table 1). Overall DIPP is the largest dataset with a frequent follow up compared to the overall T1DI cohort, and consequently had the smallest missing value ratios for both height and weight in our study. However, the two US datasets (DAISY and DEW-IT) had lower missing value ratios (for both height and weight) when compared to the other two EU sites (BABYDIAB and DiPiS). These differences are primarily due to differences in the individual study protocols and the frequency of follow up within them.

Table 1: Characteristics of data and missingness by dataset

Dataset	Number of Subjects	Height Missing Value Ratio Median [IQR]	Weight Missing Value Ratio Median [IQR]
BABYDIAB	2340	0.4 [0.25-0.50]	0.38 [0.25-0.50]
DAISY	2148	0.18 [0.09-0.50]	0.0 [0.0-0.06]
DEW-IT	2830	0.13 [0.0-0.73]	0.0 [0.0-0.50]
DiPiS	4227	0.67 [0.43-0.73]	0.64 [0.43-0.73]
DIPP	11656	0.0 [0.0-0.05]	0.0 [0.0-0.05]

Imputations in combined T1DI cohort: Figure 2 illustrates the MAPE and NRMSE values for missing value imputations in weight and height for the combined dataset (from 5 data sources) using the CDC growth charts. The MAPE values are plotted by the age group (6 groups from birth to 20-years as covered by CDC) as labeled on the y-axis. For the 0-2 year age group, the largest MAPE value for weight is 4.8% and the largest error for height is 1.7%.

The NRMSE values of the imputation errors are illustrated in Figure 2(b). The largest value is 6.0% for weight in age group 18-20 years and 2.1% for height in 0-2 years. From Figure 2(a) and 2(b), we find that the imputation errors for height are overall smaller than those for weight (in terms of both MAPE and NRMSE). In addition, for the same growth measure (height or weight), the MAPE values are smaller than the NRMSE values because NRMSE puts more weight on larger absolute measurement errors.

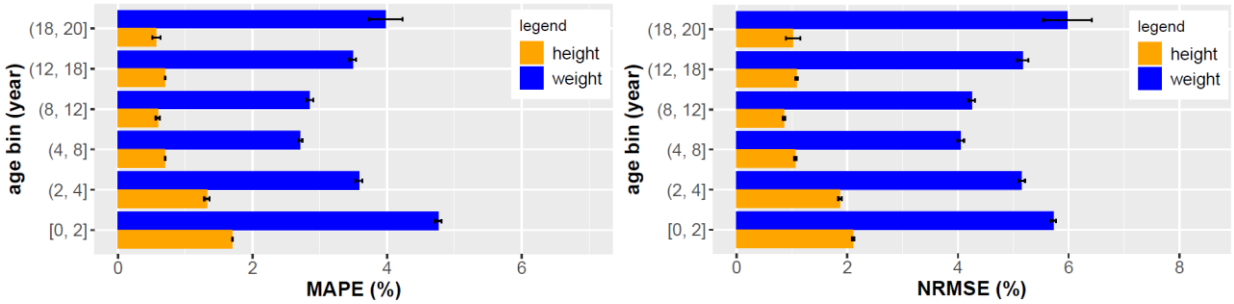

*Figure 2. Imputation errors in weight and height for **T1DI cohort** using CDC (a) MAPE; (b) NRMSE*

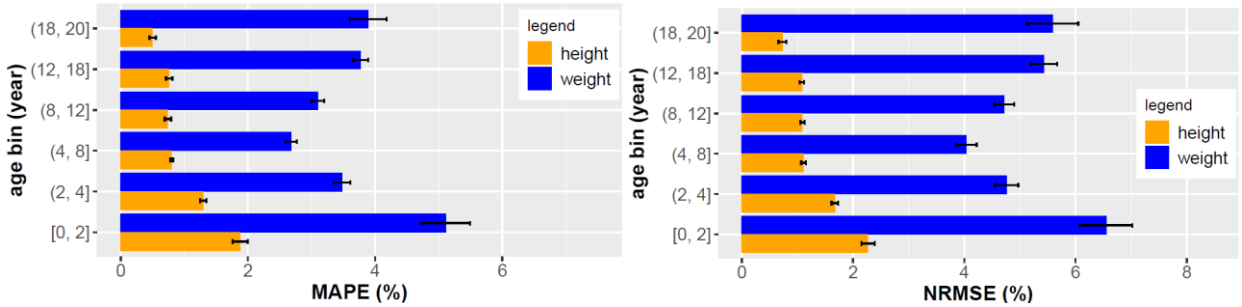

*Figure 3-1. Imputation errors in weight and height for **DAISY** using CDC (a) MAPE; (b) NRMSE*

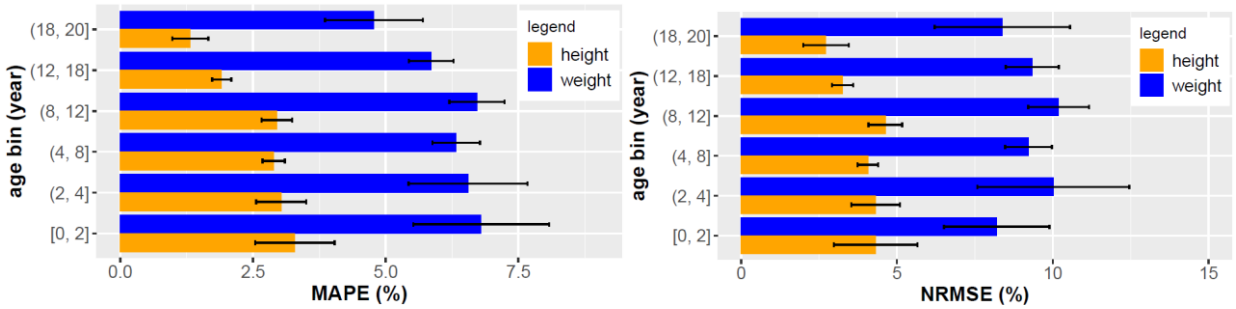

*Figure 3-2. Imputation errors in weight and height for **DEW-IT** using CDC (a) MAPE; (b) NRMSE*

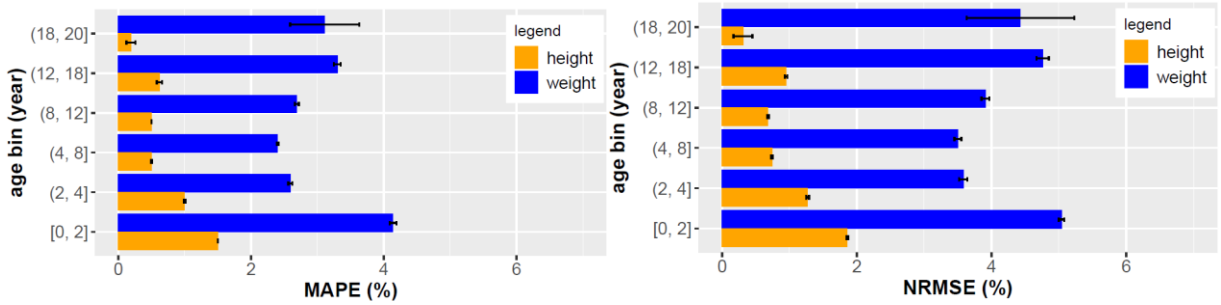

*Figure 3-3. Imputation errors in weight and height for **DIPP** (a) MAPE; (b) NRMSE*

Imputations in US datasets:

The MAPE and NRMSE values for two US data sources (DAISY and DEW-IT) are shown in Figure 3-1 (a and b) and Figure 3-2 (a and b) respectively. The imputation was implemented using CDC as reference. For DAISY, the largest MAPE values of weight are 5.1% for 0-2 years and 1.9% for height in the same age range. The largest NRMSE values for the same age group are 6.5% for weight and 2.3% for height. When compared to DAISY, the DEW-IT dataset has larger imputation errors in both height and weight. In terms of MAPE, the largest errors are 6.8% in weight and 3.3% in height for 0-2 years. The largest NRMSE values are 10.2% in weight and 4.6% in height for 8-12 years.

EU datasets and comparison to country-specific growth charts:

As the Finnish growth charts were not available for comparison, only CDC charts were used for imputing missing values in DIPP. The imputation errors in DIPP for height and weight can be found in Figure 3-3 (a) MAPE and (b) NRMSE. The largest MAPE values for weight using CDC growth charts are 4.1%, and 1.5% for height. The largest NRMSE values are 5.0% and 1.9% for weight and height, respectively.

For the other 2 EU datasets, we compared performance of imputation method using country specific growth charts. For BABYDIAB, we compared imputation method using the CDC growth charts and the German growth charts. For simplicity, we only plotted the MAPE values here on Figure 4-1 (a - height and b - weight). Since the German charts cover age up to 18 years, only the results from birth to 18 years are shown here. The largest MAPE values for weight in BABYDIAB (Figure 4-1) using CDC and German growth charts are 6.8% and 6.7% respectively, and the largest MAPE for height are 2.4 % and 2.3%, respectively. The Student's t test results computed for all ages together show no significant difference in the MAPE values of height using CDC (p =0.993), but the mean MAPE for weight using CDC is greater than that using the German charts (p < 0.001). However, the difference between them was fractional (6.02% vs. 5.81%).

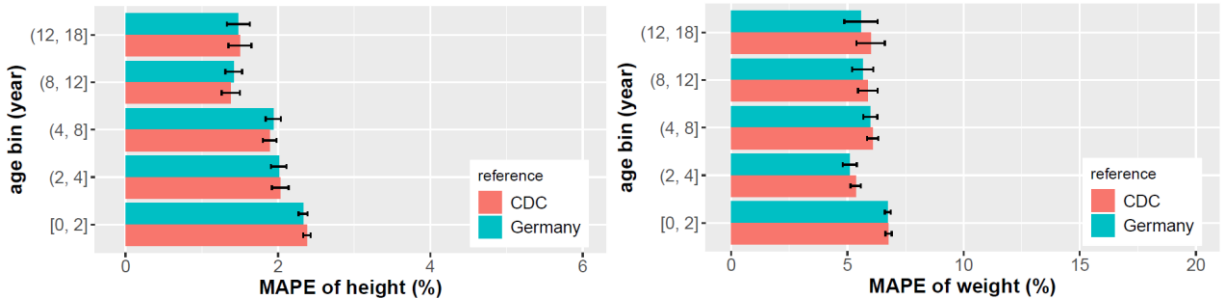

*Figure 4-1. MAPE of the imputed values in weight and height for **BABYDIAB** (a) height; (b) weight*

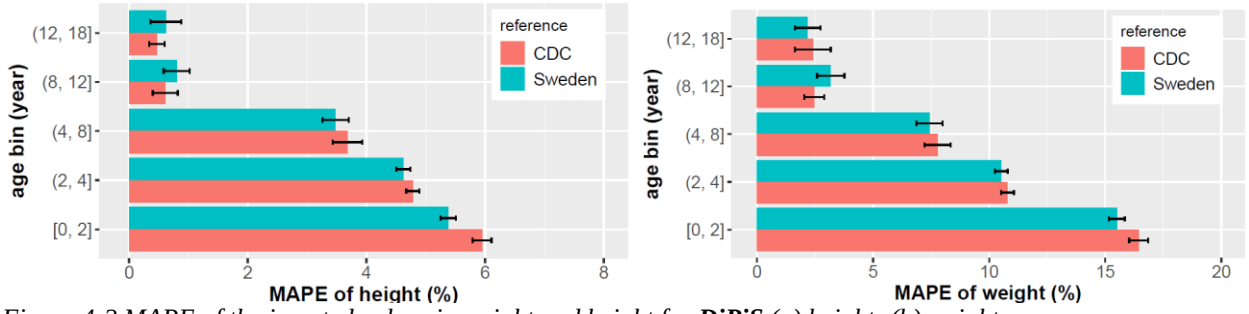

*Figure 4-2 MAPE of the imputed values in weight and height for **DiPiS** (a) height; (b) weight*

The MAPE results of DiPiS are shown in Figure 4-2 (a- height and b-weight) for which both CDC charts and Swedish charts were used. The age range of DiPiS is 0 – 12.99 years, therefore the largest age group is also 12-18 years. Due to the large missing value ratio, the largest MAPE values for weight in DiPiS using CDC and Swedish growth charts are 16.4% and 15.5% respectively, and the largest MAPE for height are 6.0 % and 5.4%, respectively.

The Student's t test results computed for all ages together show that there is no significant difference in the MAPE values of height using CDC or Swedish growth charts (p =0.37), and this is also true for weight as well (p = 0.51).

Discussion

In this study, we developed a missing value imputation method for data from multiple countries using CDC growth charts and a multiple imputation methodology. We assessed imputation performance using the large T1DI cohort and its constituent five real-world data sets. We also compared the CDC-imputed datasets to country-specific growth charts where available. From these experiments, we show that CDC reference may suffice to impute growth measurements in US and international (EU) pediatric studies. We further show that using this method, the imputation errors (as measured by MAPE and NMRSE) are within reasonable errors of margin (<10%) in the T1DI cohort. Furthermore, we found that the errors for imputed height are less than those for imputed weight across all datasets. This can be explained by the fact that height is much more tightly regulated by genes than weight which is strongly influenced by environment and eating behavior. Thus, height can be easily predicted on the basis of surrounding growth pattern, whereas weight can vary much over time.

There are a few country-specific findings. The imputation errors in BABYDIAB (Germany) and DAISY (U.S) were much smaller than those in DIPIS (Sweden). These may be due to two reasons, overall low missing value ratio in the DAISY and BABYDIAB data sets and well calibrated CDC and German growth charts. However, there is no significant difference between the error results using the CDC reference vs. the country-specific charts. The imputation errors are higher in DiPiS perhaps because of larger missing value ratios compared to the other data sets (Table 1) so that the average of the previous and subsequent observed percentiles are not a good estimate for the percentile at the age with missing growth measures.

There are several advantages of our approach. First, one common CDC growth reference for imputation in all datasets may facilitate imputations where standards don't exist or are unavailable, such as for the Finnish dataset. As the CDC reference comes from a large, diverse US sample population and has been studied extensively, it may also be the best guesstimate where growth patterns are unknown (or less studied) and almost certainly change by age groups (e.g. from early childhood to tween and adolescent age groups). Furthermore, using imputed growth measurements based on a reference model, one can derive many other features such as annual growth velocity and percentiles for downstream analyses. A second advantage is that the CDC reference covers a wide age range (0 to 20 years); therefore, one can use it to impute the missing values over the entire growth trajectory. This in comparison to a best fit approach offers significant advantages when rate of change (or slope)[22] in longitudinal data are unknown and where a change detection by age is needed. In these latter scenarios multiple models may be needed to assess different segments.[23] Third, our approach is also useful when the task at hand is to perform an association study (limiting data to a landmark age for prediction in future). In our use case of studying growth features and their potential contribution to development of islet autoimmunity[24], we found that imputations were reasonable. Finally, as the approach relies on the underlying model, our methods may be useful for other kinds of clinical data, for example for imputing missing blood pressure or glucose measurements.

As with all such studies, our approach also has some limitations. First, we base our results on data collected in the past (in many instances these studies started before 2000), i.e. before the CDC reference was widely available. However, this may have only made our results stronger in showing there is no significant difference between country-specific or CDC reference-based imputations over time. Second, it may not be the best approach to use when one needs to analyze outcomes based on observed growth to a time point (i.e. where we cannot avail ourselves of future information). In those situations, another model-based imputation method (e.g. using linear or polynomial multi-variable regression) may be a better fit. We believe the strength of our multiple imputation approach lies in where references such as the CDC are available for characterizing longitudinal pediatric growth trajectory. In addition, since the approach assumes MAR for the missing data, departures from this assumption should be assessed through sensitivity analyses[25].

Conclusion

Childhood growth variables measured during follow-up of children and adolescents up can be imputed using multiple imputation method and CDC reference parameters even when data represents geographically disparate US

and EU pediatric sites. Advantage of this approach include being able to derive many other features, such as annual growth velocity and growth percentiles; these features may be useful for downstream analyses.

Acknowledgments

We wish to thank the T1DI Study Group for their help in this work and study participants of the constituent studies. TheT1DI Study Group consists of the following members: 1) JDRF – Jessica Dunne, Olivia Lou; 2) IBM – Vibha Anand, Mohamed Ghalwash, Eileen Koski, Bum Chul Kwon, Zhiguo Li, Harry Stavropoulos, Kenney Ng; 3) DiPiS – Helena Elding Larsson, Josefine Jönsson, Åke Lernmark, Markus Lundgren, Marlena Maziarz, Lampros Spiliopoulos; 4) BABYDIAB – Peter Achenbach, Christiane Winkler, Anette Ziegler; 5) DIPP–Heikki Hyöty, Jorma Ilonen, Mikael Knip, Jorma Toppari, Riitta Veijola; 6) DEW-IT – Bill Hagopian, Michael Killian, Darius Schneider; 7) DAISY – Brigitte Frohnert, Jill Norris, Marian Rewers, Andrea Steck, Kathleen Waugh, Liping Yu. This work was supported in part by JDRF (1-IND-2019-717-I-X, 1-SRA-2019-722-I-X, 1-SRA-2019-723-I-X, 1-SRA-2019-719-I-X, 1-SRA-2019-721-I-X, 1-SRA-2019-720-I-X).

References

1. Kontopantelis E, Parisi R, Springate DA, Reeves D. Longitudinal multiple imputation approaches for body mass index or other variables with very low individual-level variability: the mibmi command in Stata. *BMC Res Notes*. 2017;10(1):41. doi:10.1186/s13104-016-2365-z

2. Rewers M, Bugawan TL, Norris JM, et al. Newborn screening for HLA markers associated with IDDM: diabetes autoimmunity study in the young (DAISY). *Diabetologia*. 1996;39(7):807-812.

3. Wion E, Brantley M, Stevens J, et al. Population-wide infant screening for HLA-based type 1 diabetes risk via dried blood spots from the public health infrastructure. *Ann N Y Acad Sci*. 2003;1005:400-403.

4. Ziegler AG, Hummel M, Schenker M, Bonifacio E. Autoantibody appearance and risk for development of childhood diabetes in offspring of parents with type 1 diabetes: the 2-year analysis of the German BABYDIAB Study. *Diabetes*. 1999;48(3):460-468. doi:10.2337/diabetes.48.3.460

5. Larsson HE. A Swedish approach to the prevention of type 1 diabetes. *Pediatr Diabetes*. 2016;17(S22):73-77. doi:10.1111/pedi.12325

6. Kupila A, Muona P, Simell T, et al. Feasibility of genetic and immunological prediction of type I diabetes in a population-based birth cohort. *Diabetologia*. 2001;44(3):290-297. doi:10.1007/s001250051616

7. Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, et al. CDC growth charts: United States. *Adv Data*. 2000;(314):1-27.

8. Kuczmarski RJ, Ogden CL, Guo SS, et al. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat 11*. 2002;(246):1-190.

9. Reference percentiles for anthropometric measures and blood pressure based on the German Health Interview and Examination Survey for Children and Adolescents 2003–2006 (KiGGS). https://www.rki.de/EN/Content/Health_Monitoring/Health_Reporting/GBEDownloadsB/KiGGS_referenzperzentile.pdf?__blob=publicationFile

10. Wikland KA, Luo ZC, Niklasson A, Karlberg J. Swedish population-based longitudinal reference values from birth to 18 years of age for height, weight and head circumference. *Acta Paediatr Oslo Nor 1992*. 2002;91(7):739-754. doi:10.1080/08035250213216

11. Mislevy RJ, Little RJA, Rubin DB. Statistical Analysis with Missing Data. *J Educ Stat*. 1991;16(2):150. doi:10.2307/1165119

12. Gemici S, Bednarz A, Lim P. A primer for handling missing values in the analysis of education and training data. *Int J Train Res*. 2012;10(3):233-250. doi:10.5172/ijtr.2012.10.3.233

13. Romero V, Salmerón A. Multivariate Imputation of Qualitative Missing Data Using Bayesian Networks. In: *Soft Methodology and Random Information Systems*. Springer Berlin Heidelberg; 2004:605-612. doi:10.1007/978-3-540-44465-7_75

14. Stekhoven DJ, Buhlmann P. MissForest--non-parametric missing value imputation for mixed-type data. *Bioinformatics*. 2012;28(1):112-118. doi:10.1093/bioinformatics/btr597

15. Buuren S van, Groothuis-Oudshoorn K. MICE : Multivariate Imputation by Chained Equations in *R*. *J Stat Softw*. 2011;45(3). doi:10.18637/jss.v045.i03

16. Honaker J, King G, Blackwell M. Amelia II: A Program for Missing Data. *J Stat Softw*. 2011;45(7). doi:10.18637/jss.v045.i07

17. Rahman MM, Davis DN. *Fuzzy Unordered Rules Induction Algorithm Used as Missing Value Imputation Methods for K-Mean Clustering on Real Cardiovascular Data*.

18. Jerez JM, Molina I, Garcá-Laencina PJ, et al. Missing data imputation using statistical and machine learning methods in a real breast cancer problem. *Artif Intell Med*. 2010;50(2):105-115. doi:10.1016/j.artmed.2010.05.002

19. Kuczmarski RJ, Ogden CL, Guo SS, et al. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat 11*. 2002;(246):1-190.

20. Moritz S, Sardá A, Bartz-Beielstein T, Zaefferer M, Stork J. Comparison of different Methods for Univariate Time Series Imputation in R. *ArXiv151003924 Cs Stat*. Published online October 13, 2015. Accessed March 1, 2021. http://arxiv.org/abs/1510.03924

21. Jin L, Bi Y, Hu C, et al. A comparative study of evaluating missing value imputation methods in label-free proteomics. *Sci Rep*. 2021;11(1):1760. doi:10.1038/s41598-021-81279-4

22. Desai M, Mitani AA, Bryson SW, Robinson T. Multiple Imputation When Rate of Change is the Outcome of Interest. *J Mod Appl Stat Methods JMASM*. 2016;15(1):160-192. doi:10.22237/jmasm/1462075740

23. Ibrahim JG, Molenberghs G. Missing data methods in longitudinal studies: a review. *Test Madr Spain*. 2009;18(1):1-43. doi:10.1007/s11749-009-0138-x

24. Li Z, Anand V, Dunne JL, et al. 1280-P: Improving Type 1 Diabetes (T1D) Prediction by Incorporating Growth Features into Landmark Models. *Diabetes*. 2020;69(Supplement 1):1280-P. doi:10.2337/db20-1280-P

25. Rezvan PH, Lee KJ, Simpson JA. The Rise of Multiple Imputation: A Review of the Reporting and Implementation of the Method in Medical Research. *BMC Med Res Methodol.* 2015;15:30. doi: 10.1186/s12874-015-0022-1