

MODELLING GROWTH PROCESSES IN INDIAN  
CHILDREN AND ADOLESCENTS

A DISSERTATION SUBMITTED TO THE DEPARTMENT OF BIOLOGY  
IN CANDIDACY FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

BY

SANDRA ARAVIND AREEKAL

REG. NO. 20173515



INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

October 25, 2023

*To amma, achan and appunni.*

# DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the institute and can also, evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

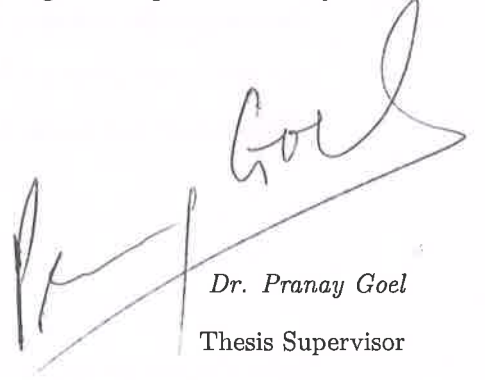
*Date: October 25, 2023*

  
*Sandra Aravind Areekal*  
*Reg No: 20173515*

# CERTIFICATE

Certified that the work incorporated in the thesis entitled "*Modelling growth processes in Indian children and adolescents*", submitted by *Sandra Aravind Areekal* was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

*Date: October 25, 2023*



Dr. Pranay Goel  
Thesis Supervisor

डॉ. प्रणय गोयल / Dr. Pranay Goel  
सहयोगी प्राध्यापक / Associate Professor  
भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान  
Indian Institute of Science Education & Research  
पुणे / Pune - 411 008, India

# Acknowledgments

I express my heartfelt gratitude to my supervisor, Dr Pranay Goel, for his invaluable support and guidance throughout this journey. His mentorship has been instrumental in shaping my research and personal growth. He has guided me in developing my writing, speaking, and thinking skills at every stage.

I am grateful to my mentors, Dr Anuradha Khadilkar and Prof. Tim J Cole, for their unwavering support and guidance throughout this journey. Their patience, expertise, and dedication to the work have been my constant source of inspiration. I am truly fortunate to have had the opportunity to learn from them.

I thank all the children and parents for their consent to participate in the study and share the data. I am thankful to the research group at the H C Jehangir Medical Research Institute Pune for their hard work in collecting high-quality data. I am also grateful to Dr Khadilkar for sharing the data with us.

I am grateful for the funding support by the Council of Scientific and Industrial Research, India.

I acknowledge the Newton Bhabha PhD Placements Program 2019-2020, jointly funded by the Department of Biotechnology, Ministry of Science and Technology, Govt. of India and British Council, New Delhi. I thank everyone at the Institute of Child Health for making me feel at home. I am thankful that I had the opportunity to interact with Ms Angie Cole and thank her for her kindness.

I acknowledge that the National Supercomputing Mission (NSM) for providing computing resources of 'PARAM Brahma' (PB) at IISER Pune, which is implemented by C-DAC and supported by the Ministry of Electronics and Information Technology (MeitY) and Department

of Science and Technology (DST), Government of India has been crucial in my research. I thank the senior technical officer (HPC), Nisha Kurkure, for patiently helping me solve all issues I faced related to PB.

I thank the IISER Pune IT systems administrator, Neeta Deo, for helping us host our webapp. I thank the bio-office and library staff for following through on all my (late) requests. I thank Ms Shweta from the AIC SEED IISER Pune, who helped me learn a lot about entrepreneurship. I thank all the library staff at IISER Pune for their support throughout my time here. I am extremely grateful to the IT and sports facilities at IISER Pune.

I acknowledge the support from the R and Stack Exchange communities for helping me troubleshoot and debug. ChatGPT has helped paraphrase and build some of the bib entries for my thesis.

I thank my friends Dr Namrata Arvind and Vishakh Narayanan for being the math teachers I never had. They have been instrumental in my understanding of mathematics, science, research and the world in general. I thank them for *extending my field*.

Thankfully, I do not regret doing a PhD. Infact, I had a lot of fun. I thank Dr Namrata Arvind for asking me to look at the moon. I am grateful for the countless perspective-shifting conversations on everything we had. Her fight in academia and badminton court inspires me every day. I am also thankful for the support group that helped me stay sane. I thank the second-best roommate in the world, Rajeshwari, for understanding me. I thank Firdousi Parvez and Dilsha for all the adventurous journeys we took, both literal and metaphorical. I thank my privilege of being friends with Ramya and Vishakh. Thanks to Neethu for all the fun times. It feels great to find people who gets you. I am grateful for all the *sulaimani* sessions with Arjun, Naveen, Amal, Gokul, Sreeram, Vrinda, Neethu, and other *jaada mallus* at IISER P. I thank my batchmates Shruti, Swapna, Saikat, Tejashree, Shirsha, Nazneen, Prachi, Mayuresh, Lokesh, and Ravi for all the fun times. I thank my amazing labmates, Arjun, Somashree, Rashmi, Prajjwal, Shreya, Suyog, and Mukta, for making research fun. Thanks to *amar shonar* Soma, Tumpa and Fido for their Bangla lessons. I am grateful to Alakananda and Vibhishan, for all the fun discussions on the Origin of Species. I thank all the members of the Zeher team for their cause of giving everyone a chance. I thank everyone who gave me a chance.

Lastly, I thank my family, friends, and teachers.

# CONTENTS

Acknowledgments	vi
1 Introduction	1
2 A physiological and public health perspective of resting metabolic rate, growth centiles and growth curves.	5
3 Modelling resting metabolic rate in Indian children	25
4 Modelling height growth in Indian children and adolescents	52
5 Height growth under a persisting metabolic insult	74
6 Discussion and Conclusion	91
A Model optimisation	97
B List of publications	138
C Ethics statement	141
D Software	142

# List of Figures

A.1	The organ mass data set for the North American children compiled by Altman and Dittmer (1962) adapted from Table 1 in Wang (2012). . . . .	98
A.2	Relative cellularity (Rc) adapted from Table 2 in Wang (2012). . . . .	101
A.3	Relative specific metabolic rate adapted from Table 3 in Wang (2012). . . . .	101
A.4	The mean measured RMR/BM ( $\pm$ SE) measured in each age group (solid line), and the Wang model predicted RMR/BM (dotted line) reproduced here for continuity. Compare to Figure 3.2 in Chapter 3 or Figures 3 and 4 in Areekal et al. (2023). Note that this figure reports SE instead of SD. . . . .	102
A.5	Optimal delta for Model 1 in Girls is 0.90 used to generate Figure 3.3 in Chapter 3 or Figure 5 in Areekal et al. (2023). . . . .	104
A.6	Optimal delta for Model 1 in Girls is 0.77. Used to generate Figure 3.3 in Chapter 3 or Figure 6 in Areekal et al. (2023). . . . .	105
A.7	Mean ( $\pm$ SE) RMR/BM measured (solid line) in each age group compared to the Model 1 predicted RMR/BM (dotted line) for the Caucasian population. Compare to Figure 3.3 in chapter 3 or Figure 5 and 6 in Areekal et al. (2023). Note that the bars represent SE here instead of SD. . . . .	106
A.8	Optimal delta for Model 2 in boys is 0.85, which is used produce Figure 3.5 in Chapter 3 or Figure 8 in Areekal et al. (2023). . . . .	107
A.9	Optimal delta for Model 2 in girls is 0.64, which is used produce Figure 3.5 in chapter 3 or Figure 9 in Areekal et al. (2023). . . . .	108
A.10	The mean ( $\pm$ SE) RMR/BM measured (solid lines) and the Model 2 predicted RMR/BM (dotted line). Comparable to Figure 3.5 in Chapter 3 or Figures 8 and 9 in Areekal et al. (2023) which reports the mean measured RMR/BM ( $\pm$ SD). . . . .	109
A.11	PSCG study data described in section 4.2.1. . . . .	110



A.12	Heights measured in the PSCG dataset described in section 4.2.1. . . . .	110
A.13	Individual growth curves in the PSCG dataset comparable to Figure 4.3 in chapter 4 and Figure 3 in Areekal et al. (2022). . . . .	111
A.14	Residuals from the initial SITAR model in boys. . . . .	113
A.15	Residuals from the initial SITAR model in girls. . . . .	114
A.16	Centiles estimated for boys under BCCGo distribution for height with age. Comparable to Figure 4.1 in chapter 4 or Figure 1 in Areekal et al. (2022). . . . .	121
A.17	Centiles estimated for girls under NO distribution for height with age. Comparable to Figure 4.1 in chapter 4 or Figure 1 in Areekal et al. (2022). . . . .	122
B.1	Copyrights information for including the article Areekal et al. (2023) in chapter 3.	139
B.2	Copyrights information for including Areekal et al. (2023) in chapter 5. . . . .	140
B.3	Copyrights information for including Areekal et al. (2022) in chapter 4. . . . .	140
D.1	A height growth monitoring app for Indian children named growHT ( <a href="https://digimed.acads.iis-erpune.ac.in/growth-charts">https://digimed.acads.iis-erpune.ac.in/growth-charts</a> ). . . . .	142

# List of Tables

2.1	Phenomenological models for BMR. E= BMR in kcal/day, a=age (year), f=female, m=male, W=weight (kg), H=height (cm), h=humidity, t=temperature . . . . .	8
A.1	The summary statistics of the MCS dataset. . . . .	98
A.2	Height measurements in the PSCG dataset. . . . .	109
A.3	Summary of the cleaned PSCG dataset. . . . .	116
A.4	Summary of the longitudinal measurements in STDM and the PSCG dataset. . .	123

# Abstract

Human growth is a complex non-linear process that occurs in several stages, including the foetal, infancy, childhood, and adolescence phases. Different patterns of growth and underlying mechanisms characterise each of these stages. In this thesis, we examined various aspects of physiological growth in childhood and adolescence in the Indian population. Initially, we looked at two measures of growth that show high levels of adaptability: metabolism and body composition. We found that Indian children have significantly lower energy expenditure (at rest) compared to Western children, and we provide a potential explanation based on differences in organ mass and body composition. We further studied height growth during adolescence and characterised the pubertal growth spurt using serial measurements of height over eight years using the SITAR (SuperImposition by Translation and Rotation) model. Finally, we investigated the impact of a metabolic disorder, Type-1 Diabetes Mellitus (T1DM), on otherwise canalised height growth programming. We found that it leads to a delay and extension of the pubertal growth spurt, resulting in compromised height in children with T1DM. However, once the resulting SITAR model is parameterized by *size*, *timing* and *intensity* of pubertal growth, the underlying growth pattern in children with T1DM was indistinguishable from those without. Our study shows how the underlying growth mechanism is robust to a persisting metabolic insult. Our study adds to the knowledge for researchers, policymakers and clinicians to understand the optimal modes for personalised interventions to address the many non-communicable diseases, including the double burden of malnutrition, while considering the long-term health implications of growth.

# CHAPTER 1

## Introduction

*“I have hitherto sometimes spoken as if the variations - so common and multiform with organic beings under domestication, and in a lesser degree with those under nature - were due to chance. This, of course is a wholly incorrect expression, but it serves to acknowledge plainly our ignorance of the cause of each particular variation.”*

*- On the origin of species, Charles Darwin*

Growth is a fundamental process inherent to all living organisms. So are the variations in the physiological features they exhibit. Moreover, both are rarely random: growth follows a definite underlying developmental program; and the biological variations are (often) due to a definite underlying cause. This thesis explores the variations in two key features, energy requirement and size, during a period of growth in humans. Specifically, we use the tools of observational studies to investigate the determinants of intra-species variation in metabolism at rest and height growth during childhood and adolescence. Metabolism is thought to be a more adaptable component whereas height growth less so. Initially, we study the changes in height growth and metabolism separately, and later look at how height growth is influenced under a metabolic disorder Type-1 diabetes mellitus (T1DM). We conclude with a discussion emphasising the sexual dimorphisms observed in metabolism and size followed by a brief discussion on the implications of the current study for future research.

The period of growth in humans spans around two decades. It is generally categorised into multiple phases, namely, fetal, infancy, childhood and adolescence. The duration of each phase and the timing of transition from each are of importance in terms of the form and physiology

of the mature adult (Tanner 1981). The timing and duration of adolescent growth play a pivotal role in shaping body composition changes, such as body fat and bone mineralization. Moreover, understanding the significance of timing and duration becomes even more essential when considering the potential link between early growth variations and later chronic disease risks through telomeric attrition, which affects growth rate, cellular senescence, and age-related degenerative processes (Cameron and Demerath 2002). Here we focus on the final stages of growth: childhood and adolescence. As early as 1994, childhood and adolescence has been identified as one of the three critical periods for the development of obesity (Dietz 1994).

Multiple physiological features undergo sudden changes during childhood and adolescence. However, here we focus on the changes in metabolism and height growth. Specifically, we study the metabolism through the resting metabolic rate (RMR) which measures the resting energy expenditure at rest (De B. Weir 1949). The RMR represents the physiology of the child in contrast to total energy expenditure (TEE) which includes the energy expenditure for physical activity, thermic effect of feeding etc Hall et al. (2011); and hence more suitable to study the underlying physiological factors. Further, height growth is analysed here based on serial measurements of height from childhood through adolescence known as height growth curves. Height growth curves captures the history of a child's growth, the path taken from childhood to adulthood. Thus, we study how RMR and height growth curves change through the milestone event of pubertal growth spurt which is a period of rapid growth driven by multiple hormones.

Earlier research focused on the *earliest* windows of interventions in case of growth disorders. Fetal, neonatal and infancy had been given a lot of attention indeed for the right reasons. Growth is thought to be programmed by the end of 2 years, with little to no windows available for later intervention. However, the last possible window of intervention also demands similar attention. Pubertal growth spurts may aid to identify the possible windows of intervention (or the last opportunity) in terms of growth disorders or overweight management in children and adolescents. Hence it is important to study how the metabolism influences pubertal growth spurt and vice versa.

There are three broad objectives that are studied in this thesis, described in detail in the next three chapters, are: i) how RMR is distributed across age and what factors contribute to the variation observed in RMR in the Indian children. ii) how does height vary with age in Indian children by characterising the average as well as individual patterns of growth. iii) how height is influenced in children diagnosed with a metabolic disorder T1DM in comparison to growth in

children without T1DM. Further, how individual variations in growth can be explained using parental height and disease severity.

**Chapter 2** offers a concise overview of the relevant literature, spanning more than a century, on three key topics: resting metabolic rate, height growth centiles, and height growth curves. Resting metabolic rate, as a measure of minimal energy expenditure for vital bodily functions, has been widely investigated in the context of understanding the energy requirements of a population. Height growth centiles and growth curves have played crucial roles in monitoring and understanding children's growth patterns. By reviewing the existing literature, this chapter establishes the foundation for the current study, identifying research gaps and emphasising the need for further investigation in the specific context of the thesis.

In **Chapter 3**, we aim to assess whether a previously established model for resting metabolic rate (RMR) in Caucasian adolescents can be applied to Indian children. This study will help determine the suitability of the existing RMR model for Indian children and contribute to better understanding metabolic differences in this population.

**Chapter 4** focuses on comparing two methods for analyzing child anthropometry: growth centiles and growth curves. Growth centiles involve plotting percentile values based on age and height, while growth curves capture the trajectory of growth over time. The objective of this research is to specifically examine height growth centiles and curves in Indian children. By comparing these two approaches, the study aims to gain insights into the growth patterns and variations in height among Indian children. The findings will contribute to a deeper understanding of child development in this population, potentially informing healthcare interventions and monitoring practices.

In **Chapter 5**, we examine the impact of Type-1 diabetes mellitus (T1DM) on height growth in Indian children. By comparing growth curves between children with T1DM and a control group, while considering parental height and disease severity, valuable insights into the effects of a persistent metabolic insult on height growth patterns can be gained.

In **Chapter 6**, we summarise the main finding of the thesis with an emphasis on the sex specific differences observed during childhood and adolescence in resting metabolic rate and height growth. We also discuss the strengths and limitations of the study followed by future prospects and a brief concluding remark.

Energy (crisis) was an important factor in the evolution of size in humans in the past (Hochberg and Albertsson-Wikland 2008); and is a leading factor in the rise of non-communicable disorders in the present (Popkin et al. 2020; Mathers and Loncar 2006). Interventions to address the problems of size, such as stunting, wasting, overweight, underweight, etc. would need to be personalised on the energy requirement of an individual in the future. Thus, understanding the inter-workings of metabolism and height is an important problem in current biology.

## Bibliography

- Cameron, N. and Demerath, E. W. (2002). Critical periods in human growth and their relationship to diseases of aging. *Am. J. Phys. Anthropol.*, 119(S35):159–184.
- De B. Weir, J. B. (1949). New methods for calculating metabolic rate with special reference to protein metabolism. *J. Physiol.*, 109(1-2):1–9.
- Dietz, W. H. (1994). Critical periods in childhood for the development of obesity. *The American Journal of Clinical Nutrition*, 59(5):955–959.
- Hall, K. D., Sacks, G., Chandramohan, D., Chow, C. C., Wang, Y. C., Gortmaker, S. L., and BA., S. (2011). Quantification of the effect of energy imbalance on bodyweight. *Lancet*, 378(9793):826–37.
- Hochberg, Z. and Albertsson-Wikland, K. (2008). Evo-devo of infantile and childhood growth. *Pediatric research*, 64(1):2–7.
- Mathers, C. D. and Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLoS medicine*, 3(11):e442.
- Popkin, B. M., Corvalan, C., and Grummer-Strawn, L. M. (2020). Dynamics of the double burden of malnutrition and the changing nutrition reality. *Lancet*, 395(10217):65–74.
- Tanner, J. M. (1981). Growth and Maturation during Adolescence. *Nutr. Rev.*, 39(2):43–55.

## CHAPTER 2

A physiological and public health perspective of resting metabolic rate, growth centiles and growth curves.



## 2.1 Resting Metabolic Rate

### 2.1.1 Intraspecific variation in RMR

Life sustains on energy released from a set of chemical reactions in the cells of an organism. The minimal amount of such chemical reactions necessary to sustain the vital body functions comprises basal metabolism and the energy output per unit time by the basal metabolism is the basal metabolic rate (BMR) (Harris and Benedict 1918). Earliest studies of human metabolism (or nutrition) reported BMR which measured the energy expenditure of an organism at rest, in wakefulness following strict, rigorous conditions prior to and during BMR measurement to ensure *basal* metabolism (McMurray et al. 2014). Later studies measured the energy expenditure under relatively milder conditions at rest and is commonly referred to as resting metabolic rate (RMR) or resting energy expenditure (REE). It is measured after 8 hrs fast (ranged between 3 to 12 hrs in previous studies), in the absence of any physical activity (8-24 h prior measurement) and diseases, minimal emotional disturbances, thermo-neutral environment (22-26°C), at rest (15-60 min prior) and in wakefulness (McMurray et al. 2014). We hitherto use RMR to refer to the energy expenditure measured at rest in the previous studies, including studies that reported REE and BMR.

RMR can be not only be measured through direct or indirect calorimetry but also estimated from predictive equations. Equation based models are typically developed for some specific population (Henry 2005). In direct calorimetry, the subject enclosed within a closed chamber and one measures the heat produced through temperature differences. Indirect calorimetry measures the oxygen consumption and carbon dioxide elimination of each subject and predicts the caloric output based on the De B. Weir (1949) equations. RMR measurement through calorimetry requires laboratory conditions, trained technicians and expensive equipments. Due to practical difficulties to measure RMR for a large number of patients, predictive equations were developed from sample populations to serve as a reference.

Here we will review the factors leading to intraspecific variation in RMR based on the physiological and phenomenological research on RMR/REE/BMR chosen from the literature. RMR and REE are equivalent to BMR but is measured under less stringent conditions than BMR (McMurray et al. 2014). The functional ecology views on RMR, the allometric scaling of metabolic rate to bodymass and the interspecies variation of RMR are not discussed here.

## **2.1.2 The evolution of RMR research over a century**

The basal metabolic rate was introduced as a clinical parameter in the nineteenth century. In 1895, Magnus Levi showed a dependence of metabolic rate on the thyroid gland secretions in humans (Henry 2005). BMR was the only reliable diagnostic measure for thyroid malfunctions till the early 19th century. Diagnosis of growth disorders, metabolic disorders like diabetes, and nutritional problems were also based on BMR measurements. After 1950, BMR usage decreased with the introduction of Iodine-metabolism based diagnostic tests in thyroid dysfunction. But, having a strict standardised measurement made it of interest to researchers as it helps in the direct comparison of energy utilisation by different organisms. In the 20th century began the use of RMR to estimate the energy expenditure of an organism. In 1922, Bedale used BMR measurements of 100 school children to estimate the energy expenditure to suggest food requirements for children at school (Bedale 1923). Later, the Food and Agriculture Organization (FAO) of the United Nations, World Health Organization (WHO) and United Nations University (UNU), proposed to estimate energy requirements for a population-based on energy expenditure of the sample sub-population (FAO/WHO/UNU 1985). Predictive equations are developed from RMR measurements of a sample subpopulation based on easy to measure anthropometric factors and later applied to individuals of similar characteristics. These FAO/WHO/UNU recommendations are used by governments and organisations under the United Nations (UN) to devise nutritional policies especially in the interest of the eradication of poverty and malnutrition. In the present, understanding the physiological variation in RMR plays a pivotal role to study and management of many non-communicable diseases.

## **2.1.3 Factors leading to the variation in RMR**

The energy expenditure of an organism is the sum of energy utilised by its living cells. This can be looked from different levels of organisation: molecular, cellular, tissue-organ or whole organism. Studies over the years on variation in RMR are categorised and presented based on the organisation level it addresses.

### **2.1.3.1 Organism level approaches in RMR studies**

The attempts to understand the metabolism of humans were always centred around clinical applications rather than a pure exploratory view. This clinical and nutritional use attracted

Authors	Model	Remarks
Aub and Du Bois (1917)	$E = A_0(W^{0.425}H^{0.715}71.84)^{2/3}$	n=10
Harris and Benedict (1918)	$E_{male} = 66.4730 + 13.7516W + 5.0033H - 6.7550a$ $E_{female} = 655.0955 + 9.5634W + 1.8496H - 4.6756a$	n=136, $R^2=0.8$ n=103, $R^2=0.6$
Kleiber (1932)	$E(m) = 71.2W^{3/4}[1 + 0.004(30 - a) + 0.010(H - 43.4)]$ $E(f) = 67.4W^{3/4}[1 + 0.004(30 - a) + 0.018(H - 43.4)]$	n=136 n=103
Quenouillie et al. (1951)	$E = 2.975H + 8W + 11.7S + 3.0h - 4.0t + 293.8$	n=8600
Energy et al. (1985)	$E_{male} = 152W^{0.73}$ , $E_{female} = 123.4W^{0.75}$	2238
Schofield (1985)	$E_{male} = 815 + 36.6W$ , $E_{female} = 580 + 31.1W$ $E_{male} = 11.472W + 873.1$ $E_{female} = 8.126 \times W + 845.6$	2238 30-60 yrs n=7173
Henry (2005)	$E_{male} = 14.2W + 593$ $E_{female} = 9.74W + 694$	30-60yrs, $R^2=0.7$ n=10552
Mifflin et al. (1990b)	$E_{male} = 10W + 6.25H - 5a + 5$ $E_{female} = 10W + 6.25H - 5a - 161$	n=251, $R^2=0.71$ n=247

Table 2.1: Phenomenological models for BMR. E= BMR in kcal/day, a=age (year), f=female, m=male, W=weight (kg), H=height (cm), h=humidity, t=temperature

large number of population based studies in BMR. Thus, different phenomenological models were created from the easy to measure anthropometric factors in those population studies. The evolution of phenomenological models are explained in this section. Table 2.1 shows a set of selected RMR predictive equations.

The early studies of BMR reasoned that the energy expenditure of an organism must depend on the energy dissipated as heat, which increases with surface area of an organism. In 1917, Aub and Du Bois (1917b,a) created a standard reference chart of BMR per unit area per unit time for men and women in different age groups (14-80 yrs). But these were based on BMR and surface area measurements of 10 subjects. Since it is difficult to measure the surface area of an individual, Du Bois and Du Bois (1989) based it on height and weight. Later several more BMR equations are reported based on surface area (Boyd 1935; Brody and Lardy 1946; Haycock et al. 1978; Breitman 1932; von Schelling 1954; Isacksson 1958). But the variation in surface area could not explain many observations like high BMR of children of age 6-18 months with less surface area and later reduction in BMR in adults and then onto elderly. There is also a difference in the BMR of individuals in different temperature zones with the same surface area. Since the surface area is not directly measurable but predicted from weight and height, this is a double-step estimation which adds up the error.

In 1918, Harris and Benedict in their classic paper titled “A biometric study of Human Basal Metabolism”, developed two equations from BMR measurements of 239 healthy subjects (136 males and 103 females) and reported that 80% variation in males and 60% variation in females

could be explained by weight, height and age (Harris and Benedict 1918). The Harris-Benedict equations for male and female BMR values,  $E$ , are given as:

$$\begin{aligned} E_{male} &= 66.4730 + 13.7516W + 5.0033H - 6.7550a, \\ E_{female} &= 655.0955 + 9.5634W + 1.8496H - 4.6756a, \end{aligned} \tag{2.1}$$

where  $W$  is the weight in kg,  $H$  is the height in centimetres and  $a$  is the age in yrs. This equation was reference standard used widely around the world for a long time. To reduce the measurement error, the average of BMR measurements taken multiple times and on multiple days were used, and pneumograph was used to confirm the absence of any muscular activity.

Later in the 1957, Durnin et al. proposed the idea of using BMR to estimate energy requirements instead of using energy intake (Energy et al. 1985). They surveyed the existing literature and presented two sets of equations to FAO/WHO/UNU as,

$$\begin{aligned} E_{male} &= 152W^{0.73} \quad \& \\ E_{female} &= 123.4W^{0.75}, \end{aligned} \tag{2.2}$$

where  $W$  is the weight of the subject. Eq. 2.2 was simplified into:

$$\begin{aligned} E_{male} &= 815 + 36.6W \quad \& \\ E_{female} &= 580 + 31.1W, \end{aligned} \tag{2.3}$$

but only applicable to the physiological range of observed covariates.

Later FAO assigned Schofield to extend the survey and develop new equations applicable to a larger population. Thus the Schofield database of 7173 subjects were formed. Schofield developed equations for different age groups based on the weight and sex of an individual, and the equations for males and females in 30-60 yr age group are shown in Table 2.1. These equations are adopted by FAO/WHO/UNU in 1985 to propose human energy requirements and are still in use.

Hayter and Henry (1993) reported that the Schofield equations are over predictive in tropical populations. An over estimation of RMR in children was reported by Ho et al. (1988) (Min and Ho 1991) and (Spurr et al. 1992) reported overestimation in children of 2-18 yrs. It was later found that more than 50% of data points ( $> 3000$ ) in Schofield database consists of Italian male subjects between the age 20-50 with high BMR values that elevated the predictions in other populations.

In 2001, due to arising concerns over the applicability of Schofield equations globally, FAO assigned CJK Henry to develop new equations for RMR predictions (Henry 2005). CJK Henry collected 10552 data points from the literature including more number of subjects from the tropics and removed all Italian subjects. Equations for male and female belonging to the 30-60 yrs age group are given in 2.1. The equations were better predictive but not significantly different from the performance of Schofield equations, hence FAO suggested continuing the use of Schofield equation until significantly different new equations are developed. The FAO equations developed over the years on a large dataset used age, sex and weight of the subjects to explain the variation between individuals, but could only explain 70% of the variation.

### **2.1.3.2 Organ-tissue level approaches in RMR studies**

Prediction BMR based on body weight assumes an equal metabolic output of its constituents. Body mass consists of different tissues and organs with different energy usage. The total body mass consists of metabolically less active fat mass (FM) and rest of the mass that is metabolically active, termed as fat-free mass (FFM). Owen et al. (Owen et al. 1986, 1987) in 1986 analysed the relationship of FFM and FM to resting metabolic rate (RMR) in 44 healthy, lean and obese women (16-65 yrs) and 60 lean and obese men. The body composition (FM and FFM) was measured through densitometry and RMR through indirect calorimetry. It was found that BMR is correlated with FFM ( $R > 0.74$ ) but still not a better predictor than the weight ( $R^2=0.8$ ). In 1990, Mifflin et al. (Mifflin et al. 1990a) analysed body composition in a larger dataset of 498 healthy subjects including normal weight (n=264) and obese (n=234) individuals. They reported FFM is the single best predictor of BMR ( $R^2=0.64$ ).

Several models based on body composition were reported later (Cunningham 1980; McNeill et al. 1987; Heymsfield et al. 1988; Ravussin and of genetics 1989; Ravussin et al. 1982; Jensen et al. 1988; Garby et al. 1988; Bernstein et al. 1983; Katch et al. 1990). Frankenfield et al. Frankenfield et al. (2005) analysed all notable predictive equations developed till 2005 and observed Mifflin- St Jeor Mifflin et al. (1990a) equations to be the best predictor for the current population, (within  $\pm 10\%$  of measured BMR in 82% of the data). It might be partly because the body composition of the population has evolved and Mifflin data points are representative of the most recent population. Further, Mifflin-St Jeor data points are heterogeneous consisting of both obese and normal individual with body mass index ranging from 17 to 42. Older RMR studies included only healthy subject data points to develop a standard reference value.

Over the past century, BMR predictive equations have evolved through improvement in equipment, sample collection, and analysis procedures. But, even the most accurate model available (Frankenfield et al. 2005; Mifflin et al. 1990a) could explain only 70-80 percentage of the variation in RMR. The remaining 20 per cent are not found to be explained by anthropometric factors. Since FFM comprises of tissues and organs with different energy utilisation, more insights can be visible from tissue-organ level approaches to understand RMR

Elia (1992) analysed the specific metabolic rate of organs and their relative contribution to the total metabolism. They measured the in vivo metabolic rate of individual tissues by, “making measurements of the arteriovenous concentration difference of oxygen across tissue in conjunction with measurements of blood flow” (Elia 1992). The metabolic rate per unit mass per unit time is defined as the specific metabolic rate of an organ ( $K_i$  values).  $K_i$  values in kcal per kg per day, for the liver is 200, the brain is 240, the heart is 440, the kidney is 440, the muscle mass is 13, the adipose tissue is 4.5, and 12 for other tissues including bone, skin, intestines, glands etc. Gallagher et al. (1998) in 1998 modelled the basal metabolic rate of an organism by summing the metabolic rates of individual organs. Using the dual-energy X-ray absorptiometry (DXA) for FM and FFM, and magnetic resonance imaging (MRI) and echocardiography, they calculated the organ and tissue masses. Further, the metabolic rate of individual organs are calculated as the product of the mass and specific metabolic rate of each organ. The specific metabolic rate of organs are taken from Elia (1992). Resting metabolic rate is measured through indirect calorimetry. The energy expenditure at rest,  $E$  (kJ/kg/day), is given as,

$$E = 1008M_{brain} + 840M_{liver} + 1848M_{heart} + 1848M_{kidneys} + 55M_{skeletalmuscle} + 19M_{adiposetissue} + 50M_{residualmass}, \quad (2.4)$$

where  $M$  is the mass of respective organs. The model is validated in only 13 subjects, but with 92% accuracy between measured and predicted resting metabolic rate. Here the limitation in this approach is the unavailability of an equipment to measure in vivo metabolic rate or specific metabolic rate of an organ.

Wang et al. 2010 reported an over estimation of the specific metabolic rate of organs by 3% in adults ( $n=37$ ) but well correlated in young and middle-age group. Later, Wang et al. 2011 compared Elia’s specific metabolic rate values in 49 men and 57 women aged 29-49 yrs. The measured resting metabolic rate is compared to Gallagher model predictions, and highly correlated with the observed BMR in men ( $R^2=0.87$ ) and women ( $R^2=0.86$ ). So assuming Elia’s reference

values are applicable, Gallagher model can explain more than 86% of the variation in BMR.

Further, the presence of brown adipose tissue (BAT), which has a high metabolic activity, was also reported to be associated with RMR. BAT is a high metabolic active tissue, and due to difficulty in characterising BAT in the adult body, variation in its presence can add up variations in BMR. Saito et al. (2009) reported the presence of BAT in 76 of 1013 and 30 of 959 men they sampled. Cypess et al. (2009) in 2009 reported a substantial amount of BAT in adults. Elia's study report only 13 kcal/kg/day  $K_i$  value for skeletal muscle mass. But, the skeletal muscle mass metabolism is shown to be a major determinant of RMR ( $R^2=0.72$ , after adjusting for FFM, FM, age, sex) by Zurlo et al. (1990) by measuring forearm oxygen uptake and comparing it with BMR measured through indirect calorimetry in 14 subjects (7 men and 7 women).

### 2.1.3.3 Cellular and molecular level approaches

Organ and tissue metabolic rate and RMR can vary between individuals if cellular and molecular determinants of RMR varies. Hence it is important to understand the smaller molecules involved in the metabolic pathways as well. Notable studies on such molecules and cellular components are presented here.

Leptin is a cell signalling protein produced by adipose tissue and its levels in the body is representative of body fat mass. Studies report both positive (Jol et al. 1998; Nicklas et al. 1997; Toth et al. 1997) and negative (Roberts et al. 1997; Kennedy et al. 1997; Nagy et al. 1997) relationship between leptin concentration and BMR. Loos et al. (2006) identified polymorphisms in the leptin receptor gene, LEPR-K6556N correlated with BMR from Single Nucleotide Polymorphism (SNP) analysis of 678 subjects. It was known from the 19th century that thyroid dysfunction leads to changes in BMR. Some studies reported links of circulating thyroxine (T4) and triiodothyronine (T3) (Svendsen et al. 1993; Astrup et al. 1992). Bernstein et al. (Bernstein et al. 1983), Westphal et al. (Bosy-Westphal et al. 2008) and Welle et al. (Welle et al. 1990) reported no dependence of T3 to RMR.

In 2004 Johnstone reported that circulating T4 explained the 25% of the the residual variation of RMR in men after adjusting for FFM,FM and age, but not significant in women. There was 26% unexplained variation in RMR after accounting for FFM, FM, age, within-subject variability and analytic error.

FFM is the main determinants of BMR at the molecular level (Owen et al. 1986; Mifflin et al.

1990a; Owen et al. 1987). Some studies reported fat mass contribution (Nelson et al. 1992) to the BMR but not all (Bogardus et al. 1986; Segal et al. 1987).

A study by Larsen et al. (2011) looked at the oxygen affinity of mitochondrial respiration ( $p50_{mito}$ ), which is defined as the the partial pressure of oxygen ( $pO_2$ ) at which mitochondrial oxygen affinity is half maximal. By measuring  $p50_{mito}$  values in isolated skeletal muscle mitochondria and RMR through indirect calorimetry in 14 subjects, they reported high correlation ( $R^2=0.66$ ) between  $p50_{mito}$  and BMR. They reported no correlation of BMR to mitochondrial density or proton leak through the membrane. And partial inhibition of cytochrome-C-oxidase (COX) lead to 5-fold increase in  $p50_{mito}$  suggesting a possible regulatory mechanism of  $p50_{mito}$  by COX.

Westphal et al. 2008 reported a significant correlation between RMR and metabolic risk factors. They analysed RMR and plasma insulin, C-reactive protein, glucose and blood pressure, in 149 families. Residual variation in RMR after adjusting FFM, FM, sex, age is found to be correlated with blood pressure, insulin resistance, plasma insulin and glucose concentrations ( $R^2 = 0.14 - 0.31$ ). Westphal also reported heritability for thyroid hormones thyrotropin, triiodothyronine, and thyroxine. A more extensive review on metabolic pathways and their components are required to understand more on cellular and molecular determinants of RMR.

#### 2.1.4 RMR during childhood and adolescence

RMR in children vary non-linearly with age. It is often thought to be because of the non-linear changes in growth, especially due to the uneven changes in mass of organs and tissues of varying metabolic rate. For a detailed review on the RMR in children and adolescents please refer to Son'kin and Tambovtseva (2012). Urlacher (2023) outlines the current views of energetics of childhood majorly from an evolutionary perspective and it's implications on human variation and health.

### 2.2 Growth centiles and growth curves

Height growth centiles (or percentiles) describe the distribution of height (as me(di)an and the dispersion) with respect to age in a population. They are drawn as percentile curves defined by the percentage of the population below each curve. As early as 1885, Galton utilised centile



curves to show an account of the distribution of anthropometric factors (Galton 1885). Centiles have since been used widely, especially in the form of charts for the clinical assessment of growth related disorders. Widely used growth charts include the World Health Organisation's (WHO) infant growth charts from birth to 5 years (World Health Organization 2006; de Onis 2006).

It is crucial to differentiate between growth *curves* and *centiles*. While they appear to serve identical functions clinically, the underlying principles of their construction are entirely different. Growth centiles are developed using cross-sectional data, which are independent measurements obtained once per subjects, together spanning the entire age range. Development of growth curves involves longitudinal data, which comprises consecutive measurements taken for each child over a prolonged duration. While growth curves describe the longitudinal trend of height (individual-specific or the average population curve), growth centiles provides an account of the height distribution (average trend and the spread around it). The methods of their construction differ owing to the fact that cross-sectional data consists independent samples, whereas the longitudinal data are correlated.

Developing growth centiles requires modelling the height progression with age as well as the spread of height distribution at every age. Early growth centiles computed means and standard deviation's (SD) of height in narrow age ranges and the resulting average growth curve was characterised using smoothing methods to model the horizontal progression. To describe evenly spaced centiles to capture the spread of data, initial methods assumed height to be normally distributed throughout childhood. This enables to specify spacing between the curves, based only on the mean and SD. However, the height is not found to be distributed normally during childhood. Moreover, the height data is found to be high skewed towards the later stages of childhood. The Lambda-Mu-Sigma (LMS) method due to Cole (1988), later modified in Cole and Green (1992), overcomes this by considering a transformations of the response variable which then follows a normal distribution. The LMS method parameterise the median, coefficient of variation and the skewness of the distribution and uses splines to smooth the centile curves. A most recent method, Generalized Additive Model for Location, Scale and Shape (GAMLSS), extends the LMS method to include a kurtosis parameter as well (Rigby and Stasinopoulos 2004, 2005, 2006).

Describing the shape of the individual or the mean growth curves were first attempted using simple polynomial or logistic models. Jenss and Bayley described growth in the first six years using a four-parameter logistic model (Jenss and Bayley 1937), and Gompertz curves (Benjamin 1825) were used to describe growth from six years to maturity. A triple-logistic model (Bock

and Thissen 1976) was able to parameterize growth from birth to maturity by modifying a previous double-logistic model (Bock et al. 1973), which partitioned growth into pre-pubertal and adolescent components (El Lozy 1978). The family of polynomial models for infant growth were proposed by Royston and Altman (1994), but they lack flexibility in shape compared to logistic model. The Infant–Childhood–Pubertal (ICP) model by Karlberg (Karlberg 1989) considers three separate phases of growth. The five-parameter Preece-Baines model that describes growth from birth to adulthood is still widely used (Preece and Baines 1978). The parameters estimated in these methods are shown to have physiological interpretability.

Other attempts to model non-linear growth curves were based on splines and kernel estimations (Gasser et al. 1984; Goldstein 1991; Altman 1992). Here, the optimal number of parameters to describe the shape and the variation of the curves are estimated from the data at hand and thus provides high flexibility but lacks biological interpretability. In the widely used mixed-effects models, a few parameters are fixed for the entire population, called the fixed effects, and some parameters are allowed to vary between the individual, called the random effects. They are able to capture the individual variation as well the population trend in a concise manner. A recent mixed-effects model named SuperImposition by Translation And Rotation (SITAR) by (Cole et al. 2010) describes individual curves as a transformation of the average growth curves using only three random effects parameters. These parameters are also found to be physiologically relevant.

Growth curves are used clinically to assess landmark physiological events during growth and development. A major event during childhood and adolescence is the pubertal growth spurt; this is prominent in the height velocity curves. Pubertal spurts are generally characterised by two components: one is the peak height velocity (PV) and the other is the age at peak height velocity (APV), with a very distinct shape in boys and girls (Tanner 1962). A smaller growth spurt is also observed before puberty in many children between 5 to 9 years (mid-growth spurt), (Tanner 1962; Tanner and Cameron 1980). Recently, Cole (2020) reported similar growth spurt in the Avon longitudinal study of parents and children (ALSPAC) (Boyd et al. 2012), in a subgroup of boys with late APV at 9 years. However, it is not clear whether this is a model artefact or a biologically relevant observation. While a few studies report a mid-growth spurt in both sexes, they are often more pronounced in boys (Molinari et al. 1980). It is useful to note that while pubertal growth peaks are well characterised by SITAR and GAMLSS (Cole 2020; Kelly et al. 2014; de Onis et al. 2011; Khadilkar et al. 2019), mid-growth spurt has not been fitted with these

models previously.

In this thesis, we construct growth curves and charts using the SITAR (Cole et al. 2010) and GAMLSS (Rigby and Stasinopoulos 2005) models using a longitudinal height data set of Indian children and adolescents. The SITAR model describes the shape of growth curve and is geared up to detect subtle changes in growth. The GAMLSS model describes how we can characterise the distribution of height in the given population.

## Bibliography

- Altman, N. S. (1992). An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *American Statistician*, 46(3):175–185.
- Astrup, A., Buemann, B., Christensen, N. J., et al. (1992). The contribution of body composition, substrates, and hormones to the variability in energy expenditure and substrate utilization in premenopausal women. *J Clin Endocrinol Metab*, 74:279–86.
- Aub, J. C. and Du Bois, E. F. (1917). Clinical calorimetry: Nineteenth paper the basal metabolism of old men. *Archives of Internal Medicine*, XIX(5\_II):823–831.
- Aub, J. C. and Du Bois, E. F. (1917a). Clinical calorimetry: nineteenth paper the basal metabolism of old men. *Archives of Internal Medicine*, XIX(5\_II):823–831. \_eprint: [https://jamanetwork.com/journals/jamainternalmedicine/articlepdf/653834/archinte\\_xix\\_5\\_ii\\_001.pdf](https://jamanetwork.com/journals/jamainternalmedicine/articlepdf/653834/archinte_xix_5_ii_001.pdf).
- Aub, J. C. and Du Bois, E. F. (1917b). Clinical calorimetry: twenty-first paper the basal metabolism of dwarfs and legless men with observations on the specific dynamic action of protein. *Archives of Internal Medicine*, XIX(5\_II):840–864. \_eprint: [https://jamanetwork.com/journals/jamainternalmedicine/articlepdf/653836/archinte\\_xix\\_5\\_ii\\_003.pdf](https://jamanetwork.com/journals/jamainternalmedicine/articlepdf/653836/archinte_xix_5_ii_003.pdf).
- Bedale, E. M. (1923). Energy expenditure and food requirements of children at school. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 94(662):368–404.
- Benjamin, G. (1825). XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. F. R. S. &c. *Philos. Trans. R. Soc. Lond.*, 115:513–583.

- Bernstein, R. S., Thornton, J. C., Yang, M. U., et al. (1983). Prediction of the resting metabolic rate in obese patients. *Am J Clin Nutr*, 37:595–602.
- Bock, R. and Thissen, D. (1976). Fitting multi-component models for growth in stature. In *Proceedings of the 9th international biometric conference*, volume 1, pages 431–442.
- Bock, R. D., Wainer, H., Petersen, A., Thissen, D., Murray, J., and Roche, A. (1973). A parameterization for individual human growth curves. *Human Biology*, pages 63–80.
- Bogardus, C., Lillioja, S., Ravussin, E., et al. (1986). Familial dependence of the resting metabolic rate. *N Engl J Med*, 315:96–100.
- Bosy-Westphal, A., Wolf, A., Bührens, F., Hitze, B., Czech, N., Mönig, H., M., and J., M. (2008). Familial influences and obesity-associated metabolic risk factors contribute to the variation in resting energy expenditure: the kiel obesity prevention study 1-3. *Am J Clin Nutr (Vol., 87)*.
- Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S., and Davey Smith, G. (2012). Cohort Profile: The ‘Children of the 90s’—the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, 42(1):111–127.
- Boyd, E. (1935). The growth of the surface area of the human body. *Institute of Child Welfare, monograph series, Minneapolis: University of Minnesota Press*, 10.
- Breitman, M. (1932). Eine vereinfachte methodik der körperoberfläche bestimmung. *Scritti Biologi*, 7:395–8.
- Brody, S. and Lardy, H. A. (1946). Bioenergetics and Growth. *The Journal of Physical Chemistry*, 50(2):168–169.
- Cole, T. J. (1988). Fitting Smoothed Centile Curves to Reference Data. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 151(3):385–418.
- Cole, T. J. (2020). Tanner’s tempo of growth in adolescence: recent SITAR insights with the Harpenden Growth Study and ALSPAC. *Ann. Hum. Biol.*, 47(2):181–198.
- Cole, T. J., Donaldson, M. D. C., and Ben-Shlomo, Y. (2010). SITAR—a useful instrument for growth curve analysis. *Int. J. Epidemiol.*, 39(6):1558.

- Cole, T. J. and Green, P. J. (1992). Smoothing reference centile curves: The lms method and penalized likelihood. *Stat. Med.*, 11(10):1305–1319.
- Cunningham, J. J. (1980). A reanalysis of the factors influencing basal metabolic rate in normal adults. *American Journal of Clinical Nutrition*, 33(11):2372–2374.
- Cypess, A. M., Lehman, S., Williams, G., Tal, I., Rodman, D., and Goldfine, A. B. (2009). ... kahn, c. *Identification and importance of brown adipose tissue in adult humans*, 360:15.
- De B. Weir, J. B. (1949). New methods for calculating metabolic rate with special reference to protein metabolism. *J. Physiol.*, 109(1-2):1–9.
- de Onis, M. (2006). WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatr.*, 95(S450):76–85.
- de Onis, M., Siyam, A., Borghi, E., Onyango, A. W., Piwoz, E., and Garza, C. (2011). Comparison of the World Health Organization growth velocity standards with existing US reference data. *Pediatrics*, 128(1):18–26.
- Du Bois, D. and Du Bois, E. F. (1989). A formula to estimate the approximate surface area if height and weight be known. 1916. *Nutrition (Burbank, Los Angeles County, Calif.)*, 5(5):303–311; discussion 312–313.
- El Lozy, M. (1978). A critical analysis of the double and triple logistic growth curves. *Ann. Hum. Biol.*, 5(4):389–394.
- Elia, M. (1992). Energy expenditure in the whole body. In *Kinney JM*, pages 19–49. Energy metabolism: tissue determinants and cellular collaries. : Raven Press Ltd, New York.
- Energy, J. F. E. C. o., Protein Requirements (1981: Rome, I., Food, Nations, A. O. o. t. U., Organization, W. H., and University, U. N. (1985). Energy and protein requirements : report of a Joint FAO/WHO/UNU Expert Consultation [held in Rome from 5 to 17 October 1981]. Pages: ita published by: Milan : Istituto Scotti Bassani Series: World Health Organization technical report series ; no. 724.
- FAO/WHO/UNU (1985). *Energy and protein requirements : report of a Joint FAO/WHO/UNU Expert Consultation*. World Health Organization.
- Frankenfield, D., Roth-Yousey, L., and Compher, C. (2005). Comparison of predictive equations

- for resting metabolic rate in healthy nonobese and obese adults: A systematic review. *Journal of the American Dietetic Association*, 105(5):775–789.
- Gallagher, D., Belmonte, D., Deurenberg, P., Wang, Z., Krasnow, N., Pi-Sunyer, F. X., and Heymsfield, S. B. (1998). Organ-tissue mass measurement allows modeling of REE and metabolically active tissue mass. *American Journal of Physiology-Endocrinology and Metabolism*.
- Galton, F. (1885). Anthropometric per-centiles. *Nature*, 31(793):223–225.
- Garby, L., Garrow, J. S., Jorgensen, B., et al. (1988). Relation between energy expenditure and body composition in man: specific energy expenditure in vivo of fat and fat free tissue. *Eur J Clin Nutr*, 42:301–5.
- Gasser, T., Muller, H. G., Kohler, W., Molinari, L., and Prader, A. (1984). Nonparametric Regression Analysis of Growth Curves. *Ann. Stat.*, 12(1):210–229.
- Goldstein, H. (1991). Nonlinear multilevel models, with an application to discrete response data. *Biometrika*, pages 45–51.
- Harris, J. A. and Benedict, F. G. (1918). A Biometric Study of Human Basal Metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, 4(12):370.
- Haycock, G. B., Schwartz, G. J., and DH., W. (1978). Geometric method for measuring body surface area a height-weight formula validated in infants, children and adults. *Minneapolis: University of Minnesota Press*, 93:62–6.
- Hayter, J. and Henry, C. (1993). Basal metabolic rate in human subjects migrating between tropical and temperate regions – a longitudinal study and review of previous work. *European Journal of Clinical Nutrition*, 47:724–34.
- Henry, C. J. K. (2005). Basal metabolic rate studies in humans: measurement and development of new equations. *Public Health Nutrition*, 8(7a):1133–1152.
- Heymsfield, S. B., Hoff, R. D., Gray, T. F., Galloway, J., and diseases, C. K. H. (1988). In: Kinney jm. In Hill, G. L. and Owen, O. E., editors, *Jeejeebhoy KN*. WB Saunders477—530, Nutrition and metabolism in patient care. Philadelphia.
- Ho, Z. C., Zi, H. M., Bo, L., and Ping, H. (1988). Energy expenditure of preschool children in a subtropical area. *World Review of Nutrition and Dietetics*, 57:75–94.

- Isacksson, B. A. (1958). simple formula for the menial arithmetic of the human body surface area. *Scand J Clin Lab Inv*, 10:283–9.
- Jensen, M. D., Braun, J. S., and HM., V. R. M. (1988). Measurements of body potassium with a whole body counter: relationship between lean body mass and resting energy expenditure. *Maw Clin Proc*, 63:864–68.
- Jenss, R. M. and Bayley, N. (1937). A mathematical method for studying the growth of a child. *Human Biology*, 9(4):556.
- Jol, J., Vahl, N., Dall, R., and JS., C. (1998). Resting metabolic rate in healthy adults: relation to growth hormone status and leptin levels. *Metabolism*, 47:1134–9.
- Karlberg, J. (1989). A biologically-oriented mathematical model (ICP) for human growth. *Acta Paediatr. Scand. Suppl.*, 350(70-94):;.
- Katch, V. L., Marks, C. C., Becque, M. D., Moorehead, C., and Rocchini, A. (1990). Basal metabolism of obese adolescents: evidence for energy conservation compared to normal and lean adolescents. *Am J Hum Biol*, 2:543–51.
- Kelly, A., Winer, K. K., Kalkwarf, H., Oberfield, S. E., Lappe, J., Gilsanz, V., and Zemel, B. S. (2014). Age-Based Reference Ranges for Annual Height Velocity in US Children. *J. Clin. Endocrinol. Metab.*, 99(6):2104–2112.
- Kennedy, A., Gettys, T. W., Watson, P., et al. (1997). The metabolic significance of leptin in humans: gender-based differences in relationship to adiposity, insulin sensitivity, and energy expenditure. *J Clin Endocrinol Metab*, 82:1293–1300.
- Khadilkar, V., Khadilkar, A., Arya, A., Ekbote, V., Kajale, N., Parthasarathy, L., Patwardhan, V., Phanse, S., and Chiplonkar, S. (2019). Height Velocity Percentiles in Indian Children Aged 5-17 Years. *Indian Pediatr.*, 56(1):23–28.
- Kleiber, M. (1932). Body size and metabolism. *Hilgardia*, 6(11):315–353.
- Larsen, F. J., Schiffer, T. A., Sahlin, K., Ekblom, B., Weitzberg, E., and Lundberg, J. O. (2011). Mitochondrial oxygen affinity predicts basal metabolic rate in humans. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 25(8):2843–2852.
- Loos, R. J. F., Rankinen, T., Chagnon, Y., Tremblay, A., Pérusse, L., and Bouchard, C. (2006).

- Polymorphisms in the leptin and leptin receptor genes in relation to resting metabolic rate and respiratory quotient in the québec family study. *International Journal of Obesity*, 30(1):183–190.
- McMurray, R. G., Soares, J., Caspersen, C. J., and McCurdy, T. (2014). Examining variations of resting metabolic rate of adults: a public health perspective. *Med. Sci. Sports Exerc.*, 46(7):1352–1358.
- McNeill, G., Rivers, G., Gpw, P., de Britto, J. J., and Abel, R. (1987). Basal metabolic rate of Indian men: no evidence of metabolic adaptation to a low plane of nutrition. *Hum Nutr Clin Nutr*, 41C:473–84.
- Mifflin, M. D., Jeor, S., T., S., Hill, L. a., Scott, B. J., Daugherty, S. a., and Koh, Y. O. (1990a). A new predictive equation in healthy individuals<sup>3</sup> for resting energy. *American Journal of Clinical Nutrition*, 51:241–247.
- Mifflin, M. D., Jeor, S. T. S., Hill, L. A., Scott, B. J., Daugherty, S. A., and Koh, Y. O. (1990b). A new predictive equation for resting energy expenditure in healthy individuals. *American Journal of Clinical Nutrition*, 51(2):241–247.
- Min, Q. and Ho, Z.-C. (1991). The basal metabolism rate of adolescent girls in the sub-tropical areas of china. *Acta Nutrimenta Sinica*, 13:252–8.
- Molinari, L., Largo, R. H., and Prader, A. (1980). Analysis of the growth spurt at age seven (mid-growth spurt). *Helv. Paediatr. Acta*, 35(4):325–334.
- Nagy, T. R., Gower, B. A., Shewchuk, R. M., and leptin and, G. M. S. (1997). and energy expenditure in children. *J Clin EndocrinolMetab*, 82:4149– 53.
- Nelson, K. M., Weinsier, R. L., Long, C. L., and Schutz, Y. (1992). Prediction of resting energy expenditure from fat-free mass and fat mass. *Am J Clin Nutr*, 56:848–56.
- Nicklas, B. J., Toth, M. J., Goldberg, A. P., and ET., P. (1997). Racial differences in plasma leptin concentrations in obese postmenopausal women. *J Clin Endocrinol Metab*, 82:315–7.
- Owen, O. E., Holup, J. L., D’Alessio, D. A., Craig, E. S., Polansky, M., Smalley, K. J., Kavle, E. C., Bushman, M. C., Owen, L. R., and Mozzoli, M. A. (1987). A reappraisal of the caloric requirements of men. *The American Journal of Clinical Nutrition*, 46(6):875–885.
- Owen, O. E., Kavle, E., Owen, R. S., Polansky, M., Caprio, S., Mozzoli, M. A., Kendrick, Z. V.,



- Bushman, M. C., and Boden, G. (1986). A reappraisal of caloric requirements in healthy women. *The American Journal of Clinical Nutrition*, 44(1):1–19.
- Preece, M. A. and Baines, M. J. (1978). A new family of mathematical models describing the human growth curve. *Ann. Hum. Biol.*, 5(1):1–24.
- Quenouille, M. H., Boyne, A. W., Fisher, W. B., and Leitch, U. (1951). Statistical studies of recorded energy expenditure of man. part i basal metabolism related to sex, stature, age, climate and race. *Commonwealth Bureau of Animal Nutrition. Technical Communication No.*, 17.
- Ravussin, E., Burnand, B., Schutz, Y., and Jequier, E. (1982). Twenty four hour energy expenditure and resting metabolic rate in obese, moderately obese and control subjects. *Am J Clin Nutr*, 35:566–73.
- Ravussin, E. and of genetics, B. C. R. (1989). age and physical fitness to daily energy expenditure. *Am J Clin Nutr*, 49:968–75.
- Rigby, R. A. and Stasinopoulos, D. M. (2004). Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. *Stat. Med.*, 23(19):3053–3076.
- Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(3):507–554.
- Rigby, R. A. and Stasinopoulos, D. M. (2006). Using the Box-Cox t distribution in GAMLSS to model skewness and kurtosis. *Statistical Modelling*, 6(3):209–229.
- Roberts, S. B., Nicholson, M., Staten, M., et al. (1997). Relationship between circulating leptin and energy expenditure in adult men and women aged 18 years to 81 years. *Obes Res*, 5:459–63.
- Royston, P. and Altman, D. G. (1994). Regression using fractional polynomials of continuous covariates: Parsimonious parametric modelling. *Journal of the Royal Statistical Society Series C*, 43(3):429–453.
- Saito, M., Okamatsu-ogura, Y., Matsushita, M., Watanabe, K., Yoneshiro, T., Nio-kobayashi, J., and Tsujisaki, M. (2009). High incidence of metabolically active brown adipose effects of cold exposure and adiposity. *Diabetes*, 58:1526–1531.

- Schofield, W. N. (1985). Predicting basal metabolic rate, new standards and review of previous work. *Human nutrition. Clinical nutrition*, 39 Suppl 1:5–41.
- Segal, K. R., Gutin, B., Albu, J., and FX., P.-S. (1987). Thermal effects of food and exercise in lean and obese men of similar lean body mass. *mJ Physiol*, 252:E110–7.
- Son'kin, V. and Tambovtseva, R. (2012). Energy metabolism in children and adolescents. *Bioenergetics*, pages 121–142.
- Spurr, G. B., Reina, J. C., and Hoffmann, R. (1992). Basal metabolic rate of colombian children 2–16 years of age: ethnicity and nutritional status. *American Journal of Clinical Nutrition*, 56:623–9.
- Svendsen, O. L., Hassager, C., and Christiansen, C. (1993). Impact of regional and total body composition and hormones on resting energy expenditure in overweight postmenopausal women. *Metabolism*, 42:1588–91.
- Tanner, J. M. (1962). *Growth at adolescence*. 2nd Edition, Blackwell Scientific Publications, Oxford.
- Tanner, J. M. and Cameron, N. (1980). Investigation of the mid-growth spurt in height, weight and limb circumferences in single-year velocity data from the London 1966–67 growth survey. *Ann. Hum. Biol.*, 7(6):565–577.
- Toth, M. J., Gottlieb, S. S., Fisher, M. L., Ryan, A. S., Nicklas, B. J., and ET., P. (1997). Plasma leptin concentrations and energy expenditure in heart failure patients. *Metabolism*, 46:450–3.
- Urlacher, S. S. (2023). The energetics of childhood: Current knowledge and insights into human variation, evolution, and health. *American Journal of Biological Anthropology*. Publisher: John Wiley & Sons, Ltd.
- von Schelling, H. (1954). Mathematical deductions from empirical relations between metabolism, surface area and weight. *Ann N Y Acad Sci*, 56:1143–64.
- Wang, Z., Heymsfield, S. B., Ying, Z., Pierson, Jr., R. N., Gallagher, D., and Gidwani, S. (2010). A Cellular Level Approach to Predicting Resting Energy Expenditure: Evaluation of Applicability in Adolescents. *American journal of human biology : the official journal of the Human Biology Council*, 22(4):476.

- Wang, Z., Ying, Z., Bosy-Westphal, A., Zhang, J., Heller, M., Later, W., and Müller, M. J. (2011). Evaluation of specific metabolic rates of major organs and tissues: comparison between men and women. *American Journal of Human Biology : The Official Journal of the Human Biology Council*, 23(3):333–338.
- Welle, S., Jozefowicz, R., and Statt, M. (1990). Failure of dehydroepiandrosterone to influence energy and protein metabolism in humans. *J Clin Endocrinol Metab*, 71:1259–64.
- World Health Organization (2006). WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development. *World Health Organization*, pages 1–312.
- Zurlo, F., Bogardus, C., Ravussin, E., Zurlo, F., Larson, K., Bogardus, C., and Ravussin, E. (1990). Skeletal muscle metabolism is a major determinant of resting energy expenditure. *J Clin Invest*, 86(5):1423–1427.

## **CHAPTER 3**

# **Modelling resting metabolic rate in Indian children**

## Commentary

In this chapter, we introduce two models describing the lower RMR/BM observed in the Indian population. Model 1 predicts a lowered relative mass of four major organs explaining the observed RMR/BM in the Indian population. The theoretical predictions for liver and kidney mass were assessed in a pilot validation study. When tested for significant difference between the Indian and Caucasian population, although the relative kidney mass was lower, the relative liver mass was not. Further, we present an alternate model (Model 2) to show that the Indian data can be explained by the changes in body composition alone. In this section we discuss some of the comments raised by an external examiner in detail.

The pilot validation study for Model 1 was carried out in children aged 6 to 8 years by measuring their liver and kidney mass using ultrasonography. Prior research indicates that when individuals experience early-life stress during critical periods of growth, for instance, during fetal growth due to maternal malnutrition or oxygen deprivation, there's often a hierarchical trade-off observed in organ growth. For instance, growth of organs such as liver, kidney or pancreas are sacrificed to spare the growth of critical organs like the brain (Hales and Barker 1992; Latini et al. 2004; Barker 2004; Baker et al. 2010). This trade-off is generally irreversible later, and becomes a part of the physiological phenotype. For example, the number of nephrons in the kidneys is established at birth, irrespective of the later growth (Luyckx and Brenner 2005). This hierarchical preservation of tissues and organs under stress, which may be pre-natal in origin, can be observed during early childhood and adolescence as well. Thus, in our pilot study, we examined the compromise in growth of sensitive organs like the liver and kidney between Indian and Caucasian children. Since the evidence so far suggests that there is a protective mechanism in place for critical organs, excluding the brain from the validation study is unlikely to substantially impact the conclusions drawn in the current study.

Further, from Figure 3.1, we observe that the relative organ mass of the kidney and liver remains relatively constant from 6 to 20 years of age, which suggests that the results from the pilot study with subjects aged 6 to 8 years (RMR-USG dataset) could be extrapolated to the older age groups (9 to 18 years) in the MCS dataset used to develop the model.

Relative cellularity is the ratio of cellularity (TBK/FFM) in children to that of adults. Fomon et al. (1982) suggests that the TBK values in children younger than 10 years differed from the

adult TBK values. However, Pierson (2005) (based on the table provided by Wang (2012)) suggests that the TBK levels for children older than 11 years and adults remained the same. Based on the data compiled by Wang (2012), it can be observed that the relative cellularity rises from 0.83 to 0.99 throughout childhood, eventually matching the adult level (i.e., 1.0) in early adolescence.

The organ mass values used in our study were obtained from Wang (2012), which, regrettably, did not provide additional characteristics of the subjects. Unfortunately, despite our diligent efforts, we were unable to access the original article by Altman and Dittmer (1962).

With respect to the Model 2, a lower RMR/BM in Indian children could arise from different proportions of lean to fat mass in the Indian population. For instance, previous studies have shown differences in the body composition of the Indian population, with a tendency for Indian adults to exhibit elevated levels of overall and abdominal body fat in individuals who may outwardly appear 'thin'. This phenotype could stem from prenatal factors and continue through childhood and adolescence, potentially increasing the risk for insulin resistance syndrome in the Indian population (Lubree et al. 2002; Yajnik et al. 2003).

From an evolutionary point of view, the higher adiposity could be a predictive adaptive response. Theories proposed by Hales and Barker (1992) and Wells (2011) suggest that this could be an adaptive response that emerges in expectation of an unfavorable adult environment that has an influence on fetal growth, which is likely to be imprinted on development during childhood.

An altered body composition with higher fat reserves is expected to play a protective role in nutrition poor situations; however, is found to have a deleterious effect in the modern obesogenic environments leading to higher risk for metabolic and cardiovascular diseases (Hales and Barker 1992; Wells 2011). For instance, Indian babies who were born small but relatively fat are found to have increased risk for insulin resistance and cardiovascular risks (Yajnik 2002).

However, future studies need to evaluate if there are any protective aspects of the fat reserves on body composition in the context of modern nutritional sources with low quality fat. Furthermore, the effect of the type of fat tissue, brown adipose tissue versus white adipose tissue also needs to be considered to ascribe a protective or deleterious role of the altered body composition.

*Published as and adapted from:*

**Sandra Aravind Areekal**, Anuradha Khadilkar, Neha Kajale, Arun S. Kinare, and Pranay Goel. 2023. “Two Novel Models Evaluating the Determinants of Resting Metabolic Rate in Indian Children”. *Human Biology and Public Health* 3 (March). <https://doi.org/10.52905/hbph2022.3.55>

### 3.1 Introduction

Malnutrition in developing countries, such as India, is often paradoxically characterized the simultaneous prevalence of undernutrition and rising overweight and obesity in children and adolescents (WHO 2016; NCD Risk Factor Collaboration (NCD-RisC) et al. 2017). One approach to studying malnourishment is through assessing the energy intake and expenditure of a population. Energy expenditure, in particular, is predominantly determined by the physiology of the individual and varies significantly, both within and across populations (Henry 2005; Johnstone et al. 2005). For instance, a 200 kcal per day difference in energy intake was sufficient to explain the excess weight of US children in 2003-2006 compared to 1976-1980 (Hall et al. 2013). It is necessary to understand factors leading to variation in energy expenditure to create personalised interventions to tackle the double burden of malnutrition. We note that the World Health Organisation’s (WHO) recommendations for energy requirements (FAO/WHO/UNU 2004) are based on studies that overestimate energy expenditure by 12% in Indian population (Henry 2005). Here we are interested in developing models that accurately describe the (resting) energy expenditure in Indian children.

A primary component of energy expenditure is the resting energy expenditure (REE) or the resting metabolic rate (RMR), which measures the energy required to maintain the vital body functions at rest. RMR is measured through direct or indirect calorimetry (De B. Weir 1949) under standard conditions such as in the post-absorptive state, in wakefulness, in the absence of any physical activity and diseases, minimal emotional disturbance and in a thermoneutral environment (22-26 °C). Phenomenological models developed on a sample population are frequently used to estimate RMR. A large number of regression models for RMR have been based on anthropometric and body composition factors for nearly a century (FAO/WHO/UNU 1985; Aub and Du Bois 1917; Cunningham 1980; Harris and Benedict 1918; Henry 2005; Katch et al. 1990; McMurray et al. 2014; Kleiber 1932; Bedale 1923; Owen et al. 1986, 1987; Schofield 1985; Mifflin et al. 1990). These models find that fat free mass (FFM) is the single largest predictor of RMR, followed by

fat mass (FM), age, and sex. However, RMR is found to be highly variable between individuals in a population (Henry 2005; Johnstone et al. 2005). Overall models based on body composition have been of limited success, as they are able to explain only about 60 – 80% variation in RMR.

An alternate strategy for modelling is to challenge the assumption that the body mass is metabolically homogeneous, as is inherent in predicting RMR from linear models of FFM or body mass. FFM or body mass is composed of organs and tissues of varying metabolic activity, which together contribute to whole-body RMR. Gallagher et al. (1998) partition RMR as the sum of metabolic rates of a number of major organs and tissues constituting the body mass. The metabolic rates of individual organs were calculated as the product of measured organ mass and the metabolic rate per unit mass (specific metabolic rate) of each organ, which was estimated in vivo by Elia (1992). The Gallagher model was able to explain 80-98% variation of RMR in several studies in adults (Bosy-Westphal et al. 2004, 2008; Wang et al. 2001, 2005, 2010a; Wang 2012; Wang et al. 2010b; Müller et al. 2011). However, the Gallagher model was found to under-predict RMR in children (Wang et al. 2010a; Hsu et al. 2003). Wang (2012) modified the Gallagher model to study how RMR/BM varies in children from birth to adulthood, and described the mean RMR/BM ( $R^2 = 0.99$ ) in a reference Caucasian dataset (Talbot 1938). Here we ask if the Wang model can describe RMR/BM in an Indian population?

Studies on metabolic rates in Indian children are scarce (Patil and Bharadwaj 2013; Swaminathan et al. 2018). Predictive equations developed for Caucasian populations (Harris and Benedict 1918; FAO/WHO/UNU 1985) have been reported to overpredict metabolic rates in Indian adults (Soares et al. 1998; Henry 2005), however, they continue to be used to predict RMR in Indian children (Srivastava et al. 2017; Esht et al. 2018; Indian Council of Medical Research (ICMR) 2010). Previous studies in Indian adults (Mason and Benedict 1931; Mason et al. 1963; Mukherjee and Gupta 1931; Krishnan and Vareed 1932; Rahman 1936; Rajagopal 1938; Niyogi et al. 1939; Kumar et al. 1961) have shown that the measured RMR per unit body surface area in Indian population is 5-18% lower than the Harris-Benedict (Harris and Benedict 1918) standards. On the other hand, Soares et al. (1998) has reported no significant difference in RMR adjusted for FFM in males and RMR adjusted for FFM and FM in females, between 18 to 30 years old Indian and Australian population. Moreover, Soares et al. (1998) had observed a *higher* RMR/FFM in Indian population than the Australian population; and the reason was speculated to be due to a higher proportion of organ mass within FFM compared to muscle mass, but this has not been verified. There is a clear absence of literature on RMR in current Indian



population. We study the influential model of RMR/BM due to Wang in Caucasian children closely to understand the determinants of RMR in Indian children.

In our study, a naive application of the Wang model clearly overestimates the mean RMR/BM observed in Indian children. We assess two major modifications of the model aimed at revealing the mechanistic basis of the lower RMR/BM. We first calibrate the relative masses of the four major organs to the observed RMR/BM, followed by a pilot study to validate organ mass predictions. Organ sizes were not found to be uniformly small, as predicted by model fits. Next we vary the residual mass, to show that this can equivalently explain whole-body RMR/BM. In other words, our paper re-evaluates the role of the relative mass of four major organs and the metabolic contribution of residual mass in determining RMR/BM in an Indian population. We conclude that either model is useful as phenomenological descriptions of RMR varying with age in Indian children. However, identifying the physiological determinants of variation in RMR continues to be an elusive problem.

## 3.2 Methods

### 3.2.1 Datasets

The following datasets were used in the study:

#### 3.2.1.1 Multi-Centre Study (MCS) dataset

MCS is a dataset on 495 healthy school going children (235 girls) aged 9 to 19 years from multiple centres in India, which is a subset of data collected as a part of a previous study (Khadilkar et al. 2019). Anthropometric, body composition and metabolic variables such as the height, weight, fat mass (FM), fat-free mass (FFM) and RMR of the subjects were measured. Fitmate GS (COSMED Srl - Italy) was used to measure RMR by indirect calorimetry. Fitmate GS has been previously validated in healthy adults by Nieman et al. (2006) and Vandarakis et al. (2013). The machine was routinely calibrated according to manufacturer recommendations, and automatic oxygen sensor calibration was carried out before each measurement. Throughout the measurement, the child remained seated, and he/she was asked to relax, and it was ensured that the child remained awake. The test was considered complete after achieving steady state. Body composition was assessed using Bioelectrical Impedance Analyzer (BIA; Tanita Model

BC-420MA), and the child was asked to void before the measurement (Chiplonkar et al. 2017; Kyle et al. 2004). The physical characteristics of the subjects are given in Table. 3.1.

Written consent was obtained from parents of the children and subjects above 18 years, and assent was obtained from children above 7 years. Deidentified data were used for all the analyses. Ethics permission for conducting this multicentric study was granted by the Ethics Committee, Jehangir Clinical Development Centre Pune. A waiver for secondary data analysis was issued by the Ethics Committee for Human Research at the Indian Institute of Science Education and Research Pune.

### 3.2.1.2 RMR-USG dataset

In this study, we measured anthropometry, RMR and organ mass (liver and kidney) of nine healthy girls and boys in the age group 6 to 8 years recruited from a school in Western India. The age group 6 to 8 years was selected so that variation in RMR due to pubertal growth spurt can be avoided. Written consent was obtained from parents of the children. Deidentified data were used for all the analyses. RMR is measured using indirect calorimetry (Fitmate GS, COSMED Srl - Italy) under the standard conditions (see above). The liver and kidney volume in the subjects were measured using ultrasonography (Voluson P8 BT 16, GE Healthcare). The liver volume was examined in the supine position and kidney volume in lateral decubitus position. The measurements were taken during deep inspiration. The measured organ volume was converted to mass as density  $\times$  volume. The density of liver and kidney in the Indian population is assumed to be 1.162 and 1.05 ( $kg/cm^3$ ) respectively (Menzel et al. 2009). A summary of RMR-USG dataset is given in Table 3.2.

The MCS (3.2.1.1) and RMR-USG (3.2.1.2) studies were carried out as per relevant guidelines and regulations.

Variables	Boys			Girls		
	<i>n</i>	Mean $\pm$ SD	Range	<i>n</i>	Mean $\pm$ SD	Range
Age ( <i>years</i> )	260	13 $\pm$ 2	9.2 - 19.8	235	13 $\pm$ 2	9 - 18.9
Weight ( <i>kg</i> )	260	43.5 $\pm$ 14.7	20.2 - 95	235	41.8 $\pm$ 11.9	18-76
Height ( <i>cms</i> )	260	152.6 $\pm$ 13.4	125.5-184.2	235	149.2 $\pm$ 9.5	122.8-173.7
BMI ( <i>kg m<sup>-2</sup></i> )	260	18.2 $\pm$ 3.9	9.8 - 31.4	235	18.5 $\pm$ 3.7	10.3 - 30.6
RMR ( <i>kcal day<sup>-1</sup></i> )	260	1212 $\pm$ 263	716 - 2370	232	1066 - 203	612 - 1846
RMR/BM ( <i>kcal kg<sup>-1</sup>day<sup>-1</sup></i> )	260	29.5 $\pm$ 6.7	15.6 - 51.5	232	26.8 $\pm$ 6.3	14.2 - 52.1
Fat mass ( <i>kg</i> )	257	8.1 $\pm$ 7.8	0.6 - 40.3	234	11.1 $\pm$ 7.1	0.5 - 38.2
Fat free mass ( <i>kg</i> )	257	35.8 $\pm$ 9.3	20.8 - 63.9	234	31.2 $\pm$ 6.0	17.5 -51.7

Table 3.1: Mean  $\pm$  standard deviation and the observed range of physical characteristics of the subjects in the MCS dataset. The sample size (*n*) is given for each variable. BMI: body mass index.

Variables	Boys ( $n=9$ )		Girls ( $n=9$ )	
	Mean $\pm$ SD	Range	Mean $\pm$ SD	Range
Age ( <i>years</i> )	7.1 $\pm$ 0.8	6.3 - 8.1	7.6 $\pm$ 0.8	6.5 - 8.7
Weight ( <i>kg</i> )	20.2 $\pm$ 3.2	17.5 - 26.5	18.3 $\pm$ 2.3	15.0 - 22.4
Height ( <i>cms</i> )	119.0 $\pm$ 7.6	109.5 - 128.4	119.7 $\pm$ 8.0	104.9 - 131.9
BMI ( <i>kg m<sup>-2</sup></i> )	14.2 $\pm$ 1.3	12.7 - 16.3	12.8 $\pm$ 0.9	11.5 - 14.2
RMR ( <i>kcal day<sup>-1</sup></i> )	1004 $\pm$ 189	771 - 1366	835 $\pm$ 187	552 - 1128
RMR/BM ( <i>kcal kg<sup>-1</sup> day<sup>-1</sup></i> )	50.5 $\pm$ 11.0	37.1 - 70.4	45.8 $\pm$ 9.1	28.9 - 58.2
Liver mass ( <i>kg</i> )	0.66 $\pm$ 0.14	0.47 - 0.92	0.53 $\pm$ 0.12	0.40 - 0.78
Kidney mass ( <i>kg</i> )	0.09 $\pm$ 0.02	0.07 - 0.12	0.08 $\pm$ 0.01	0.06 - 0.10

Table 3.2: Mean  $\pm$  standard deviation and the observed range of physical characteristics of subjects in the RMR-USG dataset.

### 3.2.1.3 Relative organ mass ( $M_i/BM$ ) data

A prominent dataset for reference physiological variables in North American population compiled by Altman and Dittmer (1962) was used for organ weights from birth to maturity. To the best of our knowledge, this was the only dataset that provided liver, brain, heart and kidney weights of children in age groups one year apart from birth to adulthood. The reference relative mass ( $M_i/BM$ ) of liver, kidney, heart and brain is illustrated in Figure 3.1.

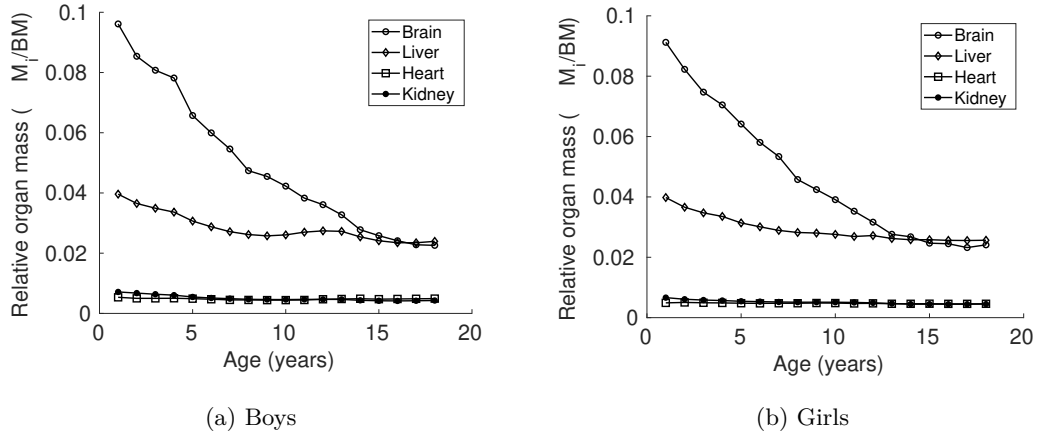


Figure 3.1: Relative organ mass ( $M_i/BM$ ) of brain, liver, heart and kidney reported by Altman and Dittmer (1962) in North American children.

## 3.2.2 Model

A mechanistic model for RMR/BM in children and adolescents due to Wang (2012) can be written as

$$\frac{RMR}{BM} = R_c \sum K_i \frac{M_i}{BM}, \quad (3.1)$$

where  $R_c$  is the relative cellularity,  $K_i$  is the specific metabolic rate of an organ ( $i$  for brain, heart, kidney, liver and the residual mass) and  $M_i/BM$  is the relative mass of the organ ' $i$ ' with respect to the body mass ( $BM$ ). Residual mass is obtained by subtracting the sum of the mass of four organs from the body mass. These physiological parameters in Eq. 3.1 are described in detail as follows:

### 3.2.2.1 Relative cellularity ( $R_c$ )

The ratio of body cell mass (BCM) to fat-free mass (FFM) is defined as the whole body cellularity, which quantifies the metabolically active portion of FFM. Whole body cellularity is thought to change in the course of life and is assumed to be smaller in children than young adults (Wang et al. 2005, 2010a). Hence the factor 'relative cellularity' ( $R_c$ ), which is defined as the ratio of  $BCM/FFM$  in children to that of adults, is incorporated in Eq. 3.1. Here,  $BCM$  is assumed to be proportional to the total body potassium (TBK) and the change in  $BCM/FFM$  in children is estimated through  $TBK/FFM$ . In the reference Caucasian adults, (Snyder et al. 1975)  $TBK/FFM$  (mmol/kg) is reported to be 68.1 for men and 64.2 for women (Forbes 1987). Thus in children,  $R_c$  is approximated as  $(TBK/FFM)/68.1$  in boys and  $(TBK/FFM)/64.2$  in girls, in a given age group. Data on  $R_c$  from birth to adulthood were compiled by Wang (Wang 2012) based on age-related changes in total body potassium (TBK) relative to FFM, from studies by Fomon et al. (1982) and Pierson (2005).

### 3.2.2.2 Specific metabolic rate ( $K_i$ )

Specific metabolic rate (kcal/kg/day) of an organ ' $i$ ' is the metabolic rate per unit mass of that organ, denoted as  $K_i$ . The specific metabolic rate ( $K_i$ ) of organs in adults was measured in vivo by Elia (1992). Elia estimated the oxygen consumption of organs in vivo by measuring the difference in arterio-venous oxygen concentration across tissue and the blood flow rate. The  $K_i$  ( $kcal\ kg^{-1}\ day^{-1}$ ) values are reported as 200 for liver, 240 for the brain, 440 for heart and kidneys, 13 for skeletal muscle mass, 4 for fat mass and 12 for residual mass in adults.  $K_i$  values are thought to be higher in children (Chugani et al. 1987; Wang et al. 2005). Hence the adult  $K_i$  values estimated in vivo by Elia (Elia 1992) are adjusted in the Wang model with an age depending factor 'relative  $K_i$ ' (Wang 2012), which is the ratio of  $K_i$  values in children to that of adults. Relative  $K_i$  values are assumed from surrogate physiological parameters (Wang 2012) such as brain oxygen consumption (Chugani et al. 1987), heartbeat rates, and other physiological parameters.

### 3.2.3 Modified model of RMR/BM in Indian children.

Eq. 3.1 suggests that relative mass of organs (and tissues) and their specific metabolic rates are the two major factors that determine the RMR/BM in children and adolescents. In this study, we look at two particular sources of variation influencing the whole body RMR/BM. First, we consider the variation in the relative mass of major organs, assuming the specific metabolic rates of organs are constant (Elia 1992). Secondly, we consider the composition of residual mass and its effect on the metabolic rate of relative residual mass and in turn on RMR/BM. We propose two models for RMR/BM in Indian children as follows:

#### 3.2.3.1 Model 1: adjusting the relative mass of high metabolic rate organs.

We modified Eq. 3.1 by adjusting the relative organ mass of four major organs (liver, kidney, brain and heart) by a fraction  $\delta_i$ . We define  $\delta_i$  as the ratio of relative organ mass ( $M_i/BM$ ) in the Indian population to the  $M_i/BM$  in the Caucasian population. Assuming  $M_i/BM$  of major organs (liver, brain, kidney, heart) are adjusted by the same fraction  $\delta$ , Eq. 3.1 can be written for the Indian population as

$$RMR_{\delta}/BM = \left( \delta \left( K_{liver} \frac{M_{liver}}{BM} + K_{heart} \frac{M_{heart}}{BM} + K_{brain} \frac{M_{brain}}{BM} + K_{kidney} \frac{M_{kidney}}{BM} \right) + K_{residual\ mass} \frac{M'_{residual\ mass}}{BM} \right) R_c, \quad (3.2)$$

where

$$\frac{M'_{residual\ mass}}{BM} = 1 - \delta \left( \frac{M_{liver}}{BM} + \frac{M_{heart}}{BM} + \frac{M_{brain}}{BM} + \frac{M_{kidney}}{BM} \right), \quad (3.3)$$

$\frac{M'_{residual\ mass}}{BM}$  is the residual mass after adjusting the relative mass of major organs by a factor  $\delta$ ,  $R_c$  is the relative cellularity, and  $K_i$  is the specific metabolic rate of an organ.

#### 3.2.3.2 Model 2: adjusting the metabolic contribution from relative residual mass

In Model 2, RMR/BM in Eq. 3.1 is modified under the assumption that the metabolic contribution from residual mass in the Indian population is lower by factor  $\delta'$  compared to the Caucasian population. Thus, the alternate model for RMR/BM can be written as

$$RMR_{\delta'}/BM = \left( K_{liver} \frac{M_{liver}}{BM} + K_{heart} \frac{M_{heart}}{BM} + K_{brain} \frac{M_{brain}}{BM} + K_{kidney} \frac{M_{kidney}}{BM} + \delta' K_{residual\ mass} \frac{M_{residual\ mass}}{BM} \right) R_c, \quad (3.4)$$

where

$$\frac{M_{residual\ mass}}{BM} = 1 - \left( \frac{M_{liver}}{BM} + \frac{M_{heart}}{BM} + \frac{M_{brain}}{BM} + \frac{M_{kidney}}{BM} \right), \quad (3.5)$$

$R_c$  is the relative cellularity,  $K_i$  is the specific metabolic rate and  $M_i/BM$  is the relative mass of respective organs.

### 3.2.4 Statistical analysis

All descriptive data are reported as the mean  $\pm$  standard deviation (SD). The measured and the theoretical values were compared using the paired t-test with the significance level set at  $\alpha = 0.05$ . The relative organ mass between the two population was compared through non-parametric Wilcoxon signed-rank test, with the significance level set at  $\alpha = 0.05$ . All the analyses were carried out using MATLAB R2019b (MATLAB 2019) and R version 3.6.2 (R Core Team 2019).

## 3.3 Results

The measured RMR per unit body mass ( $kcal\ kg^{-1}day^{-1}$ ) in Indian children is denoted as  $RMR_M/BM$ .  $RMR_T/BM$  represents the theoretical expectation calculated from the Wang model (Eq. 3.1) with the reference organ weights data reported by Altman and Dittmer (1962). Similarly, RMR/BM calculated from Model 1 (Eq. 3.2) is denoted as  $RMR_\delta/BM$  and from Model 2 (Eq. 3.4) as  $RMR_{\delta'}/BM$ . Subjects are grouped one year apart in the analysis. We employ the following notation: Children above the age of 10 years but less than 11 are denoted for brevity as age group 10, and so on.

### 3.3.1 RMR/BM in Indian children is significantly lower than Caucasian children

We studied RMR/BM in Indian children using a mechanistic model due to Wang (2012) (Eq. 3.1), which partitions total body mass into four major organs and residual mass.

The mean  $RMR_M/BM$  measured in the MCS cohort, stratified by age, was compared with the theoretical  $RMR_T/BM$  from Eq. 3.1 calculated with the relative mass of the four major

organs reported for the Caucasian population (Altman and Dittmer 1962). In Figure 3.2a and 3.2b the solid curve shows the mean measured  $RMR_M/BM$  ( $\mu \pm SD$ ); the dotted curve is the theoretical  $RMR_T/BM$  (Wang model) and is representative of the mean RMR/BM in Caucasian children (Talbot 1938; Wang 2012). In boys, the measured  $RMR_M/BM$  is significantly lower than the theoretical  $RMR_T/BM$  in the age groups 11, 13, 14, 15 and 16 years ( $P < 0.05$ ); but not at 10 and 12 years ( $P = 0.7$  and  $0.09$ , respectively). In girls,  $RMR_M/BM$  is significantly lower than  $RMR_T/BM$  in all the age groups from 10 years to 16 years;  $P < 0.05$  for 10 years and  $P < 0.001$  for 11 to 16 years.

We thus observe a significantly lower mean RMR/BM in Indian adolescents (232 girls and 260 boys) compared to the reference Caucasian adolescents (Talbot 1938), except in boys aged 9 to 11 years and 12 to 13 years.

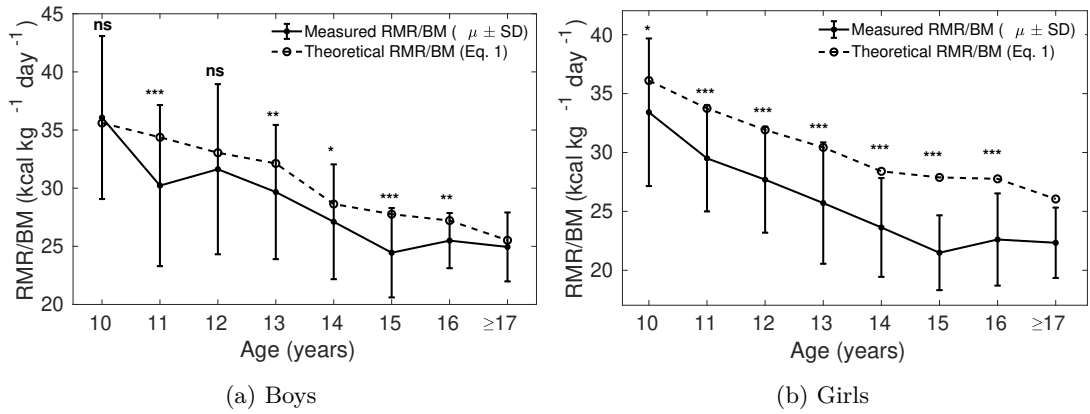


Figure 3.2: The solid curve shows the mean ( $\pm SD$ ) RMR/BM measured in each age group, and the dotted line shows the mean theoretical  $RMR_T/BM$  based on Eq. 3.1 for the Caucasian population in boys (3.2a) and girls (3.2b). ns: not significant, \*:  $P < 0.05$ , \*\*:  $P < 0.01$  and \*\*\*:  $P < 0.001$ . 9 and 10 years groups were combined for the statistical tests. The analysis is not done when the number of samples was less than 10 (17 years and above).

### 3.3.2 A modified Wang model of RMR/BM for Indian children

Measured  $RMR_M/BM$  in the MCS dataset is significantly lower than the mean  $RMR_T/BM$  in the Caucasian population. In accordance with Eq. 3.1,  $RMR_T/BM$  is determined by the relative mass of four major organs, with smaller (larger)  $M_i/BM$  leading to smaller (larger)  $RMR_T/BM$ . Thus, we hypothesise that the lower mean RMR/BM between the Indian and the Caucasian children is due to a lower mean relative mass of the four major organs in the Indian population.

We define  $\delta_i$  (see Section 3.2.3.1 below) as the ratio of relative organ mass ( $M_i/BM$ ) in the Indian population to the  $M_i/BM$  in the Caucasian population. Eq. 3.1 is modified to Eq. 3.2 by adjusting the mass of major organs by a fraction  $\delta$  (Model 1). We optimised  $\delta$  by minimising the mean squared error (MSE) between the measured and the model (Eq. 3.2), for  $\delta$  varying from 0 to 1. The optimal  $\delta$  values corresponding to the least MSE was found to be  $\delta = 0.90$  in boys and  $\delta = 0.77$  in girls.

Model 1 (Eq. 3.2) evaluated with optimal  $\delta$  was then compared with the measured  $RMR_M/BM$ , as shown in Figure 3.3. The dotted curve shows the mean  $RMR_\delta/BM$  calculated from Eq. 3.2 with  $\delta = 0.90$  in boys (Figure 3.3a) and  $\delta = 0.77$  in girls (Figure 3.3b). The solid curve shows the measured  $RMR_M/BM$  ( $\mu \pm SD$ ). We verify that the model is not significantly different from the measured values, except in the age groups 10 and 15 years in boys and 15 years in girls.

Our modified Wang model (Eq. 3.2) is thus better suited to predicting RMR/BM in Indian children compared to the naive Wang model (Eq. 3.1). Physiologically this implies that the relative organ masses in the Indian population ought to lower by a factor 0.90 in boys and 0.77 in girls compared to reference relative organ mass in the Caucasian population (Altman and Dittmer 1962).

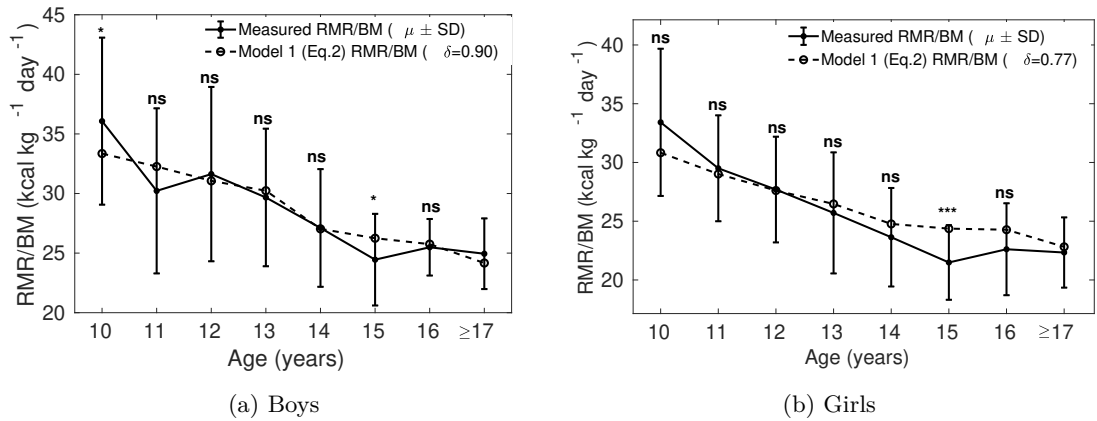


Figure 3.3: The dotted curve is the adjusted RMR/BM calculated from Eq. 3.2 assuming that the relative mass ( $M_i/BM$ ) of all the organs (liver, brain, kidney, heart) are smaller by a fraction of 0.77 in girls and 0.90 in boys compared to that of Caucasian population (Altman and Dittmer 1962), that is with  $\delta = 0.90$  and 0.77 in Eq. 3.2 in boys and girls respectively. The solid curve shows the mean measured  $RMR_M/BM$  in MCS dataset. ns: not significant, \*:  $p < 0.05$ , \*\*:  $p < 0.01$  and \*\*\*:  $p < 0.001$  (Compare Fig. 3.2)



### 3.3.3 Relative kidney mass in 6 to 8 years old Indian children is significantly lower but not relative liver mass.

Model 1 (Eq. 3.2) predicts that the relative mass of major organs in the Indian population is lower by 10% in boys and 23% in girls compared to the Caucasian population. We measured the liver and kidney masses in 9 girls and 9 boys aged 6 to 8 year in the RMR-USG children to validate the Model 1 predictions. The ratio of relative liver and relative kidney mass ( $M_i/BM$ ) measured in the RMR-USG dataset compared to the corresponding  $M_i/BM$  in the Caucasian counterparts (Altman and Dittmer 1962) are denoted as  $\delta_{liver}$  and  $\delta_{kidney}$ , respectively. Figure 3.4 shows the  $\delta_{liver}$  and  $\delta_{kidney}$  observed in the RMR-USG dataset. The mean ( $\pm SD$ ) observed  $\delta_{liver}$  is  $1.19 \pm 0.28$  in boys and  $1.0 \pm 0.16$  in girls, and  $\delta_{kidney}$  is  $0.90 \pm 0.097$  in boys and  $0.87 \pm 0.10$  in girls.

The  $\delta_{kidney}$  observed in the RMR-USG dataset is significantly lower ( $P=0.009$  in boys and  $P=0.009$  in girls; one-sided Wilcoxon signed-rank test). Consistent with Eq. 3.2 predictions, the relative kidney mass measured in Indian children is found to be lower than that of reference Caucasian children in the respective age groups. However, we failed to find any significant difference in the observed  $\delta_{liver}$  (boys  $P=0.45$  and girls  $P=0.91$ ).

It is noteworthy that the  $\delta_{kidney}$  predicted by Eq. 3.2 was close to the observed  $\delta_{kidney}$ :  $\delta_{kidney}$  was observed to be  $0.90 \pm 0.097$  compared to the prediction 0.9 in boys; in girls  $\delta_{kidney}$  was observed to be  $0.87 \pm 0.10$  compared to the predicted 0.77). However, the  $\delta_{liver}$  in both girls and boys is higher than the optimal  $\delta$  predicted by Eq. 3.2.

### 3.3.4 Alternate model of RMR/BM in Indian children based on residual mass

Residual mass (the mass remaining after subtracting liver, brain, heart and kidney mass from total body mass) constitutes a much larger part of the body mass compared to the sum of masses of four major organs. The residual mass is composed mainly of skeletal muscle mass and fat mass, along with lungs, spleen, gastrointestinal tract, connective tissue etc. Broadly, speaking, skeletal muscle mass and the fat mass are the more malleable components of the body compared to the sizes of the major organs. Moreover, the fat and muscle mass per cent in Indian children

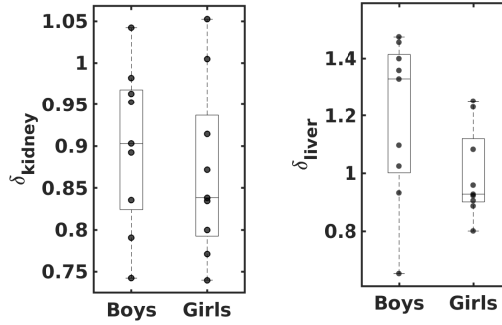


Figure 3.4:  $\delta_{kidney}$  and  $\delta_{liver}$  observed in Indian children (9 girls and 9 boys), where  $\delta_i$  denotes the ratio of the relative mass of the organ  $i$  measured in RMR-USG dataset to that of their Caucasian counterparts (Altman and Dittmer 1962). The lower and upper whiskers indicate the minimum and the maximum values; and the lower edge, middle line and the upper edge of the box indicates the 25th percentile, median and the 75th percentile values, respectively. The dots show the observed individual  $\delta$  values.

is characteristically different from the Western population (Chiplonkar et al. 2017). This can potentially account for the wide variation in RMR between children. To examine this possibility, we next studied an alternate model of RMR/BM (Model 2) which takes into account the variation in the metabolically active constituents of residual mass.

We modified Eq. 3.1 to Eq. 3.4 (Model 2), by incorporating a factor  $\delta'$  which adjusts the metabolic rate of relative residual mass in the Indian population. An optimal  $\delta'$  was obtained by minimising the mean squared error between the measured  $RMR_M/BM$  and the  $RMR_{\delta'}/BM$  calculated by Eq. 3.4 in the MCS dataset, for  $\delta'$  ranging from 0 to 1. The  $\delta'$  corresponding to the least MSE is found to be 0.85 in boys and 0.65 in girls.

In Figure 3.5, the dotted curve shows the  $RMR_{\delta'}/BM$  calculated from Model 2, with  $\delta' = 0.85$  in boys and  $\delta' = 0.65$  in girls (Figure 3.5b); and the solid curve shows the measured RMR/BM ( $\mu \pm SD$ ) in MCS dataset. In boys (Figure 3.5a), the dotted curve is not significantly different from the mean measured  $RMR_M/BM$  in the MCS dataset (solid curve), except in the age groups 11 and 15 years ( $P=0.03$  and  $0.02$ , respectively). Similarly, in girls the solid curve is not significantly different from the dotted curve, except in the age group 15 years ( $P=0.001$ ).

$\delta'$  can be interpreted physiologically as the effect of body composition differences on RMR/BM. Thus Model 2 raises the hypothesis that the metabolic contribution from the relative residual mass is reduced in the Indian children, lower by 15% in boys and 35% in girls, if the relative mass of major organs is assumed to be similar in the two populations. This indicates that variation in body composition could play a considerable role in determining RMR in Indian children.

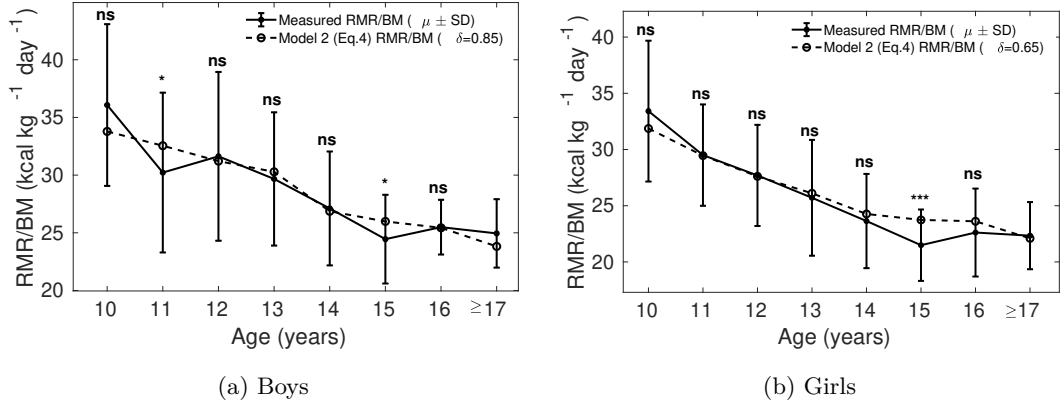


Figure 3.5: Measured RMR/BM ( $\mu \pm SD$ ) is shown as the solid curve, and the dotted curve shows the RMR/BM calculated from Eq. 3.4 with  $\delta' = 0.85$  in boys and  $\delta' = 0.65$  in girls, and reference relative organ mass for the Caucasian population (Altman and Dittmer 1962).

### 3.4 Discussion

Resting metabolic rate (RMR) is a significant factor in determining energy balance, which in turn critically influences the energy available for growth from birth to adulthood. The mean RMR per unit body mass (RMR/BM) is not uniform across populations; Indian children have significantly lower RMR/BM compared to their Caucasian counterparts. Not only are these population differences not understood from a physiological standpoint, but inter-individual variations are also poorly explained. Several models have been proposed over the years to try to explain RMR through various anthropometric variables such as height, weight, fat and fat-free mass. One such prominent model is the Katch-McArdle equation (Katch et al. 1990) which computes RMR as due to fat-free mass:  $RMR = 370 + (21.6FFM)$ . However, such models have been reported to explain only about 60-80% of the intraspecific variation. An alternate strategy is to explain the *mean* RMR of children clustered into one-year age groups. A very successful model in this class is the Wang model, which achieves an  $R^2 = 0.99$ . On the other hand, it is unclear if the Wang model is readily applicable to other populations. In particular, the Caucasian dataset modelled in the Wang study shows little variation in the age-stratified data, whereas a much wider variability is expected, in general, in Indian children. In this study, we attempt to modify the Wang class of models for application to Indian children. It is worth pointing out that using linear regression models based on body composition, such as the Katch-McArdle equation, we could only explain about 70% variation in the mean RMR/BM in an age group. We also explored several other regression models based on body composition and anthropometry, but they each explained only 30-60% of the inter-individual variation in RMR observed in Indian

children (analysis not shown).

In this work we construct two models of RMR/BM in Indian children based on the Wang model (Wang 2012), which describe the mean RMR/BM stratified by age phenomenologically. In Model 1, we assume lower organ masses are responsible for the lower observed RMR; in Model 2, residual masses are calibrated to the observed RMR. The coefficients of determination ( $R^2$ ) in explaining the mean measured RMR/BM for Model 1 and Model 2 are 0.84 and 0.85 in boys and 0.95 and 0.97 in girls, respectively. The lower accuracy of these models in describing the RMR/BM in Indian children compared to the Caucasian children ( $R^2 = 0.99$ ) is consistent with high variation in the observed RMR (ranging from 612 to 2370  $kcal\ day^{-1}$ ). It seems unlikely that larger sample sizes would substantially improve the accuracy of the model.

Next, we ask if these models provide a physiological understanding of the lower RMR/BM observed in Indian children. If the lower RMR/BM is due to a lower relative mass of four major organs (liver, kidney, heart, brain) through a modified Wang model, Model 1 (Eq. 3.2) predicted that the relative masses of the four major organs should be lower by 10% in boys and 23% in girls. Our pilot study designed to test these predictions showed the relative kidney mass was significantly lower but failed to find any significant difference in the relative liver mass. It is interesting to note that a lower relative kidney mass in Indian children is consistent with the Barker hypothesis (Almond and Currie 2011) and the observation of fewer nephrons in low birth weight babies (Wlodek et al. 2008). On the other hand, failure to observe a significant difference in relative liver weight could suggest a lower  $K_{liver}$  instead, which is consistent with lower  $K_i$  values reported in South Asian females (Shirley et al. 2019). This challenges the applicability of specific metabolic rates estimated, in particular, in Elia (1992), across racially and ethnically diverse populations. One limitation of the current study is the assumption that brain and heart masses are likely to be relatively conserved within an age group; due to practical difficulties, these were not measured in our study.

To provide further contrast, we constructed Model 2, which analyses the influence of metabolically active constituents of residual mass on RMR/BM. Model 2 predicts that the metabolic rate of residual mass is lower by 15% in boys and 35% in girls in the Indian population compared to the Caucasian population. Model 2 re-emphasizes the importance of body composition in explaining variation in RMR. It is interesting that a century-long attempt to decipher the relationship between body composition and RMR has not been very successful (FAO/WHO/UNU 1985; Aub and Du Bois 1917; Cunningham 1980; Harris and Benedict 1918; Henry 2005; Katch et al. 1990;

McMurray et al. 2014; Kleiber 1932; Bedale 1923; Owen et al. 1986, 1987; Schofield 1985; Mifflin et al. 1990; Corrigan et al. 2020). Thus, understanding the physiological underpinnings of Model 2 remains an open problem. Finally, we note that it is plausible that more complex formulations than basing RMR on either organ mass or residual mass are necessary. One attractive approach for future work is to employ data-driven machine learning strategies to discover these complex relations.

We remark on some refinements of our work that might be possible in future studies. In children, strict standard conditions for RMR measurement are difficult to achieve. The terms basal metabolic rate (BMR), resting metabolic rate (RMR) and resting energy expenditure (REE) are different measurements of the resting metabolism and are often used interchangeably; however, RMR and REE can be 3-10% higher than BMR, as they follow less stringent settings for the measurements (Psota and Chen 2013). In our study, RMR in children was not measured following fasting conditions alone; hence RMR measured in our study could be higher than the basal metabolism; such differences could be up to 100 kcal/day (Haugen et al. 2003). On the other hand, indirect calorimeters have been reported to underestimate REE in some studies (Purcell et al. 2020). We argue that the distinct patterns in boys and girls are of prime interest, and these are less likely to be explained due to measurement bias alone. In the future studies, it will be interesting to ask how resting metabolic rate measured not at a “pure resting” state will be more pragmatic measure of RMR compared to the traditional RMR in clinical settings.

The ratio RMR/BM was used to to normalise the RMR with respect to BM. This ratio has been used by several studies, including Wang (2012); Rahmandad (2014). However, other authors have critiqued this ratio due to the observation that a linear RMR–BM relationship extrapolates to a non-zero intercept (Tschöp et al. 2012; Poehlman and Toth 1995). While it is not immediately clear if RMR should be normlized by BM, in our study, we follow the Wang (2012) model closely. In other words, since our comparisons are with respect to the Wang model, the appropriate variable in our work is the normalized RMR/BM.

The significance of our study is that a lower RMR/BM in Indian children can significantly influence energy balance, and amplify the effects of lower or higher energy intakes. Swinburn and colleagues (Swinburn et al. 2006) have reported that even a 10% change in total energy expenditure (TEE; consists of RMR as a component) could lead to a 4.5% difference in mean weight of children between two populations. The implications of lower RMR/BM in Indian children on the dynamics of growth and development will be studied in the future, in particular,

using quantitative models of growth and weight changes (Hall et al. 2013). The present study has provided that basis through two phenomenological models; either of which can be used to estimate age-wise mean RMR/BM in Indian adolescents. While predicting individual RMR/BM is far from complete, the present models are likely to be referred to by clinicians and policymakers to infer energy expenditure benchmarks in Indian children. Such studies are critical to understanding the implications of a lower RMR/BM in growth, development, and life-course diseases.

### 3.5 Preprint

An earlier version of this article is available as a preprint at Areekal et al. (2021), *Two Novel Models Evaluating the Determinants of Resting Metabolic Rate in Indian Children*, 16 February 2021, PREPRINT (Version 1) available at Research Square [<https://doi.org/10.21203/rs.3.rs-196719/v1>].

### Bibliography

- Almond, D. and Currie, J. (2011). Killing Me Softly: The Fetal Origins Hypothesis. *journal of economic perspectives : a journal of the American Economic Association*, 25(3):153.
- Altman, P. L. and Dittmer, D. S. (1962). *Growth including reproduction and morphological development*. Washington : Federation of American Societies for Experimental Biology, Author.
- Areekal, S. A., Khadilkar, A., Ekbote, V., Kajale, N., Kinare, A. S., and Goel, P. (2021). Two novel models evaluating the determinants of resting metabolic rate in Indian children. *Research Square*, PREPRINT (Version 1), <https://doi.org/10.21203/rs.3.rs-196719/v1>.
- Aub, J. C. and Du Bois, E. F. (1917). Clinical calorimetry: Nineteenth paper the basal metabolism of old men. *Archives of Internal Medicine*, XIX(5\_II):823–831.
- Baker, J., Workman, M., Bedrick, E., Frey, M. A., Hurtado, M., and Pearson, O. (2010). Brains versus brawn: an empirical test of barker’s brain sparing model. *American Journal of Human Biology: The Official Journal of the Human Biology Association*, 22(2):206–215.
- Barker, D. (2004). Developmental origins of adult health and disease. *Journal of Epidemiology and Community Health*, 58(2):114.

- Bedale, E. M. (1923). Energy expenditure and food requirements of children at school. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 94(662):368–404.
- Bosy-Westphal, A., Reinecke, U., Schlörke, T., Illner, K., Kutzner, D., Heller, M., and Müller, M. J. (2004). Effect of organ and tissue masses on resting energy expenditure in underweight, normal weight and obese adults. *International Journal of Obesity and Related Metabolic Disorders : Journal of the International Association for the Study of Obesity*, 28(1):72–79.
- Bosy-Westphal, A., Wolf, A., Bührens, F., Hitze, B., Czech, N., Mönig, H., Selberg, O., Settler, U., Pfeuffer, M., Schrezenmeir, J., Krawczak, M., and Müller, M. J. (2008). Familial influences and obesity-associated metabolic risk factors contribute to the variation in resting energy expenditure: the Kiel Obesity Prevention Study. *American Journal of Clinical Nutrition*, 87(6):1695–1701.
- Chiplonkar, S., Kajale, N., Ekbote, V., Mandlik, R., Parthasarathy, L., Borade, A., Patel, P., Patel, P., Khadilkar, V., and Khadilkar, A. (2017). Reference Centile Curves for Body Fat Percentage, Fat-free Mass, Muscle Mass and Bone Mass Measured by Bioelectrical Impedance in Asian Indian Children and Adolescents. *Indian Pediatrics*, 54(12):1005–1011.
- Chugani, H. T., Phelps, M. E., and Mazziotta, J. C. (1987). Positron emission tomography study of human brain functional development. *Annals of Neurology*, 22(4):487–497.
- Corrigan, J. K., Ramachandran, D., He, Y., Palmer, C. J., Jurczak, M. J., Chen, R., Li, B., Friedline, R. H., Kim, J. K., Ramsey, J. J., Lantier, L., McGuinness, O. P., Mouse Metabolic Phenotyping Center Energy Balance Working Group, and Banks, A. S. (2020). A big-data approach to understanding metabolic rate and response to obesity in laboratory mice. *eLife*.
- Cunningham, J. J. (1980). A reanalysis of the factors influencing basal metabolic rate in normal adults. *American Journal of Clinical Nutrition*, 33(11):2372–2374.
- De B. Weir, J. B. (1949). New methods for calculating metabolic rate with special reference to protein metabolism. *J. Physiol.*, 109(1-2):1–9.
- Elia, M. (1992). Energy expenditure in the whole body. In *Kinney JM*, pages 19–49. Energy metabolism: tissue determinants and cellular collaries. : Raven Press Ltd, New York.
- Esht, V., Midha, D., Chatterjee, S., and Sharma, S. (2018). A preliminary report on physical

- activity patterns among children aged 8–14 years to predict risk of cardiovascular diseases in Malwa region of Punjab. *Indian Heart Journal*, 70(6):777–782.
- FAO/WHO/UNU (1985). *Energy and protein requirements : report of a Joint FAO/WHO/UNU Expert Consultation*. World Health Organization.
- FAO/WHO/UNU (2004). *Human energy requirements : report of a joint FAO/WHO/UNU Expert Consultation, Rome 17-24 October 2001*. Food and Agriculture Organization of the United Nations Rome.
- Fomon, S. J., Haschke, F., Ziegler, E. E., and Nelson, S. E. (1982). Body composition of reference children from birth to age 10 years. *The American Journal of Clinical Nutrition*, 35(5):1169–1175.
- Forbes, G. B. (1987). *Human body composition, growth, aging, nutrition, and activity*. New York: Springer-Verlag.
- Gallagher, D., Belmonte, D., Deurenberg, P., Wang, Z., Krasnow, N., Pi-Sunyer, F. X., and Heymsfield, S. B. (1998). Organ-tissue mass measurement allows modeling of REE and metabolically active tissue mass. *American Journal of Physiology-Endocrinology and Metabolism*.
- Hales, C. N. and Barker, D. J. (1992). Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. *Diabetologia*, 35:595–601.
- Hall, K. D., Butte, N. F., Swinburn, B. A., and Chow, C. C. (2013). Dynamics of childhood growth and obesity: development and validation of a quantitative mathematical model. *Lancet Diabetes Endocrinol.*, 1(2):97–105.
- Harris, J. A. and Benedict, F. G. (1918). A Biometric Study of Human Basal Metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, 4(12):370.
- Haugen, H. A., Melanson, E. L., Tran, Z. V., Kearney, J. T., and Hill, J. O. (2003). Variability of measured resting metabolic rate. *American Journal of Clinical Nutrition*, 78(6):1141–1144.
- Henry, C. J. K. (2005). Basal metabolic rate studies in humans: measurement and development of new equations. *Public Health Nutrition*, 8(7a):1133–1152.
- Hsu, A., Heshka, S., Janumala, I., Song, M.-Y., Horlick, M., Krasnow, N., and Gallagher, D.



- (2003). Larger mass of high-metabolic-rate organs does not explain higher resting energy expenditure in children. *Am. J. Clin. Nutr.*, 77(6):1506–1511.
- Indian Council of Medical Research (ICMR) (2010). Nutrient requirements and recommended dietary allowances for Indians In: A report of the expert group of the Indian Council of Medical Research National Institute of Nutrition Hyderabad. *Indian Council of Medical Research*.
- Johnstone, A. M., Murison, S. D., Duncan, J. S., Rance, K. A., and Speakman, J. R. (2005). Factors influencing variation in basal metabolic rate include fat-free mass, fat mass, age, and circulating thyroxine but not sex, circulating leptin, or triiodothyronine. *American Journal of Clinical Nutrition*, 82(5):941–948.
- Katch, V. L., Marks, C. C., Becque, M. D., Moorehead, C., and Rocchini, A. (1990). Basal metabolism of obese adolescents: Evidence for energy conservation compared to normal and lean adolescents. *American Journal of Human Biology*, 2(5):543–551.
- Khadilkar, A. V., Lohiya, N., Mistry, S., Chiplonkar, S., Khadilkar, V., Kajale, N., Ekbote, V., Vispute, S., Mandlik, R., Prasad, H., Singh, N., Agarwal, S., Palande, S., and Ladkat, D. (2019). Random Blood Glucose Concentrations and their Association with Body Mass Index in Indian School Children. *Indian Journal of Endocrinology and Metabolism*, 23(5):529.
- Kleiber, M. (1932). Body size and metabolism. *Hilgardia*, 6(11):315–353.
- Krishnan, B. T. and Vareed, C. (1932). Basal metabolism of young college students, men and women, in Madras. *Ind. J. Med. Res.*, 19:831–858.
- Kumar, S., Kumar, N., and Sachar, R. S. (1961). Basal metabolic rate in normal Indian adult males. *Indian Journal of Medical Research*, 49:702–709.
- Kyle, U. G., Bosaeus, I., De Lorenzo, A. D., Deurenberg, P., Elia, M., Gómez, J. M., Heitmann, B. L., Kent-Smith, L., Melchior, J.-C., Pirlich, M., Scharfetter, H., Schols, A. M. W. J., Pichard, C., and Espen (2004). Bioelectrical impedance analysis-part II: utilization in clinical practice. *Clinical Nutrition (Edinburgh, Scotland)*, 23(6):1430–1453.
- Latini, G., De Mitri, B., Del Vecchio, A., Chitano, G., De Felice, C., and Zetterström, R. (2004). Foetal growth of kidneys, liver and spleen in intrauterine growth restriction: “programming” causing “metabolic syndrome” in adult age. *Acta paediatrica*, 93(12):1635–1639.

- Lubree, H. G., Rege, S. S., Bhat, D. S., Raut, K. N., Panchnadikar, A., Joglekar, C. V., Yajnik, C. S., Shetty, P., and Yudkin, J. (2002). Body fat and cardiovascular risk factors in indian men in three geographical locations. *Food and Nutrition Bulletin*, 23(3\_suppl1):146–149.
- Luyckx, V. A. and Brenner, B. M. (2005). Low birth weight, nephron number, and kidney disease. *Kidney International*, 68:S68–S77.
- Mason, E. D. and Benedict, F. G. (1931). The basal metabolism of South Indian women. *Indian Journal of Medical Research*, 19:75–98.
- Mason, E. D., Mundkur, V., and Jacob, M. (1963). Basal energy metabolism and heights, weights, arm skinfold and muscle of young Indian women in Bombay, with prediction standards for B.M.R. *Indian Journal of Medical Research*, 51:925–932.
- MATLAB (2019). *R2019b*. The MathWorks Inc., Natick, Massachusetts, United States.
- McMurray, R. G., Soares, J., Caspersen, C. J., and McCurdy, T. (2014). Examining Variations of Resting Metabolic Rate of Adults: A Public Health Perspective. *Medicine and science in sports and exercise*, 46(7):1352.
- Menzel, H. G., Clement, C., and DeLuca, P. (2009). Icrp publication 110. realistic reference phantoms: an icrp/icru joint effort. a report of adult reference computational phantoms. *Ann ICRP*, 2:1–164.
- Mifflin, M. D., Jeor, S. T. S., Hill, L. A., Scott, B. J., Daugherty, S. A., and Koh, Y. O. (1990). A new predictive equation for resting energy expenditure in healthy individuals. *American Journal of Clinical Nutrition*, 51(2):241–247.
- Mukherjee, H. N. and Gupta, P. C. (1931). The basal metabolism of Indians (Bengalis). *Indian Journal of Medical Research*, 18:807–812.
- Müller, M. J., Langemann, D., Gehrke, I., Later, W., Heller, M., Glüer, C. C., Heymsfield, S. B., and Bosy-Westphal, A. (2011). Effect of Constitution on Mass of Individual Organs and Their Association with Metabolic Rate in Humans—A Detailed View on Allometric Scaling. *PLOS ONE*, 6(7):e22732.
- NCD Risk Factor Collaboration (NCD-RisC), Abarca-Gómez, L., Abdeen, Z. A., Hamid, Z. A., Abu-Rmeileh, N. M., Acosta-Cazares, B., et al. (2017). Worldwide trends in body-mass index,

- underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. *Lancet (London, England)*, 390(10113):2627–2642.
- Nieman, D. C., Austin, M. D., Benezra, L., Pearce, S., McInnis, T., Unick, J., and Gross, S. J. (2006). Validation of Cosmed’s FitMate in measuring oxygen consumption and estimating resting metabolic rate. *Res. Sports Med.*, 14(2):89–96.
- Niyogi, S. P., Patwardhan, V. N., and Mordecai, J. (1939). Studies on Basal Metabolism in Bombay. Part I. *Indian Journal of Medical Research*, 27:99–113.
- Owen, O. E., Holup, J. L., D’Alessio, D. A., Craig, E. S., Polansky, M., Smalley, K. J., Kavle, E. C., Bushman, M. C., Owen, L. R., and M. A. Mozzoli, E. al. (1987). A reappraisal of the caloric requirements of men. *American Journal of Clinical Nutrition*, 46(6):875–885.
- Owen, O. E., Kavle, E., Owen, R. S., Polansky, M., Caprio, S., Mozzoli, M. A., Kendrick, Z. V., Bushman, M. C., and Boden, G. (1986). A reappraisal of caloric requirements in healthy women. *American Journal of Clinical Nutrition*, 44(1):1–19.
- Patil, S. R. and Bharadwaj, J. (2013). Development of new equations for basal metabolic rate for adolescent student Indian population. *Journal of Postgraduate Medicine*, 59(1):25.
- Pierson, R. N. J. (2005). Reference body composition tables. In Heymsfield, S. B., Lohman, T. G., Wang, Z. M., and Going, S. B., editors, *Human body composition*, pages 401–410. Human kinetics, 2nd edition.
- Poehlman, E. T. and Toth, M. J. (1995). Mathematical ratios lead to spurious conclusions regarding age- and sex-related differences in resting metabolic rate. *Am. J. Clin. Nutr.*, 61(3):482–485.
- Psota, T. and Chen, K. Y. (2013). Measuring energy expenditure in clinical populations: rewards and challenges. *European journal of clinical nutrition*, 67(5):436.
- Purcell, S. A., Johnson-Stoklossa, C., Tibaes, J. R. B., Frankish, A., Elliott, S. A., Padwal, R., and Prado, C. M. (2020). Accuracy and reliability of a portable indirect calorimeter compared to whole-body indirect calorimetry for measuring resting energy expenditure. *Clinical Nutrition ESPEN*, 39:67–73.

- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rahman, S. A. (1936). The basal metabolism of young men at Hyderabad (Deccan) with a study of their physical characters. *Indian Journal of Medical Research*, 24:173–199.
- Rahmandad, H. (2014). Human Growth and Body Weight Dynamics: An Integrative Systems Model. *PLoS One*, 9(12):e114609.
- Rajagopal, K. (1938). The Basal Metabolism of Indian and European Men on the Nilgiri Hills (S. India). *Indian Journal of Medical Research*, 26:411–26.
- Schofield, W. N. (1985). Predicting basal metabolic rate, new standards and review of previous work. *Human nutrition. Clinical nutrition*, 39 Suppl 1:5–41.
- Shirley, M. K., Arthurs, O. J., Seunarine, K. K., Cole, T. J., Eaton, S., Williams, J. E., Clark, C. A., and Wells, J. C. K. (2019). Metabolic rate of major organs and tissues in young adult South Asian women. *Eur. J. Clin. Nutr.*, 73:1164–1171.
- Snyder, W., Cook, M., Nasset, E., Karhausen, L., Howells, G., and Tipton, I. (1975). Report of the task group on reference man. *ICRP Publication 23*.
- Soares, M. J., Piers, L. S., O’Dea, K., and Shetty, P. S. (1998). No evidence for an ethnic influence on basal metabolism: an examination of data from India and Australia. *British Journal of Nutrition*, 79(4):333–341.
- Srivastava, R., Batra, A., Dhawan, D., and Bakhshi, S. (2017). Association of energy intake and expenditure with obesity: A cross-sectional study of 150 pediatric patients following treatment for leukemia. *Pediatric Hematology and Oncology*, 34(1):29–35.
- Swaminathan, S., Sinha, S., Minocha, S., Makkar, S., and Kurpad, A. (2018). Are we eating too much? A critical reappraisal of the energy requirement in Indians. *Proceedings of the Indian National Science Academy*.
- Swinburn, B. A., Jolley, D., Kremer, P. J., Salbe, A. D., and Ravussin, E. (2006). Estimating the effects of energy imbalance on changes in body weight in children. *Am. J. Clin. Nutr.*, 83(4):859–863.

- Talbot, F. B. (1938). BASAL METABOLISM STANDARDS FOR CHILDREN. *Archives of Pediatrics and Adolescent Medicine*, 55(3):455–459.
- Tschöp, M. H., Speakman, J. R., Arch, J. R. S., Auwerx, J., Brüning, J. C., Chan, L., Eckel, R. H., Farese, R. V., Galgani, J. E., Hambly, C., Herman, M. A., Horvath, T. L., Kahn, B. B., Kozma, S. C., Maratos-Flier, E., Müller, T. D., Münzberg, H., Pfluger, P. T., Plum, L., Reitman, M. L., Rahmouni, K., Shulman, G. I., Thomas, G., Kahn, C. R., and Ravussin, E. (2012). A guide to analysis of mouse energy metabolism. *Nat. Methods*, 9:57–63.
- Vandarakis, D., Salacinski, A. J., and Broeder, C. E. (2013). A Comparison of Cosmed Metabolic Systems for the Determination of Resting Metabolic Rate. *Res. Sports Med.*, 21(2):187–194.
- Wang, Z. (2012). High ratio of resting energy expenditure to body mass in childhood and adolescence: A mechanistic model. *American Journal of Human Biology*, 24(4):460–467.
- Wang, Z., Heshka, S., Heymsfield, S. B., Shen, W., and Gallagher, D. (2005). A cellular-level approach to predicting resting energy expenditure across the adult years. *American Journal of Clinical Nutrition*, 81(4):799–806.
- Wang, Z., Heshka, S., Zhang, K., Boozer, C. N., and Heymsfield, S. B. (2001). Resting Energy Expenditure: Systematic Organization and Critique of Prediction Methods\*. *Obesity*, 9(5):331–336.
- Wang, Z., Heymsfield, S. B., Ying, Z., Pierson, Jr., R. N., Gallagher, D., and Gidwani, S. (2010a). A Cellular Level Approach to Predicting Resting Energy Expenditure: Evaluation of Applicability in Adolescents. *American journal of human biology : the official journal of the Human Biology Council*, 22(4):476.
- Wang, Z., Ying, Z., Bosy-Westphal, A., Zhang, J., Schautz, B., Later, W., Heymsfield, S. B., and Müller, M. J. (2010b). Specific metabolic rates of major organs and tissues across adulthood: evaluation by mechanistic model of resting energy expenditure. *American Journal of Clinical Nutrition*, 92(6):1369.
- Wells, J. C. (2011). The thrifty phenotype: an adaptation in growth or metabolism? *American Journal of Human Biology*, 23(1):65–75.
- WHO (2016). The double burden of malnutrition: Policy Brief. *WHO*, (5):12.

- Wlodek, M. E., Westcott, K., Siebel, A. L., Owens, J. A., and Moritz, K. M. (2008). Growth restriction before or after birth reduces nephron number and increases blood pressure in male rats. *Kidney Int.*, 74(2):187–195.
- Yajnik, C. (2002). The lifecycle effects of nutrition and body size on adult adiposity, diabetes and cardiovascular disease. *Obesity Reviews*, 3(3):217–224.
- Yajnik, C. S., Fall, C., Coyaji, K. J., Hirve, S., Rao, S., Barker, D., Joglekar, C., and Kellingray, S. (2003). Neonatal anthropometry: the thin–fat indian baby. the pune maternal nutrition study. *International journal of obesity*, 27(2):173–180.

## CHAPTER 4

# Modelling height growth in Indian children and adolescents

*Published as and adapted from:*

**Sandra Aravind Areekal**, Pranay Goel, Anuradha Khadilkar, Vaman Khadilkar & Tim J. Cole (2022) Assessment of height growth in Indian children using growth centiles and growth curves, *Annals of Human Biology*, 49:5-6, 228-235, DOI: 10.1080/03014460.2022.2107238

## 4.1 Introduction

Growth centiles and growth curves represent two distinct ways to assess child growth. Height *centiles* (or per-centiles (Galton 1885)) describe the distribution of height in a population with respect to age; they are typically described through evenly spaced centiles on a *growth chart*, where each centile is labelled according to the percentage of the population below it (Bowditch 1891; World Health Organization 2006). In contrast height growth *curves* show serial changes in height with age in individuals, providing information on their height velocity as well as their height attained (Tanner 1962; Komlos et al. 1992).

The two methods have evolved separately and they have different aims. Primarily they differ in the type of data used for their construction. Growth centiles are based on cross-sectional data, where the measurements—usually one per subject—are treated as independent, covering the age range under study (Cole 2012). Growth curves, on the other hand, are developed from longitudinal data, utilising repeated measurements for each child (Johnson 2015; Crozier et al. 2019).

The construction of growth centiles requires modelling the frequency distribution of the measurement at each age and smoothing the centiles across age. The World Health Organization (Borghi et al. 2006) recommends using Generalized Additive Models for Location Scale and Shape (GAMLSS) (Rigby and Stasinopoulos 2005a) to construct the centiles, which models moments of the distribution as smooth curves in age.

Growth curve modelling involves summarising the shape of individual curves as a mean curve. Early methods—known as *parametric* or structural models—used parametric functions applied to data for individuals to describe the curve shape (Jenss and Bayley 1937; Preece and Baines 1978; Karlberg 1989). Later, *semi-parametric* or non-structural models were developed using fractional polynomials or cubic splines to estimate the curve shape, optimising the number of parameters based on the data—this provides extra flexibility but can lack biological interpretability (Hauspie et al. 2004; Chirwa et al. 2014).

In the 1980s hierarchical mixed-effects models including fixed effects and random effects became available, capturing both individual variation and the population trend in a unified framework (Goldstein 1986). The growth curve model SuperImposition by Translation And Rotation (SITAR) (Cole et al. 2010) is a semi-parametric hierarchical model with three subject-



specific random effects that together explain up to 99% of the variance in longitudinal height growth.

Even though growth centiles and growth curves are different, the GAMLSS and SITAR models to fit them have some similarities; both are semi-parametric, using cubic splines to estimate curve shape, and both—in their most common form—involve three underlying summary statistics (three moments for GAMLSS and three random effects for SITAR). It is useful to fit the two models to the same data, both to compare the shapes of the resulting centiles and curves, and to emphasise their different purpose. A few previous studies have applied the two methods to the same data (Blackwell et al. 2017; Spencer et al. 2018), but none have directly compared them.

Height centiles for Indian children have been available and regularly updated since 1992 (Agarwal et al. 1992; Khadilkar et al. 2007; Indian Academy Of Pediatrics Growth Charts Committee et al. 2015; Khadikar et al. 2020). However there have been only small studies of Indian height growth curves (Hauspie et al. 1980; Satyanarayana et al. 1989; de Onis et al. 2001; Mirzaei and Sengupta 2012).

The motivation behind this study is threefold: i) to construct growth centiles and growth curves using the GAMLSS and SITAR models, respectively, using a large, recent longitudinal height dataset of Indian children aged between 6 and 18 years; ii) to compare the shapes of the resulting curves, highlighting particularly the pubertal growth spurt, summarised as the mean peak height velocity (PV) and the mean age at peak height velocity (APV), and iii) to compare the centiles with others published for Indian children.

## 4.2 Methods

### 4.2.1 Data sets

The data came from the Pune School Children Growth study (PSCG), consisting of age and height measurements of 1472 affluent urban children (798 boys and 674 girls) aged 3 to 18 years living in Pune, Western India (Khadilkar et al. 2019). The data were collected annually between 2007 and 2013, with a median of 6 (range 1 to 6) measurements per subject. Height was measured using a portable stadiometer (Leicester Height Meter; Child Growth Foundation, London, UK) calibrated with a standard height rod. Further details of the PSCG data collection can be found

in Khadilkar et al. (2019). Written informed consent was obtained from the parents of the subjects, and verbal assent was obtained from subjects aged over 7 years. The data collection was approved by the Ethics Committee of the Jehangir Clinical Development Centre Pune (dated 26th June 2007). Permission for secondary data analysis was obtained from the Ethics Committee of the Indian Institute of Science Education and Research Pune (IECHR/Admin/2021/001).

## 4.2.2 Models

### 4.2.2.1 Cross-sectional GAMLSS model for growth centiles

Growth centiles are traditionally created using cross-sectional data, independent observations,  $y_i$ , for  $i \in \{1, 2, \dots, N\}$ . The state of the art method to construct growth centiles is Generalized Additive Models for Location, Scale and Shape (GAMLSS) (Rigby and Stasinopoulos 2005a). The data are assumed to come from a distribution  $f_Y(y)$  whose first four moments are the mean ( $\mu$ ), standard deviation or coefficient of variation ( $\sigma$ ), skewness ( $\nu$ ) and kurtosis ( $\tau$ ). Classical linear regression assumes a normal distribution for  $f_Y(y)$ , with  $\mu$  linearly related to the explanatory variable (here age), constant  $\sigma$ , zero  $\nu$  and  $\tau$  fixed at 3. GAMLSS by contrast allows  $f_Y(y)$  to be selected from a wide range of available distributions, and the moments can be specified as functions of age or more generally.

The normal distribution (called NO in GAMLSS) estimates  $\mu$  and  $\sigma$  as curves in age, and ignores skewness and kurtosis. But for other distributions  $\nu$  and/or  $\tau$  can be explicitly estimated from the data. For centile estimation GAMLSS has three distributions, which all raise  $y$  to Box-Cox power  $\nu$ . To model skewness, there is the Box-Cox Cole and Green distribution (BCCG), while if both skewness and kurtosis are present, it has the Box-Cox power exponential (BCPE) and Box-Cox  $t$  (BCT) distributions (Cole and Green 1992; Rigby and Stasinopoulos 2004, 2006). Note that the BCCG distribution is equivalent to the LMS method (Cole and Green 1992), where  $\nu$  is called  $\lambda$  and transformed  $y$  is standard normally distributed. With BCPE, transformed  $y$  assumes a standard power exponential distribution, and with BCT it follows a Student's  $t$  distribution. GAMLSS fits the three distributions with identity links for  $\mu$  and  $\nu$ , and a log link for  $\sigma$ . The distributions can also be fitted with a log link for  $\mu$ , denoted by GAMLSS as BCCGo, BCPEo, and BCTo, respectively. The spline curves in age for each moment were fitted using penalised B-splines or P-splines (Eilers and Marx 1996), with the default degrees of freedom ( $df$ ) estimated by cross-validation.

Pseudo-velocity curves were constructed for the GAMLSS models by differentiating the median ( $\mu$ ) curves, and APV and PV were identified as the age and value of peak “velocity”; note that this does not represent true velocity as the data are cross-sectional, but for simplicity it is referred to here as velocity.

Despite being longitudinal, the PSCG data were treated as cross-sectional for the GAMLSS analysis. With balanced data such as the PSCG this works fine, as the centiles are unbiased. However they are less precise than for cross-sectional data, being based on fewer subjects.

#### 4.2.2.2 Longitudinal SITAR model for growth curves

The SuperImposition by Translation And Rotation (SITAR) model (Cole et al. 2010) describes the height  $y_{i,j}$  of individual  $i$  (where  $i \in \{1, 2, \dots, n\}$ ) at time  $t_j$  as,

$$y_{i,j} = a_0 + \alpha_i + H\left(\frac{g(t_j) - b_0 - \beta_i}{\exp(-c_0 - \gamma_i)}\right) + \epsilon_i, \quad (4.1)$$

where  $a_0$ ,  $b_0$ , and  $c_0$  are fixed effects,  $\alpha_i$ ,  $\beta_i$  and  $\gamma_i$  are subject-specific random effects named size, timing and intensity, respectively,  $g(t)$  denotes a link function for age such as the log or power transformation,  $H(t)$  is the population average height curve fitted using a natural cubic B-spline and  $\epsilon$  is the independent and identically distributed (i.i.d.) random error term. The random effects are assumed to be normally distributed with SD estimated from the sample. The B-spline curve  $H(t)$  has a number of knots which can be tuned by optimising their number and position on the age scale—this defines the degrees of freedom ( $df$ ) of the curve.

The three subject-specific random effects can be interpreted biologically as follows: size  $\alpha_i$  represents the subject’s offset compared to mean height, that is, by how much they are taller or shorter than average, adjusted for age. Timing of the pubertal growth spurt  $\beta_i$  measures by how much the subject’s age at peak height velocity (APV) occurs earlier or later than the population-average APV (the velocity curve is estimated as the first derivative of  $H(t)$ ). Intensity of the growth spurt  $\gamma_i$  indicates by how much the subject’s peak height velocity (PV) is higher or lower than the average PV, measured as a proportion.

### 4.2.3 Data analysis

Boys and girls were analysed separately using the packages *gamlss* (version 5.3.4) (Rigby and Stasinopoulos 2005b) and *sitar* (version 1.2.0) (Cole 2021) available in the statistical language R (version 4.0.5) (R Core Team 2021). Growth velocity in early life is much higher and more variable than later in childhood, and to avoid this dominating the analysis the data were restricted to the age range 6 to 18 years.

The data were initially cleaned by removing obvious errors using conditional plots of height on age (3 points in boys and 3 in girls). All subjects with at least one measurement were included in the analysis. A preliminary SITAR model was fitted, and points with standardised residuals beyond  $\pm 4$  SD were considered as outliers and excluded from the analysis (10 in boys and 4 in girls). The final analysis included 796 boys with 4242 measurements and 672 girls with 3539 measurements (0.3% and 0.2% excluded respectively).

Multiple plausible GAMLSS models of height on age were fitted: comparing the NO, BCCGo, BCPEo and BCTo distributions; and trying square root and log transformations for height and age. Similarly, different SITAR models were fitted varying the  $df$  of the spline curve from 4 to 9; omitting the fixed effects  $b_0$  and/or  $c_0$ , and with log transformations for height and age. The optimal model was selected by minimising the Bayesian Information Criterion (BIC) (Schwarz 1978).

Standard errors (SE) for mean APV and mean PV based on GAMLSS and SITAR were obtained with the bootstrap (500 samples). Diagnostic plots for the optimal models are given in supplementary figures 4.7 and 4.8.

## 4.3 Results

### 4.3.1 Height centiles in Indian children using GAMLSS

Height centiles for the PSCG data set are constructed using GAMLSS with the BCCGo distribution. However the distribution for girls is not skew at any age, so the Box-Cox power  $\nu$  is constrained to 1, equivalent to a normal distribution. The fitted centiles are shown by sex in Figure 4.1: the nine curves are spaced two-thirds of a z-score apart, extending from the 0.4<sup>th</sup>

to the 99.6<sup>th</sup> centile (Cole 1994). The individual heights are also shown as grey points. Each

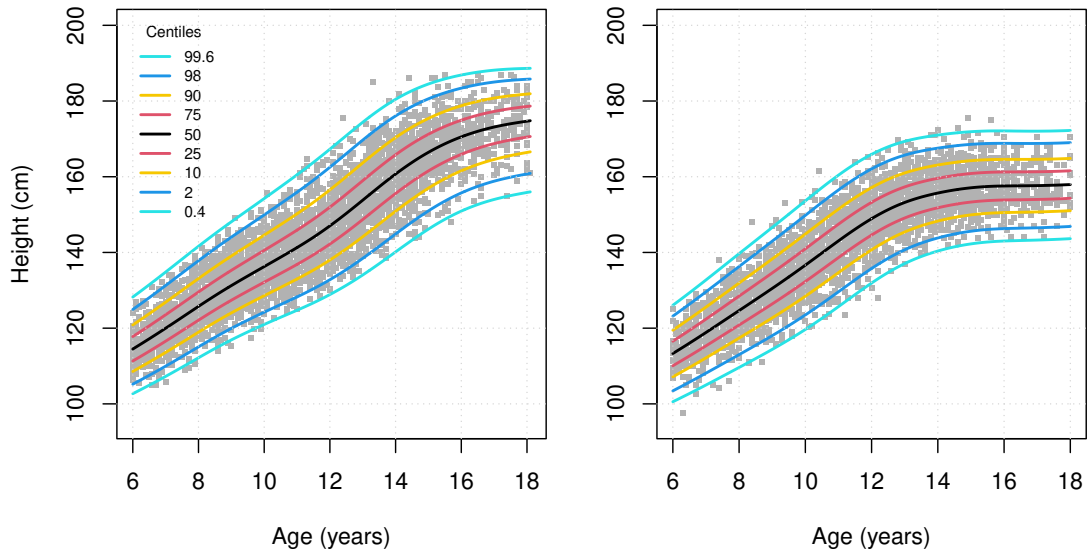


Figure 4.1: GAMLSS-fitted height centile curves based on the BCCGo distribution in boys and the normal distribution in girls. The nine centiles are equally spaced on the z-score scale. Individual heights are shown in grey ( $n = 4242$  for boys and  $n = 3539$  for girls).

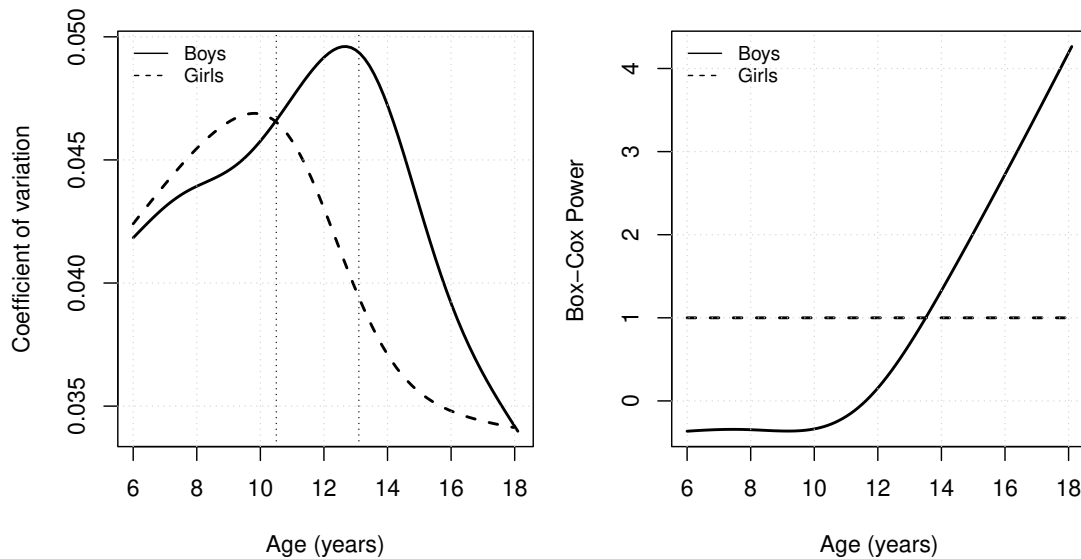


Figure 4.2: Fitted age trends for the BCCGo coefficient of variation ( $\sigma$ ) and skewness ( $\nu$ ) for boys (solid lines) and girls (dashed lines). The vertical dotted lines indicate APV by sex based on the median ( $\mu$ ) curve.

median curve corresponds to the age trend for  $\mu$ , while the age trends for  $\sigma$  and  $\nu$  are shown in Figure 4.2.

The  $\sigma$  curves peak at around 13 years in boys and 10 years in girls, similar to APV as based on the  $\mu$  curve. In boys the  $\nu$  curve rises steeply through puberty, indicating increasing left

skewness.

### 4.3.2 Height growth curves in Indian children using SITAR

Individual height growth curves in the PSCG data set are best described by SITAR models of height on log age, with 6  $df$  in boys and fixed effects for size, timing, and intensity; and 5  $df$  in girls with fixed effects for size and intensity - the models explain 98.7% of the variance in boys and 98.8% in girls. Figure 4.3 shows the individual height curves in grey. Each grey curve,

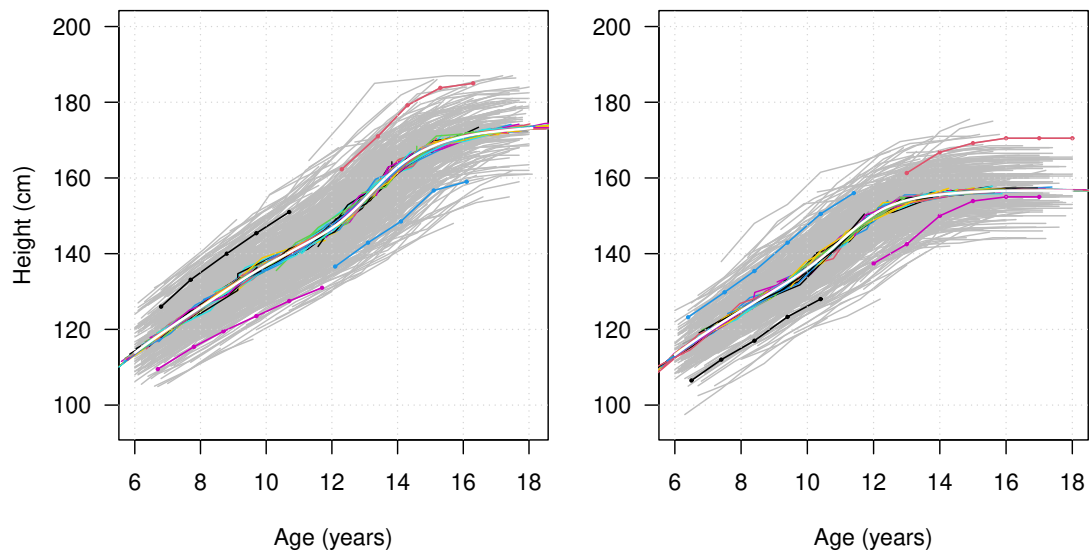


Figure 4.3: Individual height growth curves for 796 boys and 672 girls from the PSCG data set. Unadjusted curves are in grey, and curves adjusted using subject-specific SITAR random effects are in colour. The fitted mean curve is superimposed on the coloured curves and partially obscures them. Four unadjusted curves in colour show the PSCG study design.

when adjusted by the individual's fitted random effects, provides an estimate of the mean curve, and the adjusted curves appear colour-coded in Figure 4.3 with the mean curve in white. Four unadjusted curves are shown in colour to highlight the PSCG study design, where individuals had no more than six annual measurements, yet SITAR was able to estimate the entire mean curve from 6 to 18 years. The fitted mean curves by sex are shown in Figure 4.4 along with the mean height velocity curves, calculated as the first derivative of the mean curve.

In boys, mean APV (95% CI) is 13.1 (13.0, 13.3) years, and mean PV is 9.0 (8.7, 9.3) cm/year. A small peak similar to a mid-growth spurt (Tanner and Cameron 1980) is seen at 8.6 years with mean PV 5.7 cm/year. In girls, mean APV occurs at 11.0 (10.8, 11.2) years with mean PV 8.0 (7.8, 8.2) cm/year. Summary statistics of the SITAR models are shown by sex in

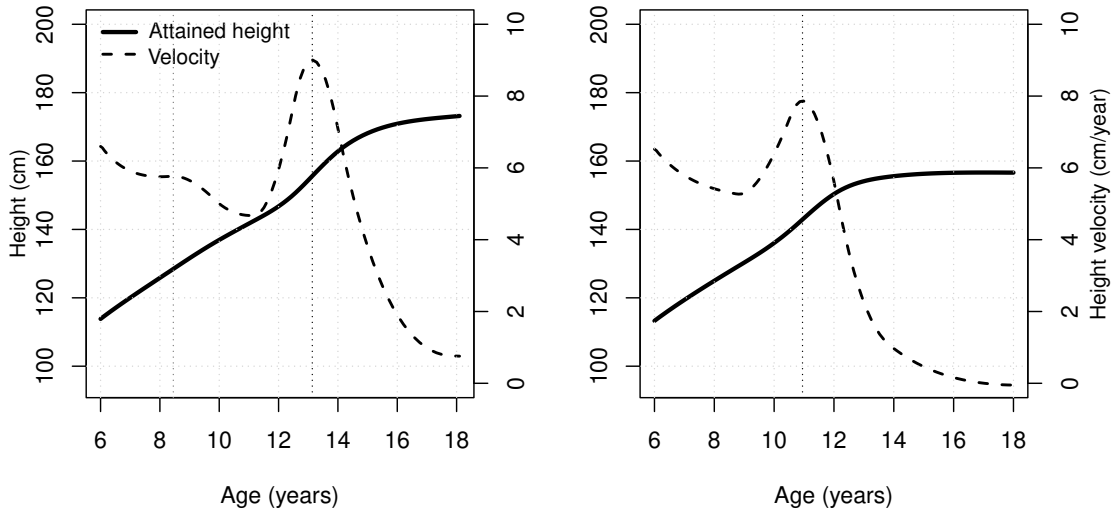


Figure 4.4: SITAR-fitted mean height growth curves (solid) and height velocity curves (dashed) by sex in the PSCG data set. The vertical dotted lines indicate the ages at peak height velocity (at 13.1 years in boys and 11.0 years in girls) and the boys’ mid-growth spurt (at 8.6 years).

Table 4.1. Quantile-quantile (Q-Q) plots of the three random effects show that size is normally distributed; however, timing and intensity deviate from normality in the lower tail (Figure 4.8). The correlations between random effects are shown in Table 4.2. A scatter plot matrix showing these correlations appears in Figure 4.9.

Table 4.1: Standard deviations of SITAR random effects and the residual standard deviation by sex.

	Boys ( $n = 796$ )	Girls ( $n = 672$ )
Size (cm)	6.85	6.14
Timing (fractional)	0.084	0.085
Intensity (fractional)	0.15	0.14
Residual (cm)	0.76	0.67

Table 4.2: Correlations between SITAR random effects by sex. See Figure 4.9 for the scatter plot matrix.

	Boys ( $n = 796$ )		Girls ( $n = 672$ )	
	Size	Timing	Size	Timing
Timing	0.31		0.34	
Intensity	0.48	0.45	0.44	0.31

Figure 4.5 shows mean height velocity curves for GAMLSS (dashed lines) and SITAR (solid lines) by sex, estimated as derivatives of the mean and median curves, respectively, with 95% confidence bands. There is a small peak in boys around 8 years in both curves. APV and PV for the curves are shown in Table 4.3; the APVs are similar, but the PVs are appreciably smaller for GAMLSS, and the confidence intervals do not overlap, due to GAMLSS being based on cross-sectionally analysed data.

Figure 4.6 compares height centiles recommended for Indian children by the Indian Academy

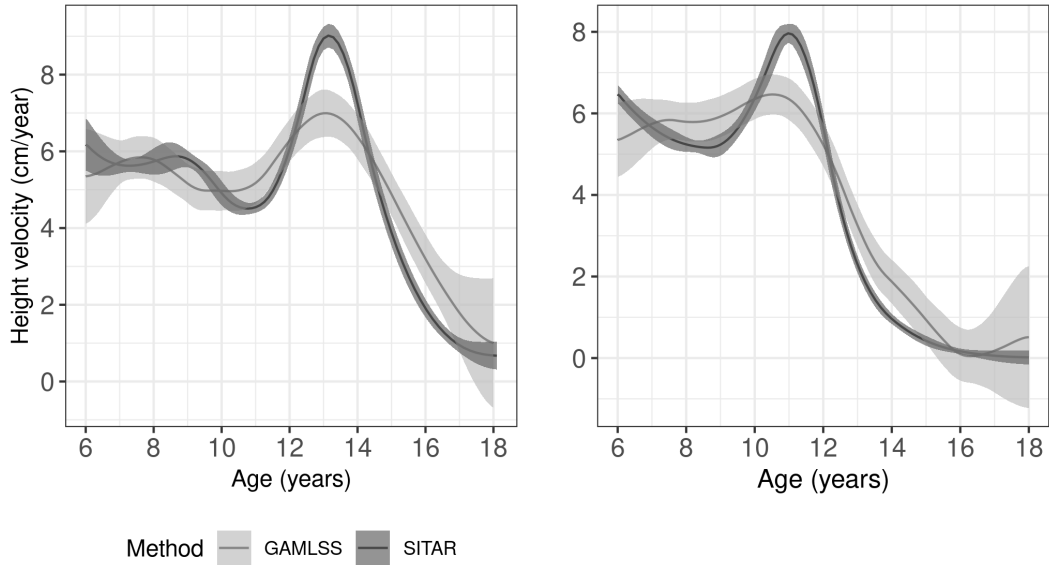


Figure 4.5: Mean height velocity curves and 95% bootstrap confidence interval bands as estimated by SITAR (dark gray) and GAMLSS (light gray) by sex (boys left, girls right).

Table 4.3: APV and PV as estimated from the SITAR and GAMLSS velocity curves by sex.

	Boys		Girls	
	SITAR	GAMLSS	SITAR	GAMLSS
APV (95% CI) years	13.1 (13.0, 13.3)	13.1 (12.4, 13.6)	11.0 (10.8, 11.2)	10.3 (8.3, 11.2)
PV (95% CI) cm/year	9.0 (8.7, 9.3)	7.1 (6.5, 7.7)	8.0 (7.8, 8.2)	6.6 (6.2, 7.0)

of Pediatrics (IAP) (Indian Academy Of Pediatrics Growth Charts Committee et al. 2015) with the GAMLSS-estimated PSCG centiles. The IAP centiles were constructed using the LMS method on a much larger cross-sectional data set. The 3<sup>rd</sup>, 50<sup>th</sup> and 97<sup>th</sup> IAP centiles are drawn as dashed red curves superimposed on the 3<sup>rd</sup>, 50<sup>th</sup> and 97<sup>th</sup> PSCG centiles shown in black. The IAP and PSCG medians match fairly well, but the outer centiles less so.

## 4.4 Discussion

Growth centiles and growth curves provide two distinct perspectives on individual and population growth. In this study, we analysed height growth in the PSCG data set to contrast the two tools in an Indian context. We constructed growth centiles with GAMLSS and growth curves with SITAR based on Indian children aged 6 to 18 years. The distribution of height in the population is well captured by the GAMLSS BCCGo centiles (Figure 4.1), which show the coefficient of variation peaking at around 13 years in boys and 10 years in girls (Figure 4.2).



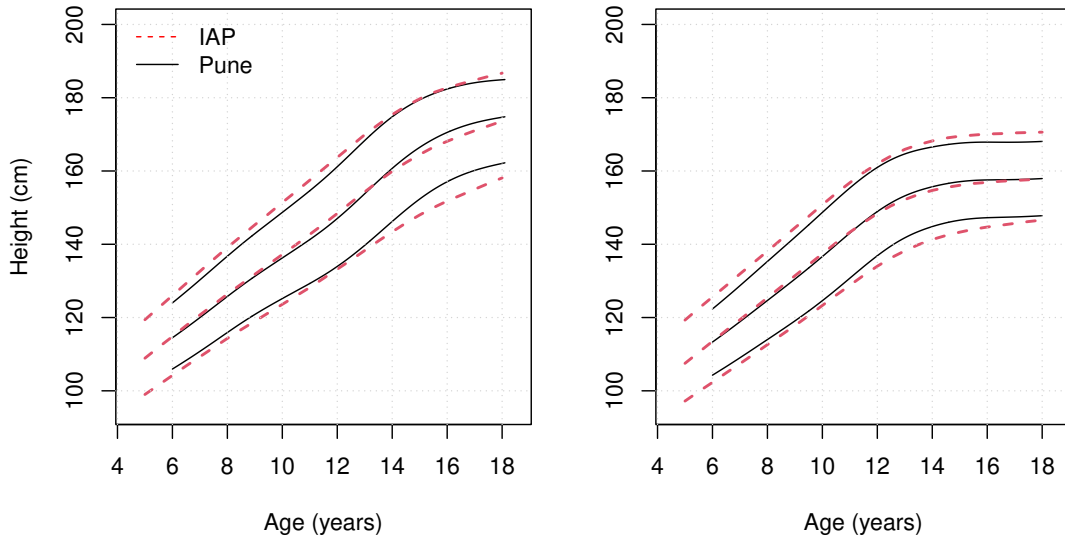


Figure 4.6: Comparing height centiles based on the PSCG data set and those recommended for the Indian population by the IAP (Indian Academy Of Pediatrics Growth Charts Committee et al. 2015). The black curves show the PSCG 3<sup>rd</sup>, 50<sup>th</sup>, and 97<sup>th</sup> centiles, while the red dashed lines are the IAP 3<sup>rd</sup>, 50<sup>th</sup>, and 97<sup>th</sup> centiles.

There is increasing left skewness in the boys' data (but not the girls') after 10 years, suggesting that a minority of boys fall behind progressively in height during puberty, extending the lower tail of the distribution. SITAR models the pubertal growth spurt effectively and explains 98.7% of the observed variance in boys and 98.8% in girls, shrinking the height SD from 7 cm to 0.7 cm. Boys also show a small mid-growth spurt, peaking at 8.6 years.

SITAR works well with the PSCG study design in that even though individuals are followed for no more than six years, the mean growth curve is estimated over the twelve-year period from 6 to 18 years—SITAR is able to combine the individual growth curves to obtain the bigger picture.

The two methods, GAMLSS and SITAR, when applied to the same data, provide distinct information that is relevant in different contexts. Centile charts are useful in *clinical medicine* as they illustrate the child's recent growth as compared to their contemporaries, and this helps the clinician to make decisions about their management. The chart can also be used to see how fast the child has been growing and predict how fast they will grow in the future, depending on their treatment.

SITAR in contrast is of little value in clinical medicine—it works with the child's entire growth curve, and so it cannot provide "real time" growth information relevant for management.

Instead it is valuable in two other contexts: *experimental medicine* and *life course epidemiology*. The SITAR mean growth curve is useful in experimental studies such as randomised clinical trials investigating the effect of a growth promoting agent such as oxandrolone in Turner Syndrome or calcium supplementation in rural Gambian children (Gault et al. 2011; Prentice et al. 2012). The SITAR model can test whether the intervention affects the size and/or timing and/or intensity of the mean curve. Separately, the SITAR subject random effects summarise individual pubertal growth patterns which can be related to individual-level stressors and adverse outcomes later in the life course (Johnson et al. 2014; Pizzi et al. 2014; Filteau et al. 2019). For example Kuh et al. (2016) have shown that early puberty is associated with better bone health in later life.

The height velocity curves estimated by GAMLSS and SITAR provide insights into the timing and intensity of the growth spurt. The APVs are comparable, but the PVs are appreciably smaller with GAMLSS. Mean APV was 13.1/10.3 years for boys/girls with GAMLSS and 13.1/11.0 years with SITAR, respectively 2.7 and 2.1 years apart (Figure 4.4). However mean PV was only 7.1/6.6 cm/year by sex with GAMLSS, some 20% smaller compared to 9.0/8.0 cm/year with SITAR. Similar results were observed by Blackwell et al. (2017), who applied the two methods to the same mixed longitudinal data set. Merrell (1931) and Cole et al. (2008) have explained this discrepancy algebraically: when individuals vary in their APV (i.e. the SD of the timing random effect is greater than zero), the mean PV based on cross-sectional data is attenuated compared to that based on longitudinal data. Note too that the velocity confidence bands for SITAR in Figure 4.4 are narrower than for GAMLSS, showing that the variance explained by SITAR is higher than for GAMLSS due to its longitudinal analysis.

Another estimate of PV came from the Indian height velocity charts of Khadilkar et al. (2019) based on annual height measurements: they reported median APV (PV) in boys and girls as 13.5 years (6.8 cm/year) and 10.5 years (PV 6.6 cm/year). The APVs are slightly later than for PSCG GAMLSS, while the PVs are similar. The Khadilkar et al. (2019) charts were constructed based on individual height velocity, i.e. year-wise differences in height used to construct the median height velocity centile, whereas GAMLSS here uses the median height centile. In practice the two should be similar, since the mean height increment is equal to the difference in the mean heights.

Reference height centile charts for Indian children have been developed by the IAP (Indian Academy Of Pediatrics Growth Charts Committee et al. 2015). They are useful for documenting the high prevalence of stunting in India (Hemalatha et al. 2020; Kumar et al. 2021). We compared the IAP charts, which are recommended for clinical use, with the GAMLSS-modelled PSCG

centiles. The median curves are in good agreement throughout the age range in girls and until puberty in boys. However, the 3<sup>rd</sup> and 97<sup>th</sup> centiles match less well, especially the 3<sup>rd</sup> centile, which is the formal cut-off to define stunting in India. The percentages of PSCG children below the IAP 3<sup>rd</sup> centile are 1.5% of boys and 1.4% of girls, about half the expected rate. This discrepancy is due mainly to the smaller  $df$  used for the IAP  $\mu$  curve, leading to the IAP centiles being stiffer and hence more linear, particularly during puberty. In addition the IAP charts were based on a nationally representative sample (unlike the PSCG children living in Pune) and hence were more heterogeneous, and this could explain the IAP outer centiles being more widely spaced.

A small growth spurt before puberty, called the mid-growth spurt, was previously reported in boys between 5.9 and 8.5 years (El Lozy 1978; Tanner 1962; Tanner and Cameron 1980; Molinari et al. 1980; Gasser et al. 1985; Remer and Manz 2001; Virani 2005). In particular Virani (2005) reported a mid-growth spurt in Indian boys at 6.2 years. There was also a mid-growth spurt at 9 years in a subgroup of boys with late puberty in the Harpenden Growth Study (Cole 2020). We observe a small peak in height velocity around 8 years in both the GAMLSS and SITAR models. Note that the SITAR confidence band in Figure 4.5 is wider at 8 years, which may indicate variation in the timing of the mid-growth spurt.

In boys, the GAMLSS median curve and the SITAR mean curve have not yet plateaued by the age of 18 years. Further, their height velocity is appreciably greater than zero at 18 years (Figure 4.5). Conversely for girls at 18 years, the height curves are flat and the mean velocity is zero. This indicates that unlike girls, boys continue growing after 18 years, and ideally the reference sample should extend into the third decade of life to document this growth period properly.

A strength of our study is that we provide a comparative analysis of two widely used methods of growth curve and growth centile analysis. We show that the height velocity curves estimated from the two methods are appreciably different. Further, while the GAMLSS model has been used to develop growth centiles in the Indian population, SITAR has not been previously applied. However there are also some limitations. Our analysis is restricted to a population in Western India, and GAMLSS is applied to longitudinal rather than cross-sectional data, which limits the number of subjects included in the analysis. The centiles created from longitudinal data should be unbiased as the measurements were made annually within a fixed protocol (Wade and Kurmanavicius 2009), but less precise.

In conclusion, we have shown different aspects of growth centile and growth curve analysis that are used widely to analyse growth from birth to maturity. The GAMLSS model captures the distribution of height by age which can be displayed as a growth chart, whereas the SITAR model estimates the shape of the mean height growth curve, which also applies to individuals. The pubertal peak in height velocity is shallower in GAMLSS centiles compared to SITAR curves. We believe that the two analyses add usefully to knowledge about growth in contemporary Indian children.

## Software

The growth centiles and curves developed here are made available as a web application accessible at <https://digimed.acads.iiserpune.ac.in/growth-charts>.

## Bibliography

- Agarwal, D. K., Agarwal, K. N., Upadhyay, S. K., Mittal, R., Prakash, R., and Rai, S. (1992). Physical and sexual growth pattern of affluent Indian children from 5 to 18 years of age. *Indian Pediatrics*, 29(10):1203–1282.
- Blackwell, A. D., Urlacher, S. S., Beheim, B., von Rueden, C., Jaeggi, A., Stieglitz, J., Trumble, B. C., Gurven, M., and Kaplan, H. (2017). Growth references for Tsimane forager-horticulturalists of the Bolivian Amazon. *American journal of physical anthropology*, 162(3):441.
- Borghini, E., de Onis, M., Garza, C., Van den Broeck, J., Frongillo, E. A., Grummer-Strawn, L., Van Buuren, S., Pan, H., Molinari, L., Martorell, R., et al. (2006). Construction of the World Health Organization child growth standards: selection of methods for attained growth curves. *Statistics in medicine*, 25(2):247–265.
- Bowditch, H. (1891). The growth of children studied by Galton’s percentile grades In 22nd annual report of the State Board of Health of Massachusetts (pp. 479–525). *Boston, MA: Wright and Potter*.
- Chirwa, E. D., Griffiths, P. L., Maleta, K., et al. (2014). Multi-level modelling of longitudinal child growth data from the Birth-to-Twenty Cohort: a comparison of growth models. *Annals of Human Biology*, 41(2):168.

- Cole, T. (1994). Do growth chart centiles need a face lift? *Bmj*, 308(6929):641–642.
- Cole, T. J. (2012). The development of growth references and growth charts. *Ann. Hum. Biol.*, 39(5):382.
- Cole, T. J. (2020). Tanner’s tempo of growth in adolescence: recent SITAR insights with the Harpenden Growth Study and ALSPAC. *Ann. Hum. Biol.*, 47(2):181–198.
- Cole, T. J. (2021). *sitar: Super Imposition by Translation and Rotation Growth Curve Analysis*. R package version 1.2.0.9000.
- Cole, T. J., Cortina-Borja, M., Sandhu, J., Kelly, F. P., and Pan, H. (2008). Nonlinear growth generates age changes in the moments of the frequency distribution: the example of height in puberty. *Biostatistics*, 9(1):159–171.
- Cole, T. J., Donaldson, M. D. C., and Ben-Shlomo, Y. (2010). SITAR—a useful instrument for growth curve analysis. *Int. J. Epidemiol.*, 39(6):1558.
- Cole, T. J. and Green, P. J. (1992). Smoothing reference centile curves: The lms method and penalized likelihood. *Stat. Med.*, 11(10):1305–1319.
- Crozier, S. R., Johnson, W., Cole, T. J., Macdonald-Wallis, C., Muniz-Terrera, G., Inskip, H. M., and Tilling, K. (2019). A discussion of statistical methods to characterise early growth and its impact on bone mineral content later in childhood. *Annals of Human Biology*, 46(1):17.
- de Onis, M., Dasgupta, P., Saha, S., Sengupta, D., and Blössner, M. (2001). The National Center for Health Statistics reference and the growth of Indian adolescent boys. *American Journal of Clinical Nutrition*, 74(2):248–253.
- Eilers, P. H. C. and Marx, B. D. (1996). Flexible Smoothing with *B*-splines and Penalties. *Statistical Science*, 11(2):89–102.
- El Lozy, M. (1978). A critical analysis of the double and triple logistic growth curves. *Ann. Hum. Biol.*, 5(4):389–394.
- Filteau, S., Kumar, G. T., Cole, T. J., et al. (2019). Steady Growth in Early Infancy Is Associated with Greater Anthropometry in Indian Children Born Low Birth Weight at Term. *Journal of Nutrition*, 149(9):1633–1641.

- Galton, F. (1885). Anthropometric per-centiles. *Nature*, 31(793):223–225.
- Gasser, T., Müller, H., Köhler, W., et al. (1985). An analysis of the mid-growth and adolescent spurts of height based on acceleration. *Ann. Hum. Biol.*, 12(2):129–148.
- Gault, E. J., Perry, R. J., Cole, T. J., et al. (2011). Effect of oxandrolone and timing of pubertal induction on final height in Turner’s syndrome: randomised, double blind, placebo controlled trial. *BMJ*, 342.
- Goldstein, H. (1986). Efficient statistical modelling of longitudinal data. *Annals of Human Biology*, 13(2):129–141.
- Hauspie, R. C., Cameron, N., and Molinari, L. (2004). *Methods in human growth research*, volume 39. Cambridge University Press.
- Hauspie, R. C., Das, S. R., Preece, M. A., and Tanner, J. M. (1980). A longitudinal study of the growth in height of boys and girls of West Bengal (India) aged six months to 20 years. *Annals of Human Biology*, 7(5):429–440.
- Hemalatha, R., Pandey, A., Kinyoki, D., et al. (2020). Mapping of variations in child stunting, wasting and underweight within the states of India: the Global Burden of Disease Study 2000–2017. *eClinicalMedicine*, 22.
- Indian Academy Of Pediatrics Growth Charts Committee, Khadilkar, V., Yadav, S., Agrawal, K. K., Tamboli, S., Banerjee, M., Cherian, A., Goyal, J. P., Khadilkar, A., Kumaravel, V., Mohan, V., Narayanappa, D., Ray, I., and Yewale, V. (2015). Revised IAP growth charts for height, weight and body mass index for 5- to 18-year-old Indian children. *Indian Pediatrics*, 52(1):47–55.
- Jenss, R. M. and Bayley, N. (1937). A mathematical method for studying the growth of a child. *Human Biology*, 9(4):556.
- Johnson, L., van Jaarsveld, C. H. M., Llewellyn, C. H., et al. (2014). Associations between infant feeding and the size, tempo and velocity of infant weight gain: SITAR analysis of the Gemini twin birth cohort. *International Journal of Obesity*, 38:980–987.
- Johnson, W. (2015). Analytical strategies in human growth research. *American Journal of Human Biology*, 27(1):69–83.

- Karlberg, J. (1989). A biologically-oriented mathematical model (ICP) for human growth. *Acta Paediatr. Scand. Suppl.*, 350(70-94.);.
- Khadilkar, V., Khadilkar, A. V., Lohiya, N. N., and Karguppikar, M. B. (2020). Extended growth charts for Indian children. *Journal of Pediatric Endocrinology & Metabolism : JPEM*, 34(3):357–362.
- Khadilkar, V., Khadilkar, A., Arya, A., Ekbote, V., Kajale, N., Parthasarathy, L., Patwardhan, V., Phanse, S., and Chiplonkar, S. (2019). Height Velocity Percentiles in Indian Children Aged 5-17 Years. *Indian Pediatr.*, 56(1):23–28.
- Khadilkar, V. V., Khadilkar, A. V., Choudhury, P., Agarwal, K. N., Ugra, D., and Shah, N. K. (2007). IAP growth monitoring guidelines for children from birth to 18 years. *Indian Pediatrics*, 44(3):187–197.
- Komlos, J., Tanner, J. M., Davies, P. S. W., and Cole, T. (1992). The growth of boys in the Stuttgart Carlschule, 1771–93. *Annals of Human Biology*, 19(2):139–152.
- Kuh, D., Muthuri, S. G., Moore, A., Cole, T. J., Adams, J. E., Cooper, C., Hardy, R., and Ward, K. A. (2016). Pubertal timing and bone phenotype in early old age: findings from a british birth cohort study. *International journal of epidemiology*, 45(4):1113–1124.
- Kumar, P., Chauhan, S., Patel, R., Srivastava, S., and Bansod, D. W. (2021). Prevalence and factors associated with triple burden of malnutrition among mother-child pairs in India: a study based on National Family Health Survey 2015–16. *BMC Public Health*, 21(1):1–12.
- Merrell, M. (1931). The relationship of individual growth to average growth. *Human Biology*, 3(1):37.
- Mirzaei, S. and Sengupta, D. (2012). Human growth curve estimation through a combination of longitudinal and cross-sectional data. *World Academy of Science, Engineering and Technology, International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering*, 6:760–765.
- Molinari, L., Largo, R. H., and Prader, A. (1980). Analysis of the growth spurt at age seven (mid-growth spurt). *Helv. Paediatr. Acta*, 35(4):325–334.

- Pizzi, C., Cole, T. J., Richiardi, L., et al. (2014). Prenatal Influences on Size, Velocity and Tempo of Infant Growth: Findings from Three Contemporary Cohorts. *PLOS ONE*, 9(2):e90291.
- Preece, M. A. and Baines, M. J. (1978). A new family of mathematical models describing the human growth curve. *Ann. Hum. Biol.*, 5(1):1–24.
- Prentice, A., Dibba, B., Sawo, Y., et al. (2012). The effect of prepubertal calcium carbonate supplementation on the age of peak height velocity in Gambian adolescents. *American Journal of Clinical Nutrition*, 96(5):1042–1050.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Remer, T. and Manz, F. (2001). The Midgrowth Spurt in Healthy Children Is Not Caused by Adrenarche. *J. Clin. Endocrinol. Metab.*, 86(9):4183–4186.
- Rigby, R. A. and Stasinopoulos, D. M. (2004). Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. *Stat. Med.*, 23(19):3053–3076.
- Rigby, R. A. and Stasinopoulos, D. M. (2005a). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(3):507–554.
- Rigby, R. A. and Stasinopoulos, D. M. (2005b). Generalized additive models for location, scale and shape,(with discussion). *Applied Statistics*, 54:507–554.
- Rigby, R. A. and Stasinopoulos, D. M. (2006). Using the Box-Cox t distribution in GAMLSS to model skewness and kurtosis. *Statistical Modelling*, 6(3):209–229.
- Satyanarayana, K., Radhaiah, G., Mohan, K. R., Thimmayamma, B. V. S., Rao, N., Rao, B. S., and Akella, S. (1989). The adolescent growth spurt of height among rural Indian boys in relation to childhood nutritional background: An 18 year longitudinal study. *Annals of Human Biology*, 16(4):289–300.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *Ann. Stat.*, 6(2):461–464.
- Spencer, P. R., Sanders, K. A., and Judge, D. S. (2018). Growth curves and the international standard: How children’s growth reflects challenging conditions in rural Timor-Leste. *American Journal of Physical Anthropology*, 165(2):286–298.



- Tanner, J. M. (1962). *Growth at adolescence*. 2nd Edition, Blackwell Scientific Publications, Oxford.
- Tanner, J. M. and Cameron, N. (1980). Investigation of the mid-growth spurt in height, weight and limb circumferences in single-year velocity data from the London 1966–67 growth survey. *Ann. Hum. Biol.*, 7(6):565–577.
- Virani, N. (2005). Growth patterns and secular trends over four decades in the dynamics of height growth of Indian boys and girls in Sri Aurobindo Ashram: A cohort study. *Ann. Hum. Biol.*, 32(3):259–282.
- Wade, A. and Kurmanavicius, J. (2009). Creating unbiased cross-sectional covariate-related reference ranges from serial correlated measurements. *Biostatistics*, 10(1):147–154.
- World Health Organization (2006). WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development. *World Health Organization*, pages 1–312.

## Supplementary Information

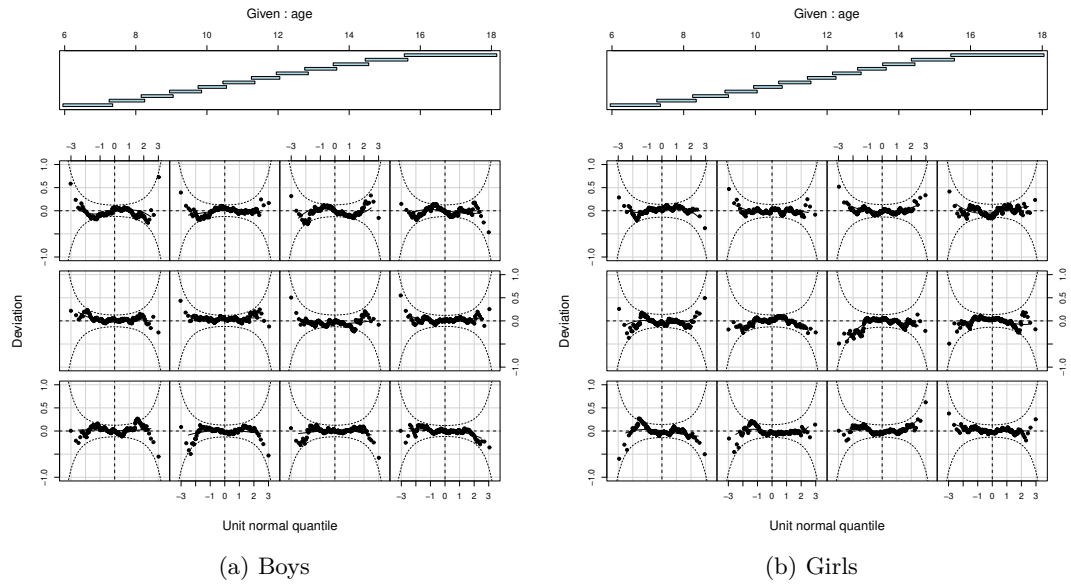


Figure 4.7: Detrended Q-Q plots (worm plots) of GAMLSS residuals in twelve age groups from 6 to 18 years by sex.

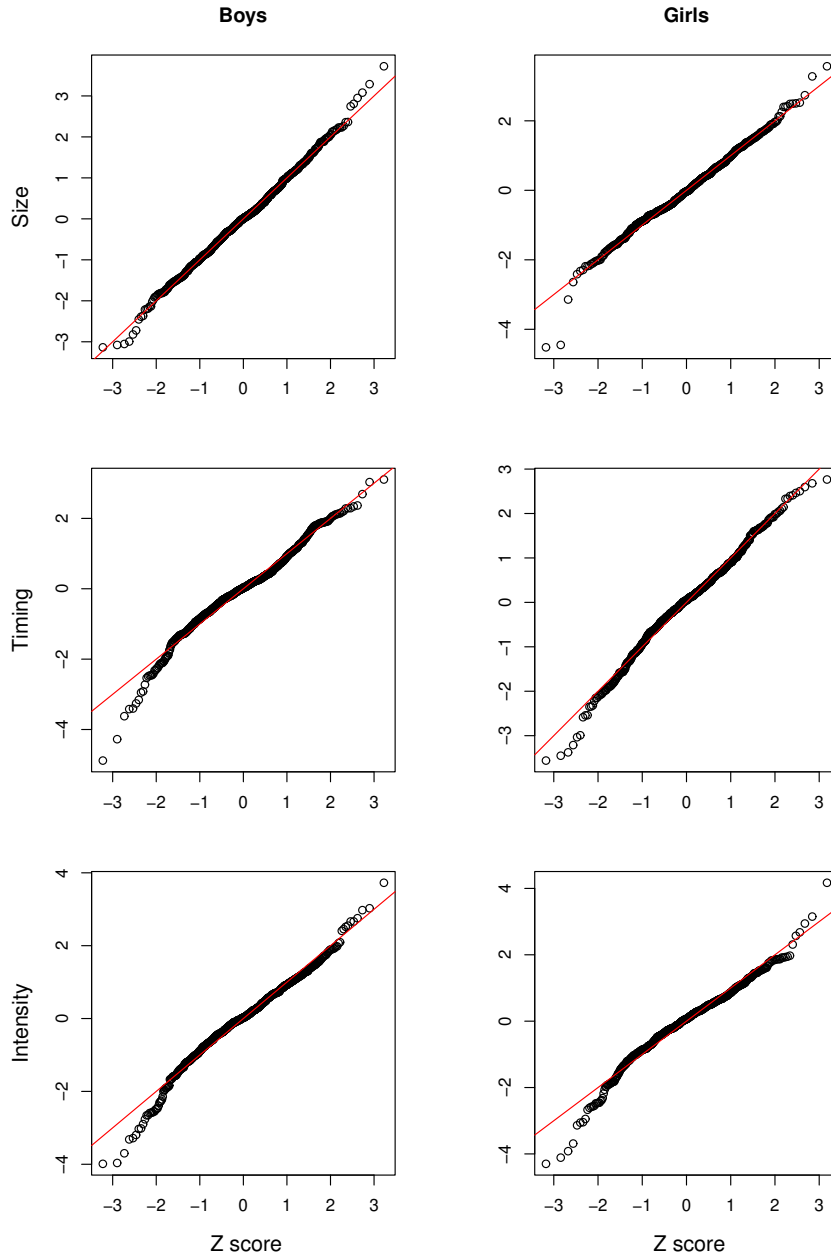


Figure 4.8: Q-Q plots of standardised SITAR random effects size, timing, and intensity by sex. A reference line (red) with unit slope and intercept 0 is given for comparison.

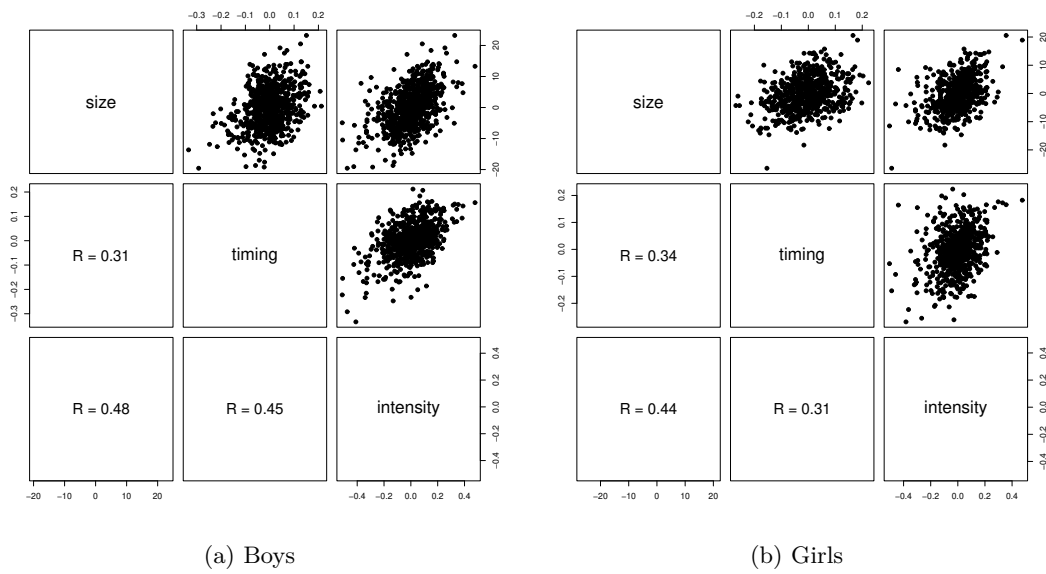


Figure 4.9: Scatter plot matrices showing the correlations between SITAR random effects size, timing and intensity by sex.

## CHAPTER 5

# Height growth under a persisting metabolic insult

*Published as and adapted from:*

**Sandra Aravind Areekal**, Anuradha Khadilkar, Pranay Goel, Tim J. Cole, “Longitudinal Height Growth in Children and Adolescents with Type-1 Diabetes Mellitus Compared to Controls in Pune, India”, *Pediatric Diabetes*, vol. 2023, Article ID 8813031, 8 pages, 2023. <https://doi.org/10.1155/2023/8813031>

## 5.1 Introduction

Type-1 diabetes mellitus (T1DM) is a chronic disorder characterised by a deficiency in insulin production. It is the major form of diabetes diagnosed in children and is commonly referred to as “childhood-onset” diabetes. The worldwide prevalence of T1DM in children under the age of 20 years in 2021 was estimated to be 1.2 million cases (149,500 incident cases), with the highest prevalence in India (Ogle et al. 2022).

Growth has previously been reported to be impaired in children with diabetes (Brown et al. 1994; Bognetti et al. 1998; Ahmed et al. 1998). However, recent reports suggest that with improved diabetes control, height growth can be within the normal range (Bizzarri et al. 2018; Santi et al. 2019; Koren 2022). Previous studies have used height-for-age z-score (HAZ) to study height growth in children with diabetes (Brown et al. 1994; Ahmed et al. 1998; Khadilkar et al. 2013; Vollbach et al. 2021). HAZ is appropriate for analysing height data treated cross-sectionally, where it adjusts for age and ensures that mean HAZ is relatively constant across age. However, for longitudinal data in individuals, it performs less well. The age adjustment fails to cater for individual differences in the timing of the pubertal growth spurt, so that for early maturers, HAZ rises with age and then falls again, while for late maturers, it falls then rises. These individual age trends in HAZ are both complex and hard to interpret.

A better approach with longitudinal data is to model untransformed height using, for example, the non-linear mixed effects model SuperImposition by Rotation and Translation (SITAR) (Cole et al. 2010). SITAR estimates i) differential trends in height growth in two groups, e.g. in children with and without diabetes, as mean differences in size, timing and intensity (Gault et al. 2011; Prentice et al. 2012), and ii) individual-level differences in the pattern of growth as subject-specific random effects for size, timing and intensity (Johnson et al. 2014; Pizzi et al. 2014; Filteau et al. 2019). These methods have not previously been applied to growth in children with diabetes.

Metabolic control is known to be affected by both behavioural and physiological changes during adolescence. Previous studies have not distinguished between behavioural and physiological factors influencing growth in children and adolescents with T1DM. Using SITAR, individual-level growth differences can be related to information relevant to each child’s condition, such as their degree of metabolic control and diabetes duration. By characterising individual variation in

size, timing and intensity of pubertal growth, SITAR provides a way to distinguish between physiological and behavioural influences on metabolically stressed growth.

Here we use the SITAR model to characterise height growth in children with diabetes compared to a local control population and to explore how i) mean height growth is affected by T1DM; and ii) individual growth during puberty is related to disease severity.

## 5.2 Methods and Subjects

### 5.2.1 Data sets

The data came from two studies in Western India, namely the Sweetlings Type-1 Diabetes Mellitus study (STDM) and the Pune School-Children Growth study (PSCG). The PSCG study here acts as a control group for the diabetes study. This study adhered to the STROBE guidelines for reporting observational studies von Schelling (1954).

#### 5.2.1.1 Sweetlings Type-1 Diabetes Mellitus (STDM)

The STDM study involved 490 children (222 boys and 268 girls) aged 1 to 26 years diagnosed with T1DM and visiting a tertiary healthcare centre in Pune, India. Each subject was seen between one and six times (median 3, interquartile range 2 to 5) between 2013 and 2021.

Height and glycated haemoglobin (HbA1c) were recorded on each occasion. Height was measured using a Seca stadiometer (Hamburg, Germany) and calibrated with standard rods, while HbA1c was measured using high-performance liquid chromatography (HPLC, BIO-RAD, Germany). Age at diabetes diagnosis was obtained from clinic records, and the mean duration of diabetes (i.e. mean age minus age at diagnosis) was calculated. Parental heights were measured at the time of the child's diagnosis, while birth weight was obtained from the birth card (where available) or parental report. Mid-parental height z-score was calculated as the mean of internally-calculated height z-scores for mother and father. The data have been analysed previously (Parthasarathy et al. 2016; Oza et al. 2022).

Written informed consent was obtained from parents, and verbal assent was obtained from children above the age of 7 years. The Ethics Committee of the Jehangir Clinical Development Centre Pune approved the study (dated March 22, 2013).

### 5.2.1.2 Pune School-Children Growth (PSCG)

The control group from the PSCG study consisted of 1472 children (798 boys and 674 girls) aged 3 to 18 years recruited from three randomly selected schools catering to middle class children from Pune, as used in a previous study (Khadilkar et al. 2019). Each child was seen annually, between one and six times (median 6, interquartile range 5 to 6) between 2007 and 2013.

Height was measured using a portable stadiometer (Leicester Height Meter; Child Growth Foundation, London) and calibrated with standard rods. Growth curves and growth centiles based on the data have been published previously (Areekal et al. 2022).

Written informed consent was obtained from parents, and verbal assent was obtained from children above the age of 7 years. The Ethics Committee of the Jehangir Clinical Development Centre Pune approved the data collection (dated June 26, 2007). In addition the Ethics Committee of the Indian Institute of Science Education and Research Pune approved secondary data analysis of the STDM and PSCG data sets (IECHR/Admin/2021/001).

Note that the STDM and PSCG data sets differed in their design; STDM had median three measurements per child, covering a period of two years, whereas PSCG had median six measurements per child, covering five years. Thus PSCG had more longitudinal information per child, leading to a more precise mean growth curve. PSCG children were recruited from private schools, which could be a potential source of bias. However, since these schools take middle class children, PSCG was a valid control for STDM.

## 5.2.2 Methods

### 5.2.2.1 SITAR model for height growth curves

Height growth curves were fitted using SuperImposition by Translation and Rotation (SITAR) (Cole et al. 2010). SITAR is a mixed effects growth curve model that estimates the mean height curve as a natural cubic B-spline with degrees of freedom ( $df$ ) chosen to optimise the fit. Subject deviations from the mean curve are captured in three subject-specific random effects: (i) *Size*, which distinguishes the final height attained, (ii) *Timing*, which captures the timing of puberty and (iii) *Intensity*, which provides information on the rate of pubertal growth.

The three random effects are assumed normally distributed and the residual error term is



considered to be independently and identically distributed (i.i.d.). The fitted mean curve is also called the height distance curve, and its first derivative is the mean height velocity curve.

### 5.2.2.2 Diabetes and control height growth curves

The mean curves for the diabetes and control children were estimated in two different ways: first by fitting a single SITAR model to the two data sets pooled, including fixed effects for size, timing and intensity to distinguish between the data sets; and second by fitting separate SITAR models to each data set. The pooled model constrains the two mean curves to be the same shape (adjusted for the fixed effect differences) whereas the separate model allows the mean curves to differ; this allows the equality of the two mean curves to be tested for.

Both ways, variants of the SITAR model were explored by considering spline curve  $df$  from 4 to 8. Models with log-transformed age and/or height scales and combinations of fixed effects (size and timing, size and intensity, and size, timing and intensity) were also considered. The optimal model was the one minimising the Bayesian Information Criterion (BIC) (Schwarz 1978). Bootstrap confidence intervals (CI) for the mean age at peak height velocity (APV in years) and peak height velocity (PV in cm/year) were obtained using 500 re-samples. The percentage of variance explained by the model was calculated as described in Patcas et al. (2022).

The optimal models fitted height versus log age, where the random and fixed effects for timing can be multiplied by 100 and viewed as percentage differences (Cole and Altman 2017). Alternatively they can be multiplied by mean APV to express them in units of months or years.

The data were analysed using the package *sitar* (version 1.2.0.9000) (Cole 2022) in the statistical language R (version 4.2.2) (R Core Team 2021). Prior data cleaning removed obviously errant points based on plots of height versus age (STDM: 10 points in boys and 9 in girls; PSCG: 11 points in boys and 3 in girls). SITAR models were then fitted, and standardised residuals exceeding 4 in absolute value were excluded (STDM: 3 in boys; PSCG: 12 in girls). 95% CI bands were obtained for the average SITAR curves using 500 bootstrap re-samples. The mean and standard deviation by age of the bootstrapped curves were summarised as smooth curves by fitting the normal distribution family in GAMLSS (version 5.4.10) (Rigby and Stasinopoulos 2005).

### 5.2.2.3 Modelling serial HbA1c measurements by age

Longitudinal trends in HbA1c were also modelled using SITAR, with the sexes pooled and a fixed effect included to distinguish between them. The random effects for timing and intensity did not improve the fit, so the model included just the random effect for size, i.e. a random intercept model.

### 5.2.2.4 Relating SITAR random effects to subject-specific covariates

The optimal SITAR height model was extended to explore how the three subject-specific random effects related to the following physiological covariates: mid-parental height z-score, birth weight, age at diabetes diagnosis, diabetes duration, and mean HbA1c (size random effect from section 5.2.2.3). For this, the SITAR height model was extended to predict each of the three random effects as linear functions of the covariates, included as fixed effects. Subjects with no HbA1c measurements were assumed to have mean HbA1c (i.e. random intercept 0). The significance level for the analysis was set at  $\alpha = 0.01$ .

## 5.3 Results

The data sets were cleaned and analysed separately by sex. The age range for the study was restricted to 4 to 19 years since the height growth pattern in infancy is distinct from that in childhood and adolescence. The final analysis included: in STDm, 460 subjects (208/252 boys/girls) with 1598 height measurements (732/866 in boys/girls); and in PSCG, 1470 subjects (797/673 boys/girls) with 8140 height measurements (4455/3685 in boys/girls).

### 5.3.1 Average growth curves

#### 5.3.1.1 Pooled models of diabetes and control growth curves

Figure 5.1 shows the SITAR mean distance and velocity curves (95% CI) for STDm and PSCG estimated from the pooled data with 6 and 5 *df* in boys and girls respectively (variance explained 99.0% and 99.1%). Table 5.1 shows the mean differences between STDm and PSCG in size, timing, and intensity. The children with diabetes were shorter than the control children, by 4.9 cm in boys and 3.8 cm in girls. Their timing of pubertal growth was also later, by 1.5 months in boys and 6.1 months in girls. Further, their intensity of pubertal growth was lower by 9.8% in

boys and 4.8% in girls.

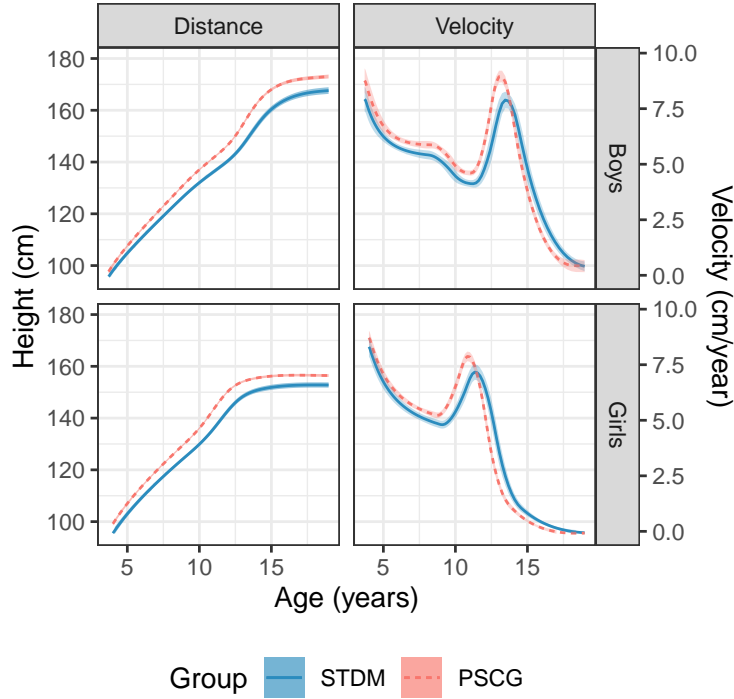


Figure 5.1: SITAR mean height growth curves and height velocity curves by sex in the diabetes (STDM) and control (PSCG) children with 95% CI bands.

Table 5.1: Differences in the fixed effects for mean size, timing and intensity in STDM compared to PSCG, with standard errors (SE) and p-values (P).

	Boys			Girls		
	Difference	SE	P	Difference	SE	P
<i>Size (cm)</i>	-4.9	0.6	0.001	-3.8	0.6	0.001
<i>Timing (%)</i>	1.0	0.9	0.3	4.6	0.9	0.001
<i>Timing (months)</i>	1.5	1.5	0.3	6.1	1.1	0.001
<i>Intensity (%)</i>	-9.8	1.7	0.001	-4.8	1.7	0.006

### 5.3.1.2 Separate models of diabetes and control growth curves

SITAR height models were also fitted to each data set separately. The optimal mean STDM curves had 6 and 5 *df* in boys and girls respectively (variance explained 99.4% and 99.5%), while the mean PSCG curves had 6 and 4 *df* in boys and girls (variance explained 98.8% in both sexes). Mean APV and PV in the two groups are given in Table 5.2. Note that the values for PSCG differ slightly from before (Areekal et al. 2022) due to the differing age ranges. The BIC of the pooled model was 29 units smaller than the sum of BIC for the two separate models in boys, but

20 units greater in girls. Thus, based on BIC, the pooled model fitted better in boys but the separate model was better in girls.

The mean STDM distance and velocity curves estimated the two ways (with 95% CI bands for the separate models) are compared in Figure 5.2, with the pooled and separate models shown as solid blue and dotted grey curves respectively. The two sets of curves are very similar in shape, showing that pooling with the control data did not materially affect the diabetes curves.

Table 5.2: Mean (95% CI) age at peak height velocity (APV) and peak velocity (PV) estimated separately for STDM and PSCG by sex.

	Boys		Girls	
	APV (years)	PV (cm/year)	APV (years)	PV (cm/year)
<i>STDM</i>	13.8 (13.4, 14.1)	8.1 (7.6, 8.7)	11.2 (10.7, 11.6)	6.8 (6.4, 7.2)
<i>PSCG</i>	13.1 (12.9, 13.2)	8.9 (8.7, 9.2)	10.9 (10.8, 11.0)	7.9 (7.8, 8.1)

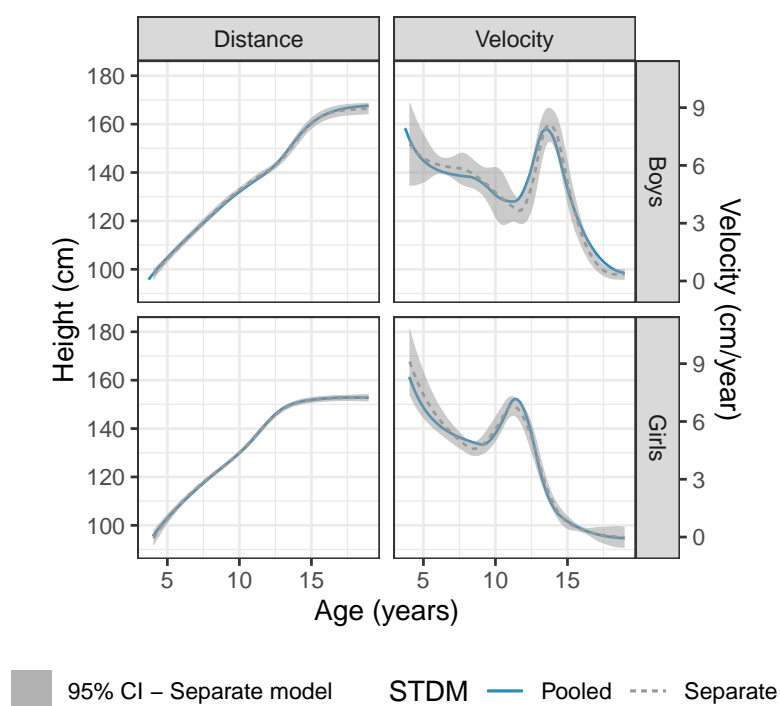


Figure 5.2: SITAR mean height growth curves and height velocity curves by sex with 95% CI bands for STDM pooled (solid blue) and separate (dotted grey).

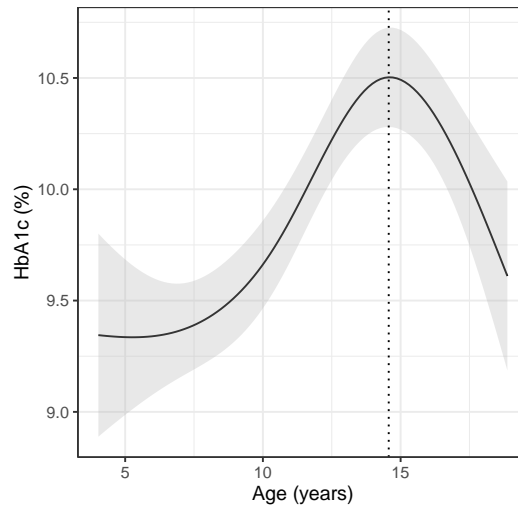


Figure 5.3: The mean HbA1c growth curve (and 95% CI band) with a peak at 10.5% at age 14.6 years.

### 5.3.2 Modelling longitudinal HbA1c measurements

The subject-specific HbA1c curves were summarised using SITAR with a random intercept and cubic B-spline with 3 *df* (variance explained 43%). The SD of the random intercept (mean individual HbA1c) was 0.14% and the residual SD was 0.15%. Figure 5.3 shows the mean curve and 95% CI band, with HbA1c rising to 10.5% at 14.6 years and then falling again. The sexes did not differ (mean difference -0.01, 95% CI -0.04 to 0.02).

### 5.3.3 Individual variation in height growth

We examined associations among the STDM children of the SITAR random effects size, timing and intensity with mid-parental height, birth weight, age at diabetes diagnosis, diabetes duration, and mean HbA1c (the random intercept from section 5.3.2). Summary statistics for the covariates are in Table 5.3, while the regression coefficients of the random effects on the covariates are in Table 5.4, where the models include all the covariates so they are mutually adjusted. Note that since birth weight proved to be unrelated to any of the random effects, it was omitted from the models.

Size was highly significantly positively associated with mid-parental height, age at diagnosis and diabetes duration in both sexes. Boys and girls with a 1 SD greater mid-parental height were respectively 3.4 and 3.5 cm taller. Among children who had had diabetes for the same length of

Table 5.3: Summary statistics of covariates in the STDM data set. N is the number of subjects with measurements and SD is the sample standard deviation.

Variable	Boys			Girls		
	N	Mean	SD	N	Mean	SD
Maternal height (cm)	184	153.8	5.2	234	153.7	5.8
Paternal height (cm)	185	166.5	6.4	206	166.8	6.7
Mid-parental height (z-score)	185	0.0	1.0	235	0.0	1.0
Birth weight (kg)	208	2.7	0.8	252	2.7	0.6
Age at Diagnosis (years)	208	7.5	4.1	249	7.5	3.9
Diabetes duration (years)	208	4.9	3.4	249	4.5	3.0

Table 5.4: Regression coefficients  $\beta$  (SE) of size, timing, and intensity on the covariates. N denotes the number of subjects included in the analysis.

	Boys ( $N = 182$ )		Girls ( $N = 226$ )	
	$\beta$ (SE)	P	$\beta$ (SE)	P
<b>Size (cm)</b>				
Mid-parental height (z-score)	3.4 (0.7)	<0.001	3.5 (0.5)	<0.001
Age at diabetes diagnosis (years)	1.6 (0.2)	<0.001	1.2 (0.2)	<0.001
Diabetes duration (years)	0.8 (0.2)	0.001	0.9 (0.2)	<0.001
Mean HbA1c (%)	-3.2 (6)	0.9	-15 (5)	0.001
<b>Timing (months)</b>				
Mid-parental height (z-score)	3.1 (1.4)	0.03	0.9 (1.0)	0.4
Age at diabetes diagnosis (years)	5.2 (0.5)	<0.001	4.2 (0.4)	<0.001
Diabetes duration (years)	5.4 (0.6)	<0.001	4.6 (0.4)	<0.001
Mean HbA1c (%)	11 (11)	0.1	-20 (9)	0.03
<b>Intensity (%)</b>				
Mid-parental height (z-score)	6.0 (2.8)	0.03	3.9 (2.3)	0.1
Age at diabetes diagnosis (years)	2.4 (0.9)	0.01	1.8 (0.9)	0.05
Diabetes duration (years)	1.7 (1.0)	0.1	1.9 (1.1)	0.09
Mean HbA1c (%)	3 (23)	0.9	-34 (24)	0.2

time, boys/girls diagnosed a year later were 1.6/1.2 cm taller. Among those diagnosed at the same age, those who had had diabetes one year longer were 0.8/0.9 cm taller. In girls, though not in boys, mean HbA1c was strongly negatively associated with mean height; a 1% higher mean HbA1c (SD 0.14%) corresponded to being 15 cm shorter, or 2 cm shorter for a 1 SD higher HbA1c.

Timing was highly significantly positively associated with both age at diagnosis and duration, with similar effects in the two sexes. Adjusted for duration, diagnosis one year later was associated with 5.2/4.2 months delay in age at peak velocity. Similarly, adjusted for age at diagnosis, one extra year of diabetes was associated with 5.4/4.6 months delay. Adjusted for diabetes duration,

being diagnosed one year later corresponded to 2.4% faster pubertal growth in boys, and rather less in girls.

All three HbA1c coefficients were larger in girls than boys, and significantly so for timing ( $P < 0.03$ ), showing the greater impact of HbA1c on growth in girls. No other coefficients in Table 5.4 differed significantly by sex. Parental heights were also analysed separately; however, the coefficients for paternal and maternal height did not differ significantly for any SITAR parameters. In addition there were no significant interactions between age at diabetes diagnosis and diabetes duration.

## 5.4 Discussion

We studied height growth in two groups of children aged 4 to 19 years from Pune, India; one diagnosed with and undergoing treatment for T1DM, and the other school-children acting as controls, to explore how a metabolic disorder such as T1DM affects height growth. Historically, Indian children are reported to have poor metabolic control (Khadilkar et al. 2013; Chowdhury 2015; Parthasarathy et al. 2016). It is possible that persistent insulin deficiency despite treatment leads to impaired growth. Previous literature on growth in Indian children with T1DM, based on HAZ comparisons, has reported short stature (Khadilkar et al. 2013) and lowered pubertal growth velocity (Parthasarathy et al. 2016). Moreover, a 15.7% prevalence of stunting was also reported in Indian children with T1DM (Bhor et al. 2022). In this work, SITAR was used to model height and estimate an average curve for each group. Children with diabetes were comparatively shorter at all ages, and their pubertal growth was both delayed and extended compared to the control group.

SITAR assumes that height growth in individuals can be summarised by their attained height (size) and the timing and intensity of their pubertal growth spurt; hence after adjusting for these the shape of the average growth curve ought to be the same in the two groups. Growth curves in diabetes were first estimated by pooling the data to have the same underlying shape as the control group curves. Next, the two sets of curves were estimated separately; this allowed the pooled and separate curves to be compared. For boys the pooled curves fitted slightly better; the opposite was true for girls. In practical terms the shapes of the diabetes curves were indistinguishable from those of the control curves after SITAR adjustment – only subtle differences were visible in the height velocity curves (Figure 2). This shows that the mechanism whereby T1DM affects

growth is accurately modelled by SITAR: the underlying growth process is invariant, robustly independent of disease status. The effect of diabetes is to slow and delay growth, which leads to reduced height overall.

A SITAR model fitted to HbA1c with age shows that it rises through childhood, peaks in puberty and then falls again (Figure 3). The pubertal peak in HbA1c is likely due not only to physiological changes, such as increased insulin resistance during puberty (Moran et al. 1999), but also to behavioural resistance to lifestyle change (Elbalshy et al. 2022).

We also analysed how individual differences in pubertal height growth relate to child-specific physiological characteristics. We found, unsurprisingly, that children with taller parents were taller. The course of disease, in particular the age at diagnosis and the time since diagnosis, affected mean height and the timing of the growth spurt – later diagnosis and longer duration were independently associated with greater height and later puberty. However there was an important sex difference: girls were affected more in terms of size, particularly with mean HbA1c, where a 1 SD increase in HbA1c was associated with being 2 cm shorter. Boys were affected more in terms of timing and intensity, which was positively associated with age at diagnosis. Thus, broadly speaking, the impact of T1DM was to reduce height in girls, whereas in boys it slowed and delayed growth. The dependence of pubertal growth on growth hormone and testosterone in girls and boys, respectively, may be important for this observed difference as suggested by Dunger et al. (2002).

SITAR has previously been used in an experimental paradigm in life course epidemiology, with growth as the exposure being related to a later life outcome (Gault et al. 2011; Prentice et al. 2012). Here we shift to a different paradigm where diabetes is the early exposure and growth in puberty is the outcome. We observe that children with diabetes grow in a different way from control children, and it is useful to put this observation into a life history framework, which deals with how individuals organise themselves in the context of scarce resources to optimise their life experiences. The transitions between different stages of growth, such as from childhood to adolescence, have been described as a period of “adaptive plasticity”, which denotes the trade-offs adopted by an organism under adverse conditions, by Hochberg and Albertsson-Wikland (Hochberg and Albertsson-Wikland 2008). An organism under stress, either external (environmental) or internal (physiological), has to make the choice whether to push for fecundity, that is, to have more children, or longevity, that is, to live longer. We see that the effect of diabetes is to make the children progressively shorter through childhood; and simultaneously to



delay the pubertal growth spurt and extend it. In this way the child with diabetes responds to the diabetic insult by investing fewer of its resources in growth.

A limitation of our study is the difference in design of the two studies, with median three measurements per child in STDM compared to six measurements in PSCG. So the individual STDM growth curves were less precisely specified than those in PSCG. However, a strength of our study is that it contrasts data on a substantial number of children with diabetes followed over time with a group of control children from broadly the same environment.

Finally, we remark that the COVID-19 pandemic signalled an urgent need for early and timely diagnosis of T1DM, flagged by studies showing a higher incidence of T1DM in children and adolescents (Guo et al. 2022) and a higher prevalence of complications during diagnosis and management such as diabetic ketoacidosis (Birkebaek et al. 2022; Shah et al. 2022). Viral infections are associated with an increased risk for T1DM in genetically susceptible children (Shi et al. 2022). These studies highlight the need for longitudinal cohorts to examine how the pandemic has affected growth in children with T1DM. As we have shown, SITAR is an excellent tool to study how diabetes affects growth.

## Bibliography

- Ahmed, M. L., Connors, M. H., Drayer, N. M., Jones, J. S., and Dunger, D. B. (1998). Pubertal Growth in ID DM Is Determined by HbA1c Levels, Sex, and Bone Age. *Diabetes Care*, 21(5):831–835.
- Areekal, S. A., Goel, P., Khadilkar, A., Khadilkar, V., and Cole, T. J. (2022). Assessment of height growth in indian children using growth centiles and growth curves. *Annals of Human Biology*, 49(5-6):228–235. PMID: 36112429.
- Bhor, S., Oza, C., Khadilkar, A., Ladkat, D., Gondhalekar, K., and Khadilkar, V. (2022). Prevalence of stunting and determinants of growth failure in children with type 1 diabetes. *Indian Journal of Child Health*, 9(9):170–174.
- Birkebaek, N. H., Kamrath, C., Grimsmann, J. M., Aakesson, K., Cherubini, V., Dovc, K., de Beaufort, C., Alonso, G. T., Gregory, J. W., White, M., et al. (2022). Impact of the covid-19 pandemic on long-term trends in the prevalence of diabetic ketoacidosis at diagnosis of

- paediatric type 1 diabetes: an international multicentre study based on data from 13 national diabetes registries. *The Lancet Diabetes & Endocrinology*.
- Bizzarri, C., Timpanaro, T. A., Matteoli, M. C., Patera, I. P., Cappa, M., and Cianfarani, S. (2018). Growth trajectory in children with type 1 diabetes mellitus: the impact of insulin treatment and metabolic control. *Hormone Research in Paediatrics*, 89(3):172–177.
- Bognetti, E., Riva, M. C., Bonfanti, R., Meschi, F., Viscardi, M., and Chiumello, G. (1998). Growth changes in children and adolescents with short-term diabetes. *Diabetes care*, 21(8):1226–1229.
- Brown, M., Ahmed, M. L., Clayton, K. L., and Dunger, D. B. (1994). Growth During Childhood and Final Height in Type 1 Diabetes. *Diabet. Med.*, 11(2):182–187.
- Chowdhury, S. (2015). Puberty and type 1 diabetes. *Indian Journal of Endocrinology and Metabolism*, 19(Suppl 1):S51.
- Cole, T. (2022). *sitar: Super Imposition by Translation and Rotation Growth Curve Analysis*. R package version 1.2.0.9000.
- Cole, T. J. and Altman, D. G. (2017). Statistics notes: Percentage differences, symmetry, and natural logarithms. *BMJ*, 358:j3683.
- Cole, T. J., Donaldson, M. D. C., and Ben-Shlomo, Y. (2010). SITAR—a useful instrument for growth curve analysis. *Int. J. Epidemiol.*, 39(6):1558.
- Dunger, D., Ahmed, L., and Ong, K. (2002). Growth and body composition in type 1 diabetes mellitus. *Horm. Res.*, 58(Suppl):1.
- Elbalshy, M., Haszard, J., Smith, H., Kuroko, S., Galland, B., Oliver, N., Shah, V., de Bock, M. I., and Wheeler, B. J. (2022). Effect of divergent continuous glucose monitoring technologies on glycaemic control in type 1 diabetes mellitus: A systematic review and meta-analysis of randomised controlled trials. *Diabetic Medicine*, page e14854.
- Filteau, S., Kumar, G. T., Cole, T. J., et al. (2019). Steady Growth in Early Infancy Is Associated with Greater Anthropometry in Indian Children Born Low Birth Weight at Term. *Journal of Nutrition*, 149(9):1633–1641.
- Gault, E. J., Perry, R. J., Cole, T. J., et al. (2011). Effect of oxandrolone and timing of pubertal

- induction on final height in Turner’s syndrome: randomised, double blind, placebo controlled trial. *BMJ*, 342.
- Guo, Y., Bian, J., Chen, A., Wang, F., Posgai, A. L., Schatz, D. A., Shenkman, E. A., and Atkinson, M. A. (2022). Incidence trends of new-onset diabetes in children and adolescents before and during the covid-19 pandemic: Findings from florida. *Diabetes*.
- Hochberg, Z. and Albertsson-Wikland, K. (2008). Evo-devo of infantile and childhood growth. *Pediatric research*, 64(1):2–7.
- Johnson, L., van Jaarsveld, C. H. M., Llewellyn, C. H., et al. (2014). Associations between infant feeding and the size, tempo and velocity of infant weight gain: SITAR analysis of the Gemini twin birth cohort. *International Journal of Obesity*, 38:980–987.
- Khadilkar, V., Khadilkar, A., Arya, A., Ekbote, V., Kajale, N., Parthasarathy, L., Patwardhan, V., Phanse, S., and Chiplonkar, S. (2019). Height Velocity Percentiles in Indian Children Aged 5-17 Years. *Indian Pediatr.*, 56(1):23–28.
- Khadilkar, V. V., Parthasarathy, L. S., Mallade, B. B., Khadilkar, A. V., Chiplonkar, S. A., and Borade, A. B. (2013). Growth status of children and adolescents with type 1 diabetes mellitus. *Indian J. Endocrinol. Metab.*, 17(6):1057–1060.
- Koren, D. (2022). Growth and development in type 1 diabetes. *Curr. Opin. Endocrinol. Diabetes Obes.*, 29(1):57–64.
- Moran, A., Jacobs Jr, D. R., Steinberger, J., Hong, C.-P., Prineas, R., Luepker, R., and Sinaiko, A. R. (1999). Insulin resistance during puberty: results from clamp studies in 357 children. *Diabetes*, 48(10):2039–2044.
- Ogle, G. D., James, S., Dabelea, D., Pihoker, C., Svensson, J., Maniam, J., Klatman, E. L., and Patterson, C. C. (2022). Global estimates of incidence of type 1 diabetes in children and adolescents: Results from the International Diabetes Federation Atlas, 10th edition. *Diabetes Res. Clin. Pract.*, 183:109083.
- Oza, C., Khadilkar, V., Karguppikar, M., Ladkat, D., Gondhalekar, K., Shah, N., and Khadilkar, A. (2022). Prevalence of metabolic syndrome and predictors of metabolic risk in Indian children, adolescents and youth with type 1 diabetes mellitus. *Endocrine*, 75(3):794–803.

- Parthasarathy, L., Khadilkar, V., Chiplonkar, S., and Khadilkar, A. (2016). Longitudinal Growth in Children and Adolescents with Type 1 Diabetes. *Indian Pediatrics*, 53(11):990–992.
- Patcas, R., Keller, H., Markic, G., Beit, P., Eliades, T., and Cole, T. J. (2022). Craniofacial growth and sitar growth curve analysis. *European Journal of Orthodontics*, 44(3):325–331.
- Pizzi, C., Cole, T. J., Richiardi, L., et al. (2014). Prenatal Influences on Size, Velocity and Tempo of Infant Growth: Findings from Three Contemporary Cohorts. *PLOS ONE*, 9(2):e90291.
- Prentice, A., Dibba, B., Sawo, Y., et al. (2012). The effect of prepubertal calcium carbonate supplementation on the age of peak height velocity in Gambian adolescents. *American Journal of Clinical Nutrition*, 96(5):1042–1050.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape,(with discussion). *Applied Statistics*, 54:507–554.
- Santi, E., Tascini, G., Toni, G., Berioli, M. G., and Esposito, S. (2019). Linear growth in children and adolescents with type 1 diabetes mellitus. *International Journal of Environmental Research and Public Health*, 16(19):3677.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *Ann. Stat.*, 6(2):461–464.
- Shah, N., Khadilkar, V., Oza, C., Karguppikar, M., Bhor, S., Ladkat, D., and Khadilkar, A. (2022). Impact of decreased physical activity due to covid restrictions on cardio-metabolic risk parameters in indian children and youth with type 1 diabetes. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 16(7):102564.
- Shi, Y., Wu, L.-Q., Wei, P., and Liao, Z.-H. (2022). Children with type 1 diabetes in covid-19 pandemic: Difficulties and solutions. *World Journal of Clinical Pediatrics*, 11(5):408.
- Vollbach, H., Auzanneau, M., Reinehr, T., Wiegand, S., Schwab, K.-O., Oeverink, R., Froehlich-Reiterer, E., Woelfle, J., De Beaufort, C., Kapellen, T., et al. (2021). Choice of basal insulin therapy is associated with weight and height development in type 1 diabetes: A multicenter analysis from the german/austrian dpv registry in 10 338 children and adolescents. *Journal of Diabetes*, 13(11):930–939.

von Schelling, H. (1954). Mathematical deductions from empirical relations between metabolism, surface area and weight. *Ann N Y Acad Sci*, 56:1143-64.

## CHAPTER 6

# Discussion and Conclusion

Metabolism and height growth are two important aspects of child growth. We studied how these two aspects are influenced during a period known to show high inter-individual variability, namely the childhood and adolescence, in the Indian population. In this thesis, we describe the curvilinear trend in metabolism and height growth in the Indian children and show that Indian children have significantly lower resting energy expenditure compared to their western counterparts. We further discuss the plausible explanations for this lowered metabolism. We also show that the underlying height growth pattern is invariant, but the children differ in the time and rate of pubertal growth and the final height. Moreover, the underlying growth pattern is not influenced in children diagnosed with Type-1 diabetes, once adjusted for the size (final height), timing and intensity of pubertal growth.

### **6.1 Sexual differences in metabolism and height growth during childhood and adolescence.**

Sexual differences in physiological and behavioural traits, apart from the trivial differences in reproductive traits, have been documented at different scales and periods during growth - from molecular to organismal (Mank and Rideout 2021) and from perinatal to adolescence (Thurstans et al. 2022). Understanding these sex difference helps to develop well targeted therapeutics and improve personalised interventions (Miguel-Aliaga 2022). Here we discuss the sex differences observed, or the lack there of, in metabolism and height growth in Indian children and adolescents.

It is known that the absolute RMR values are different between the two sexes. Previous studies have shown that the difference in absolute values of RMR between the two sexes were explained as due to their difference in lean body mass (Buchholz et al. 2001). In chapter one, we find that the RMR/BM in Indian children are lower than their Western counterparts. However, the observed differences are greater in girls at all ages compared to the boys. In the two models that have been proposed to explain the lowered RMR/BM in Indian children, we predicted differential adjustments in boys and girls: The first model predicts that the organ masses are lower by 23% in girls but only by 10% in boys; Similarly, in the second model, the residual mass -including fat mass and muscle mass - has been lowered by 35% in girls and by 10% in boys. Further, the RMR/BM reduces drastically after puberty (11 years) in girls. This needs to be considered while assessing the risk of developing overweight or obese phenotypes in adolescents females, and especially in the context of increasing incidence of metabolic disorders such as the polycystic ovarian syndrome (Bharali et al. 2022) and other non-communicable diseases.

Sexual differences in the shape of the height growth curves has been well established in previous studies (Tanner and Cameron 1980; Tanner and Whitehouse 1976), which was evident in our study as well. The pubertal growth spurt starts earlier in girls around the age of 9 years and peaks at 11 years (height velocity at 8.0 cm/year) and then decelerates and stops by the age of 18 years (height velocity = 0). In contrast, boys show a mid-growth spurt at the age of 8.6 years which decelerates until around 11 years when the pubertal growth spurt takes off. Boys peaks in height velocity is at 13.1 years (with peak height velocity = 9 cm/year), decelerates then on but continues to grow even at the age of 18 years (height velocity > 0).

Further, the boys' height distribution is more skewed after puberty compared to girls (Figure 4.1 ). This is also evident in the spacing between the centiles at the age 18 years: 0.4<sup>th</sup> to 10<sup>th</sup> centiles are wider compared to the 90<sup>th</sup> to 99.6<sup>th</sup> centiles in boys; however, the centiles evenly distributed in girls. This suggests that there is a subset of boys who's height growth have been compromised more than other. This is in line with the recent reports of differential effects in height growth in boys, compared to girls (Thurstans et al. 2022).

In chapter 3, we find a differential response to T1DM in height growth which affects the final height and timing of pubertal growth in girls, but to delay the timing and reduce the intensity of pubertal growth in boys.

Lack of any sexual differences is observed in the HbA1c growth curves. A non-linear trend in

HbA1c during adolescence has previously been reported by Clements et al. (2016) who analysed age trends in HbA1c data from a T1DM exchange directory. We observed a similar trend wherein the HbA1c growth curve rises after puberty (at 15 years in both sexes) and falls thereof. This trend is found to be independent of sex. The rise in HbA1c during puberty is attributed to both behavioural and physiological changes during puberty (Moran et al. 1999; Elbalshy et al. 2022).

Most importantly, poor metabolic control is observed to reduce final height in girls but not in boys. Our study shows that the metabolic control, measured as mean HbA1c values, is significantly associated to the size (or final height) in girls, but we failed to observe any significant relationship with any of the growth parameters in the boys. Instead, the boys were more affected in the timing and rate of pubertal growth. We also looked at the effect of individual variability in the HbA1c measurements with the height growth parameters; however, we failed to observe any significant associations.

The role of insulin as a growth hormone and the differential dependence of male and female pubertal growth on insulin vs testosterone might be of relevance in this scenario and needs to be evaluated further.

Adolescence is a period of major sexual differentiation in humans. Hence, the major changes in reproductive traits been studied very well in the past. However, the subtle sex effects in other physiological traits are only being uncovered in the recent past. It is necessary to understand such differences in metabolism and growth so that interventions for metabolic and growth disorders can be personalised well.

## **6.2 Strengths and limitations of the study**

One limitation of our study is that we could provide only correlational arguments and no causal relationship was explored here. Further, there is difference in the study designs of three major data sets analysed here: MCS is a cross-sectional study, and PSCG and STDM are longitudinal. The models developed for RMR in Indian children needs to be validated in a larger sample to provide more conclusive explanation. The difficulties in comparing body composition differences could be resolved through future studies that utilise image analysis of body composition images.

A major strength of our study is that we look at one of the largest cross-sectional data set of resting metabolic rate in children and the largest longitudinal data set of height growth in Indian



population. We are also pioneering a departure from the conventional approach of employing z-scores analysis to investigate the impact of a disease on growth utilizing SITAR. This method is unique in that it considers both individual variations and population-level differences in its key parameters: size, timing, and intensity of pubertal growth in a single framework. Furthermore, compared to other parametric models of growth which utilises a combination of exponential functions to capture the non-linearity of the shape of the growth curves (Preece and Baines 1978; Karlberg 1989), SITAR learns the shape of the growth curve from the data.

### 6.3 Conclusion

The thesis provides fresh perspectives to the study of growth offering promising avenues for further exploration in research on non-communicable diseases and the "developmental origins of health and diseases (DOHAD)" (Barker 2004).

Broadly, in this thesis we study that the variation in two physiological traits, metabolism and height, and show that they vary considerably between different ethnic groups and individuals. Every individual has a unique height growth trajectory, but once the difference in the size, timing and intensity of growth are accounted, the underlying height growth pattern is invariant or same for everyone. The underlying growth pattern of children diagnosed with T1DM are not found to be significantly different from children without diabetes, once the difference in size, timing and intensity is accounted for. It is surprising that the underlying height growth programme is robust to a persistent metabolic insult as in the case of T1DM. The role of insulin in growth as well as the metabolism is of interest here and needs to be better studied in the future.

### Bibliography

- Barker, D. (2004). Developmental origins of adult health and disease. *Journal of Epidemiology and Community Health*, 58(2):114.
- Bharali, M. D., Rajendran, R., Goswami, J., Singal, K., and Rajendran, V. (2022). Prevalence of Polycystic Ovarian Syndrome in India: A Systematic Review and Meta-Analysis. *Cureus*, 14(12):e32351.

- Buchholz, A. C., Rafii, M., and Pencharz, P. B. (2001). Is resting metabolic rate different between men and women? *British Journal of Nutrition*, 86(6):641–646.
- Clements, M. A., Foster, N. C., Maahs, D. M., Schatz, D. A., Olson, B. A., Tsalikian, E., Lee, J. M., Burt-Solorzano, C. M., Tamborlane, W. V., Chen, V., et al. (2016). Hemoglobin a1c (hba1c) changes over time among adolescent and young adult participants in the t1d exchange clinic registry. *Pediatric diabetes*, 17(5):327–336.
- Elbalshy, M., Haszard, J., Smith, H., Kuroko, S., Galland, B., Oliver, N., Shah, V., de Bock, M. I., and Wheeler, B. J. (2022). Effect of divergent continuous glucose monitoring technologies on glycaemic control in type 1 diabetes mellitus: A systematic review and meta-analysis of randomised controlled trials. *Diabetic Medicine*, page e14854.
- Karlberg, J. (1989). A biologically-oriented mathematical model (ICP) for human growth. *Acta Paediatr. Scand. Suppl.*, 350(70-94.);.
- Mank, J. E. and Rideout, E. J. (2021). Developmental mechanisms of sex differences: from cells to organisms. *Development*, 148(19).
- Miguel-Aliaga, I. (2022). Let’s talk about (biological) sex. *Nat. Rev. Mol. Cell Biol.*, 23(4):227–228.
- Moran, A., Jacobs Jr, D. R., Steinberger, J., Hong, C.-P., Prineas, R., Luepker, R., and Sinaiko, A. R. (1999). Insulin resistance during puberty: results from clamp studies in 357 children. *Diabetes*, 48(10):2039–2044.
- Preece, M. A. and Baines, M. J. (1978). A new family of mathematical models describing the human growth curve. *Ann. Hum. Biol.*, 5(1):1–24.
- Tanner, J. M. and Cameron, N. (1980). Investigation of the mid-growth spurt in height, weight and limb circumferences in single-year velocity data from the London 1966–67 growth survey. *Ann. Hum. Biol.*, 7(6):565–577.
- Tanner, J. M. and Whitehouse, R. H. (1976). Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. *Archives of Disease in Childhood*, 51(3):170–179.
- Thurstans, S., Opondo, C., Seal, A., Wells, J. C., Khara, T., Dolan, C., Briend, A., Myatt, M.,

Garenne, M., Mertens, A., Sear, R., and Kerac, M. (2022). Understanding sex differences in childhood undernutrition: A narrative review. *Nutrients*, 14(5):948.

# Appendix A

## Model optimisation

### A.1 Introduction

In this appendix section, we discuss the details of the model selection procedures used in Chapter 3, Chapter 4, and Chapter 5 as below.

1. Modelling resting metabolic rate in Indian children and adolescents
2. Modelling height growth in Indian children
3. Modelling height growth under a persisting metabolic insult.

Each appendix section briefly discusses the data set used in the study, data cleaning, model optimisation and model selection procedures.

### A.2 Modelling resting metabolic rate in Indian children and adolescents

Resting metabolic rate (RMR) a standard measure of energy expenditure at rest and allows comparison of metabolism between individuals as well as populations. Over the last century, many phenomenological models have been developed to predict RMR from anthropometry and body composition. However, they have been able to explain only 60-80% of the observed variations in RMR in adults. RMR studies in Indian children are sparse. Here we discuss the inference procedures involved in the earlier chapter titled “Modelling resting metabolic rate in Indian children”.

This chapter initially discusses the data sets analysed in the study, different linear regression models for RMR based on anthropometry and body composition and later discusses development of two novel models for RMR in Indian children by modifying the Wang model for RMR/BM in North American children.

## A.2.1 Datasets

### A.2.1.1 Multi-centre study (MCS) data set on RMR in Indian children.

The MCS data set contains anthropometric and body composition measurements of 495 children from multiple centres in India. The variables that are analysed in the present study are given below in Table A.1.

Table A.1: The summary statistics of the MCS dataset.

Characteristic	Boys, N = 260 <sup>1</sup>	Girls, N = 235 <sup>1</sup>
Age (years)	13.25 (11.67, 14.80)	13.00 (11.30, 14.55)
Height (cm)	152 (142, 164)	150 (143, 156)
Weight (kg)	42 (32, 51)	40 (33, 49)
RMR (kcal/day)	1,172 (1,030, 1,333)	1,043 (928, 1,168)
(Missing)	0	3
FM (kg)	5 (3, 12)	10 (6, 14)
(Missing)	3	1
FFM (kg)	35 (28, 43)	31 (27, 35)
(Missing)	3	1

<sup>1</sup>Median (IQR); n (%)

*Resting metabolic rate (RMR); Fat mass (FM); and Fat-free mass (FFM)*

### A.2.1.2 Altman and Dittmer Dataset

The reference organ mass reported for brain, heart, liver and kidney initially compiled by Altman and Dittmer (1962) provided in Table 1 in Wang (2012) are reproduced in Figure A.1.

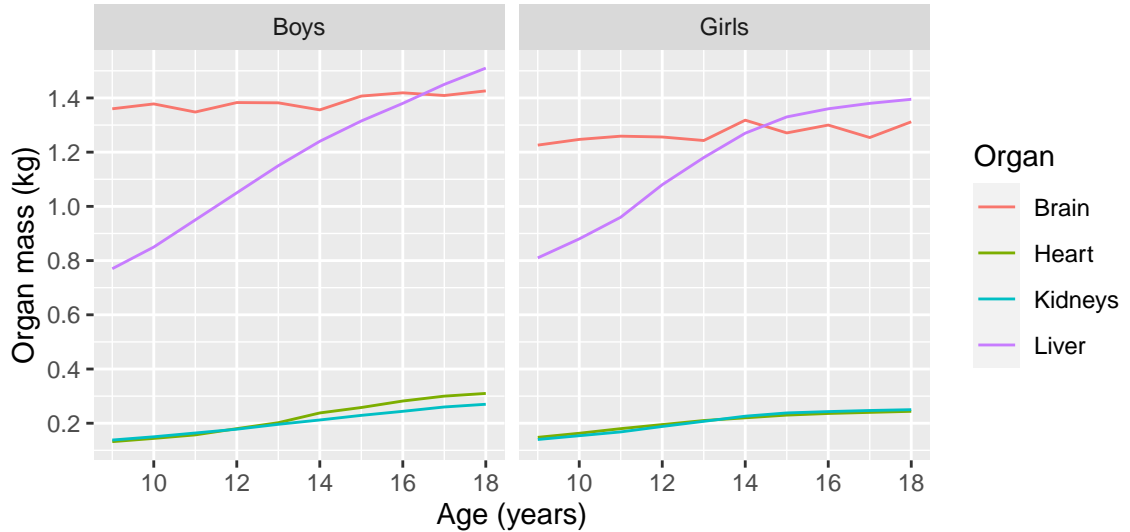


Figure A.1: The organ mass data set for the North American children compiled by Altman and Dittmer (1962) adapted from Table 1 in Wang (2012).

### A.3 Preliminary Data analysis

A large number of previous studies had modelled RMR as a linear combination of age, and anthropometric measurements such as age, height, and weight. Later studies have looked at body composition measures such as FM and FFM. We thus, initially analysed the extend of variation in RMR explained by age, weight, height, gender, fat-free mass, and fat mass.

Weight, height, age and gender are significant predictors of RMR, but can only explain 58% of the observed variation in RMR.

```
m1 <- lm(RMRkcal_day ~ Height_cms + Weight_kg + Age_yrs +
  Gender, data = MCS)
summary(m1)
```

Call:

```
lm(formula = RMRkcal_day ~ Height_cms + Weight_kg + Age_yrs +
  Gender, data = MCS)
```

Residuals:

Min	1Q	Median	3Q	Max
-447.07	-98.07	-18.60	90.11	717.42

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	182.3939	132.4192	1.377	0.169
Height_cms	5.6415	1.1924	4.731	2.93e-06 ***
Weight_kg	10.6303	0.9085	11.701	< 2e-16 ***
Age_yrs	-22.0342	4.7132	-4.675	3.81e-06 ***
GenderGirls	-114.7119	14.7101	-7.798	3.85e-14 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 160.2 on 487 degrees of freedom  
(3 observations deleted due to missingness)

Multiple R-squared: 0.5849, Adjusted R-squared: 0.5815

F-statistic: 171.5 on 4 and 487 DF, p-value: < 2.2e-16

Moreover, once FFM and fat mass (FATM) are included in the model, Height and weight are no longer significant. Further, the explained variation is marginally increased to 61%.

```
m2 <- lm(RMRkcal_day ~ Height_cms + Weight_kg + Age_yrs +
  Gender + FFM + FATM, data = MCS)
summary(m2)
```

Call:

```
lm(formula = RMRkcal_day ~ Height_cms + Weight_kg + Age_yrs +
  Gender + FFM + FATM, data = MCS)
```

Residuals:

Min	1Q	Median	3Q	Max
-427.05	-99.19	-9.11	80.02	672.19

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	870.312	174.725	4.981	8.83e-07 ***
Height_cms	-1.542	1.687	-0.914	0.361158
Weight_kg	-3.302	4.456	-0.741	0.459091
Age_yrs	-28.761	4.753	-6.052	2.89e-09 ***
GenderGirls	-61.169	17.303	-3.535	0.000447 ***
FFM	28.709	5.240	5.479	6.92e-08 ***
FATM	9.607	4.429	2.169	0.030571 *

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 154.5 on 481 degrees of freedom  
(7 observations deleted due to missingness)  
Multiple R-squared: 0.6169, Adjusted R-squared: 0.6122  
F-statistic: 129.1 on 6 and 481 DF, p-value: < 2.2e-16

### A.3.0.1 Linear model for RMR with Age, sex, FFM and FATM

Age, sex, FFM and FATM are found to explain 61% of the variation observed in RMR in Indian children.

```
m3 <- lm(RMRkcal_day ~ Age_yrs + Gender + FFM + FATM,
data = MCS)
summary(m3)
```

Call:  
lm(formula = RMRkcal\_day ~ Age\_yrs + Gender + FFM + FATM, data = MCS)

Residuals:

Min	1Q	Median	3Q	Max
-431.17	-96.14	-10.80	79.59	668.04

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	718.313	45.449	15.805	< 2e-16 ***
Age_yrs	-29.456	4.709	-6.255	8.76e-10 ***
GenderGirls	-66.240	16.397	-4.040	6.22e-05 ***
FFM	23.310	1.470	15.855	< 2e-16 ***
FATM	6.556	1.146	5.723	1.84e-08 ***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 154.4 on 483 degrees of freedom  
(7 observations deleted due to missingness)  
Multiple R-squared: 0.6158, Adjusted R-squared: 0.6127  
F-statistic: 193.6 on 4 and 483 DF, p-value: < 2.2e-16

### A.3.1 Wang model for RMR

Please refer to the Section 3.2 for the details on the Wang model, relative cellularity, relative organ mass, and relative specific metabolic rate.

### A.3.1.1 Relative cellularity (Rc)

The values of Rc is 1 for the age range 9 to 18 years as given in Figure A.2.

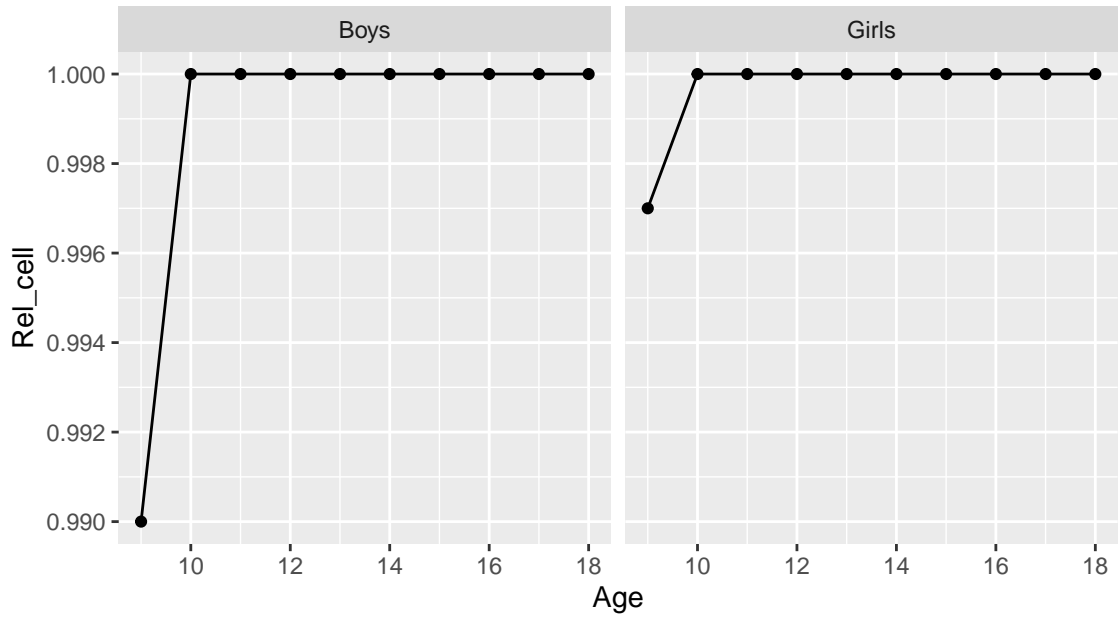


Figure A.2: Relative cellularity (Rc) adapted from Table 2 in Wang (2012).

### A.3.1.2 Relative specific metabolic rate

The Figure A.3 shows the relative specific metabolic rate for the age range 9 to 18 years.

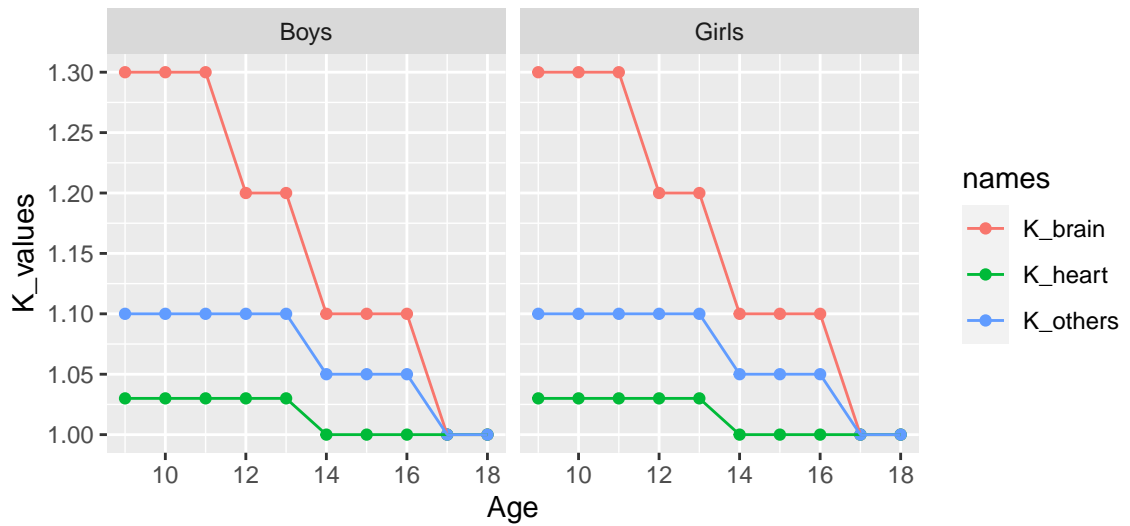


Figure A.3: Relative specific metabolic rate adapted from Table 3 in Wang (2012).

## A.3.2 Results

Wang model predictions for RMR in Indian children is given below.



```

# A function to calculate the wang predicted
# values of rmr if given the relative mass, and
# relative cellularity values for (both varies
# for each age group).

rmr_wang <- function(data) {
  data <- data
  rmr <- data$Rel_cell * ((200 * data$relmass_Liver *
    data$K_others) + (240 * data$relmass_Brain *
    data$K_brain) + (400 * data$relmass_Heart *
    data$K_heart) + (400 * data$relmass_Kidneys *
    data$K_others) + (12 * (1 - (data$relmass_Liver +
    data$relmass_Brain + data$relmass_Heart + data$relmass_Kidneys)) *
    data$K_others))
}

```

### A.3.3 Predicting RMR in MCS data set.

The age groups 9 and 10 years are considered as one for the statistical tests.

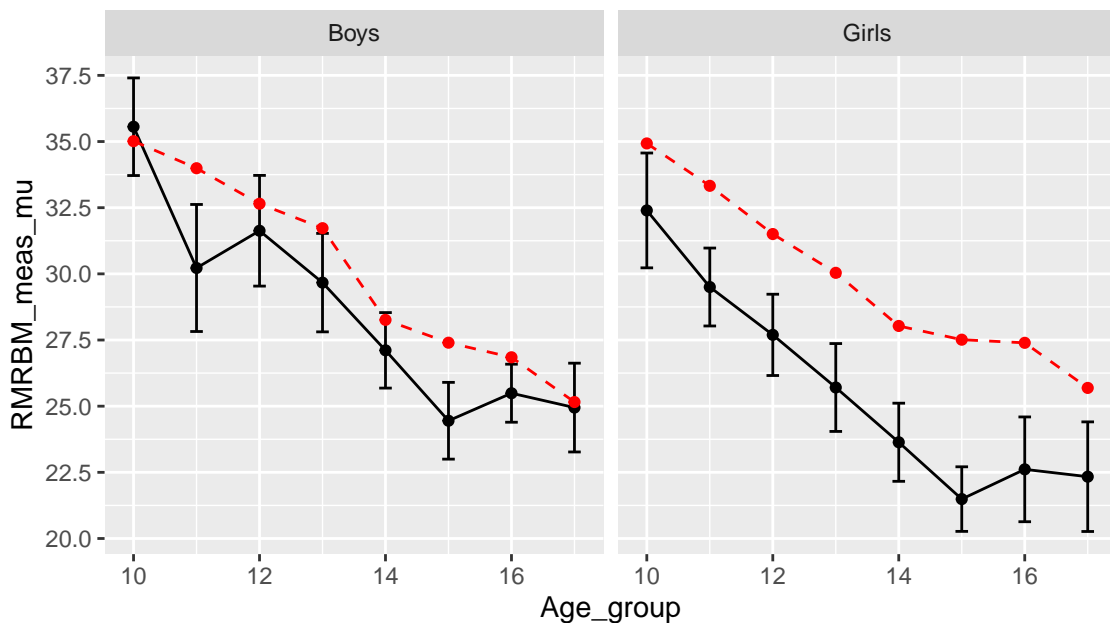


Figure A.4: The mean measured RMR/BM ( $\pm$  SE) measured in each age group (solid line), and the Wang model predicted RMR/BM (dotted line) reproduced here for continuity. Compare to Figure 3.2 in Chapter 3 or Figures 3 and 4 in Areekal et al. (2023). Note that this figure reports SE instead of SD.

The measured RMR/BM ( $\pm$  standard error(SE)) in Indian children is lower compared to the RMR/BM in North American children.

### A.3.4 Modified Wang model of RMR/BM for Indian children

Assuming that the relative organ mass is constant for a given age, we adjust the organ mass of four major organs with by a factor  $\delta$  in Model 1 (see chapter 3). The optimal delta is chosen as the one with minimum mean squared error between the Model 1 predicted RMR/BM and the measured RMR/BM.

```
# Minimize MSE for delta without residual mass
d <- seq(0, 1, 0.01)
err <- rep(0, length(d))

min_delta <- function(d, data) {

  for (i in 1:length(d)) {
    # calculate residual mass for a given delta
    # value
    rel_resid <- 1 - (d[i] * (data$relmass_Brain +
      data$relmass_Heart + data$relmass_Liver +
      data$relmass_Kidneys))
    # predict rmr based on the new residual mass
    rmrbm_new <- ((d[i] * ((data$relmass_Liver *
      data$K_others * 200) + (data$relmass_Kidneys *
      data$K_others * 440) + (data$relmass_Brain *
      data$K_brain * 240) + (data$relmass_Heart *
      data$K_heart * 440)) + (rel_resid * data$K_others *
      12)) * data$Rel_cell)

    err[i] <- sum((rmrbm_new - data$RMRBM_meas)^2)/length(data$RMRBM_meas)
  }

  #
  min_err <- min(err)
  min_err_index <- which(err == min_err)
  delta <- d[min_err_index]

  # plot(d, err, type = 'l', xlab = 'delta', ylab
  # = 'Mean square error')

  out <- list(cbind(d, err), delta)
}
```

#### A.3.4.1 Optimal delta for Model 1 in Boys

```
d <- seq(0, 1.5, 0.01)
err <- rep(0, length(d))

boys_rmrdata <- mcs_wang_agegrp2 %>%
  unnest(c(Age_group, Gender, data)) %>%
  filter(Gender == "Boys")

boys_delmod1 <- min_delta(d, boys_rmrdata)
```

```
# par(pty = 's') plot(boys_delmod1[[1]], type =
# 'l', xlab = 'delta', ylab = 'Mean squared
# error', main = 'Optimal delta for Model 1 in
# Boys') abline(v= boys_delmod1[[2]], lty = 2)
```

```
boys_delmod1[[2]]
```

```
[1] 0.9
```

```
ggplot(data = boys_delmod1[[1]] %>%
  as_tibble, aes(x = d, y = err)) + geom_line() +
  geom_point() + theme(aspect.ratio = 1) + scale_y_continuous(n.breaks = 8) +
  labs(x = "Delta", y = "Mean squared error")
```

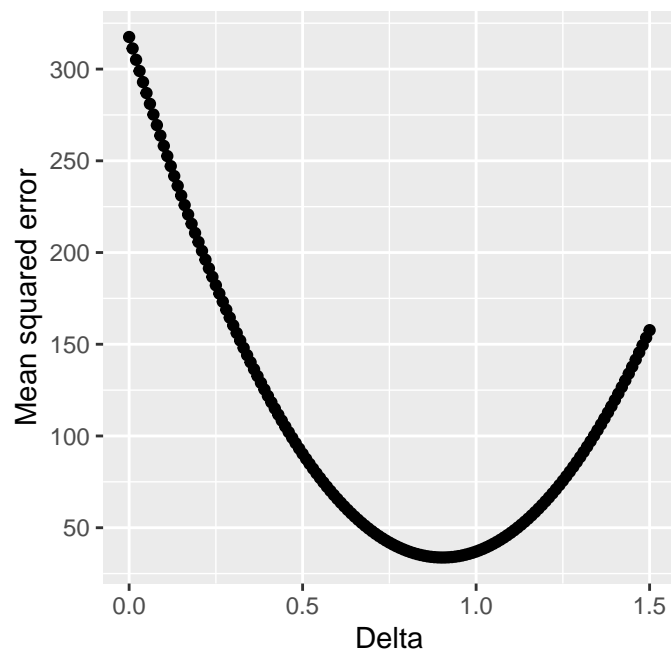


Figure A.5: Optimal delta for Model 1 in Girls is 0.90 used to generate Figure 3.3 in Chapter 3 or Figure 5 in Areekal et al. (2023).

#### A.3.4.2 Optimal delta for Model 1 in Girls

```
girls_rmrdata <- mcs_wang_agegrp2 %>%
  unnest(c(Age_group, Gender, data)) %>%
  filter(Gender == "Girls")
```

```
girls_delmod1 <- min_delta(d, girls_rmrdata)
```

```
girls_delmod1[[2]]
```

```
[1] 0.77
```

```
ggplot(data = girls_delmod1[[1]] %>%
  as_tibble, aes(x = d, y = err)) + geom_line() +
```

```
geom_point() + theme(aspect.ratio = 1) + scale_y_continuous(n.breaks = 8) +
labs(x = "Delta", y = "Mean squared error")
```

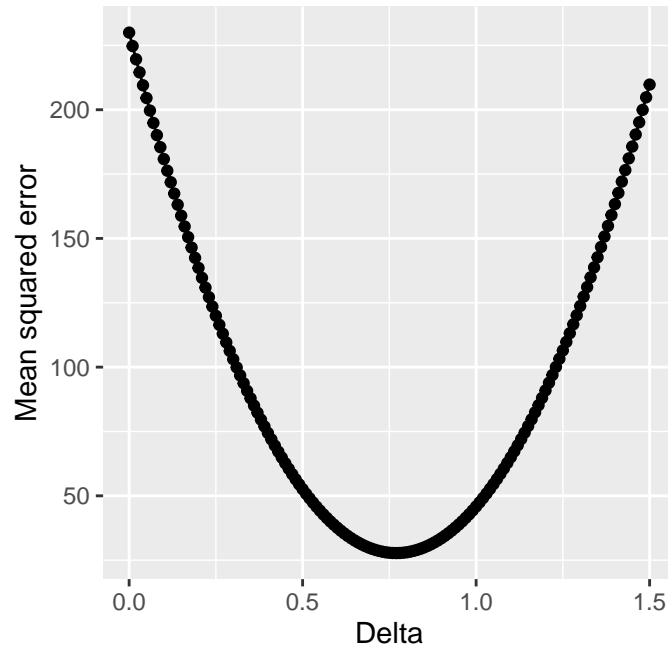


Figure A.6: Optimal delta for Model 1 in Girls is 0.77. Used to generate Figure 3.3 in Chapter 3 or Figure 6 in Areekal et al. (2023).

#### A.3.4.3 Model 1 predicted RMR/BM in Indian children

```
Model1 <- function(delta1, data) {
  # predict residual mass for the given delta
  rel_resid <- 1 - (delta1 * (data$relmass_Brain +
    data$relmass_Heart + data$relmass_Liver + data$relmass_Kidneys))

  # predict rmr based on the new residual mass
  rmrbm_new <- ((delta1 * ((data$relmass_Liver *
    data$K_others * 200) + (data$relmass_Kidneys *
    data$K_others * 440) + (data$relmass_Brain *
    data$K_brain * 240) + (data$relmass_Heart *
    data$K_heart * 440)) + (rel_resid * data$K_others *
    12)) * data$Rel_cell)
}
```

#### A.3.5 An alternate model for RMR in Indian children

```
# Minimize MSE for delta without residual mass
d <- seq(0, 1, 0.01)
err <- rep(0, length(d))
```

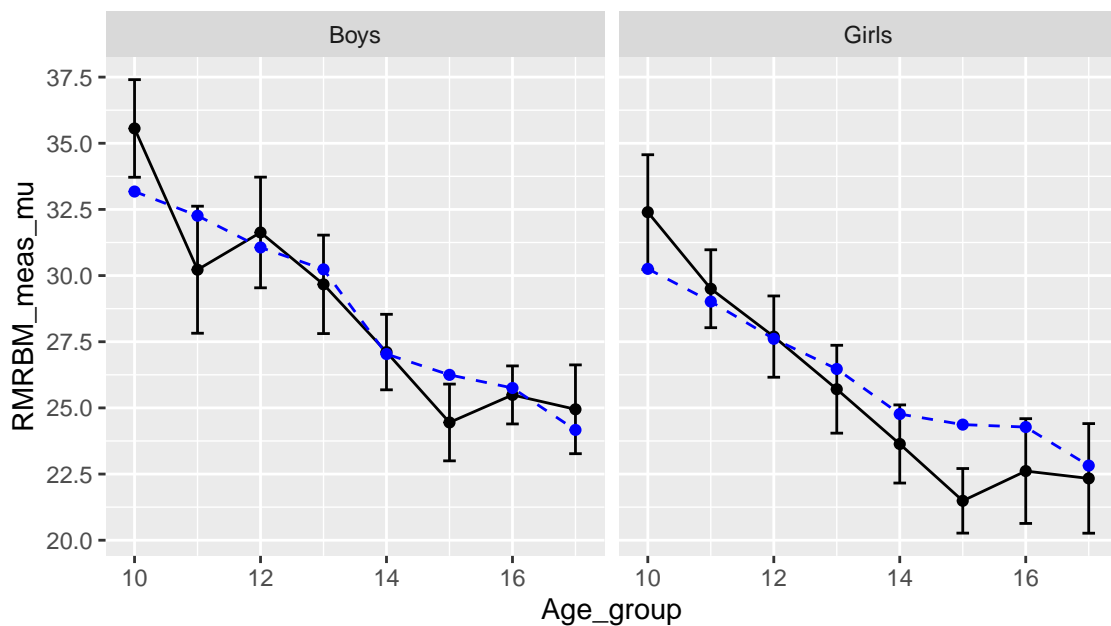


Figure A.7: Mean ( $\pm$ SE) RMR/BM measured (solid line) in each age group compared to the Model 1 predicted RMR/BM (dotted line) for the Caucasian population. Compare to Figure 3.3 in chapter 3 or Figure 5 and 6 in Arekal et al. (2023). Note that the bars represent SE here instead of SD.

```

min_delta_2 <- function(d, data) {

  for (i in 1:length(d)) {
    # calculate residual mass for a given delta
    # value
    rel_resid <- 1 - (data$relmass_Brain + data$relmass_Heart +
      data$relmass_Liver + data$relmass_Kidneys)
    # predict rmr based on the new residual mass
    rmrbm_new <- (((data$relmass_Liver * data$K_others *
      200) + (data$relmass_Kidneys * data$K_others *
      440) + (data$relmass_Brain * data$K_brain *
      240) + (data$relmass_Heart * data$K_heart *
      440) + (d[i] * rel_resid * data$K_others *
      12)) * data$Rel_cell)

    err[i] <- sum((rmrbm_new - data$RMRBM_meas)^2)/length(data$RMRBM_meas)
  }

  #
  min_err <- min(err)
  min_err_index <- which(err == min_err)
  delta <- d[min_err_index]

  # plot(d, err, type = 'l', xlab = 'delta', ylab
  # = 'Mean square error')

  out <- list(cbind(d, err), delta)
}

```

```
}
```

### A.3.5.1 Optimal delta for Model 2 in boys

```
d <- seq(0, 1, 0.01)
err <- rep(0, length(d))

boys_delmod2 <- min_delta_2(d, boys_rmrdata)

# par(pty = 's') plot(boys_delmod2[[1]], type =
# 'l', xlab = 'delta', ylab = 'Mean squared
# error', main = 'Optimal delta for Model 2 in
# Boys') abline(v= boys_delmod2[[2]], lty = 2)

boys_delmod2[[2]]
```

```
[1] 0.85
```

```
ggplot(data = boys_delmod2[[1]] %>%
  as_tibble, aes(x = d, y = err)) + geom_line() +
  geom_point() + theme(aspect.ratio = 1) + scale_y_continuous(n.breaks = 8) +
  labs(x = "Delta", y = "Mean squared error")
```

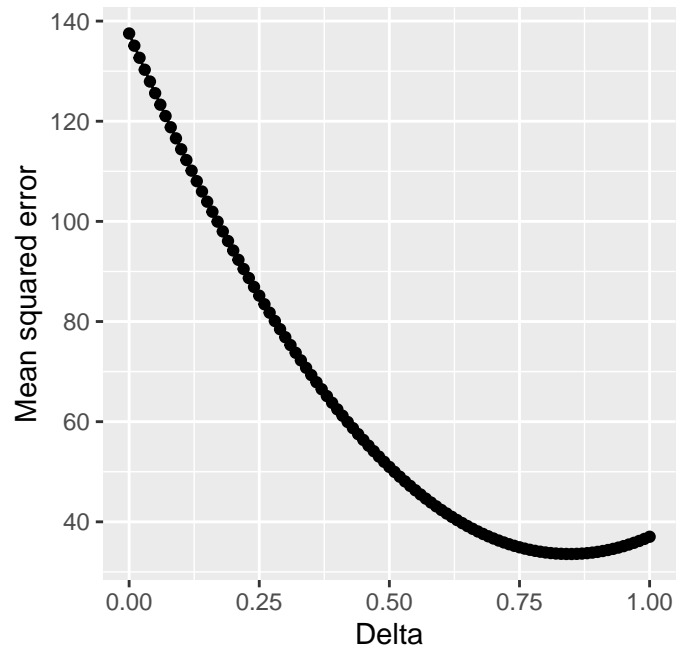


Figure A.8: Optimal delta for Model 2 in boys is 0.85, which is used produce Figure 3.5 in Chapter 3 or Figure 8 in Areekal et al. (2023).

### A.3.5.2 Optimal delta for Model 2 in Girls

```
girls_delmod2 <- min_delta_2(d, girls_rmrdata)

# par(pty = 's') plot(girls_delmod2[[1]], type =
```

```
# 'l', xlab = 'delta', ylab = 'Mean squared
# error', main = 'Optimal delta for Model 2 in
# Girls') abline(v = girls_delmod2[[2]], lty = 2)
```

```
girls_delmod2[[2]]
```

```
[1] 0.64
```

```
ggplot(data = girls_delmod2[[1]] %>%
  as_tibble, aes(x = d, y = err)) + geom_line() +
  geom_point() + theme(aspect.ratio = 1) + scale_y_continuous(n.breaks = 8) +
  labs(x = "Delta", y = "Mean squared error")
```

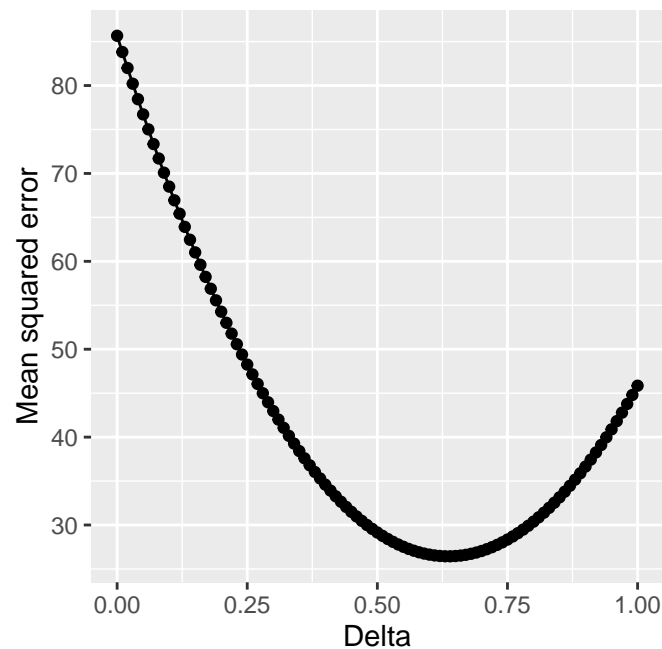


Figure A.9: Optimal delta for Model 2 in girls is 0.64, which is used produce Figure 3.5 in chapter 3 or Figure 9 in Areekal et al. (2023).

### A.3.6 Model 2 predicted RMR/BM in Indian children

```
Model2 <- function(delta2, data) {
  # predict residual mass for the given delta
  rel_resid <- 1 - (data$relmass_Brain + data$relmass_Heart +
    data$relmass_Liver + data$relmass_Kidneys)

  # predict rmr based on the new residual mass
  rmrbm_new <- (((data$relmass_Liver * data$K_others *
    200) + (data$relmass_Kidneys * data$K_others *
    440) + (data$relmass_Brain * data$K_brain *
    240) + (data$relmass_Heart * data$K_heart *
    440) + (delta2 * rel_resid * data$K_others *
    12)) * data$Rel_cell)
}
```

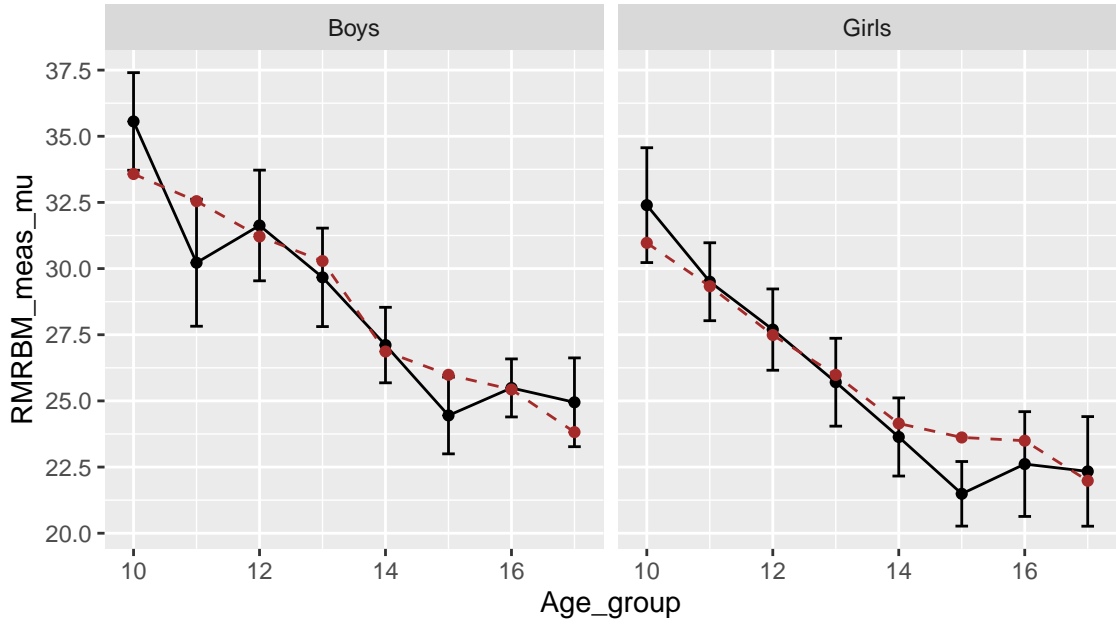


Figure A.10: The mean ( $\pm$  SE) RMR/BM measured (solid lines) and the Model 2 predicted RMR/BM (dotted line). Comparable to Figure 3.5 in Chapter 3 or Figures 8 and 9 in Areekal et al. (2023) which reports the mean measured RMR/BM ( $\pm$  SD).

## A.4 Modelling height growth in Indian children

Traditionally, height growth in children are studied using two major frameworks, namely *centiles* using cross-sectional data and *curves* using longitudinal data. Here we study the PSCG data set with longitudinal height measurements Khadilkar et al. (2019). The growth centiles are constructed using the Generalised Additive Model for Location Scale and Shape (Rigby and Stasinopoulos 2005) and growth curves using the SuperImposition by Translation and Rotation model (Cole et al. 2010).

### A.4.1 Dataset

#### A.4.1.1 PSCG data set

The age range for the analysis is restricted to 6 to 19 years. The number of measurements per subject is shown in Figure A.11. The height measurements in the PSCG data are shown in Figure A.12 and the growth curves of the subjects are shown in Figure A.13

Table A.2: Height measurements in the PSCG dataset.

Characteristic	Boys, N = 4,252	Girls, N = 3,543
Age (years)	11.30 (9.00, 13.60)	11.50 (9.20, 13.60)
Height (cm)	143 (131, 158)	145 (131, 154)



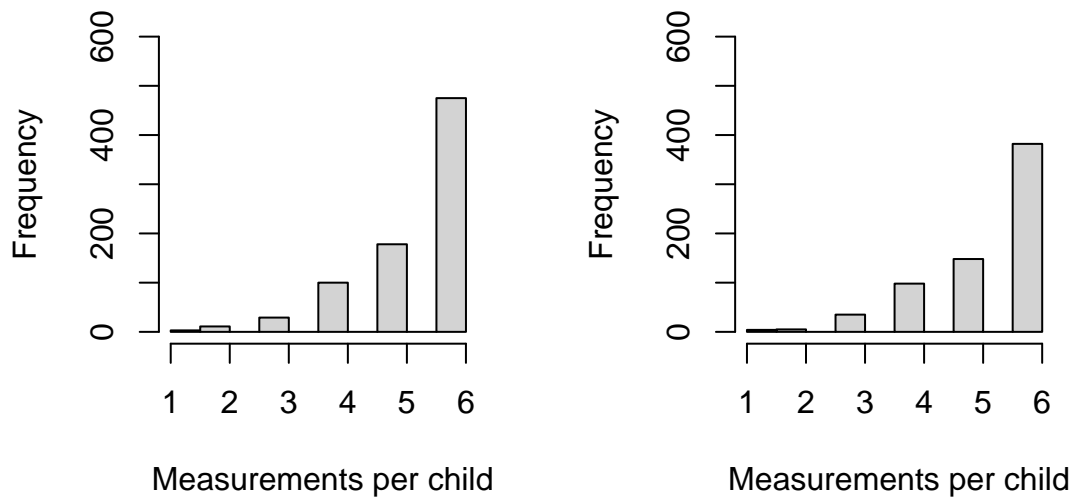


Figure A.11: PSCG study data described in section 4.2.1.



Figure A.12: Heights measured in the PSCG dataset described in section 4.2.1.

#### A.4.1.2 Initial SITAR model fitting

The `sitar()` argument from the package *sitar* (Cole 2022) takes age, height and id of each subject in the long format. The degrees of freedom of the spline curve determines its shape. The age and height axis can be scaled using logarithmic transformation to consider multiplicative nature of the growth process.

One can choose and any model with x scale as age or  $\log(\text{age})$ , y scale as height or  $\log(\text{height})$ , df ranging from 4 to 8, with or without an fixed effects. All SITAR models are fitted with the fixed effect size, denoted by  $a$  in the *sitar* package. The fixed effects  $b$  and  $c$  are assumed to be normally distributed around mean = 0 by default. If specified separately in the model, for instance as `fixed = "a+b+c"`,  $b$  and  $c$  are can be distributed around the non-zero sample mean. SITAR models with four combinations of fixed effects can be fitted: with `fixed=c("a", "a+b", "a+c", "a+b+c"`.

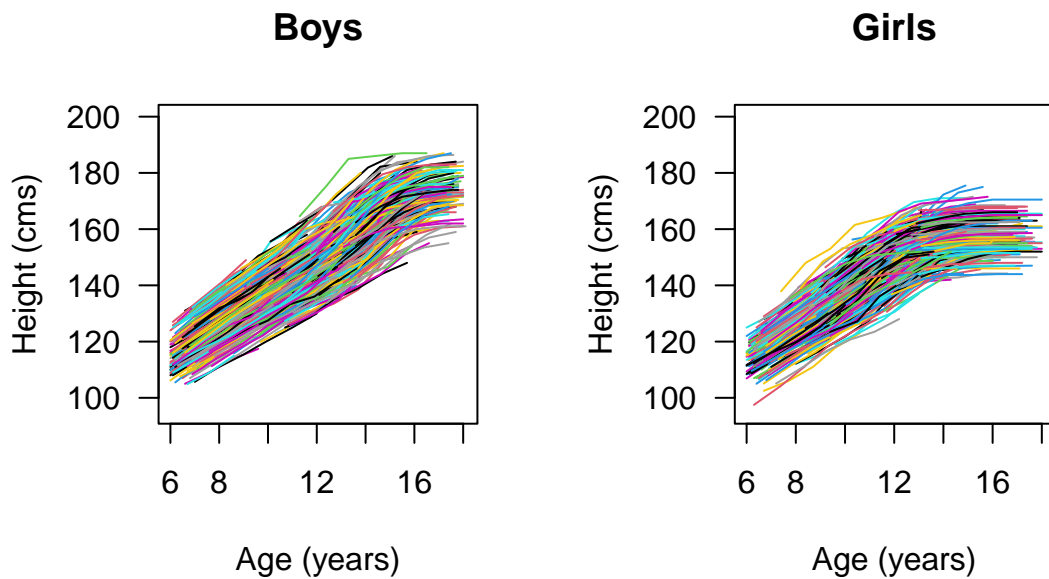


Figure A.13: Individual growth curves in the PSCG dataset comparable to Figure 4.3 in chapter 4 and Figure 3 in Aarekal et al. (2022).

The initial SITAR model in boys is fitted in  $\log(\text{age})$  and height scale with degrees of freedom (df) for the spline curves set as 6.

```
# boys
boys_2$htok <- TRUE
F1_B <- sitar(log(age), height, ID, boys_2[boys_2$htok ==
  TRUE, ], 6)
F1_B$ok <- boys_2$htok
# number of iterations
F1_B$num
```

```
[1] 16
```

```
summary(F1_B)
```

SITAR nonlinear mixed-effects model fit by maximum likelihood

```
Call: sitar(x = log(age), y = height, id = ID, data = boys_2[boys_2$htok ==
  TRUE, ], df = 6)
```

```
      AIC      BIC    logLik
17987.4 18089.09 -8977.702
```

Random effects:

```
Formula: list(a ~ 1, b ~ 1, c ~ 1)
```

```
Level: id
```

```
Structure: General positive-definite, Log-Cholesky parametrization
```

```
      StdDev      Corr
a      6.80889711 a      b
b      0.08486528 0.457
c      0.14602401 0.527 0.531
```

```
Residual 0.83054263
```

```
Fixed effects: s1 + s2 + s3 + s4 + s5 + s6 + a + b + c ~ 1
```

```
      Value Std.Error  DF  t-value p-value
```

```

s1 21.80769 0.4251856 3448 51.28982 0.0000
s2 28.65730 0.7416036 3448 38.64235 0.0000
s3 35.02615 0.6708226 3448 52.21373 0.0000
s4 53.22042 1.1859311 3448 44.87649 0.0000
s5 68.24102 1.3674500 3448 49.90385 0.0000
s6 53.92083 0.9908962 3448 54.41622 0.0000
a 113.29651 1.0342721 3448 109.54227 0.0000
b -0.01389 0.0082492 3448 -1.68322 0.0924
c 0.07689 0.0302193 3448 2.54436 0.0110

```

Correlation:

```

      s1      s2      s3      s4      s5      s6      a      b
s2 0.960
s3 0.972 0.944
s4 0.955 0.973 0.963
s5 0.963 0.987 0.959 0.984
s6 0.933 0.965 0.932 0.975 0.981
a -0.925 -0.961 -0.907 -0.946 -0.955 -0.947
b -0.778 -0.853 -0.749 -0.834 -0.841 -0.857 0.906
c -0.933 -0.963 -0.933 -0.977 -0.966 -0.966 0.966 0.889

```

Standardized Within-Group Residuals:

```

      Min      Q1      Med      Q3      Max
-6.55395059 -0.39942791 0.01271697 0.41761598 5.56313478

```

Number of Observations: 4252

Number of Groups: 796

The initial sitar model in girls is fitted in  $\log(\text{age})$  and height scale with  $df = 5$ , and fixed effects "a+c".

SITAR nonlinear mixed-effects model fit by maximum likelihood

Call:

```

sitar(x = log(age), y = height, id = ID, data = Girls_2[Girls_2$htok ==
  TRUE, ], df = 5, fixed = "a+c")

```

```

      AIC      BIC    logLik
14196.56 14282.98 -7084.28

```

Random effects:

Formula: list(a ~ 1, b ~ 1, c ~ 1)

Level: id

Structure: General positive-definite, Log-Cholesky parametrization

```

      StdDev      Corr
a      6.15008329 a      b
b      0.08494961 0.419
c      0.14218422 0.424 0.345
Residual 0.69315900

```

Fixed effects: s1 + s2 + s3 + s4 + s5 + a + c ~ 1

```

      Value Std.Error  DF  t-value p-value
s1 24.71263 0.7945272 2865 31.10357 0.0000
s2 40.28785 1.4063276 2865 28.64756 0.0000
s3 39.77952 0.9849336 2865 40.38802 0.0000
s4 55.40859 1.6059349 2865 34.50239 0.0000
s5 37.41178 0.8274197 2865 45.21501 0.0000
a 112.70934 1.0940097 2865 103.02407 0.0000

```

```

c    0.05157 0.0423617 2865    1.21743  0.2235
Correlation:
  s1    s2    s3    s4    s5    a
s2  0.990
s3  0.972  0.981
s4  0.981  0.988  0.987
s5  0.965  0.976  0.981  0.980
a  -0.970 -0.976 -0.967 -0.976 -0.957
c  -0.975 -0.987 -0.972 -0.971 -0.962  0.970

```

```

Standardized Within-Group Residuals:
      Min          Q1          Med          Q3          Max
-5.900228636 -0.377054883 -0.005894336  0.386052444  5.815822506

```

```

Number of Observations: 3543
Number of Groups: 672

```

#### A.4.1.3 Plotting outliers beyond 4 SD

```

par(pty = "s")

## boys
plot.lme(F1_B, id = pnorm(-4), idlabel = getGroups(F1_B),
         pch = "o")

## girls
plot.lme(Fit1_G, id = pnorm(-4), idlabel = getGroups(Fit1_G),
         pch = "o")

```

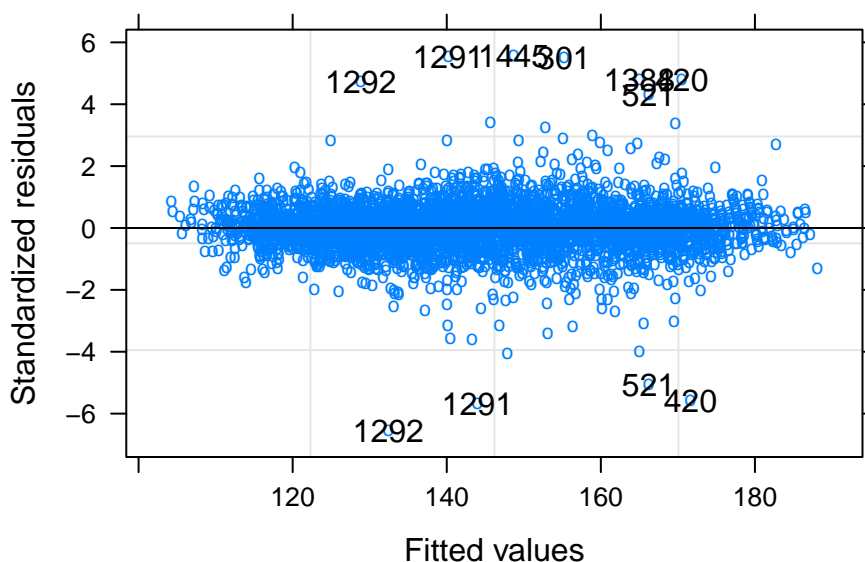


Figure A.14: Residuals from the initial SITAR model in boys.

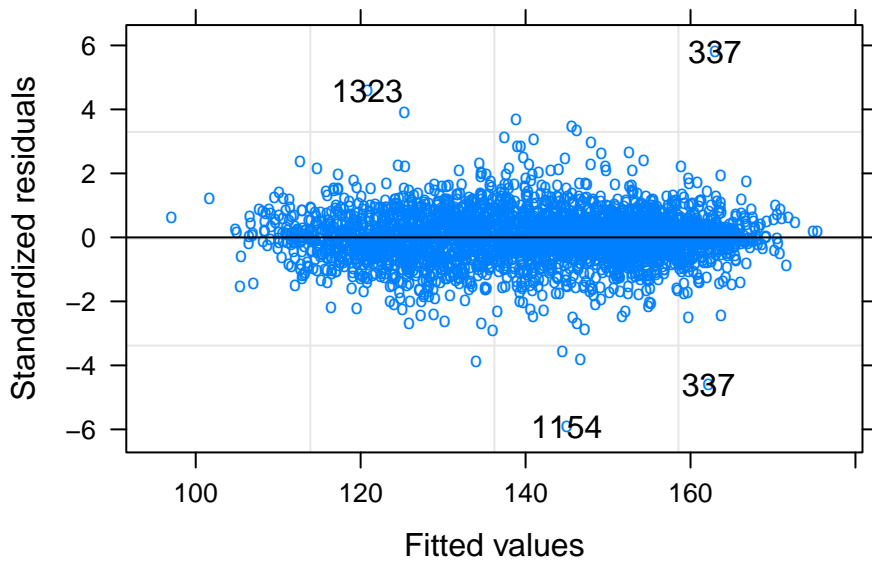


Figure A.15: Residuals from the initial SITAR model in girls.

#### A.4.1.4 Remove the points beyond 4SD and refit the SITAR model

Boys

```
boys_2$htok[boys_2$htok == TRUE] <- abs(residuals(F1_B,
  type = "p")) < 4
F2_B <- update(F1_B)
F2_B$ok <- boys_2$htok
summary(F2_B)
```

SITAR nonlinear mixed-effects model fit by maximum likelihood

Call: sitar(x = log(age), y = height, id = ID, data = boys\_2[boys\_2\$htok == TRUE, ], df = 6)

	AIC	BIC	logLik
	17495.11	17596.74	-8731.554

Random effects:

Formula: list(a ~ 1, b ~ 1, c ~ 1)

Level: id

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr		
a	6.84919618		a	b
b	0.08381081	0.461		
c	0.14648099	0.508	0.508	
Residual	0.75178565			

Fixed effects: s1 + s2 + s3 + s4 + s5 + s6 + a + b + c ~ 1

	Value	Std.Error	DF	t-value	p-value
s1	21.40214	0.3651492	3435	58.61204	0.0000
s2	27.94903	0.6480191	3435	43.12994	0.0000
s3	34.46174	0.5680840	3435	60.66311	0.0000
s4	52.34031	1.0402736	3435	50.31398	0.0000
s5	67.03076	1.1949388	3435	56.09556	0.0000
s6	53.24436	0.8960362	3435	59.42211	0.0000
a	114.18935	0.9184247	3435	124.33176	0.0000

```

b   -0.00831 0.0076221 3435  -1.08967  0.2759
c    0.09966 0.0273633 3435   3.64220  0.0003
Correlation:
  s1    s2    s3    s4    s5    s6    a    b
s2  0.954
s3  0.966  0.930
s4  0.949  0.971  0.953
s5  0.957  0.985  0.948  0.985
s6  0.928  0.964  0.919  0.976  0.982
a  -0.908 -0.951 -0.881 -0.935 -0.944 -0.938
b  -0.746 -0.833 -0.704 -0.813 -0.819 -0.837  0.894
c  -0.924 -0.959 -0.916 -0.972 -0.963 -0.962  0.959  0.874

```

```

Standardized Within-Group Residuals:
      Min          Q1          Med          Q3          Max
-3.86688200 -0.41968201  0.01611758  0.44055470  3.80117913

```

```

Number of Observations: 4239
Number of Groups: 796

```

Girls

```

Girls_2$htok[Girls_2$htok == TRUE] <- abs(residuals(Fit1_G,
  type = "p")) < 4
Fit2_G <- update(Fit1_G)
Fit2_G$ok <- Girls_2$htok
summary(Fit2_G)

```

SITAR nonlinear mixed-effects model fit by maximum likelihood

```

Call:
sitar(x = log(age), y = height, id = ID, data = Girls_2[Girls_2$htok ==
  TRUE, ], df = 5, fixed = "a+c")
      AIC      BIC   logLik
14036.62 14123.02 -7004.31

```

Random effects:

```

Formula: list(a ~ 1, b ~ 1, c ~ 1)
Level: id
Structure: General positive-definite, Log-Cholesky parametrization
      StdDev      Corr
a      6.12987170 a      b
b      0.08486755 0.415
c      0.14321221 0.414 0.340
Residual 0.66655667

```

Fixed effects: s1 + s2 + s3 + s4 + s5 + a + c ~ 1

```

      Value Std.Error  DF  t-value p-value
s1  24.65585 0.7836250 2861  31.46383 0.0000
s2  40.05495 1.3890995 2861  28.83520 0.0000
s3  39.59661 0.9788386 2861  40.45264 0.0000
s4  55.12976 1.5834952 2861  34.81523 0.0000
s5  37.25403 0.8203838 2861  45.41049 0.0000
a  112.86387 1.0821011 2861 104.30068 0.0000
c   0.05792 0.0422145 2861   1.37210 0.1701
Correlation:

```

```

      s1      s2      s3      s4      s5      a
s2  0.991
s3  0.972  0.982
s4  0.981  0.989  0.988
s5  0.966  0.977  0.982  0.981
a  -0.970 -0.976 -0.967 -0.976 -0.957
c  -0.976 -0.987 -0.972 -0.972 -0.963  0.969

```

Standardized Within-Group Residuals:

```

              Min              Q1              Med              Q3              Max
-4.0790845178 -0.3890507233  0.0004300837  0.4001396725  4.0526308530

```

Number of Observations: 3539

Number of Groups: 672

#### A.4.1.5 Summary of cleaned data set

Table A.3: Summary of the cleaned PSCG dataset.

Characteristic	Boys, N = 4,239	Girls, N = 3,539
Age (years)	11.30 (9.00, 13.60)	11.50 (9.20, 13.60)
Height (cm)	143 (131, 158)	145 (131, 154)

#### A.4.1.6 SITAR model optimisation

We fitted multiple SITAR models using logarithmic and power transformations of age and height scales. Four combinations of fixed effects are also studied. Note that all SITAR models fix with the size fixed effect. If no fixed effect is defined in the model, the mean is assumed to be distributed around 0.

The model with the least BIC is chosen as the optimal model. The BIC values for multiple models were given by the `dfpower` argument in the *sitar* package Cole (2022). The -ve values denote models that were fitted with a warning.

```

dfb <- dfpower(F2_B, df = 4:8, xpowers = 0:1, ypowers = 0:1,
  fixed = c("a", "a+b", "a+c", "a+b+c"))
dfb

df_g <- dfpower(Fit2_G, df = 4:8, xpowers = 0:1, ypowers = 0:1,
  fixed = c("a", "a+b", "a+c", "a+b+c"))
df_g

# not run here due to computational load

```

Boys:

```
dfb
```

```

, , log(height), log(age)

      a      a+b      a+c      a+b+c
4 -18403.5 -17907.9 -18346.4 -17574.5
5 -17601.8 -17600.3 -17611.4 -17541.8
6 -17551.3 -17532.7 -17556.4 -17573.4
7 -17650.4 -17662.8 -17521.6 -17602.1
8 -17605.3 -17653.8 -17637.7 -18225.1

```

```
, , height, log(age)
```

	a	a+b	a+c	a+b+c
4	-18625.7	18099.9	-18618.3	17829.8
5	17702.7	17766.3	17824.6	17740.7
6	17701.3	17623.2	17656.3	17609.9
7	17789.2	17711.9	-17709.1	17728.9
8	17747.7	17784.3	17766.7	18335.4

```
, , log(height), age
```

	a	a+b	a+c	a+b+c
4	-18231.5	-17917.5	-18162.8	-17797.4
5	-17811.4	-17825.7	-17742.6	-17746.4
6	-17771.4	-17765.4	-17667.6	-17677.4
7	-17929.7	-17791.6	-17679.1	-17702.5
8	-17815.1	-17964.1	-17844.7	-17775.1

```
, , height, age
```

	a	a+b	a+c	a+b+c
4	18438.2	-18041.6	18298.7	-17787.9
5	17843.7	17855.0	-17814.1	17728.7
6	17826.6	17792.2	17722.5	17731.5
7	-17986.6	17864.6	17789.5	17748.2
8	17868.0	-18015.9	17891.1	-18416.3

```
sort(dfb)
```

```
[1] -18625.7 -18618.3 -18416.3 -18403.5 -18346.4 -18231.5 -18225.1 -18162.8
[9] -18041.6 -18015.9 -17986.6 -17964.1 -17929.7 -17917.5 -17907.9 -17844.7
[17] -17825.7 -17815.1 -17814.1 -17811.4 -17797.4 -17791.6 -17787.9 -17775.1
[25] -17771.4 -17765.4 -17746.4 -17742.6 -17709.1 -17702.5 -17679.1 -17677.4
[33] -17667.6 -17662.8 -17653.8 -17650.4 -17637.7 -17611.4 -17605.3 -17602.1
[41] -17601.8 -17600.3 -17574.5 -17573.4 -17556.4 -17551.3 -17541.8 -17532.7
[49] -17521.6 17609.9 17623.2 17656.3 17701.3 17702.7 17711.9 17722.5
[57] 17728.7 17728.9 17731.5 17740.7 17747.7 17748.2 17766.3 17766.7
[65] 17784.3 17789.2 17789.5 17792.2 17824.6 17826.6 17829.8 17843.7
[73] 17855.0 17864.6 17868.0 17891.1 18099.9 18298.7 18335.4 18438.2
```

The lowest positive BIC (17609.9) corresponds to the model with  $x=\log(\text{age})$ ,  $y=\text{height}$ ,  $df=6$  and fixed effects  $a$ ,  $b$  and  $c$ .

Girls:

```
df_g
```

```
, , log(height), log(age)
```

	a	a+b	a+c	a+b+c
4	-14831.8	-14409.4	-14323.5	-14265.1
5	-14232.2	-14242.7	-14234.2	-14255.2
6	-14346.4	-14283.3	-14267.2	-14280.1
7	-14282.9	-14253.7	-14274.1	-14284.5
8	-14276.5	-14257.1	-14401.5	-14272.7



```
, , height, log(age)

      a      a+b      a+c      a+b+c
4 14751.3 14330.4 14219.7 -14160.4
5 -14122.4 14136.0 14123.9 -14126.4
6  14239.9 14152.9 14148.2  14145.5
7  14154.8 14158.3 14163.5  14169.6
8  14164.6 14155.4 14158.9  14165.4
```

```
, , log(height), age

      a      a+b      a+c      a+b+c
4 -14817.1 -14505.3 -14440.1 -14408.5
5 -14297.4 -14285.0 -14304.4 -14293.5
6 -14346.6 -14324.5 -14324.1 -14321.8
7 -14312.6 -14322.1 -14315.7 -14331.5
8 -14310.8 -14322.3 -14339.0 -14356.7
```

```
, , height, age

      a      a+b      a+c      a+b+c
4 14681.2 14322.0 14233.9 -14188.2
5 14147.4 14160.9 -14158.5 14176.6
6 14255.5 14205.4 14182.8 14182.9
7 -14189.5 -14198.1 14199.2 -14287.0
8 -14205.3 -14213.8 14209.8 14219.8
```

```
sort(df_g)
```

```
[1] -14831.8 -14817.1 -14505.3 -14440.1 -14409.4 -14408.5 -14401.5 -14356.7
[9] -14346.6 -14346.4 -14339.0 -14331.5 -14324.5 -14324.1 -14323.5 -14322.3
[17] -14322.1 -14321.8 -14315.7 -14312.6 -14310.8 -14304.4 -14297.4 -14293.5
[25] -14287.0 -14285.0 -14284.5 -14283.3 -14282.9 -14280.1 -14276.5 -14274.1
[33] -14272.7 -14267.2 -14265.1 -14257.1 -14255.2 -14253.7 -14242.7 -14234.2
[41] -14232.2 -14213.8 -14205.3 -14198.1 -14189.5 -14188.2 -14160.4 -14158.5
[49] -14126.4 -14122.4 14123.9 14136.0 14145.5 14147.4 14148.2 14152.9
[57] 14154.8 14155.4 14158.3 14158.9 14160.9 14163.5 14164.6 14165.4
[65] 14169.6 14176.6 14182.8 14182.9 14199.2 14205.4 14209.8 14219.7
[73] 14219.8 14233.9 14239.9 14255.5 14322.0 14330.4 14681.2 14751.3
```

The lowest positive BIC (14123.9) corresponds to the model with  $x=\log(\text{age})$  vs height,  $df=5$ , fixed effects =a+c

#### A.4.1.7 Optimal model in boys

log(age) vs height,  $df=6$ , fixed effects =a+b+c

SITAR nonlinear mixed-effects model fit by maximum likelihood

Call: sitar(x = log(age), y = height, id = ID, data = boys\_final, df = 6)

```
      AIC      BIC      logLik
17495.11 17596.74 -8731.554
```

Random effects:

Formula: list(a ~ 1, b ~ 1, c ~ 1)

Level: id

```

Structure: General positive-definite, Log-Cholesky parametrization
      StdDev      Corr
a      6.84919618 a      b
b      0.08381081 0.461
c      0.14648099 0.508 0.508
Residual 0.75178565

```

Fixed effects: s1 + s2 + s3 + s4 + s5 + s6 + a + b + c ~ 1

```

      Value Std.Error   DF   t-value p-value
s1  21.40214 0.3651492 3435  58.61204 0.0000
s2  27.94903 0.6480191 3435  43.12994 0.0000
s3  34.46174 0.5680840 3435  60.66311 0.0000
s4  52.34031 1.0402736 3435  50.31398 0.0000
s5  67.03076 1.1949388 3435  56.09556 0.0000
s6  53.24436 0.8960362 3435  59.42211 0.0000
a  114.18935 0.9184247 3435 124.33176 0.0000
b   -0.00831 0.0076221 3435  -1.08967 0.2759
c    0.09966 0.0273633 3435   3.64220 0.0003

```

Correlation:

```

      s1      s2      s3      s4      s5      s6      a      b
s2  0.954
s3  0.966  0.930
s4  0.949  0.971  0.953
s5  0.957  0.985  0.948  0.985
s6  0.928  0.964  0.919  0.976  0.982
a  -0.908 -0.951 -0.881 -0.935 -0.944 -0.938
b  -0.746 -0.833 -0.704 -0.813 -0.819 -0.837  0.894
c  -0.924 -0.959 -0.916 -0.972 -0.963 -0.962  0.959  0.874

```

Standardized Within-Group Residuals:

```

      Min      Q1      Med      Q3      Max
-3.86688200 -0.41968201  0.01611758  0.44055470  3.80117913

```

Number of Observations: 4239

Number of Groups: 796

#### A.4.1.8 The optimal model in girls

log(age) vs height, df=5, fixed effects =a+c

SITAR nonlinear mixed-effects model fit by maximum likelihood

```

Call: sitar(x = log(age), y = height, id = ID, data = girls_final,
df = 5, fixed = "a+c")

```

```

      AIC      BIC   logLik
14036.62 14123.02 -7004.31

```

Random effects:

Formula: list(a ~ 1, b ~ 1, c ~ 1)

Level: id

Structure: General positive-definite, Log-Cholesky parametrization

```

      StdDev      Corr
a      6.12987170 a      b
b      0.08486755 0.415
c      0.14321221 0.414 0.340
Residual 0.66655667

```

Fixed effects: s1 + s2 + s3 + s4 + s5 + a + c ~ 1

	Value	Std.Error	DF	t-value	p-value
s1	24.65585	0.7836250	2861	31.46383	0.0000
s2	40.05495	1.3890995	2861	28.83520	0.0000
s3	39.59661	0.9788386	2861	40.45264	0.0000
s4	55.12976	1.5834952	2861	34.81523	0.0000
s5	37.25403	0.8203838	2861	45.41049	0.0000
a	112.86387	1.0821011	2861	104.30068	0.0000
c	0.05792	0.0422145	2861	1.37210	0.1701

Correlation:

	s1	s2	s3	s4	s5	a
s1	0.991					
s2	0.972	0.982				
s3	0.981	0.989	0.988			
s4	0.966	0.977	0.982	0.981		
s5	-0.970	-0.976	-0.967	-0.976	-0.957	
a	-0.976	-0.987	-0.972	-0.972	-0.963	0.969

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-4.0790845178	-0.3890507233	0.0004300837	0.4001396725	4.0526308530

Number of Observations: 3539

Number of Groups: 672

## A.4.2 GAMLSS analysis

Please refer to the section 4.2.2.1, for more details.

GAMLSS Rigby and Stasinopoulos (2005) introduces multiple distributions capable of accommodating distributions that are asymmetric. These distributions leverage the Box-Cox transformation of a variable ( $y$ ) regulated by a power parameter  $\nu$ .

The Box-Cox Cole and Green distribution (BCCG), is employed to account for skewness in the data, enabling the modeling of asymmetric characteristics. When both skewness and kurtosis are present, two additional distributions: Box-Cox power exponential (BCPE) and Box-Cox (BCT) distributions are available. These distributions can capture the complex data shapes encountered in real-world scenarios.

The parameter estimation for these distributions assumes a linear relationship between mean ( $\mu$ ) and the Box-Cox parameter ( $\nu$ ), whereas the relationship between standard deviation ( $\sigma$ ) and  $\nu$  follows a logarithmic curve. However, an alternative approach allowing a logarithmic relationship between  $\mu$  and  $\nu$  is also explored (as in the case of BCCGo, BPEo, and BCTo distributions).

To estimate moment curves (mean, SD), the model employs penalized B-splines or P-splines. These flexible curves effectively capture intricate relationships between variables. The model employs cross-validation to find the appropriate level of degrees of freedom of the curve.

The global deviance of gamlss models under different distributions (NO, BCCGo, BPEo, BCTo) are given below (for the final iteration). The the centiles under the optimal model is shown in Figure A.16 and Figure A.17. A table showing the portion of the sample that fall under each theoretical centile is also given.

#### A.4.2.1 Boys

```
gml_b <- lms(height, age, k = log(nrow(boys_final)),
  data = boys_final)
```

```
*** Initial fit***
GAMLSS-RS iteration 2: Global Deviance = 28044.15
*** Fitting BCCGo ***
GAMLSS-RS iteration 6: Global Deviance = 27877.58
*** Fitting BCPEo ***
GAMLSS-RS iteration 8: Global Deviance = 27869.76
*** Fitting BCTo ***
GAMLSS-RS iteration 18: Global Deviance = 27877.26
```

### Centile curves using BCCGo

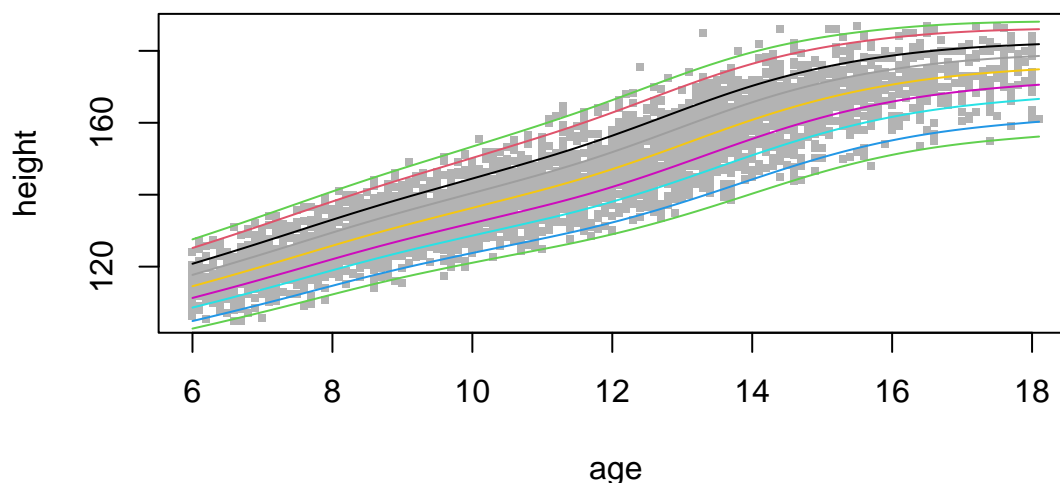


Figure A.16: Centiles estimated for boys under BCCGo distribution for height with age. Comparable to Figure 4.1 in chapter 4 or Figure 1 in Areekal et al. (2022).

	target	calib.	sample
0.4%	0.4	0.423	0.401
2%	2.0	1.654	2.005
10%	10.0	10.175	10.002
25%	25.0	24.831	25.006
50%	50.0	50.542	50.035
75%	75.0	74.551	74.994
90%	90.0	89.810	89.998
98%	98.0	98.214	97.995
99.6%	99.6	99.442	99.599

The boys data is well described by the BCCGo distribution.

#### A.4.2.2 Girls

```
gml_g <- lms(height, age, k = log(nrow(girls_final)),
  data = girls_final)
```

```
*** Initial fit***
```

```

GAMLSS-RS iteration 2: Global Deviance = 22701.42
*** Fitting BCCGo ***
GAMLSS-RS iteration 1: Global Deviance = 22684.87
*** Fitting BCPEo ***
GAMLSS-RS iteration 1: Global Deviance = 22764.89
*** Fitting BCTo ***
GAMLSS-RS iteration 12: Global Deviance = 22644.67
*** Refitting NO ***
GAMLSS-RS iteration 3: Global Deviance = 22654.87

```

### Centile curves using NO

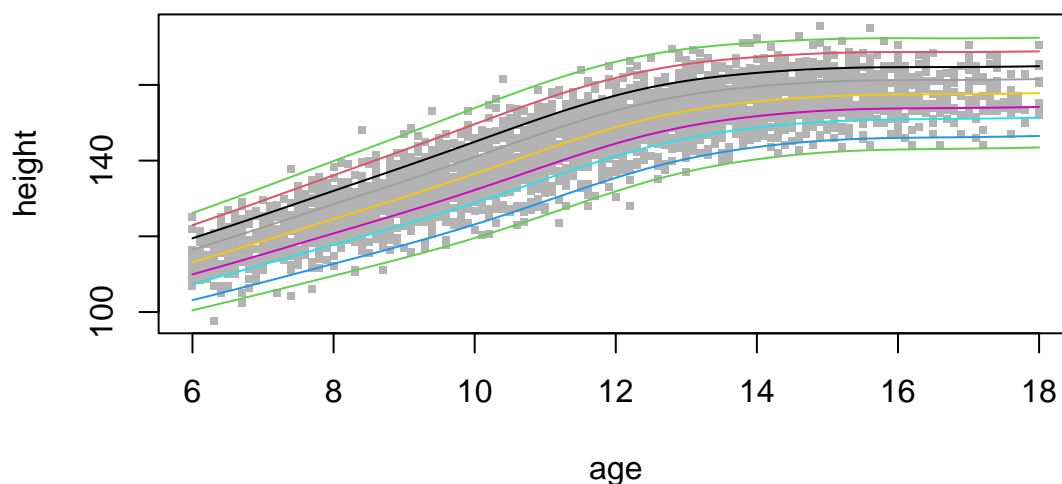


Figure A.17: Centiles estimated for girls under NO distribution for height with age. Comparable to Figure 4.1 in chapter 4 or Figure 1 in Areekal et al. (2022).

	target	calib.	sample
0.4%	0.4	0.388	0.424
2%	2.0	1.771	2.006
10%	10.0	11.276	10.003
25%	25.0	24.579	25.007
50%	50.0	49.387	50.014
75%	75.0	74.944	74.993
90%	90.0	90.266	89.997
98%	98.0	97.848	97.994
99.6%	99.6	99.640	99.576

The girls data is well described by the normal distribution (NO).

Other models with square root and log transformation of the age variable was also explored (not shown here).

## A.5 Modelling height growth in children with Type-1 diabetes

### A.5.1 Datasets

Two data sets on longitudinal height growth are analysed in this study:

1. The **Sweetlings Type-1 Diabetes Mellitus (STDM)** dataset consists of anthropometric measurements in **490 (269 girls)** subjects aged 1 to 29 years diagnosed with T1DM recruited from a tertiary healthcare centre in Pune, India.
2. The **Pune School Children Growth (PSCG)** dataset consists of longitudinal height measurements in **2949 (1268 girls)** school going children from Pune, India.

A summary of longitudinal measurements in the study is given in the Table A.4.

Table A.4: Summary of the longitudinal measurements in STDM and the PSCG dataset.

Characteristic	STDM, N = 1,598	PSCG, N = 8,158
Gender		
Boys	732 (46%)	4,461 (55%)
Girls	866 (54%)	3,697 (45%)
Age (years)	12.2 (3.8)	11.1 (3.1)
Height zscore	-1.25 (1.20)	-0.35 (0.97)
HbA1C (%)	9.80 (8.60, 11.30)	-

### A.5.2 Model Optimisation

We analysed the data in two frameworks: *Separate* and *Pooled*. In the *Separate* framework, the two separate SITAR models are fitted to each data set. In the *Pooled*, a single model is fitted to two data sets pooled.

#### A.5.3 1. Separate models

Initial models are fitted as described in the section A.4.1.2 through A.4.1.4. Multiple plausible models are explored for STDM and PSCG dataset in each sex separately.

##### A.5.3.1 BIC of plausible models in separate framework - boys

STDM:

```
stdm_boys_dfpower
```

```
, , log(Height), log(Age)
```

	a	a+b	a+c	a+b+c
4	-3555.0	-3560.4	-3549.7	3514.8
5	-3540.1	3514.6	-3507.7	-3518.9
6	-3513.2	3499.0	3547.3	-3549.9
7	-3536.2	-3544.1	-3545.1	-3560.0
8	-3552.7	3591.7	-3565.3	-3590.1

```
, , Height, log(Age)

      a      a+b     a+c    a+b+c
4 3545.4  3551.8 3536.4  3515.3
5 3499.9 -3500.3 3481.9 -3523.0
6 3494.1 -3454.1 3495.3 -3466.1
7 3492.3  3508.1 3503.1  3551.7
8 3471.6  3494.4 3555.8  3543.8
```

```
, , log(Height), Age

      a      a+b     a+c    a+b+c
4 -3541.1 -3551.5 -3529.7 -3540.5
5 -3563.6 -3551.5 -3532.8 -3552.2
6 -3540.7 -3537.7 -3524.8 -3546.8
7 -3548.2 -3570.4 -3575.9 -3577.0
8 -3532.4 -3542.3 -3553.8 -3585.1
```

```
, , Height, Age

      a      a+b     a+c    a+b+c
4  3519.6 3526.1  3491.1 3488.5
5 -3525.2 3529.7 -3486.1 3557.0
6  3509.8 3511.4  3490.9 3531.8
7  3551.4 3570.2 -3570.8 3497.0
8  3519.1 3558.2  3599.3 3581.4
```

### PSCG:

```
pscg_boys_dfpower
```

```
, , log(height), log(age)

      5      6      7      8
a    -18152.3 -17975.5 -18140.8 -18068.7
a+b  -18079.4 -17998.6 -17946.9 -18033.1
a+c  -18158.7 -17949.4 -17908.7 -17907.5
a+b+c -17972.7 -17919.4 -17940.5 -18023.9
```

```
, , height, log(age)

      5      6      7      8
a    -18219.1 18020.5 18161.0 18152.5
a+b  -18157.9 18035.9 18058.8 18051.3
a+c   18255.1 17987.6 18017.8 18068.0
a+b+c 18067.0 17976.1 18051.5 18078.1
```

```
, , log(height), age

      5      6      7      8
a    -18420.0 -18298.8 -18394.6 -18239.9
a+b  -18403.4 -18305.2 -18194.4 -18247.1
a+c  -18423.4 -18253.8 -18244.6 -18337.2
a+b+c -18350.4 -18265.8 -18219.5 -18308.6
```

```
, , height, age

          5          6          7          8
a      18385.1  18292.6  18473.6  18352.8
a+b    18403.6  18279.5  18254.8  18297.9
a+c    18415.7  18164.6  18208.5  18283.9
a+b+c  18168.5 -18202.5  18228.6  18337.1
```

### A.5.3.2 BIC of plausible models models in separate framework - girls

#### STDM:

```
stdm_girls_dfpower
```

```
, , log(Height), log(Age)

          a          a+b          a+c          a+b+c
4 -4079.2 -3974.9  3927.8 -3931.3
5 -3937.7  3944.8 -3935.2 -4105.8
6 -3952.7 -3955.4 -3952.5 -3969.6
7  3952.0 -3959.9 -3958.6  3977.5
8 -3961.0 -3982.7  4035.7 -4076.5
```

```
, , Height, log(Age)

          a          a+b          a+c          a+b+c
4  4065.2  3945.0  3881.6  3892.5
5 -3894.7  3898.5  3900.9  3908.9
6  3921.2  3910.9  3904.8 -3918.7
7  3917.2  3920.9  3910.1  3927.5
8  3924.5  3939.8  3987.5  3946.9
```

```
, , log(Height), Age

          a          a+b          a+c          a+b+c
4 -4061.3 -4012.7 -4054.5 -3970.9
5 -3976.1 -3971.0 -4208.6 -3965.2
6 -3971.6 -3967.0 -3976.8 -3994.1
7 -3978.3 -3972.0 -3986.6 -3983.3
8 -3991.8 -4096.6 -4003.5 -4012.5
```

```
, , Height, Age

          a          a+b          a+c          a+b+c
4  4004.1  3941.4  3915.0  3913.0
5  3918.0  3915.5  4122.0  4049.0
6  3928.8  3930.4  3928.5  3952.0
7  3935.4  3934.2  3946.9  3953.1
8  3940.2 -4149.2  3947.4 -3974.0
```

#### PSCG:

```
psc_girls_dfpower
```

```
, , log(height), log(age)
```



	4	5	6	7
a	-15207.1	-14502.8	-14612.9	-14516.3
a+b	-14709.2	-14484.6	-14516.4	-14518.6
a+c	-14555.7	-14498.2	-14517.8	-14524.5
a+b+c	-14499.5	-14506.9	-14525.1	-14536.0

, , height, log(age)

	4	5	6	7
a	15119.0	14368.6	14503.6	14375.1
a+b	14614.6	14374.4	14382.9	14381.4
a+c	14424.6	-14369.0	14396.9	14389.0
a+b+c	14354.2	14369.5	14388.2	14390.3

, , log(height), age

	4	5	6	7
a	-15744.3	-14583.3	-14670.0	-14606.5
a+b	-15031.4	-14591.6	-14597.0	-14607.3
a+c	-15132.4	-14576.1	-14650.8	-14598.1
a+b+c	-15109.9	-14569.6	-14604.6	-14608.4

, , height, age

	4	5	6	7
a	15457.7	14412.8	14541.6	14433.8
a+b	14897.0	14423.3	14434.8	-14446.6
a+c	14660.4	14426.4	14455.7	-14439.6
a+b+c	14554.9	14432.9	-14446.3	14460.5

### A.5.3.3 Model selection

#### A.5.3.4 Optimal model selected in boys

##### A.5.3.4.1 STDM

SITAR nonlinear mixed-effects model fit by maximum likelihood

```
Call: sitar(x = log(Age), y = Height, id = ID, data = ht_B_2[ht_B_2$htok ==
  TRUE, ], df = 6, fixed = "a+b")
      AIC      BIC    logLik
3385.26 3454.136 -1677.63
```

Random effects:

Formula: list(a ~ 1, b ~ 1, c ~ 1)

Level: id

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
a	8.10076021	a b
b	0.09683062	0.470
c	0.22690417	0.649 0.513

Residual 0.66019327

Fixed effects: s1 + s2 + s3 + s4 + s5 + s6 + a + b ~ 1

	Value	Std.Error	DF	t-value	p-value
s1	37.61776	0.8871846	514	42.40128	0
s2	45.88729	1.0074689	514	45.54711	0

```

s3 51.31230 0.9956636 514 51.53578 0
s4 66.31601 0.9392346 514 70.60643 0
s5 88.35162 1.7212197 514 51.33082 0
s6 63.94581 0.8662506 514 73.81907 0
a 92.97789 0.9770983 514 95.15714 0
b -0.10277 0.0086060 514 -11.94127 0

```

Correlation:

```

      s1      s2      s3      s4      s5      s6      a
s2 0.855
s3 0.911 0.875
s4 0.698 0.809 0.816
s5 0.864 0.898 0.881 0.783
s6 0.569 0.717 0.694 0.923 0.654
a -0.774 -0.825 -0.756 -0.631 -0.804 -0.561
b 0.198 -0.003 0.198 0.044 0.062 -0.060 0.135

```

Standardized Within-Group Residuals:

```

      Min      Q1      Med      Q3      Max
-2.9051540 -0.3569736 -0.0019481 0.3420396 2.6962918

```

Number of Observations: 729

Number of Groups: 208

#### A.5.3.4.2 PSCG

SITAR nonlinear mixed-effects model fit by maximum likelihood

Call:

```

sitar(x = log(age), y = height, id = ID, data = P_ht_B_2[P_ht_B_2$htok ==
TRUE, ], df = 6, fixed = "a+c")
      AIC      BIC    logLik
17891.56 17987.59 -8930.78

```

Random effects:

Formula: list(a ~ 1, b ~ 1, c ~ 1)

Level: id

Structure: General positive-definite, Log-Cholesky parametrization

```

      StdDev      Corr
a      6.78332199 a      b
b      0.08801212 0.467
c      0.14504449 0.509 0.550

```

Residual 0.71135506

Fixed effects: s1 + s2 + s3 + s4 + s5 + s6 + a + c ~ 1

```

      Value Std.Error  DF  t-value p-value
s1 34.21639 0.3431685 3651 99.70728 0
s2 41.47956 0.4575220 3651 90.66134 0
s3 47.88522 0.4961811 3651 96.50753 0
s4 64.88246 0.6898710 3651 94.05013 0
s5 87.02190 0.8174766 3651 106.45185 0
s6 64.20619 0.5308078 3651 120.95940 0
a 99.74442 0.4950880 3651 201.46805 0
c 0.11133 0.0125456 3651 8.87443 0

```

Correlation:

```

      s1      s2      s3      s4      s5      s6      a
s2 0.950

```

```

s3 0.955 0.936
s4 0.879 0.918 0.935
s5 0.945 0.967 0.942 0.927
s6 0.828 0.879 0.869 0.922 0.912
a -0.875 -0.892 -0.868 -0.839 -0.882 -0.809
c -0.792 -0.834 -0.860 -0.924 -0.833 -0.860 0.812

```

```

Standardized Within-Group Residuals:
      Min          Q1          Med          Q3          Max
-3.03224009 -0.45460971  0.02426882  0.47103476  3.07987956

```

```

Number of Observations: 4455
Number of Groups: 797

```

### A.5.3.5 Optimal model selected in girls

#### A.5.3.5.1 STDM

```

SITAR nonlinear mixed-effects model fit by maximum likelihood
Call: sitar(x = log(Age), y = Height, id = ID, data = .data., df = 5,
  fixed = "a+b")
      AIC      BIC    logLik
3849.882 3916.577 -1910.941

```

```

Random effects:
Formula: list(a ~ 1, b ~ 1, c ~ 1)
Level: id
Structure: General positive-definite, Log-Cholesky parametrization
      StdDev   Corr
a      7.5295989 a      b
b      0.1189044 0.309
c      0.2376369 0.558 0.319
Residual 0.5793913

```

```

Fixed effects: s1 + s2 + s3 + s4 + s5 + a + b ~ 1

```

```

      Value Std.Error DF t-value p-value
s1 39.07891 1.0682417 607 36.58246 0.0000
s2 58.26626 1.2332177 607 47.24734 0.0000
s3 53.80469 0.9688206 607 55.53628 0.0000
s4 81.62245 2.1356068 607 38.21979 0.0000
s5 49.09669 0.8121735 607 60.45099 0.0000
a 91.23600 1.2031645 607 75.83003 0.0000
b -0.03236 0.0111040 607 -2.91425 0.0037

```

```

Correlation:
      s1      s2      s3      s4      s5      a
s2 0.919
s3 0.778 0.832
s4 0.902 0.903 0.869
s5 0.698 0.794 0.921 0.766
a -0.833 -0.795 -0.796 -0.907 -0.700
b 0.060 0.097 -0.135 -0.126 -0.113 0.263

```

```

Standardized Within-Group Residuals:
      Min          Q1          Med          Q3          Max
-3.333222407 -0.335210960 -0.003927598  0.363667253  2.740354354

```

Number of Observations: 866  
Number of Groups: 253

### A.5.3.5.2 PSCG

SITAR nonlinear mixed-effects model fit by maximum likelihood

Call:  
sitar(x = log(age), y = height, id = ID, data = P\_girls\_clean2[P\_girls\_clean2\$htok ==  
TRUE, ], df = 4)  
AIC BIC logLik  
14267.19 14354.16 -7119.595

Random effects:

Formula: list(a ~ 1, b ~ 1, c ~ 1)  
Level: id  
Structure: General positive-definite, Log-Cholesky parametrization  
StdDev Corr  
a 6.00106604 a b  
b 0.09255551 0.443  
c 0.13943815 0.386 0.468  
Residual 0.63613525

Fixed effects: s1 + s2 + s3 + s4 + a + b + c ~ 1

	Value	Std.Error	DF	t-value	p-value
s1	32.12255	0.4644654	3006	69.16027	0
s2	44.77673	0.6229220	3006	71.88176	0
s3	66.06426	1.0062411	3006	65.65451	0
s4	42.51310	0.5316496	3006	79.96452	0
a	104.81674	0.6986175	3006	150.03451	0
b	-0.06241	0.0050399	3006	-12.38373	0
c	0.31502	0.0192112	3006	16.39757	0

Correlation:

	s1	s2	s3	s4	a	b
s1	0.956					
s2	0.968	0.971				
s3	0.936	0.987	0.956			
s4	-0.906	-0.912	-0.936	-0.903		
a	-0.507	-0.556	-0.581	-0.569	0.683	
b	-0.884	-0.937	-0.909	-0.927	0.900	0.663

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-3.11848674	-0.41565094	0.01146134	0.41865869	2.99682778

Number of Observations: 3685  
Number of Groups: 673

## A.5.4 2. Pooled models

### A.5.4.1 Model optimisation

Multiple plausible models are explored by pooling STDM and PSCG dataset in each sex. All the models in the pooled framework is fitted with all the fixed effects. Hence, the combinations

of possible models with varying dfs and transformations of the age and height scale are only explored.

#### A.5.4.2 Boys

```
boys_pooled_df
```

```
, , log(age)
```

```
  log(height)  height
4    -21599.3  21774.0
5    -21469.4 -21518.1
6    -21412.8 -21412.9
7    -21493.9  21460.3
8    -21564.7  21593.9
```

```
, , age
```

```
  log(height)  height
4    -22079.0 -21801.0
5    -21864.3  21651.0
6    -21760.8  21602.0
7    -21716.2  21688.0
8    -21831.2  21909.4
```

#### A.5.4.3 Girls

```
girls_pooled
```

```
, , log(age)
```

```
  log(height)  height
4    -18601.0 18295.6
5    -18500.0 18290.2
6    -18495.5 18307.8
7    -18504.2 18311.6
8    -18517.4 18829.4
```

```
, , age
```

```
  log(height)  height
4    -19262.8 -18518.9
5    -18586.3 -18374.5
6    -18619.6  18372.6
7    -19010.1  18483.2
8    -18778.3 -18667.0
```

#### A.5.4.4 Optimal models

#### A.5.4.5 Boys

SITAR nonlinear mixed-effects model fit by maximum likelihood

Call:

```
sitar(x = log(age), y = height, id = ID, data = bigdata_boys[bigdata_boys$htok ==
  TRUE, ], df = 6, a.formula = ~flag, b.formula = ~flag, c.formula = ~flag)
```

AIC      BIC      logLik  
 21288.45 21412.94 -10625.23

Random effects:

Formula: list(a ~ 1, b ~ 1, c ~ 1)

Level: id

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
a	7.07402883	a      b
b	0.08734763	0.450
c	0.15526079	0.529 0.491
Residual	0.70698564	

Fixed effects: s1 + s2 + s3 + s4 + s5 + s6 + a.flagSTDM + a + b.flagSTDM + b + c.flagSTDM + c ~ 1

	Value	Std.Error	DF	t-value	p-value
s1	33.36173	0.4181310	4162	79.78775	0.0000
s2	39.89504	0.6442640	4162	61.92342	0.0000
s3	46.35919	0.5796801	4162	79.97375	0.0000
s4	62.66738	0.9614141	4162	65.18250	0.0000
s5	84.26544	1.1031444	4162	76.38660	0.0000
s6	61.87838	0.7917075	4162	78.15813	0.0000
a.flagSTDM	-4.92872	0.6356249	4162	-7.75413	0.0000
a	101.60360	0.8386468	4162	121.15184	0.0000
b.flagSTDM	0.00971	0.0093999	4162	1.03260	0.3018
b	0.00664	0.0064440	4162	1.02964	0.3032
c.flagSTDM	-0.09856	0.0170977	4162	-5.76436	0.0000
c	0.18259	0.0220639	4162	8.27529	0.0000

Correlation:

	s1	s2	s3	s4	s5	s6	a.STDM	a	b.STDM
s2	0.963								
s3	0.972	0.944							
s4	0.931	0.967	0.946						
s5	0.962	0.984	0.953	0.971					
s6	0.903	0.952	0.905	0.967	0.964				
a.flagSTDM	-0.020	-0.012	-0.023	-0.018	-0.023	-0.022			
a	-0.904	-0.950	-0.888	-0.933	-0.941	-0.931	-0.022		
b.flagSTDM	-0.074	-0.070	-0.070	-0.068	-0.078	-0.077	0.589	0.059	
b	-0.653	-0.763	-0.631	-0.763	-0.745	-0.793	-0.014	0.842	0.032
c.flagSTDM	0.008	0.009	0.008	0.005	0.001	0.007	0.511	-0.018	0.580
c	-0.877	-0.933	-0.887	-0.960	-0.929	-0.945	0.003	0.948	0.069
	b	c.STDM							

s2  
 s3  
 s4  
 s5  
 s6  
 a.flagSTDM  
 a  
 b.flagSTDM  
 b  
 c.flagSTDM 0.002  
 c 0.843 0.003

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-3.05589880	-0.44420946	0.02340478	0.45910534	2.91864061

Number of Observations: 5178

Number of Groups: 1005

#### A.5.4.6 Girls

SITAR nonlinear mixed-effects model fit by maximum likelihood

Call:

```
sitar(x = log(age), y = height, id = ID, data = bigdata[bigdata$htok ==
  TRUE, ], df = 5, a.formula = ~flag, b.formula = ~flag, c.formula = ~flag)
```

	AIC	BIC	logLik
	18174.57	18290.18	-9069.287

Random effects:

Formula: list(a ~ 1, b ~ 1, c ~ 1)

Level: id

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
a	6.5559583	a b
b	0.0921528	0.370
c	0.1627782	0.474 0.345
Residual	0.6266788	

Fixed effects: s1 + s2 + s3 + s4 + s5 + a.flagSTDM + a + b.flagSTDM + b + c.flagSTDM + c ~ 1

	Value	Std.Error	DF	t-value	p-value
s1	34.10254	0.8496389	3612	40.13769	0.0000
s2	49.68471	1.2951733	3612	38.36144	0.0000
s3	47.53833	0.9405313	3612	50.54412	0.0000
s4	69.77077	1.5705228	3612	44.42519	0.0000
s5	44.45684	0.7870713	3612	56.48388	0.0000
a.flagSTDM	-3.76517	0.5470406	3612	-6.88280	0.0000
a	102.02282	1.0889803	3612	93.68656	0.0000
b.flagSTDM	0.04604	0.0087177	3612	5.28095	0.0000
b	0.00984	0.0039983	3612	2.46060	0.0139
c.flagSTDM	-0.04858	0.0177331	3612	-2.73939	0.0062
c	0.15815	0.0326273	3612	4.84732	0.0000

Correlation:

	s1	s2	s3	s4	s5	a.STDM a	b.STDM b
s2	0.992						
s3	0.972	0.981					
s4	0.983	0.985	0.984				
s5	0.964	0.976	0.983	0.976			
a.flagSTDM	0.007	0.006	0.004	0.006	0.000		
a	-0.964	-0.967	-0.966	-0.976	-0.954	-0.030	
b.flagSTDM	-0.005	-0.004	-0.006	-0.005	-0.014	0.501	0.006
b	-0.161	-0.171	-0.234	-0.240	-0.225	0.006	0.306 -0.008
c.flagSTDM	0.011	0.019	0.022	0.011	0.024	0.410	-0.019 0.347 -0.004
c	-0.959	-0.974	-0.965	-0.955	-0.954	-0.011	0.959 0.002 0.236
	c.STDM						

s2  
s3  
s4

```
s5
a.flagSTDM
a
b.flagSTDM
b
c.flagSTDM
c          -0.018
```

```
Standardized Within-Group Residuals:
              Min              Q1              Med              Q3              Max
-2.9479917558 -0.3917843270 -0.0005056974  0.4002058845  2.8478345017
```

```
Number of Observations: 4548
Number of Groups: 926
```

## A.5.5 Comparison of separate vs pooled models

### A.5.5.1 Boys

BICs of the models are:

	df	BIC
STDM separate	15	3454.136
PSCG separate	15	17987.586
Pooled	19	21412.942

The difference between the sum of the BICs of the two separate models and the BIC of the pooled model is:

```
[1] 28.77998
```

### A.5.5.2

	df	BIC
STDM_separate	14	3916.577
PSCG_separate	14	14354.157
Pooled	18	18290.179

The difference between the sum of the BICs of the two separate models and the BIC of the pooled model is:

```
[1] -19.44472
```

## A.5.6 Covariate analysis

### A.5.6.1 HbA1c modelling

There is no significant sex effect ( $P = 0.1$ ).

```
hb_gb <- STDM_nest %>%
  unnest(data) %>%
  dplyr::select(ID, Age, Gender, HbA1C_pc) %>%
  drop_na() %>%
  # mutate(sex = as.numeric(Gender) ) %>%
```



```
sitar(Age, log(HbA1C_pc), ID, data = ., df = 3, random = "a",
      a.formula = ~Gender)
summary(hb_gb)
```

```
SITAR nonlinear mixed-effects model fit by maximum likelihood
Call: sitar(x = Age, y = log(HbA1C_pc), id = ID, data = ., df = 3,
           random = "a", a.formula = ~Gender)
      AIC      BIC  logLik
-755.9331 -719.1837 384.9666
```

```
Random effects:
Formula: a ~ 1 | id
          a Residual
StdDev: 0.1362259 0.1527894
```

```
Fixed effects: s1 + s2 + s3 + a.GenderGirls + a ~ 1
              Value Std.Error DF t-value p-value
s1            0.1598948 0.02546180 956  6.27979 0.0000
s2            0.0338851 0.06732539 956  0.50330 0.6149
s3           -0.0047260 0.02791114 956 -0.16932 0.8656
a.GenderGirls 0.0257322 0.01591955 956  1.61639 0.1063
a              2.2452662 0.02940491 956 76.35686 0.0000
```

```
Correlation:
          s1      s2      s3      a.GndG
s2            0.321
s3            0.101  0.518
a.GenderGirls 0.043  0.025  0.027
a            -0.487 -0.939 -0.399 -0.032
```

```
Standardized Within-Group Residuals:
          Min           Q1           Med           Q3           Max
-3.267786885 -0.586210264 -0.009067741  0.562193878  3.345783891
```

```
Number of Observations: 1408
Number of Groups: 448
```

HbA1c model without sex effect:

```
hb_gb2 <- STDM_nest %>%
  unnest(data) %>%
  dplyr::select(ID, Age, Gender, HbA1C_pc) %>%
  drop_na() %>%
  # mutate(sex = as.numeric(Gender) ) %>%
sitar(Age, log(HbA1C_pc), ID, data = ., df = 3, random = "a")
summary(hb_gb2)
```

```
SITAR nonlinear mixed-effects model fit by maximum likelihood
Call: sitar(x = Age, y = log(HbA1C_pc), id = ID, data = ., df = 3,
           random = "a")
      AIC      BIC  logLik
-755.3148 -723.8152 383.6574
```

```
Random effects:
Formula: a ~ 1 | id
```

a Residual  
StdDev: 0.136591 0.1528601

Fixed effects: s1 + s2 + s3 + a ~ 1

	Value	Std.Error	DF	t-value	p-value
s1	0.1581809	0.02546163	957	6.21252	0.0000
s2	0.0312504	0.06734959	957	0.46400	0.6428
s3	-0.0058957	0.02792365	957	-0.21114	0.8328
a	2.2467471	0.02941399	957	76.38362	0.0000

Correlation:

	s1	s2	s3
s1			
s2	0.321		
s3	0.101	0.517	
a	-0.486	-0.939	-0.399

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-3.25305200	-0.57567834	-0.01735663	0.55909508	3.35774831

Number of Observations: 1408  
Number of Groups: 448

### A.5.6.2 Covariate analysis using SITAR

Summary of the covariates

```
df <- NA
df = c(6, 5)
fixed <- NA
fixed = c("a+b", "a+b")
# for covariate analysis, all three fixed effects will be
# used anyways.

data1 <- stdm_par_ln %>%
  dplyr::select(ID, Gender, meanDdur_yr, AgeDiagnosis_new,
    a.hb, mph_zz, Height, Age, htok) %>%
  filter(AgeDiagnosis_new > 0) %>%
  drop_na() %>%
  nest_by(Gender)

# bind df and fixed columns with the df
data2 <- data1
data2$df <- df
data2$fixed <- fixed

# without birth weight

data2 %>%
  rowwise() %>%
  mutate(model = list(sitar(log(Age), Height, ID, data[data$htok ==
    TRUE, ], df = df, fixed = fixed, a.formula = ~(mph_zz +
    AgeDiagnosis_new + meanDdur_yr + a.hb), b.formula = ~(mph_zz +
    AgeDiagnosis_new + meanDdur_yr + a.hb), c.formula = ~(mph_zz +
```

```
AgeDiagnosis_new + meanDdur_yr + a.hb)))) %>%
identity -> stdm_cov
```

Covariates in boys when associated with size (cm), timing (fraction), intensity (fraction).

	Value	Std.Error	DF	t-value	p-value
a.mph_zz	3.33	0.57	464	5.86	0.00
a.AgeDiagnosis_new	1.15	0.20	464	5.86	0.00
a.meanDdur_yr	0.43	0.21	464	2.06	0.04
a.a.hb	-7.90	4.76	464	-1.66	0.10
a	97.62	4.80	464	20.34	0.00
b.mph_zz	0.02	0.01	464	2.48	0.01
b.AgeDiagnosis_new	0.03	0.00	464	7.95	0.00
b.meanDdur_yr	0.03	0.00	464	8.48	0.00
b.a.hb	-0.05	0.07	464	-0.80	0.43
b	-0.13	0.02	464	-6.96	0.00
c.mph_zz	0.06	0.02	464	2.96	0.00
c.AgeDiagnosis_new	0.05	0.01	464	5.79	0.00
c.meanDdur_yr	0.04	0.01	464	4.28	0.00
c.a.hb	-0.17	0.17	464	-0.99	0.32
c	-0.09	0.06	464	-1.52	0.13

Regression coefficients of each covariates in girls when associated with size (cm), timing (fraction), intensity (fraction).

	Value	Std.Error	DF	t-value	p-value
a.mph_zz	3.43	0.49	563	6.96	0.00
a.AgeDiagnosis_new	1.12	0.17	563	6.59	0.00
a.meanDdur_yr	0.79	0.20	563	3.90	0.00
a.a.hb	-15.04	4.53	563	-3.32	0.00
a	103.41	3.09	563	33.47	0.00
b.mph_zz	0.00	0.01	563	0.45	0.65
b.AgeDiagnosis_new	0.03	0.00	563	9.62	0.00
b.meanDdur_yr	0.03	0.00	563	9.87	0.00
b.a.hb	-0.15	0.07	563	-2.23	0.03
b	-0.02	0.01	563	-2.20	0.03
c.mph_zz	0.03	0.02	563	1.58	0.11
c.AgeDiagnosis_new	0.02	0.01	563	2.12	0.03
c.meanDdur_yr	0.02	0.01	563	2.19	0.03
c.a.hb	-0.39	0.23	563	-1.71	0.09
c	0.40	0.10	563	4.06	0.00

## Bibliography

- Altman, P. L. and Dittmer, D. S. (1962). *Growth including reproduction and morphological development*. Washington : Federation of American Societies for Experimental Biology, Author.
- Areekal, S. A., Goel, P., Khadilkar, A., Khadilkar, V., and Cole, T. J. (2022). Assessment of height growth in Indian children using growth centiles and growth curves. *Annals of Human Biology*, 49(5-6):228–235.
- Areekal, S. A., Khadilkar, A., Kajale, N., Kinare, A. S., and Goel, P. (2023). Two novel models

- evaluating the determinants of resting metabolic rate in Indian children. *Human Biology and Public Health*, 3.
- Cole, T. (2022). *sitar: Super Imposition by Translation and Rotation Growth Curve Analysis*. R package version 1.2.0.9000.
- Cole, T. J., Donaldson, M. D. C., and Ben-Shlomo, Y. (2010). SITAR—a useful instrument for growth curve analysis. *Int. J. Epidemiol.*, 39(6):1558.
- Khadilkar, V., Khadilkar, A., Arya, A., Ekbote, V., Kajale, N., Parthasarathy, L., Patwardhan, V., Phanse, S., and Chiplonkar, S. (2019). Height Velocity Percentiles in Indian Children Aged 5-17 Years. *Indian Pediatr.*, 56(1):23–28.
- Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(3):507–554.
- Wang, Z. (2012). High ratio of resting energy expenditure to body mass in childhood and adolescence: A mechanistic model. *American Journal of Human Biology*, 24(4):460–467.

# Appendix B

## List of publications

1. Areekal, S. A., Khadilkar, A., Kajale, N., Kinare, A. S., & Goel, P. (2023). *Two novel models evaluating the determinants of resting metabolic rate in Indian children. Human Biology and Public Health, 3.* <https://doi.org/10.52905/hbph2022.3.55>

An earlier preprint version of the above article is published on the Research Square preprint platform. This preprint now references the final published version of the article.

Areekal, S. A., Khadilkar, A., Ekbote, V., Kajale, N., Kinare, A. S., and Goel, P. (2021). *Two novel models evaluating the determinants of resting metabolic rate in Indian children. Research Square, PREPRINT (Version 1),* <https://doi.org/10.21203/rs.3.rs-196719/v1> )

2. Areekal, S. A., Goel, P., Khadilkar, A., Khadilkar, V., & Cole, T. J. (2022). *Assessment of height growth in Indian children using growth centiles and growth curves. Annals of human biology, 49(5-6), 228–235.* <https://doi.org/10.1080/03014460.2022.2107238>

An author manuscript of the above article is available at the institute repository of University College London: <https://discovery.ucl.ac.uk/id/eprint/10156741/>

3. Areekal, S. A., Khadilkar, A., Goel, P., & Cole, T. J. (2023). *Longitudinal Height Growth in Children and Adolescents with Type-1 Diabetes Mellitus Compared to Controls in Pune, India. Pediatric Diabetes, 8 pages.* <https://doi.org/10.1155/2023/8813031>

### B.1 Copyrights and Licence

Copyrights and licence information for the three published articles:

1. The article Areekal, S. A., Khadilkar, A., Kajale, N., Kinare, A. S., & Goel, P. (2023). *Two novel models evaluating the determinants of resting metabolic rate in Indian children. Human Biology and Public Health, 3.* <https://doi.org/10.52905/hbph2022.3.55> has been published under a Creative Commons Attribution 4.0 International License (see Figure B.1);
2. The article Areekal, S. A., Khadilkar, A., Goel, P., & Cole, T. J. (2023). *Longitudinal Height Growth in Children and Adolescents with Type-1 Diabetes Mellitus Compared to*

*Controls in Pune, India. Pediatric Diabetes*, 8 pages. <https://doi.org/10.1155/2023/8813031> has been published under Creative Commons Attribution License (see Figure B.2)

3. The article Areekal, S. A., Goel, P., Khadilkar, A., Khadilkar, V., & Cole, T. J. (2022). *Assessment of height growth in Indian children using growth centiles and growth curves. Annals of human biology*, 49(5-6), 228–235. <https://doi.org/10.1080/03014460.2022.2107238>, Copyright © 2022, was reprinted by permission of Informa UK Limited, trading as Taylor & Taylor & Francis Group, <http://www.tandfonline.com> (see Figure B.3),

License

Copyright (c) 2023 Sandra Aravind Areekal,  
Anuradha Khadilkar, Neha Kajale, Arun S.  
Kinare, Pranay Goel



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

The journal is published under the Creative Commons Attribution 4.0 International License. Authors are at liberty to share their contributions, provided they give proper attribution to the original publication, without asking prior permission from the editors.

More information about this license is available at <https://creativecommons.org/licenses/by/4.0/>.

Figure B.1: Copyrights information for including the article Areekal et al. (2023) in chapter 3.

Research Article

# Longitudinal Height Growth in Children and Adolescents with Type-1 Diabetes Mellitus Compared to Controls in Pune, India

Sandra Aravind Areekal <sup>1</sup>, Anuradha Khadilkar <sup>2</sup>, Pranay Goel <sup>1</sup> and Tim J. Cole <sup>3</sup>

<sup>1</sup>Department of Biology, Indian Institute of Science Education and Research, Pune 411008, India

<sup>2</sup>Department of Growth and Endocrinology, Hirabai Cowasji Jehangir Medical Research Institute, Pune 411001, India

<sup>3</sup>University College London Great Ormond Street Institute of Child Health, London WC1N 1EH, UK

Correspondence should be addressed to Tim J. Cole; tim.cole@ucl.ac.uk

Received 20 January 2023; Revised 3 June 2023; Accepted 9 June 2023; Published 1 July 2023

Academic Editor: Kalie Tommerdahl

Copyright © 2023 Sandra Aravind Areekal et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Figure B.2: Copyrights information for including Areekal et al. (2023) in chapter 5.



Our Ref: iahb/03221684

6/30/2023

Dear Requester,

Thank you for your correspondence requesting permission to reproduce content from a Taylor & Francis Group journal content in your thesis to be posted on your university's repository.

We will be pleased to grant free permission on the condition that your acknowledgement must be included showing article title, author, full Journal title, and © copyright # [year], reprinted by permission of Informa UK Limited, trading as Taylor & Francis Group, <http://www.tandfonline.com>

**This permission does not cover any third party copyrighted work which may appear in the article by permission. Please ensure you have checked all original source details for the rights holder and if need apply for permission from the original rightsholder.**

Please note that this license **does not allow you to post our content on any other third-party websites.**

**Please note permission does not provide access to our article**, if you are affiliated to an institution and your institution holds a subscription to the content you are requesting you will be able to view the article free of charge, if your institution does not hold a subscription or you are not affiliated to an institution that has a subscription then you will need to purchase this for your own personal use as we do not provide our articles free of charge for research.

Thank you for your interest in our Journal.

With best wishes,

Taylor & Francis Journal Permissions

Web: [www.tandfonline.com](http://www.tandfonline.com)

4 Park Square, Milton Park, Abingdon, OX14 4RN

( +44 (0)20 8052 0600

**Disclaimer: T&F publish Open Access articles in our subscription priced journals, please check if the article you are interested in is an OA article and if so, which license was it published under.**

 Before printing, think about the environment.

ref:\_00D0Y35Iji.\_5007TPsS0b:ref

Figure B.3: Copyrights information for including Areekal et al. (2022) in chapter 4.

## Appendix C

# Ethics statement

The ethics committee at IISER Pune approved the study (Ref: IECHR/Admin/2019/002 dated 28th February 2019 and Ref: IECHR/Admin/2021/001 dated 17th February 2021).



# Appendix D

## Software

Figure D.1 shows a snapshot of the software described in Chapter 3.

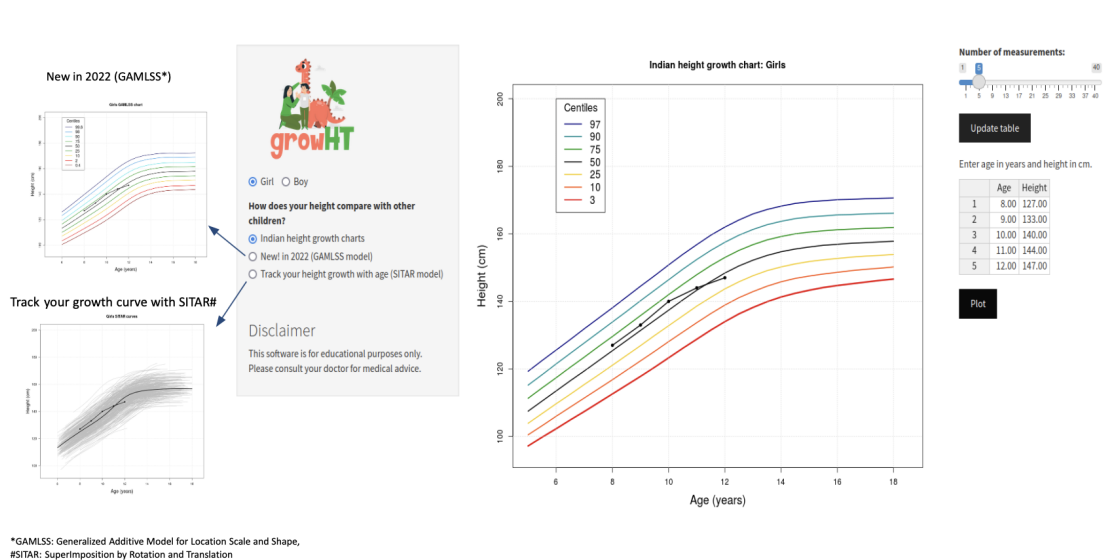


Figure D.1: A height growth monitoring app for Indian children named growHT (<https://digimed.acads.iiserpune.ac.in/growth-charts>).