

TIINA TEIVAANMÄKI

Child Growth Stunting and Development in Malawi





TIINA TEIVAANMÄKI

Child Growth Stunting
and Development in Malawi



ACADEMIC DISSERTATION

To be presented, with the permission of
the Faculty Council of the Faculty of Medicine and Life Sciences
of the University of Tampere,
for public discussion in the auditorium F114
of the Arvo building,
Arvo Ylpön katu 34, Tampere,
on 23 March 2018, at 12 o'clock.

UNIVERSITY OF TAMPERE

TIINA TEIVAANMÄKI

Child Growth Stunting
and Development in Malawi

Acta Universitatis Tamperensis 2357
Tampere University Press
Tampere 2018

ACADEMIC DISSERTATION

University of Tampere, Faculty of Medicine and Life Sciences
The National Graduate School of Clinical Investigation
Finland

Supervised by

Professor Per Ashorn
University of Tampere
Finland
Professor Yin Bun Cheung
Duke-NUS Medical School
Singapore

Reviewed by

Docent Eero Kajantie
University of Helsinki
Finland
Docent Helena Lapinleimu
University of Turku
Finland

The originality of this thesis has been checked using the Turnitin OriginalityCheck service in accordance with the quality management system of the University of Tampere.

Copyright ©2018 Tampere University Press and the author

Cover design by
Mikko Reinikka

Acta Universitatis Tamperensis 2357
ISBN 978-952-03-0672-4 (print)
ISSN-L 1455-1616
ISSN 1455-1616

Acta Electronica Universitatis Tamperensis 1863
ISBN 978-952-03-0673-1 (pdf)
ISSN 1456-954X
<http://tampub.uta.fi>

Suomen Yliopistopaino Oy – Juvenes Print
Tampere 2018



ABSTRACT

Stunting affects 159 million children under five years of age worldwide. Approximately 36% of the world's stunted children live in sub-Saharan Africa. Stunting is a measure of chronic undernutrition, and it is associated with increased mortality, morbidity and developmental problems from early childhood onwards.

This thesis is based on four studies. The aims of this thesis were to describe the timing of stunting and possible growth recovery in a rural low-income population, to examine the association between early development and later school achievements, and to examine the associations between childhood growth and later health and developmental outcomes, such as cognitive capacity and depressive symptoms in adolescence.

In this Lungwena Child Survival Study (LCSS), 813 participants were followed from the fetal period until the age of 15. Their growth was regularly monitored. In addition, a developmental assessment inventory was conducted at five years of age, mathematics test at 12 years of age and pubertal stage, cognitive capacity and depressive symptoms at 15 years of age. Cognitive capacity was assessed with Raven's Coloured Matrices score, reaction time (RT) and mathematics test, and depressive symptoms with Short Mood and Feelings Questionnaire (SMFQ). The association between development at 5 years of age and mathematics skills at 12 years, and the associations between growth and cognitive capacity and depressive symptoms at 15 years were assessed with regression models. Potential confounders were added in the models and multiple imputation (MI) was used when appropriate.

Majority of the children (80%) were stunted (height-for-age < -2 SD) at two years of age, but the prevalence of stunting declined to 37% at 15 years. Most of the children who were stunted at two years, became non-stunted during the follow-up period. Only 9% of boys and 20% of girls reached advanced puberty by the end of the follow-up period at 15 years of age. Higher developmental summary score at 5 years predicted higher percentage of correctly answered mathematics questions at 12 years. Height gain between 24 months and 15 years was statistically significantly associated with Raven's Coloured Matrices score, but not reaction time (RT) nor mathematics test. The association weakened when school education was added in the model as a potential confounder. The mean score of SMFQ in this cohort was

15. The traditional cut-off for significant depressive symptoms is 11. About 90% (95% CI 87%–92%) of the participants scored 11 or more points. Birth weight, growth, gender or pubertal maturity were not associated with SMFQ score in the primary analyses, in which the missing data was handled with multiple imputation (MI). In the sensitivity analyses with the completely observed data, birth weight was negatively and statistically significantly associated with SMFQ score.

In conclusion, I suggest that catch-up growth is possible throughout the period from birth until 15 years of age. Later pubertal development may provide an opportunity for further growth in adolescence. There is an association between early development and later school achievements, and height gain in childhood has an association with cognitive capacity later on. Part of the association between growth and intelligence seems to be mediated by schooling. The prevalence of depressive symptoms in this cohort is high, and low birth weight may contribute to depressive symptoms in adolescence.

TIIVISTELMÄ

Maailmassa 159 miljoonaa alle viisivuotiasta lasta kasvaa aliravitsemuksen vuoksi kasvuhäiriöisenä. Noin 36% näistä lapsista asuu Saharan eteläpuolisessa Afrikassa. Odotuspituudesta jääminen johtuu kroonisesta aliravitsemuksesta. Sen on osoitettu olevan yhteydessä kuolleisuuteen, sairastuvuuteen ja varhaislapsuuden kehityksen ongelmiin.

Tämä väitöskirja perustuu neljään osajulkaisuun. Väitöskirjan tavoitteina oli kuvata ja tutkia odotuspituudesta jäämisen ja mahdollisen saavutuskasvun ajoitusta matalan tulotason maissa. Lisäksi tavoitteina oli tutkia varhaisen kehityksen ja koulumenestyksen yhteyttä sekä lapsuuden kasvun ja nuoruusajan älykkyyden ja masennusoireiden yhteyttä.

Tässä Lungwena Child Survival Studyssa (LCSS) 831 lasta seurattiin sikiöajalta 15-vuotiaiksi. Tutkittavien lasten kasvu mitattiin säännöllisesti. Lisäksi lasten kehitystä tutkittiin 5 vuoden iässä, matematiikan taitoja 12 vuoden iässä ja puberteettikehitystä, älykkyyttä ja masennusoireita 15 vuoden iässä. Älykkyyttä tutkittiin Raven's Coloured Matrices-testillä reaktioajalla ja matematiikan kokeella, ja masennusoireita Short Mood and Feelings-kyselyllä (SMFQ). 5-vuotiaiden lasten kehityksen ja 12-vuotiaiden matematiikan taitojen yhteyttä tutkittiin regressioanalyysillä. Mahdolliset sekoittavat tekijät lisättiin malleihin, ja moni-imputointia (MI) käytettiin tarvittaessa.

Kaksivuotiaista kohortin jäsenistä 80%:n ja 15-vuotiaista 37%:n suhteellinen pituus oli vähemmän kuin -2 keskihajontayksikköä (SD). Suurin osa lapsista, jotka olivat kaksivuotiaina odotuspituuttaan lyhyempiä, ylittivät seuranta-aikana -2 SD:n rajan. Vain 9% pojista ja 20% tytöistä olivat 15-vuotiaana pitkällä puberteettikehityksessään. Suurempi kokonaiskehitystä kuvaavan summamuuttujan arvo 5-vuotiaana ennusti parempaa tulosta matematiikan kokeessa 12-vuotiaana. Nopeampi pituuskasvu 24 kuukauden ja 15 vuoden iän välillä oli tilastollisesti merkitsevästi yhteydessä Raven's Coloured Matrices-tulokseen, mutta ei reaktioaikaan tai matematiikan kokeen tulokseen. Tämä yhteys heikkeni kun käytyjen kouluvuosien määrä lisättiin regressiomalliin mahdollisena sekoittavana tekijänä. Keskimääräinen pistemäärä SMFQ:ssa oli 15. Tässä kyselyssä 11 pistettä tai enemmän on kliinisesti merkittävien masennusoireiden yleisesti käytetty raja-arvo.

Syntymäpaino, pituuskasvu, sukupuoli tai puberteettiaste eivät olleet yhteydessä SMFQ:n tulokseen ensisijaisessa analyysissä, jossa puuttuva aineisto paikattiin moni-imputoinnilla. Herkkyyshanalyysi tehtiin aineistolla, joka sisälsi vain ne kohortin nuoret, joilla ei ollut puuttuvaa aineistoa. Tässä analyysissä syntymäpaino oli negatiivisesti ja tilastollisesti merkitsevästi yhteydessä SMFQ:n tuloksen kanssa.

Yhteenvedonä totean, että saavutuskasvu on mahdollista koko väitöskirjan kohortin seuranta-aikana, syntymästä 15 ikävuoteen asti. Myöhäinen puberteetti saattaa tarjota mahdollisuuden saavutuskasvuun myös teini-iässä. Varhaisella kehityksellä on yhteys myöhempään koulumenestykseen ja lapsuuden pituuskasvu ennustaa älykkyydestissä menestymistä nuoruusiässä. Koulutus saattaa vaikuttaa pituuskasvuun ja älykkyyden yhteyteen. Pienellä syntymäpainolla voi olla merkitystä nuoruusiän masennusoireiden kehittämissä.

LIST OF ORIGINAL ARTICLES

The dissertation is based on the following original articles, referred to in the text by the Roman numerals I-IV.

I Teivaanmäki T, Cheung YB, Kortekangas E, Maleta K and Ashorn P: Transition between stunted and nonstunted status: both occur from birth to 15 years of age in Malawi children. *Acta Paediatrica*, Volume 104, Issue 12, pages 1278–1285, December 2015.

II Gandhi M, Teivaanmäki T, Maleta K, Duan X, Ashorn P and Cheung YB: Child Development at 5 Years of Age Predicted Mathematics Ability and Schooling Outcomes in Malawian Adolescents. *Acta Paediatrica*. Volume 102, Issue 1, pages 58-65, January 2013.

III Teivaanmäki T, Cheung YB, Pulakka A, Virkkala J, Maleta K, Ashorn P: Height gain after two-years-of-age is associated with better cognitive capacity, measured with Raven's Coloured Matrices at 15-years-of-age in Malawi. *Maternal & Child Nutrition*, 2016, doi: 10.1111/mcn.12326.

IV Teivaanmäki T, Cheung YB, Maleta K, Gandhi M, Ashorn P: Depressive symptoms are common among rural Malawian adolescents. Submitted for publication

LIST OF FIGURES AND TABLES

Figure 1. Determinants of undernutrition

Figure 2. Malawi map

Figure 3. Tanner classification

Figure 4. Example of an item in Raven's Coloured Matrices

Figure 5. Lungwena Child Survival Study (LCSS) participant flow. The flow chart shows the number of participants who attended the study at each follow-up visit.

Figure 6. Linear growth among boys ($n = 412$) and girls ($n = 394$) in the LCSS cohort. The solid lines indicate HAZ scores and the dashed lines refer to the mean deficits in centimetres compared to the reference population (14, 15). The growth curves are LOWESS-smoothed.

Figure 7. Participants shifting to a different status of stunting or remaining in the same status between age intervals. The numbers on the bars represent the percentages of stunted participants at each time point.

Figure 8. The associations between growth, development, schooling and depressive symptoms in LCSS children.

Table 1. LCSS-follow-up, collected data for each article.

ABBREVIATIONS

BDI-II	Beck Depression Inventory II
CDI	Children's Depression Inventory-II-S
CES-D	Center for Epidemiologic Studies Depression Scale
CI	Confidence interval
DSM-5	The Diagnostic and Statistical Manual of Mental Disorders 5th edition
EFW	Estimated fetal weight
FM	Fine motor
GM	Gross motor
HAZ	Height-for-age Z-score
HDI	Human development index
ICD-10	International Classification of Diseases 10
LAZ	Length-for-age Z-score
LCSS	Lungwena Child Survival Study
LOWESS	Locally weighted regression smoothed curves
MI	Multiple Imputation
MUAC	Mid-upper arm circumference

PHQ-9	Patient Health Questionnaire 9
PVT	Psychomotor vigilance test
RT	Reaction time
SD	Standard deviation
SGA	Small for gestational age
SMFQ	Short Mood and Feelings Questionnaire
WAZ	Weight-for-age Z-score
WHZ	Weight-for-height Z-score
WHO	World Health Organization
WISC-III	Wechsler Intelligence Scale for Children III

TABLE OF CONTENTS

ABSTRACT.....	3
TIIVISTELMÄ.....	5
LIST OF ORIGINAL ARTICLES.....	7
LIST OF FIGURES AND TABLES	8
ABBREVIATIONS.....	9
1 INTRODUCTION.....	13
2 LITERATURE REVIEW.....	15
2.1 Approach to the literature review	15
2.2 Definition of healthy growth and growth faltering.....	15
2.2.1 Assessment of child growth	16
2.3 Determinants of healthy growth and growth faltering.....	17
2.3.1 Epidemiology of growth faltering and catch-up growth	20
2.3.2 Consequences of growth faltering.....	20
2.4 Assessment of child development	22
2.5 Assessment of cognitive capacity.....	23
2.6 Assessment of depressive symptoms	24
2.7 Justification for the present study	24
3 AIMS OF THE STUDY	26
4 METHODS	27
4.1 Study design and study subjects	27
4.2 Study site	28
4.3 Detailed data collection methods.....	30
4.3.1 Anthropometric measurements and pubertal maturity.....	30
4.3.2 Development and school achievement	31
4.3.3 Cognitive performance.....	32
4.3.4 Depressive symptoms.....	34
4.4 Statistical analyses	34

4.5	Ethical aspects	37
5	SUMMARY OF THE RESULTS.....	38
5.1	Enrolment and follow-up	38
5.2	Growth patterns and the timing of stunting and non-stunting (Article I)	39
5.3	Association between early development and later school achievements (Article II).....	43
5.4	Association between the childhood growth and cognitive function (Article III)	44
5.5	Association between the childhood growth and depressive symptoms (Article IV).....	45
6	DISCUSSION	47
6.1	Strengths and limitations of the study	51
7	KEY FINDINGS AND CONCLUSIONS.....	54
8	FUTURE RESEARCH.....	55
9	ACKNOWLEDGEMENTS	57
10	REFERENCES	59
11	APPENDICES	67
	ORIGINAL PUBLICATIONS	69

1 INTRODUCTION

In the year 2014, there were 159 million under-five-year-old children in the world who suffered from stunting, i.e. faltered length or height gain resulting from chronic undernutrition, infections and/or environmental enteropathy, among other symptoms. The prevalence of stunting has decreased over the years, but improvement has been slow in sub-Saharan Africa, where approximately 36% of the world's stunted children live. Stunting is a measure of chronic undernutrition, and it is associated with increased mortality, morbidity and developmental – including neurodevelopmental – problems in early childhood (1).

Stunting affects child health adversely and may make a negative contribution to adult health and economic productivity. In addition, it has an intergenerational effect, resulting in low birth weight, smaller head circumference and lighter brain weight in the offspring of stunted parents. The first and the second of the Sustainable Development Goals of United Nations is to eradicate extreme poverty and hunger. (2) The timing of stunting and its relevance in preventing and treating malnutrition has been under discussion. Growth faltering has typically been considered to begin in low-income countries during the fetal period or soon after birth. It has been thought to continue for the child's first two years of life. Catch-up in linear growth has been considered to be rare after two years of age, but divergent opinions have also been proposed. (3-6)

The timing of pubertal development is essential when evaluating adolescent growth. Growth velocity is at its fastest during puberty, and late pubertal development may facilitate significant potential for late growth recovery. Assessment of pubertal development has not been included in the previous long-term follow-up studies which evaluate growth failure and recovery from stunting. Along with linear growth, brain development continues until early adulthood. This gives a theoretical rationale to suggesting that a better nutritional status may contribute to enhanced neurodevelopment throughout the childhood and into adolescence.

Disturbances in neurodevelopment compromise cognitive performance, which leads to poor education and lower later earnings (1). Besides cognitive capacity, neurodevelopmental deficits may contribute to depressive symptoms later in life. Depression caused 26.5 million disability-adjusted life years in developing countries

in 2004 (7), and it has been predicted that by year 2030, depression will be among the three leading causes of disability-adjusted life years worldwide, also in developing countries. In sub-Saharan Africa, approximately 14% of children suffer from psychological disorders. (7-10) The prevalence of depression is estimated to be higher in developing countries than in developed countries (11,12), However, child and adolescent mental health services are rarely available in the former.

This present prospective cohort study was designed to describe the timing of stunting and possible growth recovery in rural Malawian children. We also aimed to assess the association between early development and later school achievement, plus the associations between growth during both early and later childhood and cognitive capacity and depressive symptoms in adolescence.

2 LITERATURE REVIEW

2.1 Approach to the literature review

The literature review aims to provide background information for the current study. It includes definition of healthy growth and growth faltering, assessment of child growth, determinants of healthy growth and growth faltering, epidemiology of growth faltering and assessment of growth, development, cognition and depression. The literature was primarily searched from Tampere University's Nelli and Andor portals, through which there is access to all the relevant databases, such as Medline, Pubmed, and PsycInfo. The principal search words were growth/growth pattern, malnutrition/stunting, cognition, child development, child/adolescent, depression/depressive symptoms, developing countries/low-income countries/sub-Saharan Africa. Recent publications and primary data was preferred, but in addition, some review articles and reports were used.

2.2 Definition of healthy growth and growth faltering

Healthy growth means normal linear growth relative to the World Health Organization (WHO) Growth Standards. Growth may be compromised for several reasons, for example malnutrition or illness. The postnatal growth of a child is greatly affected by perinatal conditions and morbidity. Hence, fetal growth, which is affected by the nutritional status and health of the mother, should be carefully monitored. To meet the demands of fetal growth monitoring, WHO recently provided growth charts for estimated fetal weight (EFW) (13).

Postnatally child growth assessment may include the measurements of weight, length/height, mid-upper arm circumference (MUAC) and head circumference. Growth is considered adequate if the z-score values are between -1 and +2 standard deviations (SD; (1,14-16). The definition of malnutrition includes undernutrition, i.e. stunting, underweight and wasting and overweight. It also comprises deficiencies of micronutrients, including essential vitamins and minerals (14,15,17,18). Stunting is a measure of chronic or recurrent malnutrition resulting from nutritional deprivation

and/or recurrent infections or psychosocial deprivation. It means that length-/height-for-age z-score (LAZ / HAZ) of a child is below -2 SD compared to the WHO Growth Standards. Children are considered severely stunted if their LAZ / HAZ is below -3 SD from the growth standards (19). If weight-for-age (WAZ) is below -2 SD, child is underweight and if weight-for-height (WHZ) is below -2 SD, he or she is wasted. Underweight and wasting are severe if the weight is below -3 SD (14,15). It is, however, important to realize that, despite these cut-off points, growth faltering may also occur within the normal range, i.e. above -2 SD (20). Stunting reflects a failure to reach linear growth potential. Being underweight means that the body mass is low relative to chronological age. It can be influenced by both child's height and a child's weight. Therefore, underweight cannot distinguish between a child who is light in weight relative to his/her height and a child who is short relative to his/her age, but normal in weight-for-height. Wasting reveals acute undernutrition, which may be a result of recent food deprivation or illness (21). In the present study, we focus on stunting, which generally is considered a measure of chronic undernutrition and/or recurrent infections.

2.2.1 Assessment of child growth

Regular assessment of child growth provides an opportunity to detect growth failure and possible health problems that relate to growth. Without recognizing the individuals suffering from overnutrition or undernutrition, it is impossible to identify the magnitude of the problem or focus on increased nutritional or health needs. Anthropometric measures are non-invasive and easy to conduct, even in low-income settings (22).

The global consensus on methods on growth assessment should be followed. For example, linear growth should be assessed recumbent (length) for the children less than two years of age and standing (height) for the children older than two years of age (20). Length/height and weight are evaluated against the reference growth measurements in a growth chart. There are both national and global growth charts that are based on different reference populations. In this study, we used the latest WHO Growth Standards (14,15), which were developed after the LCSS cohort children were born. There was no stand-alone community health program in Malawi during the childhood of the LCSS cohort children, and this is still the case. Community health services are considered part of primary-level care within the national health system, and they are delivered to the community through the Ministry

of Health's various intervention-specific programs. They include growth monitoring (23). In addition, growth monitoring is offered by non-governmental organizations, but it still does not reach the majority of Malawian children (24). Like in many other countries, the growth charts in Malawi were based on 1978 National Center for Health Statistics (NCHS) growth references before the latest WHO Growth Standards were developed (25).

Comparison to the reference population not only shows the current growth status of the child, but deviation from the earlier growth curve also reveals growth faltering or excessive growth. Growth charts are essential and reliable tools for monitoring the growth velocity of a child when correctly used. Healthcare workers should be provided with detailed instructions about appropriate growth references and when or what actions should be taken when growth faltering is detected.

2.3 Determinants of healthy growth and growth faltering

Several determinants contribute to childhood growth, undernutrition and growth failure. Prenatally the factors are intra-uterine exposures, maternal energy and micronutrient intake, and overall nutritional status. Shorter mothers tend to have smaller babies. The mechanisms behind this intergenerational effect include similar genetic characteristics, epigenetic effects, programming of metabolic changes and reduced space for fetal growth (20). Additionally, maternal malnutrition or infections may lead to fetal growth restriction (26).

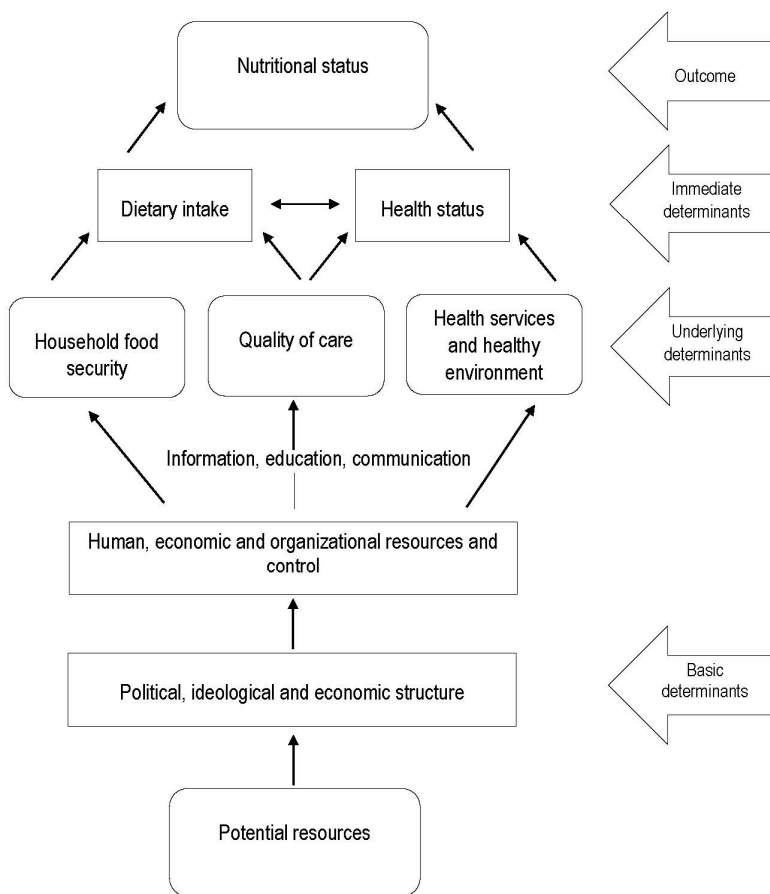
Postnatally the determinants of growth include adequate or inadequate exclusive breastfeeding and proper or unsuitable or deficiency of complementary feeding. In the developing countries, the earlier the exclusive breastfeeding has stopped and complementary feeding introduced, the greater the risk is for stunting. (27) Socio-economic background, availability of health services and vaccination, and occurrence of infectious diseases may also affect child growth. In later childhood and puberty, growth and sex hormones contribute more to healthy linear growth (28). It is known that if the birth spacing is too short or if the birth rank is high, the risk for stunting is greater. Many socio-cultural factors such as maternal literacy, nutritional knowledge and the intergenerational transmission of poverty and emotional deprivation also contribute to stunting (20).

It has been proposed that children may be stunted, even with adequate nutritional intake, because of environmental enteropathy or environmental enteric dysfunction (29). This inflammatory condition is predominantly a T-cell mediated adaptive

response to environmental challenges in tropical low-income settings. It causes villous blunting, crypt hyperplasia and lymphocytic infiltration of the intestinal epithelium and lamina propria, resulting increased intestinal permeability and decreased absorptive capacity (30-32). Despite of many biomarkers that have been tested as proxies for environmental enteropathy, proper diagnostic tools do not exist. The suggested potential interventions against environmental enteric dysfunction have been probiotics, antibiotics, anthelmintics and nutritional supplements, but they have shown limited or no effect on the condition. Therefore, there is no effective treatment for environmental enteropathy (33,34).

There are many socio-economic and political factors behind all the above-mentioned determinants of nutritional status of children (Figure 1). The causes of food insecurity and malnutrition relate both to external factors and the historical background of the society in question. Common basic causes are difficult ecological conditions and an inefficient use of technology. Economic causes include external economic dependency and maldistribution of productive assets such as land. Political causes are consumer and producer pricing structures and income policies. In addition, the possible subordination of women affects the food security. The power structure both within and among households is often legitimized and imbedded in the accepted local culture (35).

Figure 1. Figure 1. Determinants of undernutrition. Adapted from Unicef, Strategy for Improved Nutrition of Children and Women in Developing Countries, 1990 (35).



2.3.1 Epidemiology of growth faltering and catch-up growth

In compromised settings such as developing countries, growth faltering typically starts during the fetal period or soon after birth and continues during the child's first two years of life (3,5,6). The time from conception until the age of two, i.e. the first 1000 days of life, has been considered the window of opportunity for growth promotion. Growth and development are vigorous during this period, and body and organisms vulnerable to malnutrition (36).

Between two and five years of age, the mean HAZs remain relatively stable in most populations (16,37). The term 'catch-up growth' was introduced in the 1960s by Prader and his collaborators. It means rapid linear growth which allows the child to accelerate toward his/her pre-retardation growth curve (38). Catch-up in linear growth is considered rare after two years of age on a population level, but there are also divergent suggestions. Crookston and his collaborators proposed in 2013 that stunted children may demonstrate catch-up growth after the first 1000 days of life, i.e. 24 months of age. Catch-up growth in later childhood, measured as HAZs, has also been identified in other studies (3,4). This gives a rationale to explore whether there is an association between both early and late childhood growth and later health and development (37).

Catch-up growth after the first 1000 days of life would provide a possible window for growth promotion and interventions later in childhood. There are, however, some additional challenges in interventions during that time period. There is a risk for accelerated maturation and closing of the growth palates, which potentially lead to halted growth and further overweight (39). Weight gain in later childhood is associated with increased adult fat mass, and gains beyond the ideal range will increase mortality risks associated with low stature. Without concurrent reduction in child stunting and an improvement in adult stature, increasing overweight will fail to reduce mortality and increase productivity (39-42).

2.3.2 Consequences of growth faltering

Stunting may be considered a marker of many pathological disorders that are linked with increased morbidity and mortality (20). It has been proposed that childhood stunting is associated with short adult height, poor neurodevelopmental and cognitive function, elevated risk of chronic disease in adulthood, less education and reduced earnings. In most of the occupations, an evident link between height and success in work life cannot be determined. However, height may be thought of as a

distinctive marker for cognitive and behavioural development, schooling outcomes and health, which have an impact on occupational success. Stunting has, for example, been linked with late enrolment in school and grade repetition (43-45).

The mechanisms of the link between stunting and cognitive function are ambiguous. There are, however, some theories behind this association. Stunting and catch-up growth represent nutritional status. Deficiencies and imbalances of macronutrients, e.g. fatty acids and micronutrients such as iron, may have direct effects on various neuronal processes (46,47). The development of apical dendrites from the brain cortex continues after birth and is completed by two years of age. Undernutrition is thought to shorten dendrites, decrease the number of dendritic spines and increase abnormal dysplastic dendritic spines. These alterations may affect brain functions (48). In addition to cognitive development, stunting may contribute to compromised motor development, which may potentially hinder social and emotional stimulation and school education (43).

Some aspects of brain development continue in later childhood. For example, synaptic blooming and pruning in the medial prefrontal cortex has a later time course; an adult synaptic density is not obtained until adolescence. This development relates to the regulation of higher-level cognition (49). It may mean that stunting and catch-up growth have an impact on cognitive development and mental health into late childhood and adolescence.

Other than direct effects on brain development, the lack of energy or infectious diseases may reduce interactions between children and their caregivers and reduce social stimulation (46,50). In addition, many important developmental milestones occur during preschool years. Growth may affect developmental status and therefore school readiness and academic performance. It is also possible that a critical window for cognitive development does not exist, but instead, cognitive development is cumulative over the time (49). Consequently, a longer duration of positive exposures would lead to greater benefits.

Stunted children may present with more mental health problems and depression compared with children with normal linear growth (20,51-53). This may be due to altered brain structure and function, as seems to be the case in cognition. Lower intelligence is associated with more depressive symptoms (54). It is possible that stunting affects cognitive and psychological functioning concurrently, or, alternatively, poorer cognitive functioning leads to more depressive symptoms (52,54). Stunted children present a modified stress response and higher cortisol levels during stressful events compared with non-stunted children (52). Cortisol levels are also elevated in non-stunted patients with depression, which may indicate a

significant role played by this hormone in the psychological problems of stunted children (52,55). De Mola and his collaborators suggest that the association between growth failure and depressive symptoms is cumulative. They observed that the risk of depression was higher in subjects who were small for gestational age (SGA) at birth and were also stunted at the age two or four years (53).

In a short perspective, stunting increases morbidity and mortality from conditions such as pneumonia, diarrhoea and other bacterial and viral infections. This is most likely due to a generalized immune defect through reduced appetite, impaired intestinal absorption, increased catabolism and direction of nutrients away from growth and towards the immune system. Impairments in the immune system apply to both innate and acquired immune function. Furthermore, recurrent infections contribute to more stunted growth, leading to a vicious cycle of poor nutritional status (20,56,57).

Stunting has intergenerational effects. Children born to stunted women are lighter and at greater risk of morbidity, mortality and stunted growth (51,58). In addition, malnutrition and growth failure early in life may lead to later metabolic disorders such as disturbances in glucose and lipid metabolism and high blood pressure. The theory of the Developmental Origins of Health and Disease is based on Barker's hypothesis. According to this hypothesis, intrauterine growth retardation and the early post-natal environment modify genome expression and cause long-term alterations to metabolic functions. Hence, low birth weight and premature birth may associate with hypertension, coronary heart disease and type II diabetes later in life (51,59).

2.4 Assessment of child development

Developmental assessment is evaluating a child's performance compared to other children of the same age. Performance may vary across different population groups (60). There are several tools that can be used to assess child development, of which the most common ones are Bayley scales (61), Griffith's (62), McCarthy scales (63) and Denver II (64). These all are designed to be used and validated in Western countries. Standardization and validation of these tools is limited in Africa. Translation alone may not be sufficient enough to take into consideration all the local expressions and customs (65). The developmental assessment tool that was used in this study was created and standardized by Gladstone and her collaborators. It is a simple and culturally appropriate developmental assessment tool based on

various Western tools and intended for use in rural Malawi. It was mostly based on Denver II, the Denver Developmental Screening Test and Griffiths Mental Developmental Scales (65). This tool included 138 items with 34 gross motor (GM), 34 fine motor (FM), 35 language and 35 social items. Most of the items were adapted from the Denver II or Denver Developmental Screening Test.

2.5 Assessment of cognitive capacity

A variety of cross-sectional and longitudinal studies have demonstrated an association between early or concurrent stunting and cognitive capacity (43,44,66-68). The emphasis has mainly been on the consequences of the growth in early childhood (37,43,68,69). Cognitive function is considered to consist of different domains. The Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-5) defines six key domains of cognitive function: perceptual-motor function, language, executive function, learning and memory, complex attention and social cognition (70). Another traditional classification of intelligence divides it into three components: verbal ability, nonverbal reasoning ability and spatial ability (71). There is a variety of tests that can be used to assess cognitive capacity or intelligence, each with a different focus on various populations and age groups. The most commonly used tests for children and adolescents in both research and clinical work are the Wechsler Intelligence Scale for Children, Kaufman-ABC, Stanford-Binet and Raven's Standard and Coloured Matrices (72-75). Many of the tests are multidimensional, whereas Raven's Matrices was designed to measure non-verbal reasoning or general intelligence (74). These developmental domains refer to the ability to identify and infer rules for novel problems. They are independent of skills relying on previously learned knowledge (76). Most of the tests are conducted with a trained tester and a single subject, but Raven's Matrices may be used with a group and without a highly skilled tester. According to the DSM-5, each cognitive key domain consists of many subdomains. For example, visual construction is one of the subdomains of perceptual-motor function and processing speed a subdomain of complex attention. It may be suggested that Raven's Coloured Matrices measures visual construction and reaction time (RT) processing speed, which has been associated with intelligence (77). These domains should be culturally independent and not affected by education. In addition, a mathematics test (45,53) was used to assess cognitive capacity in this study (Appendix I). Success in this test is naturally

affected by the years and quality of schooling. The tests are described in detail in the section on data collection (section 4.3).

2.6 Assessment of depressive symptoms

The prevalence of depression rises in teenage years, and early interventions are crucial to prevent severe depressive symptoms. This emphasizes the need for a sensitive screening tool, which would identify even modest elevations in symptom severity for those at risk of depression. However, the selection of adolescent depression screening tools is not complete (78-80). Adolescent depression is often assessed using adult depression instruments such as the Beck Depression Inventory (BDI-II) (81), the Center for Epidemiologic Studies Depression Scale (CES-D; (82)) and the Patient Health Questionnaire (PHQ-9; (83)). The latter has received support for its applicability as a screening tool for adolescents (84). In addition, a symptom-oriented Children's Depression Inventory (CDI-II-S) for children 7-17 years old has been developed based on the Beck Depression Inventory (85).

There were no depression screening instruments that would have been validated in a rural African setting. Short Mood and Feelings Questionnaire (SMFQ) was chosen to use for assessment of reported depressive symptoms at 15 years of age (Appendix II). The SMFQ is self-administered and provides a cheap, easy and reliable measure of depressive symptoms in children at a variety of ages (86,87). The questionnaire was translated from English into the Yao language and then translated back into English. The translation was revised until inaccuracies were not detected. The SMFQ has 13 questions with three response options: true (two points), sometimes true (one point), and not true (no points). The range is 0–26 points, and a score ≥ 11 is traditionally suggested to refer to significant depressive symptoms. With this cut-off, the specificity and the sensitivity are 83% and 71% respectively for identifying those who meet the International Classification of Diseases 10 (ICD-10) criteria for depression (87).

2.7 Justification for the present study

Malnutrition remains an enormous problem in low-income countries. Despite the great efforts of local authorities as well as global organizations such as WHO, improvements in the nutritional status of children have been limited (1). The

knowledge about the late consequences of early malnutrition is contradictory. There is also a gap in the literature about the effects of stunting in late childhood. Rather established understandings about the timing of possible catch-up growth after stunting have recently been challenged. According to current knowledge, catch-up growth is considered rare after two years of age, but this theory has been criticized by some researchers, who have suggested that catch-up growth is possible also later in childhood, not only before the age of two.(3,4,88)

This study was designed to increase the existing body of knowledge about growth trajectories, timing of stunting and catch-up growth, and the relationship between these factors and cognition, schooling and mental outcomes. Information on the timing of stunting and growth recovery and the association between these factors and later health and development will help authorities implement nutritional programmes, and increase awareness of this devastating global problem.

3 AIMS OF THE STUDY

The aims of this thesis were to describe the timing of stunting and possible growth recovery in a rural low-income population, to examine the associations between childhood growth and later health and developmental outcomes such as cognitive capacity and depressive symptoms in adolescence, and to examine the association between early development and later school achievement.

The specific aims of the present study were:

1. To assess the timing and extent of stunting and non-stunting between birth and 15 years of age (Article I).
2. To assess the association between child development at five years of age and mathematics ability and schooling outcomes at 12 years of age (Article II).
3. To determine the association between height gains at different ages, including late childhood, and cognitive capacity at 15 years of age (Article III).
4. To determine the prevalence of depressive symptoms and associations between birth weight and height gains at different ages and depressive symptoms at 15 years of age (Article IV).

4 METHODS

4.1 Study design and study subjects

This study was a prospective cohort study. The Lungwena Child Survival Study (LCSS) cohort was enrolled between June 1995 and August 1996 in the Lungwena community. The original aim of the study was to describe health and its determinants among women and their children in rural Malawi. It also aimed to identify factors that might either promote or hinder healthy growth and development among pregnant women and their offspring. All pregnant women seeking antenatal care were eligible for the study, and 97% of the pregnant women in the area were enrolled. Their children were followed from the fetal period until the age of 15. The children were measured for anthropometry up to 37 times during the follow-up, in the beginning at homes of the participants and later at the study clinic in Lungwena. After enrolment, the cohort members were collected from their homes for participation at each assessment by the data collectors. A developmental assessment inventory was conducted at five years of age, the mathematics test at 12 years of age and the pubertal stage, cognitive capacity and depressive symptoms at 15 years of age, all in addition to anthropometrics. The study was divided into four sub-studies, according to which four original articles were published or submitted for publication (Table 1).

Table 1. LCSS follow-up, collected data for each article.

		LCSS follow-up			
Aims	Article I	To examine the association between child development at five years of age and mathematics ability and schooling outcomes at 12 years of age.			
	Article II	To explore the timing and extent of stunting and non-stunting between birth and 15 years of age.			
	Article III	To determine the association between height gains at different ages, including late childhood, and cognitive capacity at 15 years of age.			
	Article IV	To determine the prevalence of depressive symptoms and associations between height gains at different ages and birth weight and depressive symptoms at 15 years of age.			
Age	Collected data	Article I	Article II	Article III	Article IV
< 1 week	Birth weight*			x	x
1-18 months	Weight, length, 1-month intervals		x		x
18-60 months	Weight, length/height, 3-month intervals		x		x
5 years	A developmental assessment Inventory	x			
6 years	Weight, height		x		x
9 years	Weight, height		x		x
10 years	Weight, height		x		x
12 years	Weight, height, schooling history, a mathematics test	x			
15 years	Weight, height, pubertal stage, cognitive capacity (Raven's Coloured Matrices, RT with PVT, mathematics test), depressive symptoms (SMFQ)		x	x	x

*Children were measured before the age of 1 week

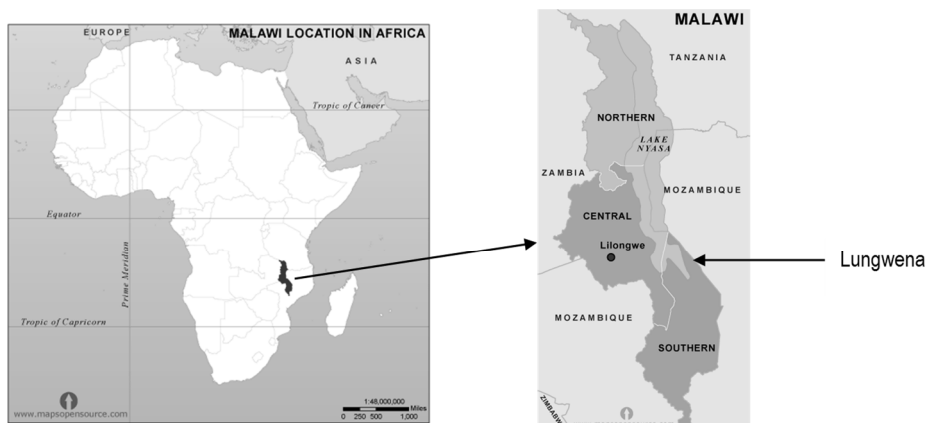
4.2 Study site

The study was conducted at Lungwena in the Mangochi District of Southern Malawi. Malawi is a small country in southeast Africa with 17.1 million inhabitants (Figure

2). Life expectancy at birth was 60.9 years in 2016 (89), and the human development index (HDI) was 0.476 in 2015 (90). The HDI is a summary measure of key dimensions of human development – namely, a long and healthy life, being knowledgeable and having a decent standard of living. The HDI is the geometric mean of the normalized indices for each of these three dimensions. It was created to emphasize people and their capabilities as the ultimate criteria for assessing the development of a country rather than just economic growth (90). Over 60% of Malawian children under five years of age are undernourished (1). In Malawi, under-five mortality has dropped from over 250 to 64 per 1000 births between 1990 and 2015. However, the country still ranks 33 in the world in under-five mortality.

The Lungwena community covers a 100 km² area comprising 26 villages and 23,000 inhabitants in about 5,200 households. Most of the inhabitants are Muslims belonging to the Yao tribe and live in matrilineal descent patterns. The literacy rate is low and the main sources of income are farming – mainly maize – and fishing. Approximately 60% of the mothers in Lungwena Child Survival Study (LCSS) gave birth at home. Others delivered in traditional birth attendant facility, health centre or at hospital. (91)

Figure 2. Map of Malawi (maps from www.mapsofopen.com).



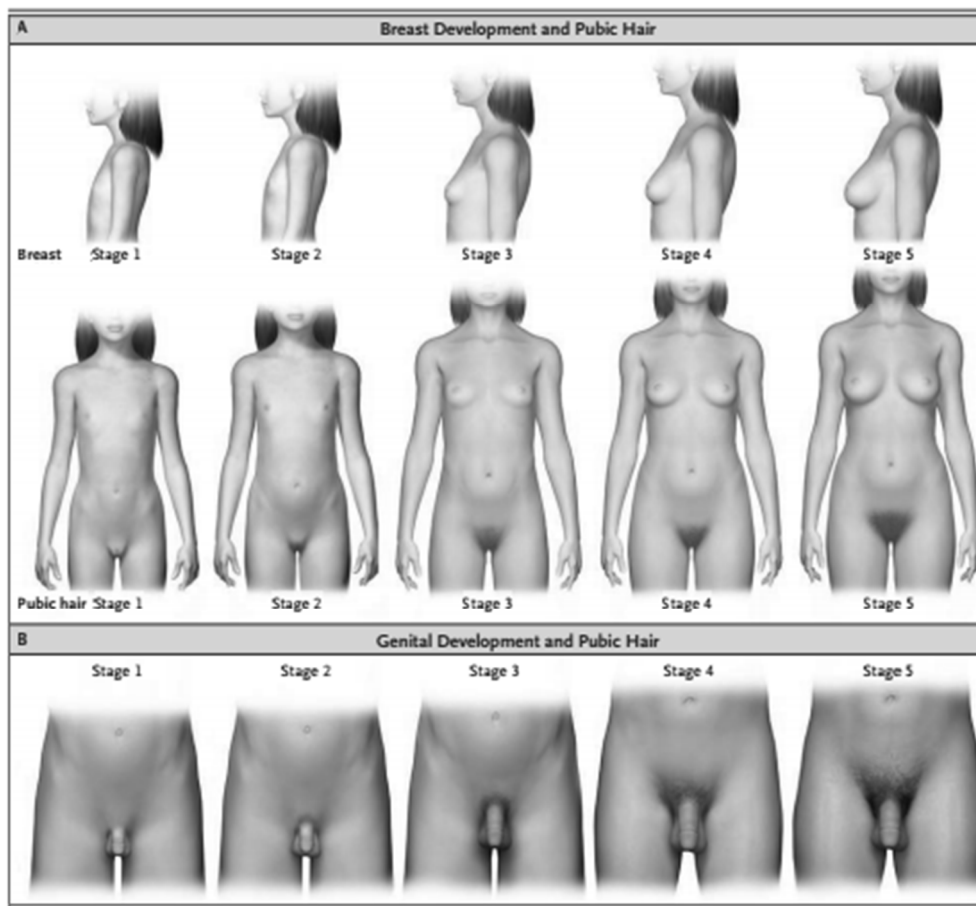
4.3 Detailed data collection methods

4.3.1 Anthropometric measurements and pubertal maturity

Research assistants measured the participants at their homes every month until they were 18 months of age and, after that, every three months until they were five years of age. After this, measurements were taken at the study clinic at 72, 108, 120, 144 and 180 months of age (15 years; Table 1). The first measurements were conducted as soon as possible after birth. During the 15-year period, the maximum number of anthropometric measurements carried out for each child was 37. The length and height of the children were measured using locally constructed length and height boards until five years of age and with stadiometers (Harpenden, Holtain Limited, UK) after that. Children were measured standing up if they were able to stand or were over 24 months of age. Those who were under the age of two and unable to stand were measured using a length board. The length and height boards and stadiometers were calibrated weekly, and no technical problems were detected with the equipment.

Research assistants assessed the pubertal development of the participants at 15 years of age using the Tanner classification (92), which includes five stages (I–V) for pubic hair development for both sexes, genitalia development for boys and breast development for girls (Figure 3). Stage I is the pre-pubertal stage and Stage V full development (92). Five different groups were formed based on the pubertal hair development to be used for comparison of growth in HAZ and mean height deficits in centimetres with the reference group.

Figure 3. The Tanner classification (Figure from Carel JC, Léger J. Precocious puberty, 2008 (93)).



4.3.2 Development and school achievement

At the age of five years, each child was invited to a developmental assessment (Table 1). The assessment was carried out by a trained research assistant. Training consisted of four two-day sessions over a two-month period, run by a paediatrician and an accredited trainer in the Griffiths Scales who also had experience in the use of other developmental screening tools. The exposure variables in Article II were the development summary score, the percentage score of each domain (gross motor, fine motor, language and social domains) separately and all four domain percentage

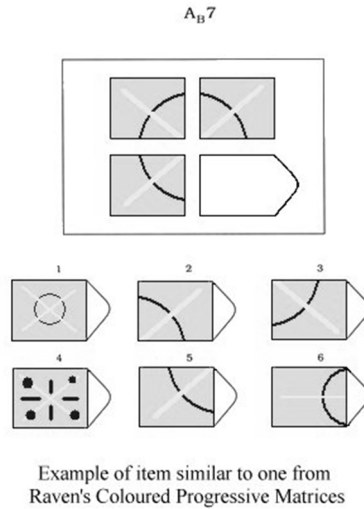
scores included at the same time. The formation of summary and percentage scores is described below in the Statistical analyses section.

At the age of 12 years, information was collected on schooling, the highest school year completed and number of times a school year was repeated, and each child was given a mathematics test (Table 1). The mathematics test that we used in the study was developed from a reference test designed for children aged 7–14 years used in the Indonesian Family Life Survey (Appendix I; (69,94). The test has previously associated with preceding development (95) and linear growth (69), which is a predictor of cognitive function in children in low-income countries (43).

4.3.3 Cognitive performance

Cognitive capacity at 15 years of age was assessed using Raven's Coloured Matrices (74), a mathematics test (69,94) and a RT test (96). Raven's Coloured Matrices has previously been found to correlate with the performance component of the Wechsler Intelligence Scale for Children (WISC-III; The Psychological Corporation 1997; (74). Raven's Matrices is considered to be a culturally unbiased test independent from language skills. It is suggested to measure general intelligence or 'Spearman's g'. General intelligence is described as the ability to solve novel problems without relying on previous knowledge or experience (97,98). Raven's Coloured Matrices was easy and rather time efficient to administer, whereas, for example, the WISC-III would have demanded more time and more data collector training. A pilot study was conducted with Raven's Standard Progressive Matrices a setting and age group similar to that of our study (99) and received poor results with noticeably little variability. Hence, Raven's Coloured Matrices was used in this study. It is originally designed and standardized for children younger than 15 years of age (74). This assessment has 36 patterns each with a piece missing. The participants choose from six alternatives the piece that they consider correct to complete the pattern (Figure 4). The best possible score was 36.

Figure 4. Example of an item in Raven's Coloured Matrices.



There is correlation between RT and intelligence (77,100). Based on a review by Sheppard and Vernon, the average correlation between RT and general intelligence is -0.26, with a range of -0.22 to -0.40 (101). RT was assessed using the computer-based Psychomotor Vigilance Task (PVT; (96)), which measures the time between a stimulus and a participant's reaction over the course of a five-minute test period with a random inter-stimulus interval of two to ten seconds. A number appears on a black display, and the participant indicates his or her reaction by pressing the enter key with a forefinger of the dominant hand. Two PVT performance metrics were evaluated in the study, the median RT (milliseconds) and the number of lapses (RTs ≥ 500 ms) in a five-minute trial. Before the test, each participant performed a test trial. In addition to Raven's Coloured Matrices and RT, the participants were given the same mathematics test they took at 12 years (69,94).

In addition to Raven's Coloured Matrices and RT, mathematics test was used to assess cognitive function of the participants at 15 years of age (45,53). It was the same test that was used in the 12-year assessment; the details of this test have been described in the previous chapter. The mathematics test is shown below as Appendix I.

4.3.4 Depressive symptoms

Depressive symptoms were screened for using the SMFQ at 15 years of age (86,87). The questionnaire was translated to local Yao language. The SMFQ measures depressive symptoms through 13 items, each with three answer options: true, sometimes true and not true. The scoring for each item is 0, 1 or 2, and the range of the total score is 0 to 26. A score of 11 or more has been traditionally been used to refer to clinically significant depressive symptoms (87). The illiterate participants were interviewed.

4.4 Statistical analyses

The data was collected by data collectors on paper forms. It was then manually entered by data clerks into a Microsoft Access database (Microsoft Corporation, Redmond, WA, USA). The manual data entry took the form of double entry. After this, all the data was converted into Excel and further into Stata files. All the paper forms were scanned and stored as pdf files.

The World Health Organization Multicentre Growth Reference Study (16) was used to derive LAZ and HAZ from absolute length and height values for children up to the age of five, and WHO Reference 2007 (17) used for children from five to 15 years of age. Reference values for length were used for the measurements before 24 months of age and height for 24 months of age or more. HAZs are calculated by dividing the absolute deviation of the measured height from the expected height by the age- and sex-specific standard deviation (SD). The deviation is a negative value if the child is stunted. SD is the square root of variance, which is the average of the squared deviations from the mean. It is calculated from the actual measurements of the reference population. One standard deviation stands for 67% and two standard deviations for 95% of the population if added and subtracted from the average (88).

$$Z\text{-score} = \frac{\text{deviation}}{\text{SD}} = \frac{\text{observed height} - \text{expected height}}{\text{SD}}$$

Locally weighted regression smoothed curves (LOWESS) were plotted for the Z-scores and mean absolute deficits in centimetres compared with those of the reference population (16,17) for both sexes (Article I). Mean absolute deficits were calculated by subtracting the median values of the reference population from the observed values and then taking the mean from those deviances at each time point (Article I). The associations between early growth and cognitive performance (Article III) and depressive symptoms (Article IV), and the association between development and schooling outcomes (Article II) were assessed with multiple logistic and linear regressions.

Chained multiple imputation was used before the regression analyses to handle missing data where appropriate. The missing values were replaced by multiple imputations (MIs) using all the variables in the regression models for the analysis of the outcomes (102,103). Imputation by chained regression is an iterative procedure to obtain multiple set of complete data values (104). Multiple imputation creates multiple predictions for each missing value by regressing the variable upon the observed and previously imputed values on other predictors. It then randomly samples from the conditional distribution and further updates the prediction for the missing values by iterating the process through the other predictors and updating the imputed values by repeating the process for multiple cycles. Standard statistical analyses were applied to each completed dataset (102). 50 sets of imputation was used and 50 sets of analytic results pooled using Rubin's rule, which accounts for the uncertainty arising from the imputation (102,104).

In Article II, the child development summary score obtained at five years of age was the main exposure variable. The outcome measures were percentage of correctly answered mathematics questions, highest school grade completed and number of repeated school grades at 12 years of age. The associations were estimated using regression analyses. Model 1 included the summary score of child development as the exposure and the three outcome variables separately. Model 2 fitted the regression models separately with each domain's percentage score. Model 3 included all four domain percentage scores simultaneously. A percentage score was calculated for each development domain: gross motor, fine motor, language and social. The number of items passed were divided by the number of items administered, and then multiplied by 100. A standardized factor score, that represents the summary score of each child's development, was derived in this study. A factor score was derived from the percentage scores from the four domains using factor analysis with the Bartlett method in order to prevent inflated type I errors due to multiple testing and in order to study the combined predictive power of all four developmental domains

(105,106). With the Bartlett approach, the sum of the squared components for the error factors is minimized. The Bartlett method produced unbiased estimates of the true underlying factor scores. Standardization was performed by subtracting the mean score from the observed score and then dividing the equation by the SD, so that the score has a mean of 0 and standard deviation of 1.

In Articles II, III and IV, possible confounders were added in the regression models where appropriate. In Article II, adjustments were made for weight-for-age Z-score near birth, gender, gestational age, father's occupation, father's literacy, mother's literacy, socio-economic level, age at the five-year assessment, HAZ at the five-year assessment; and Articles III and IV adjustments were made for gender, gestational duration (weeks), father's occupation, father's literacy, mother's literacy and a wealth index (107). The wealth index was assessed perinatally by interviewing the mothers. It summarized household ownership of radio, bicycles or tricycles, or a mattress; the number of family supporters; ownership of land per person; and number of cattle (cows, goats, sheep and chickens). It was derived from factor analysis and categorized participants into three levels: poor (below 40 percentile), middle (40–80 percentiles) and rich (top 20 percentile). Gestational duration was estimated using the nationally used chart for fundal height during the antenatal visits, because ultrasound was not available and information on the timing of menstrual periods was not reliable (108).

In Articles III and IV, Model 1 included only the exposure and the outcome variables. The exposure variables were HAZ_1, HAZ_24 and HAZ_180 in Article III, and birthweight, HAZ_1, HAZ_24 and HAZ_180 in Article IV. The outcome variables were Raven's Coloured Matrices, mathematics test score, median RT (milliseconds) and lapses in RT assessment in Article III, and SMFQ score in Article IV. In further analyses, regressions were adjusted for potential confounders and intermediate variables (Articles III and IV). Model 2 (III and IV) was adjusted for the potential confounders described above. Model 3 was further adjusted for the number of years of schooling reported at the age of 15 years (Article III) or for Raven's Coloured Matrices score and pubertal maturity (Article IV). In Article III, I also ran analyses that were otherwise identical to the primary models, but used unexplained residuals (conditional growth measures) instead of HAZs as the exposure variables, which represent child's deviation from his or her expected HAZ independent of his/her earlier HAZ (109).

All the analyses were performed using Stata 12.1 (Stata Corporation, College Station, TX, USA). The MI procedure was performed using the ICE package by Stata software (103), and the standard deviation scores (Z-scores) for

anthropometric measurements were generated using a Stata macro (18). Detailed statistical methods are found in the individual articles.

4.5 Ethical aspects

The studies included in this thesis followed the ethical guidelines of the 2008 World Medical Association's Declaration of Helsinki, which encourages the protection of the 'life, health, dignity, integrity, right to self-determination, privacy and confidentiality of research subjects' (110).

Ethical approval of the Lungwena Child Survival Study (LCSS) was obtained from the National Health Science Research Committee in Malawi (HSRC 93/94) and the College of Medicine Research and Ethics Committee. Informed consent was obtained from each guardian at the beginning of the cohort study and again from each guardian and adolescent before the visit at 15 years of age. In the event of illiteracy, the written informed consent was signed with a thumbprint.

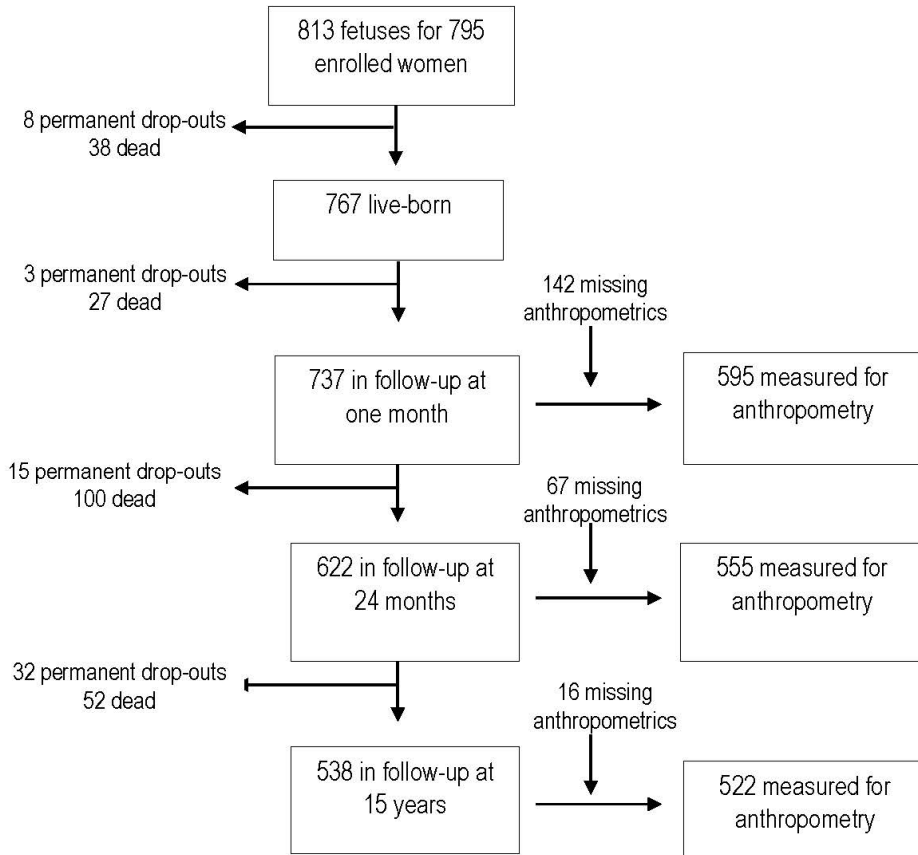
The cohort members received a small incentive, soap or body cream, and rice after participating in the examination at 15 years of age. They were also offered a lunch during the study day. Other than that, they received no benefits from participating in the study. If a medical condition was detected during the study visits, the participant was referred to the local healthcare authorities.

5 SUMMARY OF THE RESULTS

5.1 Enrolment and follow-up

The cohort was enrolled between June 1995 and August 1996, and it originally comprised 795 mothers who attended the antenatal clinic at Lungwena health centre during their pregnancies. The total number of fetuses enrolled was 813, and the number of children born live was 767 (Figure 5). These infants had a mean (SD) birthweight of 3060g (530). The proportion of newborn infants with a low birthweight (< 2500g) was 10%, and 22% of the births were preterm (< 37 weeks of gestation). The participants were measured up to 37 times during the follow-up.

Figure 5. Lungwena Child Survival Study (LCSS) participant flow. The flow chart shows the number of participants who attended the study at each follow-up visit.



5.2 Growth patterns and the timing of stunting and non-stunting (Article I)

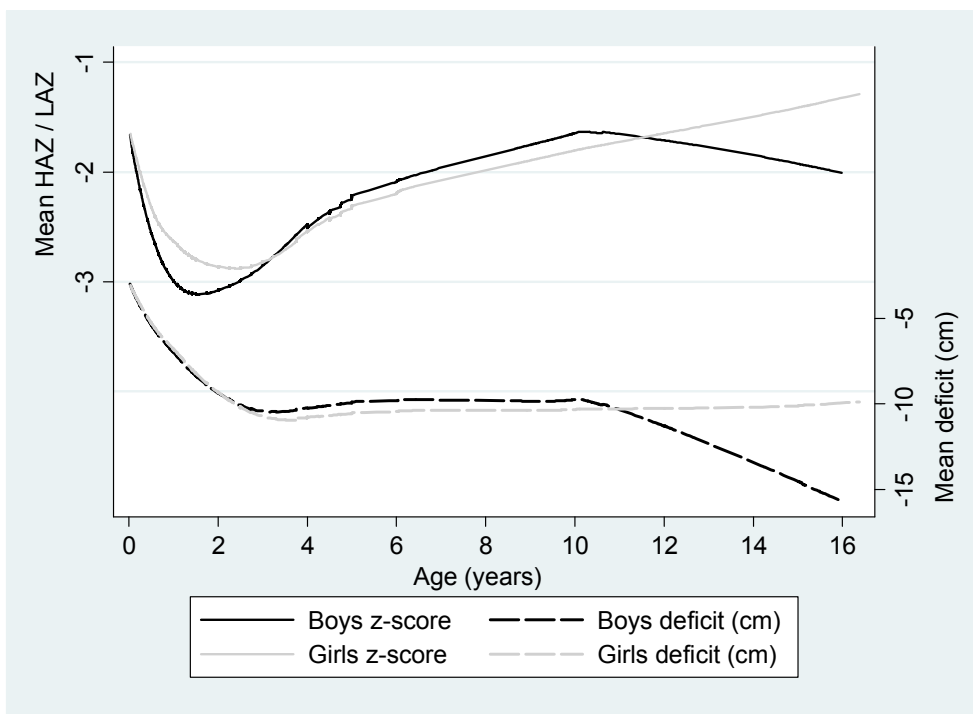
Anthropometrics were measured in 522 of the 538 participants who attended the study at 15 years of age (Figure 5). Of these participants, 258 were boys and 264 girls.

During the study, there were 20,683 individual measurements performed: an average of 27 and maximum of 37 for each participant. The linear growth of the participants from birth until 15 years of age is described in Figure 6. The mean (SD) length/height of the boys was 51 cm (2.5) at one month, 100 cm (4.8) at five years and 154 cm (8.2) at 15 years of age. The respective figures for the girls were 50 cm (2.4), 99 cm (4.6) and 153 cm (5.8). The mean LAZ declined rapidly until two years of age, after which it reached the lowest point of -3.1 in the boys and -2.9 in the girls (Figure 6.). After two years of age there was an increase in the mean Z-score that continued until the age of 15 in girls, but decreased after ten years of age in boys. At 15 years of age, the mean HAZ was approximately -2.0 in boys and -1.3 in girls (Figure 6.).

In this study, it was demonstrated, that catch-up growth and transition from non-stunted to stunted status and vice versa throughout the follow-up period from infancy to 15 years of age may occur. To address the discussion about plausibility of catch-up growth in absolute height versus LAZ/HAZ, growth both in terms of absolute height gains and deficits measured in centimetres and changes in LAZ/HAZ was presented.

At one month of age, children were on average 3.5 cm shorter than the children in the WHO reference population. The deficit in absolute length and height increased until the children were four years old and remained at about 10–12 cm in both sexes. In girls, there was a slight catch-up after the age of 10, but a further decline in the boys (Figure 6.). The mean deficit in length was 9 cm for the girls and 15 cm for the boys at 15 years of age. Notable differences in the shape of the height trajectories were not found, neither in centimetres nor in z-scores between the children who, at the age of two, were moderately (> -3 SD but ≤ -2 SD) or severely (≤ -3 SD) stunted, compared to those who were not stunted (> -2 SD). This applied to both sexes. The difference in the mean LAZ between the non-stunted and the severely stunted groups was highest at the age of two, at which point it was 2.9 Z-scores for boys and 2.4 for girls. At the age of 15, the respective differences narrowed to 1.3 and 0.9. In contrast, the difference in mean height deficit between the children who were not stunted and those who were severely stunted at two years of age increased throughout childhood in the boys and reached 10.7 cm at the age of 15. In girls, the difference in height deficit was at its largest at the age of nine (7.5 cm) and decreased to 7 cm at the age of 15 (Figure 6.).

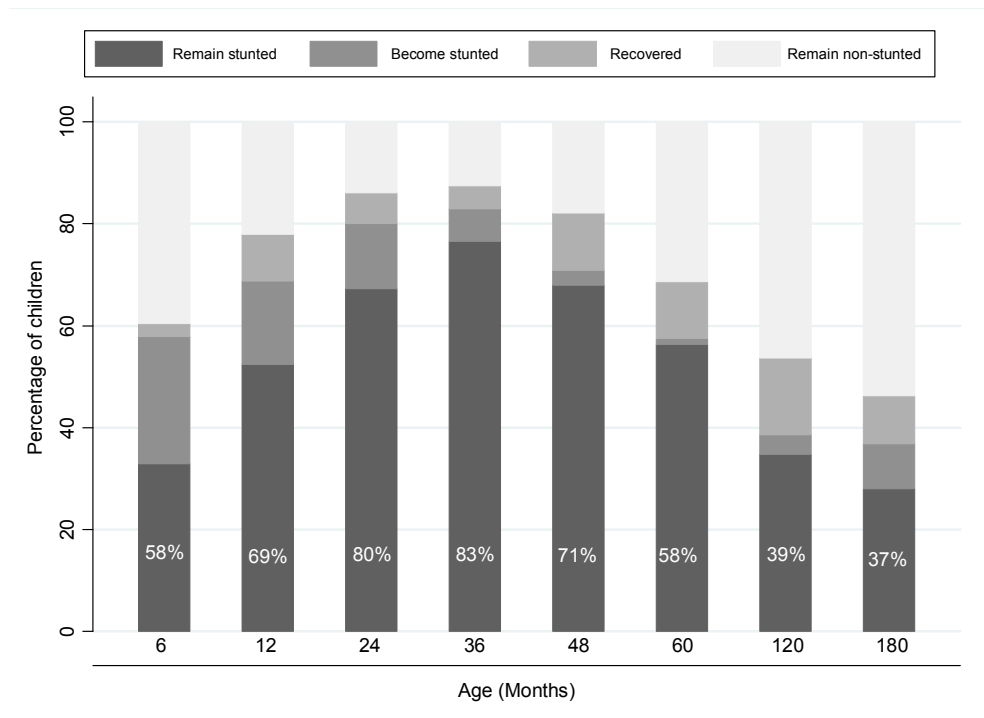
Figure 6. Linear growth among boys (n = 412) and girls (n = 394) in LCSS cohort. The solid lines indicate HAZ scores, and the dashed lines refer to the mean deficits in centimetres compared to the reference population (14,15). The growth curves are LOWESS-smoothed.



The transition status in stunting between age intervals is shown in Figure 8. A stricter definition than simply crossing the -2 SD cut-off was used to define a transition between stunting and non-stunting status to eliminate the effect of possible measurement inaccuracies. The participants who were classified as remaining stunted had LAZ/HAZ measurements of < -2 SD at the first measurement and LAZ/HAZ measurements of < -1.8 SD at the second measurement. The ones who became non-stunted made the transition from LAZ/HAZ < -2 SD to LAZ/HAZ > -1.8. Those who remained non-stunted had a LAZ/HAZ of > -2 SD at the first measurement and > -2.2 SD at the second. Those who became stunted made the transition from an LAZ/HAZ of > -2 SD to an LAZ/HAZ of < -2.2 SD. At between one and six months, 25.0% of the previously non-stunted infants became stunted and at between six and 12 months, 16.7%. The risk of stunting decreased along the age intervals. The proportion (95% CI) of stunted children (< -2 SD) was 80% (95% CI 76.5–83.5) when the children were two years old and declined to 37% (32.9–41.7) by the age of 15. When stunting was analysed every five years, the proportion of those who

became stunted was 21.3% (95% CI 17.5–25.1) between one and 60 months, 3.9% (2.2–5.6) between 60 and 120 months and 9.1% (6.5–11.7) between 120 and 180 months. The respective figures for recovering from stunting were 9.2% (95% CI 6.6–11.9), 15.0% (11.9–18.2) and 9.1% (6.5–11.7). Of the children who were moderately or severely stunted at the age of two years, 84.7% (95% CI 79.4–90.0) and 58.9% (53.0–64.8) were classified as non-stunted at least once during the rest of the follow-up.

Figure 7. Participants shifting to a different status of stunting or remaining in the same status between age intervals. The numbers on the bars represent the percentages of stunted participants at each time point.



By the age of 15, only 9.0% (95% CI 5.4–12.6%) of the boys and 19.6% (14.5–24.8%) of the girls had reached advanced puberty (Tanner Stages IV–V) as indicated by their pubic hair. There was a positive association between pubertal status and the absolute deficit in length when compared to the WHO reference population. For the adolescents in pubertal stage I at 15 years of age, the mean deficit in height was 21 cm for the boys and 13 cm for the girls. For those in pubertal stage V, the mean deficit was approximately 8 cm for both the boys and girls.

5.3 Association between early development and later school achievements (Article II)

There were 767 live born children in the cohort, of whom 415 (54%) were reached and had no missing data at 12 years of age. The association between development at five and school achievement at 12 years of age was assessed in the 415 cohort members. In Model 1, the summary score of child development was used as an exposure variable; in Model 2, the associations were assessed separately for each domain's percentage score and all four domain percentage scores simultaneously included in the analysis in Model 3. All the models were adjusted for potential confounders. The formation of percentage and summary scores are described above in the Statistical Analysis section. The outcome variables were mathematics score, highest grade completed and number of grades repeated. The background characteristics of the included participants were mostly similar compared to those who were excluded. There were slightly more preterm births (with difference of 6% and $p = 0.041$), lower mean weight-for-age near birth (with difference of 0.27 Z score and $p = 0.003$) and lower HAZ at five years of age (with difference of 0.34 Z score and $p = 0.001$) in the excluded group. In addition, the language domain score was lower in the excluded participants with an effect size (mean difference divided by SD) of about 0.22 SD ($p = 0.036$; Table 1 of Article II).

At 12 years of age, the majority (83%) of the children had attended school for at least one year, but only 4% completed more than four grades. Fifty percent of the children who had ever attended school had repeated a grade at least once. The mean (SD) of the percentage of correct mathematics questions was 36.4 (19.5). In Model 1, a positive association was found between the summary score of child development at five years of age and mathematics ability at 12 years of age (coefficient=1.77, $p=0.057$; Table 2 of Article II). MI analysis provided similar results (coefficient=1.84, $p=0.031$). In Model 2, fine motor score was positively associated with mathematics ability in both the observed (coefficient=0.41, $p=0.032$) and imputed data (coefficient=0.45, $p=0.011$). Models 2 and 3 indicated a positive but statistically non-significant association between each domain's percentage score and mathematics ability in most of the cases (Table 2 of Article II). The analyses of highest school grades and repeating did not reach statistical significance in either the actual or the imputed data (Tables 3 and 4 of Article II). In the abstract of Article II, the OR=0.834 is incorrectly reported as the p-value for the association between the

developmental summary score at 5 years of age and number of grades repeated. The correct p-value is 0.320.

5.4 Association between the childhood growth and cognitive function (Article III)

Cognitive function was assessed using Raven's Coloured Matrices and RT using the PVT and the mathematics test at 15 years of age. The mean (SD, range) values for the measures of cognitive function were 15 (6, 0–33) points with Raven's Coloured Matrices, 4 (2, 0–8) points (score) with the mathematics test, 408 (165, 229–2332) millisecond for median RT and 9 (8, 0–40) lapses in RT assessment.

The associations between height gain between one and 24 months and between 24 months and 15 years and cognitive function were assessed using three regression models. In Model 1, when LAZ/HAZ at the age of one month, 24 months and 15 years were included in the analysis, there was no statistically significantly association between height gains and Raven's Coloured Matrices score (Table 2 of Article III). In Model 2, when the regression was adjusted for the confounders, Raven's Coloured Matrices Score was positively predicted by HAZ at 15 years of age (coefficient=0.85, $p=0.03$), gender and mother's literacy, but not by HAZ at one or 24 months of age. In Model 3, when schooling was added to the adjusted analysis, schooling predicted Raven's Coloured Matrices score, but the positive association with HAZ at 15 years was weakened (coefficient=0.69, $P=0.06$; Table 2 of Article III). The average duration of school education was 2.7 years for girls and 2.6 years for boys. The Raven's Coloured Matrices score was three points lower in girls than boys ($P<0.001$). It was two points higher if the mother of the participant was literate ($P=0.02$) and one point higher with each additional year of school the participant had completed ($P<0.001$; Table 2 of Article III).

The mathematics score was statistically significantly associated only with the number of years of school education in Model 3. The mathematics score was on average 0.2 points better with each additional year completed at school ($P<0.001$; Table 3 of Article III). In Model 1, there were no statistically significant associations between LAZ/HAZ at any age and mathematics scores. In Model 2, there were statistically significant positive associations between the mathematics score, wealth index and mother's literacy.

The LAZ/HAZ and RT median did not reach statistical significance in Model 1 or in Model 2. In Model 3, schooling was positively and statistically significantly

associated with the RT median. Each additional year completed at school was associated with a 12 millisecond shorter median RT ($P=0.01$; Table 4 of Article III). In Models 2 and 3, boys had statistically significantly fewer RT lapses than girls (Table 5 of Article III). In Model 3, further adjusted for the numbers of completed years of school education, school education was statistically significantly associated with fewer RT lapses. Every additional year completed was associated with one fewer lapse ($P=0.01$).

In the analyses using growth by conditional growth measures (unexplained residuals) as predictors, the findings were similar compared to analyses using HAZ attained at ages 24 months and 15 years.

5.5 Association between the childhood growth and depressive symptoms (Article IV)

Out of the 538 participants in the 15-year assessments (Figure 5), 523 had data in one or more of the outcome variables and were included in the imputed primary analyses. Out of these 523 participants, 287 had missing values in one or more of the exposure or outcome variables. Thus, 236 participants with no missing values were included in the sensitivity analysis.

The socio-demographic baseline characteristics were substantially similar for the 523 included participants and for those who were excluded because of death, loss of follow-up or missing data (Table 1 of Article IV). The majority of the SMFQ assessments (77%) were interviewer-administered due to illiteracy in the participants. The rest self-administered the questionnaire.

In the 523 participants who were measured for anthropometrics at 15 years of age, the mean (SD, range) score of SMFQ was 15.0 (4.2, 0-26) points. The results were substantially similar for both sexes. Of all participants, 90% (95% CI 87%–92%; $n=458$) scored 11 or more points, suggesting clinically significant reported depressive symptoms. Corresponding figures for males and females separately were 92% (95% CI 88%–95%; 231) and 88% (95% CI 83%–91%; 227), respectively. Over 12% (95% CI 9%–15%; $n=62$) of the participants scored 20 or more points (Figure 2 of Article IV).

The mean LAZ/HAZ (SD) was -1.7 (1.2) at one month of age, -2.9 (1.2) at 24 months of age, -1.7 (1.0) at 120 months of age and -1.7 (0.8) at 15 years of age (Table 2 of Article IV). In the regression analyses with the imputed data, when birth weight and LAZ/HAZ at one month of age, 24 months of age, 120 months of age and 15

years of age were simultaneously included in the analysis (Model 1), birth weight (in kilograms) was not statistically significantly associated with SMFQ score (coefficient=-0,87, $p=0.133$). Height gains during different time periods were not statistically significantly associated with SMFQ score in Model 1. Similar results were found in Model 2, adjusted for the potential confounders, and in Model 3, further adjusted for Raven's Coloured Matrices and pubertal maturity. In addition, pubertal maturity or sex were also not statistically significantly associated with depressive symptoms (Table 3 of Article IV).

In the sensitivity analyses with the completely observed data ($n=236$), birth weight (in kilograms) was statistically significantly associated with SMFQ score in Model 1 (coefficient= -1.58, $p=0.014$), Model 2 (coefficient= -1.45, $p=0.034$) and Model 3 (coefficient= -1.38, $p=0.048$). Taller children had lower SMFQ scores, with a coefficient of -0.30 ($p=0.235$) at 24 months and -0.61 ($p=0.155$) at 120 months of age. By contrast taller children had a higher SMFQ score (coefficient 0.68, $p=0.044$) at 180 months of age. This association diluted and other results did not change when possible confounders and intermediate variables were included in the model (Models 2 and 3). Pubertal maturity and gender were not statistically significantly associated with depressive symptoms. The mode of questionnaire administration did not affect the results (Table III of Article IV).

6 DISCUSSION

In this thesis, it was demonstrated that both stunting and recovery from stunting occurred at all ages in rural Malawi when measured with LAZ/HAZ. There was a positive association between the summary score of child development at five years of age and mathematics ability at 12 years, and a positive association between height gain between 24 months and 15 years and cognitive capacity at 15 years of age. The prevalence of reported depressive symptoms was high among both boys and girls at 15 years of age. Lower birth weight was associated with more frequent reported depressive symptoms in the participants who had no missing values in the exposure or outcome variables, but this association was not found in the imputed data.

A significant proportion of the participants who were stunted at two years of age became non-stunted at some point. These findings most likely indicate plasticity in growth patterns throughout the follow-up period. Boys tended to be pre-pubertal at 15 years of age, and the more delayed they were, the lower their HAZ scores were. The absolute deficit in length increased until the children were four years old and plateaued at around 10–12 cm in both sexes. After the age of ten, there was a small catch-up in centimetres in the girls, but a further decline in the boys. This difference seems to be due to delayed puberty in the boys.

The discussion about the timing of stunting and catch-up growth has recently been active. The belief that catch-up in linear growth is rare after two years of age in low-income countries (5,6,111), has been challenged (3,88,112). Previously, catch-up growth was mostly investigated in terms of conditional height, i.e. HAZ, and this has raised some criticism (3,4,88). The calculation of Z-scores is explained in detail in the Methods section. Standard deviation is the denominator in the equation, and increases parallel to increasing mean height. Therefore, if the absolute length / height deficit is constant and a negative value, the HAZ will increase over time because the denominator, standard deviation, increases with age. Correspondingly, the relative deficit (Z-score) may decrease even if the absolute deficit in centimetres remains constant over time (88). This mathematical fact does not exclude HAZ from being a useful tool in measuring child growth. A risk factor may at one age induce a deficit in absolute height in centimetres, but if the adverse exposures are not repeated at subsequent ages, the relative deficit will decrease over time as each risk factor

becomes less important in a longer history of cumulative risk exposures. Nutritional and other health-promoting interventions may not be able to remove the harm already done, but they may have an impact on preventing consequences of further exposures to cumulative risk factors (113).

A positive association between the summary score of child development at five years of age and mathematics ability at 12 years of age was demonstrated in rural Malawi. This association was also found after adjustment for possible confounders such as stunting in early childhood. The summary score derived at five years of age from the gross motor, fine motor, language and social domains – rather than a single aspect of developmental impairment – was associated with later schooling outcome. This supports previous findings about the harmful effects of wide variety of different risk factors on schooling. The number of risk factors seem to be more robust predictor of outcomes than any single risk factor (43,114). The analysis was adjusted for most potential confounders and therefore presented more precise estimates of the associations. The adverse exposures behind the multiple risk factors for poor development and schooling outcomes are, among others, poverty, malnutrition, poor health and lack of stimulating environment (43). The same exposures compromise childhood linear growth, which was found in the present study to be associated with later cognitive performance. This positive association was partly mediated by schooling. The interrelationship between poverty, growth, development, schooling and cognition is complex, and the effects are multidirectional.

This thesis showed a statistically significant positive association between height gain between 24 months and 15 years and cognitive capacity assessed using Raven's Coloured Matrices score, but not mathematics score or RT median or lapses. Male sex, mother's literacy, higher wealth index and more years of school education were associated with better success in Raven's Coloured Matrices score, mathematics test and RT median and lapses at 15 years of age. Similar findings in both analyses, with HAZ scores and conditional growth measures, indicate that the results were not affected by collinearity of the exposure variables. In addition, the correlation coefficients between exposure variables HAZ at one, 24 and 180 months of age were < 0.46 , showing only a moderate correlation between them. The positive association between height gain after 24 months of age and cognitive capacity at the age of 15 appeared to be partly mediated by schooling. This could indicate that, despite criticism, the Malawian schooling system may have at least some effect (115).

Cognitive function measured using Raven's Coloured Matrices was predicted by height gain between 24 months and 15 years. This finding is in line with that from

the Young Lives study, in which catch-up growth was found to occur between eight and 15 years of age. Those who caught up also performed better in cognitive function tests (116). Unlike this multicentre study, most of the longitudinal and cross-sectional studies conducted, have mainly focused on the impact of early, but not later, childhood growth (37,43,52,67,68,117). Based on the findings in this study, it may be suggested that the focus on the critical first 1000 days of life should not limit efforts to improve the health environment in later childhood and to extend the intervention programmes to older age groups.

There was a direct positive association between the pubertal status and HAZ or the absolute deficit in length in both sexes. The onset of puberty was delayed, especially in boys. During puberty, growth velocity is at its highest in boys at pubic hair stage IV and in girls at stage III (118). Only 9% of the boys and 52% of the girls in our cohort had reached those stages at 15 years of age. The delayed pubertal development may explain the fact that the boys suffered from more persistent growth faltering than the girls. In contrast, the small catch-up in absolute height deficit in girls may have been due to the pubertal growth spurt. Based on these findings, it is proposed that there is a possibility for significant catch-up growth during puberty in both sexes, but the significance of growth failure and timing of pubertal development deserves further discussion.

The distribution of the SMFQ scores suggests high prevalence of reported depressive symptoms. This finding was similar in both sexes. Previously, in Brazil, lower birth weight and stunting at two or four years of age were associated with more depressive symptoms (53). In Jamaica, stunting before 24 months of age predicted more depressive symptoms at 17 years of age (119). However, the significance of later childhood growth on depressive symptoms is not known. In this thesis, lower birth weight was associated with more frequent reported depressive symptoms among the participants who had no missing values in the exposure or outcome variables. This negative association was consistent when possible confounders and intermediate variables were added in the model. The association was not found in the imputed data analysis. Statistically significant associations between height gains until 15 years of age, gender, pubertal maturity at the time of the assessment and reported depressive symptoms at 15 years of age were not found in either of the analyses.

The incidence of depression increases during adolescence in Western countries (120,121), but evidence from the developing world is scarce. An unexpectedly high prevalence of reported depressive symptoms was presented in our rural Malawian cohort, measured using the SMF-questionnaire. If the traditional cut-off for clinically

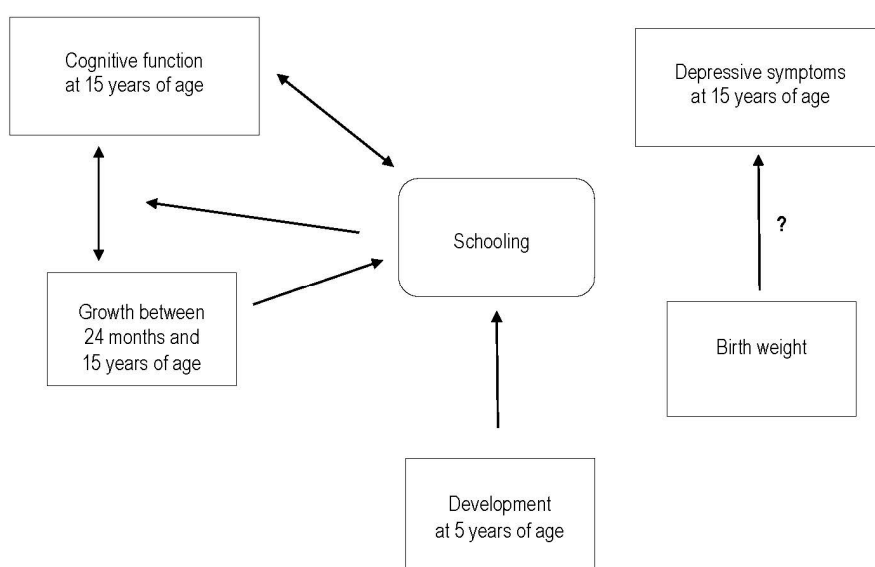
significant depressive symptoms (≥ 11) was used, approximately 90% (95% CI 87%–92%) of all participants would have suffered from this condition. The SMFQ is not validated in our study setting. The questionnaire was, however, carefully translated and translated back into its original language, and the individual items may be mainly considered unambiguous. The SMFQ is widely used, also in low-income settings and adolescents (87,122-124). It has been considered a reliable and valid instrument for the assessment of depressive symptoms in adolescents in Bangladesh (125). This study could not, however, either establish a cut-off score for clinically significant depressive symptoms in their setting. The SMFQ was designed as a screening tool, and the cut-off was set low enough to find all possible depression cases for further evaluation. However, in our study, even when the cut-off was set as high as 20 points, the percentage of the participants suffering from significant reported depressive symptoms was 12%. It has previously been estimated that 14% of children in sub-Saharan Africa suffer from psychological disorders of some kind (8,10). By comparison, a 12% prevalence of only depressive symptoms, which is estimated based on the SMFQ results in our study, is rather high.

The high prevalence of reported depressive symptoms could not be explained in this study cohort. A genetic predisposition to the disease, social disadvantage, parental mental disorders, female gender and experiences of violence are generally related to depression (8). In addition, low birth weight, growth failure and malnutrition have previously found to contribute to later depressive symptoms (53,119,126). In this study, lower birth weight was associated with more frequent reported depressive symptoms in the complete data analysis, but not in the imputed data analysis. Other than that, no associations between height gain in childhood, gender, pubertal maturity at the time of the assessment and reported depressive symptoms were not found at 15 years of age. It has been suggested that the SMFQ may overestimate the prevalence of depressive symptoms in individuals with lower educational attainment, but the overestimation in a tool has not been found to be this high (87). It is possible that the relevant determinants of depressive symptoms may be different in a rural African setting, compared to those in Western countries. For instance, orphanhood, food insecurity and death in the family may be overrepresented in this population (127,128).

Based on the literature review and the studies included in this thesis, it may be concluded that catch-up -growth in terms of HAZs during childhood may occur until adolescence. The timing of pubertal onset is important when evaluating the possibility for later growth potential. Early growth failure affects adolescent health, including cognitive capacity. The literature review showed associations between early

growth and mental health, but, in this study, only birth weight, not childhood growth, was associated with depressive symptoms. Since catch-up growth is possible in later childhood, and enhanced growth leads to healthier life, it can be suggested that the focus on the first 1000 days in growth promotion should not limit efforts to improve the nutritional environment for older children. Mental health awareness in low-income countries is poor and requires urgent attention. Figure 7 summarizes the associations that were found in this study.

Figure 8. The associations between growth, development, schooling and depressive symptoms in LCSS children.



6.1 Strengths and limitations of the study

Sample attrition is a common problem in long-term follow-up studies. In Article I, 68% of the cohort members were available, did not have missing data and were included in the analysis. The corresponding figures for the subsequent articles were 54%, 47% and 36%. In Article IV, the primary analysis was conducted using the multiple imputation (MI) method to impute missing data. Sensitivity analysis based on MI was also included in Article II, and the results were similar with both methods. The main strength of the multiple imputation (MI) method lies in it allowing the

efficient handling of missing data and uncertainty in imputations. It is relatively easy to implement and is appropriate for a wide range of data sets (102). In Article I, in the sensitivity analysis, the results were similar to those obtained with full data and with the data of the participants who remained in the study until the end of the follow-up period. The main reason of the exclusion in our study was not a large drop-out rate but, rather, death prior to age 15. Approximately 23% of the children born live died before the end of the follow-up.

Our study sample represented the target population well in all the sub-studies. The study cohort originally comprised 97% of all newborn infants in the study area during the time of enrolment. Random errors in anthropometry measurements were minimized using the mean of three individual measurements and the measurements were assessed by regularly trained data collectors. In addition, the same equipment was used for all the participants at each age, and the equipment was regularly calibrated.

The developmental assessment inventory that was used at five years of age was developed with local research workers to make the milestones culturally appropriate in Malawi (65). The mathematics test used at 12 years of age was not formally validated. However, it was modified from an instrument used in a large-scale Indonesian study that included assessment of children at same age. It had also been previously shown to be associated with height gain in children (69), which is a known predictor of cognitive ability, measured with several different developmental assessments (43).

The reliability and criterion validity of Raven's Matrices have been found to be good in Africa (129,130). In Article III, I proposed that the Raven's Matrices score, the mathematics test score and the RT median and lapses were associated with schooling, which also supports their criterion validity. There is, however, a concern that Raven's Matrices may not measure 'general intelligence' as intended in Africa, and further research is needed to clarify this (129). The validity and reliability of the RT test is relatively well documented (96,131,132). Because of difficulties in handling written tasks among the cohort members, participants were instructed to point out the correct answer in the Raven's Coloured Matrices test with their forefinger. This may have caused some problems in the validity of the test. To minimize the effect of the administration of the test, the data collectors were thoroughly trained to perform the assessment, and the quality of the measurements was regularly monitored.

The SMFQ has not been validated in our study setting, and we did not use the parent/caregiver version of the questionnaire. It was mostly assessed by an interview

instead of self-administration because of high illiteracy rate in the cohort. It is unlikely that the administration mode affected our finding – a high prevalence of reported depressive symptoms – since interviews tend to underestimate, not overestimate, the prevalence of mental health problems (133). However, the questionnaire was carefully translated and then translated back into the original language until the translations corresponded to each other, and the SMFQ is widely used, also in low-income settings and with adolescents (87,122-125).

7 KEY FINDINGS AND CONCLUSIONS

The key findings of this thesis are summarized as follows.

1. It is possible that stunting and recovery from stunting both occur throughout the period from birth until adolescence. 80% of the children were stunted at two years of age.
2. Child development at 5 years of age may be positively associated with later schooling outcomes.
3. Height gain between 2 and 15 years of age is likely to be associated with better cognitive capacity in adolescence, partly due to more school education. In the light of these results, growth promotion should not only be limited to early childhood.
4. The prevalence of reported depressive symptoms may be higher than previously thought in rural low-income adolescent populations. It is possible that a low birth weight contributes to later reported depressive symptoms. Our results highlight the importance of awareness of mental health problems in low-income countries.

8 FUTURE RESEARCH

In this study, the follow-up period ended at 15 years of age. A large part of the cohort members did not reach full puberty by the end of the study, suggesting the possibility of further growth after the age of 15. In future studies, the follow-up period should be extended until the end of adolescent growth in order to investigate the final absolute and relative height. This would also provide information about the timing of both these factors, later growth and pubertal development, and any associations. The possibility of later catch-up growth could provide an additional window of opportunity for nutritional intervention. There is a lack of evidence about the potential harm or benefit of late adolescent interventions, and this important matter should be examined.

Some of the tools used in this study were not validated in a rural African setting. The Short Mood and Feelings Questionnaire or corresponding questionnaires should be validated. It would, in addition, be useful to study the validity and feasibility of Raven's Coloured Matrices in rural low-income populations. Raven's Standard Matrices, which should have been a suitable assessment in this age group, could not be used in this study. In the pilot study, this questionnaire gave low scores with no variability. Instead, Raven's Coloured Matrices, which is designed for smaller children, was used. Compared with scores in previous studies, the scores in this cohort were also quite low. Such poor success in a test measuring general intelligence in the cohort was surprising and should be reassessed, possibly with another tool. The determinants behind the impaired cognitive development in this setting should be investigated.

The prevalence of reported depressive symptoms was high in our cohort. Low birth weight may contribute to higher depressive symptoms, but, other than that, the causes of this finding could not be defined. The study should be repeated with a locally validated tool, and efforts should be made to understand the causes of high prevalence of depressive symptoms in rural African adolescents.

This was a prospective cohort study with no interventions. Based on the study, it may be suggested that catch-up growth is possible also after the traditionally considered 1000 days of life, and actions in nutritional programmes may be proposed

until the end of the growth period. It would be useful to examine this suggestion with an interventional study.

9 ACKNOWLEDGEMENTS

This study was carried out in Tampere Center for Child Research: Global Health group, University of Tampere, Faculty of Medicine and Life Sciences. The work was supported by a grant from the Finnish Academy. I also received personal grants from Finnish Cultural Foundation, Häme Regional Fund, Foundation for Pediatric Research, Organization for Respiratory Health in Finland and Päivikki and Sakari Sohlberg Foundation. In addition, I was greatly supported by The National School of Clinical Investigation and Pediatric Postgraduate School, University of Helsinki.

There are many people who have supported me during this study. It is impossible to thank all of you individually. There are some who, however, deserve to be especially mentioned.

Professor Per Ashorn, who has patiently guided me through this journey. He deserves an official recognition of excellence for being the most tireless, uncompromising and fatherly supervisor. He always had the correct answers or the wisdom to make me find them. He has been the best possible scientific rock to lean on, and he has always been there for me. I also want to thank him and Ulla for their hospitality during these years. It has made this journey much more than just tough work.

Professor Yin Bun Cheung for being such a solid and brilliant statistical advisor. His extremely fast and thorough feedback kept me going through the challenges. He introduced me to a whole new world of statistics, which I am very grateful for. With his support, I never had to question if I was going to make it.

The preliminary examiners, Helena Lapinleimu and Eero Kajantie. Thank you for the bracing and instructive conversations, and thorough work that you did to improve my thesis. Dee Shields and Matthew James for language editing. The advisory board, prof. Matti Korppi, Peter Csónka and Lauri Lehtimäki, the meetings with you gave me inspiration and guided me to the correct direction. The staff of Lungwena Child Survival Study team. Your professionalism and joy of learning new things. You did most of the work for me. All the families and children who participated LCSS study. Without you this would of course not have been possible.

Global Health family, Anna, Ulla H, Ulla A, Yuemei, Basho, Lotta, Chiza, Kirsi, Chrissie, Charles, Heli, Austrida, Jaimie, Emma and many other friends in Finland

and Malawi. Without your friendship and peer support the work would have been so much more difficult. I also want to especially thank Anna, Ulla H, Emma and Lotta for helping me with the expected and unexpected struggles in Lungwena.

My colleagues at Helsinki University Children's Hospital. I thank all of you for your friendship. I also want to thank the Head of Pediatrics Anne Wikström and professor Markku Heikinheimo for giving me the opportunity to also concentrate on my research work during my pediatric residency.

My dearest friends from Medical School. Elina, Mitja, Lenni, Marjo, Noora and Teemu and your lovely spouses. We started a journey in medicine together, and all of us keep going strong. Your friendship is invaluable. The hilarious Team 8, friends from childhood, with you it is impossible to forget the essentials of life. I also want to thank the godfathers and godmothers of our children. Elina, Jarkko, Jaakko, Mikko, Roope, Katja, Janne, Anna-Reetta, Eetu, Ville and Satu. Thank you for being close to us and who you are.

My mother, who has been the best of help during these years both at home and in Malawi. I never could have done this without you. My brother Ville, whose keen-sighted comments are an endless source inspiration. My late dad, whose kindness, wisdom and billions of answers to my never-ending questions made me who I am. You are my hero and I miss you every day.

My deepest gratitude goes to my family. Timo, you have never questioned my decisions and have fearlessly trusted on me. Your both mental and practical support has been amazing enough to get me this far. You made me believe in myself and gave me the possibility to make my scientific dream come true. Thank you for being an efficient housewife both in Mangochi and Kytäjä. You are pretty good in that. I love you. Linnea, Louna and Edith, there are no words to describe my feelings for you. I love you to the moon and back.

10 REFERENCES

1. United Nations. The state of the world's children 2016: A fair chance for every child. United Nations Children's Fund, UNICEF, 2016.
2. United Nations. Transforming our world: The 2030 agenda for sustainable development. Human 2015.
3. Prentice AM, Ward KA, Goldberg GR, Jarjou LM, Moore SE, Fulford AJ, et al. Critical windows for nutritional interventions against stunting. *American Journal of Clinical Nutrition* 2013; 97: 911.
4. Lundeen EA, Stein AD, Adair LS, Behrman JR, Bhargava SK, Dearden KA, et al. Height-for-age z scores increase despite increasing height deficits among children in 5 developing countries. *American Journal of Clinical Nutrition* 2014; 100: 821-5.
5. Victora CG, de Onis M, Hallal PC, Blossner M, Shrimpton R. Worldwide timing of growth faltering: Revisiting implications for interventions. *Pediatrics* 2010; 125: e473-80.
6. Shrimpton R. Worldwide timing of growth faltering: Implications for nutritional interventions. *Pediatrics* 2001; 107: E75.
7. World Health Organization. The global burden of disease 2004 update. World Health Organization, 2008.
8. Patel V, Flisher A, Hetrick S, McGorry P. Adolescent health 3 - mental health of young people: A global public-health challenge. *Lancet* 2007; 369: 1302-13.
9. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS medicine* 2006; 3: e442.
10. Cortina MA, Sodha A, Fazel M, Ramchandani PG. Prevalence of child mental health problems in Sub-Saharan Africa: A systematic review. *Arch of Pediatrics and Adolescent Medicine* 2012; 166: 276-81.
11. Weissman MM, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG, et al. Cross-national epidemiology of major depression and bipolar disorder. *JAMA* 1996; 276: 293.
12. Thapar A, Stephan Collishaw, Daniel S Pine, Ajay K Thapar. Depression in adolescence. *Lancet* 2012; 379: 1056-67.
13. Kiserud T, Piaggio G, Carroli G, Widmer M, Carvalho J, Jensen L, et al. The world health organization fetal growth charts: A multinational longitudinal study of ultrasound biometric measurements and estimated fetal weight. *PLoS medicine* 2017; 14: e1002220.
14. WHO Multicentre Growth Reference Study Group. 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. Geneva: World Health Organization. 2006.
15. De Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bulletin of the World Health Organization* 2007; 85: 660-667.

16. Sudfeld CR, McCoy DC, Danaei G, Fink G, Ezzati M, Andrews KG, et al. Linear growth and child development in low- and middle-income countries: A meta-analysis. *Pediatrics* 2015; 135: e1266-75.
17. Black RE, Allen LH, Bhutta ZA, Caulfield LE, de Onis M, Ezzati M, et al. Maternal and child undernutrition: Global and regional exposures and health consequences. *Lancet* 2008; 371: 243-60.
18. Walli N, Munubhi E, Aboud S, Manji K. Vitamin D levels in malnourished children under 5 years in a tertiary care center at Muhimbili National Hospital, Dar es Salaam, Tanzania-A cross-sectional study. *Journal of Tropical Pediatrics* 2017; 63: 203-9.
19. United Nations Children's Fund, World Health Organization, The World Bank. UNICEF WHO-World Bank joint child malnutrition estimates 2012.
20. Onis M, Branca F. Childhood stunting: A global perspective. *Maternal & Child Nutrition* 2016; 12: 12-26.
21. Geneva: World Health Organization. WHO 1995 expert committee report: Physical status: The use and interpretation of anthropometry. 1995; Technical Report Series 854.
22. de Onis M, Blössner M. The World Health Organization global database on child growth and malnutrition: Methodology and applications. *International Journal of Epidemiology* 2003; 32: 518-26.
23. Advancing Partners & Communities. Country profile: Malawi community health programs. 2014.
24. Ministry of Health (MoH) [Malawi] and ICF International. Malawi service provision assessment (MSPA) 2013-14. 2014.
25. de Onis M, Wijnhoven TMA, Onyango AW. Worldwide practices in child growth monitoring. *Journal of Pediatrics* 2004; 144: 461-5.
26. Pasricha S, Biggs B. Undernutrition among children in South and South-East Asia. *Journal of Paediatrics and Child Health* 2010; 46: 497-503.
27. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley, 1987.
28. Demirchyan A, Petrosyan V, Sargsyan V, Hekimian K. Predictors of stunting among children ages 0 to 59 months in a rural region of Armenia. *Journal of Pediatric Gastroenterology and Nutrition* 2016; 62: 150-6.
29. Tanner JM. Foetus into man: Physical growth from conception to maturity. Ware: Castlemead Publications, 1989.
30. Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *The Lancet* 2009; 374: 1032-5.
31. Veitch AM, Kelly P, Zulu IS, Segal I, Farthing MJG. Tropical enteropathy: A T-cell-mediated crypt hyperplastic enteropathy. *European Journal of Gastroenterology and Hepatology* 2001; 13: 1175-81.
32. Sullivan PB, Marsh MN. Small intestinal mucosal histology in the syndrome of persistent diarrhoea and malnutrition: A review. *Acta paediatrica* (Oslo, Norway:1992). Supplement. 1992; 381: 72-7.
33. Crane R, Berkley J. Progress on growth faltering. *Lancet Global Health* 2017; 5: E125-6.
34. Crane RJ, Jones KDJ, Berkley JA. Environmental enteric dysfunction: An overview. *Food and Nutrition Bulletin* 2015; 36: S76-87.
35. Ali A, Iqbal NT, Sadiq K. Environmental enteropathy. *Current Opinion in Gastroenterology* 2016; 32: 12-7.
36. Unicef. Strategy of improving nutrition on children and women in developing countries. UNICEF 1990.

37. Piwoz E, Sundberg S, Rooke J. Promoting healthy growth: What are the priorities for research and action? *Advances in Nutrition* 2012; 3: 234-41.
38. Crookston BT, Schott W, Cueto S, Dearden KA, Engle P, Georgiadis A, et al. Postinfancy growth, schooling, and cognitive achievement: Young lives. *American Journal of Clinical Nutrition* 2013; 98: 1555-63.
39. Prader A, Tanner JM, von Harnack GA. Catch-up growth following illness or starvation: An example of developmental canalization in man. *Journal of Pediatrics* 1963; 62: 646-59.
40. Caballero B. Introduction. Symposium: Obesity in developing countries: Biological and ecological factors. *Journal of Nutrition* 2001; 131: 866S.
41. Uauy R, Kain J, Mericq V, Rojas J, Corvalán C. Nutrition, child growth, and chronic disease prevention. *Annals of Medicine* 2008; 40: 11-20.
42. Wells JCK, Hallal PC, Wright A, Singhal A, Victora CG. Fetal, infant and childhood growth: Relationships with body composition in Brazilian boys aged 9 years. *International Journal of Obesity* 2005; 29: 1192-8.
43. Victora CG, Sibbritt D, Horta BL, Lima RC, ScD TC, Wells J. Weight gain in childhood and body composition at 18 years of age in Brazilian males. *Acta Paediatrica* 2007; 96: 296-300.
44. Grantham-McGregor S, Cheung YB, Cueto S, Glewwe P, Richter L, Strupp B. Developmental potential in the first 5 years for children in developing countries. *Lancet* 2007; 369: 60-70.
45. Berkman DS, Lescano AG, Gilman RH, Lopez SL, Black MM. Effects of stunting, diarrhoeal disease, and parasitic infection during infancy on cognition in late childhood: A follow-up study. *Lancet* 2002; 359: 564-71.
46. Walker SP, Chang SM, Powell CA, Grantham-McGregor S. Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: Prospective cohort study. *Lancet* 2005; 366: 1804-7.
47. Walker SP, Wachs TD, Meeks Gardner J, Lozoff B, Wasserman GA, Pollitt E, et al. Child development: Risk factors for adverse outcomes in developing countries. *Lancet* 2007; 369: 145.
48. Schuchardt JP, Huss M, Stauss-Grabo M, Hahn A. Significance of long-chain polyunsaturated fatty acids (PUFAs) for the development and behaviour of children. *European Journal of Pediatrics* 2010; 169: 149-64.
49. Cordero ME, D'Acuña E, Benveniste S, Prado R, Nuñez JA, Colombo M. Dendritic development in neocortex of infants with early postnatal life undernutrition. *Pediatric Neurology* 1993; 9: 457.
50. Thompson RA. Developmental science and the media. Early brain development. *American Psychologist* 2001; 56: 5-15.
51. Hadders-Algra M. Variability in infant motor behavior: A hallmark of the healthy nervous system. *Infant Behavior and Development* 2002; 25: 433-51.
52. Victora CG, Adair L, Fall C, Hallal PC, Martorell R, Richter L, et al. Maternal and child undernutrition: Consequences for adult health and human capital. *Lancet* 2008; 371: 340-57.
53. Walker SP. Child development: Risk factors for adverse outcomes in developing countries. *Lancet* 2007; 369: 145-57.

54. Christian Loret de Mola, Luciana de Avila Quevedo, Pinheiro RT, Gonçalves H, Gigante DP, Vieira dos Santos Motta, et al: The effect of fetal and childhood growth over depression in early adulthood in a southern Brazilian birth cohort: e0140621. *PLoS One* 2015; 10.
55. Rutter M. The role of cognition in child development and disorder. *British Journal of Medical Psychology* 1987; 60 (Pt 1): 1.
56. Fernald LC, Grantham-McGregor SM. Stress response in school-age children who have been growth retarded since early childhood. *American Journal of Clinical Nutrition* 1998; 68: 691-8.
57. Prendergast A, Humphrey J. The stunting syndrome in developing countries. *Paediatrics and International Child Health* 2014; 34: 250-65.
58. Rodríguez L, Cervantes E, Ortiz R. Malnutrition and gastrointestinal and respiratory infections in children: A public health problem. *International Journal of Environmental Research and Public Health* 2011; 8: 1174-205.
59. Özaltın E, Hill K, Subramanian SV. Association of maternal stature with offspring mortality, underweight, and stunting in low- to middle-income countries. *JAMA* 2010; 303: 1507-16.
60. Barker D. In utero programming of chronic disease. *Clinical Science* 1998; 95: 115-28.
61. Bellman M, Byrne O, Sege R. Developmental assessment of children. *BMJ: British Medical Journal* 2013; 346: 31-5.
62. Bayley N. Bayley scales of infant development manual 1993; 2nd edition. San Antonia: Harcourt Brace, 1993
63. Griffiths R. The abilities of young children A Comprehensive System of Mental Measurement for the First Eight Years of Life. ARICD, 1984
64. McCarthy D. McCarthy scales of children's abilities. New York: Psychological Corporation, 1972.
65. Frankenburg WK, Dodds J, Archer P, Shapiro H, Bresnick B. The Denver II: A major revision and restandardization of the Denver developmental screening test. *Pediatrics* 1992; 89: 91-7.
66. Gladstone MJ. Can western developmental screening tools be modified for use in a rural Malawian setting? *Archives of Disease in Childhood* 2008; 93: 23-9.
67. Grantham-McGregor S. Linear growth retardation and cognition. *The Lancet* 2002; 359: 542.
68. Perignon M, Fiorentino M, Kuong K, Burja K, Parker M, Sisokhom S, et al. Stunting, poor iron status and parasite infection are significant risk factors for lower cognitive performance in Cambodian school-aged children. *PLoS One* 2014; 9: e112605.
69. Crookston BT, Dearden KA, Alder SC, Porucznik CA, Stanford JB, Merrill RM, et al. Impact of early and concurrent stunting on cognition. *Maternal & Child Nutrition* 2011; 7: 397-409.
70. Gandhi M, Ashorn P, Maleta K, Teivaanmäki T, Duan X, Cheung YB. Height gain during early childhood is an important predictor of schooling and mathematics ability outcomes. *Acta Paediatrica* 2011; 100: 1113-18.
71. American Psychiatric Association. DSM-5 Task Force, American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. Washington (D.C.): American Psychiatric Publishing, 2013.
72. Carroll J. Human cognitive abilities: A survey of factor-analytic studies // review. 1993; 47: 763. Canadian Psychological Association. Ottawa.

73. Kaufman, A.S., Kaufman, N.L. K-ABC: Kaufman assessment battery for children. American Guidance Service, Circle Pines, Minneapolis, 1983
74. Wechsler D. Wechsler intelligence scale for children-fifth edition. San Antonio, TX: NCS Pearson, 2014.
75. Raven J, Raven J, Court J. Manual for raven's progressive matrices and vocabulary scales. section 2: The coloured progressive matrices. San Antonio, TX: Harcourt Assessment 1998.
76. Stanford-binet intelligence scale. *Oxford University Press* 2008.
77. Horn JL, Cattell RB. Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology* 1966; 57: 253-70.
78. Vernon PA. Speed of information processing and general intelligence. *Intelligence* 1983; 7: 53-70.
79. Lewinsohn PM, Lewinsohn PM, Clarke GN, Clarke GN, Seeley JR, Seeley JR, et al. Major depression in community adolescents: Age at onset, episode duration, and time to recurrence. *Journal of the American Academy of Child and Adolescent Psychiatry* 1994; 33: 809-18.
80. Klein D. Subthreshold depressive disorder in adolescents: Predictors of escalation to full-syndrome depressive disorders. *Journal of the American Academy of Child and Adolescent Psychiatry* 2009; 48: 703-10.
81. Olino TM, Yu L, Klein DN, Rohde P, Seeley JR, Pilkonis PA, et al. Measuring depression using item response theory: An examination of three measures of depressive symptomatology. *International Journal of Methods in Psychiatric Research* 2012; 21: 76-85.
82. Beck AT, Steer RA and Brown GK. Manual for the beck depression inventory-II. San Antonio, TX: Psychological Corporation, 1996.
83. Radloff LS. The CES-D scale. *Applied Psychological Measurement* 1977; 1: 385-401.
84. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine* 2001; 16: 606-13.
85. Allgaier A, Pietsch K, Frühe B, Sigl-Glöckner J, Schulte-Körne G. Screening for depression in adolescents: Validity of the patient health questionnaire in pediatric care. *Depression and Anxiety* 2012; 29: 906-13.
86. Kovacs M. The children's depression inventory (CDI). *Psychopharmacology Bulletin* 1985; 21(4): 995-998.
87. Angold A, Costello E, Messer S, Pickles A, Winder F, Silver D. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research* 1995; 5: 237-49.
88. Turner N, Joinson C, Peters T, Wiles N, Lewis G. Validity of the short mood and feelings questionnaire in late adolescence. *Psychological Assessment* 2014; 26: 752-62.
89. Leroy JL, Ruel M, Habicht JP. Critical windows for nutritional interventions against stunting. *American Journal of Clinical Nutrition* 2013; 98: 854-5.
90. //Www.gapminder.org/data/
91. UNDP. Human development report 2016.
92. Kulmala T, Vaahtera M, Rannikko J, Ndekha M, Cullinan T, Salin M et al. The relationship between antenatal risk characteristics, place of delivery and adverse delivery outcome in rural Malawi: Maternal risks and place of delivery. *Acta Obstetrica Gynecologica Scandinavica* 2000; 79: 984-90.
93. Tanner J. Growth at adolescence. Oxford: Blackwell Scientific Publications, 1962.
94. Carel J, Léger J. Precocious puberty. *New England Journal of Medicine* 2008; 358: 2366-77.

95. Strauss J, Beegle K, Sikoki B, Dwiyanto A, Herawati Y, Witoelar F. The third wave of the Indonesia family life survey (IFLS3): Overview and field report. WR-144/1-NIA/NICHD. 2004; March.
96. Gandhi M, Teivaanmaki T, Maleta K, Duan X, Ashorn P, Cheung YB. Child development at 5 years of age predicted mathematics ability and schooling outcomes in Malawian adolescents. *Acta Paediatrica* 2013; 102: 58-65.
97. Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers* 1985; 17: 652-655.
98. Cattell RB. Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology* 1963; 54: 1-22.
99. Raven J. The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology* 2000; 41: 1-48.
100. Raven J, Raven J, Court J. Manual for raven's progressive matrices and vocabulary scales. section 3: The standard progressive matrices. San Antonio, TX: Harcourt Assessment 2000.
101. Wenger MJ, Townsend JT. Basic response time tools for studying general processing capacity in attention, perception, and cognition. *Journal of General Psychology* 2000; 127: 67-99.
102. Sheppard LD, Vernon PA. Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences* 2008; 44: 535-51.
103. Royston P. Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables. *Stata Journal* 2009; 466-77.
104. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine* 2011; 30: 377-99.
105. Johnson RA WD. Applied multivariate statistical analysis. Upper Saddle River, NJ: Prentice Hall 2002.
106. DiStefano C, Zhu M, Mindrila D. Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation* 2009; 14: 1-11.
107. Filmer D, Pritchett LH. Estimating wealth effects without expenditure data –or tears: An application to educational enrollments in states of India. *Demography* 2001; 38: 115-32.
108. Kulmala T, Vaahtera M, Ndekha M, Koivisto A-, Cullinan T, Salin M-, et al. The importance of preterm births for peri- and neonatal mortality in rural Malawi. *Paediatric and Perinatal Epidemiology* 2000; 14: 219-26.
109. Cheung YB. Statistical analysis of human growth and development. Boca Raton, FL: CRC Press, 2014.
110. World Medical Association. World medical association declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA* 2013; 310: 2191-4.
111. Bhutta ZA. What works? Interventions for maternal and child undernutrition and survival. *Lancet* 2008; 371: 417-40.
112. Hirvonen K. Measuring catch-up growth in malnourished populations. *Annals of Human Biology* 2014; 41: 67-75.
113. Luo ZC. Critical growth phases for adult shortness. *American Journal Epidemiology* 2000; 152: 125-31.

114. Stith AY, Gorman KS, Choudhury N. The effects of psychosocial risk and gender on school attainment in Guatemala. *Applied Psychology* 2003; 52: 614-29.
115. Chimombo J. Changing patterns of access to basic education in Malawi: A story of a mixed bag? *Comparative Education* 2009; 45: 297-312.
116. Fink G, Rockers PC. Childhood growth, schooling, and cognitive development: Further evidence from the young lives study. *American Journal of Clinical Nutrition* 2014; 100: 182-8.
117. Sudfeld CR, Charles McCoy D, Danaei G, Fink G, Ezzati M, Andrews KG, et al. Linear growth and child development in low- and middle-income countries: A meta-analysis. *Pediatrics* 2015.
118. Marshall WA, Tanner JM. Variations in the pattern of pubertal changes in boys. *Archives of Disease in Childhood* 1970; 45: 13-23.
119. Walker SP, Chang SM, Powell CA, Simonoff E, Grantham-McGregor SM. Early childhood stunting is associated with poor psychological functioning in late adolescence and effects are reduced by psychosocial stimulation. *Journal of Nutrition* 2007; 137: 2464-9.
120. Hankin BL, Abramson LY, Moffitt TE, Silva PA, McGee R, Angell KE. Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study. *Journal of Abnormal Psychology* 1998; 107: 128-40.
121. Lewinsohn PM, Hops H, Roberts RE, Seeley JR, Andrews JA. Adolescent psychopathology: I. prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology* 1993; 102: 133-44.
122. Ali G, Ryan G, De Silva M. Validated screening tools for common mental disorders in low and middle income countries: A systematic review. *PLoS One* 2016; 11: e0156939.
123. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the hospital anxiety and depression scale: An updated literature review. *Journal of Psychosomatic Research* 2002; 52: 69-77.
124. Allison AC, Ferreira RJ. Implementing cognitive behavioral intervention for trauma in schools (CBITS) with Latino youth. *Child Adolescent Social Work Journal* 2017; 34: 181-9.
125. Deeba F, Rapee RM, Prvan T. Psychometric properties of two measures of childhood internalizing problems in a Bangladeshi sample. *British Journal of Clinical Psychology* 2015; 54: 214-32.
126. Galler JR, Bryce CP, Waber D, Hock RS, Exner N, Eaglesfield D, et al. Early childhood malnutrition predicts depressive symptoms at ages 11-17. *Journal of Child Psychology & Psychiatry* 2010; 51: 789-98.
127. Suliman S, Mkabile SG, Fincham DS, Ahmed R, Stein DJ, Seedat S. Cumulative effect of multiple trauma on symptoms of posttraumatic stress disorder, anxiety, and depression in adolescents. *Comprehensive Psychiatry* 2009; 50: 121-7.
128. Ruiz-Casares M, Thombs BD, Rousseau C. The association of single and double orphanhood with symptoms of depression among children and adolescents in Namibia. *European Child and Adolescent Psychiatry* 2009; 18: 369-76.
129. Wicherts JM, Dolana CV, Carlson JS, van der Maasa HLJ. Raven's test performance of Sub-Saharan Africans: Average performance, psychometric properties, and the flynn effect. *Learning & Individual Differences* 2010; 20: 135-51.
130. Costenbader V, Ngari SM. A Kenya standardization of the Raven's coloured progressive matrices. *School Psychology International* 2001; 22: 258-268.

131. Roach GD, Dawson D, Lamond N. Can a shorter psychomotor vigilance task be used as a reasonable substitute for the ten-minute psychomotor vigilance task? *Chronobiology International* 2006; 23: 1379-87.
132. Loh S, Lamond N, Dorrian J, Roach G, Dawson D. The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers* 2004; 36: 339-46.
133. Waal MWd, Weele GMvd, Mast RCvd, Assendelft WJJ, Gussekloo J. The influence of the administration method on scores of the 15-item geriatric depression scale in old age. *Psychiatry Research* 2012; 197: 280-4.

11 APPENDICES

Appendix I. Mathematics test



LCSS Adolescent Health Study
Form 07. Math 15 years

Participant Code {ChildNumber} |__|__|__| : |__|
(version 2010-08-19)

1. Visit Information

- 1.1. Number (code) of visit {NumberVisit} | 0 | 2 |
 1.2. Date of visit {DateVisit} |__|__| | |__|__| | 20|__|__|
dd mm yy
 1.3. Time of visit {MathTimeVisit} |__|__| | |__|__|
hh mm

2. Test questions

Pasipa pana masamu nsano ni gatatu. Liwusyo lillilose lili ni majibu mcheche. Mwamcheche, limpepe ni lisyene. Nsagule jibu linkuganisa kuti lili lisyene ni nsungusye nambala jili ni jibu lyelele. Mtende mpela ili pasiyi. Mpaka nlembe pachipepa chine naga nli ni yejinji yakuti nlembe.

Below you will see eight mathematics questions. Each question has four answer options. Only one of the options is correct. Choose the option that you think is the correct answer and circle the letter next to it. This is illustrated in the example. You may use the blank paper provided for doing the calculation if you need to.

<p>Example) $3 - 1 =$</p> <p style="text-align: center;"> <input type="radio"/> a. 1 <input type="radio"/> c. 3 <input checked="" type="radio"/> b. 2 <input type="radio"/> d. 43 </p>	<p>Liwusyo: yindu itatu kutyosya chimo tzigale iwiri, sambano pelepa nambala jili ni jibu lisyene ni b. nsungusye "b" jikwete jibu lisyene pa liwusyoli.</p> <p><i>Explanation: $3 - 1 = 2$, so the right answer should be b. Circle the "b" to mark the choice you have made.</i></p>
---	--

<p>$68 - 26 =$</p> <p style="text-align: center;"> a. 40 c. 42 b. 41 d. 43 </p> <p style="font-size: x-small;">2.1 {MathQ 1}</p>	<p>$51 - 17 =$</p> <p style="text-align: center;"> a. 34 c. 37 b. 35 d. 38 </p> <p style="font-size: x-small;">2.5 {MathQ 5}</p>
<p>$9 \times 3 \times 2 =$</p> <p style="text-align: center;"> a. 36 c. 63 b. 54 d. 78 </p> <p style="font-size: x-small;">2.2 {MathQ 2}</p>	<p>$\frac{5}{9} - \frac{1}{3} =$</p> <p style="text-align: center;"> a. $\frac{1}{9}$ c. $\frac{2}{9}$ b. $\frac{2}{3}$ d. $\frac{4}{3}$ </p> <p style="font-size: x-small;">2.6 {MathQ 6}</p>
<p>$9 \times 11 =$</p> <p style="text-align: center;"> a. 20 c. 99 b. 90 d. 19 </p> <p style="font-size: x-small;">2.3 {MathQ 3}</p>	<p>$26 + 11 - 18 =$</p> <p style="text-align: center;"> a. 19 c. 37 b. 35 d. 38 </p> <p style="font-size: x-small;">2.7 {MathQ 7}</p>
<p>$\frac{14}{21} =$</p> <p style="text-align: center;"> a. $\frac{4}{7}$ c. $\frac{2}{3}$ b. $\frac{3}{4}$ d. $\frac{5}{6}$ </p> <p style="font-size: x-small;">2.4 {MathQ 4}</p>	<p>$(2 - \frac{2}{3}) \times 6 =$</p> <p style="text-align: center;"> a. 4 c. 8 b. 6 d. $6\frac{1}{3}$ </p> <p style="font-size: x-small;">2.8 {MathQ 8}</p>

3. Free comments

3.1. Free comments {MathFreeCom} _____

i	Collector {MthCollector}	Monitor {MthMonitor}	Entry_1 {MthEntry_1}	Entry_2 {MthEntry_2}
Code/Sign	i.1 __ __ __	i.2 __ __ __	i.3 __ __ __	i.3 __ __ __
ii Date				

Appendix II. Short Mood and Feelings Questionnaire



LCSS Adolescent Health Study

Participant Code {ChildNumber} |__|__|__| . |__|

Form 09. Short Mood and Feelings Questionnaire-Child version

(version 2010-09-10)

1. Visit Information

- 1.1. Number (code) of visit {NumberVisit} | 0 | 2 |
- 1.2. Date of visit {DateVisit} |__|__| |__|__| 20|__|__|
 dd mm yy
- 1.3. Administration of the questionnaire {SmfAdmin};
 [1] Self-administered
 [2] Interviewed

2. Mwakuyiwonela Feelings

Ayi iwusoyi yikwayana ni mwawele nli nkuyiwonela kapena kutenda mwasampano. Paliusyo lililyose chonde akumbuchile mwayiwonele kapena mwatendele mu mawiki gawili gapitega. Naga chiusyo chilli ngachikamulana, alembe nambala 1 (ngawa yakuwona). Naga ndawi sine yaliji yakuwona alembe pelembe 2 (yakuona ndawi sine). Naga chiusyocho chaliji chakuwona ndawi jejinji alembe pelembe 3 (yakuwona).

This form is about how you might have been feeling or acting recently. For each question, please answer how much you have felt or acted this way in the past two weeks. If a sentence was **not true**, mark a line in a box for **not true**. If it was **sometimes true**, mark a line in a box for **sometimes**. If a sentence was **true most of the time**, mark a line in a box for **true**.

	1. Ngawa Yakuwona Not true	2. Yakuwona ndawi sine Sometimes true	3. Yakuwona True
2.1. Naliji jwakumudwa kapena jwangasangalala. I felt miserable or unhappy. {SmfMisUnhappy}			
2.2. Pangali chachinonyeese. I didn't enjoy anything at all {SmfDidNotEnjoy}			
2.3. Nagambire kuona kupela ni nagambile kutama pangali chakutenda. I felt so tired I just sat around and did nothing. {SmfTiredSat}			
2.4. Naliji jwangatamilichika. I was very restless. {SmfRestless}			
2.5. Nayiweni mpela nganimba jwambone mpela kala. I felt I was no good anymore. {SmfFeltNoGood}			
2.6. Nalisile mnope. I cried a lot. {SmfCriedLot}			
2.7. Nayiweni yakusausya kuti nganisye chenene kapena kuti mbichile mtima. I found it hard to think properly or concentrate. {SmfHardThink}			
2.8. Naliwenjile namsyene. I hated myself. {SmfHatedMyself}			
2.9. Nayiweni mpela ndili muntu jwakusakala. I felt I was a bad person. {SmfBadPerson}			
2.10. Nayiweni mpela ndili jika. I felt lonely. {SmfLonely}			
2.11. Nayiweni mpelaga pangali kusyene syene wakunonyela une I thought nobody really loved me. {SmfNobodyLoved}			
2.12. Nayiweni mpelaga kuti ngasimba mwanache jwambone mpela wanache wane. I thought I could never be as good as other kids. {SmfGoodOther}			
2.13. Nayiweni mpela yosope yandesile yaliji yachabe I felt I did everything wrong. {SmfEveryWrong}			

3. Comments

3.1 Score _____ points {SmfScore}

3.2 Free Comments {SmfFreeCom} _____

i Code/Sign	Collector {SmfCollector} i.1 __ __ __	Monitor {SmfMonitor} i.2 __ __ __	Entry_1 {SmfEntry_1} i.3 __ __ __	Entry_2 {SmfEntry_2} i.3 __ __ __
ii Date	_____	_____	_____	_____

ORIGINAL PUBLICATIONS

REGULAR ARTICLE

Transition between stunted and nonstunted status: both occur from birth to 15 years of age in Malawi children

Tiina Teivaanmäki (tiina.teivaanmaki@gmail.com)^{1,2}, Yin Bun Cheung^{1,3}, Emma Kortekangas¹, Kenneth Maleta⁴, Per Ashorn^{1,5}

1.Department for International Health, School of Medicine, University of Tampere, Tampere, Finland

2.Department of Paediatrics, Helsinki University Hospital, Helsinki, Finland

3.Centre for Quantitative Medicine, Duke-National University of Singapore Graduate Medical School, Singapore, Singapore

4.College of Medicine, University of Malawi, Blantyre, Malawi

5.Department of Paediatrics, Tampere University Hospital, Tampere, Finland

Keywords

Low-income countries, Malawi, Puberty, South-east Africa, Stunted growth

Correspondence

Tiina Teivaanmäki MD, University of Tampere School of Medicine, Arvo-building, Tampere FIN-33014, Finland.

Tel: +358443049556 |

Fax: +358 3 35516164 |

Email: tiina.teivaanmaki@gmail.com

Received

22 January 2015; revised 28 February 2015; accepted 27 May 2015.

DOI:10.1111/apa.13060

ABSTRACT

Aim: The timing and frequency of stunting and possible catch-up growth are ambiguous in low-income settings. This study explored the timing and extent of becoming stunted and nonstunted between birth and 15 years of age in a resource-poor area of Malawi, south-east Africa.

Methods: We followed 767 children from the foetal period until 15 years of age and examined the transition between stunted and nonstunted status and the pubertal stage at 15 years of age. We also plotted smoothed curves for the mean absolute deficits in centimetres and height-for-age standard deviation scores (HAZ) according to the World Health Organization's 2006 and 2007 references.

Results: Most two-year olds (80%) were stunted (HAZ < -2 SD), but this had declined to 37% at 15 years of age. During the three five-year intervals, new stunting cases ranged from 3.9 to 21.3% and the percentage who became nonstunted was 9.1 to 15%. The majority (85%) of the children, who were moderately stunted at two years of age, became nonstunted during the follow-up period. Only, 9% of boys and 20% of girls had reached advanced puberty by the age of 15.

Conclusion: Becoming stunted and nonstunted status both occurred throughout the period from birth to 15 years of age in Malawi children. The small percentage who had reached advanced puberty by the age of 15 suggests significant further growth potential.

INTRODUCTION

Global estimates suggest that there are 162 million children under the age of five who suffer from stunting – faltered length or height gain – as a result of chronic under nutrition (1) and that more than a third (36%) live in sub-Saharan Africa. Stunting is commonly associated with increased mortality, morbidity and developmental problems in infancy and early childhood, but it also makes a negative contribution to adult health and even to economic productivity (2). Furthermore, it has an intergenerational effect, resulting in low birthweight, smaller head circumference and lighter brain weight in the offspring of stunted parents (2–5).

In low-income countries, growth faltering typically starts during the foetal period or soon after birth and continues during the child's first two years of life (6–8). Between two and five years of age, the mean height-for-age standard

deviation scores (HAZ) remain relatively stable in most populations (7,8). Two recent studies with longer follow-up periods have documented an increase in children's mean HAZ scores after five years of age. However, this upward change in relative height was associated with an increase rather than a reduction in the absolute height difference between the study population and the reference group (6,9). In the light of these results, catch-up in linear growth is considered rare after two years of age in low-income

Abbreviations

HAZ, Height-for-age Z-score; LAZ, Length-for-age Z-score; LOWESS, Locally weighted regression smoothed curves.

Key Notes

- The timing and magnitude of stunting and possible catch-up growth are ambiguous, and we studied the incidence in a resource-poor area of south-east Africa.
- Most (80%) of the 767 Malawi children studied were stunted at two years of age, but this figure had fallen to 37% by 15 years of age.
- The small percentage who had reached advanced puberty by the age of 15, 9% of boys and 20% of girls, suggests significant further growth potential.

countries (7,8,10). Relying on the mean length-for-age standard deviation scores (LAZ) and HAZ data in these studies might, however, mask new individual cases of stunting and stunting recovery, especially as most of the earlier studies have depended on very few observations for each participant (6,9,11). Furthermore, none of the reported long-term follow-up studies of adolescents have included an assessment of the participants' pubertal development and, hence, further potential for recovery. Given these study design issues, the incidence of stunting or recovery from it in later childhood remains largely an unanswered question.

Information on individual growth trajectories would often be more useful than mean growth curves in designing public health interventions to address the epidemic of stunting in low-income settings. The purpose of our study was to analyse the growth patterns and frequency of stunting and becoming nonstunted between birth and 15 years of age in a resource-poor area in the southern part of Malawi, south-east Africa. We have described the growth of a well-characterised study cohort, both in terms of absolute height gains and deficits, measured in centimetres, as well as changes in HAZ. Carrying out frequent measurements during the follow-up period meant that, in addition to presenting the mean growth of the participants, we could also identify how individual children made the transition between stunting and nonstunting status using the standard definitions of LAZ and HAZ below minus two. We also assessed the participants' pubertal stage at 15 years of age to estimate their remaining growth potential.

SUBJECTS AND METHODS

Study population

The study was conducted in the Lungwena, Mangochi District, southern Malawi, which at the time of the study covered an area of 100 km² wide area, comprising 26

villages and 23 000 inhabitants in about 5200 households. Most of the inhabitants were Muslims belonging to the Yao tribe and were living in matrilineal descent patterns. The literacy rate was low, and the main sources of income were farming, mainly maize, and fishing.

The cohort, which was enrolled between June 1995 and August 1996, originally comprised 795 mothers who attended the antenatal clinic at Lungwena health centre during their pregnancies (12,13). The total number of foetuses enrolled was 813, and the number of children born live was 767 (Fig. 1). Ethical approval for the Lungwena Child Survival Study was obtained from the National Health Science Research Committee in Malawi (HSRC 93/94) and the College of Medicine Research and Ethics Committee. Informed consent was sought from each guardian at the beginning of the study and again from each guardian and child before the visit at 15 years of age.

Anthropometric measurements

Research assistants measured the participants at their homes every month until 18 months of age and after that every three months until five years of age. After this, the measurements were taken at the study clinic, at more irregular intervals – at 72, 108, 120, 144 and 180 months of age (15 years). All the measurements were carried out between 7 a.m. and 12 noon. During the 15-year period, the maximum number of anthropometric measurements carried out for each participant was 37. Research assistants measured the children's length and height using locally constructed length and height boards with increments of five millimetres until five years of age and used stadiometers (Harpenden, Holtain Limited, UK) with reading increments of one millimetre after that (Fig. 2). They measured children standing up if the child was able to stand or was over 24 months of age. Children under the age of two who were unable to stand were measured lying down using a length board. A trained data collector positioned the child lying on

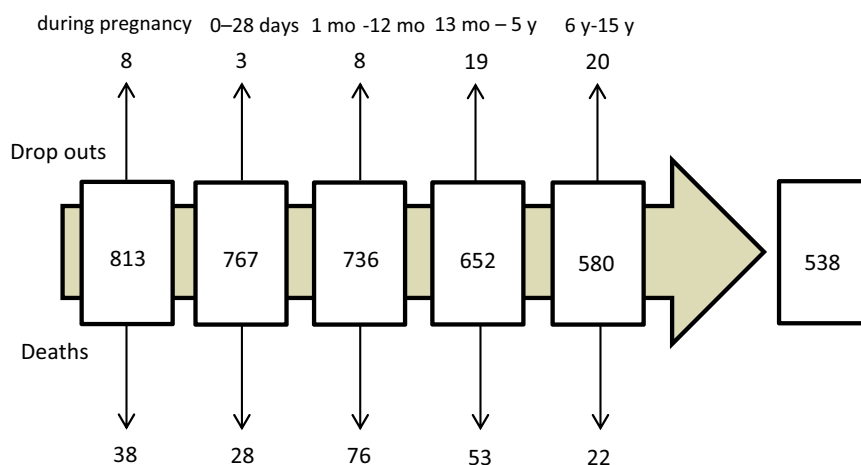


Figure 1 The flow of the study participants. There were 813 foetuses and 767 live born babies participated the study. The number of the remaining participants was 538, at 15 years of age after 179 deaths and 50 dropouts.



Figure 2 The children were measured with a stadiometer after five years of age.

the board, straightened the hips, knees and ankles and read the measurements while an assistant held the head in the correct position. All the children who were able to stand were measured using the tools described above. The participant stood barefoot on the footplate and was positioned with their heels close together, legs straight, arms at their sides and shoulders relaxed. Their heels, buttocks, shoulders and head touched the flat upright surface. The height was read from the scale with a perpendicular headpiece. These measurement techniques have previously been described by Maleta et al. (14). The length and height boards and stadiometers were calibrated weekly, and no technical problems were detected with the equipment.

Research assistants assessed the pubertal stage at 15 years of age using the Tanner classification (15), which includes five stages (I–V) for pubic hair development for both sexes and genitalia development for boys and breast development for girls. Stage I stands for prepubertal stage and stage V for full development (15).

Statistical analyses

We used the World Health Organization Multicentre Growth Reference Study (16) for children up to the age of five and the WHO Reference 2007 (17) from five to 15 years of age to derive LAZ and HAZ scores. Reference values for length were used for the measurements before 24 months of age and height for 24 months of age or more. Using both the 2006 and 2007 WHO references, we were able to cover the whole range of growth in length and height in our cohort. We performed the analyses with Stata

12.1 (Stata Corporation, College Station, TX, USA) and generated the standard deviation scores (Z-scores) for anthropometric measurements using a Stata macro (18). The first measurements that we used in the analysis were taken at one month of age. Because of the large number of home deliveries, birth measurements were not available for most of the participants.

We plotted locally weighted regression smoothed curves (LOWESS) for the Z-scores and mean absolute deficits in centimetres compared to the reference population (16,17) for both sexes. Mean absolute deficits were calculated by subtracting the median values of the reference population from the observed values and then taking the mean from those deviances at each time point. In the main analysis, we included all the measurements of all the participants enrolled in the study. We also carried out a separate analysis of the growth of the participants who took part in the examination at 15 years of age to see whether the results were confounded by participants lost to follow-up. In addition, we compared the growth of male and female participants who were, or were not, stunted at 24 months of age, describing the transition to and from stunting at various time points. We used a stricter definition than just crossing the -2 SD cut-off to define a transition between stunting and nonstunting status to eliminate the effect of possible measurement inaccuracies that would misclassify transition. The participants who remained stunted had HAZ measurements of less than -2 SD at the first measurement and HAZ measurements of less than -1.8 SD at the second measurement. The ones who became nonstunted made the transition from $\text{HAZ} < -2$ SD to $\text{HAZ} > -1.8$. Those who remained nonstunted had $\text{HAZ} > -2$ SD at the first measurement and > -2.2 SD at the second. Those who became stunted made the transition from $\text{HAZ} > -2$ SD to $\text{HAZ} < -2.2$ SD. We assessed the pubertal stage of the participants at 15 years of age and divided them into five groups according to pubic hair development. We then compared growth in HAZ and mean height deficits in centimetres to the reference group in these different groups.

RESULTS

The 767 live born infants had a mean (SD) birthweight of 3060 g (530). The proportion of newborn infants with a low birthweight, of less than 2500 g, was 10% and 22% of the births were preterm, before 37 weeks of gestation. Of the 767 infants, 50 dropped out of the study cohort and 179 died before the age of 15 (Fig. 1). Anthropometrics were recorded for 522 of the 538 participants who were still in the study at 15 years of age. Of those, 258 were boys and 264 girls. At the various time points before that, the number of participants who were measured ranged from 492 to 607. Our analyses included 20 683 measurements during the study, with an average of 27 for each live born baby.

The mean (SD) length/height of the boys was 51 cm (2.5) at one month, 100 cm (4.8) at five years and 154 cm (8.2) at 15 years of age. The respective figures for the girls were 50 cm (2.4), 99 cm (4.6) and 153 cm (5.8). The mean LAZ

declined rapidly until two years of age, at which point it reached a nadir of approximately -3.1 in the boys and -2.9 in the girls (Fig. 3). Thereafter, there was an increase in the mean Z-score that continued until the age of 15 in girls, but decreased after 10 years of age in boys. At the age of 15, the mean HAZ was approximately -2.0 in boys and -1.3 in girls (Fig. 3). A sensitivity analysis that only included the children who remained in the follow-up study until 15 years of age provided very similar results to Figure 3 (Fig. S1).

At one month of age, both the boys and the girls in the study population were, on average, 3.5 cm shorter than the children in the WHO reference population. The deficit in absolute length and height increased until the children were four years old and plateaued at around 10–12 cm in both sexes. After the age of 10, there was a small catch-up in centimetres among the girls but a further decline among the boys (Fig. 3). At the age of 15, the mean deficit in length was 9 cm for the girls and 15 cm for the boys.

The shape of the height trajectories was mostly similar for children who, at the age of two, were moderately (>-3 SD but ≤-2 SD) or severely (≤-3 SD) stunted to those who were not stunted (>-2 SD). This applied to both the boys (Fig. 4) and the girls (Fig. 5). The difference in the mean LAZ between the nonstunted and the severely stunted groups peaked at the age of two, at which point it was a mean difference of 2.9 Z-scores for boys and 2.4 for girls. At the age of 15, the respective differences narrowed to 1.3 for the boys and 0.9 for the girls. In contrast, the difference in mean height deficit between the children who were not stunted and those who were severely stunted at two years of age increased throughout childhood in the boys and

reached 10.7 cm at the age of 15 (Fig. 4). In girls, the difference in height deficit was at its largest at the age of nine (7.5 cm) and narrowed slightly to 7 cm at the age of 15 (Fig. 5).

Figure 6 shows the transition status in stunting between age intervals. Between one and six months, 25.0% of the previously nonstunted infants became stunted, and between six and 12 months, the figure was 16.7%. The risk of becoming stunted declined from 12.5% to 1.3% in the four subsequent annual intervals. When stunting was analysed every five years, the proportion of those who became stunted was 21.3% (95% CI 17.5–25.1) between one and 60 months, 3.9% (2.2–5.6) between 60 and 120 months and 9.1% (6.5–11.7) between 120 and 180 months. The respective figures for recovering from stunting were 9.2% (95% CI 6.6–11.9), 15% (11.9–18.2) and 9.1% (6.5–11.7). The proportion (95% CI) of stunted children (<-2 SD) was 80% (95% CI 76.5–83.5) when the children were two years old and declined to 37.3% (32.9–41.7) by the age of 15. Of the children who were moderately stunted at the age of two years, 84.7% (95% CI 79.4–90.0) were classified as having recovered at least once during the five subsequent age intervals. The respective proportion from those who were severely stunted at two years of age was 58.9% (CI 53.0–64.8).

By the age of 15, only 9.0% (95% CI 5.4–12.6%) of the boys and 19.6% (14.5–24.8%) of the girls had reached advanced puberty, as indicated by their pubic hair conforming to Tanner stages IV–V. Among both boys and girls, there was a direct association between the participants' pubertal status and their HAZ or the absolute deficit in length when compared to the WHO reference population

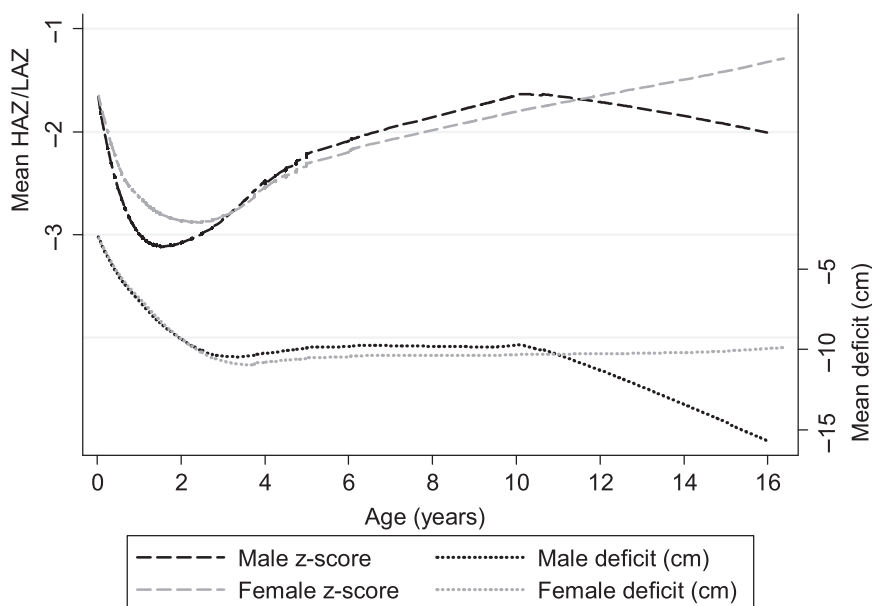


Figure 3 Linear growth among male ($n = 412$) and female ($n = 394$) participants. The dashed lines indicate HAZ scores, and the dotted lines refer to the mean deficits in centimetres compared to the reference population (16,17). The growth curves are LOWESS-smoothed.

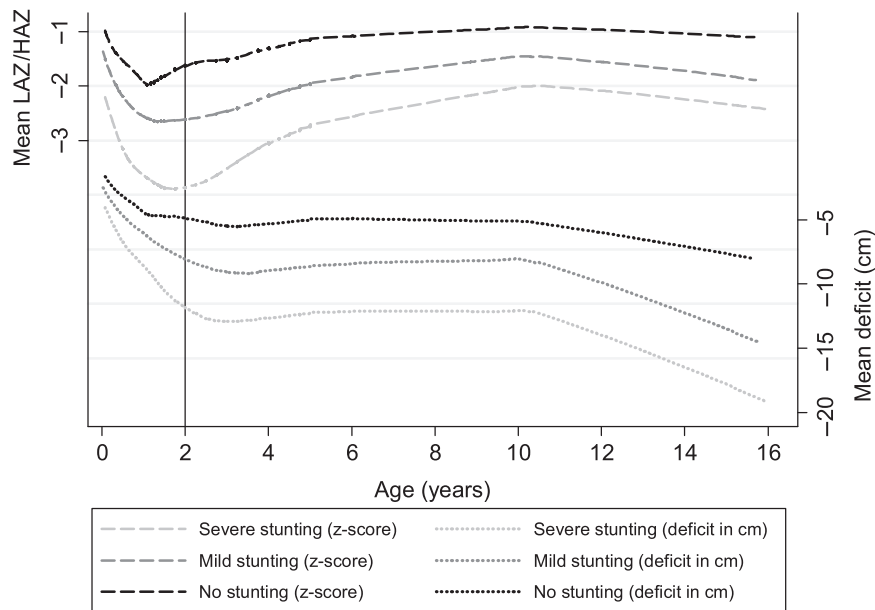


Figure 4 Linear growth among male participants, stratified by their height-for-age Z-score (HAZ) at the age of 24 months. In each set of three lines, the top one summarises data from children who were not stunted ($HAZ > -2.0$, $n = 41$) at two years of age, the middle one refers to children who were moderately stunted ($-3.0 \leq HAZ \leq -2.0$, $n = 98$) and the lowest one refers to children who were severely stunted ($HAZ < -3.0$, $n = 141$). The dashed lines indicate mean HAZ scores, the dotted lines refer to the mean deficits in centimetres compared to the reference population (16,17).

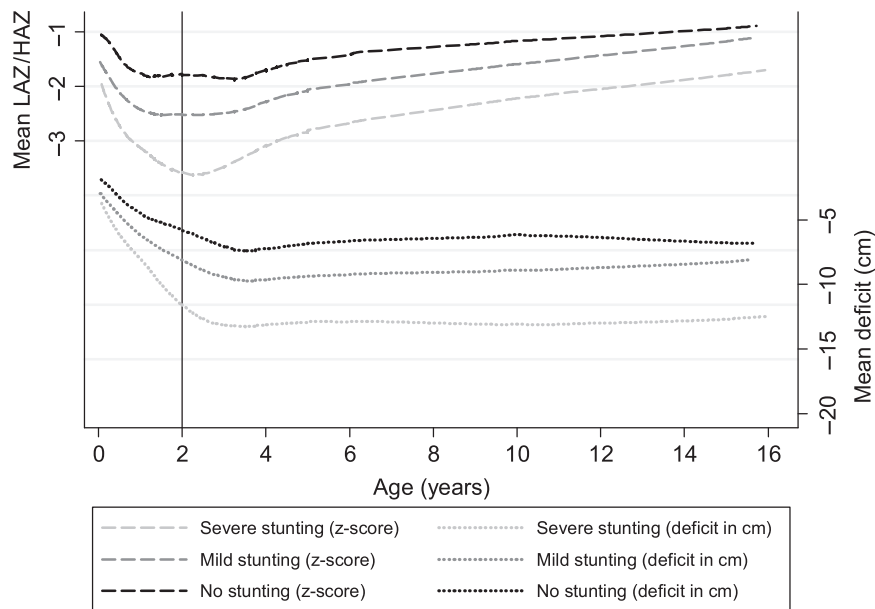


Figure 5 Linear growth among female participants, stratified by their height-for-age Z-score (HAZ) at the age of 24 months. In each set of three lines, the top one summarises data from children who were not stunted ($HAZ > -2.0$, $n = 60$) at two years of age, the middle one refers to children who were moderately stunted ($-3.0 \leq HAZ \leq -2.0$, $n = 88$) and the lowest one refers to children who were severely stunted ($HAZ < -3.0$, $n = 127$). The dashed lines indicate mean HAZ scores, the dotted lines refer to the mean deficits in centimetres compared to the reference population (16,17).

(Table 1). For the adolescents in pubertal stage I, the mean deficit in height was 21 cm for the boys and 13 cm for the girls. For those in pubertal stage V, the mean deficit was approximately 8 cm for both the boys and girls.

DISCUSSION

The main purpose of our study was to analyse the timing and magnitude of growth stunting and becoming nonstunted in a resource-poor area in Malawi, with special emphasis on

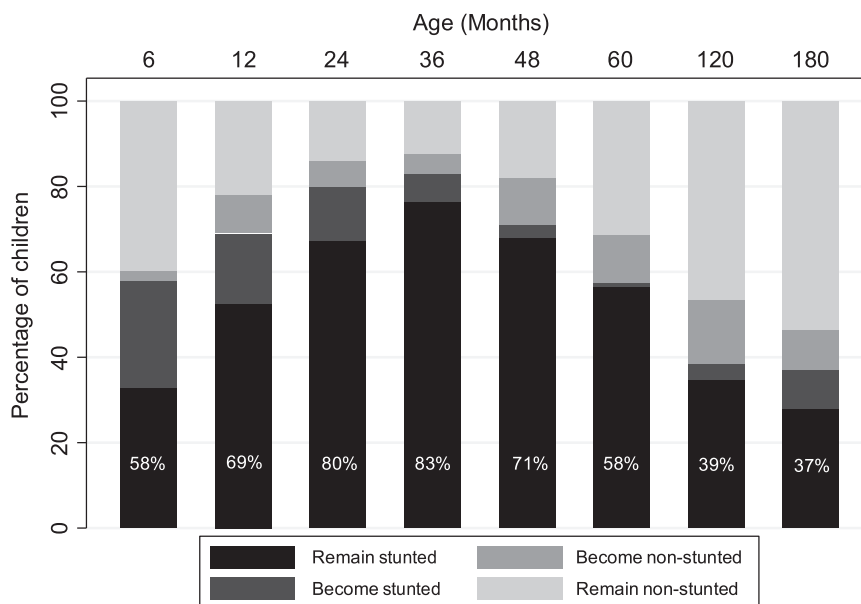


Figure 6 Participants shifting from group of nonstunted to stunted (from HAZ > -2 SD to HAZ < -2.2 SD) or from stunted to nonstunted (from HAZ < -2 SD to HAZ > -1.8 SD), and those remaining stunted (from HAZ < -2 SD to HAZ < -1.8 SD) or nonstunted (from HAZ > -2 SD to HAZ > -2.2 SD). The numbers on the bars represent the percentages of stunted participants at each time point.

Table 1 Mean height-for-age Z-score (HAZ score) and mean height deficit of the participants compared to the reference group (17), stratified by their pubertal stage at the age of 15. Mean height deficit is the mean value of individual deviances from the median of the reference population in each group at 15 years of age. Pubertal stage was evaluated by pubic hair assessment according to Tanner (15)

Gender	Pubertal stage by Tanner (pubic hair)					p-value*
	I	II	III	IV	V	
Male						
Number of participants	26 (11%)	101 (41%)	96 (39%)	21 (9%)	1 (0%)	
Mean HAZ score	-2.76	-2.32	-1.53	-1.05	-0.96	<0.001
Mean deficit in height, centimetre	-21.0	-17.5	-11.1	-7.1	-7.5	<0.001
Female						
Number of participants	33 (14%)	79 (34%)	76 (32%)	33 (14%)	13 (6%)	
Mean HAZ score	-1.92	-1.29	-1.26	-1.07	-1.05	<0.001
Mean deficit in height, centimetre	-13.0	-8.9	-8.7	-7.7	-7.6	0.001

*p-value obtained from test for trend.

events after the first two years of life. In a sample of approximately 500 intensely monitored children, we noticed that both becoming nonstunted and stunted occurred at all ages in terms of HAZ scores. A significant proportion of those stunted at two years of age became nonstunted at some point. HAZ markedly increased between two and 15 years of age in both sexes. Boys in particular tended to be prepubertal at 15 years of age and the more delayed they were, the lower HAZ scores they had.

Our sample represented the target population well. The study cohort comprised 97% of all the newborn infants in the study area during the time of enrolment. During the study, 30% of the participants died or dropped out, but in a

sensitivity analysis, the results were similar when we considered the full data and the date for those who remained in the study until the end of the follow-up period. Random measurement errors in anthropometry were minimised using the mean of three individual measurements, as assessed by trained data collectors. In addition, the same equipment was used for all the participants and it was regularly calibrated. We used the latest growth references, which are considered applicable worldwide. Our results suggest that stunting was common in our study population and becoming nonstunted occurred throughout the follow-up until 15 years of age. We believe that these findings are reliable and can be generalised to the target population.

There is limited knowledge about the transition status of individuals between age intervals (9). Most previous studies of growth faltering in developing countries (6–9,11,19,20) have used mean Z-scores or mean deficits of the population in the analyses. This may have masked some of the changes in transition status and growth recovery, which is why we included transition on an individual level in our analyses. A Consortium of Health-orientated Research in Transitioning Societies' study that also used individual measures reported similar results to ours, suggesting that recovery from stunting, defined as $HAZ < -2.0$ SD, occurs after 24 months of age (9). Transition both to and from stunting throughout the follow-up period may indicate potential plasticity in growth patterns even in late childhood.

The discrepancy in catch-up growth patterns when measured using absolute height or LAZ or HAZ has prompted discussions (6,9,19,21). We believe that the mathematical fact that standard deviation increases with age (19,21) does not exclude HAZ from being a useful tool to measure child growth. If the absolute deficit in centimetres in a child's height remains constant over time, the relative deficit (Z-score) may still decrease when compared to other children, like Prentice et al. have suggested (21). Putting another way, if a risk factor at one age induces an absolute centimetre deficit in height in a child, but this is not followed by repeated exposure to major risk factors in subsequent ages, over time the relative deficit will decrease as each single encounter with a risk factor becomes relatively less important in a longer history of cumulative risk exposure (22).

A child whose HAZ score is below minus two is generally considered to be in an adverse health state. Therefore, becoming nonstunted – in terms of HAZ rising from below minus two to above minus two – is a desired outcome, regardless of whether the absolute centimetre deficit in height remains constant. The definition for stunting is based on height-for-age standard deviation scores of more than two standard deviations below the median of the growth reference (23), although alternative definitions have also been used (24). There are no threshold values to define stunting with absolute height deficits. In summary, we believe that both absolute height deficits and HAZ are useful measures when it comes to studying growth and stunting over time. Using these two metrics, it is possible to provide a comprehensive perspective of the growth patterns of the population.

The significance of growth failure and timing of pubertal development deserves further discussion. In our data, the mean deficit in absolute length and height increased until the children were five years old. After the age of eight, there was a small catch-up in centimetres among the girls but a further decline among the boys. Correspondingly, the height-for-age Z-score increased more among the girls. The onset of puberty was delayed, on average, especially among boys. Growth velocity during puberty is at its highest in boys at the pubic hair stage IV and in girls at stage III (25). Only, 9% of boys and 52% of girls in our cohort had reached those stages at 15 years of age. The fact that boys

suffered from more persistent growth faltering than girls, with a further decline in mean absolute height deficit in late childhood, can be explained by their more delayed puberty. In contrast, the small catch-up in absolute height deficit in girls may have come from the pubertal growth spurt. Hence, we propose that the widely observed pubertal delay indicates that there is the possibility of significant further catch-up growth in both sexes.

The notion that linear growth faltering and becoming nonstunted both occur throughout childhood and adolescence may have some implications for the timing of stunting prevention programmes. The fact that a significant number of children recover from stunting after two years of age addresses the need to extend the intervention programmes to this age group. The focus on the first 1000 days should not limit improving the nutritional environment for older children. Our study was observational and hence cannot be used as direct evidence for possible intervention effects. There are, however, trials showing favourable impacts of nutritional supplementation on height gain among school-age children (26,27). The potential for later catch-up is also emphasised by data from strictly defined medical conditions, such as a specific hormone deficiency, where removal of the adverse exposure typically results in marked growth acceleration even in later childhood (28). In addition, faster height gain in late childhood may also be associated with better schooling outcomes (29,30). Therefore, while the first 1000 days is certainly a critical period, the potential importance of later efforts to promote child growth should not be ignored (6).

CONCLUSION

Our results suggest a great diversity in individual-level growth faltering and in becoming nonstunted throughout childhood. In the Malawi cohort, only a minority of the adolescents had reached advanced puberty, suggesting that the population may still have significant growth potential after the age of 15.

ACKNOWLEDGEMENTS

We would like to thank all the Lungwena Child Survival Study participants, their parents and the staff at the study site. We are also grateful to Professor Bo Eriksson and Data Manager Lotta Alho for statistical assistance.

References

1. United Nations. The millennium development goals report. New York: United Nations 2014, 2014.
2. Victora CG. Maternal and child undernutrition: consequences for adult health and human capital. *Lancet* 2008; 371: 340–57.
3. Black RE. Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet* 2008; 371: 243–60.
4. Grantham-McGregor S. Developmental potential in the first 5 years for children in developing countries. *Lancet* 2007; 369: 60–70.

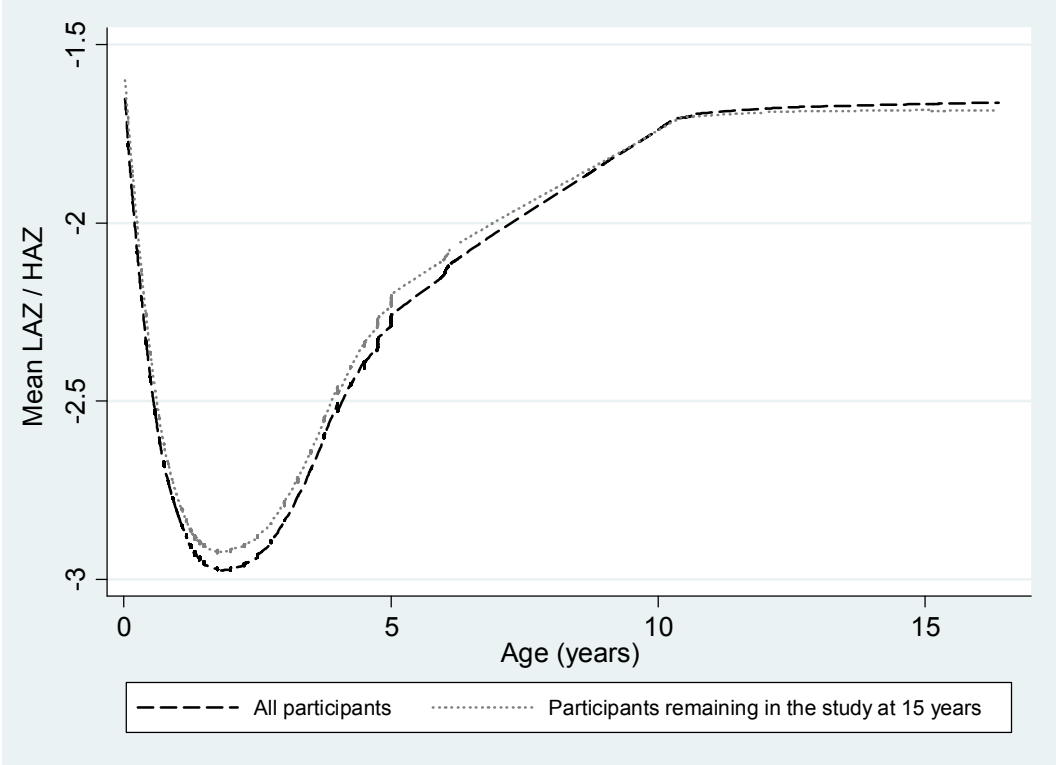
5. Lecours AR, Mandujano M, Romero G. Nutrition and brain function from infancy to old age. in memory of Joaquin Cravioto (1922-1998), Mexico City, Mexico, March 30-31, 2000. *Nutr Rev* 2001; 59: S1–101.
6. Prentice AM, Ward KA, Goldberg GR, Jarjou LM, Moore SE, Fulford AJ, et al. Critical windows for nutritional interventions against stunting. *Am J Clin Nutr* 2013; 97: 911.
7. Victora CG. Worldwide timing of growth faltering: revisiting implications for interventions. *Pediatrics* 2010; 125: e473–80.
8. Shrimpton R. Worldwide timing of growth faltering: implications for nutritional interventions. *Pediatrics* 2001; 107: E75.
9. Lundeen EA, Stein AD, Adair LS, Behrman JR, Bhargava SK, Dearden KA, et al. Height-for-age z scores increase despite increasing height deficits among children in 5 developing countries. *Am J Clin Nutr* 2014; 100: 821–5.
10. Bhutta ZA. What works? interventions for maternal and child undernutrition and survival. *Lancet* 2008; 371: 417–40.
11. Lundeen EA, Behrman JR, Crookston BT, Dearden KA, Engle P, Georgiadis A, et al. Growth faltering and recovery in children aged 1-8 years in four low- and middle-income countries: young lives. *Public Health Nutr* 2014; 17: 2131.
12. Maleta K. Childhood malnutrition and its predictors in rural Malawi. *Paediatr Perinat Epidemiol* 2003; 17: 384–90.
13. Kulmala T, Vaahtera M, Ndekha M, Koivisto A, Cullinan T, Salin M, et al. The importance of preterm births for peri- and neonatal mortality in rural Malawi. *Paediatr Perinat Epidemiol* 2000; 14: 219.
14. Maleta K. Timing of growth faltering in rural Malawi. *Arch Dis Child* 2003; 88: 574–8.
15. Tanner J. *Growth at adolescence*. Oxford: Blackwell Scientific Publications, 1962.
16. WHO multicentre growth reference study group. 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. Geneva: World Health Organization, 2006 (312 pages).
17. De Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 2007; 85: 660.
18. Vidmar SI, Cole TJ, Pan H. Standardizing anthropometric measures in children and adolescents with functions for egen: update. *Stata J* 2013; 13: 366.
19. Leroy JL, Ruel M, Habicht JP, Frongillo EA. Linear growth deficit continues to accumulate beyond the first 1000 days in low- and middle-income countries: global evidence from 51 national surveys. *J Nutr* 2014; 144: 1460–6.
20. Black RE. Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* 2013; 382: 427–51.
21. Leroy JL, Ruel M, Habicht JP. Critical windows for nutritional interventions against stunting. *Am J Clin Nutr* 2013; 98: 854–5.
22. Luo ZC. Critical growth phases for adult shortness. *Am J Epidemiol* 2000; 152: 125–31.
23. World Health Organization. WHO 1995 expert committee report: physical status: the use and interpretation of anthropometry. Geneva: World Health Organization, 1995; Technical Report Series 854
24. Kamal SA, Jamil SS. KJ-Regression model to evaluate optimal masses of extreme cases. *Int J Biol Biotech* 2014; 11: 623–48.
25. Marshall WA, Tanner JM. Variations in the pattern of pubertal changes in boys. *Arch Dis Child* 1970; 45: 13–23.
26. Orr JB. Milk, consumption and the growth of school-children. Preliminary report on tests to the Scottish board of health. *Lancet* 1928; 211: 202.
27. Lampl M. The effects of protein supplementation on the growth and skeletal maturation of New Guinean school children. *Ann Hum Biol* 1978; 5: 219–27.
28. de Wit CC. Patterns of catch-up growth. *J Pediatr* 2013; 162: 415–20.
29. Gandhi M, Ashorn P, Maleta K, Teivaanmäki T, Duan X, Cheung YB. Height gain during early childhood is an important predictor of schooling and mathematics ability outcomes. *Acta Paediatr* 2011; 100: 1113.
30. Cheung YB. Continuation of linear growth failure and its association with cognitive ability are not dependent on initial length-for-age: a longitudinal study from 6 months to 11 years of age. *Acta Paediatr* 2010; 99: 1719–23.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Linear growth among all participants and only those who remained in the follow-up study until 15-years-of-age.

Supplemental Figure 1



REGULAR ARTICLE

Child development at 5 years of age predicted mathematics ability and schooling outcomes in Malawian adolescents

Mihir Gandhi^{1,2}, Tiina Teivaanmaki³, Kenneth Maleta⁴, Xiaolian Duan³, Per Ashorn^{3,5}, Yin Bun Cheung (yinbun.cheung@scri.edu.sg)^{1,2,3}

1.Centre for Quantitative Medicine, Office of Clinical Sciences, Duke-NUS Graduate Medical School, Singapore

2.Biostatistics, Singapore Clinical Research Institute, Singapore

3.Department of International Health, University of Tampere Medical School, Tampere, Finland

4.College of Medicine, University of Malawi, Malawi

5.Department of Paediatrics, Tampere University Hospital, Tampere, Finland

Keywords

Cognition, Fine and gross motor development, Language development, Social behavior

Correspondence

Yin Bun Cheung, Singapore Clinical Research Institute, 31 Biopolis Way, Nanos #02-01, Singapore City 138669, Singapore.

Tel: +65 6508 8310 |

Fax: +65 6508 8317 |

Email: yinbun.cheung@scri.edu.sg

Received

22 March 2012; revised 6 August 2012; accepted 3 September 2012.

DOI:10.1111/apa.12021

ABSTRACT

Aim: This study aimed to examine the association between child development at 5 years of age and mathematics ability and schooling outcomes at 12 years of age in Malawian children.

Methods: A prospective cohort study looking at 609 rural Malawian children. Outcome measures were percentage of correctly answered mathematics questions, highest school grade completed and number of times repeating school grades at 12 years of age. A child development summary score obtained at 5 years of age was the main exposure variable. Regression analyses were used to estimate the association and adjust for confounders. Sensitivity analysis was performed by handling losses to follow-up with multiple imputation (MI) method.

Results: The summary score was positively associated with percentage of correctly answered mathematics questions ($p = 0.057$; $p = 0.031$ MI) and with highest school grade completed ($p = 0.096$; $p = 0.070$ MI), and negatively associated with number of times repeating school grades ($p = 0.834$; $p = 0.339$ MI). Fine motor score at 5 years was independently associated with the mathematic score ($p = 0.032$; $p = 0.011$ MI). The association between child development and mathematics ability did not depend on school attendance.

Conclusion: Child development at 5 years of age showed signs of positive association with mathematics ability and possibly with highest school grade completed at 12 years of age.

INTRODUCTION

Early child development and education may play an important role in preventing intergenerational transmission of poverty and promote adult well-being (1,2). Blair (3) proposed a developmental neurobiological model in which not only cognitive ability but also self-regulatory skills that relate to emotionality influence children's school readiness. A review of studies have found that a range of abilities, including physical and cognitive abilities, measured during kindergarten were predictive of early elementary school outcomes (4,5). The 'lag hypothesis' assumes that children who have poor developmental status in early life are slower in development but they would eventually catch-up and become proficient. In contrast, the 'deficit hypothesis' assumes that these children will suffer a long-term failure in functions and skills (6). A deficit is more of a public health concern than a lag. Without sufficient long-term follow-up data, it is impossible to know which hypothesis is more accurate and whether promotion of early child development may have long-term impact on improving lives.

Many of the earlier studies about early child development and later academic performance were oriented toward empirical prediction instead of studying plausible risk or protective factors (4,5,7). Hence, they did not adjust for potential confounders. Growth stunting and developmental delays are related phenomena (8,9). Stunting is a known risk factor of inadequate child development and school achievement (10,11). Without controlling for growth stunting, among others, it is not clear whether previous studies were showing the impact of poor early development or the impact of early growth stunting.

There is limited amount of information about early developmental status of children and its influence on later

Key notes

- Child development during early childhood is likely to be positively related to mathematics ability and school grade completed at 12 years of age.

cognitive ability and school performance in developing countries. A Brazilian study showed that child development assessed with Griffith Mental Development Scales at 4 years of age was associated with grades attained at 18 years of age ($N = 152$), having adjusted for mother's education and a wealth index (1). A Guatemalan study suggested not only that early biological risk factors and cognitive ability predicted psycho-educational test scores in adolescence but also that education could buffer the negative impact of early risk factors (12). However, this study also showed that preschool cognition explained only a small amount of variation in the psycho-educational test scores despite statistical significance. Furthermore, only 222 of the original over 2300 cohort members were analyzed. In another analysis of the same study involving 333 of the original cohort members, the risk factors were also found associated with number of years of schooling (13). Useful although they are, the common problems of loss to follow-up in long-term studies, the limited sample size, and the limited amount of evidence from a low-resource setting, where education quality may be different, suggests the needs for further studies.

One of the United Nations Millennium Development Goals is to ensure by the year 2015 that all children complete primary schooling (14). In 1999, Mozambique and Malawi, Southern Africa, had the world's highest (72%) and second highest (63%), respectively, rate of primary school dropouts (15). In 2006/2007, Mozambique improved to 56% while Malawi worsened to 64% (15). Hence, development and education for Malawian children are major international concerns.

Using data from the Lungwena Child Survival Study (LCSS), rural Malawi, we aimed to assess association between child development at 5 years of age and mathematics ability and schooling outcomes at 12 years of age. In addition to having a relatively large proportion of cohort members being successfully followed, we also used a multiple imputation (MI) method to assess the impact of loss to follow-up. Height-for-age at the age of 5 years and other covariates were adjusted for as potential confounders as the purpose of this study was to examine the relationship between the outcomes and poor early development, not the predictors of poor early development.

METHOD

Study overview

The LCSS is an ongoing prospective cohort study of 813 children looking at the health and development of rural Malawian infants and children (16,17). Lungwena is an area in southern Malawi where a government health center serves about 100 km² rural area with some 20 000 inhabitants. The original cohort for the LCSS was enrolled between June 1995 and August 1996. All pregnant women presenting for antenatal care were eligible for the LCSS and 97% of the pregnant women in the area, at that time, were enrolled. Anthropometric measurements (height/length and weight) were collected regularly from birth. At the age

of 5 years, each child was invited to a developmental assessment, carried out by a trained research assistant. Training in the use of the developmental assessment tool consisted of four 2-day sessions over a series of 2 months run by a pediatrician who was an accredited trainer of the Griffiths Scales, and had experience in the use of other developmental screening tools such as the Denver II. The training consisted of lectures regarding normal child development and practical training with normal children in Lungwena. The research assistant was then assessed over a 2-month period regularly by the pediatrician. At age 12 years, each child was assessed for health status, anthropometry, schooling performance and given a mathematics test. Ethical approval for the LCSS was obtained from the National Health Science Research Committee in Malawi (HSRC 93/94) while the additional data collection at 12 years of age visit was approved by the College of Medicine Research and Ethics Committee.

Variables

A developmental assessment inventory was developed based on Denver II, Denver Developmental Screening Test and Griffiths Mental Developmental Scales, with modifications based on focus groups with local research workers to make the milestones culturally appropriate in Malawi (17). The inventory originally had 138 items with 34 gross motor, 34 fine motor, 35 language and 35 social items. Feedback on their face validity and content validity were collected from local pediatricians, research assistants, medical students, and a language expert. After revision/addition/deletion based on the interviews, the set of items were administered to 1130 children. Logistic regression was used to assess whether there was an expected association between each item and age and unexpected association with gender. Reliability for each item was tested using two subsamples of 46 (inter-observer) and 25 (intra-observer) randomly selected children who were seen at 7 and 14 days after initial assessment. Inappropriate items were removed after a consensus meeting, resulting in a final set of 110 items. Association between the developmental scores based on the 110 items and height-for-age Z score, a known predictor of child development (1), was demonstrated in a sample of Malawian children age 3–6 years (9). Items were administered until the child failed seven consecutive items in the same domain. After this, it was assumed that the child would not attain any further milestones in that domain. Details of the items in the four domains are listed in the previous report (17). Percent scores and the summary score derived from 5-years development assessments results were used as independent variables. The percent score was derived for each development domain (gross motor, fine motor, language and social). It was calculated by number of items passed divided by number of items administered (i.e. not 'don't know'), multiplied by 100. To prevent inflated type I error due to multiple testing and to study the combined predictive power of all four development domains, a factor score was derived from the percent scores of the four domains using factor analysis with the Bartlett method

(18,19). The Bartlett factor score minimizes the sum of the squared components for the error factors. This method produces unbiased estimates of the true underlying factor scores for the participants. The score was then standardized, that is $(\text{score} - \text{mean})/\text{SD}$, so that it has mean 0 and standard deviation 1. We call this standardized factor score the summary score of child development.

The outcomes in the present analysis were based on age 12-years assessments: percentage of correctly answered mathematics test questions, highest school grade completed, and number of times a school grade was repeated. In the analysis, highest school grade completed was defined as 0 for children who never attended school, and only children who had ever attended school were included in the analysis of number of times a school grade was repeated. The mathematics test had eight questions, each with four answer choices, and was designed with reference to the mathematics test for children aged 7–14 years used in the Indonesian Family Life Survey (20,21). The test has previously been shown to associate with changes in height-for-age (22), which is an established predictor of cognitive ability in children in low-income countries (1). The association supported the criterion validity of the test. Potential confounders that were adjusted by regression analysis were age and height-for-age Z score at 5-years development assessment, weight-for-age Z score near birth, gender, gestational duration (preterm/term), father's occupation, father's literacy, mother's literacy, and a wealth index. Height-for-age Z score and weight-for-age Z score were derived using WHO 2006 Child Growth Standards (23). Gestational duration was derived using fundal height measurements (24). It was dichotomized as preterm (<37 weeks) and term (≥ 37 weeks). Wealth index summarized household ownership of radio, bicycle/tricycles, mattress, number of family supporters, ownership of land per person, and number of cattle (cow, goats/sheep, and chickens) (25). The wealth index was assessed perinatally and categorized into three levels: poor (below 40 percentiles), middle (40–80 percentiles), and rich (top 20 percentiles).

Analyses

Cohort members with no missing values in any exposure variables at 5 years of age, outcome variables at 12 years of age, and covariates were included in the analyses. To address the potential bias due to missing values, missing values were replaced by MIs using all the aforementioned variables measured near birth and at 5 and 12 years as predictors (26,27). The imputation by chained regression is an iterative procedure to get one set of imputed values (28). It updates the prediction for the missing values for one variable by (a) regressing the variable upon the observed and tentatively imputed values on other predictors and then randomly samples from the conditional distribution, (b) iterating the process through the other predictors, and (c) updating the imputed values by repeating the process for multiple cycles so that the conditional distribution stabilizes. Then multiple sets of imputed values were obtained separately, and pooled analysis was performed taking into

account the uncertainty in the imputations (26). The MI procedure was performed using ICE package by Stata software (27). Considering the high proportion of missing values in the data set, 50 sets of imputation were used (28). The children who died by 12 years assessment were not included in the imputed analysis.

The association between the developmental scores at 5 years of age and percentage of correctly answered mathematics questions at 12 years of age was analyzed by multiple linear regression. The associations between the developmental scores and highest school grade completed till 12 years of age (0, 1, 2, 3, ≥ 4 grades) were analyzed by ordinal logistic regression. Similar analysis was conducted with number of times school grades repeated till year 12 (0, 1, ≥ 2 times) as the dependent variable. The above associations were assessed in three sets of models. Model I fitted the regression model using the summary score of child development. Model II fitted the regression models separately with each domain's percent score. Model III simultaneously included all four domain percent scores. Model I would allow studying overall child development. Model II would enable us to study the association between outcome variable and each domain separately. Model III would evaluate association between outcome variable and each domain after adjusting the effect of remaining domains. The model I based on the summary score of child development was used as the main analysis to prevent inflated type I error due to multiple testing for four developmental domains. To assess the interaction between early development and schooling attendance, an additional multiple linear regression model was performed for mathematics percent score including interaction between summary score of child development and ever attended school. The potential confounders were force-entered into each of the above regression models. All the models were performed using observed data as well as imputed data sets. All analyses were performed in Stata/SE 11.2 for Windows (StataCorp, College Station, TX, USA).

RESULTS

Of 813 children, 489 (60%) children were evaluated at 12-years follow-up, 204 (25%) children died by the age 12 year visits, and majority of the remaining 120 (15%) children were not contactable as their family had relocated to other places. From the 489 evaluated children, 74 children were excluded from the analysis due to missing 5-years development assessments or baseline demographic data required in the analysis. That is, 415 (51%) of 813 children were included in the analysis.

Table S1 describes the background characteristics and development scores at 5 years of age of the cohort members included in the analysis. The proportion of preterm birth was 18.3% and the mean weight-for-age near birth and height-for-age at 5 years were -0.58 and -0.205 , respectively. The 415 cohort members included in the analysis and who were excluded were comparable in most aspects, except for slightly more proportion of preterm birth (difference = 6%; $p = 0.041$), and lower mean weight-for-age near

birth (difference = -0.27 Z score; p = 0.003) and height-for-age at 5-years assessment (difference = -0.34 Z score; p = 0.001) in the group excluded from the analysis. Language domain score was lower in the excluded group, with an effect size (mean difference divided by SD) of about 0.22 SD (p = 0.036).

Table 1 tabulates the summary statistics of the outcome variables related to schooling at 12 years of age. Majority of the children had attended school for at least 1 year (83.4%) but only 4.1% of the children completed grade 4 and above. Close to half of the children who had ever attended school had repeated a grade at least once (49.7%).

Table 2 shows multiple linear regression coefficients of percentage of correctly answered mathematics questions at 12 years of age controlling for potential confounders. Model I showed a positive association between the summary score of child development and mathematics ability in observed data (coefficient = 1.774, p = 0.057). MI analysis gave similar results (coefficient = 1.841, p = 0.035). Model II revealed that the fine motor score was positively associated with mathematics ability in both the observed (coefficient = 0.412, p = 0.032) and imputed data (coefficient = 0.445, p = 0.011). However, models II and III indicated positive but statistically non-significant association (each p > 0.05) between each domain's percent score and mathematics ability in most of the cases.

Table 3 and 4 show the ordinal logistic regression analysis results for highest school grade completed and number of school grades repeated by 12 years, controlling for potential confounders. As indicated by the summary score (model I), the higher the child development status at 5 years, the higher the school grade completed years (OR = 1.218; p = 0.096), and the lower the odds of repeating school grades by 12 years (OR = 0.834; p = 0.320). However, all

Table 1 Summary of school attendance, highest school grade completed, number of school grades repeated and percentage of correct mathematics questions at year 12 (N = 415)

Variables	Summary statistics
School attendance, n (%)	
Ever attended school	346 (83.4)
Never attended school	69 (16.6)
Highest school grade completed, n (%)	
0	100 (24.1)
1	122 (29.4)
2	113 (27.2)
3	52 (12.5)
≥4	17 (4.1)
Not reported	11 (2.7)
Number of school grades repeated n (%)*	
0	171 (49.4)
1	124 (35.8)
≥2	48 (13.9)
Not reported	3 (0.9)
Percentage of correct mathematics questions, mean (SD)	36.4 (19.51)

*Percentages are based on 'Ever attended school' (n = 346).

Table 2 Summary of multiple linear regression for percentage of correct mathematics questions at year 12*

Regressor	Model I		Model II†		Model III‡	
	Regression coefficient (p) [95% CI]		Regression coefficient (p) [95% CI]		Regression coefficient (p) [95% CI]	
	Observed (N = 415)	Imputed (N = 609)	Observed (N = 415)	Imputed (N = 609)	Observed (N = 415)	Imputed (N = 609)
Summary score	1.774 (0.057) [-0.055, 3.602]	1.892 (0.031) [0.178, 3.606]	-	-	-	-
Gross motor percent score	-	-	0.206 (0.176) [-0.092, 0.503]	0.184 (0.216) [-0.108, 0.475]	-0.034 (0.874) [-0.461, 0.392]	-0.066 (0.742) [-0.461, 0.329]
Fine motor percent score	-	-	0.412 (0.032) [0.036, 0.789]	0.445 (0.011) [0.104, 0.787]	0.395 (0.140) [-0.130, 0.921]	0.432 (0.068) [-0.032, 0.895]
Language percent score	-	-	0.160 (0.158) [-0.063, 0.383]	0.180 (0.072) [-0.016, 0.376]	0.052 (0.705) [-0.219, 0.324]	0.080 (0.507) [-0.156, 0.316]
Social percent score	-	-	0.122 (0.357) [-0.139, 0.383]	0.074 (0.570) [-0.183, 0.331]	-0.001 (0.997) [-0.307, 0.306]	-0.032 (0.834) [-0.330, 0.266]

*All models were adjusted for weight-for-age Z score near birth, gender, gestational duration, father's occupation, father's literacy, mother's literacy, socio-economic level, age at 5-years assessment, height-for-age Z score at 5-year assessment.

†Model II: regression models fitted separately each domain score as the independent variable.

‡Model III: regression model fitted simultaneously all four domain scores as the independent variables.

Table 3 Summary of ordinal logistic regression for highest school grade completed by year 12*

Regressor	Model I		Model II†		Model III‡	
	Odds ratio (p) [95% CI]		Odds ratio (p) [95% CI]		Odds ratio (p) [95% CI]	
	Observed (N = 404)	Imputed (N = 609)	Observed (N = 404)	Imputed (N = 609)	Observed (N = 404)	Imputed (N = 609)
Summary score	1.218 (0.096)	1.246 (0.070)	–	–	–	–
	[0.966, 1.536]	[0.982, 1.580]				
Gross motor percent score	–	–	1.028 (0.138)	[0.991, 1.067]	1.021 (0.380)	[0.975, 1.069]
Fine motor percent score	–	–	1.029 (0.153)	[0.990, 1.070]	1.029 (0.241)	[0.981, 1.078]
Language percent score	–	–	1.011 (0.319)	[0.989, 1.033]	0.997 (0.786)	[0.972, 1.022]
Social percent score	–	–	1.020 (0.130)	[0.994, 1.047]	1.01 (0.560)	[0.980, 1.039]

*All models were adjusted for weight-for-age Z score near birth, gender, gestational duration, father's occupation, father's literacy, mother's literacy, socio-economic level, age at 5-years assessment, height-for-age Z score at 5-year assessment.

†Model II: regression models fitted separately each domain score as the independent variable.

‡Model III: regression model fitted simultaneously all four domain scores as the independent variables.

Table 4 Summary of ordinal logistic regression for number of school grades repeated by year 12*

Regressor	Model I		Model II†		Model III‡	
	Odds ratio (p) [95% CI]		Odds ratio (p) [95% CI]		Odds ratio (p) [95% CI]	
	Observed (N = 343)	Imputed (N = 493)	Observed (N = 343)	Imputed (N = 493)	Observed (N = 343)	Imputed (N = 493)
Summary score	0.834 (0.320)	0.852 (0.339)	–	–	–	–
	[0.584, 1.192]	[0.614, 1.183]				
Gross motor percent score	–	–	0.991 (0.677)	[0.948, 1.035]	0.995 (0.828)	[0.953, 1.039]
Fine motor percent score	–	–	0.957 (0.152)	[0.901, 1.016]	0.965 (0.217)	[0.913, 1.021]
Language percent score	–	–	0.988 (0.445)	[0.959, 1.018]	0.983 (0.217)	[0.959, 1.010]
Social percent score	–	–	1.012 (0.560)	[0.972, 1.053]	1.017 (0.391)	[0.978, 1.057]

*All models were adjusted for weight-for-age Z score near birth, gender, gestational duration, father's occupation, father's literacy, mother's literacy, socio-economic level, age at 5-years assessment, height-for-age Z score at 5-year assessment.

†Model II: regression models fitted separately each domain score as the independent variable.

‡Model III: regression model fitted simultaneously all four domain scores as the independent variables.

the analyses concerning highest school grades and repeating did not reach conventional level of statistical significance (each $p > 0.05$). Analysis of MI data gave similar results.

The linear regression models for mathematics percent score with interaction between summary score of child development and ever attended school showed some positive but statistically non-significant interactions ($p = 0.941$). Among those who never attended school, every 1 SD increase in the summary score of child development was associated with 1.6% ($p = 0.131$) higher score in the mathematics test, whereas among those who attended school the increase was 1.8% ($p = 0.339$) in the observed data. Similar results were observed with the imputed values (details not shown).

DISCUSSION

We aimed to assess relation between child development at 5 years of age and mathematics ability and schooling achievements at 12 years of age. We demonstrated positive associations between the summary score of child development at 5 years of age and mathematics ability at 12 years of age in a rural Malawian cohort, with adjustment for a range of covariates including growth stunting in early life. The analysis of the observed data did not reach conventional level of statistical significance ($p = 0.057$), but the multiply-imputed data gave similar regression coefficient and a stronger level of statistical significance ($p = 0.031$) allowing for losses to follow-up. The results appeared to be stable despite the use of different methods and sample data, suggesting the reliability of the findings. It was estimated that for one SD increase in the summary score of child development, the mathematics score increased by about 0.1 SD. Taking plus and minus one (or two) SD in the summary score as indicators of slightly (or moderately) high and low levels of child development, the difference in mathematics score would be about 0.2 (or 0.4) SD. This would mean a small-to-moderate effect as per Cohen's suggestion (29). Similarly, the odds of having higher school grades completed differed by about 50% (100%) for the slightly (moderately) high and low levels of child development. This again appears to be practically significant.

We also assessed associations between individual domain of child development at 5 years of age and the outcomes at 12 years of age. We found that individual domain had a weak association with outcome measures. Only fine motor development had significant association with mathematics score when assessed individually. But there was no clear evidence of one domain being more important than the others in models for highest grade completed and number of times repeating grades at 12 years of age. This suggests that a wide spectrum, instead of a single aspect, of developmental impairment is associated with long-term outcomes. This supports the previous findings about the harmful effects of multiple risk factors on schooling (13). In contrast to the Guatemalan study, school attendance has not shown significant interaction with summary score for child development on mathematics ability.

Our findings are similar to those in the previous studies conducted in developed countries and in Brazil and Guatemala, where early development predicted later academic abilities (1,12). However, the present study analysis adjusted for potential confounders including growth stunting, eliminating the effect of known risk factors and hence provides more accurate estimates of the associations between early development and long-term outcomes. In the present study, there was indication of association between development at 5 years of age and mathematical skills. We did not see a strong enough association between early child development and highest school grade completed and number of times repeating school grades although the observed associations were in line with the hypothesis of poor child development leading to a deficit in schooling outcomes. This finding might be because schooling outcomes could be affected by environmental issues like poverty, illnesses or other practical issues. The prevalence of ever repeating a school grade was high in this cohort. We would hypothesize that being in the rural area is itself a risk factor but the precise factors are unknown. Further research into its determinants will be useful. The data suggest that the plausibility of a long-term deficit associated with poor development in early life. The 'lag hypothesis' was a potential reason for people to disregard poor early child development. But the findings appeared to support the 'deficit hypothesis' although the p -values did not clearly confirm it. The United Nations' Millennium Development Goals included not only completion of a full course of primary schooling for all children but also elimination of poverty. In Malawi, completion of primary schooling was uncommon. Education and functional skills such as mathematics may affect income and the risk of poverty (1). Early child development can be promoted, and programs to promote it are within reach (30). The results here suggest that early childhood development is an area that may have implications to not only short-term behaviors but also long-term development and the achievement of the Millennium Development Goals.

Sample attrition is a common problem in long-term follow-up studies. The Guatemalan study, for example, had only about 10% to 15% of the cohort members included in the analyses of long-term outcomes (12,13), whereas the Brazilian study was able to include only about 40% of the cohort members in the long-term analysis. In the present study, over 50% of the cohort members were included in the assessment and analysis. The main reason of exclusion was death prior to age 12, not insufficient follow-up efforts. To make a valid 'prediction' about their academic outcomes if they had survived is a question different from to make a valid assessment of the actual association among the survivors. This study only aimed at assessing the actual association. Furthermore, we have also included sensitivity analysis based on MI, an advanced statistical method for handling missing data, and found that the results are robust. There is no evidence of bias despite the attrition rate. Nevertheless, further research with planned long-term follow-up is warranted to confirm the present findings.

There are several limitations to our study. Firstly, the sample size was not powered for the present analysis, and some practically significant associations were statistically non-significant. Nevertheless, the use of MI method somewhat alleviate the issue of sample size and also showed that the results were robust in relation to loss to follow-up. Secondly, test results on a developmental inventory may be presented as *Z* score or developmental quotient if there are well-established norms from a reference population. The Malawian inventory is a newly adapted, culturally appropriate measure that has not yet developed such a norm. A simple percentage score was calculated for each domain to represent the results. However, a recent study has established that in terms of estimation of association, the choice of scoring approach has little influence on the results (9). Thirdly, the mathematics test used at 12 years of age was not a formally validated instrument. However, it was modeled on an instrument used in a large-scale Indonesian study that included assessment of children at a similar age, the instrument possesses face validity, and it had been shown previously to be associated with height gain in children (22), which is a known predictor of cognitive ability (1). They provide evidence about its validity.

In summary, there were signs that child development at age 5 years was positively associated with mathematics ability and schooling outcomes at age 12 years, even after adjusting for growth stunting and other potential confounders. The strength of association was not negligible. Early child development programs preceding the school admission age may have an impact on long-term cognitive and educational outcomes. Further studies with stronger statistical power and higher follow-up rate are needed to confirm the findings.

ACKNOWLEDGEMENTS

We are thankful to Dr. Melissa Gladstone for conducting the developmental assessment training. The last author (YBC) was supported by Singapore Ministry of Health's National Medical Research Council under its Clinician Scientist Award.

References

- Grantham-McGregor S, Cheung YB, Cueto S, Glewwe P, Richter L, Strupp B, et al. Developmental potential in the first 5 years for children in developing countries. *Lancet* 2007; 369: 60–70.
- Walker SP, Wachs TD, Gardner J, Lozoff B, Wasserman GA, Pollitt E, et al. Child development: risk factors for adverse outcomes in developing countries. *Lancet* 2007; 369: 145–57.
- Blair C. School readiness: integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *Am Psychol* 2002; 57: 111–27.
- Tramontana MG, Hooper SR, Selzer SC. Research on the pre-school prediction of later academic achievement: a review. *Dev Rev* 1988; 8: 89–146.
- Pianta RC, McCoy SJ. The first day of school: the predictive validity of early school screening. *J Appl Dev Psychol* 1997; 18: 1–22.
- Francis DJ, Shaywitz SE, Stuebing KK, Shaywitz BA, Fletcher JM. Developmental lag versus deficit models of reading disability: a longitudinal, individual growth curves analysis. *J Educ Psychol* 1996; 88: 3–17.
- Stevenson HW, Newman RS. Long-term prediction of achievement and attitudes in mathematics and reading. *Child Dev* 1986; 57: 646–59.
- Cheung YB, Yip PS, Karlberg JP. Fetal growth, early postnatal growth and motor development in Pakistani infants. *Int J Epidemiol* 2001; 30: 66–72.
- Cheung YB, Gladstone M, Maleta K, Duan X, Ashorn P. Comparison of four statistical approaches to score child development: a study of Malawian children. *Tropical Med Int Health* 2008; 13: 987–93.
- Walker SP, Chang SM, Powell CA, Grantham-McGregor SM. Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: prospective cohort study. *Lancet* 2005; 366: 1804–7.
- Chang SM, Walker SP, Grantham-McGregor S, Powell CA. Early childhood stunting and later fine motor abilities. *Dev Med Child Neurol* 2010; 52: 831–6.
- Gorman KS, Pollitt E. Does schooling buffer the effects of early risk? *Child Dev* 1996; 67: 314–26.
- Stith AY, Gorman KS, Choudhury N. The effects of psychosocial risk and gender on school attainment in Guatemala. *Appl Psychol* 2003; 52: 614–29.
- United Nations. Road map toward the implementation of the United Nations Millennium Declaration. New York: United Nations, 2001; United Nations General Assembly Document A/56/326.
- UNESCO. EFA Global monitoring report 2011. Education for All. The hidden crisis: Armed conflict and education. Paris: UNESCO, 2011.
- Maleta K, Virtanen S, Espo M, Kulmala T, Ashorn P. Timing of growth faltering in rural Malawi. *Arch Dis Child* 2003; 88: 574–8.
- Gladstone MJ, Lancaster GA, Jones AP, Maleta K, Mtitimila E, Ashorn P, et al. Can western developmental screening tools be modified for use in a rural Malawian setting? *Arch Dis Child* 2008; 93: 23–9.
- Johnson RA, Wichern DW. Applied multivariate statistical analysis, 5th ed. Upper Saddle River, NJ: Prentice Hall, 2002.
- DiStefano C, Zhu M, Míndrilá D. Understanding and using factor scores: considerations for the applied researcher. *Pract Assess Res Eval* 2009; 14: 1–11.
- Cheung YB. Growth and cognitive function of Indonesian children: zero-inflated proportion models. *Stat Med* 2006; 25: 3011–22.
- Strauss J, Beegle K, Sikoki B, Dwiyanto A, Herawati Y, Witoeilar F. The Indonesia Family Life Survey (IFLS): overview and field report, WR-144 = 1–NIA = NICHHD. Santa Monica, CA: RAND Corporation, 2004.
- Gandhi M, Ashorn P, Maleta K, Teivaanmaki T, Duan X, Cheung YB. Height gain during early childhood is an important predictor of schooling and mathematics ability outcomes. *Acta Paediatr* 2011; 100: 1113–8.
- WHO. Child growth standards based on length/height, weight and age. *Acta Paediatr Suppl* 2006; 450: 76–85.
- Andersson R, Bergstrom S. Use of fundal height as a proxy for length of gestation in rural Africa. *J Trop Med Hyg* 1995; 98: 169–72.
- Filmer D, Pritchett LH. Estimating wealth effects without expenditure data – or tears: an application to educational enrollments in states of India. *Demography* 2001; 38: 115–32.
- Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley, 1987.

27. Royston P. Multiple imputation of missing values: further update of ice, with an emphasis on categorical variables. *Stata J* 2009; 9: 466–77.
28. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011; 30: 377–99.
29. Cohen J. *Statistical power analysis for the behavioral sciences*, 2nd ed. New York: Academic Press, 1988.
30. Engle PL, Black MM, Behrman JR, Cabral de Mello M, Gertler PJ, Kapiriri L, et al. Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *Lancet* 2007; 369: 229–42.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1 Summary of covariates and development scores (N = 415).

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Table 1: Summary of Covariates and Development Scores (N=415)

Variables	Summary statistics
Covariates measured near birth	
Preterm (<37 weeks gestation), n (%)	76 (18.3)
Gender, Male, n (%)	205 (49.4)
Weight-for-age Z score, mean (SD)	-0.58 (1.15)
Socio-economic level, n (%)	
Poor	154 (37.1)
Middle	170 (41.0)
Rich	91 (21.9)
Father's occupation, n (%)	
Farmer	81 (19.5)
Fisherman	169 (40.7)
Trader	80 (19.3)
Other	79 (19.0)
Unemployed/Not alive	6 (1.5)
Father's literacy, n (%)	165 (39.8)
Mother's literacy, n (%)	52 (12.5)
Child growth and development scores based on 5-years assessment	
Height-for-age Z score, mean (SD)	-2.05 (0.915)
Gross motor percent score (33 items) , mean (SD)	96.5 (6.40)
Fine motor percent score (27 items) , mean (SD)	97.1 (5.09)
Language percent score (32 items) , mean (SD)	94.2 (8.72)
Social percent score (18 items) , mean (SD)	92.3 (7.39)
Summary score (4 domains) , mean (SD)	0.02 (1.05)

Original Article

Height gain after two-years-of-age is associated with better cognitive capacity, measured with Raven's coloured matrices at 15-years-of-age in Malawi

Tiina Teivaanmäki^{1,2}, Yin Bun Cheung^{1,5}, Anna Pulakka[§], Jussi Virkkala⁷, Kenneth Maleta⁶ and Per Ashorn^{1,3}

¹University of Tampere School of Medicine, Department for International Health, Tampere, Finland, ²Helsinki University Hospital, Department of Paediatrics, Helsinki, Finland, ³Tampere University Hospital, Department of Paediatrics, Tampere, Finland, [§]University of Turku, Department of Public Health, Finland, ⁵Duke-National University of Singapore Graduate Medical School, Centre for Quantitative Medicine, Singapore, Singapore, ⁶School of Public Health and Family Medicine, College of Medicine, University of Malawi, Malawi, and ⁷Sleep Laboratory, Finnish Institute of Occupational Health, Finland

Abstract

Stunting is a measure of chronic undernutrition, and it affects approximately 160 million children worldwide. Cognitive development of stunted children is compromised, but evidence about the association between height gain in late childhood and adolescent cognitive capacity is scarce. We aimed to determine the association between height gains at different ages, including late childhood, and cognitive capacity at 15-years-of-age. We conducted a prospective cohort study in a rural African setting in Southern Malawi. The study cohort was enrolled between June 1995 and August 1996. It originally comprised mothers of 813 fetuses, and the number of children born live was 767. These children were followed up until the age of 15 years. The anthropometrics were measured at one and 24-months-of-age and 15-years-of-age, and cognitive capacity of participants was assessed at 15-years-of-age with Raven's Coloured Matrices score, mathematic test score, median reaction time (RT) (milliseconds) and RT lapses. The associations between growth and the outcome measures were assessed with linear regression. Raven's Coloured Matrices score was predicted by height gain between 24 months and 15-years-of-age (coefficient 0.85, $P=0.03$) and (coefficient 0.69, $P=0.06$), but not by earlier growth, when possible confounders were included in the model. The association weakened when school education was further added in the model (coefficient = 0.69, $P=0.060$). In conclusion, in rural Malawi, better growth in late childhood is likely to lead to better cognitive capacity in adolescence, partly through more school education. In light of these results, growth promotion should not only be limited to early childhood.

Keywords: growth, stunting, undernutrition, adolescent, cognitive capacity, school education.

Correspondence: MD Tiina Teivaanmäki, University of Tampere School of Medicine, Arvo-building, FIN-33014 Tampere, Finland. E-mail: tiina.teivaanmaki@gmail.com

Introduction

Stunting is a measure of chronic undernutrition. The prevalence of stunting has decreased over the years worldwide. Recently, it was estimated that 160 million, i.e. 25% of all the children under five-years-of-age are stunted (UN 2015). The improvement has been slow in Sub-Saharan Africa, where the proportion of stunted children is 37% (UN 2015). Stunting is associated with increased mortality, morbidity and developmental problems including neurodevelopment in infancy and early childhood (Black *et al.* 2008; Chang *et al.* 2002;

Grantham-McGregor *et al.* 2007; Walker *et al.* 2007; Perignon *et al.* 2014; Sudfeld *et al.* 2015). It also has intergenerational effects (Walker *et al.* 2015). The children of shorter mothers have lower birth weight, grow slower and have smaller head circumference and lighter brain weight (Victora *et al.* 2008; Black *et al.* 2008; Grantham-McGregor *et al.* 2007; Lecours *et al.* 2001).

A large number of cross-sectional studies have demonstrated an association between stunting and cognitive capacity (Grantham-McGregor *et al.* 2007; Berkman *et al.* 2002; Grantham-McGregor 2002). In a recent Cambodian study, school aged children who

were stunted at the time of the assessment got lower scores in intelligence tests compared with those with better growth (Perignon *et al.* 2014). A strong association was found between concurrent stunting and cognitive skills in Peruvian children entering school (Crookston *et al.* 2011). In addition to the cross-sectional studies, some longitudinal studies, that assess the association between early childhood growth and later cognitive development, have been conducted. These studies have shown a relationship between faster growth in early childhood and better cognitive capacity later in life (Crookston *et al.* 2011; Grantham-McGregor *et al.* 2007; Walker *et al.* 2007; Crookston *et al.* 2013; Gandhi *et al.* 2011; Fink & Rockers 2014). However, except a recent multicentre study (Fink & Rockers 2014), they have mainly focused on the impact of early childhood growth. In contrast to earlier suggestions (Victoria *et al.* 2010) it has recently been proposed that stunted children may also demonstrate catch-up growth after the first 1000 days of life, i.e. 24-months-of-age (Teivaanmäki *et al.* 2015; Crookston *et al.* 2013). This gives a rationale to explore whether there is an association between height gain in late childhood and later cognitive development of children, and whether the circumstances leading to better growth also enhance the cognitive development.

School education may buffer the risks for poor development and is associated with better cognitive function in stunted high-risk populations (Gorman & Pollitt 1996; Sudfeld *et al.* 2015; McCoy 2015). In low-income settings, children who grow better are more likely to be healthier and get more school education (Grantham-McGregor *et al.* 2007). Hence,

it is possible that the association between better growth and better cognitive capacity is mediated by schooling.

Our primary study question was whether the growth between birth and 24-months-of-age and between 24 months and 15-years-of-age (180 months) were associated with cognitive development in adolescence. Because cognitive development may be affected by several exposures, our secondary aim was to find out if any observed association was mediated through other exposures, such as education. We hypothesized that better growth during any of the age periods assessed would be associated with better cognitive capacity at 15-years-of-age.

Participants and methods

Study design

This prospective cohort study was conducted in Lungwena, Mangochi District, Southern Malawi. The study area covered approximately 100 km², and included 26 villages and 23 000 inhabitants in about 5200 households. Most of the inhabitants were Muslims of the Yao tribe, and the family organization was matrilineal. The literacy rate was low, and the main sources of income were farming and fishing. The study cohort was enrolled between June 1995 and August 1996. It originally comprised 795 mothers who attended antenatal clinic at Lungwena health centre during their pregnancies (Maleta *et al.* 2003). These pregnant

Key messages

- The impact of the timing of stunting and catch-up growth on cognitive capacity later in life is not clear. Recent evidence on possible catch-up growth until adolescence emphasizes the need to explore the association between height gain at different age periods and cognitive development.
- In a resource poor area in rural Malawi, height gain between 24 months and 15 years was associated with better cognitive capacity at 15-years-of-age. This association was partially mediated through school attendance.
- Besides the child's height gain and education, better cognitive performance in adolescence was predicted by male gender and maternal educational attainment.

women carried altogether 813 fetuses, and the number of children born live was 767 (Supplemental Fig. 1). These children were followed up until the age of 15 years.

Ethical approval for the Lungwena Child Survival Study (LCSS) was obtained from the National Health Science Research Committee in Malawi (HSRC 93/94) and the College of Medicine Research and Ethics Committee. Informed consent was obtained from each guardian in the beginning of the cohort study and again from each guardian and adolescent before the visit at 15-years-of-age.

Data collection

We studied the association of height gain between birth and 24 months, and between 24 months and 15 years, respectively, with cognitive capacity at 15-years-of-age with three different regression models. The first anthropometric measurements were taken at one-month-of-age, and these measurements were used as a proxy for weight and height at birth. Subsequent measurements were taken at 24 months and 15-years-of-age. At one-month-of-age, the data collectors measured the participants' length at their homes with locally constructed length boards with reading increments of 5 mm. The data collectors were thoroughly and repeatedly trained and retrained by the investigators with intervals of one to 36 months. In addition, their work was regularly monitored. The measurements at 24 months-of-age were taken at participants' homes with a self-made height board with reading increments of 5 mm. The measurements at 15-years-of-age were done at the study clinic with a stadiometre (Harpenden, Holtain Limited, UK) with reading increments of 1 mm.

The techniques for the measurements have been described in detail earlier by Maleta *et al.* (2003). The study team calibrated all length and height boards and stadiometres weekly.

We used WHO Multi-centre Growth Reference Study (WHO Multicentre Growth Reference Study Group 2006) for children at one and 24-months-of-age and WHO Reference 2007 (De Onis *et al.* 2007) at 15-years-of-age to derive length-/height-for-age

z-scores (LAZ/HAZ). Reference values for length were used for the measurements at one and height for ≥ 24 -months-of-age. By using the two closely aligned WHO references we were able to cover the whole range of growth in length/height in the cohort. We generated the *z*-scores for anthropometric measurements using a Stata macro (Vidmar *et al.* 2013).

We used Raven's Coloured Matrices to assess cognitive performance at 15-years-of-age (Raven *et al.* 1998). We conducted a pilot study in the same setting and age group with Raven's Standard Progressive Matrices (Raven *et al.* 2000), and received poor results with very little variability. Hence, we used Raven's Coloured Matrices, which is originally designed and standardized for children younger than 15-years-of-age (Raven *et al.* 1998). Raven Coloured Matrices has been found to correlate with the performance component of the Wechsler Intelligence Scale for Children (WISC-III) (The Psychological Corporation 1997).

This assessment had 36 patterns each with a piece missing, and the best possible score was 36. The participants chose the piece that they considered correct to complete the pattern from six alternatives.

The mathematics test had eight questions, each with four answer alternatives. It was developed from a reference test designed for children aged 7–14 years used in the Indonesian Family Life Survey (Strauss *et al.* 2004; Gandhi *et al.* 2011). The test has previously associated with preceding development (Gandhi *et al.* 2013) and changes in HAZ (Gandhi *et al.* 2011), which is an established predictor of cognitive ability in children in low-income countries (Grantham-McGregor *et al.* 2007).

We used computer based Psychomotor Vigilance Task (PVT) (Dinges & Powell 1985) to assess reaction time (RT), which has correlated with intelligence (Wenger & Townsend 2000; Vernon 1983). RT test measures the time between a stimulus and participant's reaction during a 5-min test period with random inter-stimulus interval of 2 to 10 s. A number appears on a black display, and the participant indicates his or her reaction by pressing the enter key with a forefinger of the dominant hand. We evaluated two PVT performance metrics in our study, the median RT (milliseconds) and the number of lapses (RTs ≥ 500 ms) in a 5-min trial. Before the test each participant performed a test

trial to familiarize himself or herself with the procedure.

Statistical analyses

We formulated three regression models to assess the associations between growth between birth and 24 months, and between 24 months and 15 years, and cognitive capacity at 15-years-of-age. Model 1 simultaneously included HAZ at one (HAZ_1) and 24 months (HAZ_24) and 15-years-of-age (HAZ_180) as independent variables without any adjustments. Model 2 was adjusted for gender, gestational duration (weeks), father's occupation, father's literacy, mother's literacy and a wealth index (Filmer & Pritchett 2001). Wealth index was assessed perinatally by interviewing the mothers. It summarized household ownership of radio, bicycle / tricycles, mattress, number of family supporters, ownership of land per person and number of cattle (cow, goats, sheep and chickens). It was derived from factor analysis, and categorized into three levels: poor (below 40 percentile), middle (40–80 percentiles) and rich (top 20 percentile). The gestational duration was estimated by using the nationally used chart for fundal height during the antenatal visits, because ultrasound was not available and the information about the timing of menstrual periods was not reliable (Kulmala *et al.* 2000). Model 3 was further adjusted for the number of years of schooling reported at the age of 15 years. In all three models, the associations between HAZ scores and Raven's Coloured Matrices score, mathematics score, median RT (ms) and RT lapses were assessed with multiple linear regression. Both theoretical and simulation studies showed that the least squares regression is robust in analysis of non-normal data even when the sample size is smaller than the present study's (Cheung 2014; Sullivan & Sr D'Agostino 2003). To further evaluate the robustness of the chosen statistical method, we ran sensitivity analysis with ordered logistic regression.

We included all the cohort members with no missing values in exposure variables, outcome variables at 15-years-of-age and confounders in the analyses. Possible selection bias caused by missing values, loss

to follow-up and death data was assessed by comparing the background characteristics and anthropometric measurements of the included and excluded participant with Chi-square test and *t*-test. Previous scholars have discussed that sample size (or power) calculation after completion of statistical analysis is inappropriate (Feinstein & Concato 1998). Instead, they have suggested focusing on confidence interval. As such, for all regression coefficients we included confidence intervals to demonstrate the precision level of the estimates.

We also ran the analyses with HAZ_1 and 'unexplained residual' HAZ at 24 months (rHAZ_24) and 15-years-of-age (rHAZ_180) (Cheung 2014). These residual height values represent the child's deviation from his or her expected HAZ independent of his/her earlier HAZ. The three aforementioned models were also fitted for unexplained residuals of HAZ. In all of the analyses, regression coefficients with *P*-value smaller than 0.05 were considered statistically significant. We performed all the analyses with Stata 12.1 (Stata Corporation, College Station, TX, USA).

Results

Of the 767 live-born infants, 179 (23%) died and 50 (7%) dropped out during the follow-up. Of the 538 who remained in the study at the age of 15 years, 180 had missing values in one or more of the exposure or outcome variables. Thus, 358 formed the final study groups and were included in the analysis (Supplemental Fig. 1).

Among these 358 participants, the mean (SD) duration of pregnancy was 39 (3) weeks, and the newborn weight was and 3140 (525) grams. Twenty per cent of the infants were born preterm (<37 completed gestation weeks) and 9% presented with low newborn weight (<2500 g). The participants who survived but were excluded because of missing data had rather similar socio-demographic baseline characteristics to those included (Table 1). Twenty-seven per cent of the participants had not completed any years at school by the age of 15 years, and 87% of the mothers of the participants were illiterate.

Raven's Coloured Matrices

The mean (SD) HAZ of the included children was -1.61 (1.24) at one-month-of-age, -2.92 (1.23) at 24-months-of-age and -1.59 (0.96) at 15-years-of-age. The mean (SD, range) values for the outcome variables were 15 (6, 0–33) points for Raven's Coloured Matrices, 4 (2, 0–8) points for mathematics test, 408 (165, 229–2332) ms for median RT and 9 (8, 0–40) lapses in RT assessment.

When HAZ at the age of one month, 24 months and 15 years were included in the analysis (Model 1), none of them was statistically significantly associated with the Raven's Coloured Matrices score (Table 2). In Model 2, adjusted for the confounders, Raven's Coloured Matrices Score was predicted by HAZ at 15-years-of-age (coefficient = 0.85, $P = 0.03$), gender and mother's literacy, but not by HAZ at one or

24-months-of-age. In Model 3, further adjusted for the completed years of school education, schooling predicted Raven's Coloured Matrices score, but the association with HAZ at 15 years was weakened (coefficient = 0.69, $P = 0.06$) (Table 2). Raven's Coloured Matrices score was three points lower in females than males ($P < 0.001$) (Table 2). It was two points higher if the mother of the participant was literate ($P = 0.02$) and one point higher with each additional year that the participant had completed at school ($P < 0.001$) (Table 2).

Mathematics score

In Model 1 without any adjustments, there were no significant associations between HAZ at any age and mathematics scores. In Model 2, adjusted for the most important confounders, there were statistically significant associations between the mathematics score, wealth index and mother's literacy (Table 3). In Model 3, further adjusted for the number of years of school education, the mathematics score was statistically significantly associated only with the number of years of school education (Table 3). The mathematics score was 0.2 points better with each additional year completed at school ($P < 0.001$).

Reaction time

HAZ and RT median did not reach statistical significance in Model 1 or in Model 2 (Table 4). In Model 3, further adjusted for number of years of school education, schooling was statistically significantly associated with RT median (Table 4). Each additional year completed at school was associated with a 12 milliseconds shorter median RT ($P = 0.01$).

In the non-adjusted Model 1 there were no statistically significant associations between HAZ at any age and RT at 15-years-of-age (Table 5). In the adjusted Model 2, there was a statistically significant association between gender and RT lapses (Table 5). In Model 3, further adjusted for the numbers of completed years of school education, male gender and the school education were statistically significantly associated with

Table 1. Participant characteristics

Variables	Complete data ($n = 358$)	Excluded ^a ($n = 238$)	P -value
Gender, male % (n)	51 (182)	46 (107)	0.22
Mean newborn weight ^b , g (SD)	3140 (525)	3020 (470)	0.03
Low newborn weight ^b %, (n)	7.8 (20)	10.7 (14)	0.35
Mean gestational age, weeks (SD)	39.2 (3.3)	39.1 (3.2)	0.65
Preterm births (<37 gestational weeks), % (n)	19.8 (71)	18.3 (41)	0.65
Wealth index, % (n)			
Poor	37.2 (133)	39.0 (85)	0.81
Middle	38.6 (138)	39.0 (85)	
Rich	24.3 (87)	22.0 (84)	
Father's occupation, % (n)			
Fisherman	40.5 (145)	35.5 (81)	0.16
Other	23.5 (84)	19.7 (45)	
Trader	19.0 (68)	21.1 (48)	
Farmer	17.0 (61)	23.7 (54)	
Father's literacy, literate, % (n)	43.3 (155)	41.7 (95)	0.70
Mother's literacy, literate, % (n)	12.9 (46)	17.1 (39)	0.15

^aExcluded because of loss to follow-up or missing values. The final numbers of participants in the comparison vary between 386 and 596.

^bMeasured at ≤ 7 days of age.

Table 2. The association between growth and Ravens Coloured Matrices score at 15-years-of-age (180 months)

Regressor	Raven's coloured matrices					
	Model 1 ^a		Model 2 ^b		Model 3 ^c	
	Coef. [95% CI]	<i>P</i> -value	Coef. [95% CI]	<i>P</i> -value	Coef. [95% CI]	<i>P</i> -value
HAZ_1 ^d	0.01 [-0.54,0.56]	0.96	-0.01 [-0.55,0.53]	0.97	-0.02 [-0.53,0.49]	0.95
HAZ_24 ^d	0.03 [-0.56,0.62]	0.93	-0.12 [-0.68,0.44]	0.67	-0.02 [-0.56,0.51]	0.93
HAZ_180 ^d	0.37 [-0.37,1.10]	0.33	0.85 [0.09,1.61]	0.03	0.69 [-0.03,1.41]	0.06
Female sex			-2.69 [-3.94,-1.43]	<0.001	-2.62 [-3.81,-1.43]	<0.001
Gestational duration, weeks			-0.03 [-0.22,0.15]	0.07	-0.07 [-0.24,0.10]	0.43
Wealth index			0.67 [-0.12,1.47]	0.25	0.24 [-0.52,1.01]	0.53
Father's occupation			-0.43 [-0.99,0.12]	0.13	-0.36 [-0.88,0.17]	0.18
Father's literacy			0.08 [-1.21,1.22]	0.99	-0.31 [-1.46,0.85]	0.60
Mother's literacy			3.43 [1.64,5.21]	<0.001	2.11 [0.37,3.85]	0.02
School education, years					0.89 [0.62,1.17]	<0.001

^aModel 1 included three HAZ measures as independent variables, adjusted for each other.

^bModel 2 further adjusted for gender, gestational duration, wealth index, father's occupation, father's literacy and mother's literacy.

^cModel 3 further adjusted for completed years of school education.

^dHAZ=height-for-age z-score at one and 24-months-of-age and 15-years-of-age (180 months).

less RT lapses (Table 5). Every additional year completed at school was associated with one less lapse ($P=0.01$).

In sensitivity analyses ran with ordered logistic regression, growth between 24-months and 15-years-of-age was statistically significantly associated with Raven's score (coefficient 0.28, $P=0.024$). Also female sex (coefficient 0.90, $P<0.001$) and mother's

literacy (coefficient 1.10, $P=0.002$) were statistically significantly associated with Raven's score as they were in the analysis by least square regression. Similar to the results from least square regression model, there was no association between growth and mathematics score (each $P>0.38$) nor growth and the number of lapses in RT test (each $P>0.34$). Mother's literacy was statistically significantly associated with mathematics score

Table 3. The association between growth and mathematics score at 15-years-of-age (180 months)

Regressor	Mathematics					
	Model 1 ^a		Model 2 ^b		Model 3 ^c	
	Coef. [95% CI]	<i>P</i> -value	Coef. [95% CI]	<i>P</i> -value	Coef. [95% CI]	<i>P</i> -value
HAZ_1 ^d	0.01 [-0.18,0.19]	0.95	0.02 [-0.17,0.21]	0.86	0.02 [-0.17,0.20]	0.87
HAZ_24 ^d	-0.03 [-0.23,0.17]	0.78	-0.03 [-0.23,0.17]	0.78	-0.001 [-0.19,0.19]	0.99
HAZ_180 ^d	0.23 [-0.16,0.48]	0.07	0.21 [-0.06,0.47]	0.13	0.16 [-0.10,0.42]	0.22
Female sex			-0.02 [-0.47,0.42]	0.92	-0.005 [-0.43,0.42]	0.98
Gestational duration, weeks			-0.03 [-0.09,0.04]	0.38	-0.04 [-0.10,0.02]	0.23
Wealth index			0.30 [0.02,0.58]	0.04	0.18 [-0.10,0.45]	0.20
Father's occupation			-0.06 [-0.25,0.14]	0.56	-0.04 [-0.23,0.15]	0.70
Father's literacy			-0.23 [-0.66,0.20]	0.30	-0.31 [-0.73,0.10]	0.14
Mother's literacy			0.68 [0.05,1.31]	0.03	0.32 [-0.31,0.95]	0.32
School education, years					0.24 [0.14,0.34]	<0.001

^aModel 1 included three HAZ measures as independent variables, adjusted for each other.

^bModel 2 further adjusted for gender, gestational duration, wealth index, father's occupation, father's literacy and mother's literacy.

^cModel 3 further adjusted for completed years of school education.

^dHAZ = height-for-age z-score at one and 24-months-of-age and 15-years-of-age (180 months).

Table 4. The association between growth and RT median at 15-years-of-age (180 months)

Regressor	RT median					
	Model 1 ^a		Model 2 ^b		Model 3 ^c	
	Coef. [95% CI]	<i>P</i> -value	Coef. [95% CI]	<i>P</i> -value	Coef. [95% CI]	<i>P</i> -value
HAZ_1 ^d	-9.07 [-24.66,6.52]	0.25	-5.77 [-21.81,10.26]	0.48	-5.73 [-21.62,10.17]	0.48
HAZ_24 ^d	1.21 [-15.37,17.78]	0.89	4.17 [-12.56,20.89]	0.62	2.87 [-13.74,19.47]	0.73
HAZ_180 ^d	0.24 [-20.59,21.08]	0.98	-3.77 [-26.28,18.75]	0.74	-1.63 [-24.00,20.74]	0.89
Female sex			26.63 [-10.73,63.99]	0.16	25.71 [-11.33,62.74]	0.17
Gestational duration, weeks			-3.56 [-9.02,1.90]	0.20	-3.07 [-8.49,2.34]	0.27
Wealth index			-14.14 [-37.76,0.29]	0.24	-8.51 [-32.27,15.25]	0.48
Father's occupation			-9.39 [-25.86,7.08]	0.26	-10.39 [-26.73,5.95]	0.21
Father's literacy			-9.90 [-46.08,26.28]	0.59	-5.76 [-41.74, 30.22]	0.75
Mother's literacy			-19.76 [-73.01,33.48]	0.47	-2.42 [-56.67,51.83]	0.93
School education, years					-11.74 [-20.26,-3.21]	0.007

^aModel 1 included three HAZ measures as independent variables, adjusted for each other.

^bModel 2 further adjusted for gender, gestational duration, wealth index, father's occupation, father's literacy and mother's literacy.

^cModel 3 further adjusted for completed years of school education.

^dHAZ = height-for-age *z*-score at one and 24-months-of-age and 15-years-of-age (180 months).

(coefficient 0.76, $P=0.002$), and female participants scored statistically significantly more lapses than male participants (coefficient 0.85, $P < 0.001$), as was in the case of least square regression.

In supplementary analyses, we applied an alternative set of regression models that were otherwise identical to the models in Tables 2 to 5 but used conditional growth measures (unexplained residuals) at 24 months

and 15 years instead of HAZ at 24 months and 15 years as the exposure variables. Similar findings were obtained that (1) HAZ at one month and conditional growth at 24 months did not show statistically significant association with any of the four outcomes with or without adjustment for the potential confounders in Model 2 of Tables 2 to 5 (each $P > 0.29$); (2) conditional growth at 15 years was associated with Raven's

Table 5. The association between growth and RT lapses at 15-years-of-age (180 months)

Regressor	RT lapses					
	Model 1 ^a		Model 2 ^b		Model 3 ^c	
	Coef. [95% CI]	<i>P</i> -value	Coef. [95% CI]	<i>P</i> -value	Coef. [95% CI]	<i>P</i> -value
HAZ_1 ^d	-0.55 [-1.31,0.21]	0.15	-0.56 [-1.33,0.21]	0.15	-0.56 [-1.32,0.21]	0.15
HAZ_24 ^d	0.21 [-0.59,1.02]	0.60	0.35 [-0.45,1.15]	0.39	0.29 [-0.51,1.09]	0.48
HAZ_180 ^d	0.62 [-0.39,1.63]	0.23	0.14 [-0.94,1.22]	0.80	0.24 [-0.84,1.32]	0.66
Female sex			2.72 [0.92,4.51]	0.003	2.68 [0.90,4.46]	0.003
Gestational duration, weeks			0.06 [-0.20,0.32]	0.65	0.08 [-0.18,0.34]	0.53
Wealth index			-0.04 [-1.18,1.09]	0.94	0.22 [-0.92,1.36]	0.71
Father's occupation			0.06 [-0.73,0.85]	0.89	0.01 [-0.77,0.80]	0.98
Father's literacy			-1.45 [-3.20,0.28]	0.10	-1.27 [-3.00, 0.47]	0.15
Mother's literacy			-1.01 [-3.57,1.54]	0.44	-0.21 [-2.82,2.40]	0.88
School education, years					-0.55 [-0.96,-0.14]	0.009

^aModel 1 included three HAZ measures as independent variables, adjusted for each other.

^bModel 2 further adjusted for gender, gestational duration, wealth index, father's occupation, father's literacy and mother's literacy.

^cModel 3 further adjusted for completed years of school education.

^dHAZ = height-for-age *z*-score at one and 24-months-of-age and 15-years-of-age (180 months).

Coloured Matrices score (coefficient = 0.79, $P = 0.029$) after adjustment for confounders but not schooling; and (3) the regression coefficient for conditional growth at 15 years weakened and the P -values became larger (coefficient = 0.65, $P = 0.07$) after further adjustment for schooling.

Discussion

The primary aim of our study was to assess the association between HAZ at one-month-of-age and height gain during two subsequent age periods and cognitive development at 15-years-of-age. In addition, we investigated whether there are some exposures that would confound or mediate these associations. In the cohort of 358 children we found a significant association between HAZ at 15-years-of-age and Raven's Coloured Matrices score, but not with mathematics score or RT median or lapses. Because the model was adjusted for HAZ at 24 months, the coefficient for HAZ at 15 years represents the association between height gain between 24 months to 15 years and the cognitive outcomes (Cheung 2014). Thus, the finding indicated that height gain after the age of 24 months had an association with the score. Male sex, mother's literacy, higher wealth index and more years of school education were associated with better success in Raven's Coloured Matrices score, mathematics test and RT median and lapses at 15-years-of-age. The results were similar in the models which included HAZ scores and conditional growth measures. These similar findings indicate that the results were not affected by collinearity of the exposure variables. In addition, the correlation coefficients between exposure variables HAZ_1, HAZ_24 and HAZ_180 were <0.46 , showing only moderate correlation between them.

The key finding of the present study is that height gain between age 24 months and 15 years predicted scores on the Raven's Coloured Matrices. A secondary finding, based on comparison of regression coefficients without and with adjustment for schooling, is that this association appeared to be partly mediated by schooling. Theoretically, this might indicate a causal pathway, linking good growth to good cognitive development through increased exposure to education. However,

given the observational nature of our study, we cannot rule out a possibility of reverse causality, i.e. cognition either co-existing with or influencing later schooling, as documented earlier by Grantham-McGregor at her collaborators (Grantham-McGregor *et al.* 2007). If that were the case, the regression model with adjustment for schooling might be an over-adjustment. Nevertheless, this uncertainty about the causality interpretation does not affect the key finding that height gain predicted the adolescents' Raven's scores even after adjustment for perinatal factors. There is no way to formally establish a causal effect of schooling on cognition in observational studies. There has been recent discussion that the effect of schooling may be influenced not just by years of schooling but also by quality of schooling (Frost & Little 2014). All of the outcome measures used in this study were independently associated with number of years of school education. This may suggest reasonable quality of schooling in rural Malawi, although that was not particularly investigated in this study. Education itself may lead to better cognitive function (Gorman & Pollitt 1996), but the children with experience from school may also be more familiar with the school-work-kind of tasks, and therefore perform better. Altogether, our findings support the view that growth improvement after the first 1000 days of life may also contribute to better development of cognitive capacity (Crookston *et al.* 2010; Fink & Rockers 2014; Cheung & Ashorn 2010). In contrast to some previous studies (Victora *et al.* 2008; Sudfeld *et al.* 2015), we did not find any significant associations between growth before 24 months and cognitive capacity later in life.

The reliability and criterion validity of Raven's Matrices have been found to be good in Africa (Wicherts *et al.* 2010; Costenbader & Ngari 2001). Our finding that Raven's score and the other three outcome measures were associated with schooling also supported their criterion validity. However, there is a concern that, in Africa, it may not measure 'general intelligence', or 'g', as intended (Wicherts *et al.* 2010). More detailed psychometric assessment and analysis will be needed to clarify its properties. Previous research has shown that education of the mother, but not that of the father, is correlated with child development (Kong *et al.* 2015). Our analysis has also demonstrated this pattern for Raven's and the

mathematics tests, indicating their convergent and divergent validity. While the validity and reliability of Raven's Coloured Matrices (Raven *et al.* 1998) are relatively well documented compared with the mathematics test (Gandhi *et al.* 2011; Strauss *et al.* 2004) and the RT test (Roach *et al.* 2006; Dinges & Powell 1985; Loh *et al.* 2004), we cannot exclude the possibility that the absence of association between the other measures and child growth was because of insufficient level of measurement properties. Further studies that employed more detailed and locally validated measures are needed.

Female gender and mother's illiteracy were associated with lower score in Raven's Coloured Matrices score, and female gender also with more RT lapses in the model with school education included. This suggests independent associations between gender and cognitive capacity, which are not mediated by schooling. There is previous evidence about gender differences and importance of mother's literacy for cognitive skills in low-income populations (Green *et al.* 2009; Casey *et al.* 2011; Abubakar *et al.* 2010), and these associations were further solidified in our study in a rural African setting. The association between wealth index and mother's literacy and the mathematics score were diluted when number of years of schooling was added in the model. This suggests that higher socioeconomic status leads to better education and better skills in mathematics.

The study cohort comprised 97% of all the newborn infants in the study area during the time of enrolment, and our sample represented the target population well. Approximately one quarter of the original cohort died and one quarter had missing values in some of the variables. Only 7% dropped out. Hence, altogether approximately half of the original cohort members were excluded from the study. Participant characteristics were, however, similar in the included and the excluded groups except that the excluded participants were born smaller. The average birth weight was within a normal range in both groups, and there was no significant difference in the proportion of children with low birth weight. Furthermore, the regression analysis results are unlikely to be biased materially by sample selection because the estimates were conditional on HAZ at one month, which correlated

strongly with newborn weight. According to a statistical theory, factors that are conditioned on cannot cause a bias in regression findings (Cheung 2014; Fairclough 2010).

We used globally applicable WHO Multi-centre Growth References (WHO Multicentre Growth Reference Study Group 2006 (312 pages), De Onis *et al.* 2007). The height measurements were taken by trained data collectors with a proper, regularly calibrated equipment, at each age. Earlier research has shown that (1) while some anthropometric measures such as skinfold thickness may have low level of reliability, length and height measures tend to have reliability levels well above 0.9, and (2) while technical measurement errors may change over age, reliability tends to be stable (Ulijaszek and Kerr 1999). Measurement errors lead to under-estimation of regression coefficients by a factor of $(1 - \text{reliability})$ (Montgomery *et al.* 2001). As such, we would expect that our findings were neither materially nor differentially affected by measurement errors of HAZ at different ages, even though formal reliability tests were not conducted. In contrast, in the absence of detailed information on measurement error, we cannot positively exclude the possibility that our conclusion on the absence of association between growth and other outcome measures was biased. Further research is therefore needed to confirm this aspect of the findings. However, because the impact of imperfect reliability is under-estimation, not over-estimation, of association, the key finding that change in HAZ from 24 months to 15 years was positively associated with Raven's scores is unaffected.

Instead of writing, the participants were instructed to point out the correct answer with their finger in Raven's Coloured Matrices test. This may have caused some problems in validity of the test. However, the data collectors were thoroughly trained to perform all the assessments, and the quality of the measurements was regularly monitored. The weaknesses in our study, aforementioned answering technique, the possible problems in validity of Raven's Coloured Matrices in uneducated populations and rather large exclusion rate, are not likely to cause bias in the results. Hence, we believe that the results of this study may be generalized to the target population of rural East African adolescents.

In conclusion, our results support a hypothesis that in rural Malawi height gain between 24 months and 15 years is associated with better cognitive capacity at 15-years-of-age regardless of the child's previous growth. Some of this effect may be mediated through school education. Mother's illiteracy, female gender and less school education were risk factors for worse performance. This emphasizes the importance of the education of girls. In light of these results, it may be further suggested, that in low income settings growth promotion should not be limited to the first 1000 days of life.

Acknowledgements

We are grateful for all the LCSS participants, their parents and the staff at the study clinic.

Source of funding

All the direct research expenses for designing and conducting the study collection, management, analysis and interpretation of the data, preparation, review, approval of the manuscript and decision to submit the manuscript for publication were covered by an Academy of Finland grant (grant number 252075) to Professor Per Ashorn in the University of Tampere. TT received salary from University of Tampere and National School of Clinical Investigation and personal stipends awarded to her by Finnish Cultural Foundation, Häme Regional Fund, Foundation for Pediatric Research and The Research Foundation of the Pulmonary Diseases.

Conflict of interest

The authors declare that they have no conflicts of interest.

Contributor

The authors' responsibilities were as follows. PA and KM formed the original cohort and conducted the early data collection. TT, YBC and PA were responsible for

the current research design and TT, AP and JV conducted the study. TT and YBC performed and are responsible for the data analysis. TT wrote the first draft of the manuscript; YBC, AP, JV, KM and PA were involved in data interpretation and writing the final version of the manuscript. All authors read and approved the final manuscript. TT had primary responsibility for the final content.

References

- Abubakar A., Holding P., Van de Vijver F.J.R., Newton C. & Van Baar A. (2010) Children at risk for developmental delay can be recognised by stunting, being underweight, ill health, little maternal schooling or high gravidity. *Journal of Child Psychology and Psychiatry* **51** (6), 652–659.
- Berkman D.S., Lescano A.G., Gilman R.H., Lopez S.L. & Black M.M. (2002) Effects of stunting, diarrhoeal disease, and parasitic infection during infancy on cognition in late childhood: a follow-up study. *Lancet* **359** (9306), 564–571.
- Black R.E., Allen L.H., Bhutta Z.A., Caulfield L.E., de Onis M., Ezzati M. et al. (2008) Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet* **371** (9608), 243–260.
- Casey B.M., Dearing E., Vasilyeva M., Ganley C.M. & Tine M. (2011) Spatial and numerical predictors of measurement performance: the moderating effects of community income and gender. *Journal of Educational Psychology* **103** (2), 296–311.
- Chang S.M., Walker S.P., Grantham-McGregor S. & Powell C. A. (2002) Early childhood stunting and later behaviour and school achievement. *Journal of Child Psychology & Psychiatry & Allied Disciplines* **43** (6), 775–783.
- Cheung Y.B. (2014) *Statistical Analysis of Human Growth and Development*. CRC Press: Boca Raton, FL.
- Cheung Y.B. & Ashorn P. (2010) Continuation of linear growth failure and its association with cognitive ability are not dependent on initial length-for-age: a longitudinal study from 6 months to 11 years of age. *Acta Paediatrica* **99** (11), 1719–1723.
- Costenbader V. & Ngari S.M. (2001) A Kenya standardization of the Raven's coloured progressive matrices. *School Psychology International* **22** (3), 258–268.
- Crookston B.T., Dearden K.A., Alder S.C., Porucznik C.A., Stanford J.B., Merrill R.M. et al. (2011) Impact of early and concurrent stunting on cognition. *Maternal & Child Nutrition* **7** (4), 397–409.
- Crookston B.T., Penny M.E., Alder S.C., Dickerson T.T., Merrill R.M., Stanford J.B. et al. (2010) Children who recover from early stunting and children who are not stunted demonstrate similar levels of cognition. *Journal of Nutrition* **140** (11), 1996–2001.

- Crookston B.T., Schott W., Cueto S., Dearden K.A., Engle P., Georgiadis A. *et al.* (2013) Postinfancy growth, schooling, and cognitive achievement: young lives. *American Journal of Clinical Nutrition* **98** (6), 1555–1563.
- De Onis M., Onyango A.W., Borghi E., Siyam A., Nishida C. & Siekmann J. (2007) Development of a WHO growth reference for school-aged children and adolescents. *Bulletin of the World Health Organization* **85** (9), 660–667.
- Dinges D.F. & Powell J.W. (1985) Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers* **17** (6), 652–655.
- Fairclough D.L. (2010) *Design and Analysis of Quality of Life Studies in Clinical Trials*, 2nd edn. Chapman & Hall/CRC: Boca Raton, Florida.
- Feinstein A.R. & Concato J. (1998) The quest for “power”: contradictory hypotheses and inflated sample sizes. *Journal of Clinical Epidemiology* **51** (7), 537–545.
- Filmer D. & Pritchett L.H. (2001) Estimating wealth effects without expenditure data—or tears: an application to educational enrollments in states of India. *Demography* **38** (1), 115–132.
- Fink G. & Rockers P.C. (2014) Childhood growth, schooling, and cognitive development: further evidence from the Young Lives study. *American Journal of Clinical Nutrition* **100** (1), 182–188.
- Frost M. & Little A.W. (2014) Children’s learning practices in Ethiopia: observations from primary school classes. *Oxford Review of Education* **40** (1), 91–111.
- Gandhi M., Ashorn P., Maleta K., Teivaanmäki T., Duan X. & Cheung Y.B. (2011) Height gain during early childhood is an important predictor of schooling and mathematics ability outcomes. *Acta Paediatrica, International Journal of Paediatrics* **100** (8), 1113–1118.
- Gandhi M., Teivaanmäki T., Maleta K., Duan X., Ashorn P. & Cheung Y.B. (2013) Child development at 5 years of age predicted mathematics ability and schooling outcomes in Malawian adolescents. *Acta Paediatrica (Oslo, Norway: 1992)* **102** (1), 58–65.
- Gorman K.S. & Pollitt E. (1996) Does schooling buffer the effects of early risk? *Child Development* **67** (2), 314–326.
- Grantham-McGregor S. (2002) Linear growth retardation and cognition. *The Lancet* **359** (9306), 542.
- Grantham-McGregor S., Cheung Y.B., Cueto S., Glewwe P., Richter L. & Strupp B. (2007) Developmental potential in the first 5 years for children in developing countries. *Lancet* **369** (9555), 60–70.
- Green C.M., Berkule S.B., Dreyer B.P., Fierman A.H., Huberman H.S., Klass P.E. *et al.* (2009) Maternal literacy and associations between education and the cognitive home environment in low-income families. *Archives of Pediatrics & Adolescent Medicine* **163** (9), 832–837.
- Kong F., Chen Z., Xue S., Wang X. & Liu J. (2015) Mother’s but not father’s education predicts general fluid intelligence in emerging adulthood: behavioral and neuroanatomical evidence. *Human Brain Mapping* **36** (11), 4582–4591.
- Kulmala T., Vaahtera M., Ndekha M., Koivisto A., Cullinan T., Salin M. *et al.* (2000) The importance of preterm births for peri- and neonatal mortality in rural Malawi. *Paediatric and Perinatal Epidemiology* **14** (3), 219–226.
- Lecours A.R., Mandujano M. & Romero G. (2001) Nutrition and brain function from infancy to old age. In memory of Joaquin Cravioto (1922–1998), Mexico City, Mexico, March 30–31, 2000. *Nutrition Reviews* **59** (8), S1–S101.
- Loh S., Lamond N., Dorrian J., Roach G. & Dawson D. (2004) The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers* **36** (2), 339–346.
- Maleta K., Virtanen S., Espo M., Kulmala T. & Ashorn P. (2003) Timing of growth faltering in rural Malawi. *Archives of Disease in Childhood* **88** (7), 574–578.
- McCoy D.C. (2015) Poverty, physical stature, and cognitive skills: mechanisms underlying children’s school enrollment in Zambia. *Developmental Psychology* **51** (5), 600–614.
- Montgomery D.C., Peck E.A., Vining G.G. (2001) *Introduction to Linear Regression Analysis*. John Wiley & Sons Inc.: USA, 3rd Edition.
- Perignon M., Fiorentino M., Kuong K., Burja K., Parker M., Sisokhom S. *et al.* (2014) Stunting, poor iron status and parasite infection are significant risk factors for lower cognitive performance in Cambodian school-aged children. *PLoS ONE* **9** (11), e112605.
- Raven J., Raven J. & Court J. (2000) *Manual for Raven’s Progressive Matrices and Vocabulary Scales. Section 3: The Standard Progressive Matrices*. Harcourt Assessment: San Antonio, TX.
- Raven J., Raven J. & Court J. (1998) *Manual for Raven’s Progressive Matrices and Vocabulary Scales. Section 2: The Coloured Progressive Matrices*. Harcourt Assessment: San Antonio, TX.
- Roach G.D., Dawson D. & Lamond N. (2006) Can a shorter psychomotor vigilance task be used as a reasonable substitute for the ten-minute psychomotor vigilance task? *Chronobiology International* **23** (6), 1379–1387.
- Strauss, J., Beegle, K., Sikoki, B., Dwiyanto, A., Herawati, Y. & Witoelar, F. 2004, “The Third Wave of the Indonesia Family Life Survey (IFLS3): Overview and Field Report. WR-144/1-NIA/NICHD”. vol. March.
- Sudfeld C.R., McCoy D.C., Danaei G., Fink G., Ezzati M., Andrews K.G. *et al.* (2015) Linear growth and child development in low- and middle-income countries: a meta-analysis. *Pediatrics* **135** (5), e1266–e1275.
- Sullivan L.M. & Sr D’Agostino R.B. (2003) Robustness and power of analysis of covariance applied to ordinal scaled data as arising in randomized controlled trials. *Statistics in Medicine* **22** (8), 1317–1334.
- Teivaanmäki T., Cheung Y.B., Kortekangas E., Maleta K. & Ashorn P. (2015) Transition between stunted and

- nonstunted status: both occur from birth to 15 years of age in Malawi children. *Acta Paediatrica* **104** (12), 1278–1285.
- The Psychological Corporation (1997) *Wechsler Adult Intelligence Scale. 3rd ed. Technical Manual*. The Psychological Corporation: San Antonio.
- Ulijaszek S.J. & Kerr D.A. (1999) Anthropometric measurement error and the assessment of nutritional status. *British Journal of Nutrition* **82** (3), 165–77.
- UN 2015 The State of World's Children 2015: Reimagine the future.
- Vernon P.A. (1983) Speed of information processing and general intelligence. *Intelligence* **7** (1), 53–70.
- Victora C.G., Adair L., Fall C., Hallal P.C., Martorell R., Richter L. et al. (2008) Maternal and child undernutrition: consequences for adult health and human capital. *Lancet* **371** (9609), 340–357.
- Victora C.G., de Onis M., Hallal P.C., Blossner M. & Shrimpton R. (2010) Worldwide timing of growth faltering: revisiting implications for interventions. *Pediatrics* **125** (3), e473–e480.
- Vidmar S.I., Cole T.J. & Pan H. (2013) Standardizing anthropometric measures in children and adolescents with functions for egen: update. *Stata Journal* **13** (2), 366–378.
- Walker S.P., Chang S.M., Powell C.A., Simonoff E. & Grantham-McGregor S.M. (2007) Early childhood stunting is associated with poor psychological functioning in late adolescence and effects are reduced by psychosocial stimulation. *Journal of Nutrition* **137** (11), 2464–2469.
- Walker S.P., Chang S.M., Amika W., Osmond C. & Grantham-McGregor S.M. (2015) Early childhood stunting is associated with lower developmental levels in the subsequent generation of children. *Journal of Nutrition* **145** (4), 823–828.
- Wenger M.J. & Townsend J.T. (2000) Basic response time tools for studying general processing capacity in attention, perception, and cognition. *Journal of General Psychology* **127** (1), 67–99.
- WHO Multicentre Growth Reference Study Group (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: Geneva.
- Wicherts J.M., Dolana C.V., Carlson J.S. & van der Maasa H. L.J. (2010) Raven's test performance of sub-Saharan Africans: average performance, psychometric properties, and the Flynn Effect. *Learning & Individual Differences* **20** (3), 135–151.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's website.

Supplemental Figure 1. The flow of study participants.

