Situation Analysis of the Nutrition Sector in Ethiopia

2000-2015

Methodological Report

The Federal Democratic Republic of Ethiopia
Ministry of Health

TABLE OF CONTENT

List of Figures

List of Tables

EXECUTIVE SUMMARY

Early in 2013, the European Union Commission along with the 20 EU Member States represented in Ethiopia plus Norway (EU+) endorsed the EU+ Joint Cooperation Strategy for Ethiopia to ensure a coherent and cohesive response to Ethiopia's development challenges, to improve alignment, harmonization, results-based approach, predictability and transparency, whilst avoiding overlapping or fragmented interventions.

This process is expected to lead progressively towards a framework for Joint Programming in Ethiopia by the year 2016. In preparation for the joint programming status by 2016, the EU+ partners agreed to explore the interest and possibility to launch a pilot joint action in a cluster sector of common interest. The theme of nutrition was selected as one of the pilot actions to test the feasibility of joint, collaborative programming.
In view of the above considerations, the EU+ partners in Ethiopia contacted the UNICEF Ethiopia Country Office requesting them to undertake an extended analysis of the nutrition situation in Ethiopia with special focus on trends and determinants of nutrition developments during 2000-2014. UNICEF in turn contracted Tulane University and an independent consultant, to carry out the study.

This methodological reports combines the methodology of these two reports. The main report with the results and main findings can be found at http://www.unicef.org/ethiopia/nutrition.html

Data were used from four Ethiopia Demographic and Health surveys (EDHS) (2000-2014), with focus on the larger ones of 2000 and 2011. Data on resource flows from donors were from the Organisation for Economic Co-operation and Development / Development Assistance Committee (OECD/DAC), and from a study on nutrition stakeholder mapping, 2013-15. Weighing programme and evaluation data were used. Data errors from EDHS surveys included substantial age heaping and length mismeasurement of children standing when they should have been lying and vice versa; these differed between surveys affecting comparability and were taken into account, for length mismeasurement by eliminating wrong cases; sensitivity analyses were done. A total of 10.8% of the EDHS-derived sample was lost because of these errors. Samples from EDHS surveys were comparable from similarity of long term factors expected to be stable, such as respondents' heights.

Effects on results of seasonality and year-to-year production changes from drought were considered. Details of methods, such as geocoding, mapping, and resource estimates, are given in the respective sections. Please note that these are detailed methods for all analyses; the accompanying compilation report summarizes important findings, which are further detailed in complementary, extensive analytical reports.

LIST OF ACRONYMS AND ABBREVIATIONS

AGP	Agricultural Growth Program
ANC	Antenatal Care
A&T	Alive and Thrive
AEWs	Agriculture Extension Workers
B-G	Benishangul-Gumuz
CBN	Community Based Nutrition
CFI	Chronic Food Insecure
CMAM	Community-based Management of Acute Malnutrition
DRS	Developing Regional States
ECHO	The European Commission's Humanitarian Aid and Civil Protection department
EDHS	Ethiopian Demographic and Health Surveys
EHNRI	Ethiopia Health and Nutrition Research Institute
ENA	Essential Nutrition Action
ENGINE	Empowering New Generations to Improve Nutrition and Economic opportunities
EOS	Enhanced Outreach Strategy
EPHI	Ethiopia Public Health Institute
EU	European Union
EU+	European Union Member States represented in Ethiopia plus Norway
FAO	Food and Agriculture Organization
FIES	Food Insecurity Experience Scale
FMoH	Federal Ministry of Health
GAM	Global Acute Malnutrition
GDP	Gross Domestic Product
GMP	Growth Monitoring and Promotion
GoE	Government of Ethiopia
HABP	Household Asset Building Program
HDA	Health Development Army
HC	Health Centre
HEP	Health Extension Program
HEW	Health Extension Workers
HHFS	Household Food Security
HP	Health Post
iCCM	Integrated Community Case Management
IFA	Iron Folic Acid
IFPRI	International Food Policy and Research Institute
IMNCI	Integrated Management of Newborn and Child Illness
IMR	infant mortality rate

IP	Implementing Partner
IRT	Integrated Refresher Training
IYCF	Infant and Young Child Feeding
JAP	Joint Action Plan
KAP	Knowledge, Attitudes and Practices
LRRD	Linking Relief and Recovery to Development
M&E	Monitoring and Evaluation
M2M	Mother-to-Mother Support Group
MAM	Moderate Acute Malnutrition
MDD-W	Minimum Dietary Diversity – Women
MIS	Management Information System
MoA	Ministry of Agriculture
MoCYWA	Ministry of Children, Youth and Women Affairs
MoE	Ministry of Education
MoH	Ministry of Health
MoLSA	Ministry of Labour and Social Affairs
NCB	Nutrition Coordination Body
NCBNP	National Community Based Nutrition Protocol
NGO	Non-Governmental Organization
NNP	National Nutrition Program
NTC	Nutrition Technical Committee
ODA	Official Development Assistance
OFSP	Other Food Security Program
OTP	Outpatient Therapeutic Program
OWNP	One WASH National Program
PLW	Pregnant and Lactating Women
ppt	Percentage point
PSNP	Productive Safety Nets Program
REACH	Renewed Efforts Against Child Hunger and undernutrition
SAM	Severe Acute Malnutrition
SBCC	Social Behavioural Change Communication
SITAN	Situation Analysis
SNNPR	Southern Nations, Nationalities, and Peoples' Region
SUN	Scale-Up Nutrition
TSF	Targeted Supplementary Feeding
UN SCN	The United Nations Standing Committee on Nutrition
UNICEF	The United Nations Children's Fund
WASH	Water, Sanitation and Hygiene
WB	The World Bank
WHO	World Health Organization

INTRODUCTION

The Government of the Federal Democratic Republic of Ethiopia, in introducing the 2013-2015 National Nutrition Programme (NNP) stated that '… attainment of positive nutrition outcomes will be achieved through evidence based programming …'. The European Union Commission (EU) with UNICEF has sponsored a situation analysis of the nutrition sector, to assess the evidence to be '… used to develop an EU+ Joint Nutrition Strategy and Joint Action Framework for Ethiopia … '; the present report contracted by UNICEF aims to complement current capacity in Ethiopia to complete this task.

The EU in 2013 put forward a roadmap for joint nutrition programming for EU and partners with GoE, in support of the NNP. Part of this involved a situation analysis, which UNICEF agreed to implement on behalf of the EU. The overall intent of this, in sum, was to propose steps that could be followed in support of an EU+ Joint Programming on Nutrition, and an analysis of the nutrition situation was seen as one key component. UNICEF requested Tulane University to undertake a situation analysis of the nutrition sector, with three main tasks: nutrition trend and correlational analysis; mapping of nutrition interventions (including assessment of programmes and resources); and analysis of gaps and opportunities. An independent consultant was hired to contribute to the gap and opportunity analysis, with special emphasis on policy and programme options, and summarizing findings for all components . The main report with the results and main findings along with the briefs can be found athttp://www.unicef.org/ethiopia/nutrition.html. "Supplementary materials" will be put on the website, http://tulane.edu/publichealth/internut/ethiopia-nutrition-project.cfm.

Figure 1. The components of situation analysis

[a] EU. Roadmap for eu+ joint programming on nutrition.
http://capacity4dev.ec.europa.eu/joint-programming/document/eu-joint-cooperation-strategy-ethiopia-27012013-0

[b] Mason JB, Potts KS, Crum J, Hofer R and Saldanha L. Analysis of the Nutrition Sector in Ethiopia. A report to UNICEF and EU. Tulane School of Public Health and Tropical Medicine, Department of Community Health and Behavioral Sciences. New Orleans: 2015.

[c] Ljungqvist, B and Asmare E. A Situation Analysis of Nutrition in Ethiopia. Policy and Program Options. November 2015

METHODS FOR THE NUTRITION SITUATION ANALYSIS

This methodological reports combines the methodology of two reports;

- Mason JB, Potts KS, Crum J, Hofer R and Saldanha L. Analysis of the Nutrition Sector in Ethiopia. A report to UNICEF and EU. Tulane School of Public Health and Tropical Medicine, Department of Community Health and Behavioral Sciences. New Orleans: 2015

- Ljungqvist, B and Asmare E. A Situation Analysis of Nutrition in Ethiopia. Policy and Program Options. November 2015.

Data were used from four Ethiopia Demographic and Health surveys (EDHS) (2000-2014), with focus on the larger ones of 2000 and 2011. Data on resource flows from donors were from the Organisation for Economic Co-operation and Development / Development Assistance Committee (OECD/DAC), and from a study on nutrition stakeholder mapping, 2013-15. Weighing programme and evaluation data were used. Data errors from EDHS surveys included substantial age heaping and length mismeasurement of children standing when they should have been lying and vice versa; these differed between surveys affecting comparability and were taken into account, for length mismeasurement by eliminating wrong cases; sensitivity analyses were done. A total of 10.8% of the EDHS-derived sample was lost because of these errors. Samples from EDHS surveys were comparable from similarity of long term factors expected to be stable, such as respondents' heights. Effects on results of seasonality and year-to-year production changes from drought were considered. Details of methods, such as geocoding, mapping, and resource estimates, are given in the respective sections. Please note that these are detailed methods for all analyses; the accompanying compilation report summarizes important findings, which are further detailed in complementary, extensive analytical reports Data sources Several different types of data relevant to nutrition were used, and the sources and development of these data are described below. These are:

1. Anthropometric and related data.

a. Household surveys: The Ethiopian Demographic and Health Surveys (EDHSs) 2000, 2005, 2011 and 2014 (2014 data were obtained later in the analysis). Most of the analysis was done on EDHSs 2000through 2011, with correlationalanalysis using EDHS 2011 primarily.

b. Data from the National Nutrition Programme (NNP) weighing programme: initially as analyzed up to end-2012,thenusing the MS Access database in UNICEF Addis.

c. Data from NNP evaluation surveys.

2. Resources and Programmes:

a. DAC (Development Assistance Committee of OECD) database: 2012 was initially available, and 2013 become available later.

b. Programme data assembled by the Federal Ministry of Health (FMOH) (and REACH); also combined with EDHS data.

c. Programme data obtained from qualitative interviews conducted at woreda and kebele levels using semi-structured questionnaires.

Note on use of 2014 Mini DHS (MDHS) data: The 2014 MDHS data became available towards the completion of the situation analysis, and thus were assessed for inclusion. The data offered limited sample size, and presented a similar issue of mismeasurement of child length/height based on reported age, as seen in the other surveys (described further on p. 5). Various methods used for handling of errors in measurement of child anthropometry for 2014 resulted in different conclusions in undernutrition prevalence estimates. Thus, the confidence with which results can be reported for the relatively short time period of 2011-2014 is limited. Due to the smaller sample size of 2014 data, 2011 survey data was more appropriate for in-depth assessment of indicators requiring limited age groups (e.g. exclusive breastfeeding among 0-5 month children). For these reasons, analysis of 2014 data is limited to preliminary trend assessment in indicators of child nutrition from 2000-2014).

EDHS DATA

The EDHS 2011 children's recode dataset was used for preliminary assessment. EDHSs from 2000, 2005, 2011, and 2014 were subsequently merged to give a single child-level dataset. Unique identifying variables were created within all datasets prior to the merge to allow for ease of matching variables later on. Identifiers were created at the level of survey year, cluster, household, woman/mother, and child (described further in Annex). Details below (e.g. for outcome variables) are given primarily for EDHS 2011, the one mainly used; procedures for the other three surveys were analogous. Where variable names are used in the report, this is for ease of reference. Original EDHS variables are identified by the letter "V" preceding the number; thus providing the basis from which additional variables were derived. All data handling, merging, and analysis was done in SPSS version 22.

DATA STANDARDIZATION AND CLEANING

Outcome variables

Child outcomes

The data cleaning began by determining the number of cases with valid measurements for the indicators of child nutritional status: height-for-age z-score (HAZ), weight-for-height z-score (WHZ), and weight-for-age z-score (WAZ). Child's age, height, and weight are used to calculate the three indicators. There were 11654 total cases in the EDHS 2011 dataset, but only 10480 with a value for child's age in months, which ranged from 0 to 59 months, or less than 5 years. The height and weight variables were then checked for missing values and some additional values were converted to missing such as those coded 999 or 9999 (called 99xx later), labelled "out of normal range". Once these were relabelled as missing, the total number of cases with valid height, weight, and age was 9892; this matched the number of cases with a non-missing HAZ, WHZ, and WAZ, which were already calculated in the EDHS data (variables HW70, HW71, HW72). The initial and final sample sizes are shown in Table 1. Cases that were missing on HAZ, WHZ and WAZ had not been measured for height and weight (variable V113) due to the child being dead, not found, refusing measurement, or not present: frequencies are in Table 2.

Table 1. Cases with non-missing height, weight, and age by survey year (un-weighted)
** Missing on variable HW15 – height: lying or standing (no cases missing on this variable in EDHS 2014).*

	Total cases	Non-missing: age (0-59 mos) / height	ZHAZCL Valid (range: -5.0 to 4.5)	Cases remaining / removed by Mgood = 0
EDHS2000(2000)	10,873	9,560 / 9,084	8,549	8,022 / 512
EDHS 2005 (2005)	9,861	4,455 / 4,206	3,789	3,456 / 324
EDHS 2011 (2011)	11,654	10,480 / 9,893	9,480	8,123 / 1,287
EDHS 2014 (2014)	5,579	5,068 / 4,858	4,358	4,062 / 296
Total	37,967	29,563 / 28,041	26,176	23,663 / 2,419 (+94 missing listwise)*

Table 2. Frequency table of variable 'result of measurement – height/weight' (HW13), excluding

	Frequency	Percent	Valid Percent	Cumulative Percent
Dead	846	50.2	50.2	50.2
Not present	161	5.9	5.9	59.7
Refused	252	19.5	19.5	74.7
Other	99	5.9	5.9	
No measurement found in household	328	100.0	100.0	100.0
Total	1686	100.0	100.0	

The primary outcome variable used for this analysis is the HAZ score. The next step was to check the range of the HAZ variable. The dataset already limited the range from extreme height measurements by coding cases with heights below 45 cm and above 120 cm as 9996 for HAZ, labelled "out of plausible height range." Similarly, the WAZ variable was coded 99xx for those cases withweights below 2.3 kg. The range of HAZ scores was -6.00 to +5.95 with a standard deviation of 1.76. This range is considered too wide to be accurate, and taking account of the distribution, a smaller range was created.

Considering the unlikeliness that a child's HAZ score is below -5, this was selected as the lower limit since many cases with scores less than this are probably erroneous. The upper limit was set to +4.5, so that extreme cases at the high end of the data would not drive the results, but avoided eliminating as many cases as possible. These decisions are somewhat arbitrary but necessary to accurately assess what is happening with the majority of the data. Descriptive statistics of the HAZ variable (HAZCl) are in Table 3.

Table 3. Descriptive statistics of new HAZ variable (cleaned from HW70) restricted to range of -5.0 to 4.5, EDHS 2011 only, by age group

Age group	N	Minimum	Maximum	Mean	Std. Deviation
0 to 23	3451	-4.98	3.99	-9319	1.72600
24 to 59	4582	-5.00	3.58	-1.72275	1.38852
Total: 0 to 59	9357	-5.00	3.99	-1.5480	1.63536

The merged dataset with EDHSs 2000, 2005 and 2011 required the application of the WHO macro to create the HAZ, WAZ, and WHZ variables since EDHSs 2000 and 2005 had used the old reference standards.1 The special SPSS version of the macro for EDHS individual data was employed; the links to download the syntax can be found here: http://www.who.int/childgrowth/software/en/. This was done to the EDHS 2011 as well, even though the existing z-scores used the new WHO reference population. The resulting z-score variables in EDHS 2011 were slightly different from those in the original 2011 dataset (variable HW70 and HAZCL, described above), but very minimally so.

Woman outcomes

In addition to child HAZ, Women's BMI was assessed from the EDHS 2011, as prevalences <18.5(kg/m2). Anaemia was calculated from the same survey using the haemoglobin levels in the data, not adjusted for altitude (thus directly comparable across regions), with the main indicator being anaemia in non-pregnant women (cut point 12.0 g/dl).

Lying/standing, length/height measurement errors

Many children were improperly measured for height or length for their reported age, based on the variable that indicates this measurement type, V115. This is important because adjustments are made to correct for incorrect measurement for age during the calculation of height dependent z-scores. The WHO standards for assessing anthropometry indicate that recumbent length should be used for the measurement of children less than 24 months and standing height should be used for children 24 months and older.2 A cross tabulation of child's height with the variable indicating whether height was measured lying or standing clearly illustrates this problem, see Table 4. WHO also states that recumbent length is approximately 0.7 cm greater than standing height due to gravitational and compressive forces;2 this is the adjustment made during z-score calculation, i.e. 0.7 cm is added to a child's height if measured standing when it should have been lying, and 0.7 cm is subtracted from a child's height which if it was measured lying when it should have been standing. However, applying such a correction in this dataset would not suffice to correct the problem since cases measured standing are, on average, 3 cm taller than those measured lying down at every age (in months) for children ages 21 to 27 months. The error was more extensive in the direction of incorrectly measured lying among the 24 months and older: overall 19.5% of measurements were lying when they should have been standing, according to reported age (see Table 5), with 4.9% the other way. Probably these errors were in part because of miscommunication or misreporting of age. In any event, the differences are so large that they constitute a worrying source of possible errors in analysis, and in proper calculation of child's HAZ.

Table 4. Mean height in millimetres by month of age and height measurement type (EDHS 2011, un-weighted)

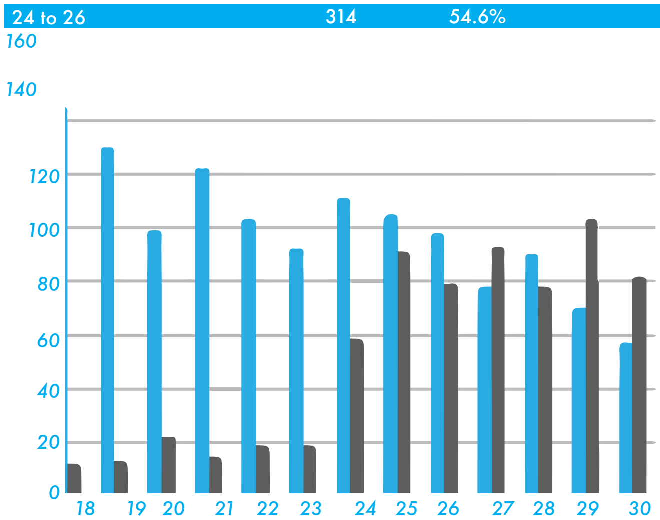
Age in months	Measurement type	Mean	N	Std. Deviation
20	Lying	775.05	99	49.874
	Standing	822.18	22	72.370
	Total	783.62	121	57.281
21	Lying	782.05	122	52.007
	Standing	804.60	15	55.229
	Total	784.52	137	52.634
22	Lying	778.18	103	44.341
	Standing	829.53	19	61.476
	Total	786.18	122	50.686

[d]The difference is believed to be due to a difference in rounding.

23	Lying Standing Total	788.92 808.68 792.31	92 19 111	48.091 59.200 50.425
24	Lying Standing Total	793.71 833.24 807.43	111 59 170	48.937 60.781 56.416
25	Lying Standing Total	796.14 835.81 814.56	105 91 196	56.577 59.892 61.288
26	Lying Standing Total	791.46 837.23 812.01	97 79 176	53.082 66.242 63.424
27	Lying Standing Total	798.04 828.52 814.61	78 93 171	46.028 58.476 55.154
28	Lying Standing Total	803.47 838.54 819.75	90 78 168	46.749 67.684 59.873
29	Lying Standing Total	801.39 837.52 822.90	70 103 173	49.565 68.269 63.764
30	Lying Standing Total	808.84 863.29 841.20	56 82 138	44.495 52.367 56.003

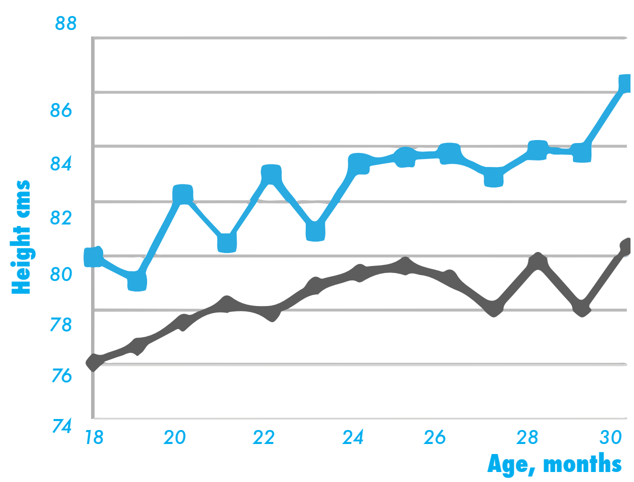
Table 5. Frequency of cases with height measured lying and standing by age group (EDHS 2011, un-weighted)

	Lying		Standing		
Age Group(months)	Frequency	Percent	Frequency	Percent	N
Total sample					
0 to 23	3658	88.5%	204	4.9%	3862
24 to 59	1241	19.5%	1241	74.4%	5964
Ages around 2 years					
21 to 23	317	83.0%	53	13.9%	370
24 to 26	314	54.6%	229	39.8%	543



Lying
Standing

Figure 2. Number of childrenby age (18-30 months) measured lying or standing (EDHS 2011)



Lying
Standing

Figure 3. Heights by agemeasured lying or standing (2011 survey)

The numbers of children in EDHS 2011 around 24 months measured lying or standing are illustrated in Figure 2: this should have no children measured standing under 24 months, and none measured lying at 24 months and older. The error is greater in the 24 months and older group. As shown in Figure 3, the differences are up to 5 cm or more, which makes a big difference to the HAZ calculation. Possibly the age (or date of birth) from the questionnaire was recorded at a different time than the length/height measurement, so the age given when deciding whether the child should be lying or standing was different to that recorded; or the guidelines were not followed. Given the large differences in HAZ, it appears that the mismeasured children's ages were probably wrong, and the actual source of error.

The difference in HAZ, for example, at 24 months is between -2.14 (n=106) measured lying (incorrectly) and -0.98 (n=57) measured standing. If wrong age is a significant source of this problem, then adding a correction in terms of cm would be incorrect anyway – it is not that a child's length is really differing by 5 cm between lying and standing, but that children are wrongly categorized by age. Moreover this error varies by survey (see Table 6). In view of this we calculated both excluding and including those wrongly measured. A variable (Mgood) was created to = 1 when the child was correctly measured, and = 0 when not.

Further, as we progressed, the checks for comparability (see below) were applied to the Mgood = 0 or 1 groups as well, to check whether excluding Mgood = 0 biased the results.

Considering the severity of this measurement error, analysis was usually done on separate groups: those less than 24 months who were measured lying down, excluding 4.9% of this age group (2011 data), and those 24 months and older who were measured standing, excluding 19.5% of this age group (2011 data).

Table 6. Percent mismeasured for height/ length by survey year (un-weighted)

	mismeasured length/height			
Survey	standing when should be lying		lying when should be standing	
	<24 mo	18-23 mo	>=24 mo	24-29 mo
EDHS 2000	7.1% (3,632)	16.2% (834)	5.8% (5439)	17.3% (1019)
EDHS 2000	7.1% (3,632)	16.2% (834)	5.8% (5439)	17.3% (1019)
EDHS 2011	5.0% (4,071)	12.3% (813)	19.8% (6266)	50.2% (1099)
EDHS 2014	11.4% (1,859)	30.2% (344)	4.7% (3208)	13.6% (639)

AGE HEAPING

Preference for exact ages (12, 24, 36 … months; also, but less so for 18, 30 …) is common to all surveys where indices rely on age, as with HAZ. There is no good work-around for wrong ages, and the best solution is taking enough time during the survey with a local calendar to find out the birth date; even then some age heaping is found. Thus the usual approach is to hope that the age heaping does not change between surveys, and especially that the rounding down usually experienced stays the same (e.g. a child of 26 months will often be called a 2-year old). Both these features of age heaping can be checked.

In the Nepal report3 we introduced a metric of the extent of age heaping. Briefly, this fits a regression line to the age distribution (e.g. a nearly horizontal line through the vertical bars in Figure 4, derives the residuals (how far each age frequency is from the mean), removes their signs, adds them together, and divides by the total N (i.e. mean absolute residual, MAR). This gives an arbitrary measure that is comparable across different sample sizes and surveys, for the same age range; the value is typically around 0.1, and the lower the better. The frequencies by age for the EDHSs used here are shown in Figure 4. The age heaping got slightly worse with time – the 2000 survey is the best – and the extent of age heaping is for example greater than in Nepal DHS surveys where the MAR was about 0.1.3

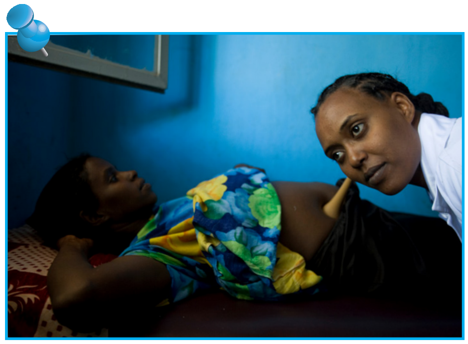
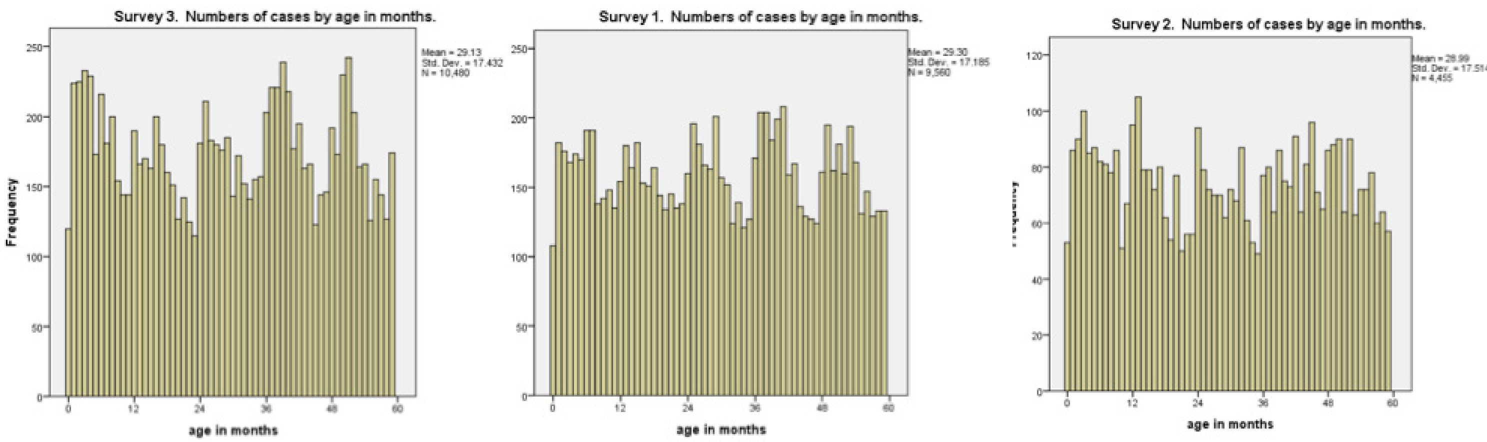


Figure 4. Age heaping by number of cases and age in months by survey



The errors detected by the length/height issue are considered to be a combination of age misreporting and failure to follow the length/height protocol, but these two cannot be distinguished from the present data.

Effects of removing mismeasured cases
Table 7 summarizes the measurement estimates at 18-30 months of age. As was seen in Figure 3, standing measurement are always substantially higher than lying, at all ages. In the 18-23 month group, a minority of children were wrongly measured (96/791 = 12.1%), thus the effect of removing this group on the mean HAZ and prevalence is fairly small, reducing mean HAZ by 0.123 (difference between -2.024 and -1.901). However, in the 24-30 month range, half the length/height measurements are done wrong (608/1193 = 51.0%), so removing those incorrectly measured improves the HAZ much more (0.508, = difference between -1.938 and -1.430). Thus on balance removing the incorrectly measured children improves the HAZ (in 18-30 months children) from -1.923 to -1.752, equivalent to prevalences from 50.1% to 45.0% (these differences are all $p<0.001$). This is the main reason that cleaning the data by removing incorrectly measured children would give an estimate of better nutritional status as height-for-age in the 2011 survey than in the original data.

A different approach is to use weight-for-age, which is not subject to the mismeasurement problem (lying/standing), but is vulnerable to age misreporting. The results reported later show similar trends to HAZ, with a steady improvement of about -1.5 percentage points (ppts)/year, when we remove the mismeasured children, who are also suspected of having misreported ages.

Table 7. Height/length measurements 18-23 and 24-30 months, EDHS 2011

Group	Measurement	HAZ	Height (cms)	Prevalence	N
18-23 months	Wrong (standing)	-1.053	81.1	35.4	96
18-23 months	Correct (lying)	-2.024	77.3	52.0	662
	Total	-1.901	77.8	49.9	758
24-30 months	Wrong (lying)	-2.435	79.7	63.4	571
	Correct (standing)	-1.430	83.9	36.7	559
	Total	-1.938	81.8	50.2	1130
18-30 months	Wrong (lying)	-2.237	79.9	59.4	667
	Correct (standing)	-1.752	80.4	45.0	1221
	Total	-1.923	80.2	50.1	1888

INDEPENDENT VARIABLES

All variables were checked for missing values and appropriate range. Details of all the variables created from existing variables for use in analyses can be found in the Annex. The Annex includes the original EDHS variable name, new variable name, derivation or formula used to recode or calculate, and any notes on missing values, etc. The original EDHS variables used differed, at times, across surveys by name and/or values and categorization. As a result, some variables are created for each survey independently, and then joined together. Merging the dataset prior to re-categorizing variables required careful consideration of what the values meant in the original datasets. The merged dataset took variable values and labels from the 2011 EDHS, which, at times have different meanings than the identical variables and values in the other survey years.

For infant and young child feeding (IYCF) practices, definitions provided by WHO were used, and variables recoded as needed to match these.4 These definitions are in Table 8.

Table 8. Definitions of WHO indicators for assessing infant and young child feeding and definition alterations used for this analysis

Indicator	WHO Definition	Population Recommended disaggregation	Definition used in this analysis (if altered from WHO definition)
Core Indicators			
1. Early initiation of breastfeeding	Proportion of children born in the last 24 months who were put to the breast within one hour of birth.	Children born in the last 24 months. 0 to 12 months 12 to 24 months	
2. Exclusive breastfeed ing under 6 months	Proportion of infants 0 to 5 months of age who are fed exclusively with breastmilk.	Infants 0 to 5 months of age. 0 to 1 months 2 to 3 months 4 to 5 months 0 to 3 months	
3. Continued breast feeding at 1 year	Proportion of children 12 to 15 months of age who are fed breastmilk.	Children 12 to 15 months of age.	
4. Introduction of solid, semi-solid or soft foods	Proportion of infants 6 to 8 months of age who receive solid, semi-solid or soft foods.	Infants 6 to 8 months of age	
5. Minimum dietary diversity	Proportion of children 6 to 23 months of age who receive foods from 4 or more food groups of the following food groups: 1.Grains, roots and tubers 2.Legumes and nuts 3.Dairy products 4.Flesh foods 5.Eggs 6.Vitamin A rich foods 7.Other fruits and vegetables	Children 6 to 23 months of age. 6 to 11 months 12 to 17 months 18 to 23 months report breastfed and non-breastfedseparately	Proportion of children 6 to 23 months of age who receive foods from 3 or more food groups of the following 6 food groups: 1.Grains, roots and tubers 2.Legumes and nuts 3.Dairy products 4.Eggs and flesh foods 5.Vitamin A rich foods 6.Other fruits and vege tables

7. Minimum acceptable diet	Proportion of children 6 to 23 months of age who receive a minimum acceptable diet (apart	Calculate from two separate fractions for breastfed children and non-breastfed	Proportion of children 6 to 23 months of age who receive a minimum acceptable diet.
5. Minimum dietary diversity	from breastmilk). Different definitions for breastfed and non-breastfed: *breastfed children who had at least the minimum dietary diversity and the minimum meal frequency the previous day *non-breastfed children who received at least 2 milk feedings, had the minimum dietary diversity not including milk feeds and had the minimum meal frequency the previous day	children 6 to 23 months of age. 6 to 11 months 12 to 17 months 18 to 23 months	Does not take into account milk feeds. *breastfed and non-breastfed children who had at least the minimum dietary diversity (as defined above in this column) and the minimum meal frequency for their age group (as defined above in this column) the previous day
8. Consumption of iron-rich or iron-fortified foods	Proportion of children 6 to 23 months of age who receive an iron-rich foods or iron-fortified food that is specially designed for infants and young children or that is fortified in the home.	Children 6 to 23 months of age. 6 to 11 months 12 to 17 months 18 to 23 months	

Some general rules applied to recoding all of the categorized variables. All responses of "don't know" or "not a de jure resident" were recoded to missing. Responses of "other" were usually recoded in multiple ways. First, "other" was kept as its own category during the first phase of consolidation to more than 2, but less than the initial number of categories. Second, during the dummy variable formation stage, two variables were created, the first treated "other" as missing, and the second treated "other" as part of the 0 category. For example, roof is the consolidated roofing variable which compressed the existing variable V129, initially 8 to 12 categories across surveys, down to 4 categories, one of which was "other." Roof converted to missing responses of "don't know" and "not a de jure resident." The dummy variable(s) for poor roof take values of 1 for poor roof, and 0 for not poor roof. In the first iteration, D_PoorRoof, "other" is converted to missing resulting in the loss of some cases. In the second iteration, D_PoorRoof2, "other" is included in the 0 category of not a poor roof. Most analysis uses the version that retains the most cases, in this case, D_PoorRoof2. Similarly, toilet facility type (variable V116) was used to assess sanitation facilities: this variable was recoded as a dummy variable where no toilet facility/ bush/ field was coded as 1 and all other toilet facility types were coded as 0, including the "other" category. Source of drinking water (variable V113) was used to assess water supply. This was recoded to a dummy variable that compared those who used surface water (including protected spring, unprotected spring, river, dam, lake, ponds, stream, canal, and irrigation channel) or rainwater as their drinking water source (coded as 1) versus all other sources (coded as 0). These recodes are given in the Annex.

As further examples, child's age and sex were included in all regression models to control for their known effects on height. These variables, HW1 and B4 from the EDHS datasets, did not require additional recoding or cleaning; there were no missing cases and no out of range values – besides the theorized misreporting of age and age heaping described previously. In addition, mother's education was included in most models to control for this always-important determinant. The variable used was highest education level of respondent (variable V106), from which was derived a dummy variable for 'no education', the lowest group alone taking the value 1, versus all other groups taking the value 0. Two variables were used to control for wealth or socioeconomic status: a dummy variable for the poorest wealth quintile (taken from variable V190) and a dummy variable for unimproved roofing (no roof or roof made of thatch, leaf or mud; from variable V129), which is viewed as a proxy for low socioeconomic status. These two variables, poorest wealth quintile and poor roof were never included in the same model due to their high level of collinearity, but rather were used to confirm conclusions of other models. Generally, poor roof was used more often.

ANALYSIS METHODS

DESCRIPTIVE ANALYSIS

Analyses of variables possibly correlated with HAZ begin with assessment of descriptive statistics of the individual predictor variable; means or percentages are presented by groups, usually in national/total, regional, livelihood, and urban/rural categories. Some topics include these descriptors for all survey years, 2000, 2005, and 2011. Tables of bivariate associations with HAZ follow these single variable tables, within the same groups: national, regional, livelihood, and urban/rural. The process was followed for all variables, and these tables are often included in the supplementary material in order to stream line the main text tables. Trends within individual predictor variables and the bivariate associations with HAZ across survey years can be gleaned from these tables.

CORRELATIONAL ANALYSIS

Tables 9 and 10 (taken from the WASH chapter in the Tulane report) display bivariate correlations between most independent/ predictor variables used within the two age groups: cases less than 24 months measured lying down, and cases 24 months and older measured standing. These show the substantial extent of collinearity between important variables, which was useful to guide model specification.

Linear regression using the ordinary least squares technique was employed to assess the relationships between several possible predictive variables and the outcome of child growth, primarily HAZ.

The models were built sequentially, not in a stepwise manner. The first model included only the independent, possibly predictive, variable of interest to clearly determine the un-weighted bivariate association (this differed from the bivariate associations given in descriptive section via tabulation since those were usually weighted). The second model added child's age and gender, to determine any confounding by these variables; for example, age was a strong confounder of the relationship of several IYCF practices, due to varying feeding practices within age groups. The next model usually included education of the mother/ respondent, using variable D_NoEd (no education). The following model included poor roofing to proxy for socioeconomic status, and the final model added mother's height. At times, poor water source and/or no toilet facility were also included, usually after poor roofing and before mother's height.
The interactions between the primary independent variable of interest and other variables in the model were always checked as each variable was included, and were kept in the subsequent models if the interaction variable had a $p<0.1$. If in subsequent models the interaction p-value rose above 0.1, it was then removed for the following model. All interaction variables were calculated by multiplying the two variables together; these are not included in the Annex but all begin with "I_" in the dataset (e.g. I_HomeEd is the interaction term of home delivery and education, which is from D_HomeDel2 * D_NoEd). If an interaction term remained significant at $p<0.1$ throughout all of the models, it was further assessed through tabulation which calculated adjusted means within category of the interaction, calculating cross cell p-values. For example, an interaction between poor water source and education that remained significant in regression analysis would be further explored by calculating the two-by-two table of mean HAZ by water source and education, along with the corresponding 4 p-values. This allowed the direction of the interaction to be understood so conclusions could be drawn. Figures accompany the two-by-two tables for ease of understanding. These two-by-two interaction tables were usually adjusted for all of the other independent variables included in the regression analysis, though some may be un-adjusted
(this is indicated in table titles).

Table 9. Associations of water supply and toilet with height-for-age by age group, EDHS 2011

Age <24 months	
Water (surface)	–0.224 (-4.170, 0.000) n=3420
Toilet (no improved toilet)	-0.150 (-2.794, 0.005) n=3421
Age >= 24 months	
Water (surface)	–0.224 (-4.170, 0.000) n=3420
Toilet (no improved toilet)	-0.150 (-2.794, 0.005) n=3421

Dependent variable: HAZ
Independent variables: age, gender, water source OR toilet dummy variables.
Mgood=1, un-weighted.

Table 10. Interaction of poor water source and no toilet on height-for-age by age group, EDHS 2011

Age <24 months	
Water * Toilet1	–0.224 (-4.170, 0.000) n=3420
Age >= 24 months	
Water * Toilet1	0.020 (0.234, 0.815) n=4551

*1. Water * Toilet indicates the interaction term of poor water source multiplied by no toilet.*
Dependent variable: HAZ
Independent variables: age, gender, poor water, no toilet, interaction of poor water and no toilet, no education, and poor roofing.Mgood=1, un-weighted.

Linear regression using the ordinary least squares technique was employed to assess the relationships between several possible predictive variables and the outcome of child growth, primarily HAZ. The models were built sequentially, not in a stepwise manner. The first model included only the independent, possibly predictive, variable of interest to clearly determine the un-weighted bivariate association (this differed from the bivariate associations given in descriptive section via tabulation since those were usually weighted). The second model added child's age and gender, to determine any confounding by these variables; for example, age was a strong confounder of the relationship of several IYCF practices, due to varying feeding practices within age groups. The next model usually included education of the mother/ respondent, using variable D_NoEd (no education). The following model included poor roofing to proxy for socioeconomic status, and the final model added mother's height. At times, poor water source and/or no toilet facility were also included, usually after poor roofing and before mother's height.

The interactions between the primary independent variable of interest and other variables in the model were always checked as each variable was included, and were kept in the subsequent models if the interaction variable had a p<0.1. If in subsequent models the interaction p-value rose above 0.1, it was then removed for the following model. All interaction variables were calculated by multiplying the two variables together; these are not included in the Annex but all begin with "I_" in the dataset (e.g. I_HomeEd is the interaction term of home delivery and education, which is from D_HomeDel2 * D_NoEd). If an interaction term remained significant at p<0.1 throughout all of the models, it was further assessed through tabulation which calculated adjusted means within category of the interaction, calculating cross cell p-values. For example, an interaction between poor water source and education that remained significant in regression analysis would be further explored by calculating the two-by-two table of mean HAZ by water source and education, along with the corresponding 4 p-values. This allowed the direction of the interaction to be understood so conclusions could be drawn. Figures accompany the two-by-two tables for ease of understanding. These two-by-two interaction tables were usually adjusted for all of the other independent variables included in the regression analysis, though some may be un-adjusted (this is indicated in table titles).

WEIGHTING

The data was weighted according to EDHS guidelines using the variable V005 divided by 1,000,000, calculating the variable named Weight. As recommended by EDHS and is common in statistical analysis, the data was weighted for descriptive statistics calculations, but remained un-weighted for all regression analyses for investigation of correlations. Interactions were usually tested both un-weighted and weighted through un-adjusted and/or adjusted two-by-two tables, then also graphed as figures (described above). The weighted interaction tables are found in the supplementary material. Weighting usually resulted in a diminished, but still present, effect in the interaction tables.
A further effect of weighting, and thereby the likely cause of diminished correlational effects, was the reduction of sample size. As a result, anytime weighting was used, the un-weighted sample size is also given so the observer can be assured of the true sample from which the statistic was derived. For un-weighted analysis, only the un-weighted sample size is given.

FACTORS AFFECTING VALIDITY AND INTER PRETATION

SAMPLE COMPARABIITY BY LONG TERM STABLE FACTORS

Although the samples were drawn to be representative of the regions, and nationally, we can check if this succeeded in producing valid through-time comparisons by checking factors that are associated with child anthropometry, but should change little over the time period studied. Variables such as height and age of respondent are clear candidates for this.

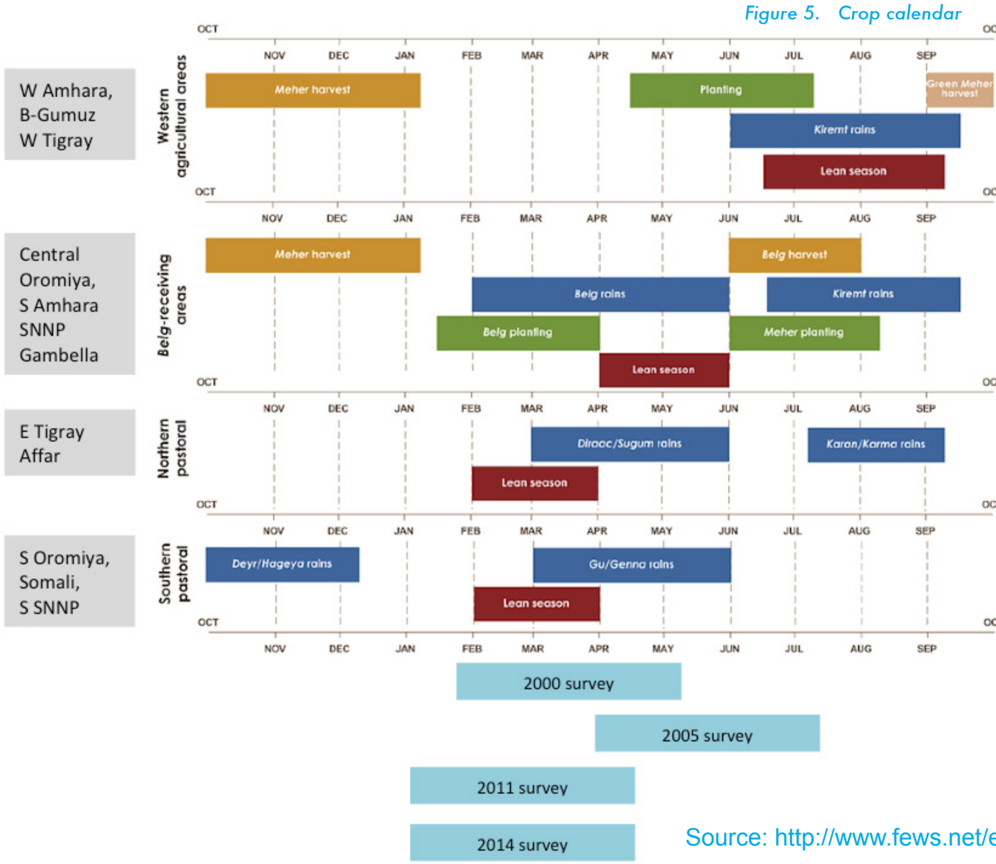
The mean respondent's height was almost identical between surveys: 2000, 156.6 cms; 2005, 156.9 cms; 2011, 156.6 cms. Age of respondents was similarly unchanged, with a mean of 29.5 years. Examining these results by region, again the results were nearly identical, with less than a cm difference across surveys (except Tigray, which had 155.4, 157.0, and 156.0 cms). This was taken as reasonable evidence that the surveys were comparable in terms of likely SES.

EFFECTS OF SEASONALITY AND DROUGHT

A crop calendar is shown in Figure 5, with tentative assignments of regions to the different cropping systems. This is needed mainly to understand how the surveys are comparable taking account of the lean season (shown in red); the survey timings are included. A map of administrative regions is included, as Figure 6. The 2000 and 2011 surveys should be comparable, taking place during the lean season in the pastoral areas (lower two calendars), before and into the start of the lean season in the 'belg-receiving areas', and before the lean season in the western agricultural areas.

The 2005 survey was after the lean season in the first two, in the lean season in belg areas, and going into the lean season in the west.

Figure 5. Crop calendar



Source: http://www.fews.net/east-africa/ethiopia

Figure 6. Administrative regions and zones of Ethiopia

Source: http://faostat3.fao.org/browse/area/238/Ea

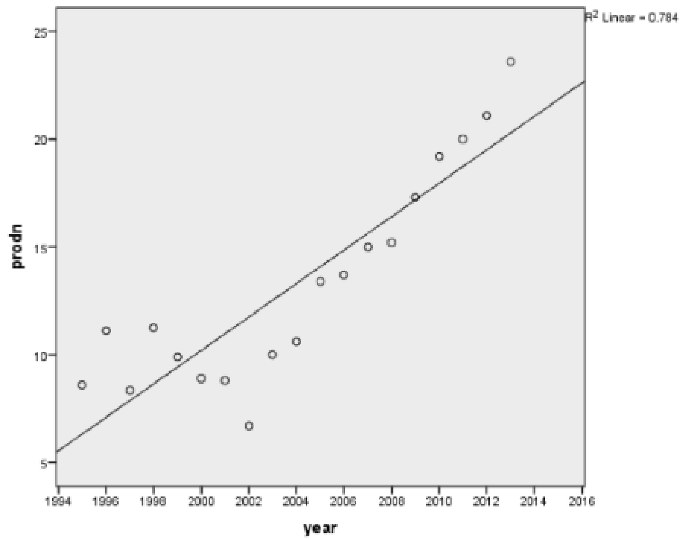
There is not much information on the expected effects of seasonality on anthropometric indices. A study conducted by Tulane University a few years ago5 (estimated that up to 5 ppts change could occur in wasting (low wt-ht) prevalence between lean and post-harvest seasons. Presumably this would be reflected in weight-for-age, but probably less, and later, in height-for-age.

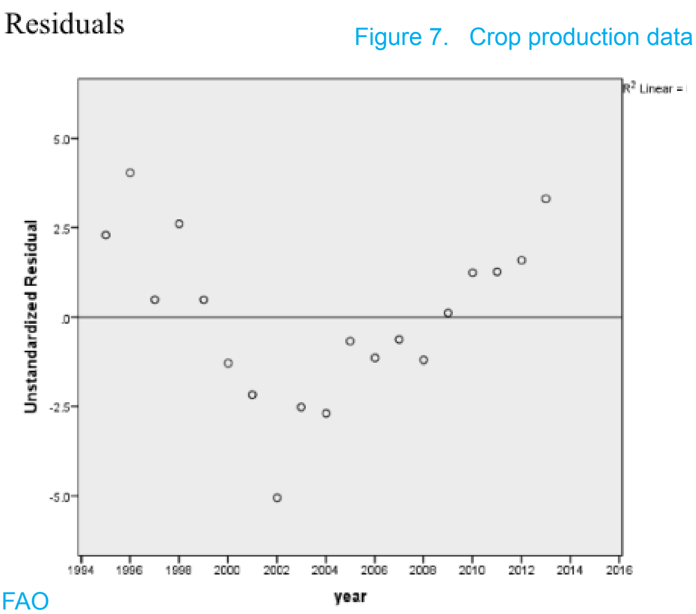
We also need to know the context of food security, in Ethiopia related particularly to drought. National data are easier to get for this than regional. Information on drought and food crop production can be used to assess the extent to which these may affect the survey results. Production estimates were evaluated in5, for the period 1989-2005, and similar procedures were used here for data retrieved through 2013, from the FAO Foodcrops and Shortages series6.

"National data are easier to get for this than regional"

The plot of these estimates is shown inFigure 7, and the deviations from the averages (residuals) help to define drought years at national level. In addition, FAO reports from around the survey time were looked up for information on areas specifically affected.

Figure 7. Crop production data





Source: FAO

The national level results (Figure 7) indicate that in the period 2000-2004 production was far below the trend, with the peak deficit in 2002, and only beginning to recover in 2005. After 2010 production increased rapidly, and was above average in the period around 2011 and even more in 2014.
Linear regression using the ordinary least squares technique was employed to assess the relationships between several possible predictive variables and the outcome of child growth, primarily HAZ.

Drought and production factors affecting 2000 survey results

Attempts were made to find more regional details around the times of the surveys, to facilitate interpretation of the nutritional data, drawing mainly on FAO 'Crop Prospects and Food Situation' reports. Harvests from 1995-9 were reported as above the average trend, but dropped below this in 2000. Since most surveys were in January-May, the effects of falling production may not have been felt by then in the highland agricultural areas. However, from the FAO report in April 2000: "In the pastoral areas of the east and south, particularly the Somali Region, which have had three consecutive years of little or no rainfall, (are) cause for serious concern. The current drought has killed large numbers of livestock and people are migrating in search of water and food." Further: "The secondary "Belg" season crop, which accounts for up to 10 percent of annual grain production, has failed. With this failure, the number of people in need of assistance has increased to more than 10 million people, including 400 000 displaced by the border war with Eritrea." (FAO, Aug 2000).

Drought and production factors affecting 2005 survey results

Harvests improved in 2004, but FAO reported in February 2005 that over 2 million people would need food assistance. In September 2005, good prospects were reported for western and central Ethiopia, less good for eastern and southern crop producing areas. Rainfall had been below average in the southeast (FAO, Apr 2006), and food problems were 'particularly serious' there. FAO estimated (Apr 2006) that 10 million people were vulnerable, of which an acute drought emergency affected 2.6 million, 1.7 million in the south-eastern pastoral areas.

Thus in mid-2005 the population in western and central Ethiopia cropping areas was recovering from a severe drought which had started in 2000. The south-eastern pastoral areas were still badly affected.

Drought and production factors affecting 2011 survey results

The year 2011 was generally good overall for harvests, but in the southern pastoral area rains failed and these areas did badly. Belg rains were also poor, but this should have been during the survey, thus not affecting its results much. The rest of the country contributed to production at better than trend levels (2010 and 2011). (From FAO reports March and June 2011.)

Wasting prevalences and increased mortality risk

In analysis of anthropometric and mortality data from the Horn of Africa, published in 2010,7 we found that the relation between wasting prevalences (WHZ < -2SDs) and child mortality was (a) non-linear, mortality only rising after a certain prevalence was reached; and (b) the wasting level at which this increased mortality started was very different between agricultural and pastoral livelihoods. This meant that the guidelines from WHO8 that gave single cut-offs (e.g. 15% wasting as an emergency) irrespective of livelihood are inappropriate. This is in line with the different growth patterns between these groups 9. We estimated approximately that for agriculturalists, the cut-point should be 10% wasting as the level at which child mortality may increase above normal; and for pastoralists, at 20% wasting child mortality may increase above normal. These cut-points are used in interpreting the data here.

RESOURCES AND PROGRAMMES

DEVELOPMENT ASSISTANCE COMMITTEE DATABASE

Information from the 2012 and 2013 DAC (Development Assistance Committee of OECD) databases was analysed to gain information on spending in Ethiopia for nutrition-specific and nutrition-sensitive programmes. Initially only the 2012 database was available and slightly different methodologies were used to analyse the two due to time restrictions and lessons learned from the first round of analysis. For the 2012 database, all projects under the DAC sector name 'Basic Health' and purpose name 'Basic Nutrition' (purpose code 12240) were analysed for relevance to the current project. A sample of projects from relevant DAC purpose codes based on the recommendations in the SUN Donor Network 'Draft Methodology and Guidance Note to Track Global Investments in Nutrition' was analysed to determine relevance (purpose codes included: 12110, 12220, 12250, 12261, 12262, 12281, 12240, 13020, 13010, 14030, 14031, 14032, 14010, 14015, 14020, 14021, 14040, 14050, 14081, 15170, 31120, 31181, 31166, 31193, 31150, 31182, 31191, 31161, 31163, 31140, 31110, 31130, 31191, 31210, 43040, 16010, 52010, 72010, 72040, 74050, 73010). For the 2013 database, only projects with over 1 million USD committed were analysed as they contributed 89.8% of the total project funds committed in Ethiopia for that year. The same purpose codes and inclusion criteria were used for each year.

PROGRAMME DATA ASSEMBLED BY FMOH

The analysis used a database assembled through a stakeholder mapping conducted in 2013 by FMOH with support from FAO, UNICEF, WFP, and WHO through the Renewed Efforts Against Child Hunger and undernutrition (REACH) initiative.10 An MS Access database was developed to document the multilateral and bilateral agencies and NGOs working in nutrition in Ethiopia; the woredas and zones where they actively supported nutrition programmes over the period 2013-2015; and to classify existing nutrition programmes into 54 types nutrition interventions identified from the 2013-2015 National Nutrition Programme (NNP) Guidelines.

The database was constructed from questionnaires and interviews with 49 organizations working in health and nutrition in Ethiopia, of which 40 organizations reported working in nutrition and 32 reported active nutrition programmes implemented at sub-national level over the period 2013-2015. Data were reported for all 802 woredas in the country. The 54 types of nutrition interventions in the database fell under the following categories: infant and young child feeding (IYCF); micronutrients; management of malnutrition; women and adolescent nutrition; water, sanitation, and hygiene (WASH); nutrition and infectious diseases; nutrition and non-communicable disease; nutrition, agriculture and food security; school health and nutrition; multisectoral approaches; capacity building, research, and monitoring/evaluation; and nutrition communications. The database also included information on the livelihood zone of each woreda; its 'hotspot' classification ; the implementation of large-scale programmes (CBN, ENGINE, and PSNP); and the duration of each programme. Data were not available on sub-woreda implementation (i.e. the coverage within the woreda). Financial resources associated with each programme were also collected in the stakeholder mapping exercise, but were not available for this analysis.

PROGRAMME DATA FROM FIELD OBSERVATION

Qualitative work was undertaken to better understand interventions in Ethiopia with the potential to address need (identified in the correlational analyses), and identify gaps and opportunities in current interventions. Case studies of two types were undertaken: 1)selected project case studies and 2)selected woreda case studies Note that the term 'project' is used for ease, as all identified support the overarching Government of Ethiopia's National Nutrition Programme (NNP). In-depth project case studies were conducted using desk review and interviews with project staff. Six projects were identified as relevant to factors associated with undernutrition based in part on the correlational analysis results, and having sufficient scale to address factors either currently, or the potential to do so in the future. Each project is described to the extent information could be obtained, highlighting important aspects of intervention area, coverage and implementation related to nutrition database. In-depth interviews and focus group discussions (FGDs), using semi-structured questionnaires, were used to collect data from frontline workers across multiple sectors (e.g. health, agriculture) and community volunteers. A total of eight woredas from five regions were selected for field assessment based on outcomes. Detailed descriptions of each project with references can be found in the Supplementary Material.

Woreda case studies were undertaken via field assessment, to better understand implementation of projects at the local level, supplementing information obtained from the FMOH/REACH review of current information on project implementation (including components), coverage, geographic location, livelihood and data availability. Woreda case studies are further described in the Supplementary Material.

MAPPING

DATA SOURCES

Using the EDHS 2011 child level data, maps of anthropometry and determinant/ situation variables were developed at the zone level. The FMOH Nutrition Stakeholder mapping database, developed by REACH,10 was used to identify programme presence, and is further described above. Using this database, maps of programmes by type or component were created at the zone level. Overlaying these maps identified zones deemed to be in highest need of certain programmes as well as those currently receiving the programmes they need.

The data used for the mapping section includes (further explanations are provided in the relevant sections):

EDHS 2011 children's recode dataset

- o Selecting only those children with the correct height measurement for their age, lying or standing (mgood=1).
- o Cases were weighted for all maps.

EDHS 2011 geographical data

- o Identifies the latitude and longitude coordinates of the approximate centre of each cluster.
- o Identifies lower levels of administrative areas, including zone and woreda, not available in primary EDHS datasets

FMOH/REACH programme database

- o Identifies programme presence at the woreda level. Includes both general-category of programmes and more specific components. For example, the WASH category includes both water and hygiene components.

GADM map shape files of Ethiopia (www.gadm.org)

- o The shape files were used to develop the maps in open source QGIS software. This is the source of the administrative boundaries depicted in all maps, which are not meant to be authoritative or exact.

CLUSTER-LEVEL GEOGRAPHICAL DATA

Cluster level geographical data is available in the form of latitude and longitude coordinates for the EDHS clusters through the DHS website. The geographical data includes information on livelihood groups, along with the zone and woreda in which the cluster falls. To protect the confidentiality of individuals, DHS states that the clusters have been randomly displaced by 0 to 10 kilometres in any direction, depending on the urban or rural location.11 The random displacement is corrected to fall within the administrative 2 boundaries in which the cluster truly lies; the zone identified by the geographical coordinates should the clusters true zone. Some clusters are missing geographical data due to the difficulty of accurately determining latitude and longitude coordinates in the some field conditions. For EDHS 2011 there are 25 clusters that are missing geographical data out of 596 total clusters. Therefore, these clusters are also missing livelihood, zone and woreda information. Within these clusters were 270 un-weighted cases, among cases whose height was correctly measured for their age (lying or standing, with anthropometry data. As a result, all information on these cases has not been included in the mapping exercise. The 270 lost cases are out of a total un-weighted sample of 8381 cases with the correct height measurement from EDHS 2011, resulting in a loss of only 3.2%. Since EDHS includes the region regardless of geographical data, it is possible to assess the regional location of the lost clusters. Table 11 illustrates the regional distribution of the lost clusters. The region of Somali experienced by far the most lost cases of 146 due to a lack of geographical data. As a result, conclusions from the mapping analysis regarding Somali should be accepted with caution.

Table 11. Regional location and number of cases lost from clusters missing geographical data

Region	Total clusters missing geo data	Total cases within clusters missing geo data	Total cases in region (mgood=1)a	Percent of cases lost due to missing geo data
Tigray	2	13	948	1.37%
Afar	2	8	811	0.99%
Amhara	5.0% (4,071)	12.3% (813)	19.8% (6266)	50.2% (1099)
Oromiya	3	22	1413	1.56%
Somali	8	146	658	22.19%
Benishangul-Gumuz	1	14	710	1.97%
SNNPR	1	7	1110	0.63%
Gambela	5	38	584	6.50%
Harari	0	0	431	0.0%
Addis Ababa	0	0	298	0.0%
Total	25	270	8381	3.22%

[a] The number of cases from the clusters that are missing geographical data are listed un-weighted, and are among only those cases with the correct height measurement (lying or standing) for their age, represented as the total cases in region, which is also un-weighted.

ZONE-LEVEL MAPPING

Zone-level mapping

The geographical EDHS data provides the latitude and longitude locations of the approximate centre of each cluster. This location is then displaced randomly by less than 10 kilometres, as described previously, but is supposedly corrected to remain in the zone that the cluster was originally in prior to displacement. This was checked using the QGIS software. First the locations of the EDHS clusters were imported into QGIS, at the individual case level. Then, the cases/clusters were captured by whatever zone they fell into on the GADM map.

The cases were then exported from QGIS with their newly captured zone from the GADM map. It was then possible to aggregate useful EDHS variables at the zonal level using the zones from the GADM map. However, prior to doing this, the spatially captured GADM zones were cross-checked with the zone names that were given in the EDHS geographical data. Possible deviations due to misspellings and name variations (or duplicate zone name across regions) were overcome by comparing the region and woreda names associated with each zone. If the zone name captured by the GADM map matched the name given by EDHS, there was no concern for possible misplacement of data. However, where they differed, a decision was made as to whether the EDHS given zone, or the spatially captured GADM zone. Although there should have been no mismatched zones, based on the EDHS correction to zone level, the mismatches could have been due to inaccurate boundary placements within the GADM shapefile and/or the consolidation of zones in GADM map as mentioned previously. Generally, the spatially captured zone was used if the zone name did not match due to its inexistence in the GADM database.

Once the process of assigning the best fitting zone to all cases concluded, it was possible to import the necessary child and programme data into the mapping software. First, the relevant child level variables, including the newly merged in programme data, were aggregated at the zone level. This aggregation process took the mean value across the zone of variables from the child data. This resulted in each zone receiving a continuous score between 0 and 1 for all variables that were dichotomous (0 or 1 for yes or no) in the child-level dataset. For example, the variable of “no toilet” from the child-level EDHS data was in the form of 0=no, 1=yes (yes means child’s household has no toilet access). In the aggregated zone-level file the no toilet variable may have a score of 0.46 meaning that 46% of children in the zone are in households with no toilet access (this is only among the children used for mapping). Most of the EDHS variables and all of the REACH programme variables were in this format – originally dichotomous.

A few EDHS variables used were originally continuous, such as the continuous z-scores for anthropometry: HAZ, WAZ, and WHZ. So the zone aggregation simply resulted in the average z-score among all of the children in the zone, and are easily interpreted as such. The zone aggregation resulted in a dataset with 72 zone cases, based on the number of zones in the GADM map. However 6 of the zones did not have any EDHS cases in them; these have been greyed out in all of the maps. In addition, several zones had fewer than 30 un-weighted EDHS cases. The complete list of zones with 0 to 29 un-weighted EDHS cases can be found in Table 2.8. The zones with EDHS cases less than 30 were still mapped (not greyed out); therefore it is important to view conclusions on these zones with caution.

DECISIONS IN MAPPING AND COMBINING GROUPS

Once the EDHS and REACH programme data were aggregated at the zone-level, this data was imported back into QGIS to create the maps. Several decisions were made in regards to displaying information.

It was decided to use three categories of each variable used so as to not oversimplify it into two, but also not create too many possible combinations in the overlay maps. The three categories were created in the form of tertiles, meaning the zones were ranked in order and split into three even groups. However, with some variables, more than one third of the zones fell into either 0% or 100%. In these cases the tertiles are not evenly enumerated. Instead, all zones with 0% are assigned tertile 1, zones greater than 0% but less than the top third of the zones are assigned tertile 2, and the top third of the zones are assigned tertile 3. With this method, the middle tertile is sometimes much smaller in number of zones than the other tertiles. This often occurred with the programme data from REACH as many programmes have a minimal presence, 0% across much of the country, or are very large presence, with 100% coverage in many zones. This method of ranking zones into tertiles is meant to compare that variable across the country, not to predefined cutoff points. The cutoffs for each tertile are indicated in the legend of each map. It is important to absorb where the cutoffs are before reading the map and making conclusions, as they vary greatly across programmes and situation variables. The tertile format is used for all of the single variable maps including: anthropometry (prevalence's), programme presence, and situation (independent EDHS variables).

The process of overlaying the maps required further decisions of how to group various categories into something meaningful. Combining three, tertiled variables, results in 27 distinct categories, which is far too many to be useful. Therefore, a grouping method was designed to indicate various levels of need.

The grouping model consisted of two pieces: need and programme presence. Need was defined by combining the anthropometry and the situational variables. It was decided at this stage to use underweight prevalence for all of the overlay maps since it captures need based on stunting and wasting. So, if a zone was in the lowest tertile (best off compared to other zones) of underweight, or the lowest tertile of the situational variable (ex. no toilet) they were deemed to be a low need zone. Zones that fell in the highest tertile (worst off) of both underweight and situation are considered highest need. So, the zones that fell in the top two tertiles of both underweight and situation, excluding the highest need combination, are deemed to be medium need.

When programme presence is added to these need combinations, 6 groups resulted: Priority A are highest need zones in the lowest (fewest) tertile of programmes; Priority B are medium need zones in the lowest tertile of programmes; Priority Care medium to highest need zones in the middle (some) tertile of programmes; Good Targeting are zones in medium need that are in the highest (most) tertile of programmes; and Best Targeting are zones in highest need with the highest tertile of programmes. Any zones not falling within one of these groups are in the lower need group. Tables presented after each map will further clarify these groups.

PRESENTATION OF MAPS

Sections of maps are based on need for certain programmes. Sections included in mapping are: WASH, PSNP, IYCF/CBN, and access to health services. Within the WASH section are several groups of maps since WASH programmes are comprised of both safe water and hygiene programmes, providing three different programme maps to use. In addition, the situation variables to use with WASH include both poor water source and no toilet. This resulted in 4 groups of maps within the WASH section: (1) WASH programmes with poor water source; (2) WASH programmes with no toilet; (3) water programmes with poor water source; and (4) hygiene programmes with no toilet. The IYCF/CBN section similarly included two groups of maps, this time using the two different programmes. Within each primary section, anthropometry results are presented first, including stunting, underweight, and wasting prevalence in zonal tertiles. These are repeated at the beginning of each main section. Within the sub-sections, the single variable maps of programme presence and situation (i.e. no toilet, poor water source, etc.) are presented in zonal tertiles.
Then, the two overlay maps are presented, starting with the map illustrating priorities A and B only to highlight the highest need zones, followed by the full overlay map with all priority groups presented. Lastly, the tables are given with the numbers of zones that fall into each priority group, along with the original 27 categories from the 3 by 3 by 3 cross tabulation. The specific zones that fall into priority A and B are given in the footnotes of the first table.

To aid in interpretation, colours have been added to the 27-cell cross tabulation table to match the colours of the priority areas depicted in the full overlay map.

Note: Select maps are presented here; the complete set of maps following the above described presentation may be found in the Tulane report.

SUPPLEMENTARY MATERIALS

"Supplementary materials" will be put on the website, http://tulane.edu/publichealth/internut/ethiopia-nutrition-project.cfm.
This contains detailed results that are not essential for the main text, but are important to keep associated with this report. Putting this on the web has the advantage that length is not a problem. It is suggested that 'Supplementary Materials' would normally be accessed when referred to in the different sections, rather than being of interest as self-standing material. The identifying codes (e.g. 2c, 3c) provided below correspond to sections in the Tulane report for ease of location.

Supplementary Materials Outline

2c. Risk factors

Supplementary material includes additional statistics results in tables and figures that are less essential than the primary tables and figures, but do hold some importance to interpretation and conclusions.
•WASH (includes all regional analysis)
•IYCF
•Access to health services

3c. Case studies

Supplementary material includes selected project summaries, woreda case study reports, and National Nutrition Project strategic objectives tables.
•Ethiopia Nutrition Project Summaries – Case Studies
1.The Agricultural Growth Program
2.Alive & Thrive
3.Empowering New Generations to Improve Nutrition and Economic Opportunities
4.UNICEF WASH programmes
5.Productive Safety Net Programme
6.Community Based Nutrition Programme

•Woreda Case Studies
1.Halaba
2.Kindo Koyisha
3.Ofla
4.Sekota
5.Bure
6.Limu bilbilo
7.& 8. Asayita and Kori

National Nutrition Programme (NNP) Strategic Objective Tables
•NNP Strategic Objective 1: Improve the nutritional status of women (15-49 years) and adolescents (10-19 years)
•NNP Strategic Objective 2: Improve the nutritional status of infants, young children and children under-5
•NNP Strategic Objective 4: Strengthen implementation of nutrition sensitive interventions across sectors
NNP Strategic Objective 4:1Agriculture sector
NNP Strategic Objective 4:2 Education sector
NNP Strategic Objective 5: Improve multisectoral coordination and capacity to ensure NNP implementation

REFERENCES

1. WHO Anthro for personal computers, version 3.2.2, 2011: Software for assessing growth and development of the world's children [programme]. Geneva WHO, 2010.

2. 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: WHO, 2006.

3. Crum J, Mason JB, Hutchinson P. Analysis of trends in nutrition of children and women in Nepal. Kathmandu, 2012

4. World Health Organization (WHO). Indicators for assessing infant and young child feeding practices: Definitions. Conclusions of a consensus meeting held 6–8 November 2007. Washington D.C., USA, 2008.

5. Chotard S, Mason JB, Oliphant NP, et al. Fluctuations in wasting in vulnerable child populations in the Greater Horn of Africa. Food and nutrition bulletin 2010;31(3 Suppl):S219-33.

6. FAO. Global Information and Early Warning System. Food Crops and Shortages; then Crop Prospects and Food Situation. Global Information and Early Warning System. http://www.fao.org/giews/english/index.htm, 1990 on.

7. Mason JB, White JM, Heron L, et al. Child acute malnutrition and mortality in populations affected by displacement in the Horn of Africa, 1997-2009. Int J Environ Res Public Health 2012;9(3):791-806.

8. World Health Organization (WHO). The management of nutrition in major emergencies. Geneva, Switzerland, 2000.

9. Tulane-UNICEF/ESARO. Assessment of child nutrition in the Greater Horn of Africa. http://tulane.edu/publichealth/internut/ to come, 2006.

10. Federal Ministry of Health (FMOH) supported by REACH. Nutrition Stakeholder Mapping 2013-15 Final Report, February 2014.

11. Perez-Heydrich C, Warren JL, Burgert CR, et al. Guidelines on the Use of DHS GPS Data. Spatial Analysis Reports No. 8. Calverton, Maryland, USA: ICF International, 2013.

12. Hijmans R. GADM database of Global Administrative Areas. Secondary GADM database of Global Administrative Areas. http://www.gadm.org.

ANNEX: DESCRIPTION OF INDEPENDENT VARIABLES

Purpose of variable(s)	New variable names	Derivation variables and values of new variable	Notes, missing, etc.
Section 1: Classifying / controlling variables			
Maternal education	D_NoEd	From V106, Values: 1=no education 0=any education (primary, secondary or higher)	No missing cases among all four surveys.
Roofing	Roof	From V129; Values: 1=poor roof (thatch/leaf/mud) 2=rudimentary roof (wood/mud, reed/bamboo, plastic sheet, cardboard) 3=finished roof (iron, cement, shingles) 4=other	"Not a dejure resident" converted to missing.
	D_PoorRoof	From Roof; Values: 1=poor roof 0=rudimentary or finished roof	"Other" converted to missing.
	D_PoorRoof2	From Roof; Values: 1=poor roof 0=rudimentary, finished, or other roof	"Other" retained in 0 category, i.e. not a poor roof.
Floor	Floor	From V127; Values: 1=poor floor (earth/sand, dung) 2=rudimentary floor (wood planks, reed/bamboo) 3=finished floor (polished wood, vinyl, cement, carpet) 4=other	"Not a dejure resident" converted to missing.
	D_PoorFloor	From Floor; Values: 1=poor floor 0=rudimentary or finished floor	"Other" converted to missing.
	D_PoorFloor2	From Floor; Values: 1=poor floor 0=rudimentary, finished, or other floor	"Other" retained in 0 category, i.e. not a poor floor.
Poverty	povcat	From D_PoorRoof2, D_SurfaceWat, and D_NoToilet; Values: 1=poor roof, surface water, and no toilet (i.e. extreme poverty) 0=else	
Livelihood	Livelihood		
Wealth	WealthQ	From V190: wealth index in quintiles Combined from all 4 surveys (imported from household dataset if not included in child-level data). EDHS 2000: imported from wealth index dataset.	
	WealthF	From V191: wealth index factor score Combined from all 4 surveys (imported from household dataset if not included in child-level data). EDHS 2000: imported from wealth index dataset.	
	Dwealth1	From WealthQ; Values: 1=wealth quintile 1 (poorest) 0=else (wealth quintile's 2-5	

	Dwealth12	From WealthQ; Values:	
	Dwealth1234	1=wealth quintile 1 and 2	
Age Groups	Age_grp	1From HW1: Age in 3 month groups; Values: 1=0 to 2 months 2=3 to 5 months 3=6 to 8 months …. 9= 24+ months	
	Age_grp2	From HW1: Age in 2 month groups; Values: 0=0 to 1 months 1=2 to 3 months 2=4 to 5 months …. 12= 24+ months	
	Age_grp6	From HW1: Age in 6 month groups; Values: 1=0 to 5 months 2=6 to 11 months 3=12 to 17 months …. 6= 36+ months	
	Age_grp12	From HW1: Age groups in years; Values: 1=0 to 11 months 2=12 to 23 months 3=24 to 35 months 4=36 to 47 months 5=48 to 59 months	
Section 2: Water & Sanitation			
Wealth	Water	From V113; Values: 1=piped 2=public tap/piped outside compound 3=well 4=surface and rainwater 5=other	"Not a dejure resident" converted to missing.
	D_SurfaceWat	From Water; Values:	"Other" included in 0 category (all other water sources).
Toilet	Dwealth1	From V116; Values: 1=flush 2=improved pit 3=open pit 4=no toilet 5=other	"Not a dejure resident" converted to missing
		From Toilet; Values: 1=no toilet 0=any other toilet type	"Other" included in 0 category (any other toilet type).

Section 3: IYCF			
Early initiation of breastfeeding	EarlyBF	From V426; Values: 1=immediately/within 1 hour of birth 0=else	Missing for EDHS 2014.
Exclusive breastfeeding	EBF	From AnyLiq, AnySSS, and BFnow; Values: 1=exclusively breastfeeding 0=not exclusively breastfeeding	Missing for EDHS 2014. If missing for any of the determining variables (i.e. AnyLiq, AnySSS, or BFnow), then missing for EBF.
Currently breastfeeding	BF05	From V404 for 2005; Values: 1=currently breastfeeding 0=not currently breastfeeding	This is for EDHS 2005 (2005) only.
	BF11	From V404 for 2011; Values: 1=currently breastfeeding 0=not currently breastfeeding	Missing for EDHS 2014.
	BFnow	From v404 (2000), BF05 (2005), and BF11 (2011); Values: 1=currently breastfeeding 0=not currently breastfeeding	Missing for EDHS 2014.
Breastfeeding status	BFstat	From BFnow and EBF; Values: 0=non-breastfeeding 1=exclusively breastfeeding 2=non-exclusively breastfeeding	Don't know recoded as missing.
Liquids Consumed	Wat	From V409, EDHS 2011; Values: 1=child drank water in last 24 hours 0=else	Don't know recoded as missing.
	Wat1	From v469a, EDHS 2000; Values: 1=child drank water in last 24 hours 0=else	Don't know recoded as missing.
	Wat2	From V469A, EDHS 2005; Values: 1=child drank water in last 24 hours 0=else	Don't know recoded as missing.
	Juice	From V410, EDHS 2011; Values: 1=child drank juice in last 24 hours 0=else	Don't know recoded as missing.
	Juice1	From v469c, EDHS 2000; Values: 1=child drank water in last 24 hours 0=else	Don't know recoded as missing.
	Juice2	From V469D, EDHS 2005; Values: 1=child drank water in last 24 hours 0=else	Don't know recoded as missing.
	Milk	From V411, EDHS 2011; Values: 1=child drank milk in last 24 hours 0=else	Don't know recoded as missing.
	Milk1	From v469h, EDHS 2000; Values: 1=child drank milk in last 24 hours 0=else	Don't know recoded as missing.
	Milk2	From V469C, EDHS 2005; Values: 1=child drank milk in last 24 hours 0=else	Don't know recoded as missing.
	Formula	From V411A, EDHS 2011; Values: 1=child drank formula in last 24 hours 0=else	Don't know recoded as missing.
	BabFood2	From V469B, EDHS 2005 (infant formula); Values:1=child drank formula in last 24 hours	Don't know recoded as missing.

	0=else		
	Soup	From V412C, EDHS 2011; Values: 1=child had soup in last 24 hours 0=else	Don’t know recoded as missing.
	Tea2	From V469E, EDHS 2005; Values: 1=child drank tea in last 24 hours 0=else	Don’t know recoded as missing.
	OthLiq	From V413, EDHS 2011; Values: 1=child drank other liquids in last 24 hours 0=else	Don’t know recoded as missing.
	OthLiq1	From v469l, EDHS 2000; Values: 1=child drank other liquids in last 24 hours 0=else	Don’t know recoded as missing.
	OthLiq2	From S469A, EDHS 2005; Values: 1=child drank other liquids in last 24 hours 0=else	“Not a dejure resident” converted to missing.
	AnyLiq	From all liquid variables above (Wat through OthLiq2); Values: 1=child drank any liquids other that breastmilk yesterday (including water) 0=else	
Grains	Grain1	From v469q, EDHS 2000; Values: 1=child had grains in last 24 hours 0=else	Don’t know recoded as missing.
	Bread2	From S470C, EDHS 2005; Values: 1=child had bread in last 24 hours 0=else	Don’t know recoded as missing.
	Porridge	From S470A, EDHS 2005; Values: 1=child had porridge in last 24 hours 0=else	
	Teff2	From S470D, EDHS 2005; Values: 1=child had teff in last 24 hours 0=else	Don’t know recoded as missing.
	BabFood2b	From V470B, EDHS 2005 (commercially fortified babyfood) ; Values: 1=child had baby food in last 24 hours 0=else	Don’t know recoded as missing.
	BabFood	From V412A, EDHS 2011 (fortified baby food – cerelac, etc.) ; Values: 1=child had baby food in last 24 hours 0=else	Don’t know recoded as missing.
	Grain	From V414E, EDHS 2011; Values: 1=child had grains in last 24 hours 0=else	
Tubers	Tubers1	From v469r, EDHS 2000; Values: 1=child had tubers in last 24 hours 0=else	Don’t know recoded as missing.
	Potatoes2	From S470E, EDHS 2005; Values:	Don’t know recoded as missing.
		1=child had potatoes in last 24 hours 0=else	
	Tubers	From V414F, EDHS 2011; Values: 1=child had tubers in last 24 hours 0=else	Don’t know recoded as missing.

Egg	Eggs2	From S470M, EDHS 2005; Values: 1=child had eggs in last 24 hours 0=else	Don't know recoded as missing.
	Eggs	From V414G, EDHS 2011; Values: 1=child had eggs in last 24 hours 0=else	Don't know recoded as missing.
Meat	MeatEggs1	From v469v, EDHS 2000; Values: 1=child had meats or eggs in last 24 hours 0=else	Don't know recoded as missing.
	Meat2	From S470K, EDHS 2005; Values: 1=child had meats in last 24 hours 0=else	Don't know recoded as missing.
	Organ2	From S470J, EDHS 2005; Values: 1=child had organ meats in last 24 hours 0=else	Don't know recoded as missing.
	Poultry2	From S470L, EDHS 2005; Values: 1=child had poultry in last 24 hours 0=else	Don't know recoded as missing.
	Fish2	From S470N, EDHS 2005; Values: 1=child had fish in last 24 hours 0=else	Don't know recoded as missing.
	Meat	From V414H, EDHS 2011; Values: 1=child had meats in last 24 hours 0=else	Don't know recoded as missing.
	Organs	From V414M, EDHS 2011; Values: 1=child had organ meats in last 24 hours 0=else	Don't know recoded as missing.
	Fish	From V414N, EDHS 2011; Values: 1=child had fish in last 24 hours 0=else	Don't know recoded as missing.
Vitamin A rich foods	Arich1	From v469o, EDHS 2000; Values: 1=child had vitamin A rich fruits in last 24 hours 0=else	Don't know recoded as missing.
	Orange2	From S470F, EDHS 2005; Values: 1=child had orange vegetables in last 24 hours 0=else	Don't know recoded as missing.
	Green2	From S470G, EDHS 2005; Values: 1=child had green vegetables in last 24 hours 0=else	Don't know recoded as missing.
	AFruits2	From S470H, EDHS 2005; Values: 1=child had vitamin A rich fruits in last 24 hours 0=else	Don't know recoded as missing.
	Orange	From V414I, EDHS 2011; Values: 1=child had orange vegetables in last 24 hours 0=else	Don't know recoded as missing.
	Green	From V414J, EDHS 2011; Values: 1=child had green vegetables in last 24 hours 0=else	Don't know recoded as missing.
	AFruits	From V414K, EDHS 2011; Values: 1=child had vitamin A rich fruits in last 24 hours 0=else	Don't know recoded as missing.

Legumes	Legumes1	From v469w, EDHS 2000; Values: 1=child had legumes in last 24 hours 0=else	Don’t know recoded as missing.
	Beans2	From S470O, EDHS 2005; Values: 1=child had beans in last 24 hours 0=else	Don’t know recoded as missing.
	Nuts2	From S470P, EDHS 2005; Values: 1=child had nuts in last 24 hours 0=else	Don’t know recoded as missing.
	Legumes	From V414O, EDHS 2011; Values: 1=child had legumes in last 24 hours 0=else	Don’t know recoded as missing.
Other fruits and vegetables	OthFV1	From v469u, EDHS 2000; Values: 1=child had other fruits or vegetables in last 24 hours 0=else	Don’t know recoded as missing.
	OthFV2	From S470I, EDHS 2005; Values: 1=child had other fruits or vegetables in last 24 hours 0=else	Don’t know recoded as missing.
	OthFruit	From V414L, EDHS 2011; Values: 1=child had other fruits in last 24 hours 0=else	Don’t know recoded as missing.
Dairy	Dairy2	From S470Q, EDHS 2005; Values: 1=child had dairy in last 24 hours 0=else	Don’t know recoded as missing.
	Dairy	From V414P, EDHS 2011; Values: 1=child had dairy in last 24 hours 0=else	Don’t know recoded as missing.
	Yog	From V414V, EDHS 2011; Values: 1=child had yogurt in last 24 hours 0=else	Don’t know recoded as missing.
Fats	Fats1	From v469y, EDHS 2000; Values: 1=child had fats in last 24 hours 0=else	Don’t know recoded as missing.
	Fats2	From S470R, EDHS 2005; Values: 1=child had fats in last 24 hours 0=else	Don’t know recoded as missing.
Other solid, semi-solid, or soft foods	OthSSS2	From S470S, EDHS 2005; Values: 1=child had any other foods in last 24 hours 0=else	Don’t know recoded as missing.
	OthSSS	From V414S, EDHS 2011; Values: 1=child had any other foods in last 24 hours 0=else	Don’t know recoded as missing.
Any solid, semi-solid, or soft foods.	AnySSS	From all solid, semi-solid or soft food variables above (Grain1 through Oth-SSS); Values: 1=child ate any solid, semi-solid or soft foods yesterday 0=else	
Food 1 (Grains)	Food1_1	From Grain1 and Tubers1; Values: 1=child ate grains, roots, or tubers 0=else	EDHS 2000 only.
	Food2_1	From Porridge2, BabFood2b, Bread2, Teff2, and Potatoes2; Values: 1=child ate porridge, baby food, bread, teff, potatoes, other grains 0=else	EDHS 2005 only.

	Food3_1	From Grain, Tubers and BabFood; Values: 1=child ate grains, tubers, baby food 0=else	EDHS 2011 only.
Food 2 (Legumes)	Food1_2	From Legumes1; Values: 1=child ate legumes 0=else	EDHS 2000 only.
	Food2_2	From Beans2 and Nuts2; Values: 1=child ate beans or nuts 0=else	EDHS 2005 only.
	Food3_2	From Legumes; Values: 1=child ate legumes 0=else	EDHS 2011 only.
Food 3 (Dairy)	Food1_3	From Milk1; Values: 1=child drank milk 0=else	EDHS 2000 only.
	Food2_3	From Milk2 and Dairy2; Values: 1=child had milk or dairy 0=else	EDHS 2005 only.
	Food3_3	From Milk, Formula, and Dairy; Values: 1=child had milk, baby formula or dairy 0=else	EDHS 2011 only.
Food 4 (Meats, poultry, fish, animal foods)	Food1_4	FFrom MeatEggs1; Values: 1=child ate meat, poultry, fish, shellfish or eggs 0=else	EDHS 2000 only.
	Food2_4	From Organ2, Meat2, Poultry2, and Fish2; Values: 1=child ate organ meats, meats, poultry, or fish 0=else	EDHS 2005 only.
	Food3_4	From Meat, Organs, and Fish; Values: 1=child ate organ meats, meat, poultry, or fish 0=else	EDHS 2011 only.
Food 5 (Eggs)	Orange2	Missing “eggs” variable in EDHS 2000	EDHS 2000 only. Missing.
	Green2	From Eggs2; Values: 1=child ate eggs 0=else	EDHS 2005 only.
	AFruits2	From Eggs; Values: 1=child ate eggs 0=else	EDHS 2011 only.
Food 6 (Vitamin A rich foods)	Food1_6	FFrom Arich1; Values: 1=child ate vitamin A rich fruits 0=else	EDHS 2000 only.
	Food2_6	From Orange2, Green2, and AFruits2; Values: 1=child ate orange or green vegetables, or vitamin A rich fruits 0=else	EDHS 2005 only.
	Food3_6	From Orange, Green, and AFruits; Values: 1=child ate orange or green vegetables, or vitamin A rich fruits 0=else	EDHS 2011 only.

Food 7 (Other fruits and vegetables)	Food1_7	From OthFV1; Values: 1=child ate other fruits or vegetables 0=else	EDHS 2000 only.
	Food2_7	From OthFV2; Values: 1=child ate other fruits or vegetables 0=else	EDHS 2005 only.
	Food3_7	From OthFruit; Values: 1=child ate other fruits 0=else	EDHS 2011 only.
Food 8 (Meats and eggs together – EDHS 2005 and 3 only)	Food2_8	From Food2_4 and Food2_5; Values: 1=child ate organ meats, meats, poultry, fish, or eggs 0=else	EDHS 2005 only.
	Food3_8	From Meat, Organs, Fish and Eggs; Values: 1=child ate organ meats, meats, poultry, fish, or eggs 0=else	EDHS 2011 only.
Total number of food groups consumed	Food_grps_tot1	From Food1_1 to Food1_7; Values: 0 through 6, summation of above variables; total number of food groups eaten. *Food1_5 is missing for EDHS 2000; so possible values are 0 through 6.	EDHS 2000 only. Out of 6 total food groups. EDHS 2000 does not have a separate "eggs" category, so only 6 food groups are possible.
	Food_grps_tot2	From Food2_1 to Food2_7; Values: 0 through 7, summation of above variables; total number of food groups eaten.	EDHS 2005 only. Out of 7 total possible food groups.
	Food_grps_tot2b	From Food2_1 to Food2_3, and Food2_6 through Food2_8; Values: 0 through 6, summation of above variables; total number of food groups eaten.	EDHS 2005 only. Out of 6 total food groups; eggs have been combined with meats.
	Food_grps_tot3	From Food3_1 through Food3_7 Values: 0 through 7, summation of above variables; total number of food groups eaten.	EDHS 2011 only. Out of 7 total possible food groups.
	Food_grps_tot3b	From Food3_1 through Food3_3, and Food3_6 through Food3_8 Values: 0 through 6, summation of above variables; total number of food groups eaten.	EDHS 2011 only. Out of 6 total food groups; eggs have been combined with meats.
Minimum Dietary Diversity	Four_of6	From Food_grps_tot1, Food_grps_tot2b, and Food_grps_tot3b; Values: 1=child ate at 4 or more foods out of 6 possible food groups 0=child ate less than 4 foods out of 6 possible food groups	
	Four_of7	From Food_grps_tot2 and Food_grps_tot3; Values: 1=child ate at 4 or more foods out of 7 possible food groups 0=child ate less than 4 foods out of 7 possible food groups	Missing for EDHS 2000.
	Three_of6	From Food_grps_tot1, Food_grps_tot2b, and Food_grps_tot3b; Values: 1=child ate at 3 or more foods out of 6 possible food groups 0=child ate less than 3 foods out of 6 possible food groups	
	Three_of7	From Food_grps_tot2 and Food_grps_tot3; Values: 1=child ate at 3 or more foods out of 7 possible food groups 0=child ate less than 3 foods out of 7 possible food groups	Missing for EDHS 2000.

Minimum Meal Frequency	TotMeals	From M39 (EDHS 2000 and EDHS 2011), and S472 (EDHS 2005); Values: 0 to 7: number of times child ate in the previous 24 hours.	“Don’t know” recoded as missing. Value 7 includes 7 or more feeding times.
	MinFreq	From BFnow, HW1, and M39 (EDHS 2000 and 2011) or S472 (EDHS 2005); Values: 1=meeting the minimum meal frequency of: • 2+ times for breastfeeding 6 to 8 month olds, • 3+ times for breastfeeding 9 to 23 month olds, and • 4+ times for non-breastfeeding 6 to 23 month olds 0=not meeting the minimum meal frequency as described above	“Don’t know” recoded as missing. Missing if M39 (EDHS 2000 and 3) or S472 (EDHS 2005) variable is missing.
	MinFreq2	From BFnow, HW1, and TotMeals; Values: 1=meeting the minimum meal frequency of: •2+ times for breastfeeding 6 to 8 month olds, •3+ times for breastfeeding 9 to 23 month olds, and •4+ times for non-breastfeeding 6 to 23 month olds 0=not meeting the minimum meal frequecy as described above	Missing if To tMeals variable is missing.
	MinFreq3	From BFnow, HW1, and TotMeals; Values: 1=meeting the minimum meal frequency of: •2+ times for breastfeeding 6 to 8 month olds, •3+ times for breastfeeding 9 to 23 month olds, and •3+ times for non-breastfeeding 6 to 23 month olds (equals the number of feedings for breastfed children since number of milk feeds are not available) 0=not meeting the minimum meal frequency as described above	Missing if TotMeals variable is missing.
Minimum Acceptable Diet	MinDiet	From MinFreq and Four_of6; Values: 1=child meets minimum meal frequency described in MinFreq and the minimum dietary diversity described in Four_of6	Missing if either component variable is missing.
	MinDiet2	From MinFreq3 and Four_of6; Values: 1=child meets minimum meal frequency described in MinFreq3 and the minimum dietary diversity described in Four_of6	Missing if either component variable is missing.
	MinDiet3	From MinFreq3 and Three_of6; Values: 1=child meets minimum meal frequency described in MinFreq3 and the minimum dietary diversity described in Three_of6	Missing if either component variable is missing.
	IronRich	From MeatEggs (EDHS 2000); Organ2, Meat2, Poultry2, Eggs2, and Fish2 (EDHS 2005); and Eggs, Meat, Organs, Fish (EDHS 2011); Vales: 1=child had iron rich foods in last 24 hours 0=else	Missing only if all of component variables are missing.
	Feed_stat	From BFnow, AnySSS, AnyLiq, and EBF; Values: 1=non-breastfeeding 2=breastfeeding and receiving solid, semi-solid, or soft foods 3=breastfeeding and receiving other liquids 4=exclusively breastfeeding	

Section 4: Access to Health Services			
Place of delivery	Del_place	From M15; Values: 1=home 2=public facility 3=private facility (private hospital/NGOs 4=other	
	D_HomeDel	From Del_place; Values: 1=home delivery 0=public or private facility delivery	"Other" converted to missing
	D_HomeDel2	From Del_place; Values: 1=home delivery 0=public, private, or other facility delivery	
Has Health Card	Hcard	From H1; Values: 1=has health card 0=does not have health card	Yes, has health card value 1 includes those reported to have card but not seen by enumerator.
Antenatal Care	PreCare11	From M2A, M2B, M2C, and M2N; Values: 0=mother saw no one for antenatal care 1= mother saw doctor for antenatal care 2=mother saw nurse or midwife for antenatal care 3=mother saw HEW for antenatal care	EDHS 2011 only.
	D_PreCare	From M2A and M2N (EDHS 2000); M2D, M2E, and 2N (EDHS 2005); and M2A, M2B, M2C, and M2N (EDHS 2011); Values: 1=mother saw any health professional, HEW or community health worker for antenatal care 0=else	
	AnteVisits	From M14; Values: 0 through 24 number of antenatal visits; Values: 0 to	"Don't know" converted to missing.
	D_AnteVisits	From M14; Values: 0=0 antenatal visits 1=1 or more antenatal visits	"Don't know" converted to missing.
Measles Immunization	Measles	From H9; Values: 0=no measles vaccination 1=yes, received measles vaccination	"Don't know" converted to missing. Yes includes: vaccination date on card, vaccination marked on card, and reported by mother.
Child recently sick	Diarrhea	From H11; Values: 1=child recently (in last 2 weeks) had diarrhea 0= child did not have diarrhea in the last two weeks	"Don't know" converted to missing.
	Fever	From H22; Values: 1=child recently (in last 2 weeks) had fever 0= child did not have fever in the last two weeks	EDHS 2005 only. Out of 7 total possible food groups.
	Cough	From H31; Values: 1=child recently (in last 2 weeks) had cough 0= child did not have cough in the last two weeks	"Don't know" converted to missing.
	FevCough	From Fever and Cough; Values: 1=child recently (in last 2 weeks) had fever or cough 0= child did not have fever or cough in the last two weeks	

	S ick	From Diarrhea, Fever, and Cough; Values: 1=child recently (in last 2 weeks) had diarrhea, fever, or cough 0= child did not have diarrhea, fever, or cough in the last two weeks	
Minimum Dietary Diversity	Four_of6	From Food_grps_tot1, Food_grps_tot2b, and Food_grps_tot3b; Values: 1=child ate at 4 or more foods out of 6 possible food groups 0=child ate less than 4 foods out of 6 possible food groups	
	Four_of7	From Food_grps_tot2 and Food_grps_tot3; Values: 1=child ate at 4 or more foods out of 7 possible food groups 0=child ate less than 4 foods out of 7 possible food groups	Missing for EDHS 2000.
	Three_of6	From Food_grps_tot1, Food_grps_tot2b, and Food_grps_tot3b; Values: 1=child ate at 3 or more foods out of 6 possible food groups 0=child ate less than 3 foods out of 6 possible food groups	
	Three_of7	From Food_grps_tot2 and Food_grps_tot3; Values: 1=child ate at 3 or more foods out of 7 possible food groups 0=child ate less than 3 foods out of 7 possible food groups	Missing for EDHS 2000.
Diarrhea treatment	DiaTreat	From H12Y and H12Z; Values: 0=no treatment sought for child's diarrhea 1=non-medical treatment sought for child's diarrhea 2=medical treatment sought for child's diarrhea	Missing for all cases with Diarrhea=0.
	D_anyDiaTreat	From H12Y; Values: 0=no treatment sought for diarrhea 1=any treatment (medical or non-medical) sought for diarrhea	Missing for all cases with Diarrhea=0.
	D_MedDiaTreat	From H12Z; Values: 0=no treatment or non-medical treatment sought for diarrhea 1=medical treatment sought for diarrhea	Missing for all cases with Diarrhea=0.
	DiaTreatType	From H12A through H12X; Values: 0=no treatment sought for diarrhea 1=non-medical treatment (shop, drug vendor, pharmacy, traditional practitioner, other) sought for diarrhea 2= government or community facility medical treatment sought for diarrhea 3=private or NGO facility medical treatment sought for diarrhea	Missing for all cases with Diarrhea=0.
Other Variables			
Child id	id	= (1*MIDX)+(10*V003)+(1,000*V002)+(1,000,000*V001)+(1,000,000,000*survey)	This is a unique child identifier by year.

Woman/ mother id	WomanID	=(10*V003)+ (1,000*V002)+ (1,000,000*V001)+ (1,000,000,000*survey)	This is a unique woman identifier by year.
Household id	HouseHoldID	=(1,000*V002)+ (1,000,000*V001)+ (1,000,000,000*survey)	This is a unique household identifier by year.
Cluster id	ClusterID	(1,000,000*V001)+ (1,000,000,000*survey)	This is a unique cluster identifier by year.
Survey year	Survey	From V000; Values: 1=2000 2=2005 3=2011 4=2014	
Weight	Weight	= V005/1000000	This is the recommended weighting variable from EDHS.
Alternate weight	JMweight	=V005/ mean of V005 for each survey	
Mothers height	R_height	From V438 (respondant's height in cm, 1 decimal) Cleaned to the correct decimal place and removed 999's.	Missing for EDHS 2014. n=223 removed to missing (999's).
	R_height2	From R_height; Removed extreme values to range of 140 to 180 cm.	Missing for EDHS 2014. n=127 to missing (out of range).
Lying/ standing	D_lying	From HW15; Values: 1=length, measured lying 0=height, measured standing	"Not measured" recoded as missing.
	mgood	From HW15 and HW1; Values 1=correctly measured for age 0=incorrectly measured for age	Note some incorrect measurements may be correct for actual age, but appear incorrect due to age misreporting. This issue is discussed further in the text.

Variables used in cluster effects analysis	
Household/individual level	Cluster level
Poverty: dummy povcat = 1 if dummies for unimproved water and toilet and roof = 1	Mean by cluster (continuous), dichotomized at mean for each survey (0.48, 0.17), gives dummy (DPovcatS1 and …S3, for surveys 1 and 3).
Water source: surface water (spring, pond, rainwater, etc), dummy (D_SurfaceWat)	Mean by cluster (continuous), dichotomized at mean for each survey (0.81, 0.58), gives dummy (DCl_SurfaceS1 and …S3, for surveys 1 and 3).
Toilet: no facility, dummy (D_NoToilet)	Mean by cluster (continuous), dichotomized at mean for each survey (0.86, 0.38), gives dummy (DCl_NoToiletSS1 and …S3, for surveys 1 and 3).
Roof: grass/thatch, dummy (D_PoorRoof2)	Mean by cluster (continuous), dichotomized at mean for each survey (0.73, 0.55), gives dummy (DCl_PoorRoof2S1 and …S3, for surveys 1 and 3).
Education: no education (V106, respondent), dummy (D_NoEd)	Mean by cluster (continuous), dichotomized at mean for each survey (0.82, 0.58), gives dummy (DCl_NoEdS1 and …S3, for surveys 1 and 3).

Situation Analysis of
The Nutrition Sector in Ethiopia
2000-2015 Methodological
Report