



Quality of anthropometric data in India's National Family Health Survey: Disentangling interviewer and area effect using a cross-classified multilevel model

Laxmi Kant Dwivedi^{a,*}, Kajori Banerjee^b, Radhika Sharma^c, Rakesh Mishra^d, Sowmya Ramesh^e, Damodar Sahu^f, Sanjay K. Mohanty^g, K.S. James^c

^a Department of Survey Research & Data Analytics, International Institute for Population Sciences, Mumbai, India

^b SVKM's Narsee Monjee Institute of Management Studies (NMIMS), Mumbai, India

^c International Institute for Population Sciences, Mumbai, India

^d UNICEF, New Delhi, India

^e Population Council, New Delhi, India

^f National Institute of Medical Statistics, Indian Council of Medical Research, New Delhi, India

^g Department of Population & Development, International Institute for Population Sciences, Mumbai, India

ARTICLE INFO

Keywords:

Anthropometric measures
Children
Data quality
Cross-classified multilevel model
Standard deviation
Team-level variation
NFHS-4
Workload of health investigators

ABSTRACT

India has adopted a target-based approach to reduce the scourge of child malnourishment. Because the monitoring and evaluation required by this approach relies primarily on large-scale data, a data quality assessment is essential. As field teams are the primary mode of data collection in large-scale surveys, this study attempts to understand their contribution to variations in child anthropometric measures. This research can help disentangle the confounding effects of regions/districts and field teams on the quality of child anthropometric data. The anthropometric z-scores of 2,25,002 children below five years were obtained from the fourth round of India's National Family and Health Survey (NFHS-4), 2015–16. Unadjusted and adjusted standard deviations (SD) of the anthropometric measures were estimated to assess the variations in measurements. In addition, a cross-classified multilevel model (CCMM) approach was adopted to estimate the contribution of geographical regions/districts and teams to variations in anthropometric measures. The unadjusted SDs of the measures of stunting, wasting, and underweight were 1.7, 1.4, and 1.2, respectively. The SD of stunting was above the World Health Organisation threshold (0.8–1.2), as well as the Demographic and Health Survey mark. After adjusting for team-level characteristics, the SDs of all three measures reduced marginally, indicating that team-level workload had a marginal but significant role in explaining the variations in anthropometric z-scores. The CCMM showed that the maximum contribution to variations in anthropometric z-scores came from community-level (Primary Sampling Unit (PSU)) characteristics. Team-level characteristics had a higher contribution to variations in anthropometric z-scores than district-level attributes. Variations in measurement were higher for child height than weight. The present study decomposes the effects of district- and team-level factors and highlights the nuances of introducing teams as a level of analysis in multilevel modelling. Population size, density, and terrain variations between PSUs should be considered when allocating field teams in large-scale surveys.

Abbreviations: SD, standard deviation; CCMM, cross-classified multilevel model; PSU, Primary Sampling Unit; NFHS, National Family Health Survey; SDGs, Sustainable Development Goals; POSHAN, Prime Minister's Overarching Scheme for Holistic Nutrition; WHO, World Health Organisation; HAZ, height-for-age z-score; WHZ, weight-for-height z-score; WAZ, weight-for-age z-score.

* Corresponding author.

E-mail addresses: laxmikant@iipsindia.ac.in (L.K. Dwivedi), kajori.b2012@gmail.com (K. Banerjee), radhikash279@gmail.com (R. Sharma), rakeshjnu31@gmail.com (R. Mishra), bbmahapatra@popcouncil.org (S. Ramesh), sahud@icmr.gov.in (D. Sahu), sanjayips@yahoo.co.in, sanjay@iipsindia.ac.in (S.K. Mohanty), ksjames@iipsindia.ac.in (K.S. James).

<https://doi.org/10.1016/j.ssmph.2022.101253>

Received 16 December 2021; Received in revised form 1 September 2022; Accepted 2 October 2022

Available online 6 October 2022

2352-8273/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Reliance on large-scale data to formulate cost-effective policies is growing rapidly in social science and public health research. Thus, it is crucial to evaluate the validity and reliability of the data used in policy decisions. Field supervision and interviewing teams are the primary modes of data collection in large-scale surveys. Several interviewer-level factors, including attitude and motivation, can affect data quality. There are two distinct types of interviewer effects: role-dependent, based on how interviewers ask questions during the survey, and role-independent, based on the individual characteristics of the interviewer, such as their educational qualifications and past work experience (Anglewicz, Adams, Obare, Kohler, & Watkins, 2009). Interviewer-respondent rapport and familiarity can also significantly influence data quality (Adida et al., 2016; Anglewicz, Akilimali, Eitmann, Hernandez, & Kayembe, 2019). In a Demographic and Health Survey (DHS) methodological report, interviewer effects on DHS data were assessed for countries that collected information on demographic and socioeconomic characteristics as well as on the work experience of investigators appointed to collect data. The report highlighted that better-qualified interviewers had a higher probability of collecting better-quality data. The report also pointed out the confounding effect of region and community on interviewer performance (Pullum et al., 2018). There is a wealth of research that points to the systematic influence of respondent characteristics on data quality issues, such as non-response, recall bias, and age misreporting (Channon et al., 2011; Durrant, Groves, Staetsky, & Steele, 2010; Johnson et al., 2009). Lengthy questionnaires were noted as occasionally tiring for respondents, resulting in poor data quality. While insights abound into how both respondent and interviewer effects can distort data quality in large-scale surveys (Amos, 2018), the contribution of interviewing teams to variations in health outcomes in large-scale surveys remains understudied.

The National Family and Health Survey (NFHS) has been at the forefront of using various biomarkers to measure the nutritional status and prevalence of anaemia and a few select morbidities through the collection of anthropometric data and blood samples. Previous studies have evaluated the quality of data provided by four of the last five rounds of NFHS, based on factors such as sample size, respondent characteristics, fertility nutrition and mortality measures (James & Rajan, 2004; Rajan & James, 2004, 2008). A reasonable amount of attention has also been paid to variations in child anthropometric measures that impact the quality of nutritional data (Harkare et al., 2021; Perumal et al., 2020). The NFHS adopts several strategies to ensure data quality, such as real-time field staff monitoring, use of standard measurement tools, and monitoring data quality indicators through field-check tables. Other strategies include training health investigators to efficiently measure the height and weight of eligible respondents, collect blood samples, and measure blood pressure. Studies have shown that the quality of anthropometric data depends on a range of survey-related factors, including inter-interviewer variability (Klipstein-Grobusch et al., 1997). It is essential to highlight that the number of biomarkers used in the NFHS has increased considerably from one round to the next over the last four rounds of the survey (Appendix Table A1). In addition to the increased number of biomarkers, the number of questions and the sample size have also increased, and the structure of interviewing teams has been reformed (Appendix Table A2). However, no study to date has assessed the role of survey workload and team-level characteristics in elucidating the quality of health outcomes, such as anthropometric measures.

In India, child malnutrition remains one of the foremost causes of premature death, and one of the biggest threats to achieving Goal 3 (good health and well-being) of the Sustainable Development Goals (SDGs). Various national and international targets have been set for the effective management of child malnutrition and its target-based systematic reduction. The SDG target of eradicating hunger and all forms of

malnutrition, including stunting, by 2030, and the World Health Organisation (WHO) target of reducing childhood stunting by 40% by 2025 (WHO, 2012) are supported by various national policies by the Indian government. The Prime Minister's Overarching Scheme for Holistic Nutrition (POSHAN), Abhiyaan, launched by the government in 2018, aims to reduce childhood stunting to 25% by 2022. However, regional variations in malnutrition indicators remain a matter of concern, with some states performing significantly poorer than others (India State-Level Disease Burden Initiative CGF Collaborators, 2019). Several studies have used multilevel approaches to assess the contribution of various geographical factors to variations in child malnutrition indicators in India and other developing countries (Amegbor et al., 2020; Jain, Rodgers, Li, Kim, & Subramanian, 2021; Marini & Gragnolati, 2006; Smith & Shively, 2019). However, most of these studies did not consider the confounding effects of team-level performance on regional data quality. Any factor that might jeopardise the data quality of anthropometric measures can have serious implications for estimates of malnutrition. It not only affects the achievement of targets set under national and international programs but also results in the misallocation of resources. Thus, we were motivated to inspect the data quality of the malnutrition indicators in India and a select few states and assess how much of the variation in these indicators is due to team-level versus geographic factors.

The central aim of this study was to assess the contribution of team-level factors to variations in the z-scores of the child anthropometric measures. The study addressed the following questions: (1) Do team-level variables, such as workload, explain any variation in child anthropometric measures? (2) If yes, to what extent do such factors contribute to the variability in anthropometric measures across the regions/districts of India?

We hypothesised that if interviewing teams are trained similarly by similar trainers and have similar supervision procedures during fieldwork, there should be no difference in performance between the different interviewing teams, or in the quality of the biometric data they collected. The results of this study can shed light on the importance of including team-level factors to predict anthropometric measures in India.

2. Materials and methods

2.1. Data source

Data from the fourth round of the NFHS were used in the present analysis. The NFHS-4, conducted in 2015–16, is a large-scale sample survey that provides information on population, health, and nutrition for all states and union territories of India. A two-stage stratified sampling method was used to collect rural and urban samples separately from each of the 640 districts in India. Using the Census-2011 framework, villages were taken as primary sampling units (PSUs) for rural areas, whereas census enumeration blocks (CEBs) were taken as the PSUs for urban areas. Details of the sample size and sampling techniques used are available in the NFHS-4 national report (Ministry of Health and Family Welfare & IIPS, 2017).

2.2. Dependent variables

Three anthropometric measures of child malnourishment were taken as the health outcome variables: height-for-age z-score (HAZ), weight-for-height z-score (WHZ), and weight-for-age z-score (WAZ) among children under five years. The z-scores from the data for these three measures are based on WHO standards (WHO Multicentre Growth Reference Study Group, 2006). Weight was measured using the Seca 874 digital scale. The height of children aged 24–59 months was measured using the Seca 213 stadiometer. The recumbent length of children under 2 years or those less than 85 cm was measured using the Seca 417 infantometer. The sample size of children below 60 months of age

included in this study was 2,59,627. Children with missing information on age, height, and weight were excluded from the statistical analysis. For the regression models, data on 2,25,002 children under 5 years were used from the NFHS-4, 2015–16.

2.3. Independent variables

The independent variables considered in the model were at individual, team, district, and state levels.

- i. **Individual-level characteristics:** Individual-level demographic and socioeconomic variables that are usually associated with predisposing a child to malnutrition were considered in the analysis. These included the child's size at birth, sex, age, birth order, any morbidity reported for the child during the two weeks prior to the survey, age-appropriate immunisation, maternal body mass index (BMI), mother's age at birth, unimproved sanitation in the household, month of interview, maternal education, place of residence, religion, caste, and household wealth status.
- ii. **Team-level characteristics:** Number of eligible children per PSU interviewed by a team
- iii. **District-level characteristics:** Workload of health investigators (multiple-respondent household-to-team ratio)
- iv. **State-level characteristics:** Type/name of agency in charge

To understand whether varying regional workloads affect the data quality of anthropometric measures, two main variables were considered: (i) multiple respondent household-to-health investigator ratio and (ii) total number of children per PSU. The first variable was constructed at the district level by dividing the number of multiple-respondent households by the total number of teams working in the district. Although each team had two health investigators, the health investigator-level codes were not available in the data. We could only identify the teams through field investigators who interviewed the child's mother. Multiple respondent households were those in which more than one eligible child was present and were measured.

The second workload variable was a straightforward count of the total number of eligible children in the PSU.

2.3.1. Ethical considerations

All standard DHS questionnaires, including those in the NFHS-4, were reviewed by the ICF Institutional Review Board (IRB) and comply with 45 code of federal regulations (CFR) 46 governing the 'Protection of Human Subjects'. More on the ethical clearance procedure and the confidentiality and privacy of the survey respondents can be found on the DHS website (<https://dhsprogram.com/methodology/Protecting-the-Privacy-of-DHS-Survey-Respondents.cfm>.)

2.3.2. Analytical strategy

Standard deviation (SD) was used among the quality check indicators for anthropometric measures in the DHS methodological report (Assaf et al., 2015). Hence, an assessment of the anthropometric z-score SDs is included in our initial statistical analysis. This provides an overview of the dispersion in existing child anthropometric data. Next, we estimated the role of team-level and geographical factors in explaining the variation in anthropometric measures of children under five years. Team-, individual-, and household-level factors were adjusted in a hierarchical model to evaluate the stepwise change in the adjusted SD of the anthropometric measures. The second part of the analysis involved constructing a cross-classified multilevel model (CCMM) to estimate the contribution of team-level factors to the total variance in anthropometric z-scores. Analysis was also performed for a few select states/regions with different malnutrition ranks and different levels of team workload to understand the contribution of team-level factors to the data quality. The selected states/regions were Bihar, western Uttar

Pradesh, central Uttar Pradesh, eastern Uttar Pradesh, western Madhya Pradesh, eastern Madhya Pradesh, Andhra Pradesh, Odisha, Punjab, West Bengal, and Tamil Nadu.

2.3.3. Estimating adjusted variance

To estimate the adjusted variance in the anthropometric measures, ordinary least squares (OLS) regression was used by taking the continuous distribution of the z-scores for children and of height/weight for adults. The OLS regression used is as follows:

$$y_i = \alpha_0 + AX' + \varepsilon_i \quad (1)$$

where y_i is the distribution of anthropometric measures, α_0 is the intercept of the model, AX' is the vector of regressors used and their related coefficients, and ε_i is the error term. From the above model, we obtain the predicted values of y_i , denoted as \hat{y}_i .

The sample standard deviation of y_i is estimated as follows:

$$SD(y_i) = \sqrt{\frac{1}{n-1} \sum (y_i - \bar{y})^2} \quad (2)$$

where \bar{y} is the mean of the distribution. Thus, the variation in y_i is the distance between the actual values and the mean.

The root mean square error (RMSE) of the OLS model is estimated as follows:

$$RMSE(y_i) = \sqrt{\frac{1}{n-1} \sum (y_i - \hat{y}_i)^2} \quad (3)$$

where $n-1$ is the degree of freedom of the error term, ε_i . Thus, the RMSE is the distance between the actual and predicted values after adjusting for regressors. We estimated the reduction in variation in the anthropometric distribution by comparing the RMSE of the model to the unadjusted SD of the distribution.

2.3.4. Cross-classified multilevel modelling

Previously published studies have used interclass correlation coefficients to understand the role of interviewers in explaining the variance in a particular outcome variable (Anglewicz et al., 2009). However, a major drawback of this method is that several individual, social, economic, and household variables cannot be controlled in the model. To overcome this issue, we used a CCMM for the anthropometric z-score distributions, as follows:

$$y_{ij(kl)} = \beta_0 + \beta_1 X_{1ij(kl)} + \beta_2 X_{2j} + \beta_3 X_{3l} + \varepsilon_{0j} + \varepsilon_{0k} + \varepsilon_{0l} + \varepsilon_{0ij(kl)} \quad (4)$$

$$\text{where } \varepsilon_{0j} \sim N(0, \sigma_{0j}^2), \varepsilon_{0k} \sim N(0, \sigma_{0k}^2), \varepsilon_{0l} \sim N(0, \sigma_{0l}^2) \text{ and } \varepsilon_{0ij(kl)} \sim N(0, \sigma_{\varepsilon}^2)$$

Here, $y_{ij(kl)}$ is the anthropometric z-score measure of the i^{th} child nested in the j^{th} PSU, measured by the k^{th} team, working in the l^{th} district of a particular state. β_0 represents the intercept of the model; that is, the overall mean outcome y across all PSUs, teams, and districts. $X_{1ij(kl)}$ is the vector of explanatory variables at the individual level and β_1 is the related coefficient for the model. X_{2j} and X_{3l} are the PSU- and district-level covariates, and β_2 and β_3 are their related coefficients, respectively. In the model, the team and district effects were cross-classified because one team worked in more than one district, and one district had more than one team. This implies that teams and districts do not follow a perfect hierarchical structure. We used a CCMM to disentangle the non-hierarchical effects (Dunn et al., 2015). A flowchart illustrating the hierarchical levels is given in Fig. 1. The residual terms at the different levels are provided in Table 1.

The same model was extended to India as a whole, after adding the state at a higher level. The variation partition coefficient at the team level from the model described in equation (4) was estimated to assess

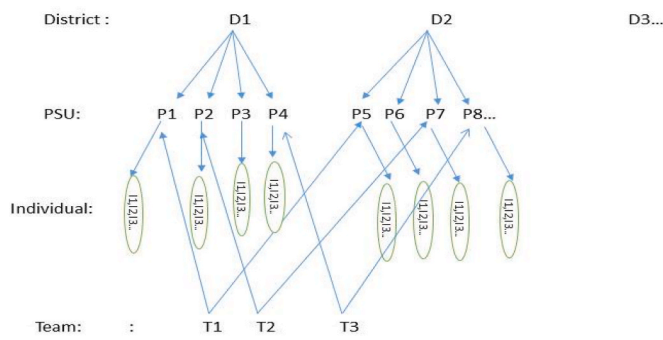


Fig. 1. Hierarchical data structure in NFHS-4, 2015-16.

Table 1 Description of error terms.

Error term	Description
ϵ_{0j}	Random effects at the PSU level (j)
ϵ_{0k}	Random effects at the team level (k)
ϵ_{0l}	Random effects at the district level (l)
$\epsilon_{0ij(kl)}$	Random effects at the individual level over different PSUs j with the combination of k teams and l districts

the contribution of team-level factors to the variation in biometric measures. To justify the use of the CCMM, the conceptual framework demonstrated in Fig. 1 is as follows. In Fig. 1, State 1 has two districts: D1 and D2. Furthermore, three teams worked in State 1. Team T1 collected information from PSU P1 in District D1 and PSU P5 in District D2. Similarly, Team T2 collected information from PSU P2 in District D1 and PSU P7 in District D2. Team T3 collected the information in a similar manner. All three teams worked in districts D1 and D2. This justifies the use of a CCMM instead of a simple multilevel model for analysis.

3. Results

The descriptive statistics of the team workload and anthropometric measures are presented in Table 2. The average number of children under five, whose height and weight were measured in a PSU, was approximately 11 at the national level. Among the selected states/regions, Bihar had the highest average number of eligible children per PSU (16 children per PSU), followed by western Uttar Pradesh, eastern Uttar Pradesh, and western Madhya Pradesh, where the average number of children per PSU ranged between 12 and 14 respectively. The number of households with multiple eligible children interviewed by each team was also high in these locations, ranging from 120 to 159. This implies that in these four states/regions, one team consisting two health investigators interviewed 120 to 159 households, where anthropometric measures were taken for multiple children. In contrast, Andhra Pradesh, Odisha, Punjab, West Bengal, and Tamil Nadu had seven to nine eligible children per PSU, and the multiple child household-to-team ratio in these states ranged from 37 to 56. This indicates that, in these five states, one team of two health investigators interviewed 37 to 56 households, where there was more than one eligible child whose anthropometric measures had to be taken.

Table 2 shows the variation in z-scores of the anthropometric measures. The national average SD of the z-scores was 1.7, HAZ, 1.4 for WHZ, and 1.2 for WAZ. This indicates that there were more deviations in height than in weight among children under the age of five. The results showed no particular SD pattern. For example, states like Tamil Nadu, with a low prevalence of stunting, did not have a lower SD than states like Bihar and Uttar Pradesh, where stunting levels were high.

Fig. 2 present the variations in z-scores by two workload variables: total number of children in a PSU and multiple child household-to-team

ratio in the states. As shown in Fig. 2, the variation measured by the SD of the HAZ scores increased with an increase in the number of eligible children in the PSUs. No such trend was observed for the SDs of the WHZ and WAZ. Fig. 3 presents a scatter plot of z-score SDs and multiple child household-to-team ratios by state/region in India. Regions such as Bihar, eastern and western Uttar Pradesh, and western and eastern Madhya Pradesh had both a high SD of the HAZ and high multiple child household-to-team ratios. In states with a low multiple child household-to-team ratio, such as Tamil Nadu, the SD of the HAZ was close to 1.8.

Fig. 4 shows the adjusted SD of the measures of stunting, wasting, and underweight as obtained from the stepwise OLS regression. The first model demonstrated a decline in the adjusted SD after the z-scores model was controlled for team-level variables, including multiple child household-to-team ratio, total number of children in a PSU, and agency-in-charge in the states. This indicates that the variation in z-scores could be explained by team-level variables. The results obtained from the regression models are presented in Appendix tables A3, A4, and A5. The analysis highlights that the multiple child household-to-team ratio, total number of children in a PSU, and agency-in-charge played a significant role in predicting the z-scores for stunting, wasting, and underweight. Although the beta coefficients were small, the multiple child household-to-team ratio as well as the total number of children per PSU had negative and significant coefficients in Models 1 and 4 for stunting. Model 1 shows that with an increase in the multiple child household-to-team ratio and the total number of children in a PSU, the HAZ scores significantly reduced by 0.01–0.02 units. This reduction remained negative and significant in the full model after controlling for socio-economic, maternal health, demographic, and household factors. Agency-in-charge also played a significant role in explaining the z-scores for the HAZ, WHZ, and WAZ.

Fig. 5 presents the variance estimates for region, district, team, and PSU levels. Three types of multilevel model were constructed. The first model (districts) did not take teams as a level, the second model (team) did not take districts as a level, and the third, which is the CCMM, took both district and team as levels to account for the non-hierarchical nature of the district and team levels. For stunting, wasting, and underweight, the crossed effects of the team ranged from 0.01 to 0.05. We found that districts contributed to a lesser extent to the variance than teams, but ignoring the district level could lead to an overestimation of team-level contribution to total variance. The full model is presented in Appendix Table A6. After controlling for an array of important socio-economic, maternal health, household, and demographic variables that predict z-scores, we found that the multiple child household-to-team ratio, total children in a PSU, and agency-in-charge continued to remain significant determinants of z-scores, especially in the case of children's HAZ.

The results from a similar analysis of the selected states are presented in Table 3. For stunting, team-level factors contributed 0.4–5.4% to the variance estimates. Their contributions ranged from 0.4 to 4.9% for wasting, and from 0.2 to 2.9% for underweight. Our findings highlight that team-level unobserved factors play a larger role in variance estimates than district-level ones. In the case of stunting, team-level factors contributed 2.5–5.4% in the states of Punjab, West Bengal, and Tamil Nadu. After individual-level factors, PSU played a significant role in explaining the unobserved variance in the z-scores of anthropometric measures in all states.

4. Discussion and conclusions

There have been several arguments regarding the standardisation of z-scores of anthropometric measures among children in India (Panagariya, 2013). Research on the accurate determinants of childhood malnutrition is ongoing in India, where malnutrition is listed as a predominant risk factor for premature death in childhood, accounting for almost 68.2% of child deaths under the age of five (India State-Level Disease Burden Initiative CGF Collaborators, 2019). Some studies have

Table 2
Descriptive statistics on team-level workload, anthropometric measures for selected states of India, 2015-16.

Selected states	Agency in charge	No. of districts	No. of PSUs	No. of teams	Total no. of households	Total no. of eligible children	Multiple child households/no. of teams	Average no. of children/PSU	% Stunted	% Wasted	% Underweight	SD for HAZ	SD for WHZ	SD for WAZ
Bihar	Academy of Management Studies (AMS)	38	1,677	48	36,772	25,437	158.3	15.7	48.38	20.87	43.93	1.70	1.27	1.18
Western Uttar Pradesh	Population Research Centre (PRC), Lucknow	26	1,531	34	32,516	18,834	157.1	13.3	44.03	16.66	37.73	1.46	1.17	1.11
Central Uttar Pradesh	Development and Research Services Pvt. Ltd. (DRS)	18	903	32	18,551	7,713	106.6	9.3	47.24	20.8	42.12	1.65	1.40	1.14
Eastern Uttar Pradesh	Goa Institute of Management (GIM), Goa	27	1,204	31	25,166	15,204	142.1	12.8	48.06	17.75	39.94	1.66	1.29	1.15
Western Madhya Pradesh	Academy of Management Studies (AMS)	28	1,421	38	31,211	15,428	119.1	11.9	43.67	25.75	44.33	1.69	1.33	1.16
Eastern Madhya Pradesh	Indian Institute of Health Management Research (IIHMR)	22	989	28	20,831	9,183	88.3	9.4	39.49	25.9	40.7	1.65	1.35	1.15
Andhra Pradesh	GFK Mode	13	494	22	10,265	3,128	36.9	6.9	31.44	17.35	32.19	1.46	1.14	1.10
Odisha	Indian Institute of Health Management Research (IIHMR)	30	1419	38	30,242	11,106	55.5	8.4	34.08	20.39	34.3	1.56	1.36	1.22
Punjab	Society for Promotion of Youth and Masses (SPYM)	20	760	23	16,449	5,216	49.3	7.4	25.75	15.69	21.65	1.47	1.35	1.20
West Bengal	Vimarsh Development Solutions Pvt Ltd	19	722	22	15,327	5,328	44.4	8.5	32.72	20.18	31.56	1.46	1.33	1.17
Tamil Nadu	EHI International	32	1216	43	26,033	7,922	39.6	7.0	27.16	19.66	23.82	1.78	1.54	1.28
India		640	28,524	779	6,01,509	2,59,627	73.87	10.63	38.37	21.04	35.73	1.68	1.39	1.22

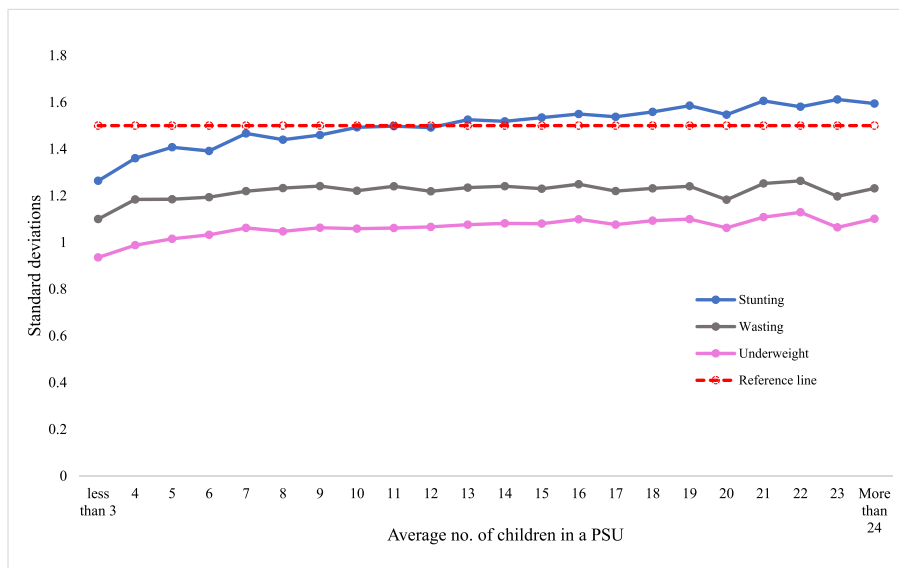


Fig. 2. Variation in child anthropometric measured by total number of children in a PSU in India, 2015-16.

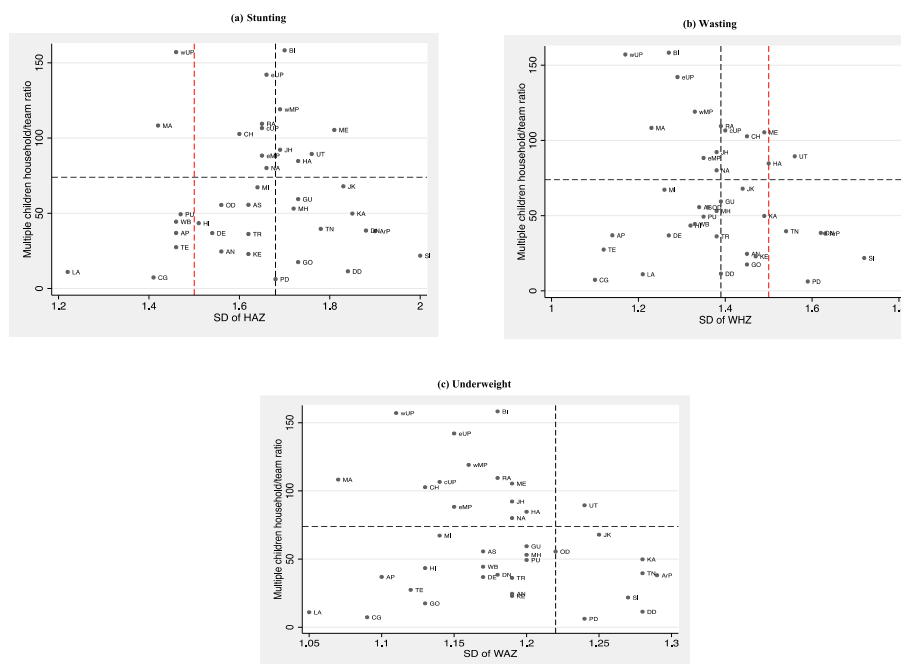


Fig. 3. Scatter plot between multiple child household to team ratio and z-scores of anthropometric measures in the different states/Union territories of India, 2015-16.

advocated district-level policy initiatives to address child malnutrition (India State-Level Disease Burden Initiative CGF Collaborators, 2020). In light of such research, our study brings attention to a vital consideration on data collection and quality that may have been overlooked. To our knowledge, this is the first study to assess the effects of team-level factors on variations in child anthropometric z-scores in India. Our study presents three major findings. First, team-level factors play a significant role in explaining variations in the z-scores of the anthropometric measures. Second, among all three anthropometric measures, the variation in z-scores was the most significant for stunting. The data quality for age measures has improved substantially over time; however, the quality of height measurements for children under five years warrants examination. Third, the contribution of team-level factors to variations in the z-scores of child anthropometric measures was larger than that of

district-level factors. The pattern is similar in states with both high and low malnourishment burden.

Random errors in the measurement of variables always increase the sample SD compared with the true population SD. This issue cannot be controlled by increasing the sample size (Grellety & Golden, 2018). This is evident from our analysis presented in Fig. 2, which shows an increase in the SD of all three malnourishment measures with an increase in the number of eligible children measured in a PSU. The SD of anthropometric measures is one of the simplest data quality indicators, and has been widely used in research studies (Corsi et al., 2017; Grellety & Golden, 2016; Mei & Grummer-Strawn, 2007). In most large-scale surveys, where several enumeration teams are deployed to gather data, systematic bias in measurement by a few teams can increase the SD of that measure (Grellety & Golden, 2018; Grellety & Golden, 2016). In a

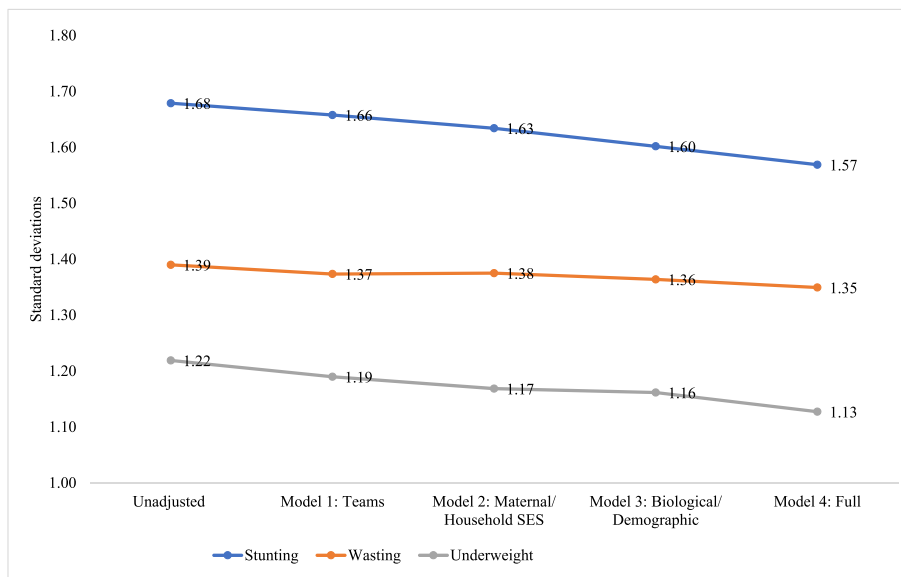


Fig. 4. Variation in child anthropometric measures after step-wise adjustments in India, 2015–16.

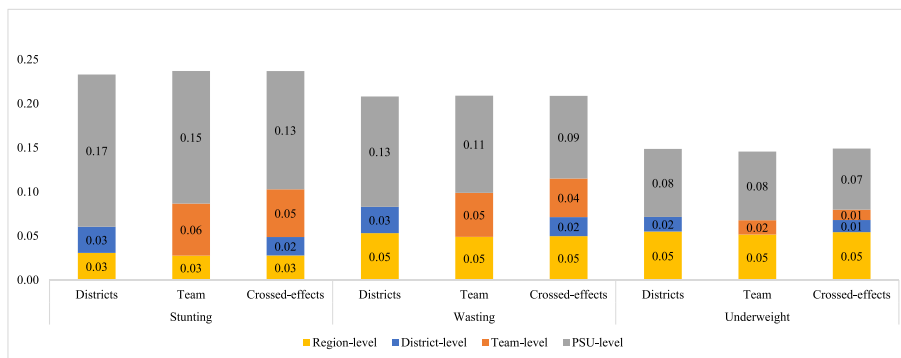


Fig. 5. Variance estimates at regional, district, team and PSU levels from only regions-districts-PSUs levels, from only regions-teams-PSUs levels, and from the cross-classified multilevel models predicting z-scores for child anthropometric measures in India, 2015-16.

Table 3

Percentage of variance explained by district, team and PSU level from cross-classified multilevel model in selected states of India, 2015-16.

Selected states	Stunting			Wasting			Underweight		
	Variance estimates (in %)			Variance estimates (in %)			Variance estimates (in %)		
	District	Team	PSU	District	Team	PSU	District	Team	PSU
Bihar	0.50	1.71	4.13	0.91	1.32	3.57	1.13	0.30	3.56
Western Uttar Pradesh	0.54	0.40	3.63	3.70	0.57	2.41	1.80	0.21	3.86
Central Uttar Pradesh	0.49	1.91	7.79	3.30	3.33	7.65	2.37	0.57	7.13
Eastern Uttar Pradesh	0.92	0.37	2.94	3.14	0.36	2.93	1.87	0.20	3.73
Western Madhya Pradesh	1.22	1.02	6.03	1.03	1.91	4.59	1.76	0.25	5.66
Eastern Madhya Pradesh	0.34	0.75	3.87	1.73	0.37	2.41	0.62	0.31	4.69
Andhra Pradesh	0.91	1.29	4.44	0.29	0.79	1.99	1.16	0.38	4.61
Odisha	1.11	1.39	4.01	1.60	1.15	3.13	3.39	1.00	5.08
Punjab	0.61	2.46	6.53	1.71	3.03	6.55	2.25	3.33	8.18
West Bengal	1.46	2.59	4.58	2.22	0.65	4.14	2.98	0.63	6.58
Tamil Nadu	0.12	5.44	6.50	1.30	4.85	7.33	0.56	2.89	7.70

single-point cross-sectional survey, the SD of an anthropometric measure should lie between 0.8 and 1.2 (World Health Organization, 2012). Our initial results indicate that the unadjusted SD was approximately 1.7 for stunting, 1.4 for wasting, and 1.2 for underweight in India, all of which are already above the recommended thresholds. After controlling for various team-level factors – such as agency in charge of the field survey in a particular state/region, multiple eligible child households in

a PSU, and total number of eligible children in a PSU whose height and weight were measured by the interviewing teams – the SD of the anthropometric measures was found to have decreased marginally. However, the unadjusted and adjusted SDs were highest for stunting. This indicates that although the data quality of anthropometric measures has improved over time in India, steps need to be taken to assess the reasons for data variance and develop tools and techniques to deal

with the challenges of a consistent measurement of child height in large-scale surveys.

Prior research has advocated regional- and district-level policy inputs using multilevel and spatial models to fit child anthropometric z-scores in India (Khan & Das, 2020; Khan & Mohanty, 2018; Li et al., 2020; Liou et al., 2020). A recent study also noted wide inter-state variations in policy effects on anthropometric measures, such as childhood stunting in India (Banerjee & Dwivedi, 2020). However, most of these studies did not consider the confounding effects of team-level performance on data quality or variations in anthropometric measures. In addition to the well-established determinants included in the child malnutrition model, such as dietary intake, socioeconomic factors, household conditions, parental health, and care (Li et al., 2020), this study sheds light on the necessity of assessing the team-level influence on anthropometric measures, especially height. Stunting is among the target indicator of the SDGs. Its inconsistent estimation, both at the national and sub-national levels, can hamper nutrition policies and flagship programs, such as the POSHAN Abhiyaan.

Several studies have demonstrated that district-level factors can explain a significant amount of variation in health outcomes. However, our study illustrates that in a CCMM, team-level factors account for a larger variance in the estimates than district-level factors do. Among the geographical levels within states – districts and PSUs – PSUs play a major role in explaining the intra-state variation in anthropometric measures, followed by team-level factors. An earlier study using data from Wave 2 of the British Household Panel Study claimed that variability in household non-refusal and non-contact is due to the influence of interviewers rather than of areas (O’Muirheartaigh & Campanelli, 1999). The confounding effects of teams and districts cannot be ignored when addressing the issue of stunting in large-scale surveys. Our results underscore the urgent need to control for team-level factors while modelling various public health outcome variables in India, especially those whose quality is objectively dependent on the interviewer/health investigator.

The findings of the study clearly demonstrate the contribution of community-level factors to the variation in anthropometric z-scores; however, they also raise caution about interpreting district-level contributions. Several on-field issues lead to team-level variations in data quality. One of the common methods that jeopardise the data collection process in large-scale surveys is the cumbersome transport of anthropometric measurement tools. This is a practical issue in regions such as western Madhya Pradesh, where the sampled PSUs are widely dispersed, requiring long travel time from one sampled PSU to another and an extended amount of time from one household to another within the PSU. In many cases, road connectivity is challenging due to the terrain. In such situations, it is difficult for field staff to carry along heavy devices. Another on-field issue is the language differences between interviewers/health investigators and respondents, which often leads to the instructions being poorly understood. Our study found that a major issue with anthropometric data lies in child height, especially those below two years whose recumbent length is taken as their height.

The results of this study are limited because of data constraints. There is a lack of data on interviewer characteristics, such as educational level and past work experience in large-scale surveys. Such data could have helped us better understand the causes behind the team-level contribution to variation in anthropometric z-scores. Nevertheless, this study can be extended to examine team-level contributions to variations in other public health outcomes, using data from future rounds of the NFHS survey, which may include interviewer-level characteristics in the raw data files. This can substantially enhance our understanding of the confounding effects of team- and community-level factors on the data quality in various regions in India.

The novelty of this study lies in its contribution to understanding the association between the workload of investigators and increasing the sample size of the NFHS. By controlling for the demographic and socioeconomic attributes of individuals and households, this study was

able to present anomalies in anthropometric measurements due to geographical and team-level characteristics. The use of the CCMM helped capture the confounding effects of team performance in various regions in India. Moreover, this study provides insights into issues with measuring height, especially for younger children (under two years) whose recumbent length is taken as their height. Stunting is one of the most important target indicators of malnourishment. Its inaccurate estimation may lead India to fail in achieving the SDG health and nutrition targets, as well as result in the misallocation of funds. The present study advocates for a thorough examination of the distribution of workload and team-level skills when evaluating anthropometric measures. We strongly recommend using a detailed list of eligible women and children in a PSU to estimate the number of days allotted to complete a survey in each PSU in future rounds of the NFHS, instead of standardising the allotted days to complete a survey in all PSUs, as is current practice. The results obtained from the skill-check tests during the training period can be used to improve on-field monitoring and evaluation intensity. Including data on certain background characteristics of interviewers and supervisors can be an effective addition to future rounds of the NFHS to help evaluate the influence of interviewers on data quality.

Ethics statement

This study is based on secondary data, and is available in public domain for research purpose. Therefore, no ethical approval was required from any institutional review board.

Authorship contributions

Conception and design of study: L. K. Dwivedi, K. Banerjee and K.S. James acquisition of data: K. Banerjee, R. Sharma and L. K. Dwivedi, analysis and/or interpretation of data: K. Banerjee, L. K. Dwivedi, R. Sharma and S. Ramesh, Drafting the manuscript: K. Banerjee and L. K. Dwivedi, revising the manuscript critically for important intellectual content: L. K. Dwivedi, K. Banerjee, R. Sharma, S. Ramesh, R. Mishra, D. Sahu, S. K. Mohanty and K. S. James, Approval of the version of the manuscript to be published: L.K. Dwivedi, K. Banerjee, R. Sharma, S. Ramesh, R. Mishra, D. Sahu, S. K. Mohanty and K. S. James.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Data availability

<https://dhsprogram.com/methodology/survey/survey-display-355.cfm>

Acknowledgement

This paper was written as part of DataQi project of the Population Council funded by Bill & Melinda Gates Foundation (grant # OPP1194597). The views expressed herein are those of the authors and do not necessarily reflect the official policy or position of the Bill & Melinda Gates Foundation and/or Population Council. The authors would like to acknowledge the support from National Data Quality Forum (NDQF) in providing feedback to initial draft of the paper and Ms. Roshni Subramanian for her editorial support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ssmph.2022.101253>.

References

- Adida, C. L., Ferree, K. E., Posner, D. N., & Robinson, A. L. (2016). Who's asking? Interviewer coethnicity effects in African survey data. *Comparative Political Studies*, 49(12), 1630–1660. <https://doi.org/10.1177/0010414016633487>
- Amegbor, P. M., Zhang, Z., Dalgaard, R., & Sabel, C. E. (2020). Multilevel and spatial analyses of childhood malnutrition in Uganda: Examining individual and contextual factors. *Nov 18 Scientific Reports*, 10(1), 20019. <https://doi.org/10.1038/s41598-020-76856-y>.
- Amos, M. (2018). Interviewer effects on patterns of nonresponse: Evaluating the impact on the reasons for contraceptive nonuse in the Indonesia and the Philippines. *Demographic Research*, 39(1), 415–430. <https://doi.org/10.4054/DemRes.2018.39.14>
- Anglewicz, P., Adams, J., Obare, F., Kohler, H.-P., & Watkins, S. (2009). The Malawi Diffusion and Ideational Change Project 2004–06: Data collection, data quality, and analysis of attrition. *Demographic Research*, 20(21), 503. <https://doi.org/10.4054/DemRes.2009.20.21>
- Anglewicz, P., Akilimali, P., Eitmann, L. P., Hernandez, J., & Kayembe, P. (2019). The relationship between interviewer- respondent familiarity and family planning outcomes in the democratic republic of Congo: A repeat cross-sectional analysis. *BMJ Open*, 9, 1–9. <https://doi.org/10.1136/bmjopen-2018-023069>
- Assaf, S., Kothari, M. T., & Pullum, T. (2015). *An assessment of the quality of DHS anthropometric data, 2005–2014*. DHS Methodological Reports No. 16 <https://www.dhsprogram.com/pubs/pdf/MR16/MR16.pdf>.
- Banerjee, K., & Dwivedi, L. K. (2020). Disparity in childhood stunting in India: Relative importance of community-level nutrition and sanitary practices. *PLoS One*, 15(9), Article e0238364. <https://doi.org/10.1371/journal.pone.0238364>
- Channon, A. A. R., Padmadas, S. S., & McDonald, J. W. (2011). Measuring birth weight in developing countries- does the Method of reporting in retrospective surveys matter. *Maternal and Child Health Journal*, 15, 12–18. <https://doi.org/10.1007/s10995-009-0553-3>
- Corsi, D. J., Perkins, J. M., & Subramanian, S. V. (2017). Child anthropometry data quality from Demographic and health surveys, Multiple Indicator cluster surveys, and national nutrition surveys in the west central Africa region- are we comparing apples and oranges? *Global Health Action*, 11(1). <https://doi.org/10.1080/16549716.2017.1328185>
- Dunn, E. C., Richmond, T. K., Milliren, C. E., & Subramanian, S. V. (2015). Using cross-classified multilevel models to disentangle school and neighborhood effects: An example focusing on smoking behaviors among adolescents in the United States. *Health & Place*, 31, 224–232. <https://doi.org/10.1016/j.healthplace.2014.12.001>
- Durrant, G. B., Groves, R. M., Staetsky, L., & Steele, F. (2010). Effects of interviewer attitudes and behaviors on refusal in household surveys. *Public Opinion Quarterly*, 74(1), 1–36. <https://doi.org/10.1093/poq/nfp098>
- Grellety, E., & Golden, M. H. (2016). The effect of random error on diagnostic accuracy illustrated with the anthropometric diagnosis of malnutrition. *PLoS One*, 11(12), Article e0168585. <https://doi.org/10.1371/journal.pone.0168585>
- Grellety, E., & Golden, M. H. (2018). Change in quality of malnutrition surveys between 1986 and 2015. *Emerging Themes in Epidemiology*, 15(8). <https://doi.org/10.1186/s12982-018-0075-9>
- Harkare, H. V., Corsi, D. J., Kim, R., Vollmer, S., & Subramanian, S. V. (2021). The impact of improved data quality on the prevalence estimates of anthropometric measures using DHS datasets in India. *Scientific Reports*, 11. <https://doi.org/10.1038/s41598-021-89319-9>
- India State-Level Disease Burden Initiative CGF Collaborators. (2019). The burden of child and maternal malnutrition and trends in its indicators in the states of India: The global burden of Disease study 1990–2017. *The Lancet Child and Adolescent Health*, 3(12), 855–870. [https://doi.org/10.1016/S2352-4642\(19\)30273-1](https://doi.org/10.1016/S2352-4642(19)30273-1)
- India State-Level Disease Burden Initiative CGF Collaborators. (2020). Mapping of variations in child stunting, wasting and underweight within the states of India: The global burden of Disease study 2000–2017. *EClinicalMedicine*, 22. <https://doi.org/10.1016/j.eclinm.2020.100317>
- Jain, A., Rodgers, J., Li, Z., Kim, R., & Subramanian, S. V. (2021). Multilevel analysis of geographic variation among correlates of child undernutrition in India. *Maternal and Child Nutrition*, 1–11. <https://doi.org/10.1111/mcn.13197>
- James, K. S., & Rajan, S. I. (2004). Respondents and quality of survey data. *Economic and Political Weekly*, 659–663. <https://www.jstor.org/stable/4414640>.
- Johnson, K., Grant, M., Khan, S., Moore, Z., Armstrong, A., & Sa, Z. (2009). *Fieldwork-related factors and data quality in the demographic health surveys program*. DHS Analytical Studies.
- Khan, J., & Das, S. K. (2020). The burden of anthropometric failure and child mortality in India. *Scientific Reports*, 10(1), Article 20991. <https://doi.org/10.1038/s41598-020-76884-8>
- Khan, J., & Mohanty, S. K. (2018). Spatial heterogeneity and correlates of child malnutrition in districts of India. *BMC Public Health*, 18(1), 1027. <https://doi.org/10.1186/s12889-018-5873-z>
- Klipstein-Grobusch, K., Georg, T., & Boeing, H. (1997). Interviewer variability in anthropometric measurements and estimates of body composition. *International Journal of Epidemiology*, 26(Suppl 1), S174–S180. https://doi.org/10.1093/ije/26.suppl_1.s174
- Li, Z., Kim, R., Vollmer, S., & Subramanian, S. V. (2020). Factors associated with child stunting, wasting, and underweight in 35 low- and middle-income countries. *JAMA Network Open*, 3(4), Article e203386. <https://doi.org/10.1001/jamanetworkopen.2020.3386>
- Liou, L., Kim, R., & Subramanian, S. V. (2020). Identifying geospatial patterns in wealth disparity in child malnutrition across 640 districts in India. *SSM Popul Health*, 10, Article 100524. <https://doi.org/10.1016/j.ssmph.2019.100524>
- Marini, A., & Gragnolati, M. (2006). *Nonlinear effects of altitude on child growth in Peru: A multilevel analysis (working paper no. 3823)*. World Bank Policy Research. <https://doi.org/10.1596/1813-9450-3823>
- Mei, Z., & Grummer-Strawn, L. M. (2007). *Standard deviation of anthropometric Z-scores as a data quality assessment tool using the 2006 WHO growth standards: A cross country analysis*. Bulletin of the World Health Organization. <https://doi.org/10.2471/BLT.06.034421>
- Ministry of Health and Family Welfare & IIPS. (2017). *National family health survey (NFHS-4) (2015-16)*. <http://rchiips.org/nfhs/NFHS-4Reports/India.pdf>.
- O'Muircheartaigh, C., & Campanelli, P. (1999). A multilevel exploration of the role of interviewers in survey non-response. *Journal of the Royal Statistical Society*, 162(3), 437–446. <https://doi.org/10.1111/1467-985X.00147>
- Panagariya, A. (2013). Does India really suffer from worse child malnutrition than Sub-Saharan Africa. *Economic and Political Weekly*, 48(18), 98–111. <https://www.jstor.org/stable/23527314>.
- Perumal, N., Namaste, S., Qamar, H., Aimone, A., Bassani, D. G., & Roth, D. E. (2020). Anthropometric data quality assessment in multisurvey studies of child growth. *American Journal of Clinical Nutrition*, 112(Suppl 2), 806S–815S. <https://doi.org/10.1093/ajcn/nqaa162>
- Pullum, T. W., Juan, C., Khan, N., & Staveteig, S. (2018). *The effect of interviewer characteristics on data quality in DHS surveys*. DHS Methodological Reports No. 24 <https://dhsprogram.com/pubs/pdf/MR24/MR24.pdf>.
- Rajan, S. I., & James, K. S. (2004). Second national family health survey: Emerging issues. *Economic and Political Weekly*, 39(7), 647–651. <https://www.jstor.org/stable/4414638>.
- Rajan, S. I., & James, K. S. (2008). Third national family health survey in India: Issues, problems and prospects. *Economic and Political Weekly*, 43(48), 33–38. <http://www.jstor.org/stable/40278234>.
- Smith, T., & Shively, G. (2019). Multilevel analysis of individual, household, and community factors influencing child growth in Nepal. *BMC Pediatrics*, 19(1), 91. <https://doi.org/10.1186/s12887-019-1469-8>
- WHO Multicentre Growth Reference Study Group. (2006). *WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development*.
- World Health Organization. (2012). *Physical status: The use and interpretation of anthropometry (WHO technical report series No. 854)*. Sixty-fifth World health assembly. <https://apps.who.int/iris/handle/10665/37003>.