Short Communication

Ex post adjustment for measurement error in stunting calculations: an illustration from Egypt

Jose Luis Figueroa[1] and Sikandra Kurdi[2],* iD

[1]Center for Health Systems Research, Instituto Nacional de Salud Pública, Cuernavaca, Mexico: [2]International Food Policy Research Institute, 1201 Eye Street NW, Washington, DC 20005, USA

Abstract

Objective: The present study provides ranges for the magnitude of bias caused by measurement error in stunting rates, a widely used a proxy for long-term nutritional status.

Design: Stunting, which is determined by the number of cases that fall below −2 SD from the mean height-for-age in the population, mechanically increases with higher variance. This variance stems from both natural heterogeneity in the population and measurement error. To isolate the effect of measurement error, we model the true distributions which could give rise to the observed distributions after subtracting a simulated measurement error.

Setting: We analyse information from three rounds of the Demographic and Health Survey (DHS) in Egypt (2005, 2008 and 2014). Egypt ranks high among developing countries with low-quality anthropometric data collected in the DHS, currently the main source of anthropometry in the country.

Participants: The study relies on re-analysis of existing DHS data, which record height, weight and age data for children under 5 years old.

Results: Under the most conservative assumptions about measurement error, the stunting rate falls by 4 percentage points for the most recent DHS round, while assuming higher levels of measurement error reduces the stunting rate more dramatically.

Conclusions: Researchers should be aware of and adjust for data quality concerns in calculating stunting rates for cross-survey comparisons or in communicating to policy makers.

The child stunting rate, defined as the share of children with a height-for-age Z-score (HAZ) of less than −2, is extensively used a proxy for average nutritional status of children in a community[1]. While having HAZ < −2 is not highly predictive of health status at the individual level, at the population level as it can serve as an indicator of the overall incidence of linear growth faltering[2]. Economists and policy makers often use the stunting rate to make cross-country comparisons or to evaluate how average child nutritional status has changed over time[3–5]. However, focusing only on differences in the stunting rate, rather than the distribution of HAZ, can be misleading. Increasing the variance of Z-scores in a population mechanically increases the stunting rate. While some

variation in height is expected due to individual genetic or environmental factors[6], the population variance may be artificially inflated due to measurement error, potentially leading to an overestimation of the actual stunting rate. In the case of Egypt, we show evidence that the measurement error in the Demographic and Health Surveys (DHS) is both high and differs from one survey round to another. Therefore, adjustment for measurement error is necessary to make statements about the evolution of stunting rates over time and how Egypt compares with other countries.

Age misreporting and inaccurate height recording are common in anthropometric surveys and lead to errors in HAZ. For example, it has been shown that age misreporting

is the major source of error in calculating HAZ values in national representative surveys in developing countries in sub-Saharan Africa and Asia[7,8] and that an error of 1·5 months for a 12-month-old child can increase the SD of HAZ by about 20 %[9]. Height measurement is also prone to measurement errors by enumerators, especially for younger children[10].

In an analysis of DHS anthropometric data quality in fifty-two countries, Egypt is one of the seven countries listed as having the most concerns raised across a variety of different quality indicators[11]. Nevertheless, the DHS stunting rate for Egypt is widely used and cited by researchers, policy makers and development organizations[12–14].

In the present communication, we show the potential ranges of stunting after removing measurement error under different assumptions about the magnitude of error for three rounds of the Egypt DHS (2005, 2008 and 2014)[15–17].

Methods

We present three scenarios which show the range of true values which could give rise to the observed data depending on how much measurement error we assume: (i) a conservative scenario; (ii) an aggressive scenario; and (iii) a moderate scenario between these two extremes. In all scenarios, we assume classical measurement error uncorrelated with the measurements and with a mean of 0. The SD of the measurement error differs in each scenario.

Conservative scenario

This approach is based on looking at the high side of the distribution. In line with the WHO recommendations from the Multicenter Growth Reference Study, we assume that the share of overly tall children should be no greater than that found in the WHO reference populations used to construct the international growth standards, so any excess must be due to measurement error.* Let H represent the proportion of children with a Z-score more than 2 SD above the mean:

$$H = \int_{1\cdot96}^{\infty} \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}}\, e^{-1/2(x/\tilde{\sigma})^2}\, dx.$$

We solve (computationally) for the SD of a normal distribution that would give this share of extremely tall children, denoted as $\tilde{\sigma}$. The difference between $\tilde{\sigma}$ and the reference population

SD of 1 is attributable to measurement error. We denote the variance of this measurement error as σ_ε^2:

$$\sigma_\varepsilon^2 = \tilde{\sigma}^2 - 1.$$

While we are confident that our conservative approach identifies a minimum amount of variance that must be due to measurement error, accounting for this measurement error still leaves the observed variance much higher than expected. So, we proceed also to show how much the stunting rate would change under more aggressive assumptions.

Aggressive scenario

As a benchmark of the maximum measurement error that could be assumed, we make the strong assumption that all variance beyond that in the reference population (where the variance is equal to 1) is attributed to measurement error. In effect, this assumes that there is no heterogeneity of malnutrition due to regional differences or other household characteristics and that any impact of malnutrition on growth is diffused equally throughout the population. This assumption of no heterogeneity is unrealistic, but by making the strongest possible assumption, we obtain the minimum true stunting rate that could be inferred from the observed data.

In this case, we define measurement error based on the difference between the variance of the observed Z-scores ($\tilde{\sigma}^2$) and 1:

$$\sigma_\varepsilon^2 = \tilde{\sigma}^2 - 1.$$

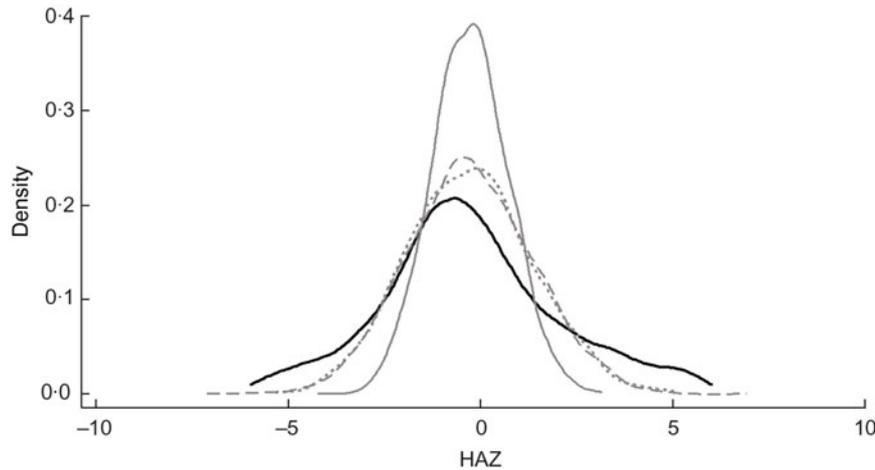### Moderate scenario: true variance inferred from a better-quality survey

Instead of the extreme assumption that the true values are distributed with a variance of 1, we can look to better-quality survey data for some guesses about the actual value of the true variance. Here, we use the variance found in a 2018 survey of households in rural areas in Upper Egypt† where we know that the enumerator training was done more carefully than for the DHS data collection. We define measurement error based on the difference between the variance of the observed Z-scores ($\tilde{\sigma}^2$) in the DHS and the variance in our higher-quality survey ($\tilde{\sigma}_S^2$):

$$\sigma_\varepsilon^2 = \tilde{\sigma}^2 - \tilde{\sigma}_S^2.$$

Simulations

We turn to simulations to model what true distributions would give rise to the observed distribution after the

---

*Recent research has found that there is more national-level diversity in growth curves for healthy children than the WHO standards imply; however, for most countries, the international growth curves are a reasonable fit[18]. In countries where it is available, using a national reference population to calculate Z-scores would eliminate concern about the comparability with the WHO reference population. We have established that for anthropometric survey data from Egypt where enumerators had better training than in the DHS (see next footnote), there is no evidence that the share of overly tall children differs substantially from the 2·5 % expected in the WHO reference population.

†The International Food Policy Institute, with funding from the US Agency for International Development, collected anthropometric information on 1255 children aged 6–59 months in May 2018. These children came from small villages in rural areas in five governorates in Egypt: Luxor, Sohag, Qena, Beni Suef and Menya.

Fig. 1 Height-for-age Z-score (HAZ) distributions based on different scenarios (——, observed distribution of Z-scores; ...., simulated under conservative scenario; – – – –, simulated under moderate scenario; ————, simulated under aggressive scenario) for children aged 6–23 months from three rounds of the Demographic and Health Survey (DHS) in Egypt (2005, 2008 and 2014)

addition of measurement error. Because we assume that errors are uncorrelated with the true values, the expected variance of this true distribution, denoted as $\hat{\sigma}^2$, is equal to the observed variance $(\tilde{\sigma}^2)$ minus the variance of the hypothesized measurement error:

$$\hat{\sigma}^2 = \tilde{\sigma}^2 - \sigma_\varepsilon^2.$$

We then define a random variable representing simulated measurement error, $\hat{\varepsilon}$, such that $\hat{\varepsilon}$ is distributed $N(0, \sigma_\varepsilon^2)$ and subtracting $\hat{\varepsilon}$ from our observed data reduces the variance to $\hat{\sigma}^2$. Because the observed measurements $(\tilde{z})$ are the sum of the true values and the measurement error, $\hat{\varepsilon}$ is positively correlated with the true value (z). The correlation (ρ) between $\hat{\varepsilon}$ and $\tilde{z}$ can be defined based on σ_ε^2 and $\tilde{\sigma}^2$:

$$\hat{\sigma}^2 = \tilde{\sigma}^2 - \sigma_\varepsilon^2 = \tilde{\sigma}^2 + \sigma_\varepsilon^2 - 2\tilde{\sigma}\sigma_\varepsilon\rho,$$

where

$$\rho = \frac{\sigma_\varepsilon}{\tilde{\sigma}}.$$

In each iteration, we randomly select values of $\hat{\varepsilon}$ correlated by ρ with the observed distribution; subtract $\hat{\varepsilon}$ from $\tilde{z}$ to generate our simulated measurement-error-free data; and calculate the resulting stunting rate. For all summary statistics, we follow the standard practice of excluding flagged values (HAZ > 6 or HAZ < −6). We present the results of 1000 iterations.

We choose values for σ_ε^2 and perform the simulations separately for children aged 6–23 months and 24–59 months because we expect the measurement error to differ between these two groups.

Results

Figures 1 and 2 illustrate of the effect of removing measurement error under the three scenarios described above. These figures show a single iteration in our simulations: the distribution of corrected Z-scores resulting from one randomly drawn vector of measurement error. Removing the error leads to tighter distributions with fewer cases of stunting.
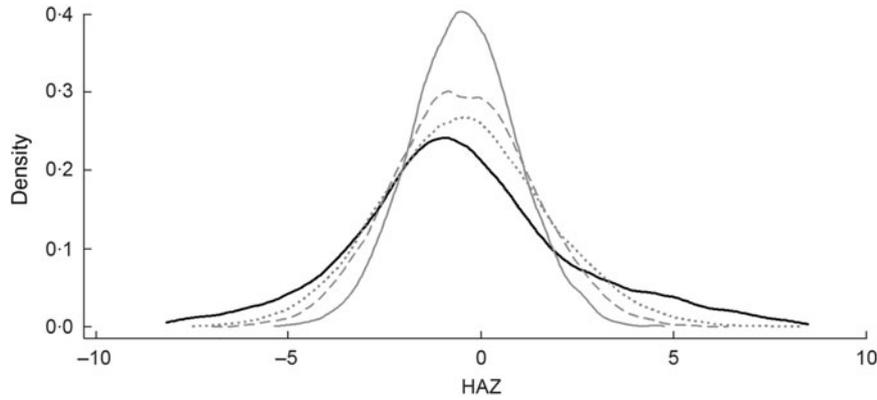
Table 1 summarizes the results of the three simulation scenarios in each of three rounds of DHS data from Egypt.

In panel A of Table 1, we summarize the DHS data on HAZ, including means, SD and the prevalence of stunting in the population of children as calculated using the Stata module zscore06, which is based on the 2006 WHO reference values[19]. The SD of the Z-scores are extremely high, ranging from a minimum of 1·71 in 2005 up to a maximum of 2·32 in 2014. As mentioned above, there is also a striking increase in stunting especially among the younger age group from 2005 to 2008 followed by a decline to 2014.

For each scenario, the corresponding panel in Table 1 shows the measurement error (σ_ε^2) that we are assuming, the empirical 95 % CI of the modelled SD of Z-scores after removing measurement error $(\hat{\sigma})$ and the empirical 95 % CI of the resulting stunting rates.

Under the conservative scenario (panel B in Table 1), the assumed measurement error accounts for between 41 and 46 % of the variance in the observed data. Removing this error leads to modelled stunting rates in the various DHS rounds from 1 to 5 percentage points lower than in the observed data.

Under the aggressive scenario (panel C in Table 1), the assumed measurement error comprises 58–68 % of the variance in the observed data. Removing this measurement

Fig. 2 Height-for-age Z-score (HAZ) distributions based on different scenarios (———, observed distribution of Z-scores; ⋯⋯⋯, simulated under conservative scenario; − − − −, simulated under moderate scenario; ————, simulated under aggressive scenario) for children aged 24–59 months from three rounds of the Demographic and Health Survey (DHS) in Egypt (2005, 2008 and 2014)

Table 1 Stunting prevalence for conservative, aggressive and moderate scenarios among children under 5 years old from three rounds of the Demographic and Health Survey (DHS) in Egypt (2005, 2008 and 2014)

Panel A. DHS data

	2005			2008			2014		
	Mean	SD ($\tilde{\sigma}$)	Stunted (%)	Mean	SD ($\tilde{\sigma}$)	Stunted (%)	Mean	SD ($\tilde{\sigma}$)	Stunted (%)
6–23 months	−1·04	2·22	31·0	−0·91	2·24	30·2	−0·28	2·32	20·4
24–59 months	−1·09	1·71	25·6	−1·25	1·86	32·0	−0·46	2·02	19·8

Panel B. Conservative scenario: CI based on data inferred from the share of overly tall children found in DHS*

	2005			2008			2014		
	σ_ε	$\hat{\sigma}$	Stunted (%)	σ_ε	$\hat{\sigma}$	Stunted (%)	σ_ε	$\hat{\sigma}$	Stunted (%)
6–23 months	1·30	1·77, 1·82	28·5, 30·7	1·38	1·72, 1·79	25·6, 28·0	1·51	1·73, 1·79	15·2, 17·0
24–59 months	0·98	1·38, 1·42	23·5, 24·9	1·16	1·43, 1·47	28·9, 30·8	1·32	1·51, 1·55	14·5, 15·7

Panel C. Aggressive scenario: CI assuming all variance is due to measurement error*

	2005			2008			2014		
	σ_ε	$\hat{\sigma}$	Stunted (%)	σ_ε	$\hat{\sigma}$	Stunted (%)	σ_ε	$\hat{\sigma}$	Stunted (%)
6–23 months	1·98	0·98, 1·02	15·7, 18·0	2·01	0·97, 1·02	12·5, 14·8	2·10	0·98, 1·02	3·7, 4·7
24–59 months	1·39	0·98, 1·01	16·9, 18·5	1·57	0·98, 1·02	21·6, 23·6	1·76	0·98, 1·01	5·6, 6·5

Panel D. Moderate scenario: CI assuming variance over other survey is due to measurement error*

	2005			2008			2014		
	σ_ε	$\hat{\sigma}$	Stunted (%)	σ_ε	$\hat{\sigma}$	Stunted (%)	σ_ε	$\hat{\sigma}$	Stunted (%)
6–23 months	1·49	1·61, 1·67	26·8, 29·1	1·53	1·61, 1·68	24·1, 26·6	1·64	1·61, 1·67	13·6, 15·3
24–59 months	1·11	1·29, 1·32	22·3, 23·9	1·32	1·28, 1·32	27·0, 29·0	1·55	1·29, 1·32	11·0, 12·1

$\tilde{\sigma}$, observed SD of Z-scores; σ_ε^2, hypothesized measurement error; $\hat{\sigma}$, modelled SD of Z-scores after removing measurement error.
*95 % CI based on 1000 simulations.

error results in a significant decrease in the estimated prevalence of stunting for all groups, particularly among children in the 2008 and 2005 samples. For example, in 2008, children aged 6–23 months would have a stunting prevalence between 12·5 and 14·8 % in contrast with the 30·2 % observed in the DHS data.

Under the moderate scenario (panel D in Table 1), the measurement error comprises 45–54 % of the variance in the observed data. The corrected stunting rates fall between the conservative and aggressive results.

For 2014, under our moderate scenario, our simulations give a mean true stunting prevalence of 14·7 % for children

aged 6–23 months and 15·3 % for children aged 24–59 months. This is dramatically lower than the DHS rate currently accepted to be true in Egypt (20·4 and 19·8 %, respectively). The results from the alternative conservative and aggressive assumptions give a range from 3·7 to 17·0 % for children aged 6–23 months, and from 5·6 to 15·7 % for older children. This shows that at the least stunting is overestimated by about 4 percentage points, that the true stunting rate could be substantially lower than even our moderate scenario suggests, and that there is a very large degree of uncertainty about the true stunting rate.

Discussion

The present study shows how measurement error leads to overestimation of the real malnutrition prevalence in Egypt, a country where issues with anthropometric data quality have been observed but are not widely appreciated. We show that measurement error has been high and variable across the three most recent rounds of the DHS and simulate a range of potential true values of the stunting rate.

Policy makers were particularly concerned by the high prevalence of stunting observed in 2008 which contrasts with the other two rounds. Previous research had suggested that this effect might be the consequence of the outbreak of avian flu that took place in Egypt in 2006, and that affected mainly rural areas in Lower Egypt, with significant consequences for household welfare and protein consumption for children[20,21]. Our simulations confirm that even accounting for measurement error, the increase in stunting in 2008 is still observed, although the level and magnitude of the change are lower.

There are important implications of these findings for policy makers in Egypt; but for the larger community, we emphasize that data quality issues in the DHS were flagged in a number of other countries. While one response is to call for improving accuracy of measurement for future rounds of collection, many researchers and policy makers are faced with the reality that the only data that exist are from surveys with data quality concerns.

We recommend that researchers in this situation use the types of adjustments described in the present communication before using DHS data from survey rounds with questionable data quality in time-series or cross-sectional analysis of stunting rates. We also recommend using this approach to communicate to policy makers about the degree of uncertainty surrounding the stunting rate when the only available data are from survey rounds with data quality concerns.

Acknowledgements

References

1. Svedberg P (2000) *Poverty and Undernutrition: Theory, Measurement, and Policy.* Oxford: Oxford University Press.
2. Perumal N, Bassani DG & Roth DE (2018) Use and misuse of stunting as a measure of child health. *J Nutr* **148**, 311–315.
3. de Onis M, Blössner M & Borghi E (2012) Prevalence and trends of stunting among pre-school children, 1990–2020. *Public Health Nutr* **15**, 142–148.
4. Garrett JL & Ruel MT (2005) Stunted child–overweight mother pairs: prevalence and association with economic development and urbanization. *Food Nutr Bull* **26**, 209–221.
5. Mukuria A, Cushing J & Sangha J (2005) *Nutritional Status of Children: Results from the Demographic Health Surveys 1994–2001. DHS Comparative Reports* no. 10. Calverton, MD: ORC Macro.
6. Jelenkoic A, Sund R, Hur Y-M *et al.* (2016) Genetic and environmental influences on height from infancy to early adulthood: an individual-based pooled analysis of 45 twin cohorts. *Sci Rep* **6**, 28496.
7. Comandini O, Cabras S & Marini E (2009) Nutritional evaluation of undocumented children: a neglected health issue affecting the most fragile people. *Eur J Public Health* **32**, 1189–1207.
8. Comandini O, Cabras S & Marini E (2016) Birth registration and child undernutrition in sub-Saharan Africa. *Public Health Nutr* **19**, 1757–1767.
9. Grellety E & Golden MH (2016) The effect of random error on diagnostic accuracy illustrated with the anthropometric diagnosis of malnutrition. *PLoS One* **11**, e0168585.
10. ICF International (2013) *MEAURE Biomarker Field Manual Demographic and Health Surveys Methodology.* Calverton, MD: ICF International.
11. Assaf S, Kothari M & Pullum T (2005) *An Assessment of the Quality of DHS Anthropometric Data, 2005–2014. DHS Methodological Reports* no. 16. Rockville, MD: ICF International.
12. Ecker O, Al-Riffai P, Breisinger C *et al.* (2016) *Nutrition and Economic Development: Exploring Egypt's Exceptionalism and the Role of Food Subsidies.* Washington, DC: International Food Policy Research Institute.
13. Kavle JA, El Zanaty F, Landry M *et al.* (2015) The rise in stunting in relation to avian influenza and food consumption patterns in Lower Egypt in comparison to Upper Egypt:

results from 2005 and 2008 Demographic and Health Surveys. *BMC Public Health* **15**, 285.

14. Rashad AS & Sharaf MF (2018) Regional inequalities in child malnutrition in Egypt, Jordan, and Yemen: a Blinder–Oaxaca decomposition analysis. *Soc Indic Res* **135**, 769–795.

15. El-Zanaty F & Way A (2006) *Egypt Demographic and Health Survey 2005*. Cairo and Rockville, MD: Ministry of Health and Population, National Population Council, El-Zanaty and Associates, and ORC Macro.

16. El-Zanaty F & Way A (2009) *Egypt Demographic and Health Survey 2008*. Cairo and Rockville, MD: Ministry of Health and Population, National Population Council, El-Zanaty and Associates, and Macro International.

17. Ministry of Health and Population, El-Zanaty and Associates, & ICF International (2015) *Egypt Demographic and Health Survey 2014*. Cairo and Rockville, MD: Ministry of Health and Population and ICF International.

18. Natale V & Rajagopalan A (2014) Worldwide variation in human growth and the World Health Organization growth standards: a systematic review. *BMJ Open* **4**, e003735.

19. Leroy J (2011) *ZSCORE06: Stata Module to Calculate Anthropometric Z-Scores Using the 2006 WHO Child Growth Standards*. Statistical Software Components S457279. Boston, MA: Boston College Department of Economics.

20. Hosny FA (2006) *The Structure and Importance of the Commercial and Village Based Poultry in Egypt*. Rome: FAO.

21. Geerlings E, Albrechtsen L & Rushton J (2007) *Highly Pathogenic Avian Influenza: A Rapid Assessment of the Socio-Economic Impact on Vulnerable Households in Egypt*. Rome: FAO.

Public Health Nutrition