USAID
FROM THE AMERICAN PEOPLE

ENHANCING NUTRITION DATA QUALITY IN THE DHS PROGRAM

DHS Qualitative Research Studies 23

November 2018

This publication was produced for review by the United States Agency for International Development. It was prepared by Sorrel Namaste, Rukundo K. Benedict, and Megan Henry.

DHS Qualitative Research Studies No. 23

Enhancing Nutrition Data Quality in The DHS Program

Sorrel Namaste[1]
Rukundo K. Benedict[1]
Megan Henry[2]

ICF
Rockville, Maryland

November 2018

[1] The DHS Program
[2] Johns Hopkins Bloomberg School of Public Health

Corresponding author: Sorrel Namaste, International Health and Development, ICF, 530 Gaither Road, Suite 500, Rockville, MD 20850, USA; telephone: +1 301-572-0466; email: Sorrel.Namaste@icf.com

Acknowledgments: We are grateful to the key informants and focus group members for sharing their expertise and insights on nutrition data quality. The authors would like to thank The DHS Program staff, namely Sunita Kishor, for her valuable feedback on the report, and Trevor Croft, Anjushree Pradhan, and Gulnara Semenov for their review and input on the recommendations. Thanks to Diane Stoy for tirelessly editing several drafts of this report. Special thanks to Erin Milner (USAID) for her input on the conceptual design and guidance throughout the development of this report.

Cover Design: Cynthia Kramer
Editor: Diane Stoy
Document Production: Joan Wardell

This study was conducted with support provided by the United States Agency for International Development (USAID) through The DHS Program (#AID-OAA-C-13-00095). The views expressed are those of the authors and do not necessarily reflect the views of USAID or the United States Government.

The DHS Program assists countries worldwide in the collection and use of data to monitor and evaluate population, health, and nutrition programs. Additional information about The DHS Program can be obtained from ICF, 530 Gaither Road, Suite 500, Rockville, MD 20850 USA; telephone: +1 301-407-6500, fax: +1 301-407-6501, email: info@DHSprogram.com, internet: www.DHSprogram.com.

Recommended citation:

Namaste, Sorrel, Rukundo K. Benedict, and Megan Henry. 2018. *Enhancing Nutrition Data Quality in The DHS Program.* DHS Qualitative Research Studies No. 23. Rockville, Maryland, USA: ICF.

CONTENTS

EXECUTIVE SUMMARY

The DHS Program is a leading source of nutrition data globally. Producing high-quality data is core to The DHS Program's mission. Throughout the survey process, there are many measures in place that ensure data quality, and continual quality improvement measurers are regularly reviewed and refined. This report describes the factors that facilitate or hinder nutrition data quality during the survey process and provides recommendations to ensure that The DHS Program continues to provide high-quality data for its users.

The nutrition topics of interest in this investigation were anthropometry, hemoglobin, and Infant and Young Child Feeding (IYCF) practices because these topics present challenges in data collection. A total of 54 key informants were interviewed (21 who were internal to The DHS Program and 33 who were external) and two focus groups were conducted with 13 staff members from The DHS Program.

Participants emphasized the unique considerations for nutrition data quality within The DHS Program. The collection of anthropometry and anemia data requires quality measures that are distinctly different from interviewing. Participants also identified IYCF as one of the more difficult topics to collect within the questionnaire because capturing dietary data is inherently challenging and also requires context-specific adaptation of the questionnaire.

A high level of capacity exists within The DHS Program. Although informants identified important challenges, they felt that collecting quality data was achievable. The informants spoke with confidence about realistic precision in large-scale surveys, and they identified specific steps needed to improve data quality. Informants who were external to the DHS surveys cited DHS's reputation and the infrastructure built and maintained by DHS over many years, which has collected difficult-to-obtain data that informs public health planning at the global, national, and local levels.

A total of 32 recommendations emerged from the key informant interviews and focus groups. Under each recommendation are steps to be considered by The DHS Program for enhancing the quality of anthropometry, hemoglobin, and IYCF data. Informants noted the critical role of country ownership of surveys and the importance of buy-in from the host country to adopt data-quality measures for nutrition.

Collecting high-quality data in some cases will require new changes to The DHS Program's existing procedures, and in other cases it is a matter of selecting data-quality activities that have worked well in some surveys and making them standard practice. In addition, different strategies to increase data quality are recommended, although not all strategies have been tested for use in the DHS surveys, thus pilots will be needed prior to wide-scale adoption. It is important to recognize that because each recommendation in this report has budget implications, the recommendations must be prioritized.

The DHS Program is the largest and most enduring survey program of its kind and has contributed to our understanding of population, health, and nutrition in low- and middle-income countries. Strengthening the quality of nutrition data in The DHS Program will, in turn, improve decision-making for nutrition and growth. Such an endeavor will require careful evaluation of the recommendations, adequate funding, piloting of strategies, and continued monitoring over several years to achieve the desired results.

KEY WORDS: nutrition, anthropometry, hemoglobin, infant and young child feeding, data quality

KEY DEFINITIONS

accuracy	The closeness of a measured value to the actual (true) value. An *accurate* measurer is one that produces, on average, measurements close to the true or expert's value.
anemia	Hemoglobin concentration below a certain threshold, which is caused by factors that affect the morphology, production, turnover, loss, or destruction of red blood cells.
anthropometry	Measurement of the human body (height and weight) as an indicator of nutritional and health status.
biomarker	Objective physical or biologic measures of health conditions.
computer-assisted personal interviewing	Electronic system used by interviewers to facilitate the interview process. Supervisors centralize the data collected by interviewers, and central office staff monitor the fieldwork operations and further process the data using Bluetooth technology to transfer data.
Demographic and Health Surveys	Household-based surveys that provide data for a wide range of indicators for population, health, and nutrition.
field check tables	Summary of data that show whether teams are collecting data correctly by checking response rates and various measures of data quality, usually run on a weekly basis during data collection.
health technicians	Data collectors hired or employed by the implementing organization to collect hemoglobin or other blood biomarkers.
host country	Country where The DHS program is conducting a survey.
informant	Key informants interviewed for this report.
interviewers	Interviewers hired or employed by the implementing organization to collect information on infant and young child feeding.
implementing agency	An institution, usually the central office of statistics in a country, that implements a survey with technical assistance from The DHS Program.

Malaria Indicator Surveys	Household-based surveys that provide data on bednet ownership and use, prevention of malaria during pregnancy, and prompt and effective treatment of fever in young children. In most cases, these surveys include biomarker tests for malaria and anemia.
measurers	Data collectors hired or employed by the implementing organization to collect anthropometry data.
precision	The closeness of measured values to each other (regardless of accuracy). A *precise* measurer is one that produces measurements that are close to each other when re-measuring the same child.
recode microdata	Publicly available datasets from surveys.
respondent	Household members who participate in the surveys.
sensitivity	Biological susceptibility to different measurement techniques.
standardization exercise	Exercise to help measurers refine and improve their measurement technique and to show that they can take accurate and precise measurements.
south-by-south consultants	Consultants hired by The DHS Program that provide technical assistance to the implementing agency.
The DHS Program	Provides technical assistance to countries for the Demographic and Health Surveys, AIDS Indicator Surveys, Malaria Indicator Surveys, and Service Provision Assessment Surveys.
The DHS Program staff	The DHS Program staff are hired by ICF and partners: Johns Hopkins Bloomberg School of Public Health Center for Communication Programs, PATH, Avenir Health, Vysnova, Blue Raster, Kimetrica, and Encompass.

ACRONYMS AND ABBREVIATIONS

CAPI	computer-assisted personal interviewing
DBS	dried blood spot
DHS	Demographic and Health Survey
IYCF	infant and young child feeding
MICS	Multiple Indicator Cluster Surveys
MIS	Malaria Indicator Survey
MoH	Ministry of Health
MOU	memorandum of understanding
TEM	technical error of measurement
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
WHO	World Health Organization

PART I BACKGROUND

Objective

Relevant high-quality data are required for informing country and global decisions on nutrition policies and programs in low- and middle-income countries. The Demographic and Health Surveys (DHS) Program provides technical assistance for the implementation of surveys in low- and middle-income countries. Such surveys provide population-level estimates of key nutrition indicators that inform national strategies and action plans designed to improve nutrition in children, women, and men. This report focuses on enhancing nutrition data quality for anthropometry, hemoglobin testing, and infant and young child feeding (IYCF) practices in surveys by The DHS Program.

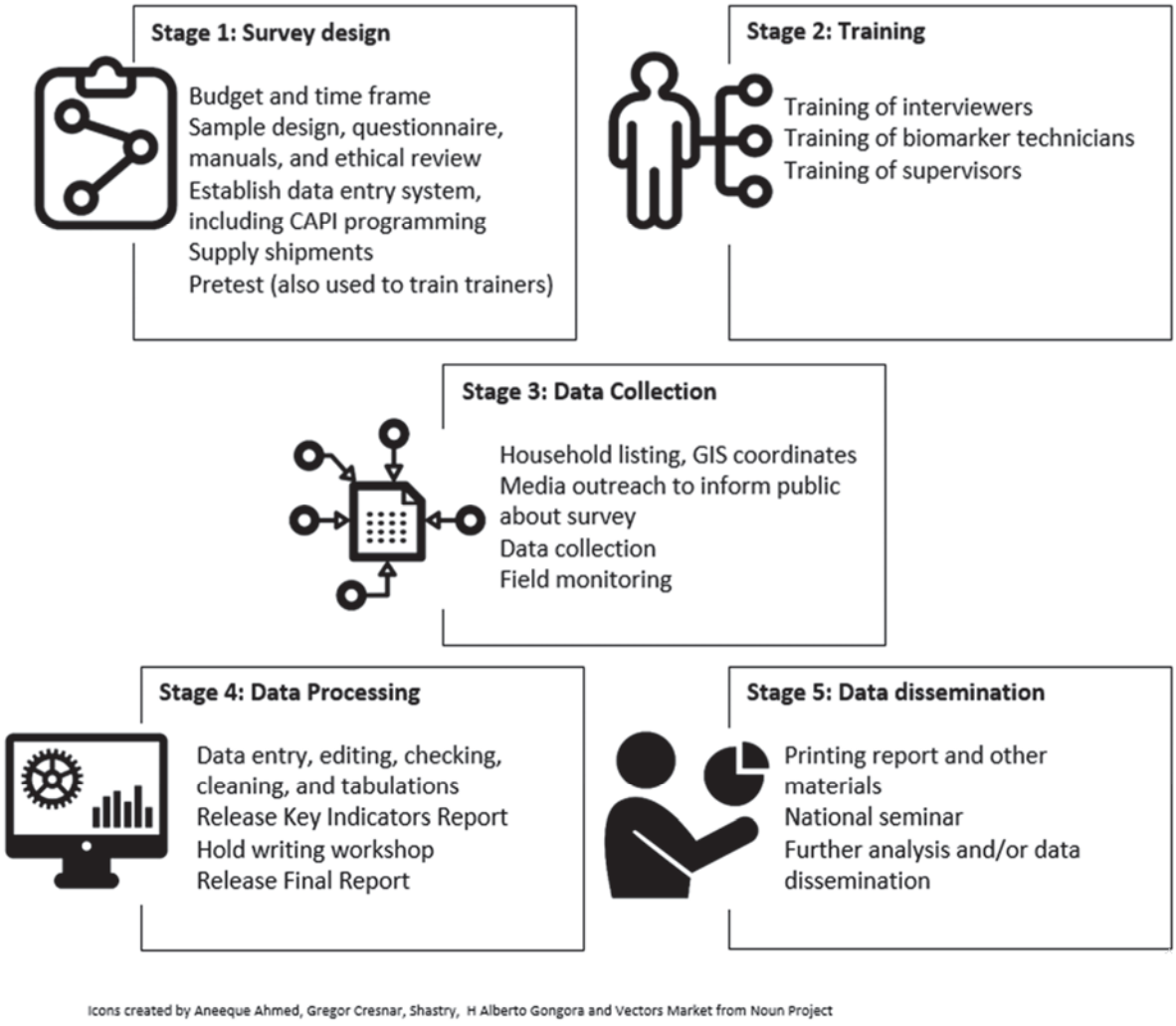
Providing reliable estimates of anthropometric data is important for monitoring global progress toward eradicating hunger, reducing health inequalities, and assessing the impact of short- and long-term nutrition and health interventions (Development Initiatives 2017). Similarly, tracking anemia, which is caused by a complex interaction among malnutrition, infectious diseases, genetics, and other factors, is important to inform the prevention of maternal and child morbidity and mortality (WHO 2011). Last, tracking optimal infant and young child feeding practices provides an opportunity to influence child cognitive development and growth and to improve the child's chances for a successful future (Kothari 2010, WHO 2008).

Within The DHS Program, two nationally representative household survey sources, the DHS surveys and Malaria Indicator Surveys (MIS), provide nutrition data. The DHS surveys provide data on a wide range of population, health, and nutrition indicators and the MIS surveys provide data that track malaria targets. The DHS surveys are the largest source of global nutrition data on anthropometry, IYCF practices, and nutrition coverage indicators. The DHS surveys, along with MIS surveys, are the primary sources of anemia data globally.

The DHS Program is committed to producing high-quality nutrition data and has quality measures in place at each survey stage: survey design, training, data collection, data processing, and data dissemination (Figure 1). Continuous quality improvement is a high priority for The DHS Program in order to provide the global health community with more accurate and expansive data.

This report identifies the factors that facilitate or hinder data quality and solutions for improving each phase of the survey process, as well as innovative approaches for collecting nutrition data. The study findings will improve the operations of The DHS Program, although the lessons learned can be utilized by a wider audience involved in the collection and use of nutrition data.

Figure 1 The DHS Program Survey Phases



Icons created by Aneeque Ahmed, Gregor Cresnar, Shastry, H Alberto Gongora and Vectors Market from Noun Project

Anthropometry, Hemoglobin, and Infant and Young Child Feeding Data

Anthropometry

The collection of anthropometric data has been a key component of The DHS Program since 1986. Anthropometric data for children are collected by measuring the height and weight of children under age 5. Height and length measurements are collected with an Infant/Child/Adult ShorrBoard®. The length of children under 24 months is measured lying down and the height of children age 24 months or older is measured standing. Weights are typically measured with a SECA digital scale that has a taring function. The measurer provides a pamphlet with the results of the height and weight measurements.

The measurers usually receive at least three days training on how to collect and record height and weight measurements (ICF International 2012). During fieldwork, multiple layers of supervision exist, and use of field check tables that monitor data quality by team are also part of the standard protocol. Field check tables are run periodically for paper-based surveys, and can be run in near real-time for surveys using the newly introduced Computer-Assisted Personal Interviews (CAPI) system.

During data processing, if a child whose age was below 24 months was measured standing up, 0.7 cm is added to the height; if a child age 24 months or older was measured lying down, 0.7 cm is subtracted from the length. This is a standard adjustment made for the WHO Growth Standards. The DHS Program then uses the 2006 World Health Organization (WHO) Growth Reference Standards to compute three nutritional or z-scores. These z-scores are HAZ (height-for-age), WAZ (weight-for-age), and WHZ (weight-for-height). After obtaining the z-scores, all data sets are cleaned by identifying cases with z-scores beyond the specified lower or upper cutoffs, eliminating extreme values that may be due to measurement or data-entry errors, and excluding them from the computation of prevalence. Low values below standard cutoffs identify stunting, underweight, and wasting.

The WHO 1995 expert committee report recommended the inclusion of specific data within the reports of anthropometry data: general characteristics of the population; sample size; measurement methods; method of determining age; percentage of excluded data; prevalence based on fixed cutoff; confidence intervals of the prevalence estimates; mean z-scores with 95% confidence intervals; standard deviation of the z-scores; and frequency distribution plots against the reference distribution. In DHS surveys, all anthropometry indicators are included in the primary survey reports, except the standard deviations of the z-scores and frequency distribution plots.

Hemoglobin

The WHO recommends hemoglobin as the biomarker that determines population-based prevalence of anemia (WHO 2011). Hemoglobin data collection has been a component of The DHS Program since the mid-1990s. In DHS and MIS surveys that include hemoglobin measurement, hemoglobin data is collected from children age 6-59 months. In DHS surveys that include hemoglobin measurements for children, most DHS surveys also include measurements for women age 15-49, and about one-third of the surveys also include measurements for men typically age 15-54 or 15-59.

The hemoglobin concentrations are measured by using a drop of capillary blood with the HemoCue 201+ and occasionally with the HemoCue 301. The DHS Program protocol specifies obtaining blood from a

finger for adults and children age 12-59 months, and from the heel of children age 6-11 months (ICF 2012). The skin is warmed by rubbing the hands or heel to increase blood flow. The hand or heel is then placed below the level of the heart, and the finger or heel is cleaned with alcohol. The third or fourth finger is selected for a finger prick. With a finger prick, the health technician uses a rolling movement of the thumb to lightly press the finger from the top knuckle toward the tip, and then maintains a gentle pressure to trap the blood. With a heel prick, the health technician applies light pressure around the heel. The skin is cleaned and then pricked with a sterile, retractable lancet to obtain the blood sample.

In general, after the first two free-flowing blood drops are wiped away with a sterile piece of gauze, the third blood drop is drawn with a microcuvette. A fourth or fifth drop of blood is sometimes used for hemoglobin measurements when other biomarkers are being tested. The blood drop is collected in the microcuvette directly from the finger or heel without touching the finger or heel. The health technician must ensure that the microcuvette is completely filled without air bubbles. The microcuvette is wiped clean and then inserted into the photometer, which displays the hemoglobin concentration reading and allows the immediate return of results to survey participants. The health technician records the hemoglobin concentration in grams per deciliter (g/dL). The health technicians provide a pamphlet and feedback to the respondent about their hemoglobin concentration, and make referrals as necessary to health facilities.

Health technicians are trained to collect hemoglobin data usually in conjunction with training for anthropometry or other biomarker data collection. The health technicians receive supervision from the team supervisors, field coordinators, and technical assistance providers. As mentioned above for anthropometry, field check tables are used to monitor field work.

Hemoglobin concentrations are adjusted for altitude and cigarette smoking because both increase hemoglobin concentrations (World Health Organization and Centers for Disease Control and Prevention 2007). The DHS Program does not identify and exclude biologically implausible hemoglobin concentrations, since there are no international guidelines for removing such cases. In DHS surveys, the cutoffs in Table 1 are applied to hemoglobin concentrations to obtain population estimates of anemia. In MIS surveys, severe anemia is reported as hemoglobin concentrations <8.0 g/dL.

Table 1 Hemoglobin concentration cutoffs used in The DHS program to define anemia at sea level

	Anemia measured by hemoglobin (g/dL)			
	Any Anemia	**Mild**	**Moderate**	**Severe**
Children age 6-59 months[1]	<11.0	10.0-10.9	7.0-9.9	<7.0
Nonpregnant women age 15-49[2]	<12.0	10.0-11.9	7.0-9.9	<7.0
Men age 15 and above	<13.0	10.0-12.9	8.0-10.9	<8.0

[1] Malaria indicator surveys define severe anemia as <8.0 g/dL.

[2] The WHO recommends slightly modified hemoglobin concentration cutoffs to define anemia for nonpregnant women: mild anemia (11.0-11.9 g/dL), moderate anemia (8.0-10.9 g/dL), and severe anemia (<8.0 g/dL).

Infant and Young Child Feeding (IYCF) Practices

The DHS surveys provide data on breastfeeding practices and complementary feeding indicators (WHO 2008, WHO 2010). The DHS Program has collected information on infant feeding practices since 1985. In DHS, questions on ever breastfeeding and bottle feeding are asked of mothers whose children were born since January of the year five years before the survey. The early initiation of breastfeeding and current breastfeeding status are only asked for the last-born child. Questions on food and liquid consumption are

asked about the youngest child born since January of the year two years before the survey and living with the mother of an interviewed woman.

The DHS survey interviewers are trained to ask mothers about their breastfeeding practices and all foods and liquids the child consumed in the previous 24 hours. Interviewers read from a list of foods grouped into categories and mark YES or NO depending on whether or not the child ate any food in specified food categories. Survey managers work with the implementing organization and others to tailor the food lists to the local context and to group the food items into categories.

The DHS Program provides information on the status of breastfeeding (early initiation, exclusive breastfeeding under age 6 months, and continued breastfeeding at age 1); introduction of solid, semi-solid, or soft foods; minimum dietary diversity; minimum meal frequency; and minimum acceptable diet (Kothari 2014), calculated according to the definitions in the 2008 IYCF indicator guide (World Health Organization 2008).

PART II METHODS

This chapter provides a description of the methods used in this study to collect and analyze information that forms the basis of recommendations for enhancing data quality in The DHS Program and other surveys.

Key informants were selected purposively. The first round of informants included individuals identified as experts in the collection and use of nutrition data. The informants were both internal and external to The DHS Program. Additional informants were selected with a snowball approach by asking key informants to identify additional participants. Key informants were contacted via email to request interviews. Interviews were conducted in person and by telephone by Dr. Sorrel Namaste of The DHS Program. Interviews were audio-recorded and then transcribed verbatim.

A semi-structured interview guide focused on the informants' perspectives of the facilitators and barriers experienced at various stages of survey implementation, and innovations in three domains: anthropometry, IYCF, and hemoglobin assessment. Key informants were asked to respond to questions relevant to their areas of expertise. Appendix I provides the initial interview guide used to explore barriers and facilitators in the data collection process, as well as nutrition innovations. The guide was modified as new concepts emerged during data collection. A survey was administered with SurveyMonkey to key informants to obtain demographic information.

Subsequently, two focus groups were conducted by Drs. Sorrel Namaste and Rukundo Benedict (also of The DHS Program) with The DHS Program survey managers to explore issues that emerged from the initial key informant interviews. The survey managers have provided technical assistance for surveys in Asia, Africa (Francophone and Anglophone), and Latin America. The focus group method increased the exchange of views and the generation of new ideas. The first focus groups, which utilized a semi-structured guide, lasted approximately two hours. A more detailed interview guide was created for the second focus group to explore key topics that emerged from the in-depth interviews and the first focus group.

The interviewers obtained verbal informed consent from the key informants and focus group participants prior to data collection. In reporting results, no statements were attributed to individuals, although the interview participants agreed to their names being listed in the report's appendix. The ICF Institutional Review Board determined that this research was not considered human subjects research, as defined by the regulations. The investigation asked experts to provide feedback that could improve data collection and they were asked only for their expert opinions.

Data Management and Analysis

Recordings and transcripts were stored on The DHS Program SharePoint site in a private folder that could be accessed only by the study investigators and study staff. The participants were identified only by a participant ID number. Transcripts were entered into NVivo 11 software (QSR International Pty Ltd, Melbourne, Victoria, Australia) for data management and content analysis.

For in-depth interviews, a codebook was developed *a priori* based upon the interview guides designed by Dr. Rukundo Benedict and Dr. Megan Henry, Research Associate, Department of Epidemiology at Johns

Hopkins Bloomberg School of Public Health, who worked as an independent nutrition consultant for this project.

Coding and content analysis of a subset of transcripts were conducted iteratively by two independent researchers over a period of five months. Inconsistencies in coding were resolved to achieve consensus on all codes.

The focus groups explored the emergent themes from the key informant interviews. During the focus groups, note cards and flip charts were used to record facilitators and barriers for each nutrition topic. This information was saved and used during analysis, along with notes taken by a designated notetaker.

Informants

A total of 57 informants were identified and contacted, and 54 were successfully interviewed between February and July 2018. Only three were unable to participate due to scheduling conflict. Appendix II provides a list of the participants interviewed. Most interviews were conducted with one interviewer and one informant, although up to four informants were present for some interviews. A total of 21 informants were internal (including a south-by-south consultant) to The DHS Program. The remaining 33 were external, and worked for a nongovernmental organization on nutrition-related work (n=7), the U.S. Federal government (n=5), some other government (n=4), an extra-governmental organization (n=6), academia (n=6), or consultancies (n=5).

A total of 49 of the 54 informants provided demographic information based on the SurveyMonkey request. Educational attainment was high among informants, with 44% having a doctorate and 52% a Masters or professional degree. The median years of survey experience was 15 (0, 47), and the median years of experience in nutrition was 14 (0, 48). The types of survey experience included a wide range of surveys under The DHS Program, as well as other surveys that focused on demographics, health, and nutrition.

Nearly all informants had visited the DHS website, which was considered a proxy for familiarity with The DHS Program. All but one informant reported using DHS data or reports for a wide range of uses such as design of other surveys (n=29), program implementation and design (n=15), monitoring and evaluation (n=14), global tracking (n=24), and research (n=30). Of the 84% who had analyzed DHS data, the majority had used the DHS recode microdata, and over half had used STATcompiler.

Focus groups were held on April 17 and May 30, 2018 with 13 DHS survey managers. There was a range of experience among survey managers with the median survey experience of 10 years (with a range from 2 to 35 years).

In this report, “informant” refers to those interviewed exclusively for this report, while “respondent” refers to household members participating in the surveys. “Measurers,” “health technicians,” and “interviewers” refer to those who are hired by or are employees of the survey implementing organization and who collect anthropometry, hemoglobin, or IYCF data. In this report, the number that follows the quoted verbatim material indicates the interview or focus group where the quote originated. Thus, one number may reflect the views of more than one informant.

PART III FINDINGS AND RECOMMENDATIONS

The DHS Program sought input on nutrition data quality from experts from The DHS Program and external experts with experience collecting nutrition data in large-scale surveys. The findings and recommendations are presented separately for anthropometry, hemoglobin, and IYCF. Each nutrition topic is discussed by the survey phases: design, training, data collection, data processing, and dissemination of the results. Since many findings for anthropometry and hemoglobin were similar, the hemoglobin discussion includes references to the anthropometry section, and the findings that are unique to hemoglobin. There is also overlap in the discussion of quality improvements in the different phases of the surveys because the activities in each phase of a survey influence one another.

The findings and recommendations from this report should be considered in the full context of all survey data in The DHS Program, and not the exclusive context of the nutrition data. The large, diverse amount of available data is a major strength of The DHS Program. However, for the nutrition-specific improvement efforts to be integrated successfully, the broader survey needs must be considered. As one informant said, "*For nutrition, of course, there are plenty of things we could do, but it's not only nutrition in the survey. We have all the other aspects in the survey, and we need to find the right balance between what we can do to improve data, improve the quality and so on, and on the other hand, keeping the budget and the organization of the survey manageable." R27*

Anthropometry

Design Phase

Improving anthropometry data quality begins at the onset of the survey. Several informants reported that parameters determined during the survey design phase are important for anthropometric data quality. Many emphasized the importance of recognizing that measuring anthropometry is distinct from interviewing, and that failing to recognize this early can affect data quality.

> *"There are sort of two extremes, there's a fear of not being able to collect it well, but then on the other side of the spectrum, there's kind of this sense of, 'Well it's only anthropometry, how hard can it be?' So I do think that there has to be a meeting in the middle and an understanding of just how standardized the techniques for collecting anthropometry need to be." R03*

> *"I think for anthropometry, some people think it is easy, some people think it is just you know, measuring kids' and womens' height and weight. They think it's simple, when it actually requires such precision and training and a lot of hands-on practice, and I think people underestimate that effort." R01*

> *"The data quality, you have to start in the very beginning…Anthropometry looks very simple to measure but to have good quality measurement, it's not that easy." R16*

As a first step in the survey process, donors, the implementing organization, and The DHS Program typically sign a memorandum of understanding (MOU) before survey design, training, and subsequent activities begin. Informants reported that the MOU is helpful in describing the different stakeholder roles and key decisions that have been agreed upon, including those for anthropometry data collection. In some surveys, however, the fact that an MOU is not in place when survey design begins can affect design decisions for anthropometry data quality if the agreements have not been determined in advance. In addition, some informants expressed the need to include more specifics in the MOU related to data quality to hold all partners accountable.

> *"An MOU was helpful just to sort of lay out the different roles of different people so that we could have everything all in one place to refer to." R15*

Engagement of host-country stakeholders in data quality is limited and has the potential to build capacity and improve quality. Informants explained that the initial visit is more about stakeholder buy-in and less about planning. As such, informants recognized this as an opportunity to establish stakeholder commitment to quality anthropometric data collection.

> *"So in terms of success, in terms of our partners…the success will come from if they have the same understanding that this is not just something you can do quickly, and it is not a light effort." R01*

Informants suggested engaging with the survey steering committee, which typically include personnel from the host-country Ministry of Health and other government organizations, the USAID mission and other U.S. government partners, bilateral organizations, and multilateral organizations. Engaging the steering committee was important to foster a better understanding of requirements for anthropometry data collection with high-level stakeholders so they can be advocates for anthropometry or at least not undermine the required inputs.

> *"Our [DHS] surveys are not done by the Ministry of Health, they are run by the Stat[istics] office. The Stat[istics] office is not always the best organization [to] look at the content of the questionnaire … So that is the reason why there is a steering committee." R27*

"I think it should be a requirement to involve nutritionists either from the Ministry of Health, or if there is a national expert committee [of nutrition experts] already set up to be involved, that would be useful…Plus, ideally, another person either from the Ministry or maybe the UNICEF office (depending on people's backgrounds) that really understand what we're trying to actually measure. It is one thing to understand nutrition, but it's another thing to really understand what we're measuring." R13

In some host countries, the nutrition expert representation on the steering committee is helpful in overcoming misconceptions about anthropometry data collection, although representation of nutrition expertise is not standard across all surveys. Many informants reported that steering committee members were often responsible for broader topics than nutrition, or were high-level officials who were often not involved in technical decisions that can influence data quality. There was also frustration expressed that there is often not nutrition-specific funding for the surveys, but there is an expectation for a wider focus on collecting this type of data.

"The same person who is responsible for maternal and child health [is also] responsible for nutrition from the Ministry of Health, so when they come, [what] they represent is very broad." R27

"And this is interesting because what I have seen in the countries is that the people who work in nutrition, they don't have necessarily a strong voice in all countries. In some countries, they are very strong, but not entirely noticed. So their participation in the initial discussion is not necessarily something that you see in every country." R21

"I feel like more often than not, the nutrition community hasn't been particularly involved in the funding or in the design, and then they certainly come out when the data are out and they want to capitalize on using it." R25

Steering Committee involvement is important for buy-in, but a technical committee may be better suited to oversee data quality. In DHS surveys, some countries have established technical committees under the steering committee. These committees can meet throughout the survey process to address technical issues. Informants found these committees useful and could envision a larger role for this type of committee in the oversight of anthropometry data quality.

"People on steering committee are very high level and they are pushing for their topic of interest so it doesn't get dropped. They are often big shots that don't know or care about [training and logistics]." FGD

"We have technical committee that meet more often than the steering committee, with more technical people. So in this committee we have people from Ministry of Health, and depending on what we have in the survey, we have nutritionists, people from malaria, HIV, and so on." R27

In some DHS surveys, informants reported the use of technical working subgroups. However, the extent to which small technical groups were established and utilized varied by survey. One informant reported that group members included representatives from the Ministry of Health, nutrition, and USAID, and that the small technical groups discussed data-quality considerations and reported back to the larger technical working committee, which in turn reported to the steering committee. The informant felt this model was helpful because the technical group discussed lessons from the previous survey that were used to improve the current survey.

"The small groups are part of the technical working committee and then we had the steering committee… [The technical group] met at the design stage…in the training they came in, collection stage, analysis and dissemination…" R29

The DHS Program technical staff can provide input at the design phase, which will increase data quality. Informants reported that engaging the technical staff from The DHS Program can help to ensure that survey design decisions made early in the process can positively influence data quality. Informants reported consulting with staff on issues such as sample size and the number of required biomarker supplies. A spreadsheet tool, developed by The DHS Program, that calculates inputs was reported to be a useful guide in this process. However, informants expressed frustration in being constrained by decisions made in the design phase that prevent them from taking the steps necessary to ensure data quality at the phase of the survey for which they are responsible.

> *"You respond to [the survey design], but we don't influence it very much." R05*

Informants expressed the need for greater involvement of technical experts in both nutrition and biomarkers during the early design phase. The DHS Program's survey inception meeting was reported to be an effective measure already in place to engage staff. However, since these inception meetings often take place after many design decisions have been made, two meetings (one pre-design and one post-design) may be needed. In addition, since these meetings focus primarily on logistics, the attendees may be unable to discuss the more technical aspects of data quality.

> *"Well, I think we can alert survey managers to the history of [data collection in] a country. [For example, if] during the previous surveys there have been issues related to data quality in the past, such as high levels of incompleteness, or heaping." R12*

Some informants also expressed concern that important aspects of survey design that will influence anthropometry data collection are sometimes neglected because the attention is focused on questionnaire adaptation, translations, ethical approvals, and other concerns that arise. One solution is to encourage good communication among regional coordinators, survey managers, and biomarker trainers to facilitate the necessary arrangements for high-quality anthropometry training and standardization.

> *"The very next visit is of course typically the questionnaire design… issues around buy-in to various specific parts of the questionnaire is…supposed to happen at the questionnaire design meeting, but…there are people with so many varied interests who are pushing for all kinds of things." R21*

Recommendation 1: Involve host-country stakeholders in decision-making and planning for high-quality anthropometry data.

Steps for consideration by The DHS Program include:

- Requiring a host-country stakeholder with expertise specifically in nutrition on the survey steering committee and technical committee who can address the necessary requirements for quality anthropometry data collection.
- Defining more instructive roles for the survey steering committee and technical committee responsible for the quality of anthropometric data.
- Developing materials for the steering committee and technical committee members to better understand the processes required for high-quality anthropometry data collection.
- Increasing the involvement of The DHS Program technical content experts (nutrition and biomarker) at the beginning of the survey design phase.

What gets budgeted gets done. The senior staff from The DHS Program develop the budget for the technical assistance costs of the survey, and the implementing agency usually calculates the local costs with the advice on line items of the senior staff from The DHS Program. Balancing requirements for high-quality nutrition data with other issues of concern is a challenge. Informants expressed varying opinions about creating an optimal balance. Some stated that cost is always going to be a constraint and that reasonable decisions must be made. Others felt that current budget constraints compromised data quality. Since the implementing agency decides its own budget, the informants emphasized the need for increasing the implementing organization's awareness and buy-in for special budget line items for activities that will ensure high-quality anthropometry data.

> *"Well, sometimes we recommend something in the budget and then the implementing agency says we don't need that much time and we need the money elsewhere, and so then they decide to cut it back." R05*

> *"Yes, everything gets rushed because of limitations on the budget." FGD*

Informants emphasized the unique requirements of funding anthropometry data collection that can be overlooked in a larger survey, such as hiring the appropriate staff and replacing staff if anthropometry data collection standards cannot be met, adequate time for training, incentives for standardization exercises, and anthropometry-specific oversight for quality control and data collection. Standardization exercises, as discussed below, are one of the most critical elements of data quality and must be explicitly accounted for in the budget. However, the implementing agency is often the host country's Bureau of Statistics, which may not have the technical knowledge to provide sufficient budget coverage that can meet the anthropometry data-collection requirements during the design visits. Informants emphasized that building nutrition capacity and commitment from the implementing agency ensures proper budgeting for activities that foster high data quality in anthropometry.

> *"There usually is the commitment to funding [anthropometry]. Whether there's always total recognition of the complexity – for instance, arranging for the training – that could be an issue." R11*

> *"Very often anthropometry is the first thing to get cut [from the DHS survey], as far as funding. 'Well we don't have funding to have the trainer be in the field to look at quality control and data collection.'" R02*

> *"For local agencies, money is what matters. They don't have it in their budget, they're not going to do it, no matter what pressure we put at that time." R21*

> *"Put it in the sub-contract for the work plan description, so like require for standardization at least 5 pairs of mother and child, or 10 pairs of mother and child. And you have to discuss it in advance because it requires money and time." R27*

> *"What works well so far is: when the people responsible for survey design, regional coordinators, when during survey design, they have conversations with different partners about the capacity-building aspect, especially pertaining to training, even during survey activities – field staff training, field interviewer training – enough budget is allocated to that. The feedback that we received that enough budget is not allocated or negotiated for anthropometry to allow for standardization exercises, for example. So what works well in most countries is if there's enough funding, then enough time and resources is allocated for that component." R25*

Subsampling could save costs and shift funds to anthropometry data quality. One recommended strategy to improve data quality, while reducing costs, was obtaining anthropometry data on a subsample. One informant explained that since DHS surveys measure total fertility rate and child mortality, the survey sample size at minimum needs to be sufficient to estimate these rare events at the subnational level. Other considerations when determining sample size are implementing capabilities, budget, and the other

indicators being measured. Prevalence estimates of nutritional status can generally be calculated with a smaller total sample size and still provide nationally and subnationally representative nutritional status estimates, as well as estimates disaggregated by demographic variables. Keeping the sample size for anthropometry data to a minimum was recommended by several informants as subsampling for anthropometry data would conserve resources and improve anthropometry data quality. Subsampling has been introduced in some surveys and was reported to work well.

> *"Actually you see for any nutrition indicators, you do not need [a] huge sample size, you can decide [to] do it in a subsample of households…" R04*

> *"…keeping the survey sample size to a necessary minimum and keeping the number of survey teams to a minimum [means] that we can meaningfully well-train and well-supervise…" R10*

Team composition plays a major role in the collection of high-quality data. Team composition varies substantially based on the questionnaires and biomarkers, the number of households visited, and the ease of navigating the geography of the host country. In the DHS surveys, the interviewer often used to collect anthropometry data, thereby saving costs and streamlining engagement with the households.

> *"A few years ago, when the surveys were less complex with less biomarkers, and the questionnaire was shorter and so on, and when we used to do survey on paper, I was pushing to not have any technician in the team, but to have the interviewer in charge of everything." R27*

However, informants identified a number of difficulties that occur when interviewers collect anthropometry, in particular insufficient time to train the measurers. Several informants suggested that a dedicated member of the data-collection team should serve as the primary measurer. This would allow for a manageable number of staff to be trained and participate in the anthropometry standardization exercise, rather than training all interviewers who have to master many learning modules. While The DHS Program staff prefer specialized individuals to collect anthropometry, this is not always done, especially if biomarkers are limited to anthropometry and hemoglobin.

> *"To look at it in terms of a potential issue of data quality: how the teams are organized, how many people we will have, if these people working in anthropometry will specialize or not, because this can change completely. In some countries, they insist to train everybody in the teams, in other[s] we are focusing only on 1-2 people that are specialized on that; and their training is parallel, in others sequential. So I think for me in terms of the design, those are the key issues." R21*

> *"Consensus here in DHS is that it's better to have people specialize by biomarker in general, not to have everybody train on everything because that makes the training longer … you cannot focus on really doing a good training." R21*

Nearly all informants noted that having two individuals to take the anthropometry measurements is critical. The assistant measurer plays an important role, especially when a child is recumbent (making sure the infant's spine is not arched and head is against board), and when a child is standing (check child is correctly positioned against board, or when the child is agitated, to keep legs in position). This also allows the mother to focus on reassuring and calming the child to reduce resistance from the child and to enable more precise measurements. Informants suggested one way to ensure members of the team works in pairs to measure children and provide only one set of height boards/weighing scales with the option of replacement in case of damage.

Careful selection and over-recruitment of measurers will help ensure appropriately staffed fieldwork. Several informants noted that high-quality data begins with the recruitment of appropriate individuals to participate in the required training. In The DHS Program, the implementing agency is responsible for all recruitment including measurers. Several informants reported that there were different practices for recruitment. Some had very structured recruitment processes while others did not follow formal guidelines. The result of these different recruitment practices is that the quality of measurers for anthropometry varies by locale.

> *"We are not recruiting interviewers, health technicians, or whatever. It's done by the implementing organization. So we have some guidelines. We say, 'We need to have people with this level of education, blah blah blah,' but at the end of the day, they do basically what they want." R27*

> *"Some countries make a big deal out of it, they put [out an] advertisement, then go through rigorous interview, they have a panel…" R27*

> *"The quality of people at the end of the day that we have in the field, it depends a lot from country to country." R27*

Taking anthropometric measurements, especially from children, can be physically taxing because it requires kneeling and bending. A few informants noted that some recruits are challenged by the physical demands of anthropometry measurements and, yet, they are still hired to undertake the data collection. At the same time, however, it is critical not to discriminate during staff recruitment.

Informants reported that it is necessary to recruit a larger group of measurers than is needed. In The DHS Program, attrition is usually built into the number of trainees recruited. However, many informants reported the level of over-recruitment budgeted for is insufficient, and the implementing agency often does not even meet the current recruitment number. As informants explained, having a larger number of recruits is helpful if they drop out because of illness, family emergency, or the inability to perform tasks. Another informant added that an advantage of a larger recruitment pool was the possibility of having reserve measurers who can either replace those that do not meet the evaluation criteria, or allow those who performed at higher levels to be promoted to a supervisory role.

> *"If for example, you have a lot of people for attrition, then you're fine: you just select them and then you use these people for reserve." FGD*

A system for supervising anthropometry data-collection needs to be established at the onset. Many informants said that establishing a system for anthropometry supervision must take into account the different supervision requirements from interviews. Rotating biomarker supervisors for anthropometry data and other biomarkers was a strategy supported by many informants.

> *"Sometimes but not automatically we add additional training for biomarker person [from The DHS Program staff] to do some field operation, but I will say it's only in a few cases because it's too costly." R27*

> *"During the design visit…we try to ensure that their number of people, duration of stay, is actually enough, so that people don't try to cut corners and that we build supervision within different levels, within the team, within the headquarters, so there is enough supervision, and especially for example, for some surveys it's kind of natural because for malaria, you have to bring slides at least within 14 days, so they have this roving teams going and collecting and at the same time, they supervise and observe. But if you don't have money for this, then it's really within the team supervisor." R27*

Informants found that setting expectations and requirements for supervision in the design phase was critical, but noted that this was not always done. Informants explained that making these early decisions with stakeholders on the steering committee was helpful later in limiting those staff who did not meet the supervisory criteria from becoming anthropometry supervisors during fieldwork.

> *"Initially we had some bad experiences with some of the [high-level government staff] being supervisors. So we say 'Okay, any supervisor has to have been in the training all time without exceptions.' And that [didn't] work very well in the beginning, but we insisted, and we said no one can be supervisor if he has not spent [time in] the training." R28*

Recommendation 2: Place anthropometry data-quality considerations at the forefront of the survey design.

Steps for consideration by The DHS Program include:

- Ensuring that funding for anthropometry data collection is adequate to meet the recommendations in this report, including a budget line item for child anthropometry training-standardization exercises.
- Collecting anthropometry data on a smaller subsample using a sample size that allows for nationally and subnationally representative nutritional status (and disaggregated by demographics) for each target population (women, children).
- Including a primary and assistant anthropometrist in each interviewing team, preferably with only one primary measurer per team.
- Over-recruiting a minimum of 15% of measurers.
- Establishing a supervision system that takes into account anthropometry-specific data-collection needs to ensure correct measurement procedures are performed.

Training Phase

Opportunities for continual learning and standardization are needed for The DHS Program staff. Survey managers are responsible for overall survey operations, and they sometimes provide technical assistance with the anthropometry training. In other cases, The DHS Program biomarker specialist staff provide technical assistance on anthropometry training, especially in surveys that include a greater number of biomarkers.[1] In the interviews and focus groups, many informants noted the varying degrees of expertise among The DHS Program staff on anthropometry measurement. The informants emphasized that greater investment in training for The DHS Program staff would improve the confidence of the staff who work on surveys

> *"Survey managers can do it [anthropometry training]. I have a medical background. You don't have the lab knowledge, but for anthropometry it's something survey managers can be trained to do." FGD*

The informants discussed the issue of determining the appropriate staff who should be trained in anthropometry measurement. Most informants recommended different levels of training. Higher-level training on the inputs required for an effective anthropometry training would be provided to survey managers, and more intensive training should be provided for any staff at The DHS Program who are directly training anthropometrists in survey trainings. Informants noted training is already provided but it is often on-the-job training and there was consensus that overall training of staff should be strengthened.

> *"At a main training, I also did biomarkers, but I had another survey manager. There is no way to do that unless you have another Survey Manager." FGD*

A mechanism raised by informants to reinforce the quality of the training was to establish eligibility requirements for staff providing technical assistance on anthropometry. The most important requirement of the anthropometrist certification raised would be achieving adequate technical error of measurement values for precision and accuracy. This would allow a trainer to serve as the gold standard measurer during the administration of the anthropometry standardization exercise (see recommendation 5). As discussed in this report, an ongoing challenge is obtaining sufficient numbers of children for the anthropometry standardization exercises or the survey trainings; this challenge also extends to training courses for The DHS Program staff. Thus, informants emphasized the importance of conducting the certification training in settings where it is feasible to conduct an anthropometry standardization exercise. Informants suggested developing an anthropometry training certificate and requiring at least one trainer on each survey to have this certificate. Informants also noted the need for refresher trainings and recertifications.

> *"One of the things that we don't do well (and have not really ever had funding, separate funding, or a lot of thought given to it), is all staff I think should go through a recertification every year, even if you have been doing it for 30 years." R21*

Consultants are an essential resource and have local knowledge, but oversight of their work is limited. In The DHS Program, biomarker consultants are frequently used for survey work. Several informants noted that the consultants are highly skilled with blood biomarkers, but are less experienced with anthropometry measurement. Since consultants encounter both biomarkers in surveys, it is important that consultants be equally skilled in both.

[1] Consultants also conduct the anthropometry training sessions, and this is covered in the next section.

"So, we have people who are more specialized in malaria or HIV or whatever. And also some biomarker people they are more specialized in the lab work, their training. But for anthropometry, basically we train them." R27

Informants recommended improving the training and supervision for biomarker consultants. The DHS Program held one biomarker workshop to train south-to-south biomarker consultants during DHS-7. However, informants stressed the need for continual training and supervision beyond a singular training workshop per DHS round.

"We did do a biomarker training with consultants…and I think that has to be a bigger part of what we do going forward because if we are using more and more, we need to be able to standardize across them." R21

"Many of the consultants we have now, we got them through this process: we would work with them in the country, we observed them, we saw that they were pretty good. And so then we had this biomarker training workshop. That's usually how it is done, how it has been done in the past, whether or not it needs to be a different mechanism going forward." R05

Informants recommended that the consultants participate in refresher training and anthropometry standardization exercises because many were trained on anthropometry measurement during The DHS Program Biomarker Workshop years ago. In addition, several informants cited the need for continuous supervision of consultants, including the development of operational guidance on the actions that are necessary for adequate oversight. The DHS Program biomarker team could provide oversight for consultants on topics such as creating agendas, planning for recruitment of children for practice sessions, designing exercises for the anthropometry standardization exercise, and reviewing standardization exercise results.

"Obviously given everything, U.S. rates and things like that, it could be more expensive. But bottom line in this is: I do think we have to have retrain – first of all a baseline training – for these people. Unfortunately for anthropometry or anemia, none of this can work online. You have to have a face-to-face…they are retrained or rechecked at that time by our own biomarker staff and they're sort of validated at that point. But the more we use people who we have not directly trained, the less we know about the quality." R21

"An issue with measurement of height: the protocol varies from trainer to trainer." FGD

"…training [consultants] once is great, but because of the complexity of biomarker activities, it would have been nice to have a refresher training." R25

Suggestions for training consultants also included biomarker consultants traveling with biomarker specialists to promote in-person training and supervision, and the possibility of remote supervision and supportive coaching.

"…at least for the first [survey], they [consultants] get to travel at least once with our experts to learn from them and the way of doing biomarker training or activities… But also in addition to that, they should have continuous support, mentoring and coaching, even if it's virtually." R25

The DHS Program Biomarker Staff and Consultants can be trained together. Regular workshops with The DHS Program staff and the biomarker consultants would be an efficient way to improve and maintain the skills necessary for quality anthropometry data collection across DHS surveys.

Host-country nutrition stakeholder involvement in training is important, but they do not always have the necessary skills for correct anthropometry measurement. Many informants reported that the attendance

of nutrition experts from the Ministry of Health and others' involvement during training was helpful, and further facilitated local ownership of the survey and buy-in for the final survey results. In general, participation by host-country stakeholders was good, although this was compromised if participation was not budgeted in advance, or if they were not invited.

Host-country nutrition stakeholders can serve as cofacilitators during training. However, all local representatives are not adequately skilled in anthropometry data collection. Thus, they may provide incorrect information during small group sessions, which causes confusion and could affect the quality of the anthropometry data collection. Informants suggested requiring host-country nutrition staff to either participate in the pre-test or hold a mini-workshop prior to the beginning of the main training to ensure nutrition host-country staff have the necessary skills.

> *"What worked for me in [country], was having UNICEF or MoH really get involved in the training…[but] Sometimes it works, and sometimes it doesn't." FGD*

> *"You would have people [country-level] who have a great experience in anthropometry…and then you have the other half that are, you know, professional pencil pushers that thinks they know what they're doing and that has just led to outright disasters." R13*

> *"There should always be a local trainer. A local trainer brings a different dynamic to the training, even in terms of the trainees: how they respond, how they interact, how they interest, and it's also good for them to the local country to feel that it has some kind of ownership in the training, rather than us going there." R05*

Recommendation 3: Provide continual training for The DHS Program staff and consultants, and strengthen the oversight of consultants.

Steps for consideration by The DHS Program include:

- Developing a learning program in which consultants have an opportunity to shadow The DHS Program staff on a survey that measures anthropometry.
- Hosting a minimum of three biomarker workshops per five-year project period, and establishing a skills certification program that includes passing an anthropometry standardization exercise.
- Allowing consultants to work independently on a survey only if they have been certified in the previous two years.
- Establishing a process for The DHS Program biomarker staff to provide oversight of consultants' work on all surveys, either remotely or face-to-face whenever possible.
- Involving the host-country nutrition experts and setting up mechanisms to ensure these stakeholders have knowledge on the correct anthropometry measurement procedures.

The needs for anthropometry training require setting expectations and clear communication. Informants said that the implementing agencies often do not realize that the training requirements for biomarkers differ from interviews. Consequently, requirements such as having all supplies available and well-organized, dividing large groups into smaller groups, sufficient space for practicing the techniques, and children on whom to practice are not in place for a successful training. Changing this misconception requires clear communication among The DHS Program survey manager, biomarker specialists, and the implementing agency.

> *"So I think the survey manager has to be in, and the biomarker person, have to be in good communication with each other. And I think the survey manager has to be very well-organized because they're the person who's responsible for communicating a lot of these requirements to the implementing agency." R05*

> *"Supplies for training need to be well-organized because you can spend a lot of your time packing and unpacking things, looking for things. Things haven't arrived. We knew they weren't going to arrive in time and they started the training anyway, or you know, there's a lot of time and effort that can be diverted from technical work on training by trying to get the supplies in the right place at the right time." R05*

The anthropometry training schedule is insufficient. Some informants reported a misconception existed that the training for anthropometry and other biomarkers could be done quickly, and that there was resistance to training schedules that were perceived as too long.

> *"…There's often a perception that this can be done much more quickly. And so [it is necessary to be] really pushing that the key part of the survey is the whole training." R15*

> *"More time for training for everyone. If you train someone and they have a lot of practice, they will be able to perform the work." FGD*

> *"One of our biggest challenges: how do we get countries to commit to the level of intense training that is required for anthropometry?" R21*

The DHS survey primary training requires between 4-6 weeks, while the training on biomarkers may last one week or more depending on the number of biomarkers. Some informants reported that three to four days of training on anthropometry data collection with two days set aside for standardization exercises was sufficient (i.e., total of 5-6 days), although this was not always the norm, with many reporting three days total, including the standardization exercise. Others thought that further guidance was needed on the amount of time required for training.

> *"We don't know exactly what is the optimal duration of training in each of these components. Some people think that for anthropometry you can train 2 days and that's enough, but I think it's not correct. It's like 3, 4, 5, 6, 7 days." R21*

Informants also reported that biomarker trainings on anthropometry and blood collection was usually done in blocks. One informant suggested alternating between topics. Often partners involved in the biomarker training are interested either in the anthropometry measurement *or* blood collection, and they often perceive the training as not focused sufficiently on their topic when delivered in the traditional way. This becomes problematic when stakeholders attempt to change the agenda in the middle of the training to increase time for their topic or for buy-in later when they say their topic was not sufficiently covered in the training.

Trainer-to-trainee ratios matter for anthropometry data quality. Increasing the ratio of trainers to trainees was recommended by informants as a way to improve anthropometry training. They explained that it was difficult to conduct quality anthropometry training when the ratio of trainers to trainees was low because it

limited the amount of direct assessment and feedback received by each trainee. One disadvantage was that the skills of the trainers are more difficult to assess when there are more trainers in the class. Informants reported that the ratio should be approximately 10 trainees per trainer.

> *"… I mentioned that sometimes anthropometry and hemoglobin are done by interviewers themselves. In that case, I would say you are training people somewhere from 70 to 200. I think that is the biggest challenge, that if in 1 week you have to train, or 2 weeks, all those 200 people on anthropometry and hemoglobin, that is really difficult." R10*

> *"Having relatively small groups training huge numbers of people – it's not easy to achieve a same level of quality…you wouldn't want to have 1 trainer to 30 people, that's too many. You can't adequately supervise and make sure the quality is good enough." R15*

> *"We need a lot of trainers, especially for anthropometry or anemia and so on, where you have to practice and look exactly at what they are doing." R27*

The pretest, which is used to replicate survey conditions and test procedures, is increasingly becoming an effective training-of-trainers tool. Informants noted that including local trainers built local capacity and energized the trainings. Informants also explained that one of the main challenges was convincing the implementing agency to commit to using the pretest participants as facilitators during the main training.

> *"In the past, I don't think that the pretest trainees took the training very seriously. But maybe in the past 3 years, 4 years, where the DHS has actually structured the pretest and made it more into a training of trainers, then I think now it is more useful." R05*

In addition to having sufficient local trainers, there were concerns that there were not enough staff from The DHS Program participating in the biomarker trainings and that this reduced the quality of the training. Other informants who conducted smaller surveys reported sending more international experts to their training sessions.

> *"I'd say we'd never send folks alone. Our standard [non-DHS organization] is at least to have [3] people for training, and that's just the external people, and then utilizing local folks who have done it before too." R15*

Different types of training techniques have been found to be effective. At the beginning of a training session, the Ministry of Health, other local nutrition staff, and the biomarker trainers often provide the public health significance rationale for collecting anthropometry data. Several informants emphasized that understanding how anthropometry data are used to improve health was a powerful motivator for the measurers.

> *"I think you really have to get buy-in from the interviewers and health investigators...by talking about how important it is for kids' health… and really making it clear to the measurers getting very, very exact quantifications is so important from a health point of view because unless the people who are doing measurements buy into that, they're probably not going to do as good a job, so it's not just the technical aspects of how to do it but the buy-in as well." R06*

> *"[There is] an instructional piece that has to happen for people to be motivated to collect high quality data." R03*

> *"...explaining the why, not just the how... why is nutrition important, and how does this relate to your data?" R15*

Informants reported that the training sessions incorporated established principles of adult learning. For example, the trainers demonstrated correct techniques and explained how poor technique results in inaccurate measurements. Lifesize dolls or mannequins (6 months, 1 year, 2 years, 3 and 5 year size dolls) have been used prior to taking measurements on people. Other instructional strategies included friendly competitions among groups of trainees, quizzes that assessed learning, and video recordings of the trainees' technique, which provided important feedback that help trainees improve their technique.

> *"…we have really made progress in terms of adult learning principles and it's really been reflected throughout now almost all the technical areas, people staff have really embraced it, and even staff meetings and partner's meetings, you do see a lot of adult learning principles applied, and really it has been proved to tremendously in a lot of countries." R25*

> *"…the other tip is videotaping or photographing actual measurements and then showing that. So a combination of role-playing but also using technology... And then a lot of quizzing along the way for the knowledge side of things, like daily tests, like daily homework, pretty vigorous." R15*

Evening sessions have been held in some surveys. This type of schedule works well when the trainees and trainers are located at the same site. However, evening sessions can be a challenge when there is a limited number of trainers

Practice with children. Rapid growth during childhood requires measurements to be exact, although exact measurement of children is difficult. There was a consensus among informants that there is no substitute for hands-on practice of anthropometry measurement with age-appropriate children. A standard part of the DHS survey training manual recommends practice measuring real children. At the end of the training, the field practice also provides an important opportunity to practice all procedures in order, such as engaging with the household, administering the questionnaire, and taking measurements.

> *"I mean I'd say anthropometry training is the hardest thing in the world. It's really hard to do right, and there are a lot of mistakes. It's just practice, practice, practice." R15*

> *"Anthropometry requires a whole different training and of course equipment...you're down there and doing something physical and explaining it in a different way." R11*

> *"Getting people out there and practicing the measures…they have a sense of what the difficulties are going to be, what the challenges are going be, and being able to troubleshoot overcoming them. So I think just getting some hands-on opportunities…even if it's not formal data collection, just opportunities to see the techniques implemented." R03*

Informants reported varying degrees of success ensuring that there were sufficient practice opportunities on children across surveys. The biggest deficiency was not having enough children under age two for practice, which was noted as the age of children where health technicians have the most difficulty.

> *"…for anthropometry, definitely you have to bring in the age that you're going to actually measure." R15*

Informants said in the DHS surveys, practice is most commonly done in health facilities. In a few surveys, trainees also receive additional practice at nearby kindergartens, or at the training venue. Some informants described that bringing children to the training site for practice worked well, although others felt bringing children to the training site was logistically difficult. Others reported using clinics or orphanages. Gaining access to on-site facilities with children presented challenges, such as securing permission for the training.

"Make sure that they secure all the permission at the on-sites facilities and will allow us to access those young kids. That is crucial. Because if we don't, it happens in so many countries, when the relationship fell apart in the field and they kick us out from their facilities, and then we couldn't finish our training. So that relationship is important." R01

"They do trainings, so the clinics were not surprised by the request and they allowed us to ask for volunteers." R23

"It's not just classroom…when we visit the facility and work with real children, that's the best. Correction from supervisors and trainers, and the more time we spend in actual training, actual hands-on, I think that's the most important…" R01

Recommendation 4: Increase individualized anthropometry training.

Steps for consideration by The DHS Program include:

- Increasing the duration of anthropometry training to six days to provide more opportunities for practice and anthropometry standardizations (see recommendation 5).
- Documenting the benefits and challenges of different training schedules, such as daily alternating of anthropometry and blood collection (if blood collection is part of the survey).
- Ensuring no more than 10 primary measurers are included in each breakout session for hands-on practice, and planning for enough trained facilitators for each breakout room.
- Hosting a minimum of three evening clinics to provide individualized training to trainees identified as having poor technique.
- Recruiting enough age-appropriate children to practice both recumbent and standing measurements.

Standardizing anthropometry measurements using children is central to data quality. Nearly all informants emphasized that the child anthropometry standardization exercises were one of the most important data-quality measures that can minimize measurement error and ensure that measurers take accurate, precise anthropometric measurements. Anthropometry standardization exercises are included in The DHS Program Biomarker Training Manual and in the training protocol, although anthropometry standardization exercises are not always completed in every survey. Barriers include budget constraints, a lack of implementation agency experience and buy-in, and the logistical challenges in recruiting children. Throughout the interviews, informants provided suggestions for overcoming these barriers, and they stressed the importance of correctly performing an anthropometry standardization exercise as the key to achieving quality data.

> *"[The standardization exercise], it's the key, because we have the data for the standardization training to compare the precision and accuracy for all the targeted age groups…particularly kids." R16*

> *"It is quite hard to do on the scale of the DHS surveys are done on, but we have done it." R13*

One informant noted that the anthropometry standardization exercise has improved anthropometric data quality in DHS surveys, even with a large number of measurers being trained. An informant cited a report[2] stating that the anthropometry standardization exercise was not performed in one survey because there was no line item in the budget. However, in the next survey in the same country, data quality improved and this was attributed to trainees participating in a standardization exercise.

Two informants raised doubts about the necessity of an anthropometry standardization exercise. For one of these informants, this doubt was related to a lack of rigor in other aspects of the training or in the anthropometry standardization exercise itself. The informant explained that poor practices are standardized when anthropometry standardization exercises are performed incorrectly. This informant emphasized the opportunity for hands-on practice measuring children.

> *"I just want to say standardization is a massive topic of which not everyone agrees to its utility…Is there agreement in the nutrition community that it's needed and has utility?" R06*

> *"Some [DHS] surveys have been criticized for standardization but I think that's a rather naïve view from the standpoint of what you care about is the end result of the quality and not whether standardization [exercise] was done. The argument would be, I guess, that standardization will improve quality. But I think there are things far more important than whether standardization [exercise] was done." R04*

Many informants found that before conducting the anthropometry standardization exercise on children, doing a practice anthropometry standardization exercise on adults was useful. This provided measurers with the opportunity to perfect their technique and understand the anthropometry standardization exercise procedures prior to participating in the anthropometry standardization exercise with children, which can be much more chaotic.

> *"The good thing about some of the DHS surveys is very often you have adult measurements. So therefore, we take standing height of adults, first …Then, they get used to body positioning, proper head position and*

[2] Source provided by key informant: Leidman et al. "Assessment of Anthropometric Data Following Investments to Ensure Quality: Kenya Demographic and Health Surveys Case Study, 2008 to 2009 and 2014.: *Food and Nutrition Bulletin*. 1-14: 2018.

coordinate the Frankfurt plane, feet position according to the mid-axial line. They get that concept down, and it's easier if you're with adults who are going to be more cooperative." R02

Informants stressed the importance of following correct anthropometry standardization exercise procedures. The first step in this process is recruiting the participants. Although adult anthropometry standardization exercises were reported to be a useful initial tool for practice, they do not present sufficient opportunity to navigate the unique requirements and challenges of measuring children. However, there were examples where the anthropometry standardization exercise was only completed with adults.

Informants also described differences in the implementation of the anthropometry standardization exercise, and they reported that the exercise has not always been designed or performed with the same level of rigor across surveys.

"The standardization we did just with adults, so we didn't end up doing standardization with kiddies now you mention it, which we could have done better." R23

"It's a very painful process. Actually the more we do it, the more we realize that it's really really difficult to do, and difficult to do well." R31

Many informants reported that anthropometry standardization exercises were not completed because it is difficult to access children for height and weight measurement. However, other informants explained that with advance preparation, it is possible to understand country rules and procedures, and to gather sufficient children for practice within those guidelines. Some informants cited the importance of buy-in and engaged implementing agencies in recruiting children, while others suggested involving the Ministry of Health in this process if it is not the implementing agency.

"There are times when it's just word of mouth, it's okay, and everything is fine. People sprinkle holy water on it, and it's done. There are other times…where we had to have written consent." R02

"I find people often give short shrift [to the standardization exercise], and that they don't plan it well.... So sometimes the whole thing falls apart and you have to go looking for new kids. So I think that one really has to be planned well ahead of time, and make sure everything is on paper, the agreement is on paper." R06

"I feel [the implementing agency] are not making an effort. It's like they don't see the importance of getting the children. To me they are not making an effort, I don't know if it's because they [do not] understand." FGD

"In the [country] DHS, the advantage of having the MoH as the implementing agent was that they knew where to get children for nutrition…In [this country], they had a [marketing event] and they had huge numbers of kids come to be tested for anthropometry and anemia. But in [another country], I suggested an event and I was told it was illegal and maybe they didn't want to do it. There is not always buy-in." FGD

Informants emphasized that recruiting children for the anthropometry standardization exercise is different from recruiting children for practice. The anthropometry standardization exercise requires repeat measurements by measurers on 10 children. These children should be younger than age 5, with half the children less than age 2 to allow for standardization on both height and length. Over-recruiting children is necessary to select the appropriate age composition. After an equal number of children over and under age 2 has been identified, some of the remaining children should be asked to stay for the entire exercise and possibly serve as a replacement if needed when some caregivers decide to leave early. The importance of providing prebudgeted incentives for caregivers was emphasized because this exercise requires a substantial time commitment from the children and their caregivers. This commitment is not the equivalent of one or

two measurements in the field, because it requires many measurements on the same children over many hours and families should be compensated for their time.

> *"So, as you know, the exercise is very tedious, so although we need minimum 10, you'd recommend maybe recruiting 15 or even 20, because sometimes children do not accept being measured at all and also sometimes mothers bring children who are too old." R10*

> *"Where are we going to find the children? Are we going to find enough younger babies, and to do both recumbent length and standing height? Especially around length measurements? You really have to work through the details on that, and it's not easy, so when we work with other collaborators that haven't surveyed, firms that haven't done that type of standardization, we find that it takes a lot of back and forth for them to say, 'Okay, this is what we mean by standardization, this is what it requires, you're going to have to find that many moms with that many babies of that month/age to be available for that much time for us.'" R31*

> *"We've tried a number of times in very different contexts, … generally orphanages are not recommended. The reason is certain children will be chosen, others will not. So why are those children chosen to be in this exercise, and they just want to go play with their peers… It's also, in orphanages, there are no parents, so it's difficult to find an adult per each child to stay with that child and take care of that child throughout the exercise...also in hospital settings we had very small rooms…space was quite a limiting factor so I would avoid health centers." R10*

> *"It's not that simple, to organize that and to make sure that those 20 kids are going to stay that period of time. This is tough to do." R10*

> *"[For an incentive], usually a drink for the mother and some sort of snack, or sometimes we've given older kids maybe a pencil or a little bracelet, like a little rubber bracelet that kids like these days." R15*

Informants acknowledged that the sheer number of measurers in surveys as large as the DHS surveys can be challenging to standardize, although there are some successful examples of standardizing hundreds of measurers. One solution is to keep the number of measurers to a minimum (as mentioned in the design phase: team composition section above). In addition, only the primary measurer must be standardized. Informants advised that there must be precautions that prevent the assistant from taking measurements in the field. Some ways to prevent this include having only one set of equipment per team, and organizing the fieldwork in such a way that taking shortcuts would not substantially reduce the team's workload or time. Another proposed solution was providing sufficient funds and human resources to conduct multiple anthropometry standardization exercises simultaneously.

> *"It is quite hard to do on the scale of the DHS surveys are done on, but we have done it." R10*

> *"Think about how many times the kid is being measured... If you have to organize multiple sets of kids to come, then do it. Otherwise you just end up with a giant mess. No kid wants to be measured 60 times." R15*

> *"It generally is easier to organize more exercises than to have so many teams with the same children. It's just a lot of stress on the children." R10*

Critical to conducting an anthropometry standardization exercise is providing a controlled environment, having the appropriate equipment and adequate room setup, and the availability of participants. Given the large number of children needed and the extended amount of time required to perform the exercise (about half a day), informants suggested that the exercise should be conducted at the training site rather than in a clinic or orphanage where the caregiver-child pairs come for reasons other than for the anthropometry standardization exercise. Minimizing the discomfort of caregivers and their children was emphasized by the informants, who suggested ensuring that children have enough space to play, toilet facilities are accessible, and toys, beverages, and food are provided.

The informants expressed diverse views about the individual or group responsible for the logistical arrangements needed for the anthropometry standardization exercise. To ensure effective anthropometry standardization, informants emphasized the need for better communication among the survey manager, implementing agency, and anthropometry trainer, because the outcome is the responsibility of all parties.

> *"I don't think it's an anthropometrist problem. It's more like the planners or partners." R01*

> *"It's the most chaotic thing you'll ever do, so be prepared for chaos and then just having things organized. I think whoever's leading the standardization needs to be just really organized… have each child labeled with a number, have all the sheets where standardization is going to be done, have all that prepped in advance, otherwise it's a mess." R15*

The anthropometry standardization exercise involves a test-retest to assess the trainee measurers' degree of precision and accuracy. This requires that all children must be measured by the trainee and an expert, who is a gold standard anthropometrist. Precision is assessed by the differences between repeated measurements taken by the same measurer, and accuracy is standardized against the expert anthropometrist. Sharing the results of the anthropometry standardization exercise with the trainee measurers was viewed as a very useful learning tool.

> *"I was in [country], sat down at a meeting…[and someone] in charge of the office of health… said to us, just me, 'Well, I understand … that once you do a standardization test in anthropometry, that you can eliminate 50 percent of your enumerators right away.' And I said, 'Excuse me, the answer is, it depends.' And the reason is very clear. The standardization testing is very often an excellent training tool." R02*

> *"But one of them gets a reading, a recording error: instead of 68.5 centimeters, he measured 78.5 centimeters, so that 10-centimeter error… So that clearly was a recording error. So he failed the test very badly. Compare that to someone who also fails the test badly, but gets the same technical error of measurement score, but always measures consistently each person higher. So the sources of error of those two people who got the same results, they both failed; they're very different. One of them needs training on reading and recording measurements, the other needs training on technique." R02*

Informants emphasized that trainers should undergo the anthropometry standardization exercise to document their ability to serve as the gold standard during the host country survey anthropometry standardization exercises. In DHS surveys, the expert is usually the anthropometry trainer. Many trainers can perform measurements well but there is no process in place to confirm the ability of the trainer to serve in this leadership role. This creates challenges when presenting the results to the trainee measurers if the trainer's measurements are substantially different from the pool of participants. Without clear documentation that confirms the trainer's skill, it is unclear if the issue is the trainee measurers' accuracy or another issue with the trainer. In these circumstances, informants reported using an average of the trainee measurers as the gold standard for accuracy, which is a less optimal situation than comparing the results to a gold standard trainer.

> *"You're comparing their results, not against the group mean or the group median, but against the 'expert', the trainer." R02*

Observation from a few surveys found that one-third to one-half of measurers failed to pass the first anthropometry standardization exercise, and that sufficient time was not always allocated for restandardization. Informants emphasized that retraining alone is not sufficient, and that a comprehensive plan is required that includes anthropometric training to improve techniques and a re-standardization exercise for measurers who do not pass the first anthropometry standardization exercise. The results of the

anthropometry standardization exercise should be used to select anthropometrists for the field. Informants also suggested selecting the high-performing trainees to assist in the further retraining of the poorer-performing trainees and to do the field monitoring.

> *"One of the reasons for doing it early…is if you do have to repeat the task with some people who did fail." R02*

The DHS Program developed an Excel tool to calculate the results of the anthropometry standardization exercise, but this has been underutilized. The tool requires improvements to make it user-friendly, including the ability to more easily modify the default setting of the 10 primary measurers. They suggested further training on the tool across The DHS Program. The tool also does not provide summary statistics on whether a measurer has passed the standardization exercise. This is usually represented by the technical error of measurement (TEM), which informants suggested be added to the tool. There is not a consensus on what is considered a passable TEM, although work is being done in this area. In the interim, use of the WHO growth standards was suggested.

> *"And then after the test is completed, the results of the test are presented back to the team so that they can see the precisions they have achieved, or each anthropometric measurements so that they can see, for example, that something like weight is much easier to measure with high precision. They can see that generally they did poor on height and they can see how they did relative to their peer. They also, through that, sense that we're paying attention to the quality of their measurements, and use that as a chance to reemphasize the importance of this." R10*

> *"Graphs are hard to interpret for survey managers. When you show the technicians the graphs [from standardization], they look nice, but how do you provide individual feedback?" FGD*

Recommendation 5: All primary anthropometry measurers should participate in the child anthropometry standardization exercise and restandardization if their performance is poor.

Steps for consideration by The DHS Program include:

- Developing checklists for The DHS Program staff on the standard operating procedures for child anthropometry exercise processes.
- Planning early for the child anthropometry standardization and restandardization exercises, including written communications with the implementing agency.
- Providing transport allowances for adults and children to travel to the training venue for practice sessions rather than trainees traveling to clinics.
- Adopting a definition of what constitutes passing an anthropometry standardization exercise by using the technical error of measurement values established by the WHO/UNICEF anthropometry data-quality working group.
- Ensuring that passing the anthropometry standardization exercise is a selection criterion for being hired as a primary measurer.
- Requiring all anthropometry trainings to have at least one trainer who has documentation of having passed an anthropometry standardization exercise in the previous two years (The DHS Program staff, consultants, survey managers).
- Revising the anthropometry standardization tool to provide passing/failing criteria, allow for easy adaptation, and produce individualized retraining plans.

- Increasing awareness and use of the anthropometry standardization tool, such as branding of the tool and developing training materials on the tool's use.
- Requiring anthropometry standardization tool results to be maintained as part of the survey files in SharePoint, with summary results presented in the survey's final report.

Data-Collection Phase

Different measurement techniques were proposed, although their feasibility and accuracy need to be piloted. There were multiple perspectives on the optimal technique for measuring height, and for balancing the optimal measurement technique with field realities. Informants stated that taking a single measurement was the easiest and most practical, although this was prone to error. Multiple measurements are more difficult to obtain, although there is no consistent method. One method is obtaining three measurements before the child is removed from the board, and using an average or median of the three measurements. This may reduce errors in reading the measuring tape, positioning of the child, and transcription. Others thought that by using this approach, measurements might become less accurate because the child became more agitated, and this could reduce accuracy.

> *"Something like length … we definitely take multiple measurements. …we would get a child into position, we would get a measurement, we would record that measurement while letting the child relax, and then, not removing them from the length board or height board…we would reposition, get another measurement, record it. Same thing: get the third measurement, record it. And by doing so, I mean we really don't leave it to the enumerator to decide if one of those measurements is incorrect, but basically with 3 measurements, then we have the ability to look at the data. If one of the measurements is really out of whack, it gets dropped. Otherwise, depending on who's using the data, we would take a mean or use the median value…because we're taking a median of 3 measures we think our central measure is going to be a reasonable estimate of what's real in a child." R03*

> *"I think that for a very good quiet child, no problem, you could measure two times, three times. For a struggling, crying baby, I think if you can't do a good measurement first time, you can't measure the second time either. So it doesn't matter how many times you measure. So to me, so that's why we in our survey we don't recommend measuring twice…we don't record twice. If you think your first measurement [is]too much off, you can remeasure, but we only record one time." R16*

Another method involved the same measurer taking two independent measurements and taking a third measurement if the difference was outside of an acceptable range. There was some concern that measurers would not follow this process, and that there could be a lack of independence between the measurements. One safeguard against cheating would involve requiring the measurer to indicate the reason for not completing three measurements. Another informant suggested having a time stamp for the amount of time taken to perform the measurements. Still another measurement approach involved two independent measurers taking measurements in the field and taking a third if the differences is outside an acceptable limit (by one of the first measurers or by a supervisor). This approach would likely result in the most accurate data, although it would also be the most resource intensive.

> *"Our [organization outside of The DHS Program] counterparts in [a country] recently were involved in a nutrition survey where they were taking two height measures. And then if there were too many, too far apart, reconciling them, but I'm not sure how realistic that really is…it's just additional time being spent on it, and sometimes the amount of time they're in the households, it is an issue for budgeting, for even movement of the entire team." R11*

> *"Normally most of our surveys we measure twice, but then if there's a certain difference between the 2 measures, then the supervisor kicks in to do the third. [Organization outside of The DHS Program]" R31*

> *"The interviewers…accompany the woman [and the child] to the first measurement station where there is anthropometry and they do measure height and weight, and then the interviewer enters this into the tablet, and the interviewer brings in the woman and the child to the second measurement team, separate team, they do the measurement, and enters the data. If there is a discrepancy … the tablet automatically tells you need to do a remeasurement, it will not let you go on. [Organization outside of The DHS Program]" R28*

Recommendation 6: Explore the use of refined measurement approaches to improve precision and accuracy of measurements.

Steps for consideration by The DHS Program include:

- Finding existing datasets (e.g., non-DHS surveys, research studies) that use different approaches to measure length/height and analyzing the data to inform the optimal measurement approach.
- Piloting measurement approaches in The DHS Program to identify the approach that will result in the most precise and accurate data, including measuring length/height:
    a. three times while the child is still in position
    b. two times (each time removing child from board) by the same measurer and obtaining a third measurement if the difference between measurements is beyond the maximum allowed differences
    c. two primary measurers taking and recording measurements independently, comparing their values, and remeasuring the child if there is inconsistency between the measurers.

Heavy clothing and elaborate hairstyles can obstruct measurements, although the extent is unknown. Several informants described how hairstyles and clothing interfere with taking accurate height and weight measurements. In The DHS Program, standard procedures include asking respondents to remove their shoes, undressing a child to their underwear (additional procedures exist for taking weight measurements of children wrapped in blankets), and unbraiding or pushing aside any hair that might interfere with the height measurement. However, informants reported that in the field, the practicalities remain a challenge.

> *"We had a lot of issues with hair in [country] there were very, very elaborate hairstyles." R15*

Informants said that within The DHS Program, there is no consensus from an ethical standpoint about how much clothing should be removed (removal of diapers or removal of clothes when a child is not wearing underwear) before weighing children and at what age. There is a lack of resolution about the practicalities of demanding removal of a child's clothing.

> *"I do remove most of the clothes, but I don't know that they should be stripped down to almost being naked." R15*

Some informants suggested that a reasonable temperature in the room and a more private setting may reduce discomfort when a child is asked to remove their clothing. Others suggested that in some country contexts, the gender of the measurer might influence the caregivers' comfort in removing a child's clothing. Others suggested noting differences in clothing or hairstyles, correcting for them, and reporting these corrections, although there was a lack of guidance on the acceptable procedures for such corrections.

> *"Another point to be considered would be also the gender composition of the survey teams given the context. So, is it appropriate that men undressed the children? Is it appropriate that women undress the children, or do they have more facility? Those are also calm conversations to be had before the survey teams go to data collection. Because that can also be hit-or-miss…how the response rate of having children being undressed (or being even measured) can be affected throughout the data collection. [Organization outside of The DHS Program]" R10*

> *"In cold climates, it's basically we have tick boxes, "Did you take off most of the clothing?" They most of the times say yes…when they don't tick the box … it's so arbitrary right? A little jersey as heavy clothing, [or] naked except for the diaper, whereas the adjustment should be bigger in a diaper than in a light jersey…I think the adjustment is really tricky. [Organization outside of The DHS Program]" R28*

> *"There was some discussion of having a pin that you kind of stick down through the hair to get down to the scalp and find out how much of an addition there was on it. I suspect that's important in places where the hairstyle significantly adds [height]. But it was stuff that was being looked at…I don't think there was anything ever published on it. [Organization outside of The DHS Program]" R18*

> *"We just do it the best we can, but we'd write it down, and document it, and then if it looks weird, we'll do something about it. [Organization outside of The DHS Program]" R15*

> *"I know that some people do the subtraction on [weight]. [Organization outside of The DHS Program]" R01*

Informants described the importance of reviewing explicit guidelines for hairstyles and clothing during training, because without guidance, issues are resolved in various ways in the field, which leads to a lack of consistency across surveys. Informants recommended greater clarity on standard procedures and improved training methods for following those procedures.

> *"It's something that we talked about in every training, but it is a problem in many countries. And it's something that needs to be discussed at length in the survey, that teams appreciate the importance of undressing the…children. It is the thing that teams are often most resistant to because it is what prompts the*

child to start crying, and if they don't appreciate the importance of it, and how undressing the child impacts the weight estimate, and how that ultimately impacts the weight-for-height score, for example." R08

"It's really important, especially in the colder climate, to be a part of the training. And having a discussion in the training about how to negotiate that. How to talk to the parents about going to a quieter room, a quieter room in the house and doing the measurement without other children observing or having a discussion about, you know, maybe the child doesn't need to be undressed for height." R10

Recommendation 7: Develop approaches to reduce systematic overestimation of measurements because of clothing and hairstyles.

Steps for consideration by The DHS Program include:

- Clarifying guidance on the amount of underclothing a child should be wearing to maintain privacy but still facilitate accurate measurements.
- Conducting simulation models on the effects of clothing and hairstyles on height and weight measurements and confirm these in pilot studies.
- Adding a question to the biomarker questionnaire that indicates when heavier clothes were not removed, and if hairstyles were not undone.
- Testing an approach to subtract the weight of heavy clothing and tall hairstyles during the data analysis phase.

Innovations are needed to reduce transcription error. Despite procedures that minimize transcription errors, some informants identified challenges measurers face when reading and recording values from the length/height boards and scales. Some informants cited problems with recording values in the paper biomarker questionnaire. Informants described trying to minimize this issue with improved training, while others suggested incorporating technology that would reduce human errors.

> *"The ideal situation is an electronic download…[results] really should electronically go somewhere…rather than some transcription because any transcription is obviously prone to human error." R09*

Informants were not familiar with anthropometry equipment equipped with the technology that can avoid transcription errors, but they were interested in the development of such technology.

> *"And the idea is to put together a prototype of a board which has everything you would want to have in the board including: durable; reasonably affordable; accurate; possibly using laser technology; storing the results so that they can always be checked if they're separately recorded; and all kinds of other things. And then putting those requirements out to a bunch of manufacturers and letting them make prototypes and then the committee will come back together and assess the different ones that they've made. And we may well have a much better product, which will be well-used for people doing weighing and measuring throughout the world." R06*

Recommendation 8: Minimize transcription error for anthropometry measurements.

Steps for consideration by The DHS Program include:

- Exploring use of length/height boards with automatic numerical reports rather than tape measures.
- Exploring Bluetooth technology that would transfer height and weight results automatically from the board and scales to tablets.
- Building and testing a CAPI program for the biomarker questionnaire, which includes double data entry of the anthropometry measurements while still in the household.

Equipment calibration is an ongoing process. Informants discussed the importance of calibrating anthropometry data-collection equipment to minimize measurement error. The two main pieces of anthropometry data-collection equipment are the ShorrBoard and SECA digital scale. A standard weight (5 kg weight) is supposed to be used to calibrate the scale, a 110 cm metal rod is supposed to be used to calibrate the ShorrBoard, and the results of the measurements are recorded.

Many informants thought calibration should be done before fieldwork and daily during fieldwork. The DHS Program manual includes instructions for calibrating boards and scales prior to fieldwork, and recommends (but does not require) measurers to calibrate scales daily during fieldwork.

> *"Recently we've had a lot more trouble with the scales. In [this country] we were having a lot of difficulties with them, even though they were brand new. So we had the standard weight and we found that every time we moved it, we were getting different results, so we were using standard weights. And I actually think I prefer it now, I feel much more confident that the measures are more accurate using the weights." R15*

> *"We calibrate our scales before going out into the field by checking them against standard weights." R03*

Some informants noted that in practice, implementation of the recommendation varies, including procuring the calibration equipment and the calibration prior to fieldwork and during fieldwork. However, informants described their willingness to adopt optimal practices. For example, in one recent DHS survey, the local implementation agency sourced hard PVC plastic rods for the ShorrBoard calibration and used the ShorrBoard as the weight for the scales. They also introduced a daily calibration and maintenance log sheet to record the results, which supervisors were required to check on a routine basis.

> *"We did a lot more quality control of a calibration of the anthropometry equipment, and I think some of the counterparts weren't as prepared for that level quality control, but I think there's definitely a lot of willingness." R15*

> *"The one additional thing we haven't mentioned during the field tests: we were also observing and making sure they are doing the calibration of the equipment properly." R10*

During training and fieldwork, equipment that does not generate the correct results should be replaced. Backup equipment must be available so that faulty equipment can be replaced expeditiously. Informants noted that in DHS surveys, two spare scales are included in the supply list because these are custom ordered. Spare measurement boards are not included because replacements usually arrive within a week or two. However, informants described spare scales and boards as not being available for immediate replacement. Without spare equipment on hand, the survey either has to continue with faulty equipment or faces a delay at a substantial financial cost.

Recommendation 9: Strengthen calibration, maintenance, and quality assurance checks of anthropometry data-collection equipment.

Steps for consideration by The DHS Program:

- Procuring aluminum or plastic rods of two different lengths and standard weights of at least 5 kg to calibrate length/height board and scales, respectively.
- Procuring extra height boards so that faulty or broken equipment can be replaced without delaying data collection.

- Training all measurers on the calibration of equipment using aluminum or plastic rods and standard weights.
- Calibrating the height board and scales using aluminum or plastic rods and standard weights prior to fieldwork and daily during fieldwork.
- Developing standard maintenance logs to be completed by measurers.

Anthropometry supervision takes time. The importance of continual supervision was discussed by the informants, who described two levels of supervision. The first is the team supervisor, who is assigned to each data-collection team and who manages the daily team activities and quality checks. The second level is a field coordinator, who rotates among teams to provide supervision.

> *"If you pass an [anthropometry] standardization test today, doesn't mean you're going to measure well tomorrow." R10*

However, the team supervisors have a heavy workload that is primarily focused on logistics and observing interviews, which limits time for supervising anthropometry measurements. In a few instances, surveys have had more than one supervisor at the team level. This has allowed for roles and responsibilities to be shared across personnel. Although this model may become more common in some settings, especially as the responsibilities associated with electronic data collection expand, this approach will remain cost prohibitive in many surveys. Similarly, the same constraints were cited for the field coordinator.

> *"So the team leader is basically, for me that's less of a supervisory role than of the team management role. A bit of all the admin hurdles, getting the local authorities to agree…schedule households for interviews. And yes, [he/she] does go from different team members, check a little bit on how they are doing. The important part of the team lead position when it comes to quality control is checking the biological forms in the evening, the paper-based, the biological forms, and going through some of the checking…where it's been correctly entered." R28*

> *"I don't know if for example the [field coordinator] who comes to supervise the interviewers will have enough time actually to observe everything, that's why we're trying to build the roving team that will have at least 1 female interviewer supervisor and 1 male, and also these people will have the competency to look at the biomarkers or have additional biomarker [supervisors]." R27*

Supervision of anthropometry data collection requires different skill sets from interviews. Many informants raised a concern that team supervisors and field coordinators lacked the technical knowledge to provide meaningful oversight for anthropometry data collection. In some cases, informants reported receiving incorrect supervision, which was more detrimental than no supervision. In some DHS surveys, the high-performing biomarker trainees have been selected as anthropometry supervisors. The informants explained that this approach was successful because the trainees followed The DHS Program methodology for anthropometry.

> *"I don't think that the [team] supervisor is an ideal person at all, unless the supervisor is strong in biomarker measurement testing. I think someone who has gone through the process of biomarker measuring and testing, who understands what to look for, who has been in the field, is the ideal person." R05*

> *"One challenge I see is sometimes is where people who haven't been trained, who are more senior people, may be in the steering committee, start coming in and doing a little field supervision. But I feel like that's always a really big risk. Because sometimes they want to make changes, or they have suggestions and it goes against the training protocol, but they have some sort of seniority perception among team members." R15*

> *"[We] created another level of supervision that I called anthropometry supervisors. We took about 7 or 8 of the best anthropometrists, made them supervisors, which again was a bit difficult because when you take the best from a group and they're equals and someone is your friend is being elevated above you, that's sensitive and has to be handled carefully. But nonetheless, that worked very well, and then we had 30 other people. So those 8 people were in charge of about 3-4 teams each, and therefore, the anthropometry trainer spoke to and checked with the 8 anthropometry supervisors, rather than 50 people individually, trying to follow 50 people in the field. So the different levels of supervision, this anthropometry supervisor tier, was extremely*

effective, [but it] very rarely can be done because of lack of funding and the time taken to implement training and implementation of that model." R02

Training supervisors is a critical to ensure supervision is correct. Informants stressed the importance of supervisors receiving adequate training in anthropometry measurement if they are to provide oversight on correct measurement techniques. Despite the recognition of the need for increased training, there are several constraints. In The DHS Program, the interviewer and biomarker trainings run in parallel. Informants explained that the team supervisors attend the interviewer training but do not participate in the anthropometry training. In a few surveys in The DHS Program, this issue has been avoided by scheduling sequential trainings for the biomarker team and the interviewers.

"Supervision is one of the biggest challenges. The implementing agency is used to monitoring fieldwork, but mostly they are used to just going and making sure the teams are in the correct place. I think it's an area that we can strengthen. It means that supervisors need to be aware of the correct procedure, but often there isn't time to provide that [training]." FDG

Several informants expressed frustration that in some surveys, the field coordinators only partially attended or did not fully engage in the anthropometry training. A few informants from outside The DHS Program also said that it was important for supervisors to complete anthropometry standardization in order to supervise others. However, this was not a majority view.

"[Supervisors] need to have gone through the same processes of standardization. If you don't know the details of how to actually do that, you're not going to be able to supervise that work." R31

Take advantage of critical windows for supervision. The timing of supervision is critical to providing remedial action. The informants noted that this is especially true during the first two weeks of field monitoring because early supervision allows for incorrect procedures in the field to be identified and corrected, because many measurers may never have worked in the field, and because intervention should take place as early as possible.

"We first went out to the field, for us to be there supervising the teams closely with the anthropometry. Because we went through the standardization, we go through the whole training, and then they get out in the field, and maybe sometimes they start kind of loosening up how they're doing it. So, to be right there to be able to correct it, any issues, especially early on, I thought, was really helpful." R15

Increasing the time the technical staff from The DHS Program (full-time or consultants) are present during the first few weeks was identified as an issue. This can be expensive and taxing on technical staff who have just completed training, although it provides an opportunity to further mentor the supervisors.

"…we can elaborate on what are the critical points, where you need to insist, issues that could come up, how can you resolve them … [then the trainer], basically just sits back and observes the supervisors, what they do, how they do their job, and then you give the feedback back to them and of course also to the teams, and we let them go." R28.

Use of supervision tools promotes the collection of high-quality data. Supervision tools that the informants considered effective included checklists, remeasurements by anthropometrist monitors, and WhatsApp groups. Some informants discussed using checklists during field monitoring. The checklists serve as a reminder for staff to conduct supervision and highlight critical areas to observe during anthropometry, and they are useful as a job aid to measure performance of field staff collecting anthropometry data. Checklists remind supervisors to visit teams and provide feedback. The checklists can also be used to assess if the

supervisors were visiting the field teams. Other informants thought there should be a rotating anthropometrist who takes measurements and compares them to the measurers. Something similar is sometimes done with interviews where a percentage of interviewers (or partial interviews) are conducted to confirm that interviewers are administering the questionnaire correctly. Other informants suggested including video and other technologies to ensure supervision visits were being conducted. Lastly, WhatsApp groups were used by supervisors to disseminate course correction information, especially for common errors, to all measurers.

> *"You oversee some percentage of the interviews that are taking place with sort of a checklist of what the ideal anthropometric measure is supposed to look like, so you can be certain over the course of time that the anthropometry training that took place at day one is being implemented six weeks down the line to the same degree as was accomplished in the training activity." R03*

> *"Creating a massive WhatsApp, that's really, really helped because then you find one mistake that a team is making, you can actually alert everybody and even send a video or a picture or whatever to say, 'Team Five was making this mistake, everybody else don't make the same kind of mistake! Do it like this.' That's actually reaching out to the teams directly." R33*

Responsive actions to supervision. Informants suggested remedial action after several supervisory visits to teams that needed to improve anthropometry data collection. Some informants emphasized the need for periodic refresher trainings of measurers at the halfway point or even earlier in data collection, depending on the problems observed during supervision. They suggested that measurers can travel to a centralized location for retraining and restandardization to improve data-collection techniques. Informants explained that this travel and training provide incentives for measurers to maintain their skill sets, as well as an opportunity for hands-on supervision before the completion of data collection. However, informants acknowledged that this was costly and would have budget implications that needed to be included during the survey design stage.

> *"… bring back teams, or if they're spread all over the country, maybe 2 or 3 groups you stop in different phases. Retraining sounds like 2 weeks, but I literally mean you make them go through the measurement, height and weight again. Are they still measuring correctly? A mini-standardization or whatever, something done at that stage. I mean these are just ideas, everything has behind it a cost, and you could do it in some countries but not all countries because we don't always have the kind of control you need to change systems in that fashion." R21*

> *"You need to pay them for all the days that they're coming in from the field …like when you're dealing with 200 person-days, the costs add up incredibly quickly. You need somewhere to meet with them, so you have to pay for a venue. Then you have transportation costs. So it's a very expensive thing, and we don't have contingency lines in our budgets, and everything that we're going to do has to be budgeted and paid for in advance…" R05*

One action of last resort was the replacement of measurers. Some informants noted there can be challenges with the implementing agency when there are proposals for removing and replacing the underperforming measurers. In some surveys, informants reported that after measurers have been replaced, replacements do not always obtain a new worker ID, which in turn can impede tracking of the replacements' performance.

> *"[We] suggested several times that they should fire several teams, period. And of course, the implementing organization refused to do it." R27*

Recommendation 10: Enhance supportive supervision for anthropometry data collection.

Steps for consideration by The DHS Program include:

- Requiring individuals who supervise anthropometry data collection to receive adequate training and pass a test on correct anthropometry measurement procedures.
- Implementing the use of a supervision biomarker checklist.
- Identifying rotating biomarker monitors who are trained to provide oversight for anthropometry data collection.
- Piloting remeasurement of anthropometry data on a subset of children by the anthropometry monitors.
- Exploring the use of video and other technologies to ensure anthropometry data-collection supervision is taking place.

Innovative quality assurance measures have great potential. Use of new technologies provides an opportunity for real-time data correction and quality assessment. Informants explained that the recent shift to the CAPI system allows for the ability to compute anthropometry z-scores in real time. The data can be corrected while data collection is in progress, rather than waiting for large amounts of data to accumulate before detecting problems. This mitigates the challenges of not being able to be on the ground in every part of the country at one time, and also provides insights into data-collection procedures in remote and insecure environments.

> *"How much accumulation do you need before you can see that there's an interviewer or team or whatever not performing well? Now the question becomes, by the time the information gets to them [field teams] do they act on it, and then are we seeing the results?" R21*

> *"What do you get from the eyes on supervision given that the eyes on supervision is so sporadic, and it's such a small percentage of the fieldwork that goes on?" R05*

Two main strategies were suggested by informants. The first involves correcting data *before* leaving the cluster, and the second is to obtain random repeat measurements on a subset of the population.

A few informants outside of The DHS Program reported using the first approach, in which they identified children with extreme anthropometric z-score values and remeasured them. The percentage of invalid scores is expected to decrease with the remeasurement of flagged cases, but data is limited on the extent this approach improves data quality.

> *"More recently we've been programming [the] ability to calculate behind the scenes… such that it will prompt for remeasurement if the child is outside of the WHO range for weight for height, height for age, or weight for age, and then it would ask them to remeasure all parameters." R10*

Assessment of the precision and accuracy in a subpopulation was also identified as a potentially useful quality assurance measure, although this has rarely been done other than in a study setting such the development of the WHO Growth Standards. Informants suggested collecting a second measurement by the same measurer and/or two independent measurers. Discrepancies may signal poor measurement techniques by the measurer(s). The data can be useful for assessing data quality post-survey and for providing feedback needed for re-training measurers with poor precision or accuracy. It is possible that measurements may be of better quality because the biomarker team may be more careful and not rush with measurements if they know that poor measurements will be exposed. Diurnal variations for height (growth spurts) and weight (liquid consumption and bowel movements) were not discussed at length, but may be important areas for future investigation.

Recommendation 11: Implement new continuous quality assurance measures during data collection.

Steps for consideration by The DHS Program include:

- Piloting different approaches for remeasuring children with flagged anthropometry data (e.g., selection of different z-score flagged values to prompt remeasurement, timing of remeasurements).
- Piloting the control of inter-observer and intra-observer variability during data collection through the following approaches:

    a. Central anthropometry data-collection monitors (national level biomarker supervisors) who rotate among clusters and conduct repeat anthropometry measurements on a subset of participants to assess accuracy, and
    b. Measurers who conduct random repeated anthropometry measurements on a subset of participants to assess precision.

- Determining how the results of inter-observer and intra-observer data quality can be used for supportive supervisory and retraining activities.
- Piloting a standardization exercise at the halfway point of data collection to assess if measurers are deviating from the standardized measuring techniques and to take corrective action.

Data Processing Phase

Field check data to monitor quality and promote corrective action. Several informants explained that the field check tables were useful in monitoring teams and that the CAPI system allows for "real-time" field monitoring. In The DHS Program, field check tables are run periodically by the implementing agency during data collection. Both The DHS Program and host-country staff review the tables and provide feedback to teams via the team supervisor about any issues identified for course correction during data collection. Informants found the field check tables to be useful because they provide information by cluster and team, and they allow identification of patterns and performance issues while the teams are still in the field.

> *"It used to be that we didn't know anything, and knew everything after the survey, right? We would get the field check tables, two, three, four weeks after the fact, with too few cases to say anything. I couldn't make any reactions until in the end…but now with CAPI, you get more, faster reactions. It helps us see things fast." R13*

> *"Thanks to field check tables, you can identify poor teams, so you can send someone to check on the team. Some teams have nightly debriefs and the supervisor can provide feedback." FGD*

> *"So even though the data teams are centrally based, they can feed that back to supervisors, who can then say, 'Okay Team One, you're like way off compared to other teams, you're all doing 0.0 on your heights, what's going on?" R15*

> *"If there's a lot of z-scores that are at the extremes, that might be a region, or a person, or something to follow up on in terms of is this real or is this a situation where we need to get in there and re-evaluate how the data is being collected in the first place?" R03*

Informants also said that field check tables can be overwhelming and difficult to interpret. A recently developed data-quality dashboard now provides results in a visual format that supports interpretation of the field check tables and the identification of poor performing teams.

Despite training on the interpretation of field check tables, the staff continue to find the anthropometry data in the field check tables confusing. A few informants expressed challenges in interpreting and translating results from the field check tables into specific actions. Guidance on detecting whether the data-quality issues are related to age or height was requested. In addition, the extent to which the data flow from the implementing agency back to the team supervisors was unclear. Informants suggested that a field-level dashboard could help them translate the data into action.

> *"We have noticed over the years that when you point out that this is a problem like the 0 and 5, or the way they measure the height or whatever it is a problem, after that you'll see that there are never any 0's or 5's. So the point is that you sent that message, so that's the only thing they're beginning to look out for, whereas what you really want is them understanding that, look it's being monitored, it's not one thing that we are looking for." R21*

There were divergent views about whether all necessary indicators were included for anthropometry data and if some indicators did not provide enough useful information to take corrective action. For instance, producing a curve of the distribution of the z-scores for each team was suggested as a useful indicator that is not currently being used. Some informants thought that there should be greater involvement across teams with expertise in data quality to determine what should and should not be included in the field check tables.

"Unfortunately as a [sampling statistician]…we are not really involved with checking the field check tables. This is something I believe only maybe only the survey managers do… there is no system where everybody should do stuff the same way…definitely we have to investigate what the field check tables look like and what are the mechanisms behind these missing data." R04

"Checking the data quality of measurements is a challenge [during survey using field checks tables]. What is an outlier and then how do you respond to technicians in the field?" FGD

Informants emphasized that many errors are not recognized until significant data have been accumulated, at which point the survey is well underway. In addition, the field check tables are utilized only to improve the collection of new data and are not to return to households for corrective action. Therefore, informants recommended that other measures, such as site visits and retraining, be used in conjunction with the field check tables.

Recommendation 12: Increase the use of field check data to monitor anthropometry data collection and promote corrective action.

Steps for consideration by The DHS Program include:

- Examining DHS survey data to identify the most useful data to include in the field check tables, and the appropriate stage in the data-collection process when these indicators can detect data issues.
- Revising field check tables to better identify deviations from standards and appropriately alert the need for remedial measures.
- Developing a field-level dashboard on anthropometry data and other topics, which presents data on team performance and alerts teams to the need for corrective actions.

More can be done to assess data quality after data processing, but before data is released. Data entry, data checking, data cleaning, and editing take place in the data-processing stage. Many informants explained that at this stage of the survey, data-quality issues *"only come to [their] attention if there's a huge disaster."* Several informants identified the need to diversify the responsibilities for data quality during the data processing phase not only to reduce workload, but also to involve content experts. One solution included establishing data-quality roundtables or committees.

> *"We don't have a proactive conversation that says, 'Hey, what do we think about the data quality? Are there any concerns?'" R25*

Informants also suggested that an expanded analysis of data quality for anthropometry and other data should be undertaken prior to writing the key indicator and final reports. This could be done by the analysis team and others in The DHS Program.

> *"I mean there are errors that can be caught by data processing and by cleaning up the data." R12*

> *"There have been at least some mumblings about the fact that that quality should be run internally and discussed before really any other work is done, before we spend countless drafts writing chapters that are either going to be adjusted or not used, and figure out how we're going to message those things." R25*

> *"I think it's a good idea to work out a way of having some kind of review of the data, and one of the things we could do is run the secondary editing type checks." R24*

To examine data quality, informants said that the review group needed criteria developed *a priori* to assess and discuss data quality. Informants indicated that the quality of anthropometric measurements cannot be captured with a single indicator. Data-quality indicators mentioned included the proportion of flagged data, heaping and digit preference, the mean and standard deviation of the z-scores, and the relationship between height-for-age and z-scores. Many informants thought the most informative indicator was the standard deviation, although others believed that further research was needed to better distinguish heterogeneity in the population from data-quality issues. Still others mentioned the importance of interpreting data quality in the context of other data sources from the country, previous DHS survey data from the country, and the program context. This would mean country-level involvement in data-quality assessment.

> *"[Is] the long-standing assumption of similar, standard deviation across populations a valid assumption? You know, it's not a research one, but I've looked at a simulations and said yeah, people talk about this issue of 'Well, we have a population where there's a lot of malnutrition and then another population that everything is perfectly fine,' so you get this bimodal issue going in that's going to drive a wider standard deviation, and the simulations I've done for those populations have to be really dramatic, very different to get to a point where you see a substantial impact on the total standard deviation. So, I think that that's an area to be explored." R18*

> *"Height-for-age and weight-for-age…are going to be affected by age measurements, but the weight-for-height measurement should not be, and it shows the same kinds of overdispersion. So that would suggest that overdispersion is with the height and the weight measurements...mainly the height." R12*

> *"'You're working in this field, you have programs on this, can you think of explanations for this?' And if they're saying, 'No, it doesn't make sense to us,' we say, 'Well then this is a topic that we need to look into further,' either do further analysis or do data-quality checks or whatever." R25*

Recommendation 13: Increase identification of data-quality issues.

Steps for consideration by The DHS Program include:

- Establishing a DHS survey data quality committee for each survey.
- Reviewing anthropometry data quality prior to the release of a key indicator report and the final report.
- Determining the criteria that will be used to assess anthropometry data quality.
- Creating a tabulation for anthropometry data quality.

Data Dissemination and Use Phase

Communications on data quality with countries and data users is important. Informants suggested that the data quality appendices included in the final reports should be revised into a more user-friendly format. Making quality data more accessible would allow those working with data from The DHS Program to better interpret the results, and the countries that use data to better understand the limitations of the data.

> *"We've talked about revisiting the data quality chapter in the final reports, as it is – I think a lot of people agree – …to help facilitate understanding by a more casual user, that chapter would have to be written very differently…So I think that there's room to really revise what that data quality chapter looks like." R25*

> *"I do think that it would be helpful within the whole data quality discussion at large to have some kind of internal assessment meeting ahead of dissemination, so that everybody's on the same page about what we think, how we feel about the various aspects of data quality, and so that everybody is prepared to answer questions…in the same way." R25*

Many informants expressed concern that more needs to be done to communicate and build the capacity of countries for the interpretation of data quality. There must be an appropriate balance between understanding that there will always be data limitations, while also supporting the valuable contribution to and influence of The DHS Program data on policies and programs.

> *"There's survey error around all of these things, and we know that. Some of it is at the data-collection stage, some of it is at the sampling stage; there are so many inputs. When the expectation is that we're supposed to give people a data point that's precise, we know that is never possible. Now if we're overestimating something by 50% or something, obviously that's a different issue, but I think sometimes funders and implementing agencies are critical on points… that are inherent in the survey." R25*

> *"We will contextualize when we put it out and say that, there are other studies that are not finding as high prevalence…when we put the results [out] we say this is what we've done and this is what we found, but it needs further investigation." R23*

> *"I don't think we have a proactive conversation about [data quality], and some missions are either unaware or not sort of technically wise enough to ask the right questions or to even know what it would mean if you said, 'Hey we have concerns about data quality here,' whereas others are." R25*

Transparency is important, although caution needs to be taken on what constitutes poor data quality. The best way to alert external users about poor quality data was seen as an ongoing challenge. There are several dissemination sources to consider. The most prominent are the final report, STATcompiler, and the recode microdata files. Different approaches ranged from suppressing information across all sources to providing cautionary notes. Many informants cited the difficulty of identifying a threshold at which there should be cautionary notes or the data should be suppressed.

> *"It's important to be transparent about it because anything we do that is not adequate quality…it is going to come back to bite you, right?" R13*

> *"In a couple of cases, I have gone back through the data files and found that some interviewers were responsible for disproportionate numbers of wasted children in some areas. I found one case in which there were two interviewers who were working in two clusters. The two of them found disproportionate numbers of wasted children, and after that they disappeared, presumably meaning that they had quit or been fired, and yet their suspicious data remained in the files." R12*

> *"So only in extreme cases should an indicator be suppressed, not given in a report…That has happened, that should be very rare, but I think it's appropriate to include some cautions, and even indicating the direction*

of a bias…It's almost as if I wish that it could be, some of these numbers could be given a gray shading, so that they would not be just copied into all kinds of international reports." R12

"It's very confusing if we say, 'Well, it's not in the final report but it's still in the dataset," [because] people play with the dataset. Once we release it, people take it as gospel, whether or not we try to put cautionary notes in." R12

Making recode microdata public was viewed as a strength of The DHS Program. Some informants considered the idea of applying adjustments to reduce bias, although many did not favor this approach. Since The DHS Program recode microdata and reports match, presenting recode data that is unadjusted can serve as a reference or standard. Adjustment approaches change as more data become available or new statistical methods are developed. This would result in inconsistency between recode microdata over time, or if the adjusted values are not included in the recode microdata, discrepancies between the recode microdata and figures in the reports. Either of these scenarios would create confusion. Many informants believed that the data from The DHS Program should serve as a reference value, since unadjusted recode data allows others to work with the data for many purposes, including for data-quality assessment studies. Another reason for not adjusting the data is the resulting delay in the release of the reports. More specifically to anthropometry data, there was concern that the adjustment approach currently available for anthropometry data could result in over-editing.

"Just like when people do models, a model changes as you add more information, or you refine the model, and so there's not a single adjusted value." R12

"Since [The DHS Program] was set up 35 years ago, I think the main principle was just to construct the data bank. They leave the other detailed analytical work to the data users." R04

"Depending on the kind of indicator it is, I prefer to use a model estimate rather than the [data from recode file] in most cases…I used to think it would be a good idea if once in a while, we could, DHS could, adjust its estimates. If we think there's a bias, then you adjust it. Now I'm convinced that that's not the right thing to do." R12

"…you can detect extreme cases [for anthropometry] but, there can be displacement or shifting within a valid range of values and you'll never detect those…So, you will be biasing the data potentially if you throw those cases out." R12

Informants viewed the availability of the recode microdata as a major strength of The DHS Program, given that many other surveys that collect anthropometry data are not accessible to the public. One informant mentioned there is not always sufficient information for interpretation of the data. For example, The DHS Program has used different scales between surveys and over time. Depending on the scale edition, the scale can read to the tenth or the nearest 100^{th} decimal place. During the data processing phase, the latter values are rounded during the creation of the final recode microdata. Similarly, the informants said that the inclusion of unnecessary variables (outdated approaches to calculating indicators) and not applying flags consistently across time were confusing for data users.

"Our data are publicly available to all, everyone can see everything that's wrong with the data. We include in the data files the number of cases that are out of range." R12

Recommendation 14: Increase usability of DHS survey datasets and interpretability of dissemination materials.

Steps for consideration by The DHS Program include:

- Publishing a data quality report to accompany each survey report that includes anthropometry data.
- Documenting in methodology section of final reports the processes that positively or negatively affect anthropometry data quality, including the results of the child anthropometry standardization exercise.
- Revising the weight variable in the final DHS survey datasets so that decimal places are not rounded when the scale is recorded to the 100^{th} decimal place, and data users can assess digit preference.
- Reporting the scale model number in DHS survey reports.
- Applying the correct definition of invalid anthropometry data in previous datasets, as part of an overall data cleaning process, as well as removing obsolete variables from the previous and future datasets in cases when the indicator definition has changed.
- Assessing the influence of invalid anthropometry data at the subnational level on the representativeness of anthropometry prevalence estimates.
- Revising anthropometry prevalence estimates in STATcompiler for the older surveys where invalid anthropometry data was handled differently from the current calculations.
- Establishing a process to ensure anthropometry (and other) data suppressed from reports under The DHS Program are not included in STATcompiler.

Hemoglobin

Informants identified the facilitators and barriers to data collection for both hemoglobin and anthropometry. The following recommendations apply to both:

- Involve host country stakeholders in decision making and planning for high-quality data.
- Place anthropometry and hemoglobin data quality considerations at the forefront of the survey design.
- Improve individualized training, with an emphasis on hands-on practice.
- Offer supervision that is customized to the correct collection of anthropometry and hemoglobin data.

These recommendations were described in the anthropometry section. As outlined below, there were similarities for hemoglobin, along with additional information that applied only to hemoglobin.

Design Phase

Many topics related to anthropometry and hemoglobin data collection were similar during the design phase. This included obtaining buy-in from host-country stakeholders for making decisions during the survey design phase that will promote high-quality hemoglobin data through steering committees, technical committees, and technical committee subgroups. As with the anthropometry data, there is the need for greater engagement of the technical staff from The DHS Program during survey design and additional planning meetings with staff who can address the quality issues with hemoglobin data.

Adequately budgeting for the components that are unique to biomarker collection included having dedicated team members to collect blood, establishing a biomarker-specific supervision system, over-recruiting to avoid sending health technicians to the field who are unable to perform the tasks, and promoting high performers into a supervisory role. One option for offsetting some of the increased costs was the possibility of subsampling both hemoglobin and anthropometry data. However, some considerations about hemoglobin data collection were distinct from anthropometry data collection.

Steering Committee involvement is important for buy-in but not necessarily for reviewing data quality. Informants emphasized the importance of engaging with the survey steering committee from the beginning and maintaining engagement throughout the survey process, because this fostered a better understanding of the technical aspects of blood collection. Technical committees and subcommittees are valuable additions that bring together those with specific interest and expertise in biomarker data collection. Engagement with technical committees should include issues related to the sensitivity of hemoglobin measurements, including those in the MIS.

Identifying the stakeholders associated with hemoglobin data collection in a country is a critical consideration. Because anemia is a multifactorial condition, identifying advocates can be difficult. Anemia is often the concern of those who work in nutrition, although its causes, treatment, and consequences extend beyond a single domain. The technical knowledge required to understand and alleviate the burden of anemia is broader than many other nutrition-related diseases. In-country capacity often exists for anthropometry and IYCF, but is often less robust for micronutrients and anemia. Consequently, even with strong

engagement by stakeholders for nutrition, the concerns about data quality with anemia may not be addressed adequately.

Blood is collected for purposes other than anemia. There may be more engagement around blood collection for malaria and HIV, which have more funding and often more expertise. Anemia needs a higher profile, and should undergo greater scrutiny for data-quality components that influence hemoglobin concentrations.

The MIS implementers do not engage with nutrition stakeholders on the quality of hemoglobin data. However, hemoglobin data derived from the MIS are available for use in both country and global databases. Global and country estimates are biased if these data are collected poorly, and this affects the effectiveness of recommendations for nutrition programming. Researchers, policy makers, and others who use the hemoglobin data may be unaware of the different survey sources and their potential influence on data quality. However, the informants noted that nutrition funding is not provided for MIS surveys. Anemia is not the highest priority among the malaria indicators, and only severe anemia needs to be detectable for the purposes of malaria tracking. The extra data-quality inputs needed for the data to be used for nutrition purposes may not be included in the malaria funding. At the same time, some informants suggested that in malaria-endemic settings, measurement of anemia without measuring its main causes, such as malaria, provides countries with limited guidance for decision making. Thus, limiting the collection of anemia in DHS surveys that do not typically measure malaria has a significant downside.

Recommendation 1: Involving host-country stakeholders in decision making and planning for high-quality hemoglobin data.

Steps for consideration by The DHS Program include:

- Requiring a host-country stakeholder with expertise specifically in nutrition on the steering committee and technical committee who can address the requirements for quality hemoglobin data collection (DHS and MIS surveys).
- Defining more instructive roles for members of the survey steering committee and the technical committee who provide oversight for hemoglobin data quality (DHS and MIS surveys).
- Developing materials for steering committee and technical committee members that will increase understanding of the requirements for high-quality hemoglobin data collection (DHS and MIS surveys).
- Increasing the involvement of The DHS Program technical content experts (nutrition and biomarker) at the beginning of the survey design phase.

The number of biomarkers selected at the design phase will influence hemoglobin data quality. There are data quality issues that arise from the increasing requests for diverse types of data. The types of biomarkers that require blood and the influence of different biomarkers on each other were considerations that informants thought required greater attention at the beginning of the survey design process.

> *"Planning phase: what's most important is to understand what people want to measure, and then we figure out the options for measuring them, and how. So this is what specimen type you going to use, what level skill you need in the person taking this specimen type, how much blood you need for all the different tests you're doing." R05*

> *"The design visit needs to be realistic about how many drops of blood [are required] for the various biomarkers." FGD*

The collection of blood is a complex technical component that is unique in the context of the surveys under The DHS Program. Clarifying this during the design phase would promote better understanding, budgeting, and planning for the downstream needs associated with blood collection.

> *"So everything that affects the budget from the technical side has to be very clear up front, and all the decisions have to be made up front." R05*

The understanding that hemoglobin measurement is prone to error must be considered in the survey design. Informants reported that using a HemoCue in a field setting was previously believed to be robust and simple to use. However, in the last several years, there has been increasing evidence that the blood-collection technique influences the accurate assessment of hemoglobin concentrations. Informants emphasized that hemoglobin, as a quantitative measurement of concentration levels, is more prone to error than biomarkers that produce qualitative results (e.g., rapid malaria diagnostic tests), and they believed that this reality was underappreciated.

> *"I know people have called HemoCue results into question, in [country] especially, because we had multiple surveys and some used HemoCue, and some did not." R31*

Hemoglobin data is collected in more than one type of survey in The DHS Program. Another key consideration involved the Malaria Indicator Surveys (MIS), which are administered by The DHS Program. Hemoglobin data are collected in the DHS surveys and the MIS surveys. Some informants believed that the type of survey influences the quality of the hemoglobin data. One informant reported a perception that the MIS was less complex because the questionnaire was shorter than a DHS survey, although the biomarker burden is equivalent in both surveys.

> *"The perception is: it is just an MIS and all you are doing is taking a couple drops of blood, anyone can do that." R35*

Subsampling needs to consider all biomarkers being collected and not just hemoglobin. One proposed strategy for improving data quality involved obtaining hemoglobin data on a subsample. As outlined in the anthropometry section, DHS surveys do not calculate sample size specifically based on nutrition indicators. Thus, it is possible to create a subsample for indicators such as anemia. Because anemia is highly prevalent in most countries, the sample size would need to be sufficient for all the biomarkers that have been selected for a country. In the case of MIS surveys, informants pointed out that the sample size is typically smaller than in DHS surveys. Subsampling in MIS surveys may not be possible but should still be assessed prior to finalizing the sample design. Informants reported that subsampling for anemia would conserve resources

and improve hemoglobin data quality. The training burden would also be reduced because fewer health technicians would be needed, which in turn would allow for more intensive hands-on training.

Blood collection requires specific considerations for team composition. Blood collection, more than any other type of data collection, requires careful thought and advanced planning for proper team composition. Informants felt that data quality was compromised by the common misconception that the collection of hemoglobin data is not very different from other types of data collection. Informants expressed frustration that this misperception resulted in improper staffing by the implementing agency.

For surveys with anthropometry and hemoglobin collection, informants recommended having a designated anthropometry measurer and blood collector. This is important with blood collection when more than one biomarker requiring blood is collected because the health technicians must use blood from the same finger prick. Some informants recommended that nurses or other medical professionals collect hemoglobin data.

> *"Normally specialized people [are used for biomarkers], because you need a nurse to take a blood sample, or a medical doctor needs to be present...so it's easier to have only one nurse who is responsible for blood-taking and then 3-4 people who do the questionnaires." R32*

> *"So what we have often as a setup is the anthropometrist assists the one doing the finger prick or the puncture…the anthropometry assistant actually puts in the cuvette and does the recording. I think that works pretty well, it requires a little bit of training on the hygiene rules but it works pretty well because the [one who is collecting blood]…is the one being stressed about [the other biomarkers]." R28*

One risk of having more than one team member collecting hemoglobin data is the sharing of supplies. There was one report of microcuvettes being removed from the sealed containers and not stored properly. Having more than one health technician per data-collection team should be avoided to prevent potential issues such as these. The teams must be monitored closely when health technicians work together in the field.

Recommendation 2: Place hemoglobin data quality considerations at the forefront of the survey design process.

Steps for consideration by The DHS Program include:

- Ensuring that funding for hemoglobin data collection is adequate to meet the recommendations in this report.
- Increasing the focus on hemoglobin data quality when planning a MIS.
- Collecting hemoglobin data on a smaller subsample that allows for nationally and subnationally representative anemia estimates for each target population (women, children, men).
- Limiting the number of individuals within a team who collect hemoglobin data to allow for adequate hands-on training.
- Over-recruiting a minimum of 15% of health technicians.
- Establishing a supervision system that takes into account hemoglobin-specific data-collection needs to ensure correct procedures are performed.

Training Phase

Many topics related to anthropometry and hemoglobin training were similar, such as the need for continual learning opportunities for The DHS Program biomarker staff and better oversight of biomarker consultants. Other suggestions included soliciting the participation of the Ministry of Health and Nutrition and multilaterals to make a presentation on the public health significance and use of anemia data, and more individualized training with smaller class sizes, more hands-on practice with adults and children, and evening sessions. Other training considerations about hemoglobin were distinct from anthropometry, and these differences are highlighted in this section.

Consultants are an essential resource with local knowledge, but oversight of their work is limited. Several informants discussed different experiences with biomarker consultants as trainers. Biomarker consultants typically have extensive experience with biomarkers that require blood collection. However, informants reported that there was a wide range of familiarity with hemoglobin data-collection procedures. Despite the existence of a standard DHS procedures and a biomarker manual that describes these procedures, some informants reported that different consultants continue to follow different hemoglobin data-collection procedures.

> *"Do you have a standardized training for anemia? The reason I ask is because the consultants were training them not to [fully] close the cuvette machine. So they were closing the device halfway. When I asked, it was because they said the blood could spill in the machine." FGD*

> *"Biomarker consultants may have slight differences on protocol, so having a standard protocol for all consultants will help." FGD*

In The DHS Program, biomarker consultants and staff are usually recruited based on their lab experience, particularly with HIV. The biomarker workshop for biomarkers collected in The DHS Program, as discussed in the anthropometry section, includes hemoglobin as well. This workshop provides an opportunity for all consultants, regardless of prior experience, to standardize hemoglobin collection procedures for The DHS Program. However, in the DHS-7, there was only one workshop. Informants emphasized that more frequent workshops were needed.

> *"[It is essential to] transfer capacity and also update the skill to the standards [of] DHS." R25*

Recommendation 3: Improve training and strengthen the oversight of consultants.

Steps for consideration by The DHS Program include:

- Developing a learning program in which consultants can work with The DHS Program staff on a survey that measures hemoglobin.
- Hosting a minimum of three biomarker workshops per five-year project period and establishing a skills certification program.
- Developing a process for The DHS Program biomarker staff to provide oversight of the consultants who work on all surveys, either remotely or face-to-face whenever possible.

Practice is key because health technicians often do not have experience in hemoglobin data collection prior to the survey. Although health technicians were generally experienced in blood collection, they have limited experience measuring hemoglobin concentrations. Informants noted that use of HemoCues is not usually a part of routine practice in countries, and is specific to The DHS Program. Informants reported the sensitivity of the measurements was underappreciated by health technicians and others involved in the survey. One informant attributed this to the reality that biomarkers, such as rapid malaria tests, do not require the same level of care to obtain correct results.

> *"Typically, what we find is in most countries, techs or nurses that work in hospitals or clinics where there are HemoCues, they don't use them routinely, so for them it's pretty new. There's been very few [settings]…where someone really knew what to do and how to do it correctly." R28*

> *"And throughout the process, we realized that the training of our technicians was crucial.…We used high qualified personnel; we started with university staff, then we used accredited laboratory technicians. But in some surveys, we trained surveyors that have no biological background, and they were not convinced that the capillary function was so crucial, and then we did validation studies that proved that when milking or when not enough training is performed, there are considerable variations within the results in the same person." R26*

> *"Q: Do you have a process of letting go of people who you don't think are going to be able to adequately collect the hemoglobin data?*
>
> *R: Yes, but for biomarkers, I would say not a formalized [process], because unlike the questionnaire, where you can give a test, and you say, 'Okay well everybody has to get an 80% and if you don't…' We don't have that. I mean, how do you conduct a test? But there are people who have very bad habits, maybe they do not fill the microcuvette properly." R05*

> *"...if you go a bit comedian, really exaggerate holding the hand down and all that, so they laugh about it but that sticks in their mind." R28*

For anthropometry data collection, there were challenges described in measuring children who were increasingly agitated during the measurement process. There is additional difficulty for hemoglobin testing because the children are being pricked, and repeated attempts to collect the blood cause additional distress. For this reason, more practice is usually done on adults for hemoglobin testing. This was generally viewed as acceptable because the procedures do not differ substantially for hemoglobin measurement between children and adults (apart from a heel prick in younger children and learning to control child's movement when agitated). Use of fellow trainees for practice is often done, although there may be resistance to having their finger pricked multiple times. Overall, there was a consensus that there need to be more opportunities for health technicians to practice on both children and adults, such as at a community outreach event.

> *"A lot people don't have enough practice pricking young kids for anemia before they go to the field, but it's hard to get, you know? You can't just repeatedly prick these little guys, and you have a lot of people who need to be trained, so you need a lot of kids, and how are you going to find a lot of kids?" R05*

> *"We usually work with adults. But we have said that the lab technicians go to a daycare center and we have obtained approval for that before, so we ensure that they are experienced with children." R26*

> *"So that is a pretty significant logistical burden on the training preparation: to figure out how you're going to get access to lots of little kids, how you're going to bring the kids and the participants together, because there's some sort of transport involved. So that does take a lot of preparation, and if the implementing agency isn't well organized, it can sometimes be too much for them because they don't have a printer, and they don't have a AV system, you know..." R05*

Different training techniques are needed for the HemoCue 301 compared to the HemoCue 201+. In most cases, The DHS Program uses the HemoCue 201+, although the HemoCue 301 has been used, particularly in the MIS surveys. There were differences in the training for the different models of the HemoCue, because HemoCue 301 uses a slightly different method to measure hemoglobin concentrations than the HemoCue 201+. The microcuvettes for the HemoCue 201+ contain an active reagent that allows a reaction to take place in the microcuvette before it is read in the analyzer. In contrast, since the HemoCue 301 microcuvettes do not contain an active reagent, the blood interacts with oxygen before being read in the analyzer. HemoCue 301 microcuvettes must be placed into the analyzer immediately because any delay results in a longer reaction and higher hemoglobin concentrations. Informants described using stopwatches in practice sessions to time trainees from the point of the finger prick to placing the cuvette into the analyzer. Although the HemoCue 301 manual suggests no more than 40 seconds, the informants aimed for no more than 20 seconds.

> *"They were just taking a long time…from the point of the finger stick to placing the cuvette in the machine, you should be aiming for no more than 20 seconds. They were taking more like 45-50 seconds, so what we spent a lot of time on is actually getting the method completely correct and then doing a lot of timed practice. And at the end of the training, all eight people were getting their cuvettes in the machine from the point of finger stick to placing within 15 seconds." R15*

Recommendation 4: Increase individualized hemoglobin training.

Steps for consideration by The DHS Program include:

- Ensuring no more than 10 health technicians are included in each breakout session for hands-on practice, and planning for enough trained facilitators for each breakout room.
- Providing more opportunities for practice on children and adults and potentially a hemoglobin standardization exercise (see recommendation 5).
- Hosting a minimum of three evening clinics that provide individualized training to those trainees who have been identified as having poor technique.
- Increasing the use of demonstrations and videos to review correct procedures and provide course corrections on techniques.
- Providing training with a timer when the HemoCue 301 is used to assure that the microcuvette is placed in the analyzer within 15 seconds of the finger prick.

Sensitivity of hemoglobin concentrations to poor measurement technique indicates that hemoglobin standardization exercise may be warranted. A few informants, external to The DHS Program, emphasized the importance of conducting a hemoglobin standardization exercise during the survey training. Standardization determines intra- and inter-health technician variation. Informants reported following the procedures set forth in the USAID-supported Burger Manual (Burger and Pierre-Louis 2003).[3] Health technician trainees measure blood drops two and three from a finger stick from volunteers. The trainer also takes a blood sample from a separate finger stick from each volunteer. The results are compared between measurements from the same finger prick by the trainee (to determine precision), as well as the difference between the trainer and trainee's measurements (to determine accuracy). Because this requires a high number of finger sticks per volunteer, the exercise should only be performed on adults. Trainees should have sufficient practice to be able to take good finger sticks and blood drops prior to the standardization exercise. It was also suggested the results of the exercise (<0.5g/dL on average) should be considered when selecting health technicians.

> *"There is a Burger manual on standardizing for the HemoCue, it's written by Susan Burger from several years ago, and that's pretty much what we use. It uses the same principles as the anthropometry standardization but usually we don't do that with babies, we just do it with adults, so everyone ends up with many pricks on their fingers. But it's exactly the same method; it depends how you enter the data in to the same kind of table to look at inter- and intra-worker reliability." R31*

The informants explained that although it is possible to observe poor technique during training, the extent of the variability from poor technique is not fully understood until the training standardization exercise is complete. Although most informants support the standardization exercise, one informant raised a concern that there may not be enough variability in hemoglobin concentrations to make the exercise worthwhile.

> *"There's a hemoglobin standardization that is in the survey toolkit, but I kind of shy away from it because I think it's a waste of time because it's not too variable." R19*

Recommendation 5: Explore the incorporation of hemoglobin standardization exercises into the biomarker training.

Steps for consideration by The DHS Program include:

- Piloting the feasibility of hemoglobin standardization exercises to review hemoglobin collection techniques and minimize errors.

[3] Source provided by key informant: Burger S, Pierre-Louis, J. A. 2003. *Procedure to Estimate the Accuracy and Reliability of HemoCue™ Measurements of Survey Workers*. Washington, DC, USA: ILSI.

Data-Collection Phase

Similar topics in anthropometry and hemoglobin data collection were the need to reduce transcription errors of results, and the importance of ensuring that equipment is standardized and maintained throughout data collection. Informants stressed the importance of supervision and they emphasized the need for supervisors to be adequately trained and for specially trained biomarker supervisors to rotate between teams, and the use of checklists and other technology that support monitoring. There were also many findings for hemoglobin data during this phase that were different from anthropometry.

Mixed views were raised on the use of the HemoCue 201+ and HemoCue 301. The DHS Program currently uses the HemoCue 201+, although in a few cases the Program uses the 301 model. Informants expressed diverse views on the use of the HemoCue 201+ versus the 301. The two models are reported to produce comparable results, although several informants raised concerns about comparability. Recent evidence points to potential differences in hemoglobin concentrations between and within models, with greater model variability for the HemoCue 301 as compared to the 201+. More research is needed to determine differences between the models.

> *"I mean ideally they [201 and 301] would be both interchangeable, so it wouldn't matter, but I don't think they are." R28*

> *"Sometimes with the 301, there can be some within model variability and 201 doesn't have as much within-model variability... we've recently published to show that it is not really meaningful but it is a little significant difference between the models" R19*

> *"With hemoglobin, we're really curious about the differences that some have found, between HB 201+ and the 301, (the HemoCue devices), and that's something that could be big topic of research that really needs to be done systematically, because a couple of papers that say there are higher hemoglobin concentrations coming from the 301 and then some of the papers saying they are exactly the same, and manufacturers saying exactly the same." R33*

Some informants noted an advantage of the HemoCue 301, which can operate at higher temperatures. There is no reagent in the cuvette and the reaction takes place in the HemoCue. With the 201+ model, the active reagent in the cuvette is affected by high temperatures. Despite this, informants recommended further research be done before recommending one model over the other.

> *"Using the HemoCue 301, with the cuvette and with the HemoCue itself, you have a wider temperature range. That's the advantage with the 301" R19*

Poor blood collection technique is a common source of measurement error. Informants described the sensitivity of hemoglobin concentrations to the blood-sampling technique. Most incorrect procedures (milking or squeezing finger, shallow puncture, not allowing alcohol to dry, not using the third drop, air bubbles in microcuvette) result in an underestimation of hemoglobin concentrations and an overestimation of anemia. There were also procedures that result in an overestimation of hemoglobin concentrations such as blood clotting, damaged or expired microcuvettes, and re-dipping the microcuvette in the blood sample.

> *"Everything bad can go wrong at collecting the sample, and you can't get any good results from a bad sample." R05*

The DHS Program uses hemoglobin collection protocols to train health technicians on the correct techniques to minimize variability. However, informants expressed a need to increase awareness about

variations in hemoglobin concentrations, not just for the surveys under The DHS Program, but also for the field of nutrition in general. Informants said health technician trainees, and sometimes trainers, are unaware of the extent of the measurement error that is introduced by poor technique. This was exacerbated by the health technicians' experience with the collection of other biomarkers such as malaria, during which squeezing the finger does not alter the test results.

> *"Before I started doing very much training at all, they all already had experience…I asked them to show me a finger stick on somebody… I found there were numerous – really, what I would consider, big – mistakes." R15*

Other factors that cannot be mitigated by improving blood collection techniques were also noted by some informants. For example, dehydration may result in higher hemoglobin concentrations.

> *"With a point-of-care instrument, and using a drop [from the finger], there are potentially many more variables. You are not taking blood from a freely flowing blood vessel. You've got a number of variables that don't apply to [venous blood]. You've got where you take the blood from (which part of the body, which finger). What is the state of hydration of the individual when you take it? Because potentially if they're dehydrated, they [are] hemo-concentrated, and you might potentially get a higher hemoglobin than is the true hemoglobin or truly reflects their truthful, normal truthful blood hemoglobin." R26*

Informants acknowledged the use of different tools to improve blood flow. For surveys in colder climates, informants suggested using hand warmers prior to blood collection. Informants reported positive experiences using hand warmers in surveys and hospital settings.

> "*You do have to warm up the hands, and hand warmers are very [helpful]. I think it's important that you have something. If you don't have hand warmers, you need to take the child's hand. I need to warm up the hand before you collect the blood because you have to get the blood circulating properly through the hand because you're not going to get a good sample if you don't do that." R15*

Informants discussed the advantages and disadvantages of the heel or finger as the site of blood collection in children age 6-12 months. The DHS Program assesses hemoglobin concentrations via heel prick in this age group. A few informants external to The DHS Program reported using finger pricks in this age group because adequate blood quantities can be obtained safely from the finger, and collecting a heel prick in younger children was viewed as adding unnecessary complexity to the survey procedures. The informants explained that WHO guidelines recommend using a finger prick for capillary blood collection in children over age 6 months (WHO, 2010).[4] Of particular concern was the trainees' lack of practice during training with heel pricks before doing the procedure in the field.

> *"I'd say from a clinical perspective, there's really no risk with a finger stick once you're past one or two months of age. So since our surveys are all starting at six months – I think if you're doing neonatal surveys, then the heel stick would come into play. So, there's no medical reason to do a heel stick, and given all the other challenges of doing it..." R15*

> *"Children are often walking around on their feet. Band-Aids come off. It's easier for dirt to get into their feet." R15*

[4] Source provided by key informant: World Health Organization. 2010. *WHO Guidelines on Drawing Blood: Best Practices in Phlebotomy*. Geneva, Switzerland: World Health Organization.

"...it's really hard to do a heel prick if someone has not done a heel prick…if they are just starting out and they don't have expertise in doing the heel prick...they are very hesitant, they don't go deep enough. You have to go very deep with heel prick to get enough blood, get a proper blood sample with a heel prick…" R19

"Sometimes you don't see a heel prick during the training." FGD

In The DHS Program, most surveys use Unistik 3 lancets that have a blade that punctures to a depth of 1.8 mm for both children and adults. In rare cases, lancets that puncture to a depth of 3.0 mm are used for adults with calloused hands. In other surveys, informants report using lancets from Becton, Dickinson and Company that puncture to 2.0 mm for both adults and children, while others reported using the 1.8 mm for pregnant women. Informants explained that the width of the lancet blade was more important than the depth because the width cuts many capillaries and is much more efficient in producing good blood flow. The WHO currently recommends using a lancet depth of 1.5 mm for children and no more than 2.4 mm for adults. The selection and role that lancets play in quality of blood specimens may need to be further explored given the range of opinions reported by the informants.

"I mean I forget the exact brands or the names…our lab has done studies where they've lined up 10 different lancets and you basically want one that tears not just punctures. You want tearing of flesh so you can get a good wound. It might be initially more painful, but it's going to cause less overall pain because you're not going to have to prick three or four times, and you're going to get the blood quicker." R15

"…2mm blade blue lancets worked great for women and kids. But it's really the blade that is the bigger factor …" R33

"The width of the lancet is much more important than the depth, because you need to cut as many capillaries as possible. As you go deeper you may cut a few more, but if you go wider, I think it's much more efficient at cutting capillaries." R33

Informants reported that countries are requesting the inclusion of more biomarkers in surveys. This brings potential data quality concerns into the sequencing of blood collection. In surveys in which hemoglobin is the only biomarker that requires blood, the first two drops are wiped away and the third drop is used, per the standard operating procedures from the HemoCue manufacturer. In some circumstances when HIV and other biomarkers are included, this procedure changes to the fourth or fifth blood drop. Informants said this could be problematic, because one laboratory study showed that the fourth drop of blood resulted in lower hemoglobin concentrations when compared to the second and third drop of blood (Whitehead 2017).[5] Further, the risk of poorer quality blood specimens would be intensified in a field setting where improper behaviors such as "milking" the finger may be used to obtain more blood. The solution proposed by informants was using the third blood drop in all surveys to avoid potential variability in hemoglobin measurement. Similarly, some informants said that using the second drop for dried blood spot (DBS) collection prior to measuring HemoCue was also not advisable because of things like repositioning of the hand to allow the drop to fall onto the filter paper.

Recommendation 6: Improve blood sampling procedures for hemoglobin measurement.

Steps for consideration by The DHS Program include:

[5] Source provided by key informant: Whitehead et al. 2017. "Effects of Preanalytical Factors on Hemoglobin Measurement: A Comparison of Two HemoCue® Point-of-Care Analyzers." *Clinical Biochemistry* 50(9): 519-530.

- Developing materials that describe how poor techniques cause false high and low hemoglobin measurements.
- Reviewing research on sources of variation in hemoglobin concentrations on a yearly basis and incorporating new findings into the data-collection procedures and materials.
- Standardizing across all surveys the use of the third blood drop for hemoglobin and not collecting DBS samples before hemoglobin.
- Modifying blood collection procedures to draw blood from a finger prick for all children age 6 to 59 months.
- Procuring hand warmers for cold climates (and potentially warmer climates).
- Investigating the most appropriate lancet size, depth, and gauge that would result in better data and adhere to within-country and global regulations.

Pooling blood samples may increase the accuracy of hemoglobin results, although pilot testing is necessary to identify the best approach. Informants described different approaches to blood collection. In The DHS Program, the wicking method uses a single drop of capillary blood that is taken directly from the finger and into the microcuvette. Although wicking was the mostly commonly cited method by those working on The DHS Program and other surveys, some informants described using the gravity method. This entails obtaining capillary blood from the finger, dropping the blood onto Parafilm (wax paper), and then filling the microcuvette. Informants described pooling capillary blood by placing a few drops of blood on the Parafilm, or using homogenized capillary blood from microtainers or venous blood from vacutainers. There was no consensus on which method would produce the most accurate results, although pooled blood was thought to be best, especially if samples were obtained by less experienced health technicians. However, further testing was recommended to confirm that pooling with capillary blood results in a better measurement compared to a single drop of capillary blood.

> *"I mean I've done it many different ways: I've done it from the Parafilm, [and] from the finger to filling a cuvette…It depends on what other indicators you're doing, [but] I think the likelihood for mistakes is less when the drop is not off the finger because it's a controlled drop and you can actually see the drop." R15*

> *"I do like the Parafilm approach. I think it minimizes the risk of re-dipping into a blood spot." R15*

> *"I've never done the wax paper thing…but I don't know, it's actually not a bad idea for the really little kids or for... kids who have really liquidy blood, it can be very hard to get the drop. Yeah, wax paper might actually be better." R33*

> *"Just putting a single drop onto a piece of wax paper, I don't think gains you a whole lot." R18*

> *"I worry it could actually work against you because if you put it on to the wax paper, you're not going to be able to recoup all of that blood." R18*

> *"I think it's easier to do it from the finger…it's more steps to put it on the parafilm, and then fill up the cuvette, just more steps. I think it's a lot easier to do it directly from the finger." R19*

> *"Collect a larger quantity into a Microtainer and then sample from the Microtainer; there's an expectation that you're going to get a much more homogeneous supply. Whereas the drop, we know that the amount of interstitial fluid could be quite variable from one drop to the next drop." R18*

> *"There's some concern that as you're pushing the drop out there to the 500 microliters…that much more interstitial fluid creates a bias. I haven't seen studies on that. I don't think it's a huge problem, but it is something that ought to be tested. But I think that that's actually a way to reduce some of the variability. If it's not creating a bias, then I think…something like that would be beneficial." R18*

Informants described that using venous blood might reduce measurement error, and could provide the opportunity to measure more biomarkers. To accommodate more countries' requests for additional biomarkers, many surveys have begun using venous specimens that allow for the collection of larger quantities of blood.

> *"…differences, marked differences between the DHS survey and the national micronutrient survey. The national micronutrient survey, they take venous blood and DHS took capillary blood." R26*

> *"We found that when carefully taken, capillary and venous did get the same results." R26*

> *"Literature is pointing to the possibility that the peripheral blood reading is not quite the same as venous blood, which we would have liked to have reported [in our final report with] an adjusted figure, because relative to other studies that have been done, the result we got [showed] high prevalence in the kids…So I think work needs to be done on that, to see whether there is a real issue or not." R23*

Informants reported that they initially anticipated that training on venous blood collection would be more challenging than capillary collection. However, they found that health technicians with previous phlebotomy experience were easy to recruit, and that their technique was of high quality because phlebotomy had been part of their routine job responsibilities. Another concern was that response rates would be lower when collecting venous blood, but informants reported this was not the case.

> *"I think collecting venous blood is something that we probably will be doing in some countries. Even if we are reluctant to do it, I think we will have to do it maybe in some cases, but I don't think that will be generalized." R21*

> *"There's much less [refusal] than we thought…almost all the surveys we've done, venous is 90% or higher." R15*

Recommendation 7: Explore pooling blood to increase accuracy of hemoglobin measurements.

Steps to be considered by The DHS Program include:

- Piloting different methods of hemoglobin blood collection.
    - Finger wicking versus gravity method on a single blood drop versus gravity method on a pooled blood drop sample versus the gravity method from a microcontainer
- Exploring the feasibility of venipuncture instead of capillary sampling.

Introducing new technology could reduce transcription errors. In the interviews, informants reported errors that were due to poor transcription of hemoglobin results. The DHS Program trains health technicians to record hemoglobin concentrations from the HemoCue machine to a paper-based biomarker questionnaire. The health technicians are taught how to read and record numbers to minimize transcription errors and to ensure standardized recording of numbers such as clearly distinguishing the numbers one and seven. One informant reported health technicians "misreading the numbers" from the HemoCue, while others described health technicians delaying the recording of the hemoglobin concentrations, or recording the hemoglobin concentrations in a notebook first and then later transferring these values to the biomarker questionnaires, which introduced greater opportunity for transcription errors. To avoid transcription errors, one informant suggested using "an electronic download of [the] result."

> *"[Hemoglobin] values are not written down immediately." FGD*

> *"Other malpractice that we have seen … is that the technician can run and put the value on some piece of paper before recording it into the official form that is provided. Sometimes we find the values in the envelopes that have the labels for the tubes. They always transcribe the value, but we encourage them to put the value directly in the form, so there is no error in transcribing the values." R26*

> *"[With] one of the most experienced health technicians that was collecting [hemoglobin], we started looking at the field check tables, and there was peak in 0 and 5…So I talk to her and say 'What's happening?' She tried to explain to me that it was because she couldn't read very well." R21*

Recommendation 8: Minimize transcription errors for hemoglobin measurement.

Steps for consideration by The DHS Program include:

- Exploring the use of Bluetooth technology to transfer hemoglobin results wirelessly from the HemoCue machine to tablets.

Procedures were identified to improve calibration and maintenance of equipment. Calibration, maintenance, and quality assurance procedures were reported by several informants. The first step described was confirming that HemoCue analyzers are operating correctly before beginning field work. Informants cited situations where the HemoCue analyzer results were not comparable across devices, even when the devices were new. To avoid introducing systematic errors in measurement, informants suggested using a venous blood sample to test and compare hemoglobin concentrations between analyzers.

> *"We [organization outside of The DHS Program] have to be able to show that we can reproduce a result. Now that's easy if you've got a whole tube of blood…" R09*

> *"We [organization outside of The DHS Program] verify each equipment that will go to the field each time the survey starts [with the reference blood samples, venous blood samples]." R26*

All informants who measure hemoglobin data in surveys outside of The DHS Program reported always using quality controls in the field to confirm that HemoCue analyzers are operating within acceptable ranges; these controls are not currently used by The DHS Program. The use of controls was highly recommended, despite the fact that HemoCue analyzers are self-calibrating. HemoTrol, an assayed hemoglobin control that mimics whole blood, has been developed for the HemoCue device. It is important to note that different controls are needed for the different HemoCue models (201+ and 301).

Informants reported that outside of The DHS Program health technicians are instructed to use three levels of control solutions to assure that the hemoglobin concentrations fall within low, medium, and high ranges. This process is completed every morning before sample collection. When the HemoCue is not measuring within the range of each of those levels, the health technicians are instructed to first check that the device was correctly cleaned and then to re-run the controls. If the controls remain out of range, the device must be replaced because the HemoCue analyzer cannot be calibrated in real time. Informants acknowledged that the control values fall within a wide range, and that this limits the ability to identify devices that are not working correctly. However, because there have been situations when devices had to be replaced in the field, informants recommended having backup devices.

> *"In terms of making sure that that machine stays like that, that it doesn't drift off, you actually have to use controls…" R09*

> *"…during the field operation, we work with the standard controls with daily runs of the quality controls." R26*

> *"The HemoCues are checked daily, we use two of three controls every day, so level 1 level 2, then level 2 level 3, and then level 3 level 1, and then we restart, and we check all the HemoCues daily." R26*

The manufacturer recommends that solutions be refrigerated in an ice box. This is a challenge for The DHS Program because maintaining a cold chain for quality controls is expensive and logistically challenging. However, one informant cited laboratory results that showed that controls were still effective after incubation, and suggested that The DHS Program could use controls without a cold chain.

> *"You can put them in an incubator for a month and they're still going to register the thing and be OK..." R19*

Informants noted that the HemoCue analyzers must be cleaned daily, which is a procedure included in The DHS Program's Biomarker Manual. Health technicians are instructed to remove and clean the cuvette holder with alcohol swabs, and to use HemoCue cleaners to clean the lens inside the analyzer. Cleaning the

analyzer, particularly the lens, reduces errors in hemoglobin measurement. However, in practice, informants reported that this process is not done consistently.

> *"It's wise to clean that, especially in dusty conditions to make sure that you're not getting dust up inside where the cuvette's being read." R19*

The informants recommended maintaining daily logs as a task that might improve the health technicians' awareness of the importance of carefully handling and maintaining the device. This task would also hold health technicians accountable for following the protocols already in place. The informants reported that the logs might include confirmation of the daily or weekly cleaning of the analyzer and its components, microcuvette expiration dates, checks that assure tight closure of the microcuvette containers, documentation of quality control results, and a listing of all anemia referrals. By maintaining logs, informants were able to document the required equipment monitoring, and could also use the information in reports and on those occasions when the data have been questioned.

> *"So, you just kind of monitor your HemoCue when you're in the field like that to make sure you're operating properly, and keep records…Records of the quality control sheets that we can actually put into the report at the end of the survey." R19*

Informants explained that health technicians are trained to handle common error messages such as the need to replace batteries or clean the analyzer. Health technicians are instructed to retain the troubleshooting guide provided from the manufacturer.

> *"[There is] a book that comes with each HemoCue and then, there is a little guide that comes. I try to keep that with each of the HemoCues, but sometimes they get lost depending on storage and all…It's a guide about if you get this error, how to correct that error, troubleshoot the HemoCue...it's always good to have backup HemoCues in the field." R19*

However, since the HemoCue error codes are not included in The DHS Program's Biomarker Manual, a job aid or inclusion in the manual is clearly necessary to properly maintain the machines. Some informants (outside of The DHS Program) reported laminating the trouble-shooting guide so it could be carried easily during household visits. Health technicians should be consulted about the most useful and easiest methods for transporting job aids, and this information should be available and standardized.

Informants reported, in some cases, health technicians returned analyzers because a problem could not be resolved, and this is another reason informants suggested having backup devices that could be used immediately. Currently, spare HemoCue analyzers are not included in the supply calculations because they can be shipped within a week or two. However, such situations can result in delays in the survey or faulty equipment being used.

Recommendation 9: Strengthen calibration, maintenance, and quality assurance checks of hemoglobin data-collection equipment.

Steps for consideration by The DHS Program include:

- Testing all HemoCue analyzers, new and recalibrated by the manufacturer, by using a venous blood sample prior to using the HemoCue analyzers in the field.
- Procuring additional HemoCues so that faulty equipment can be replaced immediately without delaying data collection.

- Testing the total system (analyzer and microcuvette) by using low, medium, and high liquid controls daily during data collection.
- Developing standard maintenance logs that would be completed each day by health technicians; these would record cleaning of the analyzer, microcuvette expiration date, and checks that assure tight closure of the microcuvette containers.
- Ensuring that all health technicians have standardized directions for responding to error codes displayed on the HemoCue analyzer.

Effective supervision on hemoglobin data collection requires dedicated staff and the ability to detect all deviations from correct procedures. As described in the previous section on anthropometry, informants emphasized that substantial human and financial resources need to be dedicated to supervision beyond that provided for interviews. Hemoglobin data collection requires dedicated biomarker supervision because hemoglobin concentrations are very sensitive to the data-collection techniques. Without careful monitoring to ensure that procedures are followed in the field, hemoglobin concentrations, on average, will be lower than the true value.

> *"I would say that everybody should be very aware that once the drop, the hemoglobin result, is obtained, there is no way of demonstrating that it was good or bad taken. So it is crucial that the training of the persona is very intensive, and the supervising keeps on looking at the good practices because it's very easy when the work is hard, to do it the easy way. If there is no time and the survey has to leave the village for a 2-hour trip, it is very easy to milk the finger to get a quick blood sample, …or if the child is crying, to abandon the cuvette and not read it directly. So the supervision helps that [training practices] are maintained over time. Even with the good training, we observe in the middle of the operation that there were some bad practices that we had to correct." R31*

> *"Sometimes, but not automatically, we add additional training for the biomarker person to do some field operation, but I will say it's only in a few cases because it's too costly." R27*

Informants acknowledged that supervision could be done at the team level or by rotation of the biomarker coordinators. Most thought that the team supervisor was not the optimal person to provide supervision because they may not have the required skill sets or enough time and facilitators to provide training on hemoglobin collection. This was especially true in the context of a DHS survey in which the team supervisor often participates in the interviewer training instead of the biomarker training. In addition, when in the field, the team supervisor often does not have the time to oversee the biomarker collection unless there is more than one team supervisor.

> *"But the team leader can be of any background, you don't need to be a doctor to lead a team in a field. That would actually probably hurt." R33*

> *"We were trying to encourage the logistics person to know about the biomarker process, be available to help with the medical waste and getting equipment there, or doing some of the recording. But I don't think that really happened; they were usually focusing on driving and getting around." R23*

One informant described an effective supervision model in which each team received a weekly field visit from an individual trained in hemoglobin collection (population-based surveys outside of The DHS Program), while other informants thought an approach with rotating biomarker supervision with frequent visits was cost prohibitive. Some described cases in which the collection of other biomarkers required the pick-up of blood specimens from the field. This presented an opportunity to provide supervision on hemoglobin collection, if the supervisors had received training on hemoglobin collection, which is typically not the case.

Supervision of hemoglobin collection requires both theoretical and practical knowledge of correct hemoglobin collection procedures. Technicians must understand each procedure and its importance, from finger capillary blood pricks to the use of the microcuvette and trouble-shooting when there are problems with the HemoCue machine. Deeper theoretical knowledge of the procedure's influence on hemoglobin concentrations is especially important because rarely can data quality problems be ascertained later based on examining the hemoglobin data collected.

As was the case for anthropometry data, many informants emphasized the importance of site visits soon after the training to evaluate and correct technique. For hemoglobin data, hands-on supervision is extra critical to identify potential errors and appropriate course corrections since issues can often not be detected when examining data and how sensitive the results or to health technician technique. However, some informants reported challenges in having sufficiently skilled supervisors to perform these monitoring visits.

> *"We have a problem finding people in-country that can provide that initial sort of supervision during early data collection to ensure that procedures are being carried out correctly in the field. So [organization] can only send two or three people, and then for a limited period of time, and we often try to find others in our collaborating organizations (UNICEF, Ministry of Health, whatever), who can assist with that early supervision. But I find that the numbers are often insufficient, so we can't provide that corrective supervision early on to the extent that might be optimal." R33*

Supervision tools can be useful when hemoglobin procedures are understood. Informants emphasized the importance of continual observations of fieldwork, while some, but not all, viewed a checklist as an effective tool for hemoglobin data collection. Use of the checklist requires understanding that even minor errors in data collection are critical, and many supervisors who use the checklists may not understand how to identify such errors. An informant from a survey outside of The DHS Program reported that team supervisors film the procedures and share the film with the central level staff who have the technical training in hemoglobin measurement to observe the data-collection procedures and provide feedback to the team. They found the filming approach to be helpful during supervision.

> *"Basically we use the checklist to make sure they go and visit the teams…I don't give much value to the actual results of the checklist." R28*

> *"You have to have eyes and you have to have people that know what to look for. You can give them a lot of checklists and so on, but if you don't have the right people who know what to look for, then it doesn't matter – you can write anything you want in the list." R05*

> *"The model works well but we have filmed/taped them for illustrative purposes and when the videos came to us, we found there were some bad practices occurring." R26*

Recommendation 10: Increase observations and enhance supportive supervision for hemoglobin data collection.

Steps for consideration by The DHS Program include:

- Rotating central biomarker monitors among clusters for the full duration of the fieldwork.
- Requiring that individuals who provide supervision on hemoglobin collection receive adequate training and pass a test on correct hemoglobin measurement procedures.
- Implementing use of supervision biomarker checklist that is administered by trained staff.
- Exploring the use of video and other technologies to support the monitoring of hemoglobin collection supervision.

Data-Processing Phase

The data-processing considerations for hemoglobin samples were distinct from those with anthropometry.

Methods are lacking to assess hemoglobin data quality, both in field check tables and final datasets. Several informants reported that field check tables were helpful for identifying teams or individuals who were experiencing problems with hemoglobin collection.

> *"….being able to look at the field check tables [allows you to] see how there are some teams having huge numbers of outliers. Or certain individuals – you can do it by interviewer or biomarker specialist – and see whether people are getting things that are way off the chart." R11*

However, some informants suggested that the primary source of measurement error is poor technique in the collection of specimens that can result in a few grams per deciliter variation. Although differences of this magnitude can cause a major shift in anemia prevalence estimates, the ability to detect these variations with statistical methods is limited.

> *"For the biomarkers, the thing about the supervision and the checks is that if you don't know what is being done incorrectly, you can't change it…by just seeing a number, we don't know, there's no way to correct it." R05*

Informants noted that outliers in hemoglobin concentrations are not common and that there are no established values for what are considered invalid hemoglobin concentrations. Tracking the percentage of invalid hemoglobin data (as in The DHS Program) may lead implementers to assume that hemoglobin data collection is going well when the information they are using to make this assessment has little value.

> *"It's very rare that you're going to have many areas that are high or very low." R19*

> *"I was alarmed with the hemoglobin study finding there were no upper and lower limits for valid values of the hemoglobin … there were apparently no established standards for limits, too low and too high. So there's nothing that the data-processing people could have been using." R12*

A few informants suggested developing metrics that could potentially assess data quality during fieldwork and post survey. One example involved the distribution of hemoglobin concentrations, although this approach is limited because it is unknown if hemoglobin concentrations should follow a normal distribution. Another possible analysis worth investigating would involve calculating the standard deviation with and without each survey team to examine if there are differences in the spread.

> *"… for hemoglobin, I think it's helpful to see the distribution, rather than just a percentage, so you can see: is it skewed, or not? You can see a lot in the distribution. We don't often do that, but I think that's something to consider…" R15*

In the absence of statistical methods to assess hemoglobin data quality, informants emphasized the need to frequently observe and carefully document hemoglobin collection procedures from the beginning to the end of the survey. However, informants described the extensive responsibilities involved in conducting a survey, which make it difficult to monitor data quality issues. Informants recommended holding routine meetings for each survey team to discuss the data quality issues for each survey. One possible platform is the recently introduced 15-minute stand-up meetings for each survey in The DHS Program. During these meetings, data-quality issues for anthropometry and other types of data are discussed. Improved integration of data-quality experts into these discussions of issues at each survey stage was suggested.

Recommendation 11: Increase the identification of hemoglobin data quality issues.

Steps for consideration by The DHS Program include:

- Developing criteria for checking hemoglobin data quality.
- Revising the field check tables to include meaningful information for hemoglobin data.
- Developing a formal documentation system of data-quality issues that arise during survey implementation that could be used when interpreting data quality and included in final reports, with a particular focus on observations of hemoglobin data collection.

Dissemination and Use Phase

Topics related to anthropometry and hemoglobin data dissemination were similar, such as the need to ensure that STATcompiler does not include any data that is suppressed in DHS reports. However, most findings with hemoglobin data at this phase were different from anthropometry.

Minor modifications to dissemination materials will help the interpretability and use of hemoglobin data. Informants recommended small changes to improve The DHS Program final survey report. First, additional context on the data-collection procedures would facilitate cross-survey comparisons, particularly by including information on the HemoCue machine model (201+ versus 301) and blood drop number used to measure hemoglobin. Second, the anemia cut-offs used by The DHS Program for nonpregnant women differ from the most recent WHO guidelines. Thus, revisions to The DHS Program tabulation plan and STATcompiler are needed. Third, since evidence on the multifactorial causes of anemia is rapidly evolving, a review of The DHS Program materials and the incorporation of new evidence will facilitate the interpretation of anemia results in the appropriate context.

Recommendation 12: Increase usability of The DHS Program reports and datasets, revise tabulation plan for anemia, and support anemia results dissemination.

Steps for consideration by The DHS Program include:

- Specifying the HemoCue model used (201+ or 301) in survey reports.
- Revising the language of the tabulation plan and nutrition chapter to reflect current evidence on the causes of anemia.
- Revising the nutrition tab plan based on the WHO-recommended hemoglobin concentration cutoffs to define anemia.
- Applying the correct definition of anemia in previous datasets, as part of an overall data cleaning process.
- Revising anemia status estimates in STATcompiler based on the correct anemia cutoff value.
- Establishing a process to ensure hemoglobin (and other) data suppressed for DHS reports is not included in STATcompiler.
- Providing materials for The DHS Program staff to use when presenting anemia results to ensure correct interpretation.

Infant and Young Child Feeding (IYCF)

Design Phase

Adding a Child questionnaire in DHS surveys must be considered very carefully. Some informants noted that DHS currently asks IYCF questions about the youngest child in the Women's questionnaire, but not of all children in the household under age 2. In other similar large-scale surveys, the IYCF questions are asked about all children under age 2 living in the household. The DHS is nearly representative of all children, because it is a rare occurrence to have a child in the household whose mother is not interviewed. Introducing a Child's questionnaire would be a major change in DHS procedures, and the quality of responses from persons who are not the mother of the child being asked about is not known.

New indicators on unhealthy eating indicators are close to being standardized. Informants discussed the possibility of adding indicators for unhealthy eating into large-scale surveys such as the DHS, and noted that an unhealthy diet is a major risk factor for a range of chronic diseases. The informants highlighted a WHO/UNICEF IYCF consultative group that was considering unhealthy eating indicators for young children.

Informants acknowledged that it was complicated to categorize healthy and unhealthy snacks and to "systematically and consistently operationalize it across the globe." However, some more obvious indicators included "sugar-sweetened beverages" that could highlight unhealthy eating practices for children under age 2.

> *"Until there's some sort of global recommendation, we're not just going to change things and make additions and changes…because it then…it might change a lot, like from year to year. So it's best to wait until there's some sort of agreed-upon standard." R13*

> *"There's some low-hanging fruit. I think sugar-sweetened beverages (with some careful thought about the definition because it's not just soda, it is also tea with sugar and so on)…I think we can start tracking that." R14*

> *"And we also wanted to know where toddlers were getting the sugars and salts. I think we needed to insist on putting that in, and I'm glad we did, but there isn't really a track record for interpreting it yet." R23*

Recommendation 1: Consider new questionnaire elements in The DHS Program.

Steps for consideration by The DHS Program include:

- Determining the feasibility of developing a Child's questionnaire separate from the Women's questionnaire; DHS currently asks IYCF questions for youngest living child instead of all children under age two; this is a proxy for global indicators.
- Adding questions on unhealthy eating practices for children under age two to the IYCF section of the questionnaire based on the 2018 IYCF consultation.

Food and liquids need to be adapted to the local context in the current core DHS questionnaire. Many sections of the DHS questionnaire, including IYCF, are adapted to reflect the population and health issues relevant to the individual countries. Informants agreed that adapting food lists to the local context was important, and some noted the need for liquid adaptation. The IYCF indicator estimates will be biased if these local items are not included, in which complementary feeding will be underestimated (food lists) and appropriate breastfeeding practices overestimated (food and liquid lists). While adaptation is a standard part of DHS procedures, a few informants noted that this does not happen for IYCF in every country.

> *"If you don't pay attention to the food lists, some of the foods are not locally adapted, even if you send it to the country." FGD*

> *"So we've done our best to do a good job…but in terms of customization, it's 50% a disaster and lots of work on our part, on consultants, on regional coordinators." R13*

Many informants emphasized the importance of engaging the right local government and key organizations to obtain buy-in and ensure accountability for the adaptation of the IYCF questions at the design phase. Some informants thought the adaptation process worked more effectively for non-nutrition topics, and they explained that the country stakeholders who represent nutrition are often individuals who represent a broader area of child health. One organization described writing a "terms of reference" for an expert group that functions in parallel with their usual technical committee. This is a dedicated nutritionists group that ensures customization is adequate for the food and drink list, but does not necessarily make the adaptations themselves.

> *"That's why I say you've got to have champions [for nutrition like] you've got for immunization…" you know that there's an entire EPI section that you have to contact…" R21*

> *"That engagement process absolutely is critical. Otherwise you won't get the kind of support that you need to have a high-quality adaptation of the IYCF questionnaire, and you won't have people who are interested in [it, and] no demand for the results…[IYCF] is one of the more demanding modules to adapt, meaning that you need more high-quality engagement in-country, in advance, to make sure that you have a really high-quality questionnaire adapted…particularly for the complementary feeding side of IYCF." R14*

Informants explained that the adaptation must specify local food and drink items but they warned that the specifications cannot be too country-specific, which would reduce the comparability of the indicator across countries. For the actual work of adaptation, informants recommended identifying regional experts who have the ideal balance of knowledge of feeding patterns and local foods, along with the necessary skill to make the appropriate adaptations. This would lead to high-quality adaptations of the IYCF sections of the questionnaire.

> *"I don't think [survey managers] have enough training on adapting the foods. [They] need more information on how the foods are also grouped." FGD*

> *"During the questionnaire design phase…make sure we're working with somebody who knows something about local feeding patterns and local foods, especially for children. So ideally a local nutritionist if that's available, to adapt the list of foods for the dietary diversity questions." R22*

> *"You don't want to adapt too much, because you want to have comparability, [it's] just one of the hallmarks of DHS…" R18*

Informants recommended regional experts should also consult with local experts who can correctly identify and label common food items in their setting. These partners know the correct terminology for foods and

dishes specific to each region, resulting in a list of items that is clearly understood by respondents during the survey. Informants also cautioned that sufficient time must be budgeted for translation to local contexts.

> *"The investment has got to be made to engage with experts in-country; to find them…including from various regions where there's wide variety and feeding practices across large countries. So that you're making sure that you have a good understanding of how people think about those food groups and that the examples are optimal and inclusive. Then you're going to get the best quality data that you can." R14*

> *"First you need to identify all the food that in the list and based on the country, and don't count only with only people at statistical office, discuss with people in nutrition, Ministry of Health, and even discuss with other people, look around what is the consumption of the people in the country…" R21*

However, one informant with extensive experience creating food lists cautioned that experts are not always sufficient to capture what is eaten at the household level, and that it is also important to talk to individuals who are representative of survey respondents. In DHS surveys, this is sometimes considered during the pretest. Trainees provide feedback on food lists in the food groups and, in many cases, the food lists are further adapted based on their input.

> *"If you need that recipe-type information, how well would those key informants be able to speak to recipes and what the main key ingredients are? That would be two professional people speaking from their own personal experience…And how representative are those two professional people? I'm not sure…They mentioned a lot of things that when we came down to the household level, or to the focus group discussions, people said, 'No, actually we don't really eat that.'" R20*

Recommendation 2: Appropriately adapt the IYCF food group examples to the local context.

Steps to be considered by The DHS Program include:

- Developing guidance on how to appropriately adapt the IYCF section of the DHS questionnaire; the guidance should be part of an overall guidance document that includes the other topics requiring adaptation in the DHS questionnaire.
- Engaging local government or multilateral staff to ensure the IYCF section of the DHS questionnaire is appropriately adapted; preferably the process would be overseen by a survey technical committee that includes representation from the host-country nutrition sector.
- Partnering with other organizations that have large-scale surveys to train regional experts on the appropriate adaptation of the IYCF section of the questionnaire, and drawing from this pool of regional experts when adapting surveys.

Training Phase

IYCF-specific guidance and training materials would be useful, but they need to be practical given the length of the full questionnaire. Several informants acknowledged that interviewers are required to learn a great deal of material for the survey, which has implications for the IYCF data. The informants described challenges with the limited duration of IYCF training, which made it demanding for interviewers to fully understand how to probe, as well as to understand all the food groups and the rationale for the questions. For example, interviewers are often unsure about why they need to ask respondents very lengthy questions about their very young infant's food and liquid consumption practices, questions that are used to calculate exclusive breastfeeding.

> *"…a concern of mine with DHS has always been the short amount of training that's available. You know, it's not anyone's fault, it's a huge multi-module survey…" R14*

> *"…[Help] enumerators understand in the training process that this module could take more time than you think it's going to take you, and to be patient and follow the instructions really well in the questionnaire. Work through it and do not just get impatient and stop asking yes/no questions…" R31*

However, informants stated that it is essential for the importance of nutrition indicators to be clearly communicated during training. Interviewers can obtain better quality data when they have a good understanding of local foods and recipes, and the relationship between feeding and optimal growth and development.

> *"I think it is a particular problem for DHS and other surveys like it that are looking at lots of indicators…[Interviewers] are getting hit with so many questions for DHS and other similar big surveys, but I think it's hard for them to maintain the importance of those questions and how to ask them correctly along with everything else they are being trained on…it's a big issue and I think it is partly training, and partly supervision." R22*

Some informants observed that it is important to recognize that the way nutritionists think about food groups is different than that of interviewers and respondents. They explained that respondents are not accustomed to thinking about nutrient components in foods and the concept of food groups is not ubiquitous across countries.

> *"People don't think about, 'What vitamin is in my vegetable?'...Even just listing a few example foods for a food group, that might not spur them to think about something that would to our mind belong in that [food group] category…it just might not be totally intuitive to think that way." R14*

During training, informants suggested various techniques that can increase familiarity with food groups such as introducing the questionnaire and its purpose with interactive presentations; learning the names of foods and dishes in multiple languages; utilizing games that help to teach the sorting of foods; and discussing the rationale for sorting food groups.

> *"It's really difficult actually to refine that list in a way, and put things into categories and capture the local understanding of examples, and then train on them. I actually think it's a pretty tricky process …and in the business of getting the survey together, I think that's often a detail that can be overlooked." R22*

Many informants reported that practice is critical, and that role-playing first with fellow interviewers and then with real survey respondents is helpful. Informants also emphasized that specific attention should be paid to training interviewers to probe for detailed answers. In the DHS surveys, role-playing and probing

for mixed-dish ingredients is sometimes included in DHS training to provide interviewers with practice, especially with complicated feeding scenarios. A few informants suggested additional training on probing for mixed-food dishes. Supervision and job aides were also recommended for maintaining data quality.

> *"…it's more of a probing exercise, like really trying to get people to remember what you did yesterday and what time did you get up. 'What did you have for breakfast? What's the first thing you do? What's the first thing you ate? Did you go out?' Those are the kinds of questions that really get more detailed. Otherwise, you get a very cursory sort of answer that may not be complete..." R20*

> *"Supervisors/trusted individuals should conduct more observation during the interviews of the child feeding sections." FGD*

Recommendation 3: Develop guidance and training materials on IYCF indicators and how to ask IYCF questions correctly.

Steps for consideration by The DHS Program include:

- Expanding information to be included in the interviewer manual on asking IYCF questions and the rationale for why questions need to be asked a certain way.
- Creating PowerPoint presentations for interviewer trainings that show how to ask IYCF questions and the rationale why questions need to be asked a certain way.
- Developing job aids for survey managers on IYCF indicators and the correct approaches for operationalizing them from the design phase through the data dissemination stage.
- Developing a job aid on the types of foods that should and should not be included in each food group.
- Making a "Do's and Don'ts" job aid for interviewers on the correct way to ask IYCF questions.

Data-Collection Phase

Some surveys use the open-recall approach and others use the list-based approach for dietary recall. Informants had different views about using the open-recall (the current MICS approach) or list-based approach (the current DHS approach) for collecting dietary recall data. The WHO IYCF guidelines recommend an open-recall approach to collect data on foods that have been consumed. With this approach, the interviewer guides the caregiver through a "free recall" of foods consumed by the child during the previous day. An alternative approach in the guideline is a list-based approach, in which the interviewer asks questions about the consumption of different foods organized into food groups.

Many informants believed that the open approach required highly skilled interviewers. Some felt that this was a positive aspect, since the burden of interpreting food groups falls on the skilled interviewers and not the survey respondents. However, they also recommended careful training and supervision to avoid artificially inflating diet diversity because open recalls seem to capture more detail, but might include insignificant quantities of foods.

> *"One thing that you do have control over is the training of your enumerators (to the extent that you have a budget and time). You can build in to your survey process to support that. Whereas what you don't have control over is your respondents, in terms of providing them the framework that they might need to respond appropriately to a questionnaire that provides a list-based approach, of food groups and being able to classify the foods they ate into the food groups when they are presented a list-based method, which might not be totally intuitive to them. Particularly among lower literacy population." R14*

> *"I've seen data using both [open recall and list] approaches, and sometimes diversity looks higher with the open recall because you capture more. And sometimes this is real and sometimes this is…a false inflation from having counted these very trivial things…" R14*

Other informants thought that the list-based approach offered better comparability across surveys. It is also more "concrete," and not an "open-end request" or a "complicated dialogue" between interviewer and respondent, and it does not rely on the interviewer to decide which foods belong in what food group.

> *"Questions are tedious to ask and answer" FGD*

> *"By the time the interviewer get to these questions, it's a long list of questions and they are tired, so may just fill in no's, or not probe, etc." FDG*

> *"Respondents are asked about a list of items and it's repetitive. If we could have a way of first marking down the things they say, and then probing for the other foods…" FDG*

> *"I personally like the list approach better because it's very concrete. You know what you asked. If people forgot things, you kind of assume that, well, they are probably forgetting them in the same way across different populations. But you're just more concrete in what words were said to the respondent." R18*

Some informants speculated that interviewers who are expected to administer the list-based approach may ultimately use the open approach, because the list-based approach places a burden on survey respondents.

> *"As they go to the field…you hear complaints like, 'The mom doesn't want us to go all the way through the list, they just want to say my kid just had these two things yesterday, and they just had porridge and water.' And interviewers are really pushing back on not wanting to go through the list, and so we really, we have to push a lot: 'Yes, you need to use this list-based approach, you need to make sure every mom is asked all of these questions.' But I am 100% sure that when there's no supervisor there watching, they don't all do it; the biggest issue is running up against the resistance from moms in the field…women getting irritated that you're asking them all of these food groups for a 3-month old baby." R22*

Ultimately, there was no consensus among informants on the best method. Some informants expressed concern about the use of the different methodologies in MICS and DHS and the differences in interpretation by countries. While there are several advantages and disadvantages to each approach, the most appropriate approach for data quality in the context of large population-based surveys has not been defined. Informants recommended further research and evaluation of the different methodologies to determine the best approach for use in the context of large-scale surveys such as the DHS.

Recommendation 4: Better align the collection of IYCF data with other large-scale surveys and global guidance.

Steps for consideration by The DHS Program include:

- Assessing whether an open-recall or list-based approach should be used in DHS surveys by comparing the two approaches in a DHS pilot.

Data-Processing Phase

Bridge the Gap between Data Collected and Data Processed. For data processing, the main challenges described by informants were related to the questionnaire adaptation. In the DHS, after food lists have been adapted, survey managers create a tabulation plan for the data processors in which different foods are grouped into the standard food categories. This requires survey managers to understand how to group the IYCF indicators so that data processors can produce the codes.

> *"One of the things the survey manager needs to do is to customize the tabulation plan to provide instructions." R24*

> *"We have to keep in mind, data processing is one of those difficult, complex topics/competencies to transfer, so it takes a long time, even with people with the right background to really acquire all the competencies required for data processors." R25*

Some informants suggested that each adaptation of the questionnaire creates an opportunity for error. One informant described a situation in which the questionnaire and the tabulation plan were adapted, but the code was not correctly amended. This established a situation prone to error, in which data processors may not have sufficient understanding of the coding process. Without proper communication and oversight, this type of situation can compromise data quality if not identified early enough. Any part of the DHS survey that requires adaptation is prone to such errors, not just IYCF indicators. A robust solution should be found that can be applied uniformly.

> *"[Data processors] don't have a clue…we don't know whether this [food] is part of Group A or Group B or Group C…" R24*

> *"Too often, countries aren't thinking about how they want to tabulate the data when they are asking for all these additional questions. It all adds up to lots of work and tons of opportunity for errors." R34*

Recommendation 5: Ensure the correct processing of IYCF data in tabulations.

Steps for consideration by The DHS Program include:

- Establishing a review process by two independent survey managers and two independent data processors for the IYCF food group tabulations and other topics that require country-specific adaptation.
- Developing instructions in the data-processing code on parts of core questionnaire that require adaptation for special attention.

Data Dissemination and Use Phase

Revise DHS standard materials based on the WHO/UNICEF 2017 and 2018 IYCF consultations. In general, many informants thought the IYCF questions are useful for both comparability over time and across settings. The WHO and UNICEF have convened many recent meetings about IYCF indicators, in which many informants in this report participated. Aligning DHS materials with the results of these careful deliberations was recommended. Changes to the indicators may be minimal, but the explanations and presentations can improve understanding of the current indicators of nutritional status of children.

One example of an item that needs to be updated is the information on breastfeeding practices. It is recommended that this information should be shown as an area graph to depict the proportion of children receiving relevant feeding practices by the child's age. The current DHS tabulation plan does not include the area graph in the format recommended by informants, but its inclusion is important to consider.

Another example includes the recent changes to the existing complementary food indicators. The Minimum Dietary Diversity (MDD) and Minimum Acceptable Diet (MAD) indicators were revised at the WHO/UNICEF 2017 IYCF meeting. In the 2018 WHO/UNICEF consultation, agreement was reached to also modify the Minimal Meal Frequency (MMF) indicator. The DHS tabulation plans for these indicators will need to be revised, along with the estimates recalculated from older surveys to illustrate trends and make comparisons over time in the STATcompiler.

Recommendation 6: Revise DHS standard materials based on the WHO/UNICEF 2017 and 2018 IYCF consultations.

Steps for consideration by The DHS Program include:

- Reviewing and revising the tabulation plan and text for breastfeeding and complementary feeding indicators; this would include an area graph on IYCF feeding practices, as recommended by WHO, in the DHS tabulation plan and a video that describes how to interpret the area graph.
- Updating the STATcompiler with revised breastfeeding and complementary feeding indicators.
- Including MDD, MAD, and MMF as calculated indicators in the recode microdata data sets.

Use of Nutrition Data to Take Action

The importance of building capacity to use the nutrition data from surveys in The DHS Program was emphasized by many informants. Informants recommended holding nutrition workshops to make data understood and accessible for use in-country purposes.

> *"Realistically, when they're making decisions on where they should emphasize programs where things aren't going well, they need to get a little bit deeper into the data, which is all there in the DHS report, but nobody wants to pick up a big thick book and try to figure it all out." R18*

> *"Our [DHS] workshops now, we cover all areas but we only touch very lightly on nutrition. There could be [more] if there's money and need and interest, there could easily be a nutrition-focused workshop, where you discuss all nuances and all these things that happen in DHS data with respect to nutrition, training the people, how to interpret it, and how to maybe produce the estimates or something like that." R12*

> *"I would try to bring nutrition to the forefront, to mainstream it along with malaria and the other topics that [DHS] is doing that for…I could see that [approach would] probably be pretty easily adapted to nutrition. You have a two-day workshop to review your nutrition results with your stakeholders in-country, what these results mean and what they imply for the actions going forward in the country..." R14*

CONCLUSION

The DHS Program undertook an assessment to identify actions that can be taken to enhance anthropometry, hemoglobin, and IYCF data quality. In this assessment, 54 key informants and 13 focus group participants elicited a variety of approaches for enhancing the quality of nutrition data. Participants included staff from within The DHS Program and external professionals with experience collecting similar data with household-based surveys. The process led to 32 recommendations that can enhance the quality of anthropometry, hemoglobin, and IYCF data. The findings and recommendations in this report summarize the experts' perspectives, and do not necessarily reflect the collective views of The DHS Program or the entire nutrition community.

Participants identified critical challenges to collecting high-quality nutrition data, and offered recommendations that could improve data quality. Some recommendations can be adopted easily, while others will require changing current systems and testing new approaches. The next step is to embark on a prioritization exercise that will identify the short-, medium-, and long-term activities for The DHS Program. Potential actions will be thoughtfully and collaboratively prioritized in partnership with The Program DHS staff, USAID, host countries, and key external stakeholders.

The criteria for prioritization will include the extent to which a recommendation improves data quality, the required level of effort, implementation capabilities in different settings, and the associated costs. Since The DHS Program surveys do not focus exclusively on nutrition, all recommendations must be considered within the context of the full survey. Ideally, enhancing nutrition data quality will strengthen the quality of the full survey, although care must be taken to prevent compromising other survey components.

The proposed strategies have the potential to improve nutrition data quality, although each strategy has funding requirements related to new equipment, training, data collection, supervision, and monitoring. Each of these recommendations also requires increased technical assistance which also increases costs. Since some recommended strategies have not been tested for use in The DHS Program, a dedicated funding stream will be needed to pilot test innovations, and to continually monitor other recommendations to ensure that their ongoing implementation leads to the expected outcomes.

A major strength of The DHS Program is the Program's focus on increasing the capacity of host country partners to collect and use data. This means that The DHS Program can recommend best practices to host countries, but does not always have the leverage to ensure that the countries adopt or follow these practices. The success of recommendations depends on the buy-in and collaboration with the implementing agency of the host country and other local stakeholders.

The DHS Program has collected difficult-to-obtain data for use worldwide to inform decision-making on nutrition in low- and middle-income countries. The high-level capacity within The DHS Program and the infrastructure built and maintained by The DHS Program staff over many years provides an opportunity to implement new measures that enhance data quality across surveys and contexts. It is hoped that many recommendations in this report will become best practices across all surveys in The DHS Program. Such an endeavor will require careful coordination, piloting of strategies, committed funding, and continued monitoring to achieve the desired results.

REFERENCES

Development Initiatives. 2017. *Global Nutrition Report 2017: Nourishing the SDGs.* Bristol, UK Development Initiatives. http://globalnutritionreport.org/the-report/.

ICF International. 2012. *Measure DHS Biomarker Field Manual*. Calverton, Maryland, USA: ICF International.
https://dhsprogram.com/pubs/pdf/DHSM7/DHS6_Biomarker_Manual_9Jan2012.pdf.

Kothari, M. and N. Abderrahim. 2010. *Nutrition Update 2010*. Calverton, Maryland, USA: ICF Macro.
https://dhsprogram.com/pubs/pdf/NUT4/NUT4.pdf.

Kothari, M., N. Abderrahim, A. Coile, and Y. Cheng. 2014. *Nutritional Status of Women and Children: A 2014 Update on Nutritional Status by Sociodemographic and Water, Sanitation, and Hygiene (WASH) Indicators Collected in Demographic and Health Surveys*. Rockville, Maryland, USA: ICF International.
https://dhsprogram.com/pubs/pdf/NUT6/NUT6.pdf.

WHO. 2008. *Indicators for Assessing Infant and Young Child Feeding Practices. Part 1 Definitions.* Conclusions of a consensus meeting held November 6-8, 2007 in Washington, DC, USA. Geneva: World Health Organization.
http://apps.who.int/iris/bitstream/handle/10665/43895/9789241596664_eng.pdf;jsessionid=B72E66B0D170B4233EE55C83C53E7998?sequence=1.

WHO. 2010. *Indicators for Assessing Infant and Young Child Feeding Practices: Part 2 Measurement*.
http://www.who.int/maternal_child_adolescent/documents/9789241599290/en/.

WHO. 2011. *Haemoglobin Concentrations for the Diagnosis of Anemia and Assessment of Severity*. Edited by Vitamin and Mineral Nutrition Information System. Geneva: World Health Organization. (WHO/NMH/NHD/MNM/11.1). http://www.who.int/vmnis/indicators/haemoglobin/en/.

WHO. 2017. *Global Nutrition Monitoring Framework: Operational Guidance for Tracking Progress in Meeting Targets for 2025*. Geneva: World Health Organization.
http://apps.who.int/iris/bitstream/handle/10665/259904/9789241513609-eng.pdf?sequence=1.

WHO and CDC. 2007. *Assessing the Iron Status of Populations: Report of a Joint World Health Organization/Centers for Disease Control and Prevention Technical Consultation on the Assessment of Iron Status at the Population Level*. 2nd Ed. Geneva: World Health Organization.
http://www.who.int/nutrition/publications/micronutrients/anaemia_iron_deficiency/9789241596107.pdf.

APPENDIX I

Key Informant Interview Guide

I have gotten in touch with you because we are writing a report on quality improvement and innovations for nutrition data within The DHS Program. The purpose of the report is to share lessons learned and present recommendations for future surveys. The report will be published and shared both internally within The DHS Program and externally to a wider audience. In order to do this successfully, it is important to understand the perspectives of people who are involved in surveys that collect nutrition data, so thank you for agreeing to talk with us today to share your experience and recommendations. During our discussion today, we will be focusing on a few core areas, including:

- Anthropometry
- Infant and Young Child Feeding
- Hemoglobin assessment

[Interviewer circle or bold ahead of time the modules that will be covered.]

Your participation in this interview is completely voluntary. If you want to stop at any time or don't feel comfortable answering a question, please let me know. Taking part in this interview is your agreement to participate.

This discussion will be recorded, and the recordings will be kept securely and only accessed by the research team. We would also like to collect demographic information about the experts we have interviewed. We will send you a separate email with a code and a link to a 10-question demographic survey online. Please fill that out when you have the time.

The information you provide as part of the interview and the survey will not be attributed to you individually but your name will appear in the appendix of the report. Do I have your permission to record the discussion and include your name in the appendix of the report? Our discussion today will take____**.** *[Enter time prior to beginning interview. Estimate 30 minutes per module.]* Do you have any questions before we start?

[Answer any questions.]

Let's begin. At the beginning, I am going to ask a few basic questions about your background, but we will quickly move on to your views about different stages of the data-collection process. Please feel free to respond to questions that you are not an expert in if you have something to share. You also are free to skip questions that you do not have an opinion on. *[START RECORDER]*

I. Demographic Information

1. What is your title or role in your institution?
2. Please describe your experience in the collection or use of survey data. Anything you would like to add on your experience in nutrition survey data specifically?

II. Data Quality

I'd like to talk about the <u>survey preparation and design</u> phase of surveys.

3. In your experience, describe what you find typically works well when preparing for and designing a survey. We are interested in things that specifically relate to the collection of
   Anthropometry:
   Infant and Young Child Feeding:
   Hemoglobin assessment:
   [Only ask questions on the topic(s) the key informant interview has expertise]

   Probes: Ask why for each item discussed.
   Probes: Ask about the sampling design (especially subsampling), supplies, financial, scope of survey, partner consensus, time, competing priorities, staff, ethical approval or confidentiality, governance.

4. In your experience, describe what you find to be the biggest challenges when preparing for and designing a survey. We are interested in things that specifically relate to the collection of
   Anthropometry:
   Infant and Young Child Feeding:
   Hemoglobin assessment:
   [Only ask questions on the topic(s) the key informant interview has expertise.]

   Probes: Ask why for each item discussed.
   Probes: Ask about the sampling design (especially subsampling), supplies, financial, scope of survey, partner consensus, time, competing priorities, staff, ethical approval or confidentiality, governance.

5. For the challenges you mentioned, what sorts of things could reduce these challenges in the future?
   We are interested in things that specifically relate to the collection of
   Anthropometry:
   Infant and Young Child Feeding:
   Hemoglobin assessment:
   [Only ask questions on the topic(s) the key informant interview has expertise.]

Next, I would to talk about the <u>training</u> phase of surveys.

6. In your experience, describe what you find typically works well when conducting a survey training. We are interested in things that specifically relate to the collection of
   Anthropometry:
   Infant and Young Child Feeding:
   Hemoglobin assessment:
   [Only ask questions on the topic(s) the key informant interview has expertise.]

   Probes: Ask why for each item discussed.
   Probes: Staff, supplies/equipment, logistics, time/duration of sessions, language.

7. In your experience, describe what you find to be the biggest challenges when conducting a survey training. We are interested in things that specifically relate to the collection of
   Anthropometry:

Infant and Young Child Feeding:
Hemoglobin assessment:
[Only ask questions on the topic(s) the key informant interview has expertise.]

Probes: Ask why for each item discussed.
Probes: Staff, supplies/equipment, logistics, time/duration of sessions, language.

8. For the challenges you mentioned, what sorts of things could reduce these challenges in the future?
   We are interested in things that specifically relate to the collection of
   Anthropometry:
   Infant and Young Child Feeding:
   Hemoglobin assessment:
   [Only ask questions on the topic(s) the key informant interview has expertise.]

   Probes: Staff, supplies/equipment, logistics, time/duration of sessions, language

Now, I would to talk about the <u>data-collection</u> phase of surveys.

9. In your experience, describe what you find typically works well during the data-collection phase.
   We are interested in things that specifically relate to the collection of
   Anthropometry:
   Infant and Young Child Feeding:
   Hemoglobin assessment:
   [Only ask questions on the topic(s) the key informant interview has expertise.]

   Probes: Ask why for each item discussed.
   Probes: Staff, supplies/equipment, logistics, time, field checks, language.

10. In your experience, describe what you find to be the biggest challenges during the data-collection phase. We are interested in things that specifically relate to the collection of
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    [Only ask questions on the topic(s) the key informant interview has expertise.]

    Probes: Ask why for each item discussed.
    Probes: Staff, supplies/equipment, logistics, time, field checks, language.

11. For the challenges you mentioned, what sorts of things could reduce these challenges in the future?
    We are interested in things that specifically relate to the collection of
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    [Only ask questions on the topic(s) the key informant interview has expertise.]

We are about halfway done. I would to talk about the <u>data-processing</u> phase of surveys.

12. In your experience, describe what you find typically works well during the data-processing phase of the survey. We are interested in things that specifically relate to the collection of
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    [Only ask questions on the topic(s) the key informant interview has expertise.]

    Probes: Ask why for each item discussed.
    Probes: Staff, logistics, time.

13. In your experience, describe what you find to be the biggest challenges during the data processing phase. We are interested in things that specifically relate to the collection of
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    [Only ask questions on the topic(s) the key informant interview has expertise.]

    Probes: Ask why for each item discussed.
    Probes: Staff, logistics, time.

14. For the challenges you mentioned, what sorts of things could reduce these challenges in the future?
    We are interested in things that specifically relate to the collection of
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    [Only ask questions on the topic(s) the key informant interview has expertise.]

Next, I would like to talk to you about <u>data dissemination and use of data</u>.

15. In your experience, describe what you find typically works well when disseminating data and promoting its use. We are interested in things that specifically relate to the collection of
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    [Only ask questions on the topic(s) the key informant interview has expertise.]

    Probes: Ask why for each item discussed.
    Probes: financial, staff, logistics, country involvement, workshops.

16. In your experience, describe what you find to be the biggest challenges when disseminating data and promoting its use. We are interested in things that specifically relate to the collection of
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    [Only ask questions on the topic(s) the key informant interview has expertise.]

    Probes: Ask why for each item discussed.
    Probes: Financial, staff, logistics, country involvement, workshops.

17. For the challenges you mentioned, what sorts of things could reduce these challenges in the future?
    We are interested in things that specifically relate to the collection of
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    [Only ask questions on the topic(s) the key informant interview has expertise.]

18. Are there any indicators that you are aware of within the DHS core survey that the nutrition community is NOT using?

For the last phase of the survey, I would like to talk to you about the <u>Research</u> on data quality for surveys.

19. In your experience, describe what you find typically works well when assessing the quality of data post-survey. We are interested in things that specifically relate to the collection of
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    [Only ask questions on the topic(s) the key informant interview has expertise.]

    Probes: Ask why for each item discussed.

20. In your experience, describe what you find to be the biggest challenges when assessing the quality of data post-survey. We are interested in things that specifically relate to the collection of
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    [Only ask questions on the topic(s) the key informant interview has expertise.]

    Probes: Ask why for each item discussed.

21. For the challenges you mentioned, what sorts of things could reduce these challenges in the future?
    We are interested in things that specifically relate to the collection of
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    [Only ask questions on the topic(s) the key informant interview has expertise.]
    Probe: Anything else until no further responses.

III. Innovation

I would now like to shift gears and discuss new innovations in nutrition data and collection.

22. Do you think there any nutrition innovations or technologies that you think are ready for use in population-based surveys or will be ready soon in the areas of:
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    Micronutrient biomarker assessment:
    Other nutrition topics:
    If so, what?

23. From your perspective, what are the benefits to using these new nutrition innovations or technologies in the areas of:
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    Micronutrient biomarker assessment:
    Other nutrition topics:

24. From your perspective, describe what you find to be the biggest challenges to using these new nutrition innovations or technologies in the areas of:
    Anthropometry:
    Infant and Young Child Feeding:
    Hemoglobin assessment:
    Micronutrient biomarker assessment:
    Other nutrition topics:

25. Can you describe the advantages and disadvantages of including new nutrition indicators or biomarkers within a:
    a. Full DHS survey?
    b. Subsampling within a DHS survey?
    c. Piggy-backing onto an existing DHS survey?

    Of all three approaches, what would you suggest and why?

IV. Closing Questions

26. Do you have anything else you would like to add that we have not already talked about?

27. Do you have any questions for me?

28. Finally, is there one person that you think it would be critical we speak with as part of this study?

Thank you for your time. I really appreciate your talking with me today.
[END RECORDER]

APPENDIX II

List of Key Informants

Hassan Ali Ahmed	Action Against Hunger Canada (ACF)
Mike Amakyi	ICF South-by-south consultant
Mary Arimond	FHI 360
Fred Arnold	ICF
Shireen Assaf	ICF
Bernard Barrere	ICF
Bo Robert Beshanski-Pedersen	United Nations Children's Fund (UNICEF)
Oleg Bilukha	Center for Global Health (CGH), Division of Global Health Protection (DGHP), Centers for Disease Control and Prevention (CDC)
Debbie Bradshaw	South African Medical Research Council
Elizabeth Britton	ICF
Kenneth H. Brown	Bill and Melinda Gates Foundation (BMGF)
Merrole Cole-Sinclair	St. Vincent's Hospital Melbourne, Australia
Trevor Croft	ICF
Anne Cross	ICF
Megan Deitchler	FHI 360
Mahmoud Elkasabi	ICF
Joy Fishel	ICF
Dean A. Garrett	PATH
Laurence Grummer-Strawn	World Health Organization (WHO)
José Miguel Guzman	ICF
Abibata Handley	Encompass
Chika Hayashi	United Nations Children's Fund (UNICEF)
Rebecca Heidkamp	Johns Hopkins School of Public Health
Christine Hotz	Independent consultant
Toni Jones	ICF
Pamela Kakande	Uganda Bureau of Statistics
Sunita Kishor	ICF
Monica Kothari	PATH
Julia Krasevec	United Nations Children's Fund (UNICEF)
Eva Leidman	Center for Global Health (CGH), Division of Global Health Protection (DGHP), Centers for Disease Control and Prevention
Dora Inés Mazariegos-Cordero	Laboratorio de Bioquímica Nutricional, Unidad de Nutrición y Micronutrientes (NM-NL) Instituto de Nutrición de Centroamérica y Panamá (INCAP)
Vrinda Mehra	United Nations Children's Fund (UNICEF)
Zuguo Mei	International Micronutrient Malnutrition Prevention and Control Program (IMMPaCt) Division of Nutrition, Physical Activity and Obesity, Centers for Disease Control and Prevention (CDC)
Purnima Menon	International Food Policy Research Institute (IFPRI)
Melinda Munos	Johns Hopkins School of Public Health
Suteera Nagavajara	ICF consultant

Erica Nybro	Johns Hopkins University
Thomas Pullum	ICF
Kia Reinis	ICF
Ruilin Ren	ICF
Elizabeth C. Rhodes	Emory University
Fabian Rohner	Groundwork
Guillermo Rojas	ICF
Victoria Sauveplane	Action Against Hunger Canada (ACF)
Kerry Schulze	Johns Hopkins School of Public Health
Gulnara Semenov	ICF
Irwin Shorr	Shorr Productions LLC
Parminder S. Suchdev	Emory University
Katie Tripp	International Micronutrient Malnutrition Prevention and Control Program (IMMPaCt) Division of Nutrition, Physical Activity and Obesity, Centers for Disease Control and Prevention (CDC)
Ann Way	ICF
Ralph Whitehead	International Micronutrient Malnutrition Prevention and Control Program (IMMPaCt) Division of Nutrition, Physical Activity and Obesity, Centers for Disease Control and Prevention (CDC)
Frank Wieringa	Institut de Recherche pour le Développement
James Wirth	Groundwork
Bradley Woodruff	Groundwork