© WHO / Isaac Rudakubana

unicef

World Health Organization

Research Agenda Brief

RESEARCH AGENDA TO ADDRESS GAPS IN DATA COLLECTION, ANALYSIS AND REPORTING ON ANTHROPOMETRIC INDICATORS IN CHILDREN UNDER 5 YEARS OLD

Age collection methodology in population-based surveys to generate anthropometric Z-scores

Statement of problem

Estimates of childhood malnutrition are largely obtained from population-based surveys. A child's age, derived using the date of birth (DOB) and date of interview/measurement (DOI), is an essential piece of information in the calculation of childhood nutrition indicators. Examples include anthropometric indices of weight-for-age and height-for-age. For anthropometric assessment, the day, month and year of birth are used to calculate exact age in days for deriving the z-scores based on the WHO child growth standards. In cases where the day of birth is not available, 'missing day of birth' should be imputed. The standard analysis approach is to use the fifteenth day of the month.[1]

Two critical components determine the accuracy of a child's age: (1) the data points collected to calculate age; and (2) the source of that information (respondent-reported, event calendars, records). The established best-practice data points are birth month and year, at a minimum, as opposed to the age of the child at the time of interview.[2] This is the case regardless of the source of information.[3]

> **Best practices regarding the information source are much less clear and are the focus of this research brief.**

Typically, in a population-based survey, a child's DOB and age are collected by asking the mother or the primary caretaker. A consistency check can be performed where the child's DOB is compared against the child's age in completed years; if the responses do not agree, further probing is done, and the age, the birth date or both are adjusted. Another source of DOB data are home-based records, such as birth certificates, child health handbooks or vaccination cards.[4] In some surveys, a local calendar of events is used (i.e., a customized calendar that provides dates of significant events for a specific geographic area) with the aim of reducing errors in estimating a child's DOB.

When using the local calendar of events, the goal is to narrow down the DOB of the child to at least the nearest month and year, and preferably, to the day. This is done by asking the caregiver a series of 'before and after' questions, with the aim of identifying two known events closest to the child's DOB, one which occurred before and one which occurred after the child was born. This technique is called the 'sandwich' because the child's estimated month and year of birth are sandwiched between two identified events in the calendar.

Settings with poor vital registration systems and cultures where the celebration of birthdays is not common present a particular challenge. For example, a large portion of children with the birth date of January 1 has been observed due to a delay between a birth and its registration and for other reasons.[5] Further evidence is needed on which information source or combinations of sources should be used to determine age and which approach is best depending on the setting. There is also a need to better understand how best to implement the collection of data using different information sources. For instance, for event calendars, there are variations in implementation that need to be resolved, such as the format of the calendar.

In anthropometric assessments for children under 5 years of age, inaccuracies in age will affect height-for-age and weight-for-age (as well as mid-upper arm circumference-for-age and BMI-for-age) but not the weight-for-height indicator. Random age errors often show up in the age distribution as 'heaping' at regular intervals, such as every six months. These errors will artificially inflate the standard deviation of the z-score distributions and can lead to a spurious relationship between birth month and average z-score.[6] Inflation of the standard deviation in the z-scores can lead to sizeable biases in anthropometric prevalence estimates that are reliant on age determination, such as stunting and underweight.[7] Directional biases in age determination have also been observed,[8] but as they generally cannot be observed

retrospectively in either the age or the z-score distribution, their historical presence and variation by context is unclear.

Biases in anthropometric prevalence estimates as a result of these various age measurement issues can be much larger than the associated statistical uncertainty, thereby threatening the validity of the results. Studies have indicated that an error in age can potentially lead to a misestimation in malnutrition prevalence ranging from 7 to 30 percent.[9,10]

This brief describes a research agenda for improving the collection of age information among children under 5 years of age in population-based surveys.[11]

© UNICEF/UN0733024/Kongmuang

Research questions and approach:

1. What best practices for the currently available DOB data collection methods should be used for children under 5 years of age in population-based surveys?

2. Are there improvements that can be made to currently available DOB data collection methods, including use of computer-assisted personal interviewing (CAPI), to improve the ease of DOB collection and the quality of DOB data?

3. What is the most accurate and reliable source of DOB data, or combination of DOB data sources, to estimate age in days for children under 5 years of age in population-based surveys and does it vary by context?
   a. Respondent-reported recall of DOB (includes consistency check between DOB and age)
   b. Event calendars
   c. Home-based records by the type of record (e.g., child health handbooks, vaccination cards, birth certificates, birth registration)
   d. Health facility-based records obtained by visits to health facilities while in the field
   e. Triangulation of approaches

4. What best practices should be used for training on DOB data collection?

5. What age-related paradata should be used and reported to understand the quality of age estimation data?

Proposed research approach for improving the collection of age information for children under 5 years of age in population-based surveys

Research topic 1

Best available practices for DOB data collection

Type of research
Mixed method review: Identify best practices for DOB based on protocols and methods used in different surveys to collect DOB.

Outcomes
Primary: Best practices for collecting DOB data and subjective field-based experience on their effectiveness.

Secondary: Estimate of how common the DOB data collection method or combination of methods is and how it has changed over time.

Data source(s)
A systematic search of peer review and grey literature. An online survey and/or key informant interviews with stakeholders.

Research topic 2

Improvements in DOB data collection methods

APPROACH 1

Type of research	Outcomes	Data source(s)
Qualitative: Undertake a Delphi method to obtain consensus on: 1) best practices for developing local event calendars using currently available technology; and 2) determine utility and define parameters for the development of artificial intelligence-based (AI) tools to automatically generate local event calendars.	**Primary (1):** Characteristics of high-quality event calendars using currently available technology from diverse perspectives; **Primary (2):** Feasibility of software-generated local event calendars and their empirical suitability for field use.	Questionnaire administered to a panel of experts in two or more rounds followed by a virtual or in-person meeting.

APPROACH 2

Type of research	Outcomes	Data source(s)
Product development: Identify generic AI-based technology that can be adapted for generation of local event calendars and/or the development of new software. Dependent on an AI-generated event calendar tool being deemed feasible based on the research outlined above.	**Primary:** AI event software developed that is ready for efficacy studies.	Local event calendars based on AI technology and the traditional method in diverse contexts.

APPROACH 3

Type of research	Outcomes	Data source(s)
Primary data collection: Efficacy and effectiveness studies comparing local event calendars developed via AI versus the traditional method. Dependent on an AI-generated event calendar tool being deemed feasible based on the research outlined above. Effectiveness studies dependent on outcomes of efficacy studies.	**Primary:** Differences in the local event calendar content between the automated and traditional event calendar tool.	Small-scale efficacy studies (multi-centre or individual) in a controlled environment. Small-scale effectiveness studies (multi-centre or individual) in a field environment from diverse settings including in low- and middle-income countries. Validations independent from the developers are required.

APPROACH 4

Type of research	Outcomes	Data source(s)
Primary data collection: Development and pilot testing of procedures to improve the ease and quality of DOB collection, including photographing home-based records, programming event calendars in CAPI, consistency checks,12 and out-of-range checks (e.g., calculation of anthropometric z-scores in CAPI and re-probing for DOB/age if z-scores are outside of acceptable limits).	**Primary:** Feasibility and utility of methods to improve the ease and quality of DOB collection.	Small-scale studies (multi-centre or individual) in a field environment from diverse settings including in low- and middle-income countries.

Research topic 3

Most accurate/reliable source/s of DOB data

APPROACH 1

Type of research

Secondary analyses: Assess the relationship between DOB data quality and the data collection approach (recognizing that this would not be a randomized study and therefore confounding factors are likely to be present).

Outcomes

Primary: Age data quality (for example, DOB completion rates and age heaping) by DOB collection method.

Secondary: Age data quality of respondent-reported recall disaggregated by respondent type.

Data source(s)

Secondary data obtained through a systematic search of peer review and grey literature, including surveys reports.

APPROACH 2

Type of research

Primary data: Effectiveness studies comparing known date of birth via either documentation with the family at the time of birth or verified birth registration data to results from six randomly-assigned study arms that will collect DOB via (a) respondent-reported recall only; (b) respondent-reported recall with the aid of event calendar prompts; (c) home records only; (d) respondent-reported recall plus home records; (e) respondent-reported recall with facility visit confirmation; (f) respondent-reported recall with the aid of event calendar prompts and facility visit confirmation; (g) home records and facility visit confirmation; (h) respondent-reported recall plus home records and facility visit confirmation; and (i) the combination of all methods.

With a cohort study, data can be collected at multiple time points to investigate the relationship between DOB accuracy and the age of the child (e.g., Do caregivers report younger kids as older? Do caregivers report older kids as younger?).

Outcomes

Primary (1): The level of systematic and random age errors present in each study arm compared to the known DOB;

Primary (2): Differences in anthropometric z-scores and prevalence estimates in each study arm compared to the known DOB.

Secondary (1): At a minimum, stratification of the primary outcomes by age. Alternatively, if multiple measurements between birth and age 5 are possible, capture patterns on how the primary outcomes may or may not vary with the child's age;

Secondary (2): Cost-effectiveness of each age collection method.

Data source(s)

Cohort field research (multi-centre or individual) in a field environment from diverse settings, including in low- and middle-income countries.

Study(s) can be nested within existing cohort-based research where age of birth is known for cost-effectiveness purposes.

Research topic 4

Best practices for effective training

Type of research

Mixed methods review: Conduct a landscape assessment to identify best practices for training on DOB collection.

Outcomes

Primary: Identification of best practices for age data collection training in population-based surveys and subjective field-based experience on their effectiveness.

Secondary: Estimate of how often various training methods are used in practice and any trends over time.

Data source(s)

A systematic search of peer review and grey literature. An online survey and/or key informant interviews with stakeholders.

Research topic 5

Understand quality of age estimation data

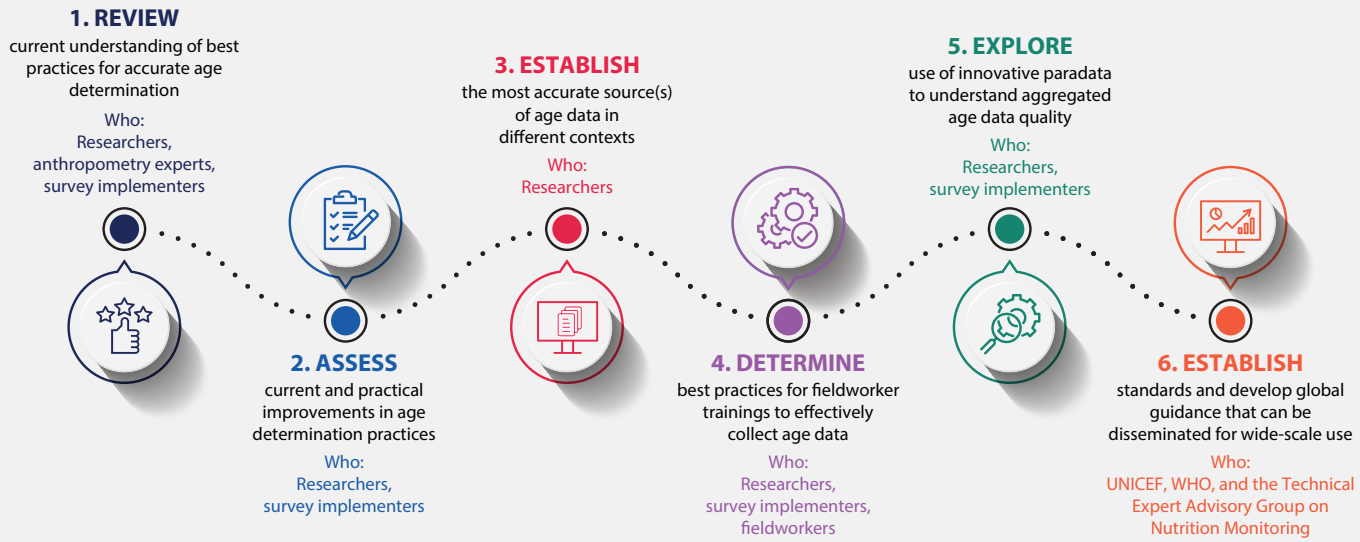
Type of research
Primary data collection: Pilot the use of DOB-related paradata, such as confidence ranking for DOB accuracy provided by the recall respondent and/or the enumerator.

Outcomes
Primary: Validity of the paradata to predict age data quality and use of paradata to make course-corrections in real-time.

Data source(s)
Small-scale studies or surveys (multi-centre or individual) in a field environment from diverse settings, including in low- and middle-income countries.

Research roadmap

Anthropometric indices of weight-for-age and height-for-age are highly sensitive to the precise age of the child; therefore, inaccuracies in the birth date of the child directly impact the rigour and reliability of childhood malnutrition estimates. Identification of best practices and improvements to these practices for DOB collection are a high priority. A roadmap towards the establishment of global guidance on this topic is presented below.



[1] Quezada AD, García-Guerra A, Escobar L. Bias correction of nutritional status estimates when reported age is used for calculating WHO indicators in children under five years of age. Salud Publica Mex. 2016 Jun;58(3):351-7. doi: 10.21149/spm.v58i3.7894. PMID: 27598932

[2] Food and Agriculture Organization of the United Nations. Guidelines for estimating the month and year of birth of young children. Rome: FAO; 2008, <https://www.unscn.org/web/archives_resources/files/Guidelines_for_estimating_the_month_463.pdf>.

[3] When seasonal calendars are used to collect data on age in months (rather than month/year of birth), the DOB should be back calculated using age in months and DOI to identify age in days after imputing a day of birth (i.e., use day 15 for all children, random allocation of a day for each child, etc.).

[4] WHO recommendations on home-based records for maternal, newborn and child health. Geneva: World Health Organization; 2018. Licence: CC BY-NC-SA 3.0 IGO305.

[5] Torun, Huzeyfe and Tumen, Semih, The Empirical Content of Season-of-Birth Effects: An Investigation with Turkish Data. IZA Discussion Paper No. 10203, SSRN: https://ssrn.com/abstract=2840147 or http://dx.doi.org/10.2139/ssrn.2840147

[6] Larsen, A.F., Headey, D. & Masters, W.A. Misreporting Month of Birth: Diagnosis and Implications for Research on Nutrition and Early Childhood in Developing Countries. Demography 56, 707–728 (2019), <https://doi.org/10.1007/s13524-018-0753-9>.

[7] Grellety E, Golden MH (2016) The Effect of Random Error on Diagnostic Accuracy Illustrated with the Anthropometric Diagnosis of Malnutrition. PLoS ONE 11(12): e0168585. doi:10.1371/journal.pone.0168585

[8] Bairagi, R., Aziz, K.M.A., Chowdhury, M.K. et al. Age Misstatement for Young Children in Rural Bangladesh. Demography 19, 447–458 (1982), <https://doi.org/10.2307/2061012>.

[9] R Bairagi, R I Ahsan, Inconsistencies in the findings of child nutrition surveys in Bangladesh, The American Journal of Clinical Nutrition, Volume 68, Issue 6, December 1998, Pages 1267–1271, <https://doi.org/10.1093/ajcn/68.6.1267>.

[10] Oshaug A, Pedersen J, Diarra M, Bendech MA, Hatløy A. Problems and pitfalls in the use of estimated age in anthropometric measurements of children from 6 to 60 months of age: A case from Mali. J Nutr 1994; 124:636-644, <www.researchgate.net/publication/15028916_Problems_and_Pitfalls_in_the_Use_of_Estimated_Age_in_Anthropometric_Measurements_of_Children_from_6_to_60_Months_of_Age_A_Case_from_Mali>.

[11] The prevalence of missing day, month and year of birth and poor data quality in population-based surveys and changes over time are addressed in the brief on Thresholds for anthropometric data quality indicators in population-based surveys. The impact of DOB quality on anthropometric estimates is addressed the brief on Random and systematic error in anthropometric estimates in population-based surveys.

[12] In Demographic and Health Surveys, the age of children in the household is asked in the household questionnaire, while the DOB followed by age are asked in the Woman's Questionnaire. A consistency check between DOB and age in CAPI are performed but there is currently no additional consistency check between age in the household questionnaire and age in the Woman's Questionnaire.

If interested in joining this effort

or if you have any questions or comments, please contact the TEAM Working Group on Anthropometric Data Quality at: nfsdata@who.int and sdg2.2@unicef.org.