Technical briefing paper

Lessons learnt from the Wasting and Stunting Technical Interest Group (WaSt-TIG)

Research methods for studies looking at the relationship between wasting and stunting

February 2022

USAID FROM THE AMERICAN PEOPLE

ENN

Irish Aid
Rialtas na hÉireann
Government of Ireland

Authors

Philip James and Tanya Khara

Acknowledgements

This technical briefing paper has been informed by inputs from, and discussion with, members of the Wasting and Stunting Technical Interest Group (WaSt-TIG). A special thanks go to all WaSt-TIG members, who have given their time and focused their experience and considerable intellects on the topic, and who have also reviewed early drafts of this paper: Jeanette Bailey, Zulfiqar A. Bhutta, André Briend, Sheila Isanaka, Marko Kerac, Natasha Lelijveld, Andrew Mertens, Mark Myatt, Sophie E Moore, Martha Mwangome, Kieran S. O'Brien, Kate Sadler, Natalie Sessions, Susan Thurstans, Patrick Webb, Jonathan C. Wells and Stephanie V. Wrottesley. Thank you to Marie McGrath for the additional review.

Thank you to all members of the WaSt-TIG whose research over the years has collectively informed the content of this paper. We also acknowledge the huge body of work that has been undertaken globally linked to or outside of the specific activities of the WaSt-TIG group. All this has contributed to our collective understanding of best practices and knowledge gaps within research practices.

We were able to undertake this work thanks to the generous support of the USAID's Bureau for Humanitarian Assistance and the Department of Foreign Affairs, Ireland. The ideas, opinions and comments included here are entirely the responsibility of the document's authors and do not necessarily represent or reflect the policies of the donors.

Designed by Chris Woodrow

More information

More information on the work of the WaSt-TIG can be found here:
https://www.ennonline.net/ourwork/reviews/wastingstunting

We encourage readers that have further specific questions and suggestions on research methods relevant to the work of the WaSt-TIG to contact the WaSt team directly at philip@ennonline.net.

Recommended citation

Philip James and Tanya Khara (2022). Research methods for studies looking at the relationship between wasting and stunting: Lessons learnt from the Wasting and Stunting Technical Interest Group (WaSt-TIG). ENN: Kidlington, UK.

Contents

Summary of key points covered in the technical brief by section

Section in technical brief	Key considerations
Cross-sectional data	· Cross-sectional data provide a snapshot of a given situation at one point in time · Relationships between exposures and outcomes can be investigated for their strength of associations, but causality cannot be determined · Repeated cross-sectional (panel) data can enable an exploration of time trends · This data can be relatively quick to collect, but does not allow for an estimation of incidence, is limited in its ability to explore seasonality, and does not enable the process of wasting and stunting to be explored · The Wasting and Stunting Technical Interest Group (WaSt-TIG) have utilised survey data in particular · Cross-sectional data has been useful to generate hypotheses that have been further explored using different study designs
Longitudinal data	· These record characteristics from a group of individuals followed over time, with repeated measurements in the same people · This type of data can capture incidence and potential predictors of outcomes, but are still subject to bias and do not guarantee that causality can be assessed · Longitudinal cohort studies allow for investigations of multiple exposures with multiple outcomes, a dynamic exploration of the interplay between exposure and outcome over time, and an investigation of seasonal trends · They are prone to loss-to-follow-up, may be expensive to conduct, and may have limitations for generalisability · The WaSt-TIG have used population cohorts and cohorts from nutrition treatment programmes
Outcomes: mortality	· The ultimate functional outcome of interest is mortality · Mortality is a rare outcome · Cause of death is often captured bluntly, making it difficult to attribute the full contribution of nutritional deficits to risk of death
Outcomes: anthropometric	· Anthropometric criteria are imperfect proxies that help screen people for risk of mortality and adverse outcomes, but in themselves are not necessarily the primary outcome (which is often mortality) · Anthropometric outcomes are markers of risk on the pathway to poor functional outcomes · It is important to consider the process that leads to an individual becoming wasted or stunted, and not just a focus on anthropometric cut-offs · To enable comparison between studies, clear case definitions are required for incident cases, relapse, persistent wasting, anthropometric recovery, episodes of wasting and stunting and concurrent wasting and stunting (WaSt) · Low weight-for-age is helpful to consider in study design as it captures children who are concurrently WaSt (those at elevated risk of mortality), including infants under 6 months of age · Anthropometric recovery does not necessarily mean that functional recovery (immune function, cognitive function, the ability to grow and stay well) has been achieved; anthropometric outcomes should be complemented by other measures of immediate and longer-term functional recovery · The timing of outcome measurement matters, as peaks in wasting, stunting, and concurrent WaSt vary

Section in technical brief	Key considerations
Choice of exposures	· Given the interplay between wasting and stunting over time, anthropometric outcomes can also act as exposures of interest · Common exposure variables, beyond anthropometry, are limited to age and sex in many datasets · Other useful exposure variables to collect include socio-economic indicators, food security, biochemical measures of nutrition status, measures of infection and inflammation, parental characteristics (especially maternal nutrition and health) and gestational age
Considering seasonality	· Considering seasonality is essential for understanding the pattern of both exposures and outcomes · There is no single pattern of seasonality; the seasonality of a particular setting can vary year on year and is constantly feeding in multiple stresses into different pathways · Seasonality can sometimes be estimated by rainfall and temperature records · Seasonal trends are fluctuations around a mean that are cyclical in nature; these can be modelled in many different ways
Study duration	· Although dependent on research questions, generally for population cohorts longitudinal birth cohorts (covering pre-natal characteristics) followed up at least to 24 months are ideal · For studies using programme data from nutrition treatment programmes, a minimum follow-up of six months is suggested to capture the period of the highest risk of relapse
Data collection frequency	· Generally, more frequent data collection is better for greater precision around the timing of exposures and outcomes · Monthly data collection is particularly useful · Increased data collection frequency is only useful if quality can be maintained · Frequency will also depend on the definition of certain exposure variables, e.g. the minimum interval required to capture an episode of growth faltering · Not all exposure variables may need to be collected at each data collection time point · Integration of study design with existing demographic surveillance systems can be beneficial if data quality is good
Pooling datasets	· Mortality, severe wasting and concurrent WaSt are examples of relatively rare outcomes, requiring proper sample size calculations to ensure their detection · Pooling datasets is a common strategy to reach the sample sizes needed and improve interpretations of generalisability of findings · Successful pooling requires standardisation of datasets and proper permissions obtained
Data cleaning	· Survey data can be cleaned by using flags for extreme, likely implausible, values · Common data cleaning protocols for surveys include World Health Organization (WHO) and Standardised Monitoring and Assessment of Relief and Transitions (SMART) flags · The type of data cleaning protocol used can influence outcome estimates · Data cleaning criteria for data from clinical and treatment programme and research settings can be less clear-cut. These datasets often contain very sick children where some 'implausible' values may well reflect the reality · Other data cleaning methods to consider include defining internal cut-offs, and pre-establishing acceptable variation between data collection timepoints · Looking at digit preference and the shape of the distribution provide further ways to assess data quality
Systematic reviews	· Systematic reviews can be very useful for summarising the existing evidence base, exploring the extent that certain observations have been seen in different contexts, and generating hypotheses for future research questions · It is recommended to build on search terms used in previous reviews because search criteria that allow for wasting and stunting terms to be searched separately (rather than being addressed within the same article) will lead to an overwhelming number of articles being returned

Introduction

The Wasting and Stunting Technical Interest Group

In 2013, the Emergency Nutrition Network (ENN) started to explore the separation between wasting and stunting in evidence generation, programming and policy. The aim was to better understand the complex relationships and associations between wasting and stunting in infants and young children, and examine whether current separations were justified or useful for achieving the goals of improving nutritional status and decreasing mortality risk in vulnerable populations. To facilitate this process, ENN set up the Wasting and Stunting Technical Interest Group (WaSt-TIG), a group of 41 volunteer expert researchers, programmers and donors in the fields of child growth, nutrition and epidemiology. The WaSt-TIG began by examining the existing evidence on the relationship between these two outcomes of undernutrition, identifying and prioritising gaps in evidence, and then set about filling in some of those evidence gaps. Since 2014, the project has gone through three phases, and is currently in its fourth phase of implementation:

- Phase 1. 2014–15: Reviewing existing evidence, defining and prioritising the gaps
- Phase 2. 2016–17: Exploring existing datasets to investigate associations between wasting and stunting and the implications for individuals
- Phase 3. 2018–19: Delving deeper and communicating what had been learnt to date
- Phase 4. 2020–date: Exploring implications for policy and practice, for both programming and research

Aim and intended audience

The aim of this technical brief is to capitalise on and share the extensive experience of the WaSt-TIG in scrutinising data through the lens of the relationship between wasting and stunting in the hope that it will inspire and facilitate others to do the same. Data used by the group include multiple datasets drawn from diverse settings and which have been collected with a variety of objectives, methodologies and study designs.

As a result of the work of the WaSt-TIG, there are a number of lessons that have been learnt that may help researchers and programmers involved in research, in their research methodologies. These are relevant not only for the design and planning of new studies, but also for the analysis of *existing datasets* that can be looked at through the lens of wasting and stunting.

Overview of technical brief

In this technical brief, we start with an overview of the main characteristics of cross-sectional and longitudinal data, and how these different types of data have been used by the WaSt-TIG. In doing so, we set out some of the strengths and limitations of both, and discuss which research questions lend themselves to the different types of data. We focus on these two types of data as these comprise the study designs that have been predominantly used by the WaSt-TIG to date. This is not to suggest that other types of studies (e.g. case-control studies, intervention studies and meta-analyses) are unimportant, and we trust that the lessons learnt will be helpful for all types of studies looking at wasting and stunting.

In the second part of the technical brief, we summarise some of the learning and best practices arising from the collective experience of the WaSt-TIG, focusing on the choice of outcomes and exposures, consideration of seasonality, duration of studies, frequency of data collection, pooling datasets and data cleaning. We finish with reflections on systematic reviews, and then provide short overall concluding thoughts.

Box 1 provides an overview of key epidemiological concepts that are referred to in this technical briefing paper.

Box 1: Key epidemiological concepts

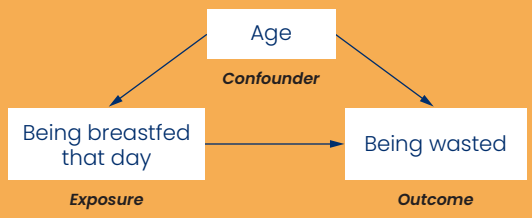
Studies can be categorised into two broad groups: those that are *observational* and those that are *experimental*.

Observational studies: These are non-experimental, meaning that no interventions are given. Participants are simply observed and/or have measurements taken without any attempt to influence the outcome. There are three main types of observational studies: cross-sectional, case-control and cohort studies.

Experimental studies: These involve the introduction of an intervention and assess its impact on the study population outcome(s). Those involved in the study are usually randomised into a group receiving the intervention and a group that does not receive it (a control group for comparison).

All study designs can be influenced by various forms of **bias**. This is particularly true of observational studies. Bias leads to a false or incomplete interpretation of the true association between an exposure and an outcome. For example, this can include **selection bias** (e.g. where selection of study participants into a cohort may not be truly random and therefore may not accurately represent the population of interest), **information bias** (where measurements of exposures and outcomes may be inaccurately measured or classified) and **confounding bias** (1).

Confounders are variables associated both with the exposure and with the outcome, and if they are not taken into account, they can distort the interpretation of the data. By way of illustration, Maria is looking at cross-sectional data taken from a survey of infants and children aged 0–59 months. She is interested in the association between being breastfed that day (the exposure) and being wasted (the outcome). Maria finds there is a positive association between being breastfed and being wasted. She knows that this may not necessarily be a causal relationship, and she decides to look at age as a potential confounding variable. She finds that age is indeed associated with the exposure (younger children are more likely to be breastfed than older children) and is also associated with the outcome (younger children are more likely to be wasted than older children). Age is therefore a confounder in this example. If Maria's sample contained more younger infants than older children, age could help to explain the association observed between being breastfed and being wasted.



There are many useful resources that go into a more detailed overview of the above concepts. Readers wanting more information are directed to resources such as Pearce (2012) (2), Delgado-Rodríguez et al. (2004) (1) and McNamee (2003) (3). There are also free online courses that cover an introduction to epidemiology.[1]

[1] Many such courses can be accessed here: https://www.coursera.org/search?query=epidemiology&

Types of data and their use by the WaSt-TIG

Cross-sectional data

Key characteristics of cross-sectional data

Cross-sectional data provides us with a snapshot of a situation within a specific population in a given location at one point in time (4). Most studies using single time point cross-sectional data are observational and enable an estimation of the prevalence of the outcome of interest at a given point in time (e.g. wasted, stunted or concurrently wasted and stunted children) (2).

If exposure variables are collected within cross-sectional studies, this allows for the investigation of the direction[2] and strength of associations between certain exposures and outcomes. These are often called 'prevalence case-control studies' (2) and look at the difference in outcomes between exposed and non-exposed study subjects. In cross-sectional studies, associations between exposures and outcomes are not necessarily causal. As with other observational studies, cross-sectional studies are prone to confounding and other types of bias. Box 2 provides further information on ways to assess the strength of evidence for a causal relationship.

[2] By 'direction' we are referring to whether the association between the exposure and outcome is positively or inversely correlated. Whether the association is positive or inverse tells us nothing about the timing of exposure and outcome, nor about the strength of claims of causality (see Box 2).

Box 2: Assessing strength of evidence for causality using the Bradford Hill criteria

The Bradford Hill criteria were set out in an address in 1965 (5) to provide researchers with a flexible set of criteria to help assess how likely it is that associations found in data, investigating a disease outcome and various exposures, are causal in nature.

1. **Strength of the association:** Small effect sizes do not necessarily imply there is no causal link between the exposure and outcome, but generally larger effect sizes are more likely to strengthen claims of causality.
2. **Consistency:** This is the extent to which findings have been repeatedly seen in different contexts (different places, investigators, circumstances and periods of time).
3. **Specificity:** The likelihood of causation is greater the more specific the outcome (disease) and the sub-group of people affected (population), and where there are few other likely pathways that might explain the association between the exposure and outcome.
4. **Temporality:** It must be considered whether the exposure of interest occurred before the measurement of the outcome, or whether there is possibility for reverse causality (i.e. the outcome affecting the exposure).
5. **Biological gradient:** This is also known as dose-response. If the likelihood of an outcome increases with greater exposure, this generally strengthens claims of causality.
6. **Plausibility:** This is the extent to which the association between the exposure and outcome is biologically plausible, with the caveat that new discoveries going beyond current knowledge are always possible.
7. **Coherence:** This is the triangulation of findings using different sources of evidence, for example, the convergence of epidemiological and laboratory findings.
8. **Experiment:** This is the extent to which an association between an exposure and outcome has been confirmed in experimental (intervention) studies.
9. **Analogy:** This is whether there are other associations between similar exposures or outcomes.

There have been many scientific advances since 1965, with examples provided by researchers on how the criteria can be applied in the face of modern molecular techniques (6), for example, but to this day they remain helpful aspects to consider when assessing causation.

Repeated cross-sectional data taken within the same population at different points in time,[3] sometimes referred to as panel data, enable a description of how change occurs over time at the population level, although not at the individual level (4). Note that this design is not restricted to observational studies, as cluster-randomised trial designs may use repeated cross-sectional data in an experimental context where subjects are assigned to different exposures.

Studies using cross-sectional data can be relatively quick to conduct, and they enable multiple outcomes and exposures to be investigated. As such, they are often used in the first stages of research to explore hypotheses of mechanisms and associations that could be further explored in other study designs, such as longitudinal or intervention studies.

Other key limitations of cross-sectional data within the context of exploring the relationship between wasting and stunting include:

- The inability to estimate incidence of wasting and stunting, which prevents estimation of the true burden of the outcome. For example, particularly in the context of wasting, the use of prevalence data alone underestimates the burden, and a context-specific incidence correction factor is required to generate more realistic estimates (7).
- The difficulty of establishing seasonal trends in wasting and stunting. The effect of seasonality on estimates of undernutrition is well-established (8-11). Although repeated cross-sectional data allow for a limited exploration of seasonality, the identification of seasonal peaks and patterns of co-variation over the year requires longitudinal datasets (12).
- The inability to explore the process of wasting and stunting, i.e. how and why anthropometric status changes over time, as cross-sectional data enables only the measurement of an individual's status at one point in time.

The use of cross-sectional data in the WaSt-TIG

The WaSt-TIG started by exploring existing cross-sectional datasets, first, because they were easily accessible to the group, and second, because the lens of the relationship between wasting and stunting was new and so it was felt that a number of questions could be at least partly answered with this type of data. Analysing cross-sectional data was felt to represent a useful starting point for the WaSt-TIG to help guide which remaining questions needed to be asked of existing longitudinal data, and then which questions needed further primary research.

Furthermore, even if a question had been looked at cross-sectionally in one context, it remained useful to look at it at in another context for the triangulation of research findings.

Much use has been made of survey data for this hypothesis-generating phase of the work. Common types of nutrition surveys using standardised sampling and data collection methodologies include Demographic and Health Surveys (DHS), Multiple Indicator Surveys (MICS), specific national nutrition surveys[4] and Standardised Monitoring and Assessment of Relief and Transitions (SMART) surveys. DHS and MICS surveys tend to have better national and sub-national representation but less precision around nutrition outcome estimates, whereas SMART surveys tend to have better precision around nutritional outcome estimates but may have limited national representativeness since they often look at sub-national or small-scale levels (13). However, national-scale SMART surveys have also been conducted (14). The WaSt-TIG have used all types of surveys in their analyses (15-17), have pooled data from DHS and MICS (15, 16), and have analysed SMART surveys separately (17).

Selected papers in which the WaSt-TIG have used cross-sectional data are summarised in **Table 1** and the main research questions that the cross-sectional data have been used to answer are presented in **Table 2**. Examples include:

- An early analysis of data from 560 nutrition surveys in Asia and Africa looked at the extent to which mid-upper arm circumference (MUAC) measurements could identify children who were stunted as well as wasted (12).
- DHS surveys from Bangladesh, the Democratic Republic of the Congo, Ethiopia, Pakistan and Nigeria were used to estimate the burden of children who were concurrently wasted and stunted in those countries (18), estimates that had previously never been calculated. Additional analyses using DHS and MICS surveys from 84 countries further explored the global burden of concurrence and were stratified by age and sex (15, 16).
- SMART surveys from 51 countries were used to examine how best to identify children who were concurrently wasted and stunted, the severity of wasting and stunting in those children, and the overlap between wasting, stunting, and underweight measures in children (17).

There is general consensus from the WaSt-TIG that while cross-sectional data were extremely useful in the first two phases of the project, generating many hypotheses that have subsequently been explored in other study designs, there is likely limited utility in continued exploration of survey data unless new hypotheses come to light.

[3] Different people may be sampled at these different time points, but the overall sample population is the same.

[4] For example, the most commonly used nutrition data in Pakistan are from the serial national nutrition surveys in 2001, 2011 and 2018.

Longitudinal cohort data

Key characteristics of longitudinal cohort data

Longitudinal cohort datasets record characteristics from a group ('cohort') of individuals that have been followed over time, with repeated measurements/observations in the same people (19). In these datasets, it is possible to investigate people who are exposed or unexposed to a certain factor, and then to estimate the subsequent risk of developing an outcome of interest.[5]

Unlike cross-sectional datasets, longitudinal datasets enable the calculation of the incidence of an outcome (new cases over time), rather than only the prevalence of an outcome. They also enable the exploration of potential predictors/drivers of an outcome, due to the ability to know that a certain exposure occurred before the outcome of interest. However, while they can account for more confounding than cross-sectional datasets, they are still subject to bias and it not possible to make definitive causal claims. As described above for cross-sectional studies, using the criteria in Box 2 can help in assessing whether it is likely that causal relationships are present (21).

Longitudinal cohort studies have the advantages of allowing the investigation of multiple exposures with multiple outcomes, allowing for a more dynamic exploration of the interplay between exposure and outcome over time, and enabling an investigation of seasonal trends if follow-up investigations capture different seasonal windows. However, they are prone to loss to follow-up of individuals (introducing potential bias if those remaining in the study are different to those who have been lost to follow-up[6]) (19). There are also important limitations related to the representativeness of the cohort to the general population. The generalisability of longitudinal cohorts to the wider population depends on several factors. For example:

- How similar those enrolled into the cohort are to the wider population: population cohorts may share more similar characteristics of the general population than, for example, cohorts comprised of sick children from nutrition treatment programmes, but even then, there may still be important limitations to generalisability. For example, some population cohorts may be limited to a specific geographical location that is more accessible to researchers, or enrolment into a population birth cohort may depend on the referral of births at health clinics rather than births occurring home.
- How the study was conducted may influence representativeness: for example, a closely monitored population cohort may involve better referral and healthcare provision for participants than in the general population, and therefore underestimate adverse outcomes compared to the general population.
- While population cohorts of entirely untreated children are very useful for exploring scenarios in the absence of interventions, and especially valuable for investigating associations between anthropometric indicators and mortality, given the widespread availability of treatment for severe wasting, these datasets are rare and relatively old.

Use of longitudinal cohort data by the WaSt-TIG

The WaSt-TIG have used two main types of longitudinal cohort datasets in their research to date:

1. **Population cohorts:** These have comprised a defined group of people, usually from community rather than clinical settings, and have often made use of existing demographic health surveillance data.
2. **Cohorts from nutrition treatment programmes:** These are often datasets from Community-based Management of Acute Malnutrition (CMAM) programmes, comprising individuals admitted for the treatment of severe wasting and monitored regularly until discharge (and in some studies then subsequently followed up post-discharge).

These two types of cohorts have been used to explore different questions pertaining to the relationship between wasting and stunting by the WaSt-TIG (selected papers and questions summarised in Table 1 and Table 2).

For example, **population cohorts** have been used to:

- investigate which anthropometric indices are independently associated with near-term mortality in untreated[7] children in Senegal (22), and in a later meta-analysis of 12 cohorts of untreated children (23);
- explore the effect of age and sex on the relationship between anthropometry and mortality (24, 25);
- explore the effect of seasonality on episodes of becoming wasted, and the interrelationships between early growth trajectories and subsequent risks of becoming wasted or stunted in The Gambia (11); and
- build hypotheses around best anthropometric criteria for identifying high-risk children for treatment in CMAM programmes, and modelling the potential implications for programme size and workload (23).

[5] Cohorts can be either prospective or retrospective. In **prospective** studies the individuals are enrolled into the cohort, have exposure variables measured at baseline and are followed up over time. At the time of enrolment, they have not yet developed the outcome of interest. In **retrospective** cohorts the individuals have already developed the outcome of interest, and the researchers look back in time to assess the exposure variables. More details on these definitions, with illustrations, can be found in Euser et al. (2009)(20).

[6] However, this potential bias can be described accurately by comparing those followed versus not.

[7] The data were collected before therapeutic feeding services for severe wasting were widely available in those contexts; 'untreated' here therefore refers to children not receiving the package used today to treat severe wasting using a CMAM approach.

Programme cohorts have been used to:

- explore whether treatment for severe wasting enables a restarting or acceleration of linear growth (12, 26);
- determine the proportion of children admitted for wasting treatment who are stunted (12); and
- explore how children with concurrent wasting and stunting (WaSt), with a focus on those who are severely underweight, respond to treatment (an analysis of their growth trajectories and outcomes) to inform future intervention study protocols aimed at testing the level of treatment intensity needed for this group (27).

Table 1: Selected work from the WaSt-TIG exploring associations between wasting and stunting

Phase	Paper description	Data sources	Type of data	Research questions investigated
1	A technical briefing paper, including a narrative review and analyses of various types of data to summarise associations between wasting and stunting in programme datasets (12)	560 nutrition surveys in Asia and Africa CMAM programme data from Malawi and the Democratic Republic of the Congo	Cross-sectional, survey Longitudinal, programme	· Can MUAC identify children who are stunted as well as wasted? · Does treatment for wasting enable a restarting or acceleration of linear growth? · What proportion of children admitted for wasting treatment are stunted?
1	An analysis of five high burden country datasets to estimate the burden of concurrent WaSt for the 2015 Global Nutrition Report (18)	DHS surveys from Bangladesh, the Democratic Republic of the Congo, Ethiopia, Pakistan, Nigeria	Cross-sectional, survey	Burden of concurrent WaSt
2	A re-analysis of DHS and MICS datasets from 84 countries to generate a pooled prevalence estimate of the burden of concurrent WaSt in those countries (15, 16)	DHS and MICS surveys from 84 countries	Cross-sectional, survey	Burden of concurrent WaSt, stratified by age, gender and United Nations region
2	An analysis of cross-sectional datasets (SMART surveys) to examine concurrence in more detail and look at how best to identify concurrently wasted and stunted children (17)	Cross-sectional surveys using SMART methodology from 51 countries	Cross-sectional, survey	· Exploring the degree of overlap between wasting, stunting and underweight children in surveys · The association between being wasted and being stunted · The severity of wasting and stunting in WaSt children · The prevalence of WaSt by age and sex · Identifying the weight-for-age z-score and mid-upper arm circumference thresholds that best detect cases of WaSt

Phase	Paper description	Data sources	Type of data	Research questions investigated
2	An analysis focusing on identifying those children at most risk of mortality using cohort data from Niakhar, Senegal (22)	Community-based cohort study, untreated children, Senegal	Longitudinal, research cohorts	· Which anthropometric indices are independently associated with near-term mortality
2	An additional paper which specifically explores the patterns of concurrent WaSt in children under five years of age in the Niakhar, Senegal cohort, highlighting the increased risk in boys (24)	Community-based cohort study, untreated children, Senegal	Longitudinal, research cohorts	· Prevalence of wasting, stunting and interaction · Effect of age, sex and anthropometric indicators on the prevalence of concurrent WaSt · Effect of age, sex, stunting, wasting, and interaction on mortality
3	Data from the Medical Research Council Unit The Gambia surveillance programme has been another key focus on the WaSt work, with detailed analysis carried out on cohorts of 0–24-month age groups in relation to seasonality and growth, wasting as a risk factor for stunting and vice versa (11)	Retrospective cohort analysis, based on growth-monitoring records spanning four decades from clinics in rural areas in The Gambia	Longitudinal, research cohorts	· Age-related differences in the prevalence of stunting, wasting or concurrence · Estimate whether infants who were wasted in their first wet season tended also to be wasted in the subsequent wet season · Early growth trajectories and subsequent risks of stunting · Individual weight-for-length trajectories as predictors of subsequent risk of stunting · Longitudinal prediction of stunting through the use of time-lagged wasting
4	A multiple cohort individual-participant data meta-analysis of the risk of mortality (within six months) by anthropometric case definition (23)	Retrospective cohort analysis, based on 12 community-based cohorts of untreated children	Longitudinal, research cohorts	· Which anthropometric diagnostic criteria best identify children with a high risk of death? Stratification by age · Furthermore, how suitable are these criteria for use as a case-finding and admission criteria for therapeutic feeding programmes? · Hypotheses for implications for programme intensity requirements by considering the risk of mortality by different anthropometric categories

Phase	Paper description	Data sources	Type of data	Research questions investigated
4	An extended analysis based on the above study (23) to further investigate variations in mortality risk associated with anthropometric deficits in by age and sex, with an aim of informing programming and policy decision making (in progress) (25)	Retrospective cohort analysis, based on 12 community-based cohorts of untreated children	Longitudinal, research cohorts	· Assess the effect of sex and age and geographical region on mortality associated with a number of anthropometric case definitions
4	An exploration of how severely underweight and wasted children respond to treatment using a pooled secondary data analysis to inform future intervention studies (in progress) (27)	17 CMAM datasets	Longitudinal, programme data	· How children with WaSt, with a particularly focus on low weight-for-age z-score (WAZ), respond during wasting treatment programmes, including those who are considered "non-responders" · To explore WAZ growth trajectories of children receiving treatment for wasting in CMAM programmes · To explore how WAZ relates to other growth trajectories (MUAC, weight-for-length z-score, length-for-age z-score) in children receiving treatment for wasting in CMAM programmes · To explore WAZ and other growth trajectories among children whose MUAC does not reach the criteria for being discharged as "cured" in order to inform future discharge criteria

Table 2: Examples of research questions looked at by the WaSt-TIG and types of datasets that can and cannot answer them

Examples of research questions	Cross-sectional	Longitudinal (population cohort)	Longitudinal (CMAM programme data)
What is the prevalence of wasting, stunting and concurrent WaSt?	✔	✔	✔
Is the prevalence of concurrent WaSt greater than could be expected if wasting and stunting were independent phenomena? In other words, is there evidence for a direct relationship between the two?	✔	✔	✖
Does wasting appear to be associated with stunting at one point in time, and vice versa?	✔	✔	✔
What are the environmental, socio-economic, demographic, household characteristics **associated with** wasting, stunting and concurrent WaSt?	✔	✔	✖
What are the environmental, socio-economic, demographic, household characteristics **predictive of (i.e. drivers of)** wasting, stunting and concurrent WaSt?	✔	✔	✖
What anthropometric (diagnostic) criteria best capture children concurrently wasted and stunted?	✔	✔	✔[8]
What is the influence of seasonality on wasting, stunting and the relationship between the two?	✔ (Repeated cross-sectional)	✔	✔
What is the association between earlier anthropometric status (birth anthropometry, process/episodes of wasting/stunting) and later experiences of nutritional deficits, morbidity and mortality?	✖	✔	✖
How does being treated for wasting affect linear growth?	✖[9]	✖[9]	✔
Could children with certain anthropometric criteria (or combinations of criteria) at admission to wasting treatment programmes be treated effectively with different intensities of intervention?	✖	✖	✔[10]
Sex differences and stratification by age can be carried out for all the above questions (dependent on power)	✔	✔	✔

[8] Note this will be a biased cohort of children already in a therapeutic feeding programme, so care should be taken with generalising to the wider population.

[9] Unless the cross-sectional or longitudinal datasets contain a subset of children who were previously treated for wasting.

[10] Programme data can help build hypotheses around the best anthropometric criteria for identifying high risk children for treatment in CMAM programmes, and modelling the potential implications for programme size and workload. However, an experimental study, such as a randomised controlled trial, would then be needed to test the hypotheses.

Learning from WaSt-TIG on research methods

Choosing and defining outcomes

As with all research studies, outcomes need to be clearly defined, with details of how they are calculated, and communicated in language that is accessible and consistent with previous studies (where possible). The primary functional outcome of interest is mortality. Anthropometric measures reflect potential exposures on the pathway to mortality and are therefore considered as outcomes in many studies. For consistency, we also refer to anthropometric outcomes for the purpose of this paper.

Mortality

Mortality is a rare outcome in many datasets, especially where sample sizes are small and where treatment is offered to individuals. For studies that do investigate mortality as an outcome, it can be challenging to obtain a cause of death (28). Most of the WaSt-TIG studies looking at mortality have only been able to consider all-cause mortality: categories of mortality causes, where they are collected, tend to be very broad. For those starting new cohorts, training on obtaining verbal autopsies will help provide the necessary context to determine the likelihood of deaths being nutrition-related or not (29, 30).

Anthropometric outcomes

The process of wasting and stunting

When considering the choice of outcomes, it is common to focus on a fixed anthropometric cut-off, such as those used to define a wasted, stunted or concurrently wasted and stunted child. There is increasing recognition, however, that it is important to consider the *process* that leads to an individual becoming wasted or stunted, i.e. the process of *wasting* and *stunting*. In the latest *Lancet* (2021) maternal and child nutrition update series, Victora et al. presented the distribution of height-for-age z-score (HAZ) and weight-for-height z-score (WHZ) curves from the DHS included in the paper's analysis, and how they compared to the WHO (2006) standards (31). In this paper, the DHS surveys had HAZ and WHZ distributions that were shifted to the left of the standard growth curves. This meant that populations represented by the DHS surveys not only had increased proportions of stunted and wasted children compared to well-nourished populations, but also that the *whole population* had a lower WHZ or HAZ than if they had been well-nourished, i.e. even those children who seemingly fell in the non-wasted or stunted categories were growing below their potential. This reinforces a need to focus on the process of wasting and stunting, on how many children grow below their potential and the implications of this, rather than solely on those who reach the classification of being wasted or stunted (see **Box 3** for an illustration of this).

Box 3: An illustration of how to consider the process of wasting

Take the case of a boy, Yusuf, who was measured at the start of July with a healthy WHZ = +2.1. Shortly afterwards, Yusuf became ill and started to lose weight. At the start of September, Yusuf was measured again and was found to have WHZ = -1.5. At this stage, he would still not be classified as being wasted, as his WHZ is above the cut-off of WHZ <-2. However, he has been through a process of wasting, potentially with implications for later linear growth or other outcomes. It is therefore the *trajectory* (and change from baseline) that could be of interest from a research and programming perspective, alongside the considerations of specific anthropometric cut-offs.

Considering incidence

Researchers should maximise the utility of longitudinal datasets by not limiting analyses to wasting and stunting at one point in time, but also by considering cumulative incidence / incidence proportion (*new cases* in a time period). In the case of the WaSt-TIG, this is important as capturing incidence enables a better estimate of the true burden of wasting and stunting in the population, compared to capturing information on prevalence alone. Having a more accurate picture of the burden enables better caseload projections for programmes, and enables a more realistic exploration of the relationships between wasting and stunting. For example, at the population level it has been shown that using only data on the prevalence of severe wasting vastly underestimates the true burden, and that the extent of the underestimation varies markedly by country (e.g. underestimating the burden by 1.3 times in Niger, and by 30.1 times in Burundi) (7).

Defining relapse, persistent wasting and episodes

To understand the experience of wasting and stunting in individual children, and the implications for different outcomes, it is important to look at whether episodes of being wasted are repeated over time (*relapse cases*). In order to do this consistently, a decision is required on the length of time children need to be 'recovered' before being considered a relapse case. For example, some studies use a minimum 60-day period of recovery before a relapsed case is recorded (**32**).

In some studies, a definition of *persistent wasting* has been adopted as a way of quantifying those children who dip in and out of being wasted (doing so more frequently than could be captured by the definition of relapse above). A definition of persistent wasting needs to be decided in advance, and depends on the study timeframe and frequency of measurements. For example, Mertens et al. (2020) in their meta-analysis of cohorts decided that persistent wasting would be defined as those that have at least half of their WHZ measurements falling below -2 over the first two years of life, where measurements were taken at least monthly (**32**). While different studies may use different definitions, the most important aspect is for the definition to be clearly stated, to facilitate interpretation of results between studies.

For population cohorts looking at episodes of wasting and stunting researchers will need to decide how to designate the start and end of a unique episode (the episode period). This will depend on how frequently measurements are collected (see section below), with some studies using the mid-point of two time points to estimate this.[11]

Concurrent WaSt

For studies looking at individuals who are concurrently WaSt (whether as an outcome or an exposure), a clear definition is particularly important, as those who are categorised as wasted by a low MUAC definition will likely not be the same individuals as those categorised as wasted by low weight-for-length/height (**33**-**35**). In most studies, the definition for concurrent WaSt has been those individuals with a weight-for-length/height z-score (WLZ/WHZ) <-2 and a length-for-age/height-for-age z-score (LAZ/HAZ) <-2.[12] **Box 4** provides the rationale for this and explains why many subsequent studies have continued to use this definition, despite the increasingly widespread use of MUAC.

While concurrent WaSt has been an important focus in many of the studies that the WaSt-TIG have supported (Table 1), a key advantage of longitudinal analyses is the opportunity to improve the understanding of how episodes of wasting contribute to the aetiology of stunting and, likewise, how stunting (already evident at birth or later episodes) contributes to the risk of later wasting. As such, concurrence of wasting and stunting may only be part of the picture to consider, and a longer-term exploration of the interplay between stunting and wasting (i.e. with both as potential outcomes and exposures) may be relevant. Indeed, the WaSt-TIG and others have gone on to explore such themes (**11**, **38**).

Box 4: The definition of concurrent WaSt within the WaSt-TIG

A key paper that influenced the thinking behind many of the topics addressed by the WaSt-TIG was the McDonald et al. 2013 meta-analysis of 10 datasets from low- and middle-income countries, which explored the effect of multiple rather than single anthropometric deficits on the risk of mortality in children (39). In this paper, the highest risk of mortality was among children who were concurrently stunted (HAZ<-2), wasted (WHZ <-2) and underweight (weight-for-age z-score [WAZ] <-2); compared with children with no anthropometric deficit, the mortality hazard ratio was 12.3 (95% CI: 7.7, 19.6). The recruitment dates for the included studies ranged from 1977 to 1995; MUAC was not included in the analyses.

Additional analyses using cross-sectional datasets comprising almost 1.8 million children showed that all children with concurrent WaSt were also underweight (17), and therefore that it was sufficient to simplify approaches by focusing on concurrent WaSt rather than the overlap of all three anthropometric deficits, i.e. a focus on concurrent WaSt does not deliberately ignore underweight. Since the WaSt-TIG have wanted to explore how best to capture children at the highest risk of mortality, the same case definition of concurrent WaSt has been commonly used since then, i.e. WHZ/WLZ <-2 and HAZ/LAZ <-2.

Studies have shown that both WAZ and MUAC measures are reasonably good at identifying children with WaSt. Furthermore, a combination of severely low MUAC (<115mm) and severely low WAZ (<-3) criteria capture children at the highest risk of death, including those with concurrent WaSt (17, 22), who would otherwise (due to the number of measurements required) be more difficult to identify.

[11] Investigators will not know the final outcomes of some individuals after the end of the study (right-censoring) (e.g. whether an individual in a current episode of wasting would go on to recover, die or not respond) and therefore a decision needs to be made in advance on how to deal with this. For example, some investigators require that datasets need to have at least half of cases recovered in the study period (32).

[12] From a research point of view, it is important to recognise that in very young (or very stunted) children with a length <45cm, WLZ is not calculated (36) and therefore the presence of concurrent WaSt cannot be determined. Furthermore, some researchers would argue that small-for-gestational age needs to be considered as a separate anthropometric category in the newborn period, with a transition to standard indicators of WHZ and WAZ measurements by three months of age. Given high rates of prematurity in some populations, this may help avoid over-diagnosing wasting in early infancy (11, 37).

Defining recovery

For studies using data from CMAM treatment programmes, as well as those using population datasets that need to define a period of 'normal' growth or 'recovery' after an episode of wasting, it is important to remember there is no gold standard definition of 'recovery'. Anthropometric criteria are imperfect proxies that help screen people for the risk of mortality and adverse outcomes, but in themselves are not outcomes. Anthropometric recovery does not necessarily mean that functional recovery (immune function, cognitive function, the ability to grow and stay well) has been achieved (**40**). It can be challenging to define functional outcomes that can be measured at the population level, both due to lack of data on what other markers would be needed to define functional recovery, and the prohibitive cost and logistics of measuring these at the population level. Where possible, a greater diversity of outcome measures beyond anthropometry will help further knowledge and practice on this issue.

When anthropometric definitions of recovery are used, these are often based on common case definitions, e.g. for children aged 6–59 months, these may include reaching MUAC >125mm and/or WHZ >-2. However, these case definitions are not always standardised across programmes and studies, meaning it is essential for researchers to clearly state their definitions. Furthermore, there is still much learn about anthropometric definitions of recovery. As described in Box 4, it has been increasingly recognised that concurrent WaSt, captured well by screening for low WAZ, identifies children with an elevated risk of mortality. Studies are therefore beginning to explore the implications of this finding for admission and discharge criteria for therapeutic feeding programmes. It might be decided, for example, that children with low WAZ (<-3) should receive therapeutic feeding; however, it might be found that some of these children have a MUAC >125mm (i.e. they are not considered wasted). What would the appropriate definition of recovery be for this subset of children? This is the topic of an upcoming trial led by the International Rescue Committee in Mali, with data collection starting in 2022.

Timing of outcome measurement

Finally, it is important to consider the timing of measurements of outcomes. For example, in a meta-analysis of 10 cohorts, the peak prevalence of wasting was in the first three months of life, whereas peak incidence of wasting was at 12–15 months, and peak stunting prevalence was at 18–21 months (**32**, **41**). Individual studies may need to examine existing data to help decide on the optimal time to measure specific outcomes.

Choice of exposures

Looking at an outcome on its own does not shed light on how to mitigate the risk of that outcome. Exploring exposure variables for both wasting and stunting helps us to understand potential risk factors (including those that are common to both), and therefore also potential pathways to impact that can inform the design and testing of programmatic improvements. Given the ability of longitudinal data to explore the interplay between wasting and stunting, many of the anthropometric outcomes listed in the above section also form exposure variables of interest. A few examples of this include the analysis of whether being wasted in the previous three months is predictive of being stunted at a given time (time-lag analysis), or whether early life trajectories of WLZ (over the first two years of life) are predictive of being stunted 20–24 months of age (**11**). There are many combinations of research questions that fit in this theme, and the above guidance on anthropometric outcomes is also relevant here in this exposure section.

Many datasets that have been used for secondary data analyses by the WaSt-TIG have had limitations in the available exposure data. For example, in the pooled dataset of cohorts informing recent analysis of mortality risk (**25**), there was no consistently available exposure data beyond anthropometry, age and sex. This has meant the investigators have been restricted in the analyses that they can do and what can be adjusted for, in turn involving assumptions being made in attributing risk to a nutritional deficit. For example, these could include assuming all deaths are the result of nutritional deficits, and not being able to quantify the true proportion of deaths caused by accidents or other morbidities unrelated to the nutritional deficits (**25**).

Across the datasets being used, the quality and diversity of exposure data varies in terms of when the data were collected, what the research question was at the time (if using research data), and what resources were available. This, of course, means that current analysis plans are necessarily dependent on the research questions and covariates given priority at the time of the original data collection. This point is particularly relevant when datasets are being pooled to increase sample size (see section below) and researchers are aiming to get consistency across all exposure and outcome variables for meta-analyses.

While the specific exposure variables of interest and other covariates collected will depend on the research question, there are some common variables that consistently appear in the limitation sections of WaSt-TIG papers as variables that would have been useful to have information

on to aid with interpretation and generalisability of findings. Beyond age and sex, which are often captured, examples of some of these additional variables include:

- **Socio-economic indicators:** e.g. household size, measures of wealth
- **Access to healthcare**, including antenatal care
- **Food security:** measures of household and individual food security
- **Nutritional status beyond anthropometry:** dietary diversity and adequacy, biochemical measures of nutrient status, where feasible
- **Measures of infection, inflammation and intestinal dysfunction**
- **Parental characteristics:** anthropometry, dietary patterns, nutritional and health status, education and livelihoods. These may be particularly useful in studies related to wasting and stunting because, for example, they may form possible explanatory factors for why there are sex differences in early childhood nutrition status (42, 43).
- **Gestational age estimates:** to be collected in birth cohorts. These especially useful for understanding growth trajectories in young infants and subsequent risk of mortality, hospitalisation and neonatal complications (44, 45).

Considering seasonality

The effects of seasonality on exposures and outcomes in nutrition studies have been well-described (8, 46-49), although there remain many gaps in understanding. Choosing variables that indicate seasonality can be challenging, especially for pooled analyses involving datasets from different agro-ecological, socio-economic and cultural areas. Often particular months of the year are known to be generally associated with wet and dry seasons (11), although rainfall patterns are not necessarily consistent year on year (46). It can be tempting to assume that there is just one seasonal pattern in a given location. However, studies have shown that while so many aspects related to early life nutrition are seasonal (e.g. maternal diet and food insecurity, maternal body mass index, neonatal anthropometry), there is no single pattern: infection risk may have one pattern, maternal workload another, heat stress in pregnancy another, cultural factors another, food insecurity another, and infant growth yet another (50). The seasonality of a particular setting can vary year on year and is constantly feeding multiple stresses into different pathways, and these can be difficult to entangle in one location, let alone in pooled analyses.

One example of how the WaSt-TIG has investigated the effect of seasonality is in a retrospective cohort study from The Gambia (11). Here, clinic-based routine growth-monitoring data was used to look at patterns of growth in children from rural villages, including an investigation of whether the season of their birth influenced wasting and stunting later in life. Growth trajectories were stratified by whether the children were born in the wet season (July–October) or in the dry season (the other months of the year). Ways in which other researchers have accounted for aspects of seasonality in pooled analyses include using actual rainfall data from the time period and location to help define seasonal windows (32). A more involved method goes beyond doing this at the broad geographical level to pairing individual anthropometric data with the most precise temperature record of the location and date available ('spatiotemporally-varying climate data') (51).

In longitudinal datasets the seasonality of exposures and outcomes will need to be visualised and accounted for in analyses. Seasonal trends are fluctuations around a mean that are cyclical in nature. There are a number of ways these can be modelled, for example using cubic splines or Fourier terms. Useful explanations and examples on how to model seasonality in data include Marshak et al. (2021) (46), Saville et al. (2021) (50), Schoenbuchner et al. (2019) (11), Mertens et al. (2020) (32), Fulford (2014) (52) and Zhu (2021) (53).

Duration of study

When setting up new studies to look at issues related to wasting and stunting, the duration of data collection in a population or programme cohort will depend on the specific research questions, as well as the resources available. Some key lessons by the WaSt-TIG have been learnt in this regard. In the past, for example, research using CMAM programme data tended to focus on the period of time under wasting treatment, with then either no follow-up subsequently, or limited follow-up post-discharge period. However, to capture the full impact of wasting on later growth trajectories, the work of WaSt-TIG members has highlighted the need for longer follow-up periods (11, 26, 32). Furthermore, previous research has demonstrated there is a high prevalence of wasting and stunting present at birth and early in life (11, 32, 41, 54), and that episodes of wasting experienced early in life are associated with later episodes of wasting and stunting (11). To gain a fuller picture of the window of exposure, for population cohorts a longitudinal birth cohort is recommended. Ideally this will be with gestational age and birthweight data as available (see section above). The duration should be from birth to as long as possible, but ideally capturing pre-natal data and continuing follow-up at least up to 24 months to capture the peak prevalence and incidence of anthropometric deficits (see section on anthropometric outcomes above).

A systematic review considering studies that followed up children after severe wasting treatment found that risk of relapse was highest in the first six months post-discharge (55). Therefore, when using data from children treated for wasting and then discharged, follow-up should be for at least six months, with regular data collection (see following section), in order to capture the risk of relapse.

Frequency of data collection

As with deciding the duration of a study, the frequency of data collection also depends on the research questions being investigated. For example, if the only outcome is near-term mortality, data collection at six-month intervals may be sufficient (provided there is good dating and classification of the death, which can be challenging). However, generally more frequent data collection is better for greater precision around the timing of exposures and outcomes. Monthly intervals have proved particularly useful in previous WaSt-TIG studies using population cohorts, and weekly data collection for studies using treatment programme data. The longer the intervals between data collection during follow-up, the harder it becomes to attribute a given outcome to a risk factor and test hypotheses of predictors of outcomes. The increased frequency, however, comes with certain caveats: more frequent data collection is beneficial only if data quality and the required depth of investigations (i.e. the number of variables being collected) can be maintained.

How exposures of interest are defined also feeds into the required data collection frequency. For example, if growth faltering is a risk factor of interest, the minimum interval needs to be specified and then data collection planned accordingly. In a detailed cohort study in The Gambia that collected anthropometric data every two days, growth faltering in individuals could be identified using weekly data collection. In that example, rates of faltering would be underestimated or missed if the frequency of assessment was reduced to twice monthly or monthly data collection (56).

The frequency of data collection by field investigators will also depend on the study design. Furthermore, the number/type of variables being collected at each data collection point may vary; a tool for a six-monthly screen may look different to the more frequent monthly or quarterly follow-ups. In some settings, it may be possible to utilise existing demographic surveillance systems for certain records (e.g. mortality outcomes and sometimes morbidity and anthropometric records, depending on the surveillance system). This might then reduce the need for regular follow-up in cases where the surveillance system is good, and therefore require planning only for the collection of data on additional exposures and outcomes at less frequent intervals.

Pooling datasets

Many of the studies the WaSt-TIG have conducted have investigated rare outcomes such as mortality, severe wasting and concurrent WaSt. These less frequent outcomes require a large sample size and stratifying the data to analyse outcomes in different categories (e.g. household factors, sex, age band) further increases the sample size requirements. One strategy the WaSt-TIG have used to overcome this constraint has been to pool datasets from comparable contexts to ensure there is a sufficient sample size for the investigations and to allow for meta-analyses.

To successfully pool the datasets, the WaSt-TIG members have first needed to standardise them. For pooling cross-sectional surveys, standardised data collection methodologies, such as those for DHS and SMART surveys, have helped with the process of pooling data (15-17). Pooling can often be harder for population cohorts (with different selections of exposures, outcomes and data collection frequency) and nutrition treatment programme datasets (with their differing admission and discharge criteria), although sometimes raw data can be used to re-calculate exposures and outcomes (27).

Obtaining permissions for pooling datasets is essential and while relatively straightforward for publicly available data (such as DHS surveys), this can be time-consuming for restricted datasets. All datasets (and even more so for older datasets, e.g. those used for mortality analyses among untreated children) require permission for use from the original principal investigators or research institutes and/or ethical committees of studies, agreeing co-authorship, assessing the compatibility of variables, negotiating different file formats, and performing recalculations using the raw data (e.g. converting z-scores using the National Center for Health Statistics (NCHS) reference to the WHO 2006 growth standards, or making a common definition of anthropometric exposures and outcomes across datasets). The considerable time required for completing these steps should be factored into planning, and for new study designs, common definitions of exposures and outcomes should be considered where possible to allow for future meta-analyses.

Data cleaning

Due to the use by the WaSt-TIG of many different types of datasets in their analyses, an issue that has been discussed at length by members has been what the most appropriate method of data cleaning is depending on the source of the data. For survey data, it is common to use a standard data cleaning protocol that excludes potential outliers from the raw data based on z-score cut-offs. There are several common methods for cleaning survey data

(57), but two of the most widely used protocols adopt the WHO (2006) flags (58) or the SMART flags (59) (see Table 3). These aim to remove extreme (i.e. likely implausible) values that are more likely to be measurement error rather than reflect true measurements, and are used particularly when it is known that survey teams cannot return to the households to verify extreme values.

Table 3: Examples of exclusion criteria for data cleaning using WHO (2006) and SMART (2013) flags

	Exclusion criteria for data cleaning		
WHO (2006) growth standards (58)	HAZ < –6	WAZ < –6	WHZ < –5
	HAZ > 6	WAZ > 5	WHZ > 5
SMART flags (59)	HAZ < –3	WAZ < –3	WHZ < –3
	HAZ > 3	WAZ > 3	WHZ > 3

Abbreviations: HAZ, height-for-age z-score; WAZ, weight-for-age z-score; WHZ, weight-for-height z-score.

The choice of cleaning protocol makes a difference to estimates of the prevalence of malnutrition (57, 60, 61), and hence must be carefully considered, especially when using pooled datasets, where one cleaning method should be applied to raw data if possible (unless it has been pre-specified that all datasets have used the same data cleaning protocol).

For data coming from clinical and research settings, especially when individual children have been measured several times by trained clinicians or anthropometrists in longitudinal datasets, the choice of data cleaning protocol is not so straightforward (and indeed is a continued area of discussion within the WaSt-TIG). Many papers using such datasets do adopt the same data cleaning protocol as with survey data, especially the use of the WHO (2006) growth standards flags. Some papers using MUAC will additionally set their own MUAC cut-offs based on the range most likely to be biologically plausible (e.g. excluding MUAC <70 mm or >240 mm in children -59 months (23)). However, some researchers are concerned that using these data cleaning protocols may exclude children who are very sick or undernourished, and genuinely have extreme anthropometric measurements. The WHO (2006) growth standards flags are broadly designed to reflect the values incompatible with life. Clinical and programme datasets often contain very sick children who are indeed at high risk of mortality due to their level of wasting and stunting. In these settings, it is therefore likely that some 'implausible' values do reflect the reality. More resources and general information of data cleaning are provided in Box 5.

Given there is no universally agreed-upon data cleaning protocol for longitudinal data, it may be that each research team needs to decide in advance what difference in measurements between time points are plausible, to try to ensure that variation between consecutive measurements represents realistic change. Given these criteria may well vary in different contexts, when it comes to data pooling, it is again best to use raw data where possible and then apply one data cleaning method to all datasets.

Box 5: More resources and information on data cleaning

The topic of data cleaning procedures and their implications have been considered in depth by several researchers. As one example, a recent paper by Woolley et al. (2020) further outlines the details about what data cleaning techniques might be inappropriate, and what some of the alternatives could be (60). Many alternatives are described in the paper, including defining internal cut-offs, using a certain number of box lengths from the 25th or 75th percentile to determine which z-scores to exclude; using various cleaning methods specific to longitudinal datasets that plot the trajectories of individuals and determine outliers based on expected trends; and using linear spline regression models, as well as the authors' novel five-step data cleaning algorithm.

Additional considerations to help assess data quality of anthropometric data, especially relevant to survey data, include:

- Digit preference (the tendency of length/height measurements to cluster at digits commonly used in rounding, for example .0 and .5, or age estimates 'heaping', e.g. at 0, 6, 18, 24 months).
- The shape of the distribution (skewness and kurtosis) and the size of the standard deviation of the z-scores (a measure of variability in the data). Generally, anthropometric variables from high-quality datasets, when converted into z-scores and compared with the reference population, would demonstrate a symmetrical distribution with a standard deviation close to 1.0 (62). A higher standard deviation may indicate a greater risk of measurement error (63). The acceptable standard deviation varies by the anthropometric indicator, and the WHO Global Database on Childhood Growth and Malnutrition contain methods guidelines describing the following acceptable ranges (58):
  - HAZ: 1.10 to 1.30;
  - WAZ: 1.00 to 1.20; and
  - WHZ: 0.85 to 1.10.

Experience of the WaSt-TIG with systematic reviews

- Systematic reviews can be very useful for summarising the existing evidence base, exploring the extent that certain observations have been seen in different contexts, and for generating hypotheses for future research questions. For example, a recent systematic review published by the WaSt-TIG highlights population-level data that shows that wasting, stunting and concurrent WaSt are all more prevalent in boys than girls, and that wasting is higher in younger children, while stunting is higher in older children (64).
- As with all systematic reviews, those on topics related to wasting and stunting must carefully consider what conclusions can be drawn without meta-analysis, what the strength of evidence and risk of bias is from individual studies, and should provide clear reflection on what research questions still remain. For future systematic reviews on related topics, it is recommended to build on search terms used in previous reviews (e.g. see methods section of reference (64)) as search criteria that allow for wasting and stunting terms to be searched separately (rather than being addressed within the same article) will lead to an overwhelming number of articles being returned. Furthermore, age categories vary and are not always easy to compare, with some studies in the systematic review pooling estimates for 6–23 months and others using 0–30 months.

Concluding reflections

The WaSt-TIG have used a variety of types of datasets to explore associations between wasting and stunting over the past eight years. This has largely focused on cross-sectional data in the early phases to explore hypotheses which were then further investigated in longitudinal datasets, using both population cohorts and nutrition treatment programme cohorts. Explorations using intervention study designs are in their early days, but are building on the lessons learnt from the wealth of analyses done using existing datasets to date. The hope is that in the near future such intervention studies (e.g. see the proposed ENN prospective cohort protocol (65)) will further strengthen the evidence base and claims around causality, as well as contribute towards potential nutritional treatment options for children at the highest risk of mortality.

We hope the practices outlined in this technical brief will be useful to those looking at related questions in existing datasets (there are plenty of rich datasets with good quality longitudinal anthropometric data that have not been fully explored), as well as for setting up new studies. The types of research questions remaining (Box 6) lend themselves to longitudinal datasets, both population and treatment programme cohorts, rather than cross-sectional data. The above practices and suggestions are far from comprehensive, and there are undoubtedly some that have been overlooked. Furthermore, data analysis is an aspect that is highly specific to the type of data, study design and research question; it has been beyond the scope of this technical brief to explore detailed statistical considerations.

We conclude this briefing paper by acknowledging the huge body of work that has been undertaken globally linked to or outside of the specific activities of the WaSt-TIG group and that has contributed to our collective understanding of good practices and knowledge gaps within research practices.

We therefore encourage readers that have further specific questions on research methods relevant to the work of the WaSt-TIG to contact the team directly at philip@ennonline.net.

Box 6: Examples of further research questions in WaSt-related work

Examples of pending research questions are found in the discussion sections of the key literature from the WaSt-TIG, outlined in Table 2. Thurstans et al. (2021) (64) and Angood et al. (2016) (66) are particularly useful for highlighting remaining topics for investigation. Other examples of remaining questions include further investigation into the relationship between concurrent WaSt and mortality. For example, we do not yet know whether it makes a difference to mortality risk if concurrent WaSt results from a child who has been stunted for months/years and then experiences rapid weight loss, compared to a persistently wasted child who slowly becomes stunted through limited linear growth. Nor do we know whether the heightened mortality risk is limited to those experiencing a tight timeframe of concurrence (i.e. restricted to simultaneously experiencing both conditions), or whether episodes of wasting and stunting falling consecutively but not necessarily simultaneously also confers an elevated mortality risk. We also do not know how those with concurrent WaSt, possibly detected more simply by severe underweight, but who are not severely wasted, should be treated. The WaSt-TIG have published a protocol for a prospective cohort study to explore this latter point (65) and the International Rescue Committee are starting a trial in Mali in 2022 which will also help explore these scenarios.

References

1 Delgado-Rodriguez M, Llorca J. Bias. J Epidemiol Community Health. 2004;58(8):635–41.

2 Pearce N. Classification of epidemiological study designs. Int J Epidemiol. 2012;41(2):393–7.

3 McNamee R. Confounding and confounders. Occup Environ Med. 2003;60(3):227–34; quiz 164, 234.

4 Wang X, Cheng Z. Cross-Sectional Studies: Strengths, Weaknesses, and Recommendations. Chest. 2020;158(1S):S65–S71.

5 Hill AB. The environment and disease: association or causation? Proc R Soc Med. 1965;58:295–300.

6 Fedak KM, Bernal A, Capshaw ZA, Gross S. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. Emerging themes in epidemiology. 2015;12(1):1–9.

7 Isanaka S, Andersen CT, Cousens S, Myatt M, Briend A, Krasevec J, et al. Improving estimates of the burden of severe wasting: analysis of secondary prevalence and incidence data from 352 sites. BMJ Glob Health. 2021;6(3).

8 Chambers R, Longhurst R, Bradley D, Feachem R. Seasonal dimensions to rural poverty: analysis and practical implications. J Trop Med Hyg. 1979;82(8):156–72.

9 Vaitla B, Devereux S, Swan SH. Seasonal hunger: a neglected problem with proven solutions. PLoS Med. 2009;6(6):e1000101.

10 Chotard S, Mason JB, Oliphant NP, Mebrahtu S, Hailey P. Fluctuations in wasting in vulnerable child populations in the Greater Horn of Africa. Food Nutr Bull. 2010;31(3 Suppl):S219–33.

11 Schoenbuchner SM, Dolan C, Mwangome M, Hall A, Richard SA, Wells JC, et al. The relationship between wasting and stunting: a retrospective cohort analysis of longitudinal data in Gambian children from 1976 to 2016. The American Journal of Clinical Nutrition. 2019;110(2):498-507.

12 Khara T, Dolan C. Technical Briefing Paper: Associations between Wasting and Stunting, policy, programming and research implications. Oxford, UK: ENN; 2014.

13 Grellety E, Golden MH. Change in quality of malnutrition surveys between 1986 and 2015. Emerging themes in epidemiology. 2018;15(1):1–13.

14 Daher J, Cassard F, Gari S, Sauveplane-Stirling V, Nyawo M, Codjia P. Implementation of nutrition surveys using SMART methodology in sub-Saharan Africa. Field Exchange 58. 2018:68.

15 Khara T, Mwangome M, Ngari M, Dolan C. Children concurrently wasted and stunted: A meta-analysis of prevalence data of children 6-59 months from 84 countries. Matern Child Nutr. 2018;14(2):e12516.

16 International Food Policy Research Institute. Global Nutrition Report 2016: From Promise to Impact: Ending Malnutrition by 2030. Washington, DC; 2016.

17 Myatt M, Khara T, Schoenbuchner S, Pietzsch S, Dolan C, Lelijveld N, et al. Children who are both wasted and stunted are also underweight and have a high risk of death: a descriptive epidemiology of multiple anthropometric deficits using data from 51 countries. Arch Public Health. 2018;76(1):28.

18 International Food Policy Research Institute. Global Nutrition Report 2015: Actions and Accountability to Advance Nutrition and Sustainable Development. Washington, DC; 2015.

19 Wang X, Kattan MW. Cohort studies: Design, analysis, and reporting. Chest. 2020;158(1):S72-S8.

20 Euser AM, Zoccali C, Jager KJ, Dekker FW. Cohort studies: prospective versus retrospective. Nephron Clin Pract. 2009;113(3):c214–7.

21 Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. PLoS Med. 2008;5(3):e67.

22 Myatt M, Khara T, Dolan C, Garenne M, Briend A. Improving screening for malnourished children at high risk of death: a study of children aged 6-59 months in rural Senegal. Public Health Nutr. 2019;22(5):862–71.

23 Khara T, Myatt M, Sadler K, Bahwere P, Berkley JA, Black RE, et al. Anthropometric criteria for best identifying children at high risk of mortality: a pooled analysis of 12 cohorts. Submitted. 2022.

24 Garenne M, Myatt M, Khara T, Dolan C, Briend A. Concurrent wasting and stunting among under-five children in Niakhar, Senegal. Maternal & Child Nutrition. 2019;15(2):e12736.

25 Thurstans S, Wrottesley SV, Fenn B, Khara T, Sadler K, Bahwere P, et al. Evaluation of anthropometric deficits and the associated risk of death by age and sex in children 6-59 months: A secondary data analysis. In progress. 2022.

26 Isanaka S, Hitchings MDT, Berthe F, Briend A, Grais RF. Linear growth faltering and the role of weight attainment: Prospective analysis of young children recovering from severe wasting in Niger. Matern Child Nutr. 2019;15(4):e12817.

27 How do children with severe underweight and wasting respond to treatment: a pooled secondary data analysis to inform future intervention studies. In progress. 2022.

28 Kallander K, Kadobera D, Williams TN, Nielsen RT, Yevoo L, Mutebi A, et al. Social autopsy: INDEPTH Network experiences of utility, process, practices, and challenges in investigating causes and contributors to mortality. Popul Health Metr. 2011;9(1):44.

29 Moyer CA, Johnson C, Kaselitz E, Aborigo R. Using social autopsy to understand maternal, newborn, and child mortality in low-resource settings: a systematic review of the literature. Global health action. 2017;10(1):1413917.

30 D'Ambruoso L, Kahn K, Wagner RG, Twine R, Spies B, Van Der Merwe M, et al. Moving from medical to health systems classifications of deaths: extending verbal autopsy to collect information on the circumstances of mortality. Global health research and policy. 2016;1(1):1–15.

31 Victora CG, Christian P, Vidaletti LP, Gatica-Dominguez G, Menon P, Black RE. Revisiting maternal and child undernutrition in low-income and middle-income countries: variable progress towards an unfinished agenda. Lancet. 2021;397(10282):1388–99.

32 Mertens A, Benjamin-Chung J, Colford J, Hubbard A, van der Laan M, Coyle J, et al. Child wasting and concurrent stunting in low- and middle-income countries. medRxiv. 2020:2020.06.09.20126979-2020.06.09.

33 WHO, UNICEF. WHO child growth standards and the identification of severe acute malnutrition in infants and children: A joint statement by the World Health Organization and the United Nations Children's Fund. Geneva: WHO; 2009.

34 Hossain MI, Ahmed T, Arifeen SE, Billah SM, Faruque A, Islam MM, et al. Comparison of midupper arm circumference and weight-for-height z score for assessing acute malnutrition in Bangladeshi children aged 6-60 mo: an analytical study. Am J Clin Nutr. 2017;106(5):1232–7.

35 Tadesse AW, Tadesse E, Berhane Y, Ekstrom EC. Comparison of Mid-Upper Arm Circumference and Weight-for-Height to Diagnose Severe Acute Malnutrition: A Study in Southern Ethiopia. Nutrients. 2017;9(3):267.

36 Assaf S, Kothari MT, Pullum T. An assessment of the quality of DHS anthropometric data, 2005-2014. DHS Methodological Reports No 16. 2015(September).

37 Oumer A, Fikre Z, Girum T, Bedewi J, Nuriye K, Assefa K. Stunting and Underweight, but not Wasting are Associated with Delay in Child Development in Southwest Ethiopia. Pediatric Health, Medicine and Therapeutics. 2022;13:1.

38 Richard SA, Black RE, Gilman RH, Guerrant RL, Kang G, Lanata CF, et al. Wasting is associated with stunting in early childhood. The Journal of Nutrition. 2012;142(7):1291–6.

39 McDonald CM, Olofin I, Flaxman S, Fawzi WW, Spiegelman D, Caulfield LE, et al. The effect of multiple anthropometric deficits on child mortality: meta-analysis of individual data in 10 prospective studies from developing countries. Am J Clin Nutr. 2013;97(4):896–901.

40 WHO. WHO Consultation on Prevention and Treatment of Wasting in Children, Geneva, 17–19 December 2019. Geneva: WHO; 2020.

41 Benjamin-Chung J, Mertens A, Colford JM, Hubbard AE, van der Laan MJ, Coyle J, et al. Early childhood linear growth failure in low- and middle-income countries. 2021.

42 Lampl M, Gotsch F, Kusanovic JP, Gomez R, Nien JK, Frongillo EA, et al. Sex differences in fetal growth responses to maternal height and weight. American Journal of Human Biology: The Official Journal of the Human Biology Association. 2010;22(4):431–43.

43 Saville NM, Harris-Fry H, Marphatia A, Reid A, Cortina-Borja M, Manandhar DS, et al. Differences in maternal and early child nutritional status by offspring sex in lowland Nepal. Am J Hum Biol. 2021:e23637.

44 Victora CG, Villar J, Barros FC, Ismail LC, Chumlea C, Papageorghiou AT, et al. Anthropometric Characterization of Impaired Fetal Growth: Risk Factors for and Prognosis of Newborns With Stunting or Wasting. JAMA Pediatr. 2015;169(7):e151431.

45 Ma S, Finch BK. Birth outcome measures and infant mortality. Population research and policy review. 2010;29(6):865–91.

46 Marshak A, Venkat A, Young H, Naumova EN. How Seasonality of Malnutrition Is Measured and Analyzed. Int J Environ Res Public Health. 2021;18(4).

47 Moore SE. Early life nutritional programming of health and disease in The Gambia. J Dev Orig Health Dis. 2016;7(2):123-31.

48 Panter-Brick C. Seasonal growth patterns in rural Nepali children. 2009.

49 Rayco-Solon P, Fulford AJ, Prentice AM. Differential effects of seasonality on preterm birth and intrauterine growth restriction in rural Africans. Am J Clin Nutr. 2005;81(1):134–9.

50 Saville NM, Cortina-Borja M, De Stavola BL, Pomeroy E, Marphatia A, Reid A, et al. Comprehensive analysis of the association of seasonal variability with maternal and neonatal nutrition in lowland Nepal. Public Health Nutr. 2021:1–16.

51 Baker RE, Anttila-Hughes J. Characterizing the contribution of high temperatures to child undernourishment in Sub-Saharan Africa. Sci Rep. 2020;10(1):18796.

52 Fulford AJ. The coefficient of cyclic variation: a novel statistic to measure the magnitude of cyclic variation. Emerging themes in epidemiology. 2014;11:15.

53 Zhu Z. Towards data science. 2021. [cited 2022]. Available from: https://towardsdatascience.com/taking-seasonality-into-consideration-for-time-series-analysis-4e1f4fbb768f.

54 Mertens A, Benjamin-Chung J, Colford JM, Coyle J, van der Laan MJ, Hubbard AE, et al. Causes and consequences of child growth failure in low- and middle-income countries. 2020.

55 Stobaugh HC, Mayberry A, McGrath M, Bahwere P, Zagre NM, Manary MJ, et al. Relapse after severe acute malnutrition: A systematic literature review and secondary data analysis. Matern Child Nutr. 2019;15(2):e12702.

56 Naumenko DJ, Dykes J, O'Connor GK, Stanley Z, Affara N, Doel AM, et al. A Novel method for the identification and quantification of weight faltering. American Journal of Physical Anthropology. 2021;175(1):282–91.

57 Crowe S, Seal A, Grijalva-Eternod C, Kerac M. Effect of nutrition survey 'cleaning criteria' on estimates of malnutrition prevalence and disease burden: secondary data analysis. PeerJ. 2014;2:e380.

58 World Health Organization. WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development. Geneva: World Health Organization; 2006.

59 SMART. Measuring mortality, nutritional status, and food security in crisis situations: SMART methodology Version 1, April 2006: SMART; 2013 [Available from: https://smartmethodology.org/.

60 Woolley CS, Handel IG, Bronsvoort BM, Schoenebeck JJ, Clements DN. Is it time to stop sweeping data cleaning under the carpet? A novel algorithm for outlier management in growth data. PloS one. 2020;15(1):e0228154.

61 Freedman DS, Lawman HG, Skinner AC, McGuire LC, Allison DB, Ogden CL. Validity of the WHO cutoffs for biologically implausible values of weight, height, and BMI in children and adolescents in NHANES from 1999 through 2012. The American journal of clinical nutrition. 2015;102(5):1000–6.

62 Harkare HV, Corsi DJ, Kim R, Vollmer S, Subramanian S. The impact of improved data quality on the prevalence estimates of anthropometric measures using DHS datasets in India. Scientific reports. 2021;11(1):1–13.

63 Grellety E, Golden MH. The Effect of Random Error on Diagnostic Accuracy Illustrated with the Anthropometric Diagnosis of Malnutrition. PLoS One. 2016;11(12):e0168585.

64 Thurstans S, Sessions N, Dolan C, Sadler K, Cichon B, Isanaka S, et al. The relationship between wasting and stunting in young children: A systematic review. Maternal & Child Nutrition. 2021.

65 ENN. Bringing New Evidence On Undernutrition And Mortality Risk Into Practice: Protocol For A Prospective Cohort Study. Oxford: ENN; 2021.

66 Angood C, Khara T, Dolan C, Berkley JA, Group WTI. Research priorities on the relationship between wasting and stunting. PloS one. 2016;11(5):e0153221.

ENN
2nd Floor, Marlborough House, 69 High Street,
Kidlington, Oxfordshire, OX5 2DN, UK

Tel: +44 (0)1865 372340

office@ennonline.net
www.ennonline.net