A prospective validation of a laser anthropometric device for measuring height and length of children less than 60 Months in Vietnam

Statistical Analysis Plan (SAP)

May 26, 2025

Diane Baik

Accurate and reliable measurement of child length and height is essential for assessing nutritional status, monitoring growth, and identifying undernutrition among children under five years of age. However, conventional anthropometric measurement methods using wooden length/height boards can be time-consuming and challenging to implement in community-based settings. This study aims to validate the accuracy, precision, reliability, and agreement of the D100 Laser Height/Length Measuring Device against the conventional ShorrBoard among children aged 1–59 months in Quang Nam Province, Vietnam. Using a prospective, paired-measurement design involving trained experts and community volunteers, the study will assess device-related and measurer-related differences, including performance in recumbent length and standing height measurements. This Statistical Analysis Plan (SAP) outlines the analytical methods for evaluating measurement agreement, technical error of measurement, and equivalence within prespecified margins, as well as the implications of measurement differences for the classification of childhood stunting and wasting using WHO Child Growth Standards. The SAP also specifies the analytical populations, statistical models, subgroup and sensitivity analyses, and data quality procedures to ensure a rigorous, transparent, and reproducible assessment of the device's performance and its potential suitability for community-based nutrition programming.

Table 1 – Summary of Data Analysis Plan and Descriptions for Study 1

Type of Error to be Analyzed	Methods of Assessments	Description	When to Use	Interpretation
1. Agreement or Concordance (Systematic Error)**:** Closeness of agreement between measurements from two methods (one being an accepted method • *Inter-rater inter-method Technical Error of Measurement (inter-TEM)* (60)	Bland-Altman Plot (62,63)	Measures agreement between the laser device measured by CHWs and ShorrBoard measured by Experts	When comparing a new device (laser device) with an established or accepted one (ShorrBoard)	Small mean difference and narrow limits of agreement = High agreement; A lower inter-TEM = High Agreement; acceptable TEM threshold ≤0.7cm difference for height/length measurement (53).
	Mean Difference (Bias)	Measures the average difference between two methods (calculate mean, SD, and absolute mean difference) (64)	For paired measurements (Laser device vs. ShorrBoard)	Mean difference close to zero = High agreement
2. **Reliability** (Random Error): Consistency between repeated measurements under similar conditions and over time	Intraclass Correlation Coefficient (ICC) (61)	Measures consistency between repeated measurements under similar conditions	**Intra-rater intra-TEM:** Comparing measurements from the same measurers using the same device over repeated trials	High ICC (>0.90) = High reliability

Type of Error to be Analyzed	Methods of Assessments	Description	When to Use	Interpretation
• *Intra-rater intra-TEM reliability, Inter-rater intra-TEM reliability, Intra-rater inter-method reliability, test-retest reliability* (61)			**Inter-rater intra-TEM:** Comparing measurements from different measurers using the same device over repeated trials **Intra-rater inter-method:** Comparing measurements from the same measurers using different devices over repeated trials **Test-retest reliability:** Measures the consistency of results when the same measurer uses the same device to measure the same individuals multiple times under the same conditions over time	
3. Precision (Random Error)**:** Closeness of repeated measurements when measuring using same device • *Intra-rater intra-TEM precision*	Coefficient of Variation (CV)	Measures the ratio of the standard deviation to the mean	For assessing the spread of repeated height measurements	Lower CV= Higher precision
	Standard Deviation	Measures the average deviation from the mean	For evaluating the spread of repeated measurements	Lower SD= Higher precision
	Technical Error of Measurement (TEM) (61)	Measures the variability between repeated measurements taken under the same conditions.	For assessing how much error is introduced by the measurer or the instrument	Lower TEM= Higher precision

Type of Error to be Analyzed	Methods of Assessments	Description	When to Use	Interpretation
	Bland Altman Plot (62,63)	Measures the range within which 95% of repeated measures fall	To quantify the level of agreement between two or more repeated measurements made by same person using same device	Narrower LoA = Higher Precision; acceptable TEM threshold ≤0.7cm difference for height/length measurement (53)
Exploratory Variables				
Incidence of WLZ/WHZ category difference when using laser device vs. ShorrBoard (categorical)	**Cross-tabulation** of WHZ category (e.g., wasted (WLZ/WHZ<-2SD)/not wasted (WLZ/WHZ≥-2SD)) for each method			
Change in WLZ/WHZ-score as assessed by laser device vs. ShorrBoard (difference in ICC) (continuous)	ICC **Bland–Altman analysis**: • Plots the mean of both methods vs. their difference • Shows bias and limits of agreement	Use **ICC(2,1)** or **ICC(A,1)** for absolute agreement (preferred when comparing methods)		
Incidence of LAZ/HAZ category difference when using laser device vs. ShorrBoard (categorical)	**Cross-tabulation** of LAZ/HAZ category (e.g., stunted (LAZ/HAZ<-2SD)/not stunted (LAZ/HAZ≥-2SD)) for each method			
Changes in LAZ/HAZ-score as assessed by laser device vs. ShorrBoard (difference in ICC) (continuous)	ICC **Bland–Altman analysis**: • Plots the mean of both methods vs. their difference • Shows bias and limits of agreement	Use **ICC(2,1)** or **ICC(A,1)** for absolute agreement (preferred when comparing methods)		

We will use the Pitman test to assess the correlation between agreement and the child's age (or length/height) (65), and the Levene test of variance homogeneity to analyze the SDs of the differences between methods (66). In addition to the commonly used ICC, we will also examine the concordance correlation coefficient (CCC) (45), which does not rely on the ANOVA assumptions like the ICC does (67). The mean absolute difference represents the average of the absolute differences between two measurements from the same subject.

4. **Secondary outcomes:** Assess the differences in weight-for-height or weight-for-length z-score and height-for-age or length-for-age z-score categorization when assessed by the two different methods.

Measured Variables

This section lists every variable recorded by the field instrument and every variable derived during analysis, including variables used only for subsetting, exclusion, or sensitivity analysis. Variables described as derived are computed in the analysis script and are not present in the raw data extract. Table 2 covers variables as collected; Table 3 covers derived variables.

Table 2 – Variables recorded in the field instrument

Domain	Variable (as recorded)	Type	Role in analysis	Notes
Identification and design	Child_ID	Identifier	Linking key	Links the four instrument-by-measurer records contributed by each child
	Enum_Team	Categorical (1–4)	Design factor	Teams 1 and 3 are expert teams; teams 2 and 4 are community volunteer teams
	Team_Names	Text	Documentation	Free-text names of team members
	Survey_Date	Date	Derivation input	Used with date of birth to compute exact age
	Zone, AP, Commune, Village	Categorical	Descriptive; clustering	Geographic identifiers within Minh Long Area Programme
	Data_check	Indicator	Data management	Field-entered verification flag
Child characteristics	child_DOB	Date	Derivation input	Date of birth
	MOA	Integer (months)	Superseded	Completed months; a floor rather than a rounding of true age. Replaced by exact age in days
	Gender, Gender_MF	Binary	Covariate; z-score input	Sex is required for all WHO z-score calculations
	Ethnic_Grp	Categorical	Descriptive; potential subgroup	Not used in primary analysis
	Oedema	Binary	Descriptive; classification input	Bilateral pitting oedema status, required for correct wasting classification

Domain	Variable (as recorded)	Type	Role in analysis	Notes
	Weight (kg)	Continuous	z-score input; positive control	Does not depend on stature, so weight-for-age must be identical between devices
Anthropometry, wooden board	W1_Height, W2_Height 2, W3_Height 3 (cm)	Continuous	Primary measurement	First, second and (where required) third board readings
	W1/W2/W3 incl. 1 mm gap (cm)	Continuous	Analysed board readings	Board readings incorporating the 1 mm headpiece correction
	Measured_3x_TrueYes	Binary	Protocol and derivation	Whether a third board reading was taken
	W_Final1, W_Final2, W_AverageFinal1_Final 2	Continuous	Supplied derived values	Reproduced independently in the analysis script as a verification step
Anthropometry, laser device	L1_Height, L2_Height, L3_Height (cm)	Continuous	Primary measurement	First, second and (where required) third laser readings
	Measured_3x	Binary	Protocol and derivation	Whether a third laser reading was taken
	L_Final1, L_Final2, L_AverageFinal1_Final2	Continuous	Supplied derived values	Reproduced independently in the analysis script
Measurement context	position	Categorical (Lying, Standing)	Prespecified stratification factor	Corresponds exactly to age group in this dataset
	Child_Crying_1YES_2NO	Binary	Effect modifier	Reported by the measuring team
	Crying_Difficult_Measuring_1YES_2NO	Binary	Effect modifier	Reported by the measuring team; may be endogenous to measurement difficulty
	Weather during measurement	Categorical	Descriptive only	Near-constant (almost all measurements indoors); insufficient variation to analyse
	Measurement time in minutes; Measurement time in minutes – Laser	Continuous	Not analysable	Missing for the large majority of occasions, structured by team, and containing a constant placeholder value in one team
Supplied z-scores	WAZ_Wood, HAZ_Wood, WHZ_Wood, WAZ_Laser, HAZ_Laser, WHZ_Laser	Continuous	Cross-check only	Recomputed independently for analysis; the supplied values contain biologically implausible entries
	WAZNutStat, HAZNutStat, WHZNutStat (Wood and Laser)	Categorical	Cross-check only	Classification labels derived from the supplied z-scores

Table 3 – Variables derived during analysis

Derived variable	Definition	Role in analysis
age_days, age_months	Survey_Date minus child_DOB, in days; months as days divided by 30.4375	Covariate, stratification, and input to all WHO z-scores
agegrp	Under 24 months versus 24 to 59 months	Prespecified stratum
measure	Length indicator for recumbent measurement, height indicator for standing	Selects the correct WHO growth standard
cadre	Expert versus community volunteer, from Enum_Team	Design factor; measurer effect
team	Enum_Team as a four-level factor	Confounding check for the measurer effect
block	Team pair A (teams 1 and 2) or B (teams 3 and 4)	Design structure; sensitivity analysis
w_r1, w_r2, l_r1, l_r2	Retained replicate pair: the two closest of three readings where a third was taken, otherwise the first two	Input to all precision estimates
w_mean, l_mean	Mean of the retained pair	Child-level summary value for each instrument-by-measurer cell
w_dif, l_dif	Difference within the retained pair	Input to intra-observer technical error of measurement
dif_w_ft, dif_l_ft	Absolute difference between the first two readings, rounded to the 1 mm instrument graduation	Determines whether a third reading was required
trig_w, trig_l	Indicator that the rounded first-pair difference exceeded 0.7 cm	Protocol adherence; burden analysis
has_w3, has_l3	Indicator that a third reading was recorded	Protocol adherence
unres_w, unres_l, unres	Third reading required but not recorded	Exclusion criterion 2
gap_add	Whether the 1 mm headpiece correction was applied to a given board record	Sensitivity analysis
haz_b, whz_b, waz_b; haz_l, whz_l, waz_l	WHO 2006 z-scores recomputed separately from board-derived and laser-derived stature	Classification outcomes and exclusion screening
biv_haz, biv_whz, biv_waz, biv_any	WHO biological implausibility flags (±5 SD for weight-for-height, ±6 SD for height-for-age, −6 to +5 SD for weight-for-age)	Exclusion criterion 1
excl_biv, excl_unres, excluded	Child-level exclusion indicators and their union	Defines the primary analysis set
C1, C2, C3, C4	Child-level paired contrasts between instrument-by-measurer cells	Accuracy outcomes
dev_eff, mea_eff, int_eff	Device main effect, measurer main effect, and their interaction, from the balanced within-child design	Primary decomposition for objective 2
stunt, sstunt, wast, swast (board and laser)	Binary classifications at the −2 and −3 z-score cut-offs	Classification outcomes
band	Distance of the board-derived height-for-age z-score from the −2 cut-off, in four categories	Prespecified subgroup analysis
sq	Quartile of expert board-measured stature	Subgroup analysis across the measurement range

Seven variables were added after project planning, and exact age in days was used in place of months completed. This is because a completed-month variable is a floor rather than a rounding of true age and

shifts height-for-age z-scores down. The Measurer cadre, team and block variables were constructed so that measurer effect could be considered separately from device, and team-level confounding could be assessed. Third reading trigger and completion variables were included to permit assessment of adherence to the protocol and for definition of the second exclusion criterion. The correction indicator for a 1 mm headpiece was retained to assess the effect of the correction on estimated bias. The first exclusion criterion was WHO biological implausibility flags. Prespecified subgroup analysis included distance from the diagnostic threshold and stature quartile. Finally, two variables describing the duration of measurement were found not to support analysis.

Statistical Models by Question

The unit of analysis is the child. Each child contributes four measurements: board and laser, each by an expert team and by a community volunteer team. Analyses are conducted overall and stratified by age group (under 24 months, measured recumbent; 24 to 59 months, measured standing). All models are specified below in sufficient detail to be reproduced from the data.

1. Did the ShorrBoard achieve gold-standard measurement precision in this setting?

The estimator is the TEM, defined as the square root of the sum of squares of within-pair differences/(2n) for each instrument and cadre separately, with 95% confidence interval (CI) calculated using the chi-square distribution on n degrees of freedom. The formula is as follows. In this case, the reference process is considered gold standard if the expert board TEM is less than 0.4 cm.

If expert board TEM is 0.4 cm or above, no accuracy can be estimated under the WHO/UNICEF interim standards, and the analysis switches to the effectiveness branch where only precision is assessed. The criterion is non-inferiority of laser TEM to board TEM. All the accuracy models below assume this condition is satisfied.

2. What are the separate device and measurer effects, and is the laser device equivalent to the board within the prespecified margin?

Four child level paired contrasts are defined: the laser minus board within experts (C1), the laser minus board within volunteers, the volunteer minus expert (board), and the volunteer (laser) minus expert (board). C1 is the primary contrast for device validity. Each contrast is summarized by a one-sample t-test of the mean difference with upper and lower 95% confidence intervals and Bland-Altman limits of agreement. Post-hoc, limits of agreement for a single future measurement are derived using the correction for replicated data, i.e. adding within-subject replicate variance for each method to the variance of the differences of the mean measurement between the two methods.

The model specification used is a balanced, complete two-by-two (device, cadre) design within each child, so that both the device and measurer effects are estimable as main effects. The model is stature(i,j,k) = $\beta 0 + \beta 1 \cdot \text{device}(j) + \beta 2 \cdot \text{cadre}(k) + \beta 3 \cdot \text{device}(j) \times \text{cadre}(k) + u(i) + e(i,j,k)$. The parameters are defined as in the text, where i is the child index, j is the instrument index (board is reference), k is the cadre index (expert is reference), and u(i) and e(i,j,k) are child random intercept and residual error terms, with variances $\sigma^2 u$ and $\sigma^2 e$, respectively. For balanced, orthogonal designs such as this, the fixed effects equal linear combinations of the child-level cell means estimated directly via one-sample t-tests. Thus, the device main effect is the mean of (laser minus board) averaged over the cadres, the measurer main effect is the mean of (volunteer minus expert) averaged over the instruments, and the interaction is C2 minus

C1. This analysis is preferred as it makes no distributional assumptions about the random intercept term. The mixed model is fit as a check and should agree to within rounding.

Equivalence is tested with a two one-sided tests procedure procedure with a margin of δ=0.2 cm. and α=0.05. The two treatments are considered equivalent if the 90% confidence interval for the mean difference is completely within $\pm\delta$. A secondary margin of 0.4 cm is reported. All four contrasts are equivalence tested with C1 as the primary.

Lin's concordance correlation coefficient and the intraclass correlation coefficient for absolute agreement, ICC(A,1) are also shown for comparative purposes. Neither of these is used to test whether the device meets the prespecified requirements: both of these are measures for the ratio of the between-child variance to the total variance so are based upon the age range of the sample as much as the functioning of the instrument.

3. Does device performance differ between recumbent length and standing height?

To test systematic difference, contrast(i) = $\gamma 0 + \gamma 1 \cdot \text{position}(i) + \varepsilon(i)$ were separately fitted to C1 and C2 with position contrast between recumbent and standing. A Welch's t-test assuming unequal variances between strata was used to compare means. Dispersion was tested using the Brown-Forsythe test. This is a variant of ANOVA, in which absolute deviations from the median of each stratum are used. This test was used instead of the mean-based Levene's test or the F test because it is not dependent on the normality assumption. The ratio of technical errors of measurement was computed as the square root of the variance ratio, with boundaries determined from the F distribution. Summed for each instrument and cadre, the same test was applied to the intra-observer replicate differences to discriminate a position effect on between-method agreement from a position effect on the repeatability of the measurements of each instrument.

To test proportional bias, we regress the difference between the two methods on their means, $\text{difference} = \alpha 0 + \alpha 1 \cdot \text{mean} + \varepsilon$. We can test the null hypothesis that $\alpha 1 = 0$. For heteroscedasticity, we regress the absolute value of the residuals from the regression against the mean, and all four contrasts are fitted. Performance across the measurement range is described in terms of quartiles of expert-board-measured stature.

4. What are the consequences of residual measurement disagreement for nutritional classification?

Height-for-age, weight-for-height, and weight-for-age Z-scores in board-derived stature and laser-derived stature were computed from the WHO 2006 Child Growth Standards using the actual days since birth and the indicator for measurement position (supine or standing). Stunting (moderate and severe) and wasting (moderate and severe) were computed as the Z-score crossing the -2 and -3 cut-off, respectively. Comparative binary classifications between the 2 devices were tested with McNemar's test with continuity correction. Agreement was summarized in terms of the number and direction of clashing classifications, Cohen's kappa coefficient, and difference in estimated prevalence. Confidence intervals for the prevalence difference were obtained using a cluster bootstrap with 2000 resamples at the level of the child in order to respect the non-independence due to the fact that each child has two measurement occasions.

Subgroup and follow-up analysis. Discordance is assessed as the distance of the board-derived HAZ from the cut-off of −2, as four prespecified classes of < 0.25 SD, 0.25 to < 0.5 SD, 0.5 to < 1.0 SD and > = 1.0 SD. All classification analyzes are repeated within strata by age group.

For fewer than five clashing pairs, McNemar's test is reported as not estimable, because the test statistic is not stable. Kappa is also estimated, but its results may be influenced by aggregation and are presented with caution. Severe wasting is expected to fall into this category.

Controls, checks and assumption testing

Since weight-for-age is height independent, the weight-for-age value calculated from board height must be the same as that from laser height. If this is not the case, it indicates a problem in the z-score pipeline, and analysis is halted.

The final-measurement variables derived from the raw measurements must be reproduced completely and correctly, for otherwise the derivation rule has been misunderstood, and one must re-establish the derivation rule before proceeding to the analysis.

As we expect the device-by-measurer interaction to be zero, any device effect found to depend on whether an expert or a volunteer, say, is using it, is reported separately for each cadre; we do not report an overall device effect.

We will check normality of differences using quantile-quantile plots and the Shapiro-Wilk test. In the event of a material deviation from normality, the median and the 10% trimmed mean with a bootstrap confidence interval will be reported as the primary measure of the treatment effect. The mean will be reported as a secondary measure of the treatment effect. In all cases, strong estimates will be given in addition to the parametric estimates. The assumption of homoscedasticity will be assessed as in research question 3 by the absolute-residual regression, and in the presence of heteroscedasticity, the limits of agreement will be given as range-dependent. Independence is handled by analyzing at the child level and performing bootstrap resampling at the child level.

For mixed models, restricted maximum likelihood is used, alongside the bobyqa optimiser with 200,000 function evaluations as a maximum. For singular fits, where the estimated variance component is estimated to equal (or be very close to) 0, the random effect is removed from the model. If the mixed model does not converge, the algebraically equivalent paired contrast estimate is used. Since the paired contrast is the primary one, no conclusions change based on whether the mixed model does or does not converge.

Only C1 is a primary contrast; all other contrasts, subgroup analyzes, and classification results are prespecified secondary analyzes for estimation only. Multiplicity is not corrected for, and this is explicitly acknowledged in the results wherever secondary analyzes are reported.

Sensitivity analyses

Four prespecified sensitivity analyses are reported alongside the primary estimates: replicate selection (the two closest of three readings versus the first two protocol readings); the 1 mm headpiece correction applied versus not applied; the exclusion criteria (both applied, biological implausibility only, or none); and the protocol-adherent analysis set versus the full analysis set retaining children whose discrepant replicate pairs were never resolved.

Software

Analyses are conducted in R version 4.5.2. WHO 2006 z-scores are computed with the anthro package; mixed models with lme4. The analysis script is version-controlled and will be archived with the study materials so that the analysis can be reproduced exactly.

References

1. WHO. WHO guideline on the prevention and management of wasting and nutritional oedema (acute malnutrition) in infants and children under 5 years. Geneva: WHO; 2023.
2. Bhutta ZA, Berkley JA, Bandsma RHJ, Kerac M, Trehan I, Briend A. Severe childhood malnutrition. Nat Rev Dis Primer. 2017;3(1):17067.
3. UNICEF. No time to waste: Improving diets, services and practices for the prevention, early detection and treatment of wasting in early childhood. New York: UNICEF; 2021.
4. Bailey J, Lelijveld N, Marron B, Onyoo P, Ho LS, Manary M, et al. Combined Protocol for Acute Malnutrition Study (ComPAS) in rural South Sudan and urban Kenya: study protocol for a randomized controlled trial. Trials. 2018 Apr 24;19(1):251.
5. UNICEF/WHO. Levels and trends in child malnutrition: key findings of the 2021 edition of the joint child malnutrition estimates. Geneva: WHO; 2021.
6. FAO. Codex Alimentarius: Standard for the labelling of and claims for foods for special medical purposes CODEX STAN 180-1991.
7. FAO. Codex Alimentarius: General standard for the labelling of and claims for prepackaged foods for special dietary uses CODEX STAN 146-1985.
8. FAO. Codex Alimentarius: Codes of Practice.
9. Hossain MI, Nahar B, Hamadani JD, Ahmed T, Brown KH. Effects of community-based follow-up care in managing severely underweight children. J Pediatr Gastroenterol Nutr. 2011 Sep;53(3):310–9.
10. Javan R, Kooshki A, Afzalaghaee M, Aldaghi M, Yousefi M. Effectiveness of supplementary blended flour based on chickpea and cereals for the treatment of infants with moderate acute malnutrition in Iran: A randomized clinical trial. Electron Physician. 2017;9(12):6078–86.
11. Nikièma L, Huybregts L, Kolsteren P, Lanou H, Tiendrebeogo S, Bouckaert K, et al. Treating moderate acute malnutrition in first-line health services: an effectiveness cluster-randomized trial in Burkina Faso. Am J Clin Nutr. 2014;100(1):241–9.
12. Vanelli M, Virdis R, Contini S, Corradi M, Cremonini G, Marchesi M, et al. A hand-made supplementary food for malnourished children. Acta Biomed. 2014;85(3):236–42.
13. Padhani ZAO, Cichon BAO, Das JK, Salam RAO, Stobaugh HC, Mughal MAO, et al. Systematic Review of Management of Moderate Wasting in Children over 6 Months of Age. LID - 10.3390/nu15173781 [doi] LID - 3781. (2072-6643 (Electronic)).
14. Griswold SP, Langlois BK, Shen Y, Cliffer IR, Suri DJ, Walton S, et al. Effectiveness and cost-effectiveness of 4 supplementary foods for treating moderate acute malnutrition: results from a cluster-randomized intervention trial in Sierra Leone. Am J Clin Nutr. 2021 Sep 1;114(3):973–85.
15. Peter Rohloff, Subodh Gupta, Waleska López Canu, Wendy Rodríguez Gómez, Shela Sridhar, Anahi Venzor. New WHO guideline on the prevention and management of acute malnutrition in infants and young children: remaining challenges. BMJ Paediatr Open. 2024 May 7;8(1):e002471.
16. World Vision International. Nutrition Marketplace [Internet]. Toronto: WVI; 2010 [cited 2025 Mar 6]. Available from: https://www.wvi.org/sites/default/files/Nutrition_MarketplaceV3.pdf
17. Nutrition Working Group, Child Survival Collaborations and Resources Group (CORE). Positive Deviance/Hearth: A resource guide for sustainably rehabilitating malnourished children [Internet]. Washington, D.C.; 2002 [cited 2025 Mar 6]. Available from: https://coregroup.org/wp-content/uploads/2017/09/Positive-Deviance-Hearth-Resource-Guide.pdf
18. Prasad V, Holla R, Gupta A. Should India Use Commercially Produced Ready to Use Therapeutic Foods (RUTF) for Severe Acute Malnutrition (SAM). Soc Med. 2009 Feb 8;4(1):52–5.

19. Bhandari N, Mohan SB, Bose A, Iyengar SD, Taneja S, Mazumder S, et al. Efficacy of three feeding regimens for home-based management of children with uncomplicated severe acute malnutrition: a randomised trial in India. BMJ Glob Health. 2016;1(4):e000144.
20. Vietnam MoH. Vietnam National Nutrition Strategy 2021-2030 with Vision to 2040. Hanoi; 2021.
21. Global Nutrition Report. Global Nutrition Report: Country Nutrition Profile for Vietnam.
22. de Onis M, Borghi E, Arimond M, Webb P, Croft T, Saha K, et al. Prevalence thresholds for wasting, overweight and stunting in children under 5 years. Public Health Nutr. 2019 Jan;22(1):175–9.
23. Mondon C, Tan PY, Chan CL, Tran TN, Gong YY. Prevalence, determinants, intervention strategies and current gaps in addressing childhood malnutrition in Vietnam: a systematic review. BMC Public Health. 2024 Apr 4;24(1):960.
24. Nguyen PH, Saha KK, Ali D, Menon P, Manohar S, Mai LT, et al. Maternal mental health is associated with child undernutrition and illness in Bangladesh, Vietnam and Ethiopia. (1475-2727 (Electronic)).
25. Ngoc Hien N, Ngoc Hoa N. Nutritional Status and Determinants of Malnutrition in Children under Three Years of Age in Nghean, Vietnam. Pak J Nutr. 8(7):958–64.
26. Ethnic minority development in Vietnam: The Journal of Development Studies: Vol 43 , No 7 - Get Access [Internet]. [cited 2025 Feb 8]. Available from: https://www.tandfonline.com/doi/full/10.1080/02673030701526278
27. Baulch B, Chuyen TTK, Haughton D, Haughton J. Ethnic minority development in Vietnam. J Dev Stud. 2007 Oct 1;43(7):1151–76.
28. Lavin T, Preen DB, Newnham EA. Access to Obstetric Care and Children's Health, Growth and Cognitive Development in Vietnam: Evidence from Young Lives. Matern Child Health J. 2017;21:1277–87.
29. Bennett IM, Schott W, Krutikova S, Behrman JR. Maternal mental health, and child growth and development, in four low-income and middle-income countries. (1470-2738 (Electronic)).
30. Ministry of Health Vietnam. National Guidelines: Diagnostic and Treatment of Acute Malnutrition in Children from 0 to 72 months old. Hanoi: Ministry of Health Vietnam; 2016.
31. World Vision Vietnam. Nutrition Clubs in Vietnam. 2016.
32. Baik D, Klaas N. World Vision's Positive Deviance/Hearth Training of Facilitators Manual and Handouts [Internet]. Toronto: World Vision International; 2021 [cited 2025 Feb 10]. Available from: https://figshare.com/articles/book/World_Vision_s_Positive_Deviance_Hearth_Training_of_Facilitators_Manual_and_Handouts/13615310
33. Baik D, Ndolo R. Nourishing Children: Menus for Recovery [Internet]. London: WVI; 2023 [cited 2025 Mar 6]. Available from: https://www.wvi.org/sites/default/files/2023-10/Nourishing%20Children_Menus%20for%20Recovery%202023_0.pdf
34. Baik D. World Vision's Positive Deviance/Hearth programme: multi-country experiences. FEX [Internet]. 2019 Jul [cited 2025 Feb 20];(60). Available from: https://www.ennonline.net/sites/default/files/FEX-60-Web_final_11-14.pdf
35. Bisits Bullen PA. The positive deviance/hearth approach to reducing child malnutrition: systematic review. Trop Med Int Health. 2011 Nov;16(11):1354–66.
36. Bolles K, Speraw C, Berggren G, Lafontant JG. Ti Foyer (Hearth) community-based nutrition activities informed by the positive deviance approach in Leogane, Haiti: a programmatic description. Food Nutr Bull. 2002 Dec;23(4 Suppl):11–7.
37. Mackintosh UA, Marsh DR, Schroeder DG. Sustained positive deviant child care practices and their effects on child growth in Viet Nam. Food Nutr Bull. 2002 Dec;23(4 Suppl):18–27.
38. Triatmaja NT, Mahmudiono T, Mamun AA, Abdullah NA. Effectiveness of Positive Deviance Approach to Reduce Malnutrition among under Five Children: A Systematic Review and Meta-Analysis of Interventional Studies. Nutrients. 2023 Apr 19;15(8).

39. Young MF, Baik D, Reinsma K, Gosdin L, Rogers HP, Oy S, et al. Evaluation of mobile phone-based Positive Deviance/Hearth child undernutrition program in Cambodia. Matern Child Nutr. 2021 Oct;17(4):e13224.
40. Anino OC, Were GM, Khamasi JW. Impact evaluation of positive deviance hearth in Migori County, Kenya. Afr J Food Agric Nutr Dev. 2015;15(5):10578–96.
41. Kang Y, Kim S, Sinamo S, Christian P. Effectiveness of a community-based nutrition programme to improve child growth in rural Ethiopia: a cluster randomized trial. Matern Child Nutr. 2017 Jan;13(1).
42. Grellety E, Golden MH. Weight-for-height and mid-upper-arm circumference should be used independently to diagnose acute malnutrition: policy implications. BMC Nutr. 2016 Feb 5;2(1):10.
43. Kesari A, Noel JY. Nutritional Assessment. In: StatPearls. Treasure Island (FL); 2024.
44. Laar ME, Marquis GS, Lartey A, Gray-Donald K. Reliability of length measurements collected by community nurses and health volunteers in rural growth monitoring and promotion services. BMC Health Serv Res. 2018 Feb 17;18(1):118.
45. Bougma K, Mei Z, Palmieri M, Onyango D, Liu J, Mesarina K, et al. Accuracy of a handheld 3D imaging system for child anthropometric measurements in population-based household surveys and surveillance platforms: an effectiveness validation study in Guatemala, Kenya, and China. Am J Clin Nutr. 2022 Jul 6;116(1):97–110.
46. Conkle J, Keirsey K, Hughes A, Breiman J, Ramakrishnan U, Suchdev PS, et al. A collaborative, mixed-methods evaluation of a low-cost, handheld 3D imaging system for child anthropometry. Matern Child Nutr. 2019 Apr;15(2):e12686.
47. Huang S, Conkle J, Homer CSE, Kounnavong S, Phongluxa K, Vogel JP. Comparing the accuracy of an ultrasound height measurement device with a wooden measurement board among children aged 2-5 years in rural Lao People's Democratic Republic: A methods-comparison study. PLoS One. 2023;18(11):e0289514.
48. Mayol-Kreiser SN, Garcia-Turner VM, Johnston CS. Examining the utility of a laser device for measuring height in free-living adults and children. Nutr J. 2015 Sep 8;14:93.
49. Soller T, Huang S, Horiuchi S, Wilson AN, Vogel JP. Portable digital devices for paediatric height and length measurement: A scoping review and target product profile matching analysis. PLoS One. 2023;18(7):e0288995.
50. Cichon B, Das JK, Salam RA, Padhani ZA, Stobaugh HC, Mughal M, et al. Effectiveness of Dietary Management for Moderate Wasting among Children > 6 Months of Age-A Systematic Review and Meta-Analysis Exploring Different Types, Quantities, and Durations. Nutrients. 2023 Feb 21;15(5):1076.
51. 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 Apr;22(5):862–71.
52. Odei Obeng-Amoako GAO, Stobaugh H, Wrottesley SV, Khara T, Binns P, Trehan I, et al. How do children with severe underweight and wasting respond to treatment? A pooled secondary data analysis to inform future intervention studies. (1740-8709 (Electronic)).
53. de Onis M, Onyango AW, Van den Broeck J, Chumlea WC, Martorell R. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. Food Nutr Bull. 2004 Mar;25(1 Suppl):S27-36.
54. Bauman A, Ernst K, Hayden M, Roe DJ, Murray R, Agawo M, et al. Assessing Community Health: An Innovative Tool for Measuring Height and Length. J Trop Pediatr. 2018 Apr 1;64(2):146–50.
55. ACHI. BMI Screening Program: Height and Weight Measurement Training Manual [Internet]. Arkansas, USA: Arkansas Center for Health Improvement; 2018 Aug [cited 2025 Feb 8] p. 23. Available from: https://achi.net/wp-content/uploads/2018/01/Height_and_Weight_Measurement_Training_Manual.pdf

56. Sundar R, Sudha C. Thermal Transport and Thermal Diffusivity by Laser Flash Technique: A Review. Int J Thermophys. 2024 Dec 30;46(1):13.
57. World Health Organization, United Nations Children's Fund (UNICEF). Recommendations for data collection, analysis and reporting on anthropometric indicators in children under 5 years old [Internet]. Geneva: World Health Organization; 2019 [cited 2025 Mar 9]. 144 p. Available from: https://iris.who.int/handle/10665/324791
58. Guest G, Namey E, McKenna K. How Many Focus Groups Are Enough? Building an Evidence Base for Nonprobability Sample Sizes. Field Methods. 2017 Feb 1;29(1):3–22.
59. World Health Organization, United Nations Children's Fund (UNICEF). Recommendations for data collection, analysis and reporting on anthropometric indicators in children under 5 years old [Internet]. Geneva: World Health Organization; 2019 [cited 2025 Feb 21]. 144 p. Available from: https://iris.who.int/handle/10665/324791
60. Viviani C, Arezes PM, Bragança S, Molenbroek J, Dianat I, Castellucci HI. Accuracy, precision and reliability in anthropometric surveys for ergonomics purposes in adult working populations: A literature review. Int J Ind Ergon. 2018 May 1;65:1–16.
61. Goto R, Mascie-Taylor CGN. Precision of measurement as a component of human variation. J Physiol Anthropol. 2007 Mar;26(2):253–6.
62. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet Lond Engl. 1986 Feb 8;1(8476):307–10.
63. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999 Jun;8(2):135–60.
64. Zubir S, Ikhwan MN, Abd. Rahman S, Hamadan H, Salim N, Haji Baruji TsHME, et al. Technical Error Measurements, Reliability, and Validity of Customized Anthropometric Grid. J Occup Saf Health. 2023 Jun 1;20:51–8.
65. Norton K, Olds T. Anthropometrica: A Textbook of Body Measurement for Sports and Health Courses. UNSW Press; 1996. 438 p.
66. Gastwirth J, Gel Y, Miao W. The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice. Stat Sci. 2010 Oct 2;24.
67. Chen CC, Barnhart HX. Comparison of ICC and CCC for assessing agreement for data without and with replications. Comput Stat Data Anal. 2008 Dec 15;53(2):554–64.
68. World Health Organization. Integrated Management of Childhood Illness: distance learning course [Internet]. Geneva: World Health Organization; 2014 [cited 2025 Feb 20]. 15 p. Available from: https://iris.who.int/handle/10665/104772
69. Hoang NTD, Orellana L, Le TD, Gibson RS, Worsley AF, Sinclair AJ, et al. Anthropometric Status among 6–9-Year-Old School Children in Rural Areas in Hai Phong City, Vietnam. Nutrients. 2018 Oct 4;10(10):1431.
70. RANDOM.ORG - Sequence Generator [Internet]. [cited 2025 Feb 21]. Available from: https://www.random.org/sequences/
71. Marshak A, Venkat A, Young H, Naumova EN. How Seasonality of Malnutrition Is Measured and Analyzed. Int J Env Res Public Health. 2021 Feb 13;18(4).
72. World Health Organization, United Nations Children's Fund (UNICEF). Recommendations for data collection, analysis and reporting on anthropometric indicators in children under 5 years old [Internet]. Geneva: World Health Organization; 2019 [cited 2025 Mar 7]. 144 p. Available from: https://iris.who.int/handle/10665/324791