Title: Agreement in infant growth indicators and overweight/obesity between community and clinical care settings

ACCEPTED FOR PUBLICATION IN THE JOURNAL OF THE ACADEMY OF NUTRITION AND DIETETICS

Keywords: infant growth; Special Supplemental Nutrition Program for Women, Infant and Children; pediatric clinical care; anthropometry; overweight and obesity

Author contact information

Holly A. Harris, PhD, APD: Postdoctoral Fellow, Center for Childhood Obesity Research, 129 Noll Laboratory, The Pennsylvania State University, University Park, PA 16802, USA.

Postdoctoral Fellow, Generation R Study Group, Erasmus Medical Center, PO Box 1738, 3000 DR, Rotterdam, The Netherlands.

Samantha M. R. Kling, PhD, RD: Postdoctoral Fellow, Obesity, Geisinger Obesity Institute, Epidemiology and Health Services Research Geisinger, Danville, PA 17822, USA.

Research Fellow and Assistant Research and Teaching Professor, Center for Childhood Obesity Research, Department of Nutritional Sciences, The Pennsylvania State University, University Park, PA 16802, USA.

Quantitative Research Scientist, Evaluation Sciences Unit, Department of Medicine, School of Medicine, Stanford University, Palo Alto, CA 94305, USA.

Michele Marini, MS: Statistician, Center for Childhood Obesity Research, The Pennsylvania State University, University Park, PA 16802, USA.

Sandra G. Hassink, MD: Medical Director, American Academy of Pediatrics Institute for Healthy Childhood Weight, 2602 Pennington Dr, Wilmington DE 19810, USA.

Lisa Bailey-Davis, DEd, RD: Assistant Professor, Geisinger Obesity Institute, Population Health Sciences, Geisinger, 100 N. Academy Ave, Danville, PA 17822, USA.

Jennifer S. Savage, PhD: Associate Professor and Director of the Center for Childhood Obesity Research, 129 Noll Laboratory, The Pennsylvania State University, University Park, PA 16802.

Author contributions: HAH, MM and SMRK designed the current study, conducted the statistical analyses and interpreted the results; SGH, LBD and JSS assisted with the interpretation of the results; HAH wrote the first draft of the manuscript, all authors contributed to manuscript preparation, and read and approved the final manuscript.

Corresponding author: JSS

Funding/ financial disclosures: This project is supported by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) under grant number R40MC28317, Maternal and Child Health Field-initiated Innovative Research Studies Program. Funding for REDCap was provided by The Penn State Clinical & Translational Research Institute, Pennsylvania State University CTSA, NIH/NCATS Grant Number UL1 TR002014).

Title: Agreement in infant growth indicators and overweight/obesity between community and clinical care settings

RESEARCH SNAPSHOT

Research Question: What is the agreement in infant growth assessments between community (Women, Infants, and Children [WIC]) and clinical (primary care providers) care settings?

Key Findings: Measurements which include length (e.g. Weight-for-Length) tend to show greater disagreement between community and clinical settings than those that do not (e.g. Weight-for-Age). Findings indicate the need to investigate the techniques, standards, and training protocols for measuring infant weight and length in these settings to identify effective strategies to improve measurement practices, accuracy, and communication of consistent messaging.

ABSTRACT

Background: Infants from low-income backgrounds receive nutrition care from both community and clinical care settings. However, mothers accessing these services have reported receiving conflicting messages related to infant growth between settings, although this has not been examined quantitatively.

Objective: Describe the agreement in infant growth assessments between community (Women, Infants, and Children [WIC]) and clinical (primary care providers [PCP]) care settings.

Design: A cross-sectional, secondary data analysis of infant growth measures abstracted from electronic data management systems.

Participants/ setting: Participants included a convenience sample of infants (*n*=129) from northeastern Pennsylvania randomized to the WEE Baby Care study from July 2016 to May 2018. Infants had complete anthropometric data from both community and clinical settings at 6.2±0.4 months old. Average time between assessments was 2.7±1.9 weeks.

Main outcome measures: Limits of agreement (LOA) and bias in Weight-for-Age, Length-for-Age, Weight-for-Length, and BMI-for-Age *z* scores; and cross-context equivalence in weight status between care settings.

Statistical analysis performed: Bland-Altman analyses were used to describe the LOA and bias in *z* scores between care settings. Cross-context equivalence was examined by dichotomizing infants' growth indicators at the 85th and 95th percentile cut-points and cross-tabulating equivalent and discordant categorization between settings.

Results: Strongest agreement was observed for Weight-for-Age *z* scores (95% LOA: -0.41 to 0.54). However, the LOA intervals for growth indicators which included length were wider, suggesting weaker agreement. There was a high level of inconsistency for classification of overweight/obesity using Weight-for-Length *z* scores, with 15.5% (85th percentile cut-point) and 11.6% (95th percentile cut-point) discordant categorization between settings, respectively.

Conclusions: Infant growth indicators that factor in length could contribute to disagreement in the interpretation of infant growth between settings. Further investigation into the techniques, standards, and training protocols for obtaining infant growth measurements across care settings is required.

INTRODUCTION

Rapid weight gain in infancy is a risk factor for later obesity,[1] and obesity is more prevalent in children from disadvantaged backgrounds.[2] Frequent and accurate monitoring of infant growth is fundamental to effective preventive pediatric health care.[3] Low-income families are eligible for support covering multiple community and clinical care services, such as the Special Supplemental Nutrition Program for Women, Infants and Children (WIC) and well-child visits (WCVs) at pediatric Primary Care Providers (PCPs), with scheduled visits overlapping in the first year of life. Anthropometric measures are routinely collected and documented in both of these settings to evaluate nutritional status and growth, which guides clinical decision making, educational messages and nutrition services.[4] However, mothers have reported receiving conflicting messages related to infant growth from the two settings, leading to confusion and potentially undermining healthcare providers' advice.[5] Examining anthropometric measures obtained across sectors during a sensitive growth period in infancy could contribute to optimizing childhood obesity prevention and management,[6] particularly in vulnerable populations.[7]

Mixed educational messaging from medical professionals and WIC nutritionists is a barrier for mothers' adherence to recommended responsive feeding or nutrition advice.[8,9] One source of conflicting messaging may be the assessment and communication of infant growth. In a qualitative study, mothers reported that the assessment or advice related to child growth were inconsistent between WIC nutritionists and PCPs in the clinical care setting.[5] The authors concluded that parent confusion in child growth status could be driven by different measurement techniques, equipment used or time lapsed between appointments.[5] However, there is little evidence to support the (dis)agreement of infant growth assessment between community and clinical care settings.

Infant anthropometric data collected in clinical settings have typically been validated against "gold-standard" researcher-measured assessments.[10-12] Such studies indicate greater variability in measurements of infant length, perhaps due to the logistical challenges associated with assessing recumbent length.[13] In the WIC setting, validity of anthropometric data has only be examined in preschool-aged children,[14] therefore limiting the generalizability of findings. Nevertheless, these studies fail to evaluate the agreement in growth measurement between community and clinical settings, where providers may use different practices to collect, evaluate and record data. There is a need to identify the agreement of the measures between settings as well as differences that call for improvements in measurement or data collection processes. This is particularly important in low-income settings, where families may access multiple services and would benefit from multidisciplinary care to reduce inequities.[15]

To understand whether infant growth assessment practices across settings need to be evaluated and standardized, the first aim of this study was to describe the limits of agreement and potential bias in infant growth indicators measured between WIC and pediatric PCP care settings in the WEE Baby Care study. A second aim was to establish whether measurements obtained from this sample were comparable between settings (i.e., cross-context equivalence) based on percentile cut-points which are often used to classify weight status in practice.

MATERIALS AND METHODS

Study Design and Participant Description

The current cross-sectional study examines a convenience subsample of infants from the WEE Baby Care pragmatic trial with complete weight and length data at approximately 6 months of age from both the pediatric PCP and WIC visits.[15] Details of the study design, recruitment, inclusion and exclusion criteria, and CONSORT diagram have been published.[15]

Briefly, mother-newborn dyads were recruited from July 2016 to May 2018 in northeastern Pennsylvania, an area geographically characterized by the Health Services and Resources Administration as Medically Underserved.[16] Dyads were randomized into a responsive parenting intervention group to reduce rapid infant weight gain or a standard care control group; however, there were no study group differences on growth outcomes at any point in the study, or participants' frequency of WIC visits or WCV(data not published). This study was approved by the Institutional Review Boards of The Pennsylvania State University and Geisinger. All participants provided written informed consent.

Data collection procedures

Anthropometric assessments. Mother-infant dyads are recommended to attend 3 WIC visits and 5 well-child visits at PCPs from birth to 6-months. Measurements from each setting were anticipated at approximately infant age 2-, 4- and 6-months. However, many mothers in the sample did not keep to the schedule as recommended (i.e., skipped visits, rescheduled). Therefore, a convenience sample of infants with anthropometric measures from both settings occurring at approximately the 6-month visit were examined (n=129). Infant birth date and anthropometric assessment date were used to calculate infant age in months. Infants in the sample were (Mean ± SD) 6.2±0.4 months old. The average time between WIC visits and WCVs for each participant was 2.7±1.9 weeks, and 81% of the sample had visits that occurred within one month of each other.

The clinical and community settings trained staff measured weight and length, recorded the data, and monitored fidelity as per their standard practices. The Bureau of WIC training manual[17] described that infants were weighed once to the nearest ¼ lb on a double beam infant scale or electronic digital scale with light or no clothing except for a dry diaper. Infant recumbent length was measured to the nearest 1/8 inch using an infant measuring board with a fixed headboard and adjustable foot board, in duplicate until two measures agreed within ¼ inch. Anthropometric measure policies at pediatric PCP sites[18] followed similar procedures and used similar equipment to that of WIC providers. However, weight was measured in duplicate until the measurements agreed within 0.1 kg (or ¼ lb). If an infant measuring board was not available at the PCP clinic, a recumbent mat was used. At the PCP clinics, the nurse administrator was responsible for ensuring that clinical staff were trained and implemented the standard procedure. In the WIC setting, WIC clinic managers were responsible for managing the measurement techniques of the staff. Patient data was recorded electronically via WIC's electronic data management system (EDMS), QuikWIC. Pediatric PCPs recorded patient data from WCVs in the Epic® Electronic Health Record (Verona, WI). Weight and length data were securely and systematically extracted from each provider's EDMS by research study staff. However, the research staff did not check that these procedures were followed during the study.

World Health Organization (WHO) growth standards[19] were used to calculate indicators of growth (*z* scores and percentiles) at each setting: Weight-for-Age (WFA), Length-for-Age (LFA), Weight-for-Length (WFL) and Body Mass Index-for-Age (BMI). Standardized growth indicators were examined, as opposed to raw weight and length, because these are preferred measures to assess growth, and are adjusted for child sex and age (exception for WFL, which is not adjusted for child age). Mean differences in *z* scores were calculated by subtracting measurements taken from the clinical setting (PCPs) from the community (WIC) setting (i.e., community *z* score minus clinical *z* score) and averaging the mean differences in the sample. A mean *z* score difference of zero indicates perfect agreement between settings, whereas positive values reflect a higher score from the community setting and negative values reflect a higher score from the clinical setting.

Demographic characteristics. Enrolled participants completed surveys online through REDCap electronic survey sytem[20] or paper questionnaires. Demographic variables were collected from mothers at enrollment, including age, marital status, highest level of education attained, employment status, household income and number of people living in the household. Infant sex, gestational age and birth weight were obtained from hospital medical records. Mothers reported their infant's race and ethnicity.

Statistical analysis

Data were analyzed using SAS (version 9.4).[21] Differences in participant demographic characteristics and growth indicators by study group were examined using independent samples t-tests (continuous variables) and chi-square tests (categorical variables). 'Agreement' was examined through the lens of comparing one measure to another with the assumption that neither measure is expected to be more accurate than the other.[22] To examine the agreement in *z* score growth indicators between community and clinical settings, procedures outlined by Bland and Altman[23] were used. The main purpose of the Bland-Altman analysis is to compare two different methods by calculating and plotting the difference between a pair of measurements against the mean for each subject.[24] Patterns of bias and error between methods were assessed visually using a Bland-Altman plot, which plots the bias (difference between community and clinical measures) against the magnitude (mean of the community and clinical measures) of differences. In this step, mean differences were calculated and limits of agreement were constructed (mean and 95% prediction interval of the difference between community and clinical measures). The mean difference scores for each indicator were examined for normality using Shapiro-Wilk tests, which all showed that the data were normally distributed ($Ps \geq 0.19$). Paired t-tests were used to test whether this bias was significantly different from zero.[23] An acceptable limit of agreement for each indicator was determined *a priori* and set at <0.67 SD, as this value represents crossing at least one centile band on a standard infant growth chart.[25] Frongillo et al.[22] defined *cross-context equivalence* to mean that a measure "performs consistently across contexts" (p.1821). This definition was adopted in the current study to examine equivalence of categorization of individuals into percentile cut-points which are often used to classify weight status in practice. Infant growth indicators were dichotomized above or below the 85th percentile, and equivalence (both settings categorized an infant ≥85th or <85th percentile) and discordance (infant categorized as ≥85th percentile at community setting and <85th percentile at clinical setting or vice versa) were cross-tabulated. These steps were also repeated using the ≥95th *vs.* <95th percentile category. These categories were chosen because The American Academy of Pediatrics defined infant overweight as WFL ≥95th percentile and cautions practitioners to monitor the rate of weight gain as infants approach 85th and 95th percentiles.[6]

RESULTS

Demographic characteristics for mothers and their infants are shown in **Table 1**. There were no study group differences in demographic characteristics, and there were no study group differences in child growth indicators between community and clinical settings.

Agreement in anthropometric data collected from the community and clinical settings (shown in **Table 2**) was strongest for WFA *z* scores, while the limits of agreement for indicators that included infant length were wider (i.e., less agreement). However, each indicator's limits of agreement did not fall within the *a priori* acceptable limit. There was a statistically significant (*P*=0.002) bias toward WFA *z* scores from the community setting to be greater than those from the clinical setting (**Figure 1A**). LFA *z* scores from the community setting were less than those from the clinical setting (**Figure 1B**), although this bias was not statistically significant (*P*=0.07). There was a statistically significant (*P*=0.002) bias toward

WFL *z* scores from the community setting to be greater than those from the clinical setting (**Figure 1C**). Similar results emerged for BMI *z* scores (**Figure 1D**), and this bias was also statistically significant (*P*=0.002). There was no association between bias and the magnitude of differences between community and clinical settings for each growth indicator.

Cross-context equivalence in percentile categories for growth indicators across settings are shown in **Table 3**. Across all indicators, categorization of infants' WFA at the 85th and 95th percentile cut-points was generally equivalent between community and clinical settings (96.2% and 94.6%, respectively). The WFL indicator showed the least cross-context equivalence between community and clinical settings compared to other indicators, with 15.5% and 11.6% of infants discordantly categorized at the 85th and 95th percentile cut-points, respectively. The community setting tended to categorize children above the 85th and 95th percentile for WFL more frequently than scores derived from the clinical setting. Compared to WFL, discordant categorization of infants' BMI between settings was marginally lower using the 85th and 95th percentile cut-points (14.7% and 7.8%, respectively).

DISCUSSION

Low-income families whose infants are at higher risk of obesity[1] typically receive care from both WIC nutritionists and PCPs in community and clinical settings, respectively. However, low-income mothers, WIC nutritionists, and pediatricians have qualitatively reported conflicting messaging on infant and child growth across care settings.[5] The current study tested whether infant growth assessments and classification of overweight/obesity agreed between community (WIC nutritionist) and clinical (pediatric PCPs) settings in an effort to identify opportunities for improvement in the assessment of infant growth. Results revealed a general lack of agreement in infant growth indicators between community and clinical settings, according to the acceptable limits of agreement which were determined *a* *priori*. This appears to be clinically meaningful, because the acceptable limits of agreement was a clinically informed cut-off range (<0.67 SD) to represent crossing at least one centile band on a standard infant growth chart.[25] The limits of agreement varied depending on the indicator examined. Indicators which factored in length, notably WFL, which are typically used to categorize infants with "overweight",[6] had a wide limit of agreement compared to indictors which did not factor in length (i.e., WFA). Furthermore, there was evidence of a lack of cross-context equivalence in overweight/obesity classification based on percentile categories using these growth indicators. These findings indicate a need to investigate the techniques, protocols and standards of measuring infant length, and how growth is communicated to parents using community and clinical care services.

To the authors' knowledge, this is the first study to examine agreement in infant anthropometric measures between community and clinical care settings. While conclusions cannot be made about the validity of measures, the current findings would suggest a need to investigate the accuracy of growth assessments in both settings. For example, routinely-collected infant anthropometric measures from clinical settings have been compared to "gold-standard" researcher-measured data.[10-12] Generally, agreement appears to be stronger for infant weight than for length when comparing clinical data to researcher-measured data.[10,12] Similarly, one study of older children aged 2-5 years showed stronger agreement for weight than for height when comparing community (WIC) to researcher-measured data.[14] Future research should examine agreement between community, clinical and researcher-measured "gold standard" data to pinpoint sources of inaccuracies and opportunities to improve assessment practices.

Wide limits of agreement for growth indicators that include infant length reinforce the need to improve assessment practices across settings.[26] Sources of error may emerge from variability in care setting policies, experience and techniques of individual practitioners, and the accuracy and precision of equipment used. While anthropometric assessment policies are somewhat similar across the WIC and PCP settings in the current study,[17,18] adherence to these procedures is unknown. Lipman et al.,[26] for example, showed that 48% of practitioners marked exam room paper to obtain infant length rather than using a measuring board. But even with precise and calibrated equipment, incorrect technique can yield inaccurate measurements. There are many logistical challenges associated with measuring length, like stretching an infant's body with both legs extended out on a measuring board and the assistance of a second observer, often an untrained parent, to position the infant.[13] Yet, routine training has shown to improve practitioners' infant length measurement techniques.[26] To further minimize measurement error across settings, the WHO recommends the standardization of procedures, regular auditing for adherence to procedures and data quality monitoring,[13] which could be promoted by contracted personnel with appropriate accreditation and expertise in health care settings.

Findings from the current study align with Bailey-Davis et al.'s qualitative study[5] in which mothers reported conflicting child growth assessments and messaging between care settings. In the current study, approximately 12% and 16% of infants were placed in discordant WFL categories based on the 85th and 95th percentile cut-points, respectively. This means that up to 16% of infants may receive contradictory assessments of weight status, resulting in false positives (i.e., one care setting identifies a child as at risk of overweight when they are in fact normal weight) or false negatives (where one care setting classifies an infant as not at risk of overweight, when they are in fact, at risk). However, it is not possible to rule out that a health professional's interpretation and communication of infant growth assessments may be different across settings, despite documented 'agreement' in anthropometric measures. Disagreement in weight status classification also represents a missed opportunity for providers to implement early intervention strategies to improve infant growth trajectories and reinforce nutrition education messages. Misclassification of weight status between providers, regardless of setting, may also be important when considering the integration of routine health data collected to interpret RCT results.[27]

Strengths and limitations

The current study examined a vulnerable sample of infants from low-income backgrounds, who are at higher risk of developing overweight and whose mothers may use health care across multiple settings.[4] Limitations included the inability to ascertain whether one measure was more accurate than the other, and the inability to control variability in time between anthropometric assessments.[22] However, time varying between WIC appointments and WCVs likely reflects real-world practices. Recommendations suggest that infant growth is measured frequently by pediatricians in the first year of life[3] and regular collection of weight and length/height data is mandatory at WIC,[28] but appointments may be nonadjacent across settings. Providers could communicate to families or other health professionals that anthropometric assessments are based on the day that they are taken, which may vary since the last clinical or WIC visit because infant growth is dynamic. Differences in measurement time points, however, are common in literature investigating agreement in anthropometrics across sources, for example, parent-reported versus researcher measured infant length and weight.[29] However in the current study, growth indicators that were adjusted for children age, including WFA, LFA and BMI, were compared across settings. The use of these age-adjusted indicators accounts for some variation due to differences in measurement time points. Furthermore, research staff did not monitor fidelity to standardized procedures for assessment of infant anthropometrics, although this was undertaken using in-house protocols in each setting. Research staff independently assessing fidelity to standardized procedures in each care setting could contribute to a greater understanding of the strengths and shortfalls of adherence to procedures in future research.

CONCLUSION

Disagreement in infant growth indicators between community and clinical care settings may prevent the adoption of recommended responsive feeding advice, especially for low-income families. While community and clinical care settings generally agree on infant growth indicators, there is greater disagreement for growth indicators which factor in length. Furthermore, some disagreement in classification of infants' overweight/ obesity status was observed between sectors. This may be the basis for conflicting health messages between health providers, resulting in confusion for parents. These findings indicate the need to evaluate the accuracy of assessments across settings and investigate how effective strategies to improve measurement practice can be implemented.

References

1. Taveras EM, Rifas-Shiman SL, Belfort MB, et al. Weight status in the first 6 months of life and obesity at 3 years of age. *Pediatrics.* 2009;123(4):1177-1183.
2. Ogden CL, Fryar CD, Hales CM, et al. Differences in obesity prevalence by demographics and urbanization in US children and adolescents, 2013-2016. *JAMA.* 2018;319(23):2410-2418.
3. Committee On Practice and Ambulatory Medicine, Bright Futures Periodicity Schedule Workgroup. 2020 Recommendations for Preventive Pediatric Health Care. *Pediatrics.* 2020;145(3).
4. Koleilat M, Whaley SE, Esguerra KB, et al. The Role of WIC in Obesity Prevention. *Curr Pediatr Rep.* 2017;5(3):132-141.
5. Bailey-Davis L, Kling SMR, Cochran WJ, et al. Integrating and coordinating care between the Women, Infants, and Children Program and pediatricians to improve patient-centered preventive care for healthy growth. *Transl Behav Med.* 2018;8(6):944-952.
6. Daniels SR, Hassink SG, Committee On Nutrition. The Role of the Pediatrician in Primary Prevention of Obesity. *Pediatrics.* 2015;136(1):e275-292.
7. Andrea SB, Hooker ER, Messer LC, et al. Does the association between early life growth and later obesity differ by race/ethnicity or socioeconomic status? A systematic review. *Ann Epidemiol.* 2017;27(9):583-592.e585.
8. Savage JS, Neshteruk CD, Balantekin KN, et al. Low-Income women's feeding practices and perceptions of dietary guidance: A qualitative study. *Matern Child Health J.* 2016;20(12):2510-2517.

9. Savage JS, Birch LL, Marini M, et al. Effect of the INSIGHT responsive parenting intervention on rapid infant weight gain and overweight status at age 1 year: A randomized clinical trial. *JAMA Pediatrics.* 2016;170(8):742-749.

10. Bryant M, Santorelli G, Fairley L, et al. Agreement between routine and research measurement of infant height and weight. *Arch Dis Child.* 2015;100(1):24.

11. Furlong KR, Anderson LN, Kang H, et al. BMI-for-age and weight-for-length in children 0 to 2 years. *Pediatrics.* 2016;138(1):e20153809.

12. Howe LD, Tilling K, Lawlor DA. Accuracy of height and weight data from child health records. *Arch Dis Child.* 2009;94(12):950.

13. de Onis M, Onyango AW, Van den Broeck J, et al. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food Nutr Bull.* 2004;25(Sup):S27-S36.

14. Crespi CM, Alfonso VH, Whaley SE, et al. Validity of child anthropometric measurements in the Special Supplemental Nutrition Program for Women, Infants, and Children. *Pediatr Res.* 2012;71(3):286.

15. Savage JS, Kling SMR, Cook A, et al. A patient-centered, coordinated care approach delivered by community and pediatric primary care providers to promote responsive parenting: pragmatic randomized clinical trial rationale and protocol. *BMC Pediatr*. 2018;18(1):293.

16. MUA Find [Internet]. dataHRSA.gov.

17. Anthropometric Training Manual In: Bureau of Women, Infants and Children (WIC), Pennsylvania, USA 2018.

18. Geisinger Health Service. Policy and procedures for pediatric anthropometric measures. 2019.

19. 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.

20. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377-381.

21. SAS [computer program]. *Version 9.4.* Cary, NC: SAS Institute Inc; 2012.

22. Frongillo EA, Baranowski T, Subar AF, et al. Establishing Validity and Cross-Context Equivalence of Measures and Indicators. *J Acad Nutr Diet.* 2019;119(11):1817-1830.

23. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;327(8476):307-310.

24. Kalra A. Decoding the Bland–Altman plot: basic review. *J Prac Cardiovasc Sci.* 2017;3(1):36.

25. Ong KK, Loos RJF. Rapid infancy weight gain and subsequent obesity: Systematic reviews and hopeful suggestions. *Acta Paediatrica.* 2006;95(8):904-908.

26. Lipman TH, Hench KD, Benyi T, et al. A multicentre randomised controlled trial of an intervention to improve the accuracy of linear growth measurement. *Arch Dis Child.* 2004;89(4):342-346.

27. McCord KA, Salman RA-S, Treweek S, et al. Routinely collected data for randomized trials: promises, barriers, and implications. *Trials.* 2018;19(1):29.

28. USDA Food and Nutrition Service. 7 CFR 246.4 (a). State Plan. 2017. 2017; https://www.gpo.gov/fdsys/pkg/CFR-2011-title7-vol4/pdf/CFR-2011-title7-vol4-sec246-4.pdf. Accessed December 9, 2019.

29. Hazrati S, Hourigan SK, Waller A, et al. Investigating the accuracy of parentally reported weights and lengths at 12 months of age as compared to measured weights and lengths in a longitudinal childhood genome study. *BMJ Open.* 2016;6(8).

Table 1: Demographic characteristics of mothers and their infants (*n*=129, age 6-months) from the WEE Baby Care Study (northeastern Pennsylvania) with anthropometric data available from both the community[a] and clinical[b] setting

Demographic Characteristic	
Infant	
Male, n (%)	67 (52)
Gestational age (weeks), mean ± sd	39.0 ± 1.0
Birthweight (kg), mean ± sd	3.5 ± 0.5
Infant race, n (%)	
Black	18 (14)
White	70 (54)
Other	41 (32)
Hispanic/ Latino, n (%)[c]	40 (33)
Mother	
Age at infant birth (years), mean ± sd	27.8 ± 5.5
Primiparous, n (%)	41 (32)
Marital status, n (%)	
Married and/or living with partner	63 (52)
Single	48 (40)
Divorced/separated	8 (7)
Other	2 (2)
Education, n (%)	
Some high school or less	15 (13)
High school graduate	60 (50)
Some college	38 (31)

College graduate or greater	8 (7)
Annual household income, n (%)	
< $10,000	32 (26)
$10,000-$24,999	37 (31)
$25,000-$49,999	39 (32)
$50,000-$74,999	2 (2)
Do not know/ Refuse to answer	11 (9)
Average size of household (persons), mean ± sd	3.1 ± 1.4

[a]the Special Supplemental Nutrition Program for Women, Infant and Children (WIC)
[b]pediatric Primary Care Providers (PCP)
[c]data available for *n*=121 participants

Table 2: Limits of agreement in infant growth assessments between community[a] and clinical[b] care settings from the WEE Baby Care Study (northeastern Pennsylvania; n = 129)

Community setting	Clinical setting	Mean difference[c]	
Mean ± standard deviation			95% Limits of agreement
WFA *z* score[d]			
0.26 ± 1.12	0.19 ± 1.13	0.06 ± 0.24**	-0.41 to 0.54
LFA *z* score[e]			
-0.28 ± 1.11	-0.16 ± 1.09	-0.18 ± 0.74	-1.59 to 1.35
WFL *z* score[f]			
0.66 ± 1.04	0.48 ± 1.11	0.18 ± 0.62**	-1.07 to 1.42
BMI *z* score[g]			
0.56 ± 1.06	0.38 ± 1.13	0.18 ± 0.61**	-1.05 to 1.41

*P<0.05, **P<0.01, ***P<0.001;

[a]the Special Supplemental Nutrition Program for Women, Infant and Children (WIC)

[b]pediatric Primary Care Providers (PCP)

[c]Mean difference calculated by subtracting mean WIC *z* scores from PCP *z* scores

[d]Weight-for-Age

[e]Length-for-Age

[f]Weight-for-Length

[g]Body Mass Index-for-Age

Table 3: Cross-context equivalence of growth indicators between community[a] and clinical[b] settings according to 85th and 95th percentile categories for growth indicator for *n*=129 infants (age 6-months)

	WFA[c]	LFA[d]	WFL[e]	BMI[f]
Equivalent categorization for 85th percentile, %				
Both settings <85th	72.9	83.0	58.9	64.3
Both settings ≥85th	23.3	7.0	26.6	20.9
Discordant categorization for 85th percentile, %				
Community ≥85th, clinical <85th	3.1	4.6	6.2	5.4
Community <85th, clinical ≥85th	0.8	5.4	9.3	9.3
Equivalence categorization for 95th percentile, %				
Both settings <95th	85.3	92.3	76.0	80.6
Both settings ≥95th	9.3	3.1	12.4	11.6
Discordant categorization for 95th percentile, %				
Community ≥95th, clinical <95th	2.3	1.6	8.5	6.2
Community <95th, clinical ≥95th	3.1	3.1	3.1	1.6

[a]Community setting is the Special Supplemental Nutrition Program for Women, Infant and Children (WIC)

[b]Clinical setting is clinical pediatric Primary Care Providers (PCP)

[c]Weight-for-Age

[d]Length-for-Age

[e]Weight-for-Length

[f]Body Mass Index-for-Age

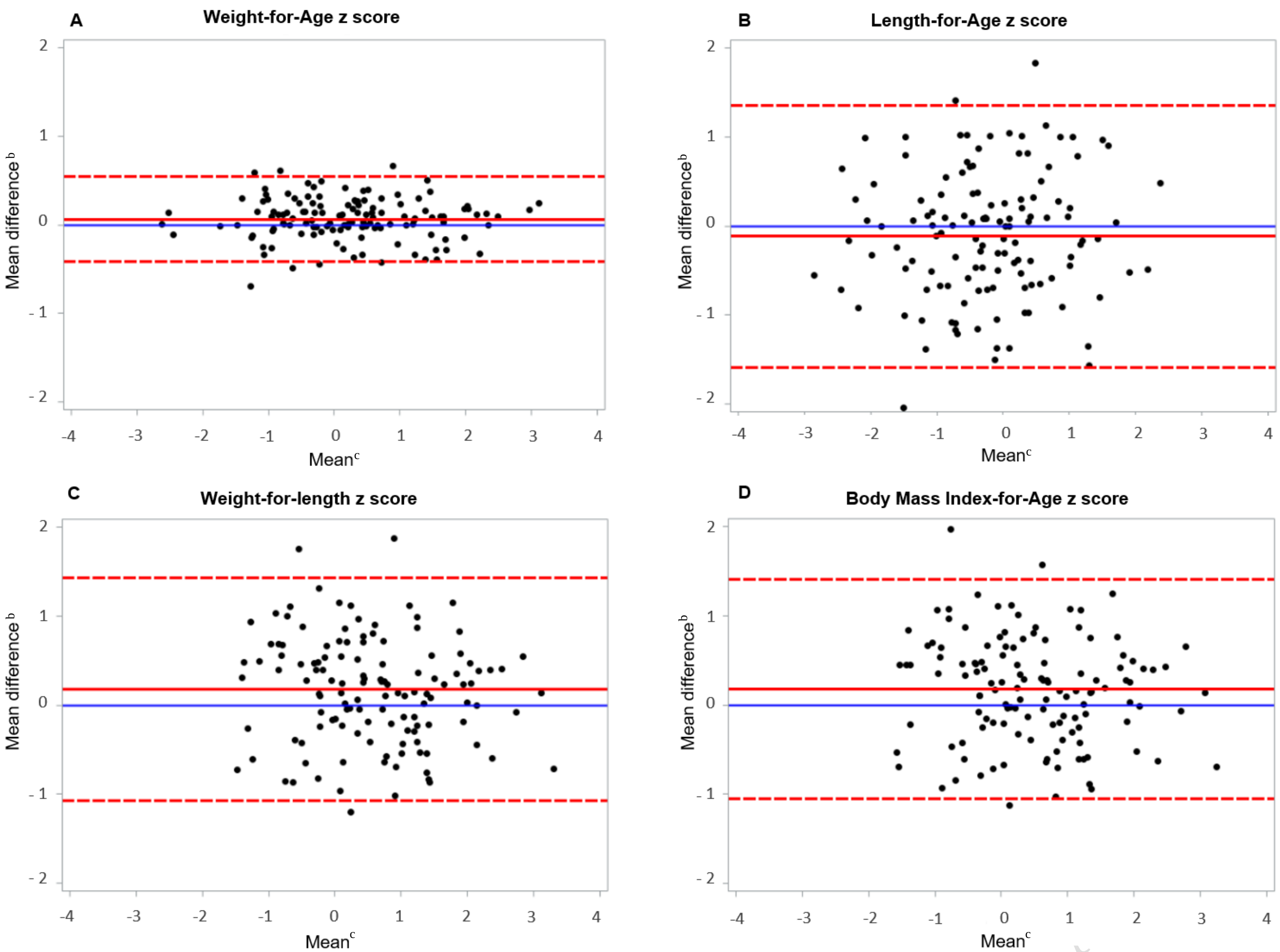
A
Weight-for-Age z score
B
Length-for-Age z score
C
Weight-for-length z score
D
Body Mass Index-for-Age z score
Mean difference b
Mean c
2
1
0
-1
-2
-4
-3
-2
-1
0
1
2
3
4

Figure 1: Bland-Altman plots of the potential bias of infant growth measurements[a] taken at community and clinical settings for the WEE Baby Care study (*n*=129 infants)

Footnotes: [a] Calculated into *z* scores from the World Health Organization growth standards[19]

[b]Mean difference between settings is calculated by subtracting *z* scores derived from measurements taken at the clinical setting (pediatric PCPs) from the community setting (i.e., WIC) and averaging these differences (shown by red solid line). The red dashed lines show the 95% limits of agreement interval. The solid blue line shows a theoretical mean difference of 0 (i.e., perfect agreement).

[c]Mean *z* scores (*x* axis) are calculated by averaging the *z* scores across settings.