Title

Intermachine differences in DXA measurements vary by skeletal site, and impact the assessment of low bone density in children.

Authors

Zemel, Babette
Wasserman, Halley
Kelly, Andrea
et al.

Publication Date

2020-12-01

DOI

10.1016/j.bone.2020.115581

Peer reviewed

HHS Public Access

Author manuscript

Bone. Author manuscript; available in PMC 2021 December 01.

Published in final edited form as:

Bone. 2020 December ; 141: 115581. doi:10.1016/j.bone.2020.115581.

Intermachine Differences In DXA Measurements Vary By Skeletal Site, And Impact The Assessment Of Low Bone Density In Children.

Babette S. Zemel[1,2], **Halley Wasserman**[3,4], **Andrea Kelly**[2,5], **Bo Fan**[6], **John Shepherd**[7], **Joan Lappe**[8], **Vicente Gilsanz**[9], **Sharon Oberfield**[10], **Karen K. Winer**[11], **Heidi J. Kalkwarf**[12,4]

[1]Division of Gastroenterology, Hepatology and Nutrition, The Children's Hospital of Philadelphia, Philadelphia

[2]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia

[3]Division of Endocrinology, Cincinnati Children's Hospital Medical Center, Cincinnati

[4]Department of Pediatrics, University of Cincinnati College of Medicine

[5]Division of Endocrinology and Diabetes, The Children's Hospital of Philadelphia, Philadelphia

[6]Department of Radiology, University of California San Francisco, San Francisco

[7]University of Hawaii Cancer Center, University of Hawaii (Manoa)

[8]Division of Endocrinology, Department of Medicine, Creighton University, Omaha

[9]Departments of Orthopaedic Surgery and Radiology, Children's Hospital Los Angeles, Los Angeles, CA

[10]Division of Pediatric Endocrinology, Diabetes, and Metabolism, Department of Pediatrics, Columbia University Medical Center, New York

[11]Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD

[12]Division of Gastroenterology, Hepatology and Nutrition, Cincinnati Children's Hospital Medical Center

Abstract

Corresponding author: Babette S. Zemel, PhD, Division of Gastroenterology, Hepatology and Nutrition, The Children's Hospital of Philadelphia, 2716 South Street, rm 14471, Philadelphia, PA 19146, Phone: 215-590-1669, zemel@email.chop.edu.

Credit Author Statement: Babette S. Zemel: conceptualization, formal analysis, data curation, methodology, writing original draft; Halley Wasserman: conceptualization, writing review and editing; Andrea Kelly: data curation, conceptualization, writing review and editing; Bo Fan: data curation, writing review and editing; John Shepherd: data curation, conceptualization writing review and editing; Joan Lappe: data curation, writing review and editing; Vicente Gilsanz: data curation; Sharon Oberfield: data curation, conceptualization, writing review and editing; Karen K. Winer: data curation, conceptualization, project administration, funding acquisition, writing review and editing;, Heidi J. Kalkwarf: conceptualization, formal analysis, data curation, methodology, writing original draft.

Background: Bone mineral content (BMC) and areal-bone mineral density (aBMD) measurements of the lumbar spine (LS) and whole body less head (WBLH) by dual energy x-ray absorptiometry (DXA) are recommended for bone health assessment in children. Intermachine differences were not considered previously in formulating these recommendations.

Methodology: DXA measurements of the LS, WBLH, total hip, femoral neck and distal 1/3 radius from the Bone Mineral Density in Childhood Study were examined. Healthy children, ages 6 to 16 years, from five clinical centers participated. The same spine, whole body, and femur phantoms were measured on each Center's DXA machine. Percentage of individuals with low BMC or aBMD (Z-score<−1.5) was determined. Clinical center differences were evaluated by analysis of covariance adjusting for height and BMI Z-score, calcium intake, physical activity, Tanner stage and bone age. Logistic regression assessed odds of low BMC or aBMD across clinical centers.

Results: Significant differences among Clinical Centers ($p<0.05$) were evident in adjusted mean BMC and aBMD Z-scores (n=1503) for all skeletal sites. WBLH BMC and aBMD Z-scores had the greatest range across centers (−0.13 to 0.24, and −0.17 to 0.56, respectively). The percentage of children with Z-scores less than −1.5 varied among Clinical Centers from 1.9 [95%CI 0.8, 4.5] to 8.1 [95%CI 5.7, 11.3] for WBLH BMC, 1.1 [95%CI 0.4, 3.5] to 6.3 [95%CI 3.8, 10.1] for WBLH aBMD, and from 4.4 [95%CI 2.8, 7.0] to 12.6 [95%CI 9.3, 16.9] for distal 1/3 radius aBMD. For each skeletal site except total hip aBMD and femoral neck BMC, at least one center had significantly lower odds of low bone density.

Conclusions: By design, our reference ranges capture intermachine variability. Most clinical centers don't know where their machine falls within the range of intermachine variability, and this may affect diagnosis of children evaluated for conditions that threaten bone health. Total hip scans showed the least, and whole body scans showed the most intermachine variability. Pediatric bone health assessment recommendations should recognize intermachine differences and address this important issue.

Keywords

Bone density; dual energy x-ray absorptiometry; children; intermachine variability

1. Introduction:

Assessment of areal bone mineral density (aBMD) by dual energy x-ray absorptiometry (DXA) is an essential component of bone health evaluation in children with medical conditions that threaten bone mineral accrual. DXA is recommended for use in children because of the very low radiation exposure, rapid scan times, and wide availability. Current DXA technology can accurately measure bone density even in infants and young children. The International Society for Clinical Densitometry (ISCD) Pediatric Positions state that the preferred skeletal sites for bone health assessment in pediatric patients are the lumbar spine (LS) and the whole body less head (WBLH), although alternative skeletal sites can be used when these are not feasible[1, 2]. The WBLH and LS skeletal sites were selected as preferred sites for several reasons. First, robust reference data are available, which is essential for determining age- and sex-specific Z-scores for bone health assessment[3, 4].

Scan precision (i.e., the closeness of two or more measurements to each other) at both skeletal sites is excellent, and skeletal landmarks for scan analysis are easily identified.

Other technical issues were not considered in the pediatric recommendations. Routine monitoring of DXA equipment is performed by daily scanning of a machine-specific spine phantom. The machine-specific spine phantom is also used at the time of equipment repairs to confirm proper calibration of the device. However, beyond the use of the spine phantom, no standard procedures in clinical practice are in place to assure calibration for other scan types. Moreover, intermachine differences in bone mineral content (BMC) and aBMD measurements have not been considered previously as a criterion for selecting optimal scan sites for bone health assessment in children. It is known that measurements of the same individual on different machines yield different results [5], but it is unknown whether these differences are consistent across skeletal sites.

To address this knowledge gap we evaluated the magnitude of intermachine differences from the multi-center Bone Mineral Density in Childhood Study (BMDCS). The BMDCS enrolled healthy children at five clinical centers in the United States. The study generated robust reference data for use with Hologic DXA equipment that allow clinicians to determine whether aBMD or BMC is appropriate for the child's age, sex and race[3, 4]. Here we compared unadjusted and adjusted BMC and aBMD-Z, as well as risk of having "low" BMD-Z across the clinical centers to assess the potential impact of intermachine differences on 1) BMC and aBMD Z-score measures and 2) classification of children as having low bone density at the lumbar spine, total hip, femoral neck, whole body or forearm. We also compared measurements of the same spine, whole body and femur phantoms scanned at each clinical center as an independent measure of machine calibration.

2. Sample and Methods:

2.1 Sample and study design

Study participants and the methods for acquiring DXA scans in the BMDCS have been described previously[3]. Healthy children (n=1554), ages 6 to 16 years, were enrolled at five clinical centers (Children's Hospital of Los Angeles, Creighton University, Cincinnati Children's Hospital Medical Center, The Children's Hospital of Philadelphia, Columbia University). By design, these five centers were selected to capture geographic variability, and the resulting reference ranges did not account for center differences in sample characteristics or intermachine differences. Children were evaluated annually for up to seven years. An additional cohort of children, 5 and 19 years of age (n=450), were later added to the sample, resulting in over 10,000 observations.

For participants less than 18 years of age, a parent/guardian provided written informed consent and the participant provided assent. Participants 18 years of age or older provided written informed consent. All study protocols and procedures were approved by the Institutional Review Board for Human Subjects at each clinical center.

For the purposes of this analysis, only data from the first year of data collection, for participants 6 to 16 years of age, were included for the following reasons: first, this approach

avoids the non-independence of observations that would result from including multiple observations per subject from the longitudinal dataset. Second, it permits comparison with the phantom calibration results acquired at study initiation. Observations were restricted to individuals who had complete data from the initial study visit for all DXA measurements and covariates listed below (n=1503). These observations represent ~15% of the scans used to generate the final reference curves.

2.2 Data collection

Data collection included height and weight measurements, dietary calcium intake and physical activity by questionnaire, Tanner stage of sexual maturation by exam, and race/ethnicity by self-report. Bone age was determined from hand-wrist radiographs evaluated by a single expert pediatric radiologist and scored using the Greulich and Pyle Atlas[6]. Height, weight and BMI (weight/height^2) Z-scores were calculated using the CDC 2000 growth curves[7]. DXA scans of the whole body, lumbar spine, hip and non-dominant forearm were acquired. Four clinical centers used the Hologic 4500A model and one center used the Hologic 4500W model (Hologic, Inc, Malborough, MA). All technicians were experienced in scan acquisition in children, and followed standard procedures. Scans were analyzed centrally (University of California, San Francisco). Age, sex and ancestry-specific reference ranges for the LS, total hip, femoral neck, WBLH and distal 1/3 radius from this study have been published[4], and were used to calculate BMC and aBMD Z-scores.

2.3 DXA calibration

Intermachine calibration across the five clinical centers was assessed at study onset using a single set of phantoms that included the European Spine Phantom (QRM, Moehrendorf, Germany)[8], and Hologic femur[9] and whole-body phantoms (Hologic, Inc., Marlborough, MA)[10] that were circulated among the clinical centers. Each phantom was scanned 10 times at each clinical center. Phantom scans were analyzed for aBMD and BMC.

2.3 Statistical analysis

Analysis of Variance (ANOVA) or the Chi^2 test (for categorical variables) was used to test for mean differences in participant characteristics and BMC and aBMD Z-scores among clinical centers. Evidence of intermachine differences was assessed using analysis of covariance (ANCOVA) to account for covariates known to be associated with BMC and aBMD and that could possibly contribute to bone Z-score differences among clinical centers: height Z-score, BMI Z-score, calcium intake, physical activity, Tanner stage and bone age. Age, sex, and population ancestry were not included in these models as these effects are accounted for in the calculation of BMC and aBMD Z-scores. To examine the effect of intermachine differences on identification of individuals with low aBMD, we calculated the percentage of participants with BMC or aBMD Z-scores less than −1.5, which represents 6.7% of a normal distribution. We chose a cut-off of −1.5 because BMC and aBMD reference ranges were established based on this study, so there were very few individuals with Z-scores less than −2. We used simple and multivariable logistic regression to estimate the odds ratios (OR) and 95% confidence intervals [CI] of having a low BMC or aBMD Z-score relative to Center 2. Center 2 was selected because among the four centers that used 4500A DXA models, mean Z-scores for this Center were most consistently similar

to the reference ranges across multiple skeletal sites. The multivariable logistic regressions accounted for the covariates listed above.

3. Results:

Participant characteristics for the overall sample and by clinical center are shown in Table 1. There were significant differences among clinical centers in population ancestry, Tanner stage, weight, BMI, calcium intake, physical activity, and bone age. There were no significant differences among centers for age, sex, weight Z-score, height, height Z-score or BMI Z-score. The mean (unadjusted) age-, sex- and race-specific Z-scores differed among centers for WBLH BMC and aBMD, and 1/3 radius BMC and aBMD, but not for the other bone Z-scores (Figure 1 and supplementary Table).

To further explore clinical center differences in BMC and aBMD Z-scores, we used ANCOVA to adjust for potential confounders (height and BMI Z-scores, dietary calcium intake, total physical activity, Tanner stage and bone age). The unadjusted and adjusted mean Z-scores [95% Confidence Interval] by clinical center are shown in Figure 1 (and Supplementary Table). In general, the adjusted mean Z-scores had more narrow confidence intervals than the unadjusted Z-scores. Significant differences among centers ($p<0.05$) were noted in adjusted BMC and aBMD Z-scores for all scan sites. The range in adjusted mean BMC Z-scores was greater for WBLH (−0.14 to 0.24) and distal 1/3 radius (−0.21 to 0.22) compared to spine (−0.13 to 0.03), total hip (0.06 to 0.20) and femoral neck (−0.05 to 0.07). The range in adjusted aBMD Z-scores was also larger for WBLH (−0.17 to 0.56) and distal 1/3 radius (−0.20 to −0.05) compared to spine (−0.15 to 0.06), total hip (−0.04 to 0.10) and femoral neck (−0.13 to 0.07). No clinical center was consistently higher or lower for all skeletal sites. For example, Center 1 tended to have higher Z-scores compared to other clinical centers for distal 1/3 radius BMC and WBLH aBMD and BMC, whereas Center 5 tended to have higher Z-scores for the total hip and femoral neck. The distribution of Z-scores, both unadjusted and adjusted for covariates, was significantly different from zero for spine BMC for Center 2; for WBLH BMC for Center 1 and 4; for WBLH aBMD for Center 1, 2, and 4; for total hip BMC at Center 5; distal 1/3 radius BMC for Center 1, 2 and 4; and for Distal 1/3 Radius aBMD for Center 5.

The percentage of the sample with an unadjusted Z-score <−1.5 was not significantly different among clinical centers for most skeletal sites (Table 2). Significant differences were evident for WBLH BMC where percentages ranged from 1.9 [95% CI 0.8, 4.5] to 8.1 [95% CI 5.7, 11.3], WBLH aBMD where percentages ranged from 1.1 [95%CI 0.4, 3.5] to 6.3 [95% CI 3.8, 10.1], femoral neck aBMD where percentages ranged from 2.7 [95% CI 1.4, 5.4] to 7.8 [95% CI 5.5, 11.0], and distal 1/3 radius aBMD where percentages ranged from 4.4 [95% CI 2.8, 7.0] to 12.6 [95% CI 9.3, 16.9]. We used logistic regression to adjust for covariates to estimate the odds ratios for having a low Z-score (Table 3). Adjusted odds ratios were significantly different from from the reference group for at least one clinical center for all skeletal sites except for total hip aBMD and femoral neck BMC. Lower odd ratios correspond to centers that had a higher adjusted mean Z-score. For example, WBLH BMC and aBMD Z-scores generated from scans at Center 1 were significantly less likely to

be <−1.5 than the reference group, and spine BMC and aBMD Z-scores from Center 5 were significantly less likely to be <−1.5.

BMC and aBMD results from the circulating phantoms are presented in Table 4. The differences between the highest and lowest values were smallest for the femur phantom (1.9 and 2.4% of group mean for total hip BMC and aBMD, respectively) and largest for BMC of the whole body phantom (14% of the group mean). Of note, one clinical center had a 4500W model whereas the other centers used 4500A models. Whole body scan acquisition with the 4500W model is different from the 4500A, but the scan acquisition for other skeletal sites is the same across models. Excluding the 4500W model, the difference across centers in whole body BMC goes from 14% to 3.7%. The inter-site coefficient of variation was 0.9% for total hip and 5.8% for whole body BMC when the 4500W was included in the comparison. When the 4500W was excluded from this comparison the coefficient of variation was 1.7% for whole body BMC and 0.5% for total hip BMC.

4. Discussion

This study demonstrated significant intermachine differences in BMC and aBMD outcomes using *in vivo* (i.e., comparison of cohorts) and *in vitro* (i.e., comparison of phantoms) approaches that are relevant to the assessment and diagnosis of compromised bone health in children. The 2007 and 2013 ISCD Pediatric Position consensus statements established criteria for the diagnosis of osteoporosis in children, and stated that DXA measurements of BMC and/or aBMD are an integral part of pediatric bone health assessment[1, 11]. The recommended sites for clinical assessment were the lumbar spine and WBLH, with additional sites for consideration under a variety of circumstances[1, 2]. While the ISCD recommendations were cautious to note that the diagnosis of osteoporosis in children and adolescents should not be made on the basis of densitometric criteria alone, DXA measurements are a major contributing factor in diagnosis and are particularly important for monitoring treatment and disease effects on bone health. The recommendations also emphasized the importance of good scan acquisition technique in children, knowledge of the least significant change (LSC), and adequacy of pediatric reference data. An additional dimension that was not previously considered is the measurement agreement between DXA devices at different skeletal sites, especially with those used to generate reference data. As we have shown, differences between DXA devices can impact the identification of low bone density in children.

Our analyses from the multi-center BMDCS show significant intermachine variability in BMC and aBMD results on Hologic DXA systems, especially for the clinically recommended skeletal site of the WBLH. After adjusting for relevant covariates, two out of four clinical centers had significantly lower odds of reporting a WBLH BMC Z-score <−1.5. For all other scan sites, except the total hip aBMD and femoral neck BMC, at least one clinical center had a significantly lower odds of an aBMD Z-score <−1.5. Of note, total hip aBMD is recommended as an alternative site by the most recent ISCD Pediatric Positions[2] when a spine or total body scan are not feasible, or for continuity in the transition to adulthood. Our results suggest that hip scans provide the most consistent results across Hologic DXA devices.

The differences between clinical centers were also apparent from the results of phantom scans. Most notably, there were large differences in whole body aBMD across centers. One center used a 4500W DXA model, while all others used a 4500A, and this accounts for large differences in whole body BMC values across sites. The difference as a percent of the overall mean was 14.0% when the 4500W is included, and 3.7% when the values from this device were exclude. For whole body phantom aBMD measurements, there was a 5.5% difference across clinical centers excluding the 4500W device. The 4500W is a narrow fan beam system that takes seven passes to cover the whole body. Each fan pass overlaps the neighboring passes. The 4500A is a wide fan beam system that only takes 3 passes to cover the whole body without overlap. These differences in scan acquisition likely account for the large variance between centers in whole body results.

Other phantom results showed less variability across clinical centers, yet there was still 4% or more variability for aBMD at both the spine and the whole body, the clinically recommended scan sites for children. Differences between DXA machines are known to occur [5]. However, as we show here, this degree of variability is concerning because it may result in major differences between centers in the identification of children with low bone density. A given child if measured at one facility may be diagnosed with low bone density, whereas if measured at another facility, the results may fall into a normal range. For example, a 12 year old non-black female with a WBLH BMC of 748 g from one DXA device would have a Z-score of −2.0. The same child measured on a device that produces results that are 14% higher value (864 g) would have a Z-score of −1.26. These findings have implications for diagnosing osteoporosis in children as well as for monitoring BMC and aBMD. For example, the European Cystic Fibrosis Society recommends obtaining DXA scans every 5 years in patients with an aBMD Z-score ≤−1.0, every 2 years in those with Z-scores between −1.0 and −2.0, and annually for those patients with Z-scores <−2.0 [12]. Thus, even a small difference in aBMD measurement could potentially impact frequency of follow-up and medical decision making.

The data used to create the BMDCS pediatric reference curves were not adjusted for clinical center differences, with the recognition that some machine variability exists. An individual clinical DXA center may not know the degree to which their DXA device deviates from those used to generate the reference curves. Indeed, discrepancies in aBMD Z-scores generated by different reference data bases have been reported[13, 14] and are most extreme for whole body BMD Z-scores[13]. Based on our findings reported here, it is likely that even when BMC or aBMD results are compared to reference data acquired on the same DXA models, there will be discrepancies due to specific intermachine differences between the clinical DXA and the DXA devices used to generate the reference ranges.

Sources of variability in DXA measurements include intermachine differences, within-scanner precision errors, such as quantum noise, variations in patient positioning, and inhomogeneities in soft tissue composition[15]. Dowthwaite et al. [16] compared measurements on 130 females ages 8 to 24y obtained on Hologic 4500A and Discovery models. They observed a high concordance between measurements from the two scanners, but systematic differences were evident. Combined, these findings highlight the pitfall of using a measurement cut-off designed to make diagnosis or treatment decisions, especially if

based on a single measurement. Disparities in Z-scores generated from different reference databases adds to the concern over the use of a measurement cut-off in diagnosis or treatment[14, 17]. Examination of trends over time in an individual measured on the same DXA machine may be useful, even though Z-score values may be higher or lower than a traditional cut-off [18].

As yet, unidentifed sample characteristics may explain the differences we observed between clinical centers, despite having controlled for height and BMI Z-scores, calcium intake, physical activity, Tanner stage and bone age. Differences in the prevalence of osteoporosis and osteopenia in adults were reported for the National Health and Examination Surveys in 1988–94 vs. 2005–2006. These differences could not be explained by changes in scanners, BMI distribution of the samples, or bisphosphonate use[19]. We also observed differences between the participants characteristics at our clinical centers in population ancestry, indicators of biological maturation (Tanner stage and bone age), weight Z-score and lifestyle measures. We adjusted for these factors in the comparisons between centers, yet significant center differences remained. The measurement of phantoms avoided these possible sources of bias, and demonstrated differences between devices, especially for the whole body DXA measures.

Our findings also have implications for research studies of pediatric diseases that do not include a control group of otherwise healthy children. In the absence of a control group, studies that evaluate disease effects by comparing values for a patient group to published reference ranges could potentially lead to erroneous conclusions. Inclusion of a study group of healthy children measured on the same device would provide a reference for how the device performs relative to the reference ranges, as well as possible regional differences. For example, Meeuwes et al reported LS aBMD Z-scores <−2.0 in 33% of females and 47% of males with sickle cell disease from a single center [20]. As no control group was enrolled in this study, it is unknown whether these results accurately represent the bone health status of children with sickle cell disease or are skewed because the DXA machine characteristics differ from those of the device(s) used to acquire the specific reference data on which the Z-scores are based.

Our results further emphasize the importance of acquiring follow-up scans on the same device. The least significant change (LSC) between two DXA aBMD measurements on the same machine for children and adolescents in this study, calculated from previously reported percisions errors, ranged from 2.6% for the spine and total hip aBMD to 4.6% for the 1/3 radius aBMD. The added intermachine variability in spine and WBLH, raises doubts about any meaningful interpretation of change in bone health status if follow-up scans are obtained on a different DXA device. The ISCD Adult Position Statement emphasizes that it is not possible to quantitatively compare aBMD measurements between facilities without ensuring cross-calibration [21]. Our study reiterates this point in a pediatric population.

The ISCD Pediatric Positions recommended acquiring scans of the spine and whole body based on the reproducibility of these measurements, the ease of identifying skeletal landmarks, and the availability of normative data. In addition to publishing normative data[4], we previously reported the coefficient of variation (precision) for the total hip and

femoral neck aBMD as 0.85% and 1.29%, respectively[22]. These values are comparable to the values obtained for spine and WBLH aBMD (0.85% and 0.95%, respectively), even among children in the 6 to 9 year age group. Also, studies have demonstrated the responsiveness of the total hip to vitamin D supplementation[23] and weight-bearing exercise (see for example and [24]). The consistency in total hip measurements across clinical centers provides further evidence that this skeletal site should be considered for bone health assessment in children.

The main limitation of this study is that participants were not measured on each of the DXA devices to conduct a direct comparison of the machines. We made the assumption that by adjusting for covariates, we should have similar average Z-scores at each site. It is possible that other sample characteristics not accounted for might contribute to clinical center differences. However, we included results from phantoms that were measured on each machine to confirm the overall pattern of differences between DXA devices.

Our findings provide crucial evidence for the pediatric clinical care community to consider intermachine differences in pediatric recommendations and multi-center studies. Numerous sources of variability in DXA measurements should be considered when using a cut-off value to categorize low bone density in children. DXA manufacturers need to reduce intermachine variability and thereby improve the ability to detect low bone density and monitor treatment in children. Our findings also suggest that total hip measurements should be considered as a skeletal site for bone health assessment in children due to the consistency in measurements across clinical centers.

5. Conclusions:

Significant intermachine variability exists and has the potential to affect diagnosis of children being evaluated for conditions that threaten bone health. The implications of our findings include: (1) intermachine differences create statistically significant differences between centers in the likelihood of being diagnosed with low bone density; (2) multi-center studies should assess intermachine differences as part of study design and analysis; (3) whole body measurements have the largest and total hip aBMD have the smallest intermachine differences; (4) in order to assess the magnitude of deficits when studying pediatric disease groups, it is imperative to include a local control group as it is unknown how their machine deviates from the reference population; (5) pediatric recommendations for scan sites should recognize intermachine differences and make efforts to address this important and frequently overlooked issue.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Support: This work was funded by the National Institute of Child Health and Human Development (NICHD) contracts NO1-HD-1–3228, −3329, −3330, −3331, −3332, and −3333, the Clinical and Translational Research Center Grant 5-MO1-RR-000240 and UL1 RR-026314, the Nutrition Center of the Children’s Hospital of

Philadelhia, and the Children's Hospital of Philadelphia Research Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Crabtree NJ, et al., Dual-energy X-ray absorptiometry interpretation and reporting in children and adolescents: the revised 2013 ISCD Pediatric Official Positions. J Clin Densitom, 2014 17(2): p. 225–42. [PubMed: 24690232]
2. Weber DR, et al., The Utility of DXA Assessment at the Forearm, Proximal Femur, and Lateral Distal Femur, and Vertebral Fracture Assessment in the Pediatric Population: 2019 ISCD Official Position. J Clin Densitom, 2019 22(4): p. 567–589. [PubMed: 31421951]
3. Kalkwarf HJ, et al., The bone mineral density in childhood study: bone mineral content and density according to age, sex, and race. J Clin Endocrinol Metab, 2007 92(6): p. 2087–99. [PubMed: 17311856]
4. Zemel BS, et al., Revised reference curves for bone mineral content and areal bone mineral density according to age and sex for black and non-black children: results of the bone mineral density in childhood study. J Clin Endocrinol Metab, 2011 96(10): p. 3160–9. [PubMed: 21917867]
5. Jankowski LG, et al., Cross-calibration, Least Significant Change and Quality Assurance in Multiple Dual-Energy X-ray Absorptiometry Scanner Environments: 2019 ISCD Official Position. J Clin Densitom, 2019 22(4): p. 472–483. [PubMed: 31558404]
6. Greulich WW and Pyle SI, Radiographic atlas of skeletal development of the hand and wrist. 1959: Stanford University Press.
7. Kuczmarski RJ, et al., CDC growth charts: United States. Adv Data, 2000(314): p. 1–27.
8. Kalender WA, et al., The European Spine Phantom--a tool for standardization and quality control in spinal bone mineral measurements by DXA and QCT. Eur J Radiol, 1995 20(2): p. 83–92. [PubMed: 7588873]
9. Slosman DO, et al., The use of different dual X-ray absorptiometry brands in a multicenter clinical trial: consequences and limits. J Clin Densitom, 1999 2(1): p. 37–44. [PubMed: 23547312]
10. Blunt BA, et al., Hologic whole body phantom: using in vitro data to correct in vivo whole body data. . J Bone Mineral Res, 2000 15(1): p. SU278.
11. Gordon CM, et al., Dual energy X-ray absorptiometry interpretation and reporting in children and adolescents: the 2007 ISCD Pediatric Official Positions. J Clin Densitom, 2008 11(1): p. 43–58. [PubMed: 18442752]
12. Sermet-Gaudelus I, et al., European cystic fibrosis bone mineralisation guidelines. J Cyst Fibros, 2011 10 Suppl 2: p. S16–23. [PubMed: 21658635]
13. Kocks J, et al., Z-score comparability of bone mineral density reference databases for children. J Clin Endocrinol Metab, 2010 95(10): p. 4652–9. [PubMed: 20668038]
14. Ma J, et al., The choice of normative pediatric reference database changes spine bone mineral density Z-scores but not the relationship between bone mineral density and prevalent vertebral fractures. J Clin Endocrinol Metab, 2015 100(3): p. 1018–27. [PubMed: 25494661]
15. Blake GM and Shepherd JA, Effect of random BMD measurement errors on diagnostic classification using T-scores. J Clin Densitom, 2007 10(4): p. 415–7; author reply 417–8. [PubMed: 17993402]
16. Dowthwaite JN, et al., Cross-Calibrated Dual-Energy X-Ray Absorptiometry Scanners Demonstrate Systematic Bias in Pediatric and Young Adult Females. J Clin Densitom, 2018 21(2): p. 281–294. [PubMed: 28258886]
17. Leonard MB, et al., Discrepancies in pediatric bone mineral density reference data: potential for misdiagnosis of osteopenia. J Pediatr, 1999 135(2 Pt 1): p. 182–8. [PubMed: 10431112]
18. Ward LM, et al., A Contemporary View of the Definition and Diagnosis of Osteoporosis in Children and Adolescents. J Clin Endocrinol Metab, 2019.
19. Looker AC, et al., Prevalence and trends in low femur bone density among older US adults: NHANES 2005–2006 compared with NHANES III. J Bone Miner Res, 2010 25(1): p. 64–71. [PubMed: 19580459]

20. Meeuwes M, et al., Bone mineral density, growth, pubertal development and other parameters in Brazilian children and young adults with sickle cell anaemia. Trop Med Int Health, 2013 18(12): p. 1539–46. [PubMed: 24134458]

21. Schousboe JT, et al., Executive summary of the 2013 International Society for Clinical Densitometry Position Development Conference on bone densitometry. J Clin Densitom, 2013 16(4): p. 455–66. [PubMed: 24183638]

22. Shepherd JA, et al., Optimal monitoring time interval between DXA measures in children. J Bone Miner Res, 2011 26(11): p. 2745–52. [PubMed: 21773995]

23. El-Hajj Fuleihan G, et al., Effect of vitamin D replacement on musculoskeletal parameters in school children: a randomized controlled trial. J Clin Endocrinol Metab, 2006 91(2): p. 405–12. [PubMed: 16278262]

24. Mackelvie KJ, et al., A school-based exercise intervention augments bone mineral accrual in early pubertal girls. J Pediatr, 2001 139(4): p. 501–8. [PubMed: 11598595]

Highlights

- Intermachine differences in DXA measurements in children vary according to skeletal site
- Whole body measurements showed the greatest and total hip measurements showed the least intermachine variation.
- Intermachine variation can have a significant impact on the identification of low bone density in children.
- Recommendations for pediatric clinical bone health assessment by DXA should consider the effect of intermachine variability at different skeletal sites.

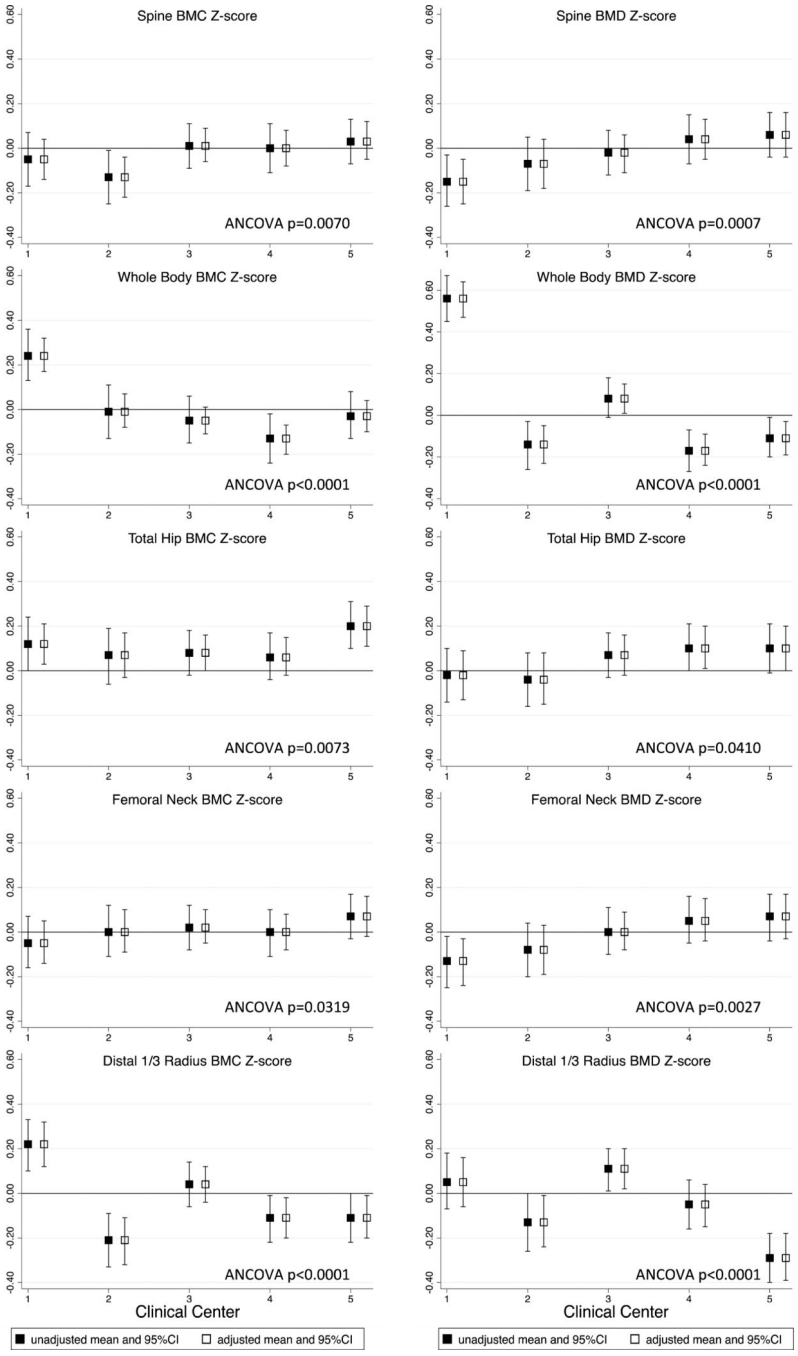

Figure 1.

Mean [95% CI] BMC and aBMD Z-Scores For Each Clinical Center Unadjusted and Adjusted for Covariates [Height Z-score, BMI Z-score, Calcium Intake and Physical Activity]

Table 1

Participant characteristics by clinical center.

	Clinical Center						
	All [n=1503]	**Center 1 [n=263]**	**Center 2 [n=240]**	**Center 3 [n=383]**	**Center 4 [n=324]**	**Center 5 [n=293]**	**p-value[1]**
Demographics	**n [%]**	**n [%]**	**n [%]**	**n [%]**	**n [%]**	**n [%]**	
Sex, Female[2]	764 [51]	138 [53]	122 [51]	191 [50]	163 [50]	150 [51]	ns
Population Ancestry[2]							p<0.0001
African	361 [24]	90 [34]	47 [20]	65 [17]	115 [36]	44 [15]	
Hispanic	256 [17]	11 [4]	72 [30]	153 [40]	3 [1]	17 [6]	
European	696 [46]	144 [55]	87 [36]	57 [15]	188 [58]	220 [75]	
Other	190 [13]	18 [7]	34 [14]	108 [28]	18 [6]	12 [4]	
Tanner Stage[2]							p<0.0001
Stage 1	654 [44]	124 [47]	107 [45]	146 [38]	137 [42]	140 [48]	
Stage 2	203 [13.5]	17 [8]	28 [14]	38 [19]	68 [34]	52 [26]	
Stage 3	153 [10]	27 [18]	20 [13]	42 [27]	33 [22]	31 [20]	
Stage 4	205 [14]	39 [19]	44 [21]	50 [24]	39 [19]	33 [16]	
Stage 5	288 [19]	56 [21]	41 [17]	107 [28]	47 [15]	38 [13]	
	mean [sd]	**mean [sd]**	**mean [sd]**	**mean [sd]**	**mean [sd]**	**mean [sd]**	
Age, y	11.0 [3.1]	11.0 [3.1]	10.8 [2.9]	11.3 [3.2]	10.7 [3.0]	10.9 [3.0]	ns
Height, cm	144.8 [18.2]	145.5 [18.3]	144.1 [17.3]	146.8 [19.1]	143.6 [17.5]	143.7 [18.3]	ns
Weight, kg	41.2 [15.7]	41.3 [15.4]	40.2 [15.1]	43.5 [16.6]	40.1 [15.2]	40.0 [15.3]	0.013
BMI, kg/m^2	18.8 [3.3]	18.7 [3.1]	18.6 [3.2]	19.3 [3.5]	18.6 [3.3]	18.6 [3.2]	0.011
Height Z-Score	0.1 [0.8]	0.2 [0.8]	0.1 [0.8]	0.1 [0.9]	0.2 [0.8]	0.1 [0.8]	ns
Weight Z-Score	0.3 [0.8]	0.3 [0.8]	0.3 [0.8]	0.4 [0.8]	0.3 [0.8]	0.2 [0.8]	0.09
BMI Z-Score	0.3 [0.8]	0.3 [0.8]	0.3 [0.8]	0.4 [0.9]	0.3 [0.9]	0.3 [0.8]	ns
Calcium intake, mg/d	921 [564]	850 [493]	919 [610]	921 [631]	872 [545]	1042 [492]	0.0004
Physical activity, h/wk	17.7 [14.7]	23.0 [17.5]	17.4 [13.1]	18.1 [14.6]	13.4 [12.0]	17.4 [14.3]	p<0.0001
Bone age, y	11.1 [3.4]	11.1 [3.3]	10.9 [3.4]	11.7 [3.7]	10.9 [3.2]	10.8 [3.4]	0.007

Author Manuscript

Author Manuscript

Author Manuscript

[1] Overall p-value from analysis of variance or

[2] Chi2 test for differences among groups.

Table 2

Percentage of children with low Z-scores [Z < −1.5].

	Clinical Center						
	All	**Center 1**	**Center 2**	**Center 3**	**Center 4**	**Center 5**	
Percentage with low Zscores [< −1.5]	**Percentage [95% CI]**	**Percentage [95% CI]**	**Percentage [95% CI]**	**Percentage [95% CI]**	**Percentage [95% CI]**	**Percentage [95% CI]**	**P for Center Differences**
Spine BMC	6.3 [5.1, 7.7]	7.2 [4.7, 11.1]	7.5 [4.8, 11.6]	**7.8** [5.5, 11.0]	5.9 [3.8, 9.0]	**3.1** [1.6, 5.8]	ns
Spine aBMD	6.4 [5.3, 7.7]	7.6 [5.0, 11.5]	**8.3** [5.4, 12.6]	7.0 [4.9, 10.1]	6.2 [4.0, 9.4]	**3.1** [1.6, 5.8]	ns
WBLH BMC	5.6 [4.5, 6.9]	**1.9** [0.8, 4.5]	6.3 [3.8, 10.1]	**8.1** [5.7, 11.3]	7.7 [5.3, 11.2]	2.7 [1.4, 5.4]	0.0003
WBLH aBMD	4.8 [3.9, 6.1]	**1.1** [0.4, 3.5]	**6.3** [3.8, 10.1]	6.0 [4.0, 8.9]	6.5 [4.3, 9.7]	3.8 [2.1, 6.7]	0.0035
Total Hip BMC	4.9 [3.9, 6.1]	4.6 [2.6, 7.9]	**7.1** [4.4, 11.1]	5.7 [3.8, 8.6]	4.9 [3.0, 7.9]	**2.4** [1.1, 4.9]	ns
Total Hip aBMD	5.1 [4.1, 6.4]	5.3 [3.2, 8.8]	**6.3** [3.8, 10.1]	6.3 [4.2, 9.2]	4.0 [2.3, 6.8]	**3.8** [2.1, 6.7]	ns
Femoral Neck BMC	5.0 [4.0, 6.2]	5.3 [3.2, 8.8]	4.6 [2.6, 8.1]	**6.8** [4.7, 9.8]	4.6 [2.8, 7.5]	**3.1** [1.6, 5.8]	ns
Femoral Neck aBMD	6.3 [5.1, 7.6]	7.6 [5.0, 11.5]	7.1 [4.4, 11.1]	**7.8** [5.5, 11.0]	5.9 [3.8, 9.0]	**2.7** [1.4, 5.4]	0.0340
1/3 Radius BMC	6.1 [5.0, 7.5]	**4.6** [2.6, 7.9]	**8.8** [5.8, 13.1]	4.7 [3.0, 7.3]	7.1 [4.8, 10.5]	6.1 [3.9, 9.5]	ns
1/3 Radius aBMD	8.1 [6.8, 9.6]	7.6 [5.0, 11.5]	9.2 [6.1, 13.5]	**4.4** [2.8, 7.0]	8.0 [5.5, 11.5]	**12.6** [9.3, 16.9]	0.0036

Bolded values represent the lowest and highest values across Centers

It is expected that 6.7 percent of the sample should have a Z score < −1.5

Bone. Author manuscript; available in PMC 2021 December 01.

Table 3

Unadjusted and adjusted odds ratios of low BMC or aBMD Z-score [Z < −1.5] by clinical center.

		Clinical Center				
		Center 1	**Center 2**	**Center 3**	**Center 4**	**Center 5**
Skeletal Site		**Odds Ratio [95% CI]**	**Reference Group**	**Odds Ratio [95% CI]**	**Odds Ratio [95% CI]**	**Odds Ratio [95% CI]**
Spine BMC	unadjusted	0.97 [0.50,1.90]	1.00	1.06 [0.58, 1.95]	0.78 [0.40, 1.52]	**0.39 [0.17, 0.90]**
	adjusted[1]	0.97 [0.45, 2.07]	1.00	1.15 [0.58, 2.29]	0.68 [0.32, 1.45]	**0.30 [0.12, 0.74]**
Spine aBMD	unadjusted	0.92 [0.48, 1.75]	1.00	0.85 [0.46, 1.54]	0.73 [0.39, 1.40]	**0.35 [0.16, 0.79]**
	adjusted[1]	1.04 [0.52, 2.08]	1.00	0.86 [0.45, 1.64]	0.62 [0.31, 1.23]	**0.31 [0.13, 0.71]**
WBLH BMC	unadjusted	**0.29 [0.11, 0.82]**	1.00	1.34 [0.71, 2.54]	1.27 [0.65, 2.47]	0.43 [0.18, 1.02]
	adjusted[1]	**0.26 [0.08, 0.81]**	1.00	1.41 [0.67, 2.99]	1.18 [0.54, 2.58]	**0.32 [0.12, 0.85]**
WBLH aBMD	unadjusted	**0.18 [0.05, 0.61]**	1.00	0.97 [0.50, 1.90]	1.05 [0.53, 2.09]	0.59 [0.27, 1.31]
	adjusted[1]	**0.15 [0.04, 0.57]**	1.00	0.95 [0.45, 1.99]	0.89 [0.42, 1.93]	0.46 [0.19, 1.10]
Total Hip BMC	unadjusted	0.64 [0.30, 1.36]	1.00	0.81 [0.42, 1.56]	0.69 [0.34, 1.40]	**0.32 [0.13, 0.80]**
	adjusted[1]	0.59 [0.25, 1.35]	1.00	0.67 [0.33, 1.37]	0.59 [0.27, 1.28]	**0.26 [0.10, 0.66]**
Total Hip aBMD	unadjusted	0.85 [0.40, 1.81]	1.00	1.02 [0.52, 1.98]	0.64 [0.30, 1.36]	0.59 [0.27, 1.31]
	adjusted[1]	0.80 [0.36, 1.76]	1.00	0.96 [0.48, 1.93]	0.53 [0.24, 1.17]	0.48 [0.21, 1.10]
Femoral Neck	unadjusted	1.19 [0.53, 2.66]	1.00	1.54 [0.74, 3.17]	1.02 [0.46, 2.27]	0.67 [0.27, 1.63]
BMC	adjusted[1]	1.09 [0.45, 2.60]	1.00	1.42 [0.66, 3.07]	0.87 [0.37, 2.05]	0.53 [0.21, 1.37]
Femoral Neck	unadjusted	1.09 [0.56, 2.14]	1.00	1.13 [0.61, 2.10]	0.83 [0.42, 1.63]	**0.37 [0.16, 0.88]**
aBMD	adjusted[1]	1.07 [0.52, 2.20]	1.00	1.02 [0.53, 1.98]	0.73 [0.35, 1.49]	**0.28 [0.12, 0.70]**
1/3 Radius BMC	unadjusted	0.51 [0.24, 1.05]	1.00	0.52 [0.27, 1.00]	0.81 [0.44, 1.50]	0.69 [0.36, 1.33]
	adjusted[1]	0.54 [0.25, 1.19]	1.00	**0.48 [0.24, 0.97]**	0.72 [0.37, 1.41]	0.61 [0.30, 1.24]
1/3 Radius aBMD	unadjusted	0.83 [0.44, 1.56]	1.00	**0.47 [0.24, 0.90]**	0.88 [0.48, 1.59]	1.45 [0.83, 2.53]
	adjusted[1]	0.80 [0.41, 1.55]	1.00	**0.48 [0.24,**	0.88 [0.47, 1.64]	1.41 [0.78, 2.53]

Adjusted for height Z-score, BMI Z-score, physical activity, calcium intake, Tanner stage and bone age

Bolded values are those where the 95% CI excludes 1.0.

Table 4

Circulating phantom aBMD and BMC results by clinical center.

	Clinical Center	Center 1	Center 2	Center 3	Center 4	Center 5	Difference (highest to lowest)	Difference as percent of overall mean	Coefficient of Variation [SD/ mean]
European Spine	BMC [g]	30.94	31.10	30.93	**32.09**	**30.62**	1.47	4.7%	1.8%
	aBMD [g/cm^2]	1.027	**1.026**	1.027	**1.069**	1.032	0.04	4.1%	1.8%
Whole body	BMC [g]	**739.0**	737.5	**639.6**	729.4	712.1	99.4 27 *	14.0% 3.7% *	5.8% 1.7% *
	aBMD [g/cm^2]	1.180	**1.192**	1.146	**1.128**	**1.128**	0.06	5.5%	2.6%
Femur	BMC [g]	38.31	**38.75**	**38.03**	38.38	38.56	0.72	1.9%	0.7%
	aBMD [g/cm^2]	**0.787**	0.793	0.791	**0.806**	0.798	0.02	2.4%	0.9%

Each center scanned the same set of 3 phantoms [European Spine, Femur, Whole Body] 10 times. Results are the mean value.

Lowest and highest values are in bold

* the variability excluding Center 3 which used a 4500w device. Whole body aBMD was not affected by the 4500w since it did not affect the range