This is a repository copy of Amalgamated Reference Data for Size-Adjusted Bone Densitometry Measurements in 3598 Children and Young Adults-the ALPHABET Study.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/134697/

Version: Accepted Version

Article:

https://doi.org/10.1002/jbmr.2935

© 2016 American Society for Bone and Mineral Research. This is the peer reviewed version of the following article: Crabtree, NJ, Shaw, NJ, Bishop, NJ et al. (10 more authors) (2017) Amalgamated Reference Data for Size-Adjusted Bone Densitometry Measurements in 3598 Children and Young Adults-the ALPHABET Study. Journal of Bone and Mineral Research, 32 (1). pp. 172-180, which has been published in final form at https://doi.org/10.1002/jbmr.2935. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
AMALGAMATED REFERENCE DATA FOR SIZE-ADJUSTED BONE DENSITOMETRY MEASUREMENTS IN 3598 CHILDREN AND YOUNG ADULTS – THE ALPHABET STUDY

Nicola J Crabtree (PhD)1, Nicholas J Shaw (MB ChB)1, Nicholas J Bishop (MB ChB MD)2, Judith E Adams (MB BS) 3, M Zulf Mughal (MB ChB) 4, Paul Arundel (MB BS) 2, Mary S Fewtrell (MD)5, S Faisal Ahmed (MD) 6, Laura A Treadgold (PhD) 7, Wolfgang Högler (PD MD) 1, Natalie A Bebbington (MSc)1, Kate A Ward (PhD)8,9 on behalf of the ALPHABET Study Team *

Affiliations:

1Department of Endocrinology and Diabetes, Birmingham Children’s Hospital, Birmingham, UK

2Academic Unit of Child Health, University of Sheffield, Sheffield UK

3Radiology and Manchester Academic Health Science Centre, Manchester Royal Infirmary, Central Manchester University Hospitals NHS Foundation Trust and University of Manchester, Manchester, UK

4Department of Endocrinology, Central Manchester University Hospitals NHS Foundation Trust and University of Manchester, Manchester, UK

5Childhood Nutrition Research Centre, University College London, Institute of Child Health, London, UK

6Developmental Endocrinology Research Group, School of Medicine, University of Glasgow, Royal Hospital for Sick Children, Glasgow, UK
7 Division of Biomedical Imaging, Leeds Institute of Cardiovascular and Metabolic Medicine, University of Leeds, Leeds, UK

8. Nutrition and Bone Health, MRC Human Nutrition Research, Cambridge, UK

9. MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK

- The ALPHABET study team: R Ashby, J Anderson, D Chapman, I Hodgkinson, M Machin, B Oldroyd, S Shepherd, J Walford, J Williams

Corresponding author:
Dr Nicola Crabtree
Department of Endocrinology
Birmingham Children’s Hospital
Steelhouse Lane
Birmingham
B4 6NH
UK
Tel:  0121 333 8517
Email: Nicola.crabtree@bch.nhs.uk

Grant Supporters:

NJC, NJS, NJB, NB, KW - Arthritis Research UK - Grant Reference 19714
NJB - NIHR; Orthopaedic Research UK
PA, NJB – Sheffield Children's Hospital Charity
KW - MRC as part of the Nutrition and Bone Health research programme (Programme number U105960371)
NJS - National Osteoporosis Society and SPARKS

MZM, JEA - National Osteoporosis Society and Research Endowment in Central Manchester University Hospitals NHS Foundation Trust

WH, MSF, LAT, SFA – None

**Supplemental information included:** 25 pages

**Disclosures:** The authors have nothing to disclose
Abstract

The increasing use of dual energy X-ray absorptiometry (DXA) in children has led to the need for robust reference data for interpretation of scans in daily clinical practice. Such data need to be representative of the population being studied and be ‘future-proofed’ to software and hardware upgrades. The aim was to combine all available paediatric DXA reference data from seven UK centres to create reference curves adjusted for age, sex, ethnicity and body size to enable clinical application, using in-vivo cross calibration and making data back- and forward-compatible.

Seven UK sites collected data on GE-Lunar or Hologic Scanners between 1996 and 2012. Males and females aged 4 to 20 years were recruited (n=3598). The split by ethnic group was: White Caucasian 2887; South Asian 385; Black Afro-Caribbean 286; mixed heritage 40. Scans of the total body and lumbar spine (L1-L4) were obtained. The European Spine Phantom was used to cross-calibrate the 7 centres and 11 scanners. Reference curves were produced for L1-L4 bone mineral apparent density (BMAD) and total body less head (TBLH) and L1-L4 areal bone mineral density (aBMD) for GE Lunar Prodigy and iDXA (sex-and ethnic-specific) and for Hologic (sex-specific). Regression equations for TBLH BMC were produced using stepwise linear regression. Scans of 100 children were randomly selected to test backwards and forwards compatibility of software versions, up to version 15.0 for GE Lunar, and Apex 4.1 for Hologic.

For the first time, sex and ethnic-specific reference curves for lumbar spine BMAD, aBMD and TBLH aBMD are provided for both GE-Lunar and Hologic scanners. These curves will facilitate interpretation of DXA data in children using methods recommended in ISCD guidelines. The databases have been created to allow future updates and analysis when more
definitive evidence for the best method of fracture prediction in children is agreed.

Keywords: DXA; paediatric; BMD; BMC; reference; lean mass
Introduction

The increasing availability and use of dual energy X-ray absorptiometry (DXA) technology in children has brought to the fore the need for robust reference data for all DXA manufacturers. Although manufacturer reference databases are available, they are often not population based nor representative of the individual population being studied (1). Such databases may also have wide variability due to small numbers, with limited power to model rapid skeletal changes during different phases of growth. A further limitation for their use in daily practice is the widespread use of multiple generations of hardware and acquisition and analysis software that may distort the output. There is a need to enable transition between them when monitoring skeletal health in individual patients or undertaking longitudinal research studies.

In 2013 the International Society for Clinical Densitometry (ISCD) updated their 2007 Pediatric Bone Densitometry Guidelines for bone assessment in children (1-3). The committee concluded that DXA is the preferred method for assessment of areal bone mineral content (BMC) and density (aBMD) and that estimating aBMD should be part of the overall assessment for children at elevated risk of a clinically significant fracture (1-3). Measurements of total body less head (TBLH) and/or posterior-anterior lumbar spine aBMD or BMC are recommended; in conjunction with a history of clinically significant fractures, these can be used to indicate the diagnosis of osteoporosis in children and adolescents (1-3). In children with short stature or growth delay, the measurements should be size-corrected using appropriate methods (4-7). The guidelines also acknowledge that adjustment for soft-tissue measurements may be useful in children with malnutrition or in those with muscle and/or skeletal deficits, as has been shown previously (8-11). Despite these guidelines, there are still inconsistencies in the management of children with low BMD and bone fragility around
the world. The lack of robust reference data in a format that permits the diagnostic
application of ISCD recommendations is a source of inconsistency.

The primary aim of the current study was to combine all available paediatric DXA reference
data from seven UK centres to create age-, sex-, ethnic- and size-corrected reference curves
for use in clinical practice and prediction equations for the assessment of the muscle and bone
relationship, and a database which is in-vivo cross calibrated and back- and forward-
compatible.

Methods

Subjects

Three thousand five hundred and ninety eight healthy, community dwelling children aged 4
to 20 years were recruited from 7 UK centres (Birmingham, Leeds, London, Glasgow,
Sheffield, Middlesbrough, Manchester) using centre-specific protocols, from 1996 to
2012 (Supplementary Table 1). Participants were a self-selected convenience sample from
across each study region, recruited through advertisement in local schools and colleges,
general practice surgeries and youth groups. Children of White Caucasian, South-Asian and
Black Afro-Caribbean /African descent were included in the study, depending on centre-
specific inclusion and exclusion criteria. Ethnicity was defined by participants’ self-reporting
both parents being of identical ethnic origin; where this was not the case, data were excluded.
All centres recruited healthy children without known metabolic bone disease, confirmed
through centre-specific screening questionnaires (Supplementary Table 1); abnormal results
were followed-up and excluded if metabolic bone disease was suspected. Children were
included who had had one or more moderate or high trauma fractures (12). At all centres,
local research Ethics Committees approved the studies. All research was carried out in
accordance with the Declaration of Helsinki.
Anthropometric measurements

Height and weight were measured according to centre specific protocols and body mass index (BMI) calculated as weight/height² (kg/m²). To describe the population at each centre, height, weight and BMI measurements were transformed to standard deviation scores (Z-Scores) using the 1990 British growth reference data (13-15).

Scan acquisition

Children were scanned at each centre on either a GE Lunar™ DPX-L, Prodigy or iDXA scanner (GE Medical Systems, Madison, Wisconsin, US) in Birmingham, Leeds, London, Glasgow, Sheffield, Middlesbrough or on a QDR Discovery Hologic™ scanner (Hologic, Bedford, MA, US) in Manchester. Total body, lumbar spine and proximal femur scans were obtained; since the femur is not currently a recommended site according to the current ISCD guidelines (2) only total body and lumbar spine are reported. Standard operating procedures were followed in each centre. All scans were analysed centrally in Birmingham by two Clinical Scientists and were scored for quality of scan acquisition and analysis. DPX-L scans were analysed using software version 4.6c, Prodigy and iDXA scans using Encore version 15.0 (Basic and Enhanced) and Hologic scans using Apex 4.1. Spine bone mineral apparent density (BMAD) was calculated using an adapted method of Carter et al. (g/cm³) (4, 16, 17).

Lumbar spine BMAD (g/cm³) = \( \frac{BMC_1 + BMC_2 + BMC_3 + BMC_4}{V_1 + V_2 + V_3 + V_4} \)

Where \( V_n \) is the volume of the \( n^{th} \) individual vertebra = \( AP_n^{1.5} \) (\( AP_n \) = Projected vertebral area of the \( n^{th} \) vertebra)

\( BMC_n \) is the bone mineral content of the \( n^{th} \) vertebrae
Prediction equations were generated for GE Lunar (Prodigy, iDXA) and Hologic (Discovery) for predicted total body less head bone mineral content (TBLH-BMC) by linear regression analysis of log transformed, lean mass, fat mass, height and age (9, 18).

Centre cross-calibration:
The European Spine Phantom (ESP) was used to cross-calibrate bone measurements at 7 centres and 11 scanners. (19, 20). The phantom was measured once at each centre 10 times without repositioning. For practical purposes this process was not repeated and therefore we relied on local monitoring of scanner operation to verify machine stability. Birmingham was used as the reference centre and all sites cross-calibrated to these measurements. Additional measurements were taken on the iDXA and Hologic scanners using the Leeds Paediatric Spine Phantom, developed by The University of Leeds (in-house).

In-vivo cross calibration:
In-vivo cross calibration was performed in Birmingham, firstly for DPX-L to Prodigy in healthy children (n=105) and then for Prodigy to iDXA in children undergoing scans for clinical purposes (n=70). Both studies were approved by South Birmingham Ethics Committees. Cross-calibration equations were produced using linear regression analysis of absolute values. Machine differences were tested using paired t-test and machine bias with Bland and Altman (Supplementary table 2). The equations were used to transform data from the other GE-Lunar centres to Birmingham for lumbar spine DPX-L to Prodigy Basic and iDXA; and for total body DPX-L to Prodigy basic, Prodigy enhanced and iDXA ¹. In-vivo cross-calibration was not performed between Hologic and GE-scanners for bone or soft tissue measurements.

¹ Prodigy Enhanced is an option only available for total body scans.
Back- and forward compatibility

Scans of 100 children were selected from each of the GE Lunar and Hologic databases to create equations for back- and forwards-compatibility of the reference curves. Within each cohort of 100 children, 20 children per age-band (5-7, 8-10, 11-13, 14-16, 17-19 years) were selected at random (10 male, 10 female) from each of the manufacturer specific datasets. Total body and lumbar spine scans were analysed on software versions: GE-Lunar 10, 11, 13, 14, 15; Hologic 12.4, Apex 2.4, 3.1, 4.1. This sub-set of scans remains available for analysis for future software versions.

Statistical analysis

The Lambda-Mu-Sigma (LMS) method was used to produce age reference curves for Lumbar Spine BMAD, L1-L4 aBMD and TBLH BMD. The LMS curves were generated using the method described by Cole and Green (21) (LMSchartmaker Pro version 2.54 © 1997-2011 Medical Research Council, UK). In brief, reference centile curves describe the distribution of the dependent variable as it varies with the independent predictor covariate, here being age. The curves are fitted using the parametric approach of the penalised log likelihood method as cubic splines by non-linear regression. The degree of smoothing required for the curves is expressed in terms of the equivalent degrees of freedom (edf) (21). The resulting model for the dependent variable, generated from the raw data, is summarised by three parameters, namely: L the Box-Cox power transformation needed to remove any skewness from the distribution, M the median, and S the coefficient of variation. The LMS models were fitted using the “Loop” analysis function in the software, setting the maximum edf’s for the cubic
splines at 3, 6 and 3 and the minimum edf’s at 0, 1 and 1, for L, M and S respectively. The reference model choice was guided by the Schwarz Bayesian Criterion and visual inspection of the curves, resulting in a parsimonious model. Goodness of fit was investigated using the detrended Q-Q plots and ensuring the Q-test statistic was less than 2 (22-24). Standardized residuals were tested for normality and the distribution of subjects within the expected centiles was calculated.

Figures 1-3 and Supplemental Figures 3-5 highlight the age-related mean with the 5th and 95th confidence intervals with each sex and ethnic group fitted separately. Standard deviation scores (Z-scores) are calculated from the LMS parameters using the equation:

\[ Z = \left( \frac{y}{M} \right)^L - 1 \right) / L * S \]

Z = Z-score, y = measured value, M = estimated mean, L = skewness, S = distribution

The need for ethnic specific curves was tested using a one-sided t-test of the Z-scores calculated from the gender specific white data. Where, a significant difference from zero was observed, ethnic specific curves were generated. The goodness of fit of the curves is described by comparing expected versus observed Z-score centile distributions in Supplemental Tables 7a-j.

Regression equations for TBLH-BMC were produced using stepwise linear regression; covariates in the initial model were log-transformed total body lean, total body fat, height and age, only significant covariates were used. Residual plots were inspected for normality to check for skewness and bias in the prediction models.

Results

A total of 3598 scans from children and young adults aged 4 to 20 years-old were included in this study (1820 female, 1778 male). The split by ethnic group was: White Caucasian 2887;
South Asian 385; Black African/ Afro Caribbean 286 and 40 mixed heritage. One hundred
and one subjects were excluded (61 extreme body size [either height, weight or BMI SDS < -
3.5 or > 3.5SD]; 40 mixed heritage), leaving a total of 3497 subjects for the generation of
reference data (Table 1). Descriptive data by centre are shown in Table 2. There were small,
significant centre differences in height, weight and BMI SDS. Subjects were generally taller,
heavier with greater BMI than the 1990 UK-reference population (13-15).

Manufacturer differences

Phantom cross calibration: Using the ESP and with Birmingham as the reference centre there
were no significant differences between all 11 scanners in phantom BMC and aBMD
(including Hologic). In contrast, BA was more variable between the centres but the only
significant difference was observed between the Hologic scanner and all GE scanners
(p=0.010) (Supplemental Figure 1).

We explored these differences further using the Leeds Paediatric Spine Phantom scanned on
a Hologic Discovery and GE-Lunar iDXA scanners. There were no significant differences in
aBMD however BMC and BA were significantly different between the two (p<0.001), with
Hologic giving increasingly higher values compared to the iDXA with increasing BMC and
BA. Therefore, transformation equations were produced. However, when we applied these to
the in-vivo data there were still systematic differences between the Hologic and GE-Lunar
datasets. Consequently, we could not combine different manufacturer scan data and thus
needed to generate brand-specific reference data for use in clinical practice.

In-vivo cross-calibration: In-vivo cross-calibration data were only available for the GE-Lunar
scanners (25, 26). The strong linear relationships between scanners from a single
manufacturer enabled successful transformation of the in-vivo reference datasets collected
from three generations of GE-Lunar scanners. Once successfully transformed, the Bland
Altman tests showed no residual bias. Consequently, this allowed the pooling of all the GE-Lunar data.

Software differences – backwards and forwards compatibility

For GE Lunar, there were no differences in any parameter measured using the basic analysis from version 10 onwards (Prodigy). Version 14.0 included an enhanced total body analysis to try and make Prodigy total body results comparable with the newly introduced iDXA. Whilst there were no differences between the basic analysis, it is not surprising that there were differences between the basic and enhanced total body analyses for all measured parameters (aBMD, BMC, BA, lean and fat) (Supplemental Figure 2).

For Hologic there were no differences between software versions 12.4 through Apex 4.1. It is important to note that this is only true if the same analysis option is used; for this study NHANES BCA was selected throughout.

Reference curve generation (Figures 1-3, Supplementary data S3-5)
Because of the known differences in development between boys and girls their data were separately analysed for BMAD, aBMD and TBLH-BMC.

Size-adjusted lumbar spine (Supplemental tables 4a-c)
Small, but significant differences were found for BMAD between White and Asian, and White and Black children, (Figure 1). In girls, the mean difference in Z-score, calculated using White as the referent group, was 0.25 (0.88), p<0.0001 and 0.62 (1.18) p<0.0001 for South Asian and Black Caribbean girls respectively (Supplemental Table 7a-b). In boys, the mean difference in Z-score, again calculated using White as referent group, was 0.24 (0.96),
p=0.001 and 0.46 (0.98) p<0.0001 for South Asian and Black Caribbean’s respectively (Supplemental Table 7a-b). When Z-scores were recalculated using ethnic-specific LMS data they were no longer significantly different from 0. LMS data were therefore generated for each ethnic group separately.

Figure 3 shows inter-scanner curve comparisons for males and females separately. Despite cross-calibrating the Hologic BMC and BA values to GE Lunar using the ESP, highly significant differences between the scanners remained confirming the differences described earlier. The result of these differences was that calculated BMAD was lower from the Hologic scanner. We explored whether this was due aBMD, BMC or BA. BMC and aBMD were not different but BA was greater in Hologic. Using log-log transformation, (27) the relationship between BA and BMC differed between scanners: for Prodigy, iDXA and DPX-L this was $BA^{1.7}$ (expected $BA^{1.5}$ (4)), whereas for the QDR Discovery it was $BA^{1.9}$.

Lumbar spine and total body less head areal BMD (Supplemental Tables S5-6) In contrast to the BMAD findings there were no significant differences in South Asian children when compared to the white group. Differences remained for black compared to white girls (lumbar spine 0.69 (1.14) p<0.001; TBLH 1.04 (1.08), p<0.0001) and boys (lumbar spine 0.56 (0.97) p<0.0001; TBLH 0.93 (1.06), p<0.0001) (Supplemental Tables S7d,e, 7e, h). We therefore combined the data for White and South Asian children, and re-checked the distribution of Z-scores to check for normality and to ensure differences were not significantly different from 0, they were not confirming the appropriateness of combining data.

Total body less head BMC (Tables 3-6)
ANOVA was performed with TBLH-BMC as the dependent variable and lean body mass, fat body mass, height, age, gender and ethnicity as co-variates or factors in the model.

Significant effects were noted for all covariates and factors. Total body lean mass was the greatest predictor of TBLH-BMC, closely followed by total body fat mass, age and height.

Significant interactions were noted for all covariates between genders and ethnic groups (p<0.001). Girls had greater TBLH-BMC than males for the same lean mass, fat mass, height and age. For the same gender, Afro-Caribbean children had greater TBLH-BMC for the same covariate values (data not shown). Consequently, using stepwise linear regression analysis with parsimonious variable selection of the log-transformed parameters, individual predictor models were generated for each manufacturer, each ethnic group and each gender (Table 3a-d). Individual Z-scores can be produced from by inputting age, height, lean and fat mass into the prediction equation. The predicted value can then be used to calculate the Z-score by using the following equation:

\[ Z - score = \frac{Measured \ value - Predicted \ value}{Predicted \ value \times SEE} \]

**Discussion**

For the first time, DXA measurements in children and young adults aged 4-20 years combining data collected across multiple generations of GE-Lunar and Hologic DXA scanners and software have been collated. Reference data are presented using some of the recently recommended methods by ISCD for clinical use. We provide reference curves for age- and size-adjusted lumbar spine and total body bone densitometry up to the age of 20 years. We also give prediction equations for size- and body composition-adjusted TBLH-BMC measurements. These data enable calculation of sex-specific Z-scores for three ethnic
groups from 4 years-of-age through to the children switching to adult transition services.

Looking ahead, our random dataset of 100 healthy children provides forwards compatibility of software, which allows us testing of future software updates.

Scanner differences

The strong linear relationships between the in-vivo cross-calibration of the reference datasets enabled pooling of all of the GE-Lunar scanners after applying machine specific (i.e. Prodigy, i-DXA) in-vivo transformation equations (Supplementary Table 2a-b). Unfortunately, only data from in-vitro phantoms were available for cross-calibration between the two scanner manufacturers. The observed BA differences were due to varying projectional errors of the fan-beam (Hologic) versus narrow-fan (GE-Lunar) technology. Since the phantom consists of an anthropomorphic spine set in a fixed position it cannot account for differences in body thickness or spine depth which introduces significant errors in measurement when scanning in-vivo. For this reason we were unable to cross-calibrate Hologic to GE-Lunar data. Our findings confirm the inappropriate nature of using phantoms to cross-calibrate between hardware with different properties, i.e. pencil → narrow-fan → fan beam (28,29).

Software differences

The data presented here are for the latest software version of each manufacturer; Encore 15.0 (GE Lunar) and Apex 4.1 (Hologic). With simple transformations it is possible to interpret the DXA results using any version of software going back to GE Lunar Encore 10.0 and Hologic 12.4. Our findings confirm that for both manufacturers it is necessary to always use software specific reference data. It should be noted that for both, it is essential to ensure that when comparing results from different software versions the same analysis options are selected. For GE-Lunar this means selecting enhanced or basic analysis, and for Hologic
Apex software the NHANES BCA analysis should be switched on (30). For older, pre-Apex
versions of Hologic, the ‘auto whole body analysis’ should be used.

Reference data and their use in fracture prediction
Our study presents age- (TBLH-aBMD, spine aBMD) and size-adjusted data for bone
densitometric variables (BMAD, TBLH-BMC) previously shown to best predict fractures in
healthy or chronically ill children (31); these also represent some of the methods currently
recommended by ISCD (1, 2). In over 450 children with chronic disease the diagnostic odds
ratio for predicting vertebral fractures was 9.3 (5.3-14.9) for lumbar spine BMAD; for
predicting long bone fractures the odds ratio was 6.5 (4.1-10.2) for TBLH-BMC for lean
mass (31). BMAD has also been shown to be the best size-adjustment method for prediction
of fractures in healthy children (32). Current understanding is that when interpreting
paediatric bone density results it is preferable to use a size-adjustment method, such as
BMAD or a height-adjusted Z-score(1), however a firm consensus regarding the most
appropriate size-adjustment technique has yet to be established and for this reason the use of
age-adjusted aBMD is still recommended by ISCD (2). Unlike previous studies, some of
which are described below, that present reference data from a single manufacturer and using
one software version (7, 16, 33, 34) the data presented here can easily be applied to different
software versions and manufacturers. If necessary, data can be regenerated using newer size-
adjustment methodology.

The Bone Mineral Density Childhood Study (BMDCS) multi-center study generated robust
US-population-derived reference data for Hologic scanners (software version 12.3 for
baseline and Apex 2.1 for follow-up scans) from over 10 000 measurements in over 2000
individuals of TBLH and lumbar spine BMC and aBMD measurements in 5 to 20-year olds
(6, 6). Size-adjusted prediction equations using height for age Z-scores were also generated
and verified using an independent dataset. No data have yet been published to show whether this method of adjustment significantly improves fracture prediction. Reference data were also generated from the NHANES study; to date only LMS data for total body composition have been published (33). It should be noted that both the NHANES and the BMDCS studies generate Hologic reference data and are from much larger population samples than the UK database presented here.

In contrast to the current study, NHANES data have been cross-calibrated from Hologic to GE-Lunar. Data generated on Hologic 4500 scanners (software version Apex 3.0) were cross calibrated to GE Lunar iDXA values (Software version 14.0) (29, 34). However, despite being the largest published database (approximately 20 000 measurements), only data for total body measurements were presented. Since reductions in TBLH-BMC only predict long bone and not vertebral fracture risk (31), isolated total body data may have limited clinical use. Another possible limitation of the NHANES reference database translation to GE measurements is that pragmatic cross-calibration was performed using data from a native Chinese population and then applied to transform a much larger dataset of a North American US population (34).

Limitations

There are several limitations to this study. The previously discussed differences in phantom measurements between the scanners due to projection error and table height differences (Figure 3) and subsequent lack of in-vivo data for cross-calibration meant that we were unable to create a single combined dataset, applicable to both manufacturers’ scanners. The data were all collected in UK centres, but are applicable for use worldwide provided the same software and scan protocols are used. Caution should be applied when using the data in
populations in which there may be differences in growth rates or body habitus and robust
testing should be employed. In our study the sample size for the South Asian and Afro-
Carribean populations were considerably smaller than the White population and recruited
mostly from one centre and as such we cannot be certain that this is fully representative of the
population. We cannot rule out recruitment bias in any of the centres but as can be seen from
**Supplementary Table 1** protocols and sampling strategies were broadly the same.

Although we cannot confirm that the differences between GE Lunar and Hologic reference
data were not due to population differences, it is likely that the differences are due to
differences in scanner technology. We believe the cross-calibration procedure is as robust as
it can be, since collecting repeated measurements on scanners across the country is neither
ethical nor feasible. Because only one centre collected Hologic data, in one ethnic group,
there are fewer subjects and the Hologic dataset did not include different ethnic groups.
Despite this, we have made this Hologic dataset robust to software updates and increased the
utility of the data previously published in 2007 (16). Finally, we have focussed on testing
the data based on bone measurements only, clearly repeating this work for body composition
would be an advantage (29, 34).

**Conclusion**

In conclusion, we present backwards- and forward-compatible ethnic- and sex specific
reference data for size-adjusted bone density in children and young adults, generated from
measurements in over 3500 individuals using GE and Hologic scanners. These data have
been produced using methods included in the most recent ISCD guidelines and for the first
time present curves for lumbar spine BMAD and prediction equations for TBLH-BMC taking
into account lean mass and body size, together with age-and gender-specific curves for
lumbar spine and TBLH aBMD. This reference database data has been specifically designed to allow future updates and analysis when more definitive evidence for the best method of fracture prediction in children is agreed.
Acknowledgements

The ALPHABET study team includes, R Ashby, J Anderson, D Chapman, I Hodgkinson, M Machin, B Oldroyd, S Shepherd, J Walford and J Williams.

The amalgamation of the reference data was supported by an Arthritis Research UK Clinical Studies Pilot/Feasibility Grant (Grant Reference 19714). We thank the Arthritis Research UK Paediatric Clinical Steering group for supporting the study. NJC was also funded by a National Institutes of Health Research Clinical Development Fellowship (HCS/P10/009).

KW was funded by MRC as part of the Nutrition and Bone Health research programme (Programme number U105960371). The initial collection of the Birmingham GE-Lunar database was funded through a project grant from the National Osteoporosis Society and SPARKS. The initial collection of the Hologic reference database in Manchester was funded through a project grant and a Linda Edwards Memorial PhD Fellowship from the National Osteoporosis Society and subsequently from Research Endowment in Central Manchester University Hospitals NHS Foundation Trust. We would also like to thank Dr Will Johnson, MRC Human Nutrition Research, Cambridge and Professor Tim Cole, University College London, for their advice in preparing the revised version of this manuscript.

This paper presents independent research funded in-pat by the National Institute for Health Research (NIHR). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Authors’ roles: Current study design: NJC, KW, NJS, NJB; Study conduct: NJC Data collection: All authors and the ALPHABET study team. Data analysis: NJC Data interpretation: NJC, KW Drafting manuscript: NJC, KW. Revising manuscript content: All authors. Approving final version of manuscript: All authors; Data integrity of manuscript: NJC
References:


Figure legends

**Figure 1** Comparison of GE Lunar iDXA™ lumbar spine BMAD LMS reference curves between the three different ethnic groups. (A) BMAD (g/cm$^3$) for girls; (B) BMAD (g/cm$^3$) for boys. Solid black line represents the mean for White Caucasian Children (± 95% Confidence interval -dotted black line). Dark grey dashed line represents the mean for Black Afro-Caribbean Children; Dashed light grey line represents the mean for South Asian Children.

**Figure 2** Comparison of lumbar spine BMAD LMS reference curves between males and females (A) GE Lunar iDXA; (B) Hologic Discovery. Solid black line represents males (mean ± 95% Confidence interval). Dashed line represents females (mean ± 95% Confidence interval).

**Figure 3** Comparison of lumbar spine BMAD LMS reference curves between manufacturers, GE Lunar iDXA™ compared to Transformed Hologic Discovery (Hologic data transformed using cross calibration equations generated from the European Spine Phantom). (A) Females; (B) Males. Solid black line represents GE Lunar iDXA™ (mean ± 95% Confidence interval). Dashed line represents Hologic Discovery (mean ± 95% Confidence interval).
Table 1 Distribution of subjects used for the generation of reference data

<table>
<thead>
<tr>
<th></th>
<th>GE Lunar Prodigy</th>
<th></th>
<th>GE Lunar iDXA (including transformed Prodigy)</th>
<th></th>
<th>Hologic Discovery</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Male</td>
<td>1245</td>
<td></td>
<td>Male</td>
<td>1411</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>1302</td>
<td></td>
<td>Female</td>
<td>1499</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>White Caucasian</td>
<td>925</td>
<td></td>
<td>1091</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>South Asian</td>
<td>192</td>
<td></td>
<td>192</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Black Afro Caribbean</td>
<td>128</td>
<td></td>
<td>128</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>South Asian</td>
<td>184</td>
<td></td>
<td>184</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Black Afro Caribbean</td>
<td>148</td>
<td></td>
<td>148</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>White Caucasian</td>
<td>325</td>
<td></td>
<td>325</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>South Asian</td>
<td>262</td>
<td></td>
<td>262</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Black Afro Caribbean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2 Patient anthropometric data. Mean (SD)

<table>
<thead>
<tr>
<th>Centre</th>
<th>Number</th>
<th>Mean (SD) Height Z-score</th>
<th>Mean (SD) Weight Z-score</th>
<th>Mean (SD) BMI Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birmingham</td>
<td>935</td>
<td>0.20 (1.09)</td>
<td>0.45 (1.24)</td>
<td>0.46 (1.25)</td>
</tr>
<tr>
<td>Middlesbrough</td>
<td>390</td>
<td>0.35 (0.97)</td>
<td>0.41 (0.96)</td>
<td>0.31 (1.00)</td>
</tr>
<tr>
<td>Leeds</td>
<td>171</td>
<td>0.34 (1.00)</td>
<td>0.42 (1.00)</td>
<td>0.31 (1.11)</td>
</tr>
<tr>
<td>Glasgow</td>
<td>212</td>
<td>0.15 (1.02)</td>
<td>0.34 (1.07)</td>
<td>0.36 (1.02)</td>
</tr>
<tr>
<td>London</td>
<td>372</td>
<td>0.11 (1.03)</td>
<td>0.29 (1.10)</td>
<td>0.27 (1.12)</td>
</tr>
<tr>
<td>Sheffield</td>
<td>830</td>
<td>0.40 (1.05)</td>
<td>0.59 (1.11)</td>
<td>0.51 (1.15)</td>
</tr>
<tr>
<td>Manchester</td>
<td>587</td>
<td>0.30 (0.96)</td>
<td>0.47 (1.01)</td>
<td>0.41 (1.03)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>3497</td>
<td>0.28 (1.03)</td>
<td>0.46 (1.11)</td>
<td>0.42 (1.14)</td>
</tr>
<tr>
<td>Centre Differences (p value)</td>
<td>&lt;0.001</td>
<td>0.001</td>
<td>0.003</td>
<td></td>
</tr>
</tbody>
</table>

Using a one-sided t-test all Z-scores were significantly (p<0.0001) greater than zero. Centre differences were compared using ANOVA.
Table 3a Prediction Equations for Total body less head bone mineral content (TBLH-BMC (g)) for lean mass (g), fat mass (g), height (cm) and age (1 decimal place) for the GE Lunar Prodigy™ Software version Encore 15.0.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Ethnicity</th>
<th>GE Prodigy</th>
<th>$r^2$</th>
<th>SEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls</td>
<td>White Caucasian</td>
<td>TBLH-BMC = 3.77x10^{-4} x LEAN^{0.845} x FAT^{0.130} x Height^{0.928} x Age^{0.179}</td>
<td>0.966</td>
<td>0.0988</td>
</tr>
<tr>
<td></td>
<td>South Asian</td>
<td>TBLH-BMC = 2.24x10^{-4} x LEAN^{0.603} x FAT^{0.122} x Height^{1.535} x Age^{0.216}</td>
<td>0.970</td>
<td>0.0935</td>
</tr>
<tr>
<td></td>
<td>Black Afro-Caribbean</td>
<td>TBLH-BMC = 1.02x10^{-3} x LEAN^{0.941} x FAT^{0.100} x Height^{0.543} x Age^{0.311}</td>
<td>0.967</td>
<td>0.1002</td>
</tr>
<tr>
<td>Boys</td>
<td>White Caucasian</td>
<td>TBLH-BMC = 2.93x10^{-4} x LEAN^{0.939} x FAT^{0.073} x Height^{0.930} x Age^{0.079}</td>
<td>0.972</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>South Asian</td>
<td>TBLH-BMC = 1.47x10^{-4} x LEAN^{0.978} x FAT^{0.060} x Height^{1.060}</td>
<td>0.978</td>
<td>0.0932</td>
</tr>
<tr>
<td></td>
<td>Black Afro-Caribbean</td>
<td>TBLH-BMC = 1.94x10^{-3} x LEAN^{0.983} x FAT^{0.048} x Height^{1.018}</td>
<td>0.973</td>
<td>0.0883</td>
</tr>
</tbody>
</table>

Total body less head BMC = TBLH-BMC; Total body lean mass = LEAN; Total body fat mass = FAT; SEE = Standard error estimate

Z-Score = (Measure Value – Predicted Value) / (Predicted Value x SEE)
Table 3b Prediction Equations for Total body less head bone mineral content (TBLH-BMC (g)) for lean mass (g), fat mass (g), height (cm) and age (1dp) for the GE Lunar Prodigy™ using the ENHANCED analysis mode - Software version Encore 15.0.

<table>
<thead>
<tr>
<th></th>
<th>GE Prodigy-Enhanced</th>
<th>$r^2$</th>
<th>SEE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Girls</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White Caucasian</td>
<td>TBLH-BMC = $4.24 \times 10^{-3} \times \text{LEAN}^{0.682} \times \text{FAT}^{0.079} \times \text{Height}^{0.905} \times \text{Age}^{0.122}$</td>
<td>0.967</td>
<td>0.0818</td>
</tr>
<tr>
<td>South Asian</td>
<td>TBLH-BMC = $6.04 \times 10^{-3} \times \text{LEAN}^{0.511} \times \text{FAT}^{0.106} \times \text{Height}^{1.110} \times \text{Age}^{0.185}$</td>
<td>0.937</td>
<td>0.0809</td>
</tr>
<tr>
<td>Black Afro-Caribbean</td>
<td>TBLH-BMC = $9.01 \times 10^{-3} \times \text{LEAN}^{0.744} \times \text{FAT}^{0.103} \times \text{Height}^{0.545} \times \text{Age}^{0.234}$</td>
<td>0.961</td>
<td>0.0910</td>
</tr>
<tr>
<td><strong>Boys</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White Caucasian</td>
<td>TBLH-BMC = $1.47 \times 10^{-3} \times \text{LEAN}^{0.813} \times \text{FAT}^{0.055} \times \text{Height}^{0.949}$</td>
<td>0.974</td>
<td>0.0839</td>
</tr>
<tr>
<td>South Asian</td>
<td>TBLH-BMC = $5.06 \times 10^{-3} \times \text{LEAN}^{0.883} \times \text{FAT}^{0.044} \times \text{Height}^{0.586}$</td>
<td>0.979</td>
<td>0.0775</td>
</tr>
<tr>
<td>Black Afro-Caribbean</td>
<td>TBLH-BMC = $3.81 \times 10^{-3} \times \text{LEAN}^{0.856} \times \text{FAT}^{0.047} \times \text{Height}^{0.692}$</td>
<td>0.974</td>
<td>0.0735</td>
</tr>
</tbody>
</table>

Total body less head BMC = TBLH-BMC; Total body lean mass = LEAN; Total body fat mass = FAT; SEE = Standard error estimate

Z-Score = (Measure Value – Predicted Value) / (Predicted Value x SEE)
Table 3c Prediction Equations for Total body less head bone mineral content (TBLH-BMC (g)) for lean mass (g), fat mass (g), height (cm) and age (1dp) for the GE Lunar iDXA™ - Software version Encore 15.0.

<table>
<thead>
<tr>
<th>Girls</th>
<th>GE Lunar iDXA</th>
<th>$r^2$</th>
<th>SEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Caucasian</td>
<td>TBLH-BMC = 1.85 x 10^{-3} x LEAN^{0.736} x FAT^{0.077} x Height^{0.950} x Age^{0.135}</td>
<td>0.965</td>
<td>0.0843</td>
</tr>
<tr>
<td>South Asian</td>
<td>TBLH-BMC = 2.58 x 10^{-3} x LEAN^{0.538} x FAT^{0.110} x Height^{1.210} x Age^{0.192}</td>
<td>0.967</td>
<td>0.0836</td>
</tr>
<tr>
<td>Black Afro-Caribbean</td>
<td>TBLH-BMC = 4.27 x 10^{-3} x LEAN^{0.787} x FAT^{0.105} x Height^{0.594} x Age^{0.239}</td>
<td>0.962</td>
<td>0.0931</td>
</tr>
<tr>
<td>Boys</td>
<td>GE Lunar iDXA</td>
<td>$r^2$</td>
<td>SEE</td>
</tr>
<tr>
<td>White Caucasian</td>
<td>TBLH-BMC = 5.88 x 10^{-4} x LEAN^{0.827} x FAT^{0.055} x Height^{1.095}</td>
<td>0.974</td>
<td>0.0849</td>
</tr>
<tr>
<td>South Asian</td>
<td>TBLH-BMC = 2.01 x 10^{-3} x LEAN^{0.906} x FAT^{0.047} x Height^{0.708}</td>
<td>0.980</td>
<td>0.0798</td>
</tr>
<tr>
<td>Black Afro-Caribbean</td>
<td>TBLH-BMC = 1.78 x 10^{-3} x LEAN^{0.887} x FAT^{0.051} x Height^{0.765}</td>
<td>0.975</td>
<td>0.0754</td>
</tr>
</tbody>
</table>

Total body less head BMC = TBLH-BMC; Total body lean mass = LEAN; Total body fat mass = FAT; SEE = Standard error estimate

Z-Score = (Measure Value – Predicted Value) / (Predicted Value x SEE)
Table 3d Prediction Equations for Total body less head bone mineral content (TBLH-BMC (g)) for lean mass (g), fat mass (g), height (cm) and age (1dp) for the Hologic Discovery – Software version Apex 4.1.

<table>
<thead>
<tr>
<th></th>
<th>Hologic Discovery</th>
<th>²</th>
<th>SEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls White Caucasian</td>
<td>TBLH-BMC = 1.20 x10⁻² x LEAN(^{0.704}) x Height(^{0.717}) x Age(^{0.235})</td>
<td>0.954</td>
<td>0.0871</td>
</tr>
<tr>
<td>Boys White Caucasian</td>
<td>TBLH-BMC = 4.77 x10⁻³ x LEAN(^{1.343}) x FAT(^{-0.046}) x Height(^{0.398})</td>
<td>0.960</td>
<td>0.0962</td>
</tr>
</tbody>
</table>

Total body less head BMC = TBLH-BMC; Total body lean mass = LEAN; Total body fat mass = FAT; SEE = Standard error estimate

Z-Score = (Measure Value – Predicted Value) / (Predicted Value x SEE)