**External validation of a** **childhood fat mass prediction model: individual participant data meta-analysis of predictive performance in 19 countries**

Hudda MT (Senior Research Fellow)1\*, Wells JCK (Professor)2, Adair LS (Professor)3, Alvero-Cruz JRA (Full Professor)4, Ashby-Thompson MN (Postdoctoral Research Scientist)5, Ballesteros-Vásquez MN (Scientific Researcher)6, Barrera-Exposito J (Senior Research Fellow)7, Caballero B (Emeritus Professor)8, Carnero EA (Research Scientist)9, Cleghorn GJ (Emeritus Professor)10, Davies PSW (Honorary Professor)10, Desmond M (Honorary Research Fellow)2, Devakumar D (Associate Professor)11, Gallagher D (Professor)12, Guerrero-Alcocer EV (Professor)13, Haschke F (Visiting Professor)14, Horlick M (Retired)15, Ben Jemaa H (Assistant)16, Khan AI (Scientist)17, Mankai A (Assistant Professor)16, Monyeki MA (Professor)18, Nashandi HL (Lecturer)19, Ortiz-Hernandez L (Professor)20, Plasqui G (Associate Professor)21, Reichert FF (Professor)22, Robles-Sardin AE (Associate Researcher)6, Rush E (Emeritus Professor)23, Shypailo RJ (Instructor)24, Sobiecki JG (Doctoral student)25, ten Hoor GA (Assistant Professor)26, Valdés J (Professor)27, Wickramasinghe VP (Professor)28, Wong WW (Professor)24, Riley RD (Professor)29, Owen CG (Professor)1, Whincup PH (Professor)1, Nightingale CM (Lecturer)1

*\*Corresponding author*

1 Population Health Research Institute, St George’s University of London, Cranmer Terrace, London, England, SW17 0RE

2 Population, Policy and Practice Programme, UCL Great Ormond Street Institute of Child Health, London, UK

3 Department of Nutrition, University of North Carolina Schools of Public Health and Medicine, North Carolina, United States

4 Faculty of Medicine, University of Malaga, Málaga, Spain. Former affiliation: Biodynamic and Body Composition Laboratory, Faculty of Education Sciences, University of Málaga, Málaga, Spain

5 Dept. of Pediatrics, New York Nutrition Obesity Research Center, Columbia University Medical Center,

1150 St. Nicholas Ave, New York, NY 10032, United States

6 Coordinación de Nutrición, Centro de Investigación en Alimentación y Desarrollo, A.C. Carretera Gustavo Enrique Astiazarán Rosas, no. 46, Col. La Victoria, Hermosillo, Sonora MÉXICO, C.P. 83304

7 Biodynamic and Body Composition Laboratory, Faculty of Education Sciences, University of Málaga, Málaga, Spain

8 Department of International Health, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe St, Room 2041, Baltimore, MD 21205, United States

9 Translational Research Institute, Advent health Orlando, 301 E. Princeton St., Orlando, FL 32804, USA

10 Child Health Research Centre, University of Queensland, Brisbane, Australia

11 UCL Institute for Global Health, 30 Guilford Street, London, WC1N 1EH

12 Dept. Medicine & Institute Human Nutrition, Division of Endocrinology, New York Nutrition Obesity Research Center, Columbia University Medical Center, 1150 St. Nicholas Ave, New York, NY 10032, United States

13 Centro Universitario UAEM Amecameca. Universidad Autónoma del Estado de México. Carretera Amecameca-Ayapango km 2.5. CP. 56900. Col Centro, Amecameca de Juárez, Estado de México

14 Department of Pediatrics, PMU Salzburg, Muellner Hauptstrasse 48, 5020 Salzburg, Austria

15 Body Composition Unit, St. Luke's-Roosevelt Hospital, New York, NY, USA

16 Nutrition Department, Higher School of Health Sciences and Techniques, University of Tunis El Manar, Tunis, Tunisia

17 International Centre for Diarrheal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh, 68 Shaheed Tajuddin Ahmed Sarani, Mohakhali, Dhaka 1212, Bangladesh

18 Physical Activity, Sport and Recreation Research Focus Area (PhASRec), Faculty of Health Sciences, North-West University , Potchefstroom, 2520, South Africa

19 School of Nursing and Public Health, Faculty of Health Sciences and Veterinary Medicine, University of Namibia, Windhoek, Namibia

20 Departamento de Atención a la Salud, Universidad Autónoma Metropolitana Xochimilco, Mexico City, Mexico

21 Maastricht University, Dept. of Nutrition and Movement Sciences, NUTRIM. P.O. Box 616, 6200 MD, Maastricht, The Netherlands

22 Postgraduate Program in Physical Education, Federal University of Pelotas, Pelotas, Brazil

23 Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland 0640, New Zealand

24 Baylor College of Medicine, USDA/ARS Children's Nutrition Research Center, 1100 Bates Avenue, Houston, TX , USA

25 MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge, CB2 0QQ, UK

26 Dept. Work and Social Psychology, Maastricht University, P.O. Box 616, 6200 MD Maastricht, the Netherlands

27 Departamento de Bioquímica, Centro de Investigación y de Estudios Avanzados del I.P.N., Mexico City, Mexico

28 Faculty of Medicine, University of Colombo, Sri Lanka

29 Centre for Prognosis Research, School of Medicine, Keele University, Staffordshire, UK

Correspondence to:

Dr Mohammed Hudda

Population Health Research Institute,

St George’s, University of London,

Cranmer Terrace,

London SW17 0RE

Tel: (+44) 208 725 5553

Email: mhudda@sgul.ac.uk

ORCID ID: 0000-0001-7894-1159

Twitter: @mohammedhudda

Word Count: 339 (abstract), 4147 (main body)

**External validation of a childhood fat mass prediction model: individual participant data meta-analysis of predictive performance in 19 countries**

**Abstract**

**Objective**

To evaluate the performance of a recent UK-based prediction model for estimating childhood fat-free mass (FFM) (and indirectly fat mass [FM])

**Design and Setting**

Individual participant data meta-analysis of cross-sectional data from 19 countries

**Participants**

5,693 children (49.7% male) aged between 4 and 15 years with complete data on the predictors included in the UK-based model (weight, height, age, sex and ethnicity) and on the independently assessed outcome measure (FFM determined by deuterium dilution assessment).

**Main Outcome Measures**

The outcome of the UK-based prediction model was natural log-transformed FFM (lnFFM). Predictive performance statistics of R2, calibration slope, calibration-in-the-large and root mean square error (RMSE) were assessed in each of the 19 countries and then pooled via random-effects meta-analysis. Calibration plots were also derived for each country including flexible calibration curves.

**Results**

The model showed good predictive ability in non-UK childhood populations, providing R2 values of >75% in all countries and >90% in 11 of the 19 countries and with good calibration (i.e., agreement) of observed and predicted values. RMSE values (on FFM scale) were <4kg in 17 of the 19 settings. Pooled values of R2, calibration slope and calibration-in-the-large were 88.7% (95%CI: 85.9 to 91.4%), 0.98 (95%CI: 0.97 to 1.00) and 0.01 (95% CI: -0.02 to 0.04), respectively. There was evidence of heterogeneity in the R2 and calibration-in-the-large values across settings but not in the calibration slope. Model performance did not vary markedly between boys and girls, age, ethnic and national income groups. To further improve the accuracy of the predictions, the model equation was re-calibrated in terms of the intercept in each setting so that country-specific equations are available for future use.

**Conclusion**

The UK-based prediction model, which is based on readily available measures, provides childhood FFM predictions, and hence FM, in a range of non-UK settings that explain a large proportion of the variability in observed FFM, and exhibit very good calibration performance, especially after re-calibration of the intercept for each population. The model demonstrates good generalisability both in low/middle-income and high-income populations of healthy children.

**Print Abstract**

**Study Question:** To externally validate the performance of a UK-based prediction model for estimating childhood fat-free mass (FFM) (and indirectly fat mass [FM]) in 19 non-UK childhood settings.

**Methods:** Individual participant datawere obtained on 5693 individuals with complete data on the predictors included in the original UK-based model (weight, height, age, sex and ethnicity) and on the outcome measure (FFM determined by deuterium dilution assessment). Predictive performance statistics of R2, calibration slope, calibration-in-the-large and root mean square error (RMSE) were assessed in each of the 19 countries and then pooled via random-effects meta-analysis. Calibration plots were also derived for each country including flexible calibration curves.

**Study answer and limitations:** The model showed good predictive ability in all non-UK childhood populations, providing R2 values of >75% in all settings, with excellent calibration of observed and predicted values (see figure). RMSE values (on FFM scale) were <4kg in 17 of the 19 settings. Pooled values of R2, calibration slope and calibration-in-the-large were 88.7% (95%CI: 85.9 to 91.4%), 0.98 (95%CI: 0.97 to 1.00) and 0.01 (95% CI: -0.02 to 0.04), respectively. To improve the accuracy of the predictions, the model equation was re-calibrated in terms of the intercept in each setting so that accurate country-specific equations are available for use. Study limitations include the limited global representation, with only a small number of children included from East Asia and none from the Middle East.

**What this study adds:** This external validation study demonstrates strong predictive performance of the UK-based model at estimating FFM (and thus FM) in a range of non-UK childhood settings. The equation, based on readily available markers, performed consistently well in both low/middle- and high-income settings, demonstrating its wider generalisability. The re-calibrated country-specific equations improve the accuracy of predictions and are recommended for future use for effective clinical and public health obesity surveillance, prevention and management.

**Funding:** British Heart Foundation (FS/17/76/33286), the Wellcome Trust (204809/Z/16/Z) and the NIHR ARC South London

**Print Abstract Figure:**

**Introduction**

The high global prevalence of childhood obesity poses a major global public health challenge. Recent data from the World Health Organization estimate that ~18% of all children and adolescents aged 5-19 years worldwide (i.e., over 340 million individuals) are affected by overweight or obesity.1 However, the most widely used marker of high childhood adiposity, body mass index (BMI), has serious limitations.2-4 BMI is poorly correlated with fat mass (FM) in childhood and, crucially, as a weight-for-height index, is unable to discriminate between FM and fat-free mass (FFM), which can both vary markedly in individuals with a given BMI.3 This is of considerable importance as childhood FM has been shown to be more strongly associated with long term type 2 diabetes risk than body fatness markers based on overall weight (of which BMI is a such marker).5 The availability of simple, accurate methods for assessment of FM could represent an important advance in the field of adiposity assessment over the use of BMI.4 6 In vivo techniques for the assessment of FM exist, such as bioelectrical impedance analysis (BIA), dual energy x-ray absorptiometry (DXA), and MRI scanning, but may lack accuracy,4 7 8 and (in the case of DXA and MRI) are inappropriate for general use. An alternative accurate method for FM assessment which has been developed and validated within the UK childhood and adolescent population,6 9 is based on the prediction of FFM (and indirectly FM, as FM = weight - FFM) using simple assessments of height, weight, sex, age and ethnicity. Ethnicity was also included as a predictor within the developed model to allow for established ethnic-differences in childhood body fatness.10-12 In order to maximise the accuracy of the predictions obtained from this approach, the model was developed using childhood FM data obtained from the deuterium dilution method, a reference standard method of adiposity assessment, which provides accurate, safe, and minimally invasive measurements of total body water (and FFM) with an error of less than 1%.13 14 The developed model was also shown to predict FM levels as accurately as DXA and BIA in UK children.9 However, the predictive performance of the model has not so far been examined in childhood populations outside the UK.15 Therefore, in this study, we have conducted an external validation of the UK-based FM prediction model to assess its predictive performance in a wide range of non-UK childhood populations aged 4-15 years with comparable FM assessment from the reference-standard deuterium dilution method.

**Methods**

*Data Sources and study population*

A Medline literature search was carried out through PubMed to identify all published studies which used the deuterium dilution method13 14 to assess FFM and FM in at least 100 healthy children or adolescents aged between 4 and 15 years and also included measurements of weight and height and basic demographic information on age and sex (search dates: May 2020, search terms: “deuterium dilution”, “study”, “children OR adolescents”). Twenty-four such studies were identified; study authors were contacted and invited to contribute data to this investigation. Of these, 14 agreed to participate, 4 had data sharing restrictions and were unable to participate, no response was received from the remaining 6 study investigators. Study collaborators provided data from a further 4 studies (1 with unpublished deuterium dilution data 16 and 2 of which were small17 18). Further information is available in the PRISMA individual participant data (IPD) flow diagram (supplementary file). Each of the 18 studies provided data on weight, height, age, sex, ethnicity, and deuterium dilution assessed total body water (TBW), FM and FFM. The earliest study was conducted in 1981-198219 and the most recent in 2017-2019;20 most studies (83%) were conducted since 2000 (Supplementary Table 1). In total 5,715 generally healthy children aged 4-15 years were included from 18 studies spanning 19 countries: Australia17, Austria19, Bangladesh21, Brazil22, China23, Mexico24 25, Namibia20, Nepal26, Netherlands27, New Zealand10, the Philippines23, Peru18, Poland16, Russia23, South Africa20 23, Spain28, Sri Lanka29, Tunisia30 and USA31 32 (Supplementary Table 1). Whilst data from 5 of these countries (Peru N=56, Spain N=92, China N=95, Philippines N=80, and Australia N=42) had smaller participant numbers than anticipated,17 18 23 28 these were included to avoid data wastage.

*Outcome and predictor assessment*

The outcome was natural log-transformed FFM (lnFFM), ascertained using the deuterium dilution reference method,13 14 which was also used as the outcome for the previously developed UK-based prediction model.6 Five children with missing information on the outcome (lnFFM), one child with an implausible weight value and 16 children with implausible FFM values (i.e., FFM > weight) were excluded. There were no other missing data on the predictors of weight, height, age, sex or ethnicity (classified as either White, Black, South Asian, Other Asian or Other). Following these exclusions, 5,693 children aged 4-15 years of age were included in analyses.

**Statistical Analysis**

All statistical analyses were conducted in Stata (version 17). We followed the TRIPOD (transparent reporting of a multivariable model for individual prognosis or diagnosis) guideline for the reporting of studies validating a multivariable prediction model.15

*Evaluation of overall performance of UK-based prediction model equation*

The UK-based model equation6 (Box 1) was applied to all children within this external validation dataset to obtain a prediction of lnFFM from weight, height, age, sex and ethnicity. The predictive performance of the model equation was assessed within each country by comparing the deuterium dilution assessed lnFFM value with the predicted lnFFM value obtained from the UK-based model. Model performance was assessed by examining established predictive performance measures of:

1. R2 – the percentage of variance in deuterium dilution observed lnFFM explained by predicted lnFFM estimated by the UK-based prediction model
2. Root mean square error (RMSE) – the average difference between predicted lnFFM from the UK-based model and the deuterium dilution observed lnFFM. RMSE indicates the absolute fit of the model to the data (i.e., how close the model’s predicted values of lnFFM are to the deuterium dilution observed values of lnFFM).
3. Model Calibration – assessed collectively by means of three measures; the slope, calibration-in-the-large (CITL) and the calibration plot:
	1. Slope – obtained from the model regressing deuterium dilution observed values of lnFFM on model predicted values of lnFFM. The calibration slope reports on the accuracy of the predictions across the range of lnFFM values by evaluating the spread of the predicted values and has a target value of 1. A slope < 1 suggests that predicted values are too high for children with low observed values and too low for children with high observed values. A slope > 1 suggests the opposite.
	2. CITL – intercept term obtained from a linear regression model of deuterium dilution observed values of lnFFM on model predicted values of lnFFM, where the slope is constrained to be the ideal of 1 (with a CITL of 0 being ideal). CITL measures the overall agreement between average model predicted values of lnFFM and average deuterium dilution observed values of lnFFM (i.e., it tells you about the systematic bias of predicted lnFFM obtained from the model when compared to the observed values of lnFFM).
	3. Calibration plot – graph of deuterium dilution observed lnFFM plotted against model predicted values of lnFFM with a local regression (loess) smoother fitted across all individuals to produce a flexible calibration curve (created using pmcalplot on Stata).

To summarise results across all countries, the country-specific performance measures of R2, calibration slope and CITL were pooled via a random-effects meta-analysis to obtain an estimate of average performance and between-country heterogeneity. The random-effects meta-analysis model was fitted using the restricted maximum likelihood method33 using the Hartung-Knapp34 approach to adjust the standard error of the pooled performance measures. Heterogeneity was summarised for each of the performance measures of R2, calibration slope and CITL using the estimate of between-study variance (tau2) and a 95% prediction interval which provides a range of values one would expect the performance measure to fall within for a new individual setting.35 For the random-effects meta-analysis, the variance and confidence interval for R2 was estimated using the Wald-type method outlined by Tan,36 and 95% prediction intervals were capped at the maximum values of 100%. The model performance was also assessed within sub-groups of sex, three-year age groups (4-6, 7-9, 10-12 and 13-15 years), and ethnic origin. Additionally, the model performance was assessed separately in low/middle-income and high-income childhood populations, categorized using the World Bank Income Classification.37 The country-specific classifications were obtained for the year of study commencement such that Australia, Austria, Netherlands, New Zealand, Spain, USA, and Poland were classified as high-income and the remaining 12 countries were classified as low/middle-income.

While the primary results focus on the model performance measures assessed comparing deuterium dilution observed values of lnFFM to model predicted lnFFM, as the UK-based model was developed to predict FFM on the log scale, RMSE and Calibration plots are also presented on the FFM scale to improve the interpretability of the results, comparing deuterium dilution observed FFM with model predicted FFM (obtained by exponentiating the model predictions of lnFFM).

*Country-specific re-calibration of UK-based prediction model equation*

Where systematic error in the UK-based model’s prediction of lnFFM was observed across countries, the intercept term of the UK-based equation was re-calibrated to provide updated country-specific equations. Within each country, the linear predictor portion of the UK-based model was first estimated for all individuals before fitting a linear regression model with deuterium dilution observed values of lnFFM as the dependent variable and the linear predictor as the independent variable. The slope of this model was constrained to be one and the constant term from the model was used as the updated intercept term for that country. The performance of the country-specific re-calibrated equations were re-examined to assess the predictive performance of the updated model equations.

*Patient involvement and dissemination of research findings*

For this external validation study based upon secondary data analysis, no patients were directly involved in setting the research question, outcome measures, study design or implementation. However, previous focus groups carried out including members of the public, had indicated the need to develop and validate new methods for childhood body fatness assessment, which informed the development of this work. No patients were involved in the interpretation or writing up of results. We plan to disseminate these research findings to relevant stakeholders by presenting our findings at relevant obesity-related conferences, by sharing the findings with The Office for Health Improvement and Disparities (formerly Public Health England) who are responsible for the English National Child Measurement Programme and the WHO Childhood Obesity Surveillance Initiative Steering Group members, and by making plain-language summaries available on social media.

**Results**

A summary of the key characteristics of all 5,693 included participants is presented by country in Table 1. The median age of all participants was 10.8 years, slightly higher than in the development population6 (median age 9.6 years). All studies included both males and females, except for one male only study.19 Average levels of deuterium dilution assessed FM and FFM were, as expected, higher in studies of older children, who were also generally taller and heavier (Table 1). Median FM in this validation dataset (8.5 kg) was very similar to that of the development population6 (8.4 kg), though FFM was higher in this external validation dataset (27.8 kg) than in the development population (24.8 kg).6 The representation of different ethnic groups varied across the settings, with 15 countries including data from a single ethnic group (Table 1).

*Assessment of* *UK-based model performance in non-UK settings*

The UK-based model equation, when applied to the external validation data, produced high R2 values greater than 75% in all countries and values greater than 90% in 11 of the 19 countries (Table 2). When the country-specific R2 values were pooled via random-effects meta-analysis, the overall pooled R2 value was 88.66% (95%CI: 85.91 to 91.41%) (Figure 1). There was evidence of some between-country heterogeneity in the R2 values (tau2=28.62), and the 95% prediction interval for the R2 value (i.e. the expected range of R2 values one would expect to obtain from applying the model to a previously unstudied child population) was 77.04% to 100.00%. RMSE values were generally low (≤0.11 in all countries except for Russia and Sri Lanka) (Table 2). The model demonstrated high levels of calibration in terms of the slope in each of the countries. The observed calibration slope estimates ranged between 0.91 in Spain and 1.05 in Australia and South Africa (Table 2) with the country-specific 95% confidence intervals around the respective slope term containing the ideal value of 1 in 15 of the 19 countries (Table 2). The pooled calibration slope was 0.98 (Figure 1; 95%CI: 0.97 to 1.00) with no evidence of between-country heterogeneity in the calibration slopes (tau2=0), and the 95% prediction interval for the calibration slope you would expect to observe in a new country of 0.92 to 1.05. Although the country-specific CITL values were close to the ideal value of 0 for most countries, ranging from -0.12 in Russia to 0.10 in Bangladesh, the associated 95% confidence intervals failed to contain the ideal value of 0 in any of the settings (Table 2). When pooled via random-effects meta-analysis, the overall pooled CITL value was almost equal to the ideal value of zero (pooled CITL = 0.01) (Figure 1; 95% CI: -0.02 to 0.04). There was some evidence of heterogeneity in the CITL values (tau2=0.0031), with a 95% prediction interval for the CITL which would likely be observedin a new country of -0.11 to 0.13. Graphically, the calibration plots also demonstrated the good levels of calibration across the range of lnFFM values within each of the countries (Figures 2a & 2b), including at the lower- and upper-ends of the distribution of lnFFM, with the flexible calibration curve close to the ideal 45 degree line of perfect calibration. However, the graphs further demonstrated evidence of some systematic error in the prediction of lnFFM across setting (Figures 2a & 2b).

The generally high model performance was maintained when assessed across sub-groups of sex (Supplementary Figure 1), three-year age groups (Supplementary Figures 2), ethnic groups (Supplementary Figure 3) and national income level (Table 3). The model also demonstrated low levels of heterogeneity in the performance statistics across the three-year age groups (Supplementary Figure 4) and ethnic groups (Supplementary Figure 5).

On the FFM scale, the country-specific RMSE values ranged between 1.32kg and 4.83kg, with a RMSE value of <4kg in 17 of the 19 countries (Supplementary Table 2). Calibration plots demonstrated good levels of agreement between deuterium dilution observed FFM and model predicted FFM within each of the countries, including at the lower- and upper-ends of the distribution of FFM, with the flexible calibration curve close to the ideal 45 degree line of perfect calibration (Supplementary Figures 6a & 6b).

As a result of the observed systematic error in the UK-based model’s prediction of lnFFM, the equation was re-calibrated in terms of the intercept term to provide updated country-specific equations (Box 1) and calibration plots (Supplementary Figures 7a & 7b). After re-calibration, the CITL and RMSE values were closer to the ideal values of 0 in all settings, with all of the country-specific 95% CIs for the CITL now containing the ideal value (Supplementary Table 3).

**Discussion**

*Principal findings*

We externally validated the predictive performance of a recently proposed model using weight, height, age, sex, and ethnicity to estimate lnFFM, developed and validated within a UK childhood and adolescent population, in 19 other countries from several regions of the world. The developed model equation generally showed very good predictive ability in these new settings, with good calibration of observed and predicted values, demonstrating the generalisability of the model in childhood populations outside the UK. The model equation produced high R2 values of >80% in all settings, with RMSE values (expressed in terms of FFM for interpretability) of lower than 4kg in the majority of settings. The RMSE of 0.1 for Ln FFM indicates an error of 10% on predicted FFM values, which takes into account that the absolute error in any individual child will depend on the magnitude of their FFM. Notably, the predictive performance was consistently high among both low/middle-income and high-income countries. While the calibration slope statistics were close to the ideal value of 1 for all countries, the CITL values suggested a small systematic error in the prediction of lnFFM across settings. Therefore, the model equation was re-calibrated in terms of the intercept term to provide updated country-specific prediction equations. Following re-calibration, the model performance showed improvement in the CITL and RMSE values in all settings as expected.

*Comparison with other studies*

To our knowledge, this is the first attempt to validate the UK-based prediction model equation6 in settings outside the UK. A small number of previous studies have developed models to estimate body fatness in childhood populations outside the UK,38-43 producing R2 values of >80% comparable to those observed in the present study. However, direct comparisons of the performance of those models with the model being validated in this study are difficult for the following reasons. Firstly, while those models also produced high R2 values of >80%, their outcomes (FM%) are different from the outcome of the current model (absolute FFM values). While FM needs to be standardised for height before interpretation or comparisons between individuals (typically expressed as a fat mass index [FMI] or as FM%), the use of absolute values (of either FM or FFM) as the outcome of the prediction model produces more accurate and precise predictions than estimating FM% which is derived from FM and weight, and is more variable than the absolute values. Furthermore, as discussed previously,6 we chose to estimate absolute FM values indirectly (by predicting FFM from the model and subtracting estimates from weight to obtain predicted FM) rather than directly (using FM as the outcome of the model), as the variability in FFM with height (one of the strongest predictors of body composition) was more homogeneous than for FM and thus the indirect approach resulted in more precise FM predictions. We believe this modelling decision was one of the main reasons for the high predictive performance observed. Secondly, neither the calibration slopes nor the CITL values were assessed for these models. Thirdly, these models largely used DXA as the reference method for assessing body fatness (i.e., as the outcome for model development). However, DXA suffers from low levels of accuracy in its estimation of FM,4 7 8 44 45 which varies considerably by body shape, sex,8 44 and different DXA devices/software. Finally, most models used additional measurements including skinfold thickness, waist circumference and/or bioelectrical impedance in order to estimate FM rather than being based on readily available anthropometric and demographic predictors.39-43 One such study, which used the same data on Tunisian children used in our present study, also developed a prediction model to estimate FFM.30 However, the model was not based on readily available predictors as it relied on resistance from bioimpedence analysis and was not corrected for model optimism, which may explain why the model equation produced a R2 value of 91.8% compared with the value of 81.0% obtained in the present study.

*Strengths and limitations*

The current investigation had several strengths in its approach to validating the developed UK-based prediction model. Crucially, body fatness assessments from each of the 19 countries were made using the same reference standard deuterium dilution method that the UK-based model was based upon,6 which provides accurate, safe, and minimally invasive measurements of TBW (and FFM) with very low error.13 14 Most studies included in this external validation were conducted recently and were sufficiently large to provide accurate estimates of the country-specific prediction performance statistics. The pooled data used for this external validation across the 19 settings spanned a wide age range of 4-15 years, allowing us to obtain an accurate picture of the model performance across childhood and adolescence in the new settings. The maintained high predictive performance of the model at both the lower- and upper-ends of the distribution of FFM indicates the model’s potential utility for population-based obesity surveillance (for example in the English Child Measurement Programme and the WHO Childhood Obesity Surveillance Initiative). The populations included approximately equal numbers of both high- (56%) and low/middle-income countries, which allowed for accurate sub-group assessment by income classsification. The study also had a few limitations to note. Firstly, there was limited global representation, with only a small number of children included from East Asia and none from the Middle East. Secondly, five countries had fewer than 100 participants, which impacted the representativeness of the results and the precision of estimates obtained from these populations, although predictive performance statistics from these populations were consistent with those of the larger countries. Finally, deuterium dilution assessment involves estimating TBW which is then converted into an estimate of FFM using a chosen hydration constant.46 The hydration constants used across the individual studies were those originally applied by individual study authors. However, using a consistent hydration constant published by Wells et al.46 for all studies did not materially affect the study results (results not presented) and therefore the choice of hydration constants is unlikely to have affected the results appreciably.

*Implications for clinicians and policymakers*

Although widely used due to its simplicity and reliance solely on measures of height and weight, BMI has serious limitations as a marker of childhood body fatness, which have long been recognized.2-4 For example, BMI is unable to discriminate between FM and FFM and it has been demonstrated that even at a given BMI among children of the same age and sex, both FM and FFM can vary substantially,3 which is particularly concerning in light of evidence that, independent of height, childhood FM is strongly associated with emerging cardiometabolic risk47-56 and is more strongly related with adult type 2 diabetes risk than overall weight (which BMI is directly based upon).5 Furthermore, it has been suggested that FFM, independent of FM, also has implications on adverse health conditions. This emphasises the needed for a shift away from proxy weight-for-height indices such as BMI, which produce inaccurate childhood body fatness assessment12 57-59 and result in a large proportion of children being misdiagnosed with overweight or obesity, and towards more direct and accurate assessment of FM. The availability of the extensively validated prediction model provides a major advancement within this area of childhood body fatness assessment. The accuracy and simplicity of the model, relying solely on readily available non-invasive measurements to assess FFM and thus FM (as FM = weight – exponential[predicted lnFFM]), has implications for its wider applicability both in routine healthcare practice and in population-wide obesity surveillance, monitoring, and prevention initiatives where more complex measurements of body fatness (such as waist circumference and skinfold thickness) are not so readily available. The consistency of the model performance across settings also strengthens the conclusions that the model, particularly after local-level re-calibration, minimises the bias in FFM predictions across a wide range of settings. However, the UK-based model prior to local country-specific re-calibration of the intercept terms, was shown to have good predictive ability across the settings, both in high-income and low/middle-income populations, and thus can be implemented in childhood settings where a re-calibrated equation is not provided.

Body fatness markers require standardisation for height to minimize their correlation with height,60 in order to provide a consistent measure across age and sex, allowing for efficient monitoring and tracking of childhood body fatness. Two common approaches to provide appropriate standardisation of FM for height, are to convert FM to FM% (FM/weight \*100) or to FMI (FM/heightp, where p is the power of height needed to obtain maximum height independence in the population of interest). FM assessment (using either the updated country-specific model equation provided where available or the original UK-based model), following height standardisation (either as FM% or as FMI), can then be used in conjunction with respective established reference curves61-63 for improved child obesity surveillance, management and prevention. A MS Excel calculator has been developed (supplementary file) to allow for simple calculation of FFM, FM and FM% from the relevant predictor variables. Assessments of FM can also be made readily available to clinicians by embedding the validated equation within existing computer software used by general practitioners and paediatricians. This approach would be consistent with the use of other prediction algorithms in clinical practice such as; the QRISK3 or Framingham Risk Score (CVD).64 65 It would also be possible and straightforward to apply the algorithms within existing childhood obesity surveillance initiatives such as the English National Child Measurement Programme and the WHO Childhood Obesity Surveillance Initiative, to provide assessments of FM, FFM and FMI or FM%.

*Further Research*

Further external validation of the model in countries/regions not included within this study would be of value. Additionally, the development of sex- and age-specific FM reference values, based upon prospectively associated risks of diseases associated with obesity, could allow individuals to be classified into groups based on future disease risk attributable to their current FM levels, as opposed to current centile-based approaches to classify FM.61

**Summary**

**What is already known on this topic**

Improvements to the assessment of childhood body fatness, currently based upon body mass index (BMI), are required.

Assessment of childhood fat mass, which is more strongly related with adult type 2 diabetes risk than weight (which BMI is based upon), could provide improvements

A prediction model which accurately estimates fat mass levels in healthy children and adolescents has been developed and validated for the UK childhood population but its performance in other populations is unknown.

**What this study adds**

This external validation study demonstrates strong predictive performance of the UK-based model at estimating fat-free mass in a wide range of non-UK settings.

The equation, which is based on readily available markers of height, weight, age, sex and ethnic group, performed consistently well in both low/middle- and high-income settings demonstrating its wider generalisability.

The re-calibrated model equations for each of the 19 countries further improve the accuracy of fat-free mass predictions and are recommended for future use

**Contributors:** MTH, RDR, CGO, PHW and CMN designed the study. LSA, JRAA, MNA, MNB, JB, HBJ, BC, EAC, GJC, PSWD, MD, DD, DG, EVG, FH, MH, IK, AM, MAM, LO, HLN, GP, FFR, AER, ER, RJS, JGS, GHtH, JV, JCKW, VPW collected the data. MTH, RDR, CMN analysed the data. MTH, RDR, CGO, PHW and CMN interpreted the data. MTH, RDR, CGO, PHW and CMN drafted the manuscript. All authors critically evaluated and revised the manuscript. The corresponding author (MTH) attests that all listed authors meet authorship criteria, that no others meeting the criteria have been omitted and serves as the guarantors for the contents of this paper.

**Acknowledgements:** We are grateful to Professor John Reilly for his advice on data sources and data access. We are also grateful to Cara L Eckhardt, Josephine Avila, Igor Y Kon, Jinzhong Wang from the Eckhardt et al. (2003) study and all staff involved in recruitment and data collection from all the studies included. Data gathered from South Africa was supported by South Africa Medical Research Council (SAMRC) and National Research Foundation (NRF). We would like to thank the peer reviewers and editors for their helpful feedback and comments which helped improve the article.

**Funding:** MH was funded by a British Heart Foundation PhD studentship (FS/17/76/33286) in the early stages of this investigation and is supported by St. George’s, University of London. CMN was supported by the Wellcome Trust Institutional Strategic Support Fund (204809/Z/16/Z) awarded to St George’s, University of London. This research was supported by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration South London (NIHR ARC South London). This publication is the work of the authors who will serve as guarantors for the contents of this paper. The views expressed in this paper are those of the authors and not necessarily those of the NIHR, the Department of Health and Social Care or other funding agencies. The funders had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

**Competing interests:** All authors have completed the ICMJE uniform disclosure form and declare: this research was supported by grants from the British Heart Foundation (FS/17/76/33286), the Wellcome Trust (204809/Z/16/Z) and the NIHR ARC South London; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

**Ethical approval:** Ethical approval was not required for this study based on secondary data analysis.

**Data Sharing:**  Data from each of the individual studies are available upon request from the respective principal investigators.

**Transparency:** The lead author (MTH) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

**Provenance and peer review:** Not commissioned; externally peer reviewed.

**Licencing: This manuscript must be published under a CC BY NC licence to prevent the content from being exploited commercially.**

**FIGURE LEGENDS:**

Figure 1:

Title - External validation predictive performance statistics based on lnFFM, by country and overall

Footnote - Performance based on ln(fat-free mass). Overall estimates from random-effect REML model with Hartung-Knapp standard errors. Green line around the ‘overall’ diamond indicates the 95% prediction intervals. Upper limit of the prediction interval for R2 capped at 100%

Figure 2a:

Title - Calibration assessment of the model based on lnFFM in the Americas and European countries

Footnote - Calibration based on ln(fat-free mass). Dashed line represents line of equality. Blue line is a loess smoother through the individual data points. Histogram is the distribution of predicted ln(fat-free mass). Slope = Calibration Slope and CITL = Calibration-in-the-Large.

Figure 2b:

Title - Calibration assessment of the model based on lnFFM in the African, Asian and Australasian countries

Footnote - Calibration based on ln(fat-free mass). Dashed line represents line of equality. Blue line is a loess smoother through the individual data points. Histogram is the distribution of predicted ln(fat-free mass). Slope = Calibration Slope and CITL = Calibration-in-the-Large.

Box 1:

Title - Re-calibrated country-specific model equations for the prediction of natural log-transformed FFM

Footnote - Country-specific constant term for UK obtained from equation provided in Hudda MT et al. Development and validation of a prediction model for fat mass in children and adolescents: meta-analysis using individual participant data. BMJ. 2019;366:l4293.

**References**

1. World Health Organisation. Obesity and overweight 2018 [Available from: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight> accessed 09/12/2020.

2. Hall DM, Cole TJ. What use is the BMI? *Arch Dis Child* 2006;91:283-86.

3. Wells JC. A Hattori chart analysis of body mass index in infants and children. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity* 2000;24:325-29.

4. Wells JC, Fewtrell MS. Measuring body composition. *Arch Dis Child* 2006;91:612-17.

5. Hudda MT, Aarestrup J, Owen CG, et al. Association of Childhood Fat Mass and Weight With Adult-Onset Type 2 Diabetes in Denmark. *JAMA Netw Open* 2021;4:e218524.

6. Hudda MT, Fewtrell MS, Haroun D, et al. Development and validation of a prediction model for fat mass in children and adolescents: meta-analysis using individual participant data. *BMJ* 2019;366:l4293.

7. Williams JE, Wells JC, Wilson CM, et al. Evaluation of Lunar Prodigy dual-energy X-ray absorptiometry for assessing body composition in healthy persons and patients by comparison with the criterion 4-component model. *Am J Clin Nutr* 2006;83:1047-54.

8. Reilly JJ, Gerasimidis K, Paparacleous N, et al. Validation of dual-energy x-ray absorptiometry and foot-foot impedance against deuterium dilution measures of fatness in children. *International journal of pediatric obesity : IJPO : an official journal of the International Association for the Study of Obesity* 2010;5:111-15.

9. Hudda MT, Owen CG, Rudnicka AR, et al. Quantifying childhood fat mass: comparison of a novel height-and-weight-based prediction approach with DXA and bioelectrical impedance. *International journal of obesity* 2021;45:99-103.

10. Rush EC, Puniani K, Valencia ME, et al. Estimation of body fatness from body mass index and bioelectrical impedance: comparison of New Zealand European, Maori and Pacific Island children. *European journal of clinical nutrition* 2003;57:1394-401.

11. Shaw NJ, Crabtree NJ, Kibirige MS, et al. Ethnic and gender differences in body fat in British schoolchildren as measured by DXA. *Arch Dis Child* 2007;92:872-5.

12. Hudda MT, Nightingale CM, Donin AS, et al. Body mass index adjustments to increase the validity of body fatness assessment in UK black African and South Asian children. *International journal of obesity* 2017;41:1048-55.

13. Deurenberg P, Yap M. The assessment of obesity: methods for measuring body fat and global prevalence of obesity. *Baillieres Best Pract Res Clin Endocrinol Metab* 1999;13:1-11.

14. Wells JC, Fuller NJ, Dewit O, et al. Four-component model of body composition in children: density and hydration of fat-free mass and comparison with simpler models. *Am J Clin Nutr* 1999;69:904-12.

15. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1-73.

16. Desmond MA, Sobiecki JG, Jaworski M, et al. Growth, body composition, and cardiovascular and nutritional risk of 5- to 10-y-old children consuming vegetarian, vegan, or omnivore diets. *Am J Clin Nutr* 2021;113:1565-77.

17. Wickramasinghe VP, Cleghorn GJ, Edmiston KA, et al. Validity of BMI as a measure of obesity in Australian white Caucasian and Australian Sri Lankan children. *Annals of human biology* 2005;32:60-71.

18. Danysh HE, Gilman RH, Wells JC, et al. El Niño adversely affected childhood stature and lean mass in northern Peru. *Climate Change Responses* 2014;1:7.

19. Haschke F. Body composition of adolescent males. Part I. Total body water in normal adolescent males. Part II. Body composition of the male reference adolescent. *Acta Paediatr Scand Suppl* 1983;307:1-23.

20. Diouf A, Adom T, Aouidet A, et al. Body mass index vs deuterium dilution method for establishing childhood obesity prevalence, Ghana, Kenya, Mauritius, Morocco, Namibia, Senegal, Tunisia and United Republic of Tanzania. *Bulletin of the World Health Organization* 2018;96:772-81.

21. Khan AI, Hawkesworth S, Hawlader MD, et al. Body composition of Bangladeshi children: comparison and development of leg-to-leg bioelectrical impedance equation. *J Health Popul Nutr* 2012;30:281-90.

22. Reichert FF, Wells JC, Ekelund U, et al. Prospective Associations Between Physical Activity Level and Body Composition in Adolescence: 1993 Pelotas (Brazil) Birth Cohort. *Journal of physical activity & health* 2015;12:834-9.

23. Eckhardt CL, Adair LS, Caballero B, et al. Estimating body fat from anthropometry and isotopic dilution: a four-country comparison. *Obesity research* 2003;11:1553-62.

24. Ballesteros-Vásquez MN, Guerrero-Alcocer EV, Grijalva Haro MI, et al. Effect of ingestion and excretion of fluids in determining body composition with deuterium dilution method in school children. *Nutricion hospitalaria* 2015; 32(3).

25. Mendoza Pablo PA, Valdés J, Ortiz-Hernández L. Accuracy of body mass index for age to diagnose obesity in Mexican schoolchildren. *Nutr Hosp* 2015;31:2668-75.

26. Devakumar D, Grijalva-Eternod CS, Roberts S, et al. Body composition in Nepalese children using isotope dilution: the production of ethnic-specific calibration equations and an exploration of methodological issues. *PeerJ* 2015;3:e785.

27. Ten Hoor GA, Rutten GM, Van Breukelen GJP, et al. Strength exercises during physical education classes in secondary schools improve body composition: a cluster randomized controlled trial. *The international journal of behavioral nutrition and physical activity* 2018;15:92.

28. Alvero-Cruz JR, Alvarez Carnero E, Fernández-García JC, et al. Validity of body mass index and fat mass index as indicators of overweight status in Spanish adolescents: Esccola Study. *Medicina clinica* 2010;135:8-14.

29. Wickramasinghe VP, Lamabadusuriya SP, Cleghorn GJ, et al. Validity of currently used cutoff values of body mass index as a measure of obesity in Sri Lankan children. *The Ceylon medical journal* 2009;54:114-19.

30. Ben Jemaa H, Mankai A, Khlifi S, et al. Development and validation of impedance-based equations for the prediction of total body water and fat-free mass in children aged 8-11 years. *Clinical Nutrition (Online)* 2018:227-33.

31. Horlick M, Arpadi SM, Bethel J, et al. Bioelectrical impedance analysis models for prediction of total body water and fat-free mass in healthy and HIV-infected children and adolescents. *Am J Clin Nutr* 2002;76:991-9.

32. Shypailo RJ, Wong WW. Fat and fat-free mass index references in children and young adults: assessments along racial and ethnic lines. *The American Journal of Clinical Nutrition* 2020;112:566-75.

33. Viechtbauer W. Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *Journal of Educational and Behavioral Statistics* 2005;30:261-93.

34. Hartung J, Knapp G. On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Stat Med* 2001;20:1771-82.

35. Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;342:d549.

36. Tan LJ. Confidence Intervals for Comparison of the Squared Multiple Correlation Coefficients of Non-nested Models. The University of Western Ontario, 2012.

37. The World Bank. The world by income 2020 [Available from: <https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html> accessed 22/12/2021.

38. Dugas LR, Cao G, Luke AH, et al. Adiposity is not equal in a multi-race/ethnic adolescent population: NHANES 1999-2004. *Obesity (Silver Spring)* 2011;19:2099-101.

39. Stevens J, Cai J, Truesdale KP, et al. Percent body fat prediction equations for 8- to 17-year-old American children. *Pediatr Obes* 2014;9:260-71.

40. Stevens J, Ou FS, Cai J, et al. Prediction of percent body fat measurements in Americans 8 years and older. *International journal of obesity* 2015;40:587.

41. Stevens J, Truesdale KP, Cai J, et al. Nationally representative equations that include resistance and reactance for the prediction of percent body fat in Americans. *International journal of obesity* 2017;41:1669-75.

42. Truesdale KP, Roberts A, Cai J, et al. Comparison of Eight Equations That Predict Percent Body Fat Using Skinfolds in American Youth. *Childhood obesity (Print)* 2016;12:314-23.

43. Lohman TG, Caballero B, Himes JH, et al. Estimation of body fat from anthropometry and bioelectrical impedance in Native American children. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity* 2000;24:982-8.

44. Sinaiko AR, Donahue RP, Jacobs DR, et al. Relation of Weight and Rate of Increase in Weight During Childhood and Adolescence to Body Size, Blood Pressure, Fasting Insulin, and Lipids in Young Adults. *Circulation* 1999;99:1471-76.

45. Wong WW, Hergenroeder AC, Stuff JE, et al. Evaluating body fat in girls and female adolescents: advantages and disadvantages of dual-energy X-ray absorptiometry. *Am J Clin Nutr* 2002;76:384-9.

46. Wells JC, Williams JE, Chomtho S, et al. Pediatric reference data for lean tissue properties: density and hydration from age 5 to 20 y. *Am J Clin Nutr* 2010;91:610-8.

47. Steinberger J, Jacobs DR, Raatz S, et al. Comparison of body fatness measurements by BMI and skinfolds vs dual energy X-ray absorptiometry and their relation to cardiovascular risk factors in adolescents. *International journal of obesity* 2005;29:1346-52.

48. Lindsay RS, Hanson RL, Roumain J, et al. Body mass index as a measure of adiposity in children and adolescents: relationship to adiposity by dual energy x-ray absorptiometry and to cardiovascular risk factors. *J Clin Endocrinol Metab* 2001;86:4061-7.

49. Falaschetti E, Hingorani AD, Jones A, et al. Adiposity and cardiovascular risk factors in a large contemporary population of pre-pubertal children. *Eur Heart J* 2010;31:3063-72.

50. Whincup PH, Gilg JA, Donald AE, et al. Arterial distensibility in adolescents: the influence of adiposity, the metabolic syndrome, and classic risk factors. *Circulation* 2005;112:1789-97.

51. Samouda H, de Beaufort C, Stranges S, et al. Adding anthropometric measures of regional adiposity to BMI improves prediction of cardiometabolic, inflammatory and adipokines profiles in youths: a cross-sectional study. *BMC Pediatr* 2015;15:168.

52. Gutin B, Islam S, Manos T, et al. Relation of percentage of body fat and maximal aerobic capacity to risk factors for atherosclerosis and diabetes in black and white seven- to eleven-year-old children. *The Journal of pediatrics* 1994;125:847-52.

53. Wells JC, Cole TJ. Height, adiposity and hormonal cardiovascular risk markers in childhood: how to partition the associations? *International journal of obesity* 2014;38:930-5.

54. Weber DR, Leonard MB, Shults J, et al. A Comparison of Fat and Lean Body Mass Index to BMI for the Identification of Metabolic Syndrome in Children and Adolescents. *The Journal of Clinical Endocrinology & Metabolism* 2014;99:3208-16.

55. Gishti O, Gaillard R, Durmus B, et al. BMI, total and abdominal fat distribution, and cardiovascular risk factors in school-age children. *Pediatr Res* 2015;77:710-8.

56. Nightingale CM, Rudnicka AR, Owen CG, et al. Influence of adiposity on insulin resistance and glycemia markers among U.K. Children of South Asian, black African-Caribbean, and white European origin: child heart and health study in England. *Diabetes Care* 2013;36:1712-19.

57. L'Abee C, Visser GH, Liem ET, et al. Comparison of methods to assess body fat in non-obese six to seven-year-old children. *Clinical nutrition (Edinburgh, Scotland)* 2010;29:317-22.

58. Freedman DS, Sherry B. The validity of BMI as an indicator of body fatness and risk among children. *Pediatrics* 2009;124 Suppl 1:S23-34.

59. Simmonds M, Llewellyn A, Owen CG, et al. Simple tests for the diagnosis of childhood obesity: a systematic review and meta-analysis. *Obesity reviews* 2016;17:1301-15.

60. Benn RT. Some mathematical properties of weight-for-height indices used as measures of adiposity. *Br J Prev Soc Med* 1971;25:42-50.

61. McCarthy HD, Cole TJ, Fry T, et al. Body fat reference curves for children. *International journal of obesity* 2006;30:598-602.

62. Wells JC, Williams JE, Chomtho S, et al. Body-composition reference data for simple and reference techniques and a 4-component model: a new UK reference child. *Am J Clin Nutr* 2012;96:1316-26.

63. Weber DR, Moore RH, Leonard MB, et al. Fat and lean BMI reference curves in children and adolescents and their utility in identifying excess adiposity compared with BMI and percentage body fat. *Am J Clin Nutr* 2013;98:49-56.

64. Wilson PW, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97:1837-47.

65. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *Bmj* 2017;357:j2099.

**TABLES**

Table 1: Basic Summary Statistics of the analysis population, by country

|  |  |  |
| --- | --- | --- |
|  | **Median (25th – 75th centile)** |  |
| **Region and Country** | **Age (years)** | **Height (m)** | **Weight (kg)** | **DD Fat Mass (kg) \*** | **DD Fat-free Mass (kg) \*** | **Males (N, %)** | **Ethnic Group (N, %)** |
| **White** | **Black** | **South Asian** | **Other Asian** | **Other** |  |
| *North America* |  |  |  |  |  |  |  |  |  |  |  |
| Mexico (N=330) | 8.4 (7.4 - 9.9) | 1.30 (1.25 - 1.40) | 31.6 (24.6 - 40.8) | 7.6 (5.2 - 12.4) | 22.7 (19.6 - 28.4) | 170 (51.5) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 330 (100) |
| USA (N=1810) | 10.6 (8.3 - 13.0) | 1.44 (1.30 - 1.58) | 39.2 (28.2 - 53.6) | 9.1 (5.7 - 15.0) | 28.5 (21.3 - 38.8) | 867 (47.9) | 571 (32) | 457 (25) | 2 (0.1) | 283 (16) | 497 (27) |
|   |  |  |  |  |  |   |  |  |  |  |  |
| *South America* |  |  |  |  |  |   |  |  |  |  |  |
| Brazil (N=450) | 13.3 (13.1 - 13.6) | 1.58 (1.52 - 1.63) | 49.5 (42.5 - 56.3) | 10.1 (6.6 - 15.2) | 38.1 (34.0 - 43.4) | 236 (52.4) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 450 (100) |
| Peru (N=56) | 11.0 (8.5 - 13.5) | 1.46 (1.32 - 1.55) | 40.2 (32.4 - 50.6) | 9.4 (5.8 - 14.6) | 30.9 (23.9 - 35.5) | 25 (44.6) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 56 (100) |
|  |  |  |  |  |  |   |  |  |  |  |  |
| *Europe* |  |  |  |  |  |   |  |  |  |  |  |
| Austria (N=107) | 12.0 (11.1 - 13.6) | 1.49 (1.42 - 1.60) | 40.1 (34.0 - 48.0) | 5.5 (4.2 - 8.5) | 33.2 (28.5 - 40.6) | 107 (100.0) | 107 (100) | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Netherlands (N=716) | 13.5 (12.8 - 14.1) | 1.62 (1.57 - 1.68) | 51.3 (44.9 - 58.9) | 11.9 (8.8 - 16.9) | 38.6 (34.0 - 43.6) | 342 (47.8) | 716 (100) | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Poland (N=174) | 7.3 (6.1 - 8.7) | 1.25 (1.18 - 1.34) | 24.6 (21.0 - 28.0) | 4.3 (3.3 - 5.8) | 20.0 (17.5 - 22.8) | 81 (46.6) | 174 (100) | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Russia (N=197) | 10.8 (8.9 - 13.7) | 1.47 (1.34 - 1.62) | 37.9 (28.8 - 51.5) | 11.3 (8.3 - 17.3) | 25.8 (20.0 - 33.1) | 97 (49.2) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 197 (100) |
| Spain (N=92) | 14.0 (13.0 - 15.2) | 1.61 (1.55 - 1.68) | 55.8 (48.1 - 63.0) | 13.7 (8.8 - 17.9) | 42.7 (37.0 - 49.2) | 46 (50.0) | 89 (97) | 2 (2) | 0 (0) | 0 (0) | 1 (1) |
|  |  |  |  |  |  |   |  |  |  |  |  |
| *North Africa* |  |  |  |  |  |   |  |  |  |  |  |
| Tunisia (N=155) | 9.0 (8.0 - 10.0) | 1.38 (1.31 - 1.44) | 32.0 (27.0 - 36.0) | 8.0 (6.4 - 11.1) | 23.1 (20.2 - 27.0) | 80 (51.6) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 155 (100) |
|  |  |  |  |  |  |   |  |  |  |  |  |
| *Sub-Saharan Africa* |  |  |  |  |  |   |  |  |  |  |  |
| Namibia (N=151) | 10.0 (9.0 - 11.0) | 1.38 (1.33 - 1.46) | 33.3 (27.2 - 43.4) | 8.6 (6.1 - 14.5) | 23.6 (20.6 - 28.1) | 66 (43.7) | 0 (0) | 114 (76) | 0 (0) | 0 (0) | 37 (25) |
| South Africa (N=411) | 8.0 (7.0 - 8.8) | 1.24 (1.17 - 1.31) | 23.7 (20.4 - 28.3) | 5.5 (4.4 - 7.6) | 17.6 (15.3 - 21.2) | 175 (42.6) | 0 (0) | 411 (100) | 0 (0) | 0 (0) | 0 (0) |
|   |  |  |  |  |  |   |  |  |  |  |  |
| *South Asia* |  |  |  |  |  |   |  |  |  |  |  |
| Bangladesh (N=187) | 5.1 (5.0 - 7.1) | 1.10 (1.02 - 1.17) | 15.8 (14.1 - 18.4) | 2.0 (1.4 - 2.8) | 13.6 (12.2 - 16.2) | 93 (49.7) | 0 (0) | 0 (0) | 187 (100) | 0 (0) | 0 (0) |
| Nepal (N=100) | 8.6 (8.3 - 9.0) | 1.23 (1.16 - 1.29) | 21.8 (18.0 - 25.8) | 4.2 (3.3 - 5.8) | 17.6 (14.6 - 20.2) | 49 (49.0) | 0 (0) | 0 (0) | 100 (100) | 0 (0) | 0 (0) |
| Sri Lanka (N=288) | 10.0 (7.6 - 12.2) | 1.37 (1.24 - 1.49) | 31.5 (23.0 - 41.1) | 9.1 (5.2 - 15.3) | 21.5 (16.2 - 27.1) | 162 (56.3) | 0 (0) | 0 (0) | 288 (100) | 0 (0) | 0 (0) |
|  |  |  |  |  |  |   |  |  |  |  |  |
| *East Asia* |  |  |  |  |  |   |  |  |  |  |  |
| China (N=95) | 10.0 (9.4 - 10.6) | 1.38 (1.33 - 1.43) | 32.2 (27.4 - 37.0) | 6.2 (3.9 - 8.7) | 25.6 (22.8 - 28.2) | 48 (50.5) | 0 (0) | 0 (0) | 0 (0) | 95 (100) | 0 (0) |
| Philippines (N=80) | 15.4 (15.1 - 15.7) | 1.57 (1.52 - 1.64) | 48.9 (43.4 - 54.5) | 13.8 (9.5 - 17.2) | 34.9 (31.3 - 40.8) | 32 (40.0) | 0 (0) | 0 (0) | 0 (0) | 80 (100) | 0 (0) |
|  |  |  |  |  |  |   |  |  |  |  |  |
| *Australasia* |  |  |  |  |  |  |  |  |  |  |  |
| Australia (N=42) | 8.2 (7.0 - 10.9) | 1.33 (1.19 - 1.43) | 27.2 (23.1 - 37.1) | 6.6 (4.2 - 11.0) | 21.9 (16.7 - 27.5) | 27 (64.3) | 0 (0) | 0 (0) | 42 (100) | 0 (0) | 0 (0) |
| New Zealand (N=252) | 10.1 (7.4 - 12.4) | 1.42 (1.27 - 1.54) | 36.5 (28.6 - 50.0) | 9.6 (6.3 - 15.6) | 27.2 (21.7 - 37.1) | 124 (49.2) | 82 (33) | 0 (0) | 0 (0) | 0 (0) | 170 (67) |
|   |  |  |  |  |  |  |  |  |  |  |  |

Footnote: DD = Deuterium Dilution

Table 2: External validation predictive performance statistics based on lnFFM, by country

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Region and Country** | **N** | **R2 (%)** | **Calibration Slope** | **Calibration-in-the-Large** | **RMSE** |
| *North America* |  |  |  |  |  |
| Mexico | 330 | 92.95 (91.49 to 94.42) | 1.01 (0.98 to 1.04) | 0.05 (0.04 to 0.05) | 0.08 |
| USA | 1810 | 93.32 (92.72 to 93.91) | 1.00 (0.99 to 1.01) | 0.02 (0.01 to 0.02) | 0.10 |
|   |  |  |  |  |  |
| *South America* |  |  |  |  |  |
| Brazil | 450 | 76.69 (72.92 to 80.46) | 0.96 (0.91 to 1.01) | 0.05 (0.04 to 0.06) | 0.10 |
| Peru | 56 | 92.29 (88.41 to 96.17) | 0.94 (0.87 to 1.01) | 0.04 (0.02 to 0.06) | 0.09 |
|  |  |  |  |  |  |
| *Europe* |  |  |  |  |  |
| Austria | 107 | 91.47 (88.37 to 94.56) | 0.96 (0.90 to 1.02) | 0.06 (0.05 to 0.07) | 0.09 |
| Netherlands | 716 | 81.53 (79.09 to 83.97) | 1.00 (0.97 to 1.04) | -0.03 (-0.03 to -0.02) | 0.09 |
| Poland | 174 | 93.28 (91.36 to 95.21) | 0.96 (0.92 to 0.99) | 0.04 (0.03 to 0.05) | 0.07 |
| Russia | 197 | 91.30 (88.98 to 93.62) | 0.93 (0.89 to 0.97) | -0.12 (-0.13 to -0.10) | 0.15 |
| Spain | 92 | 80.85 (73.82 to 87.89) | 0.91 (0.82 to 1.00) | 0.03 (0.01 to 0.05) | 0.10 |
|  |  |  |  |  |  |
| *North Africa* |  |  |  |  |  |
| Tunisia | 155 | 80.98 (75.59 to 86.37) | 1.02 (0.94 to 1.10) | -0.02 (-0.03 to -0.01) | 0.08 |
|  |  |  |  |  |  |
| *Sub-Saharan Africa* |  |  |  |  |  |
| Namibia | 151 | 90.14 (87.16 to 93.13) | 0.93 (0.88 to 0.98) | -0.06 (-0.07 to -0.05) | 0.09 |
| South Africa | 411 | 91.95 (90.46 to 93.44) | 1.05 (1.02 to 1.08) | -0.05 (-0.05 to -0.04) | 0.08 |
|   |  |  |  |  |  |
| *South Asia* |  |  |  |  |  |
| Bangladesh | 187 | 89.50 (86.65 to 92.35) | 0.99 (0.94 to 1.04) | 0.10 (0.09 to 0.10) | 0.11 |
| Nepal | 100 | 91.66 (88.53 to 94.79) | 0.99 (0.93 to 1.04) | 0.06 (0.05 to 0.07) | 0.08 |
| Sri Lanka | 288 | 83.11 (79.56 to 86.67) | 0.99 (0.94 to 1.04) | -0.06 (-0.07 to -0.04) | 0.16 |
|   |  |  |  |  |  |
| *East Asia* |  |  |  |  |  |
| China | 95 | 85.10 (79.57 to 90.63) | 0.98 (0.90 to 1.07) | 0.08 (0.07 to 0.10) | 0.11 |
| Philippines | 80 | 81.03 (73.55 to 88.51) | 1.04 (0.93 to 1.15) | -0.05 (-0.07 to -0.03) | 0.09 |
|  |  |  |  |  |  |
| *Australasia* |  |  |  |  |  |
| Australia | 42 | 96.64 (94.65 to 98.64) | 1.05 (0.98 to 1.11) | 0.03 (0.01 to 0.05) | 0.07 |
| New Zealand | 252 | 92.96 (91.29 to 94.64) | 0.98 (0.94 to 1.01) | 0.03 (0.02 to 0.04) | 0.10 |

Footnote: Performance based on ln(fat-free mass). RMSE = Root Mean Square Error

Table 3: External validation predictive performance statistics based on lnFFM, by World Bank Income Classifications

|  |  |  |
| --- | --- | --- |
|  | **Low/Middle- Income Group****N=2,473** | **High-Income Group****N=3,193** |
| **R2 (%)** | 92.19 (91.60 to 92.78) | 93.64 (93.21 to 94.07) |
| **Calibration Slope** | 0.98 (0.97 to 0.99) | 0.97 (0.97 to 0.98) |
| **Calibration-in-the-Large** | -0.00 (-0.01 to 0.00) | 0.01 (0.01 to 0.01) |
| **RMSE** | 0.11 | 0.10 |

Footnote: Performance based on ln(fat-free mass). RMSE = Root Mean Square Error. Income Classifications ascertained for the initial calendar year the study began. Low-, Lower Middle- and Upper Middle-Income groups have been combined into a Low/Middle-Income group. Low/Middle-Income group income group includes Bangladesh, Brazil, China, Mexico, Namibia, Nepal, Peru, the Philippines, Russia, South Africa, Sri Lanka and Tunisia. High-income group includes Australia, Austria, Netherlands, New Zealand, Spain, USA and Poland

**Data from Figure 1 forest plot**

|  |  |  |  |
| --- | --- | --- | --- |
| **Country** | **R2** | **Calibration Slope** | **Calibration-in-the-Large** |
| Australia | 96.64 (94.65 to 98.64) | 1.05 (0.98 to 1.11) | 0.03 (0.01 to 0.05) |
| Austria | 91.47 (88.37 to 94.56) | 0.96 (0.90 to 1.02) | 0.06 (0.05 to 0.07) |
| Bangladesh | 89.50 (86.65 to 92.35) | 0.99 (0.94 to 1.04) | 0.10 (0.09 to 0.10) |
| Brazil | 76.69 (72.92 to 80.46) | 0.96 (0.91 to 1.01) | 0.05 (0.04 to 0.06) |
| China | 85.10 (79.57 to 90.63) | 0.98 (0.90 to 1.07) | 0.08 (0.07 to 0.10) |
| Mexico | 92.95 (91.49 to 94.42) | 1.01 (0.98 to 1.04) | 0.05 (0.04 to 0.05) |
| Namibia | 90.14 (87.16 to 93.13) | 0.93 (0.88 to 0.98) | -0.06 (-0.07 to -0.05) |
| Nepal | 91.66 (88.53 to 94.79) | 0.99 (0.93 to 1.04) | 0.06 (0.05 to 0.07) |
| Netherlands | 81.53 (79.09 to 83.97) | 1.00 (0.97 to 1.04) | -0.03 (-0.03 to -0.02) |
| New Zealand | 92.96 (91.29 to 94.64) | 0.98 (0.94 to 1.01) | 0.03 (0.02 to 0.04) |
| Peru | 92.29 (88.41 to 96.17) | 0.94 (0.87 to 1.01) | 0.04 (0.02 to 0.06) |
| Philippines | 81.03 (73.55 to 88.51) | 1.04 (0.93 to 1.15) | -0.05 (-0.07 to -0.03) |
| Poland | 93.28 (91.36 to 95.21) | 0.96 (0.92 to 0.99) | 0.04 (0.03 to 0.05) |
| Russia | 91.30 (88.98 to 93.62) | 0.93 (0.89 to 0.97) | -0.12 (-0.13 to -0.10) |
| South Africa | 91.95 (90.46 to 93.44) | 1.05 (1.02 to 1.08) | -0.05 (-0.05 to -0.04) |
| Spain | 80.85 (73.82 to 87.89) | 0.91 (0.82 to 1.00) | 0.03 (0.01 to 0.05) |
| Sri Lanka | 83.11 (79.56 to 86.67) | 0.99 (0.94 to 1.04) | -0.06 (-0.07 to -0.04) |
| Tunisia | 80.98 (75.59 to 86.37) | 1.02 (0.94 to 1.10) | -0.02 (-0.03 to -0.01) |
| USA | 93.32 (92.72 to 93.91) | 1.00 (0.99 to 1.01) | 0.02 (0.01 to 0.02) |
| ***Overall*** | ***88.66 (85.91 to 91.41)*** | ***0.98 (0.97 to 1.00)*** | ***0.01 (-0.02 to 0.04)*** |