**Development and validation of a prediction model for fat mass in children and adolescents: an individual participant data (IPD) meta-analysis**

Mohammed T Hudda (PhD Student)1, Mary S Fewtrell (Professor of Paediatric Nutrition) 2, Dalia Haroun (Associate Professor in Nutrition)3, Sooky Lum (Senior Research Fellow)4, Jane E Williams (Research Nurse)2, Jonathan C K Wells (Professor of Anthropology and Paediatrics)2, Richard D Riley (Professor of Biostatistics)5, Christopher G Owen (Professor of Epidemiology) 1, Derek G Cook (Professor of Epidemiology)1,Alicja R Rudnicka (Professor of Statistical Epidemiology)1, Peter H Whincup (Professor of Cardiovascular Epidemiology)1, Claire M Nightingale (Lecturer in Medical Statistics and Epidemiology)1

1 Population Health Research Institute, St George’s, University of London, London, UK

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

3 College of Natural and Health Sciences, Department of Public Health and Nutrition, Zayed University, Dubai, UAE

4 Respiratory, Critical Care & Anaesthesia section of III Programme, UCL Great Ormond Street Institute of Child Health, London, UK

5 Centre for Prognosis Research, Research Institute for Primary Care and Health Sciences, Keele University, Staffordshire, UK

**Correspondence to:**

Dr Claire Nightingale

Population Health Research Institute,

St George’s, University of London,

Cranmer Terrace,

London SW17 0RE

Tel: (+44) 208 725 2342

Email: [cnightin@sgul.ac.uk](mailto:cnightin@sgul.ac.uk)

ORCID ID: 0000-0002-4803-7617

Twitter: @mohammedhudda

**Word Count:** 3831

**An individual participant data (IPD) meta-analysis to develop and validate a prediction model for fat mass in children and adolescents**

**ABSTRACT**

**Objective**

To develop and validate a prediction model for fat mass (FM) in children and adolescents aged 4-15 years using routinely available risk factors of height, weight and demographic information without the need for more complex forms of assessment

**Design and Setting**

Individual participant data (IPD) meta-analysis of four cross-sectional studies with a fifth (unavailable at time of model development) for external validation; all studies included data on anthropometry and deuterium dilution assessments of FM.

**Participants**

All participants with complete data on anthropometric measurements and deuterium dilution measured body fatness from the four studies and from the validation study.

**Main Outcome Measure**

Multivariable linear regression analysis, using backwards selection for inclusion of predictor variables and allowing non-linear relationships, was used to develop a prediction model for fat free mass (and subsequently fat mass by subtracting resulting estimates from weight) based on the four studies. Internal validation and then internal-external cross-validation was used to examine overfitting and generalisability of the model’s predictive performance within the four development studies; finally external validation followed using the fifth dataset.

**Results**

Model derivation was based on a multi-ethnic population of 2,375 children (48% boys) aged 4-15 years. The final model containing predictor variables of height, weight, age, sex and ethnicity had extremely high predictive ability (R2adjusted: 94.8% [95%CI: 94.4 to 95.2]) with excellent calibration of observed and predicted values. The internal validation demonstrated minimal overfitting and good model generalisability, with excellent calibration and predictive performance. External validation in 176 individuals aged 11-12 years demonstrated promising generalisability of the model (R2: 90.0% [95%CI: 87.2 to 92.8]) with good calibration of observed and predicted FM (slope: 1.02 [95%CI: 0.97 to 1.07]). The mean difference between observed and predicted FM was -1.29 kg (95%CI: -1.62 to -0.96 kg).

**Conclusion**

The developed model accurately predicted levels of FM in children and adolescents. The prediction model is based on simple anthropometric measures without the need for more complex forms of assessment and could improve the accuracy of childhood body fatness assessments (compared to those provided by body mass index) for effective clinical and public health obesity surveillance, prevention and management.

**INTRODUCTION**

With the increasing prevalence of childhood obesity both globally1 and in the UK, where overweight and obesity together affect about a third of children aged 2-15 years,2 high childhood body fatness (BF) represents a serious public health problem. High BF levels in childhood have been shown to be associated with adult overweight and obesity and with elevated risks of non-communicable diseases in adulthood, notably type 2 diabetes and cardiovascular diseases.3-7

Accurate and practical methods for quantifying childhood BF are essential for effective monitoring, prevention and management of high BF, overweight and obesity in childhood.8 9 Body mass index (BMI, weight/height2), the most widely used marker of childhood BF in clinical and public health practice, has serious limitations as a marker of BF in children.9-11 Firstly, as a weight-based measure, it does not discriminate between lean (fat free mass; FFM) and fat mass (FM) which can vary markedly in individuals with a given BMI.10 Secondly, height2 provides poor height standardisation of weight in children; a higher power is needed to obtain height standardisation.12-14 Finally, childhood BMI is not a consistent marker of BF across different ethnic groups. In the UK and the US it has been shown to overestimate BF in black African children and underestimate BF in children of Asian origin.15-19 Similar problems have been reported in other settings; BMI under estimates BF in South Asian girls and over estimates BF in Pacific Island girls in New Zealand. 20

Whilst imaging (by dual-energy x-ray absorptiometry (DXA) or MRI), densitometric and isotope dilution methods are available and accurate, they are unsuitable for routine clinical or public health BF assessment.11 21 Simple methods of BF assessment, based on routinely available measurements (particularly weight and height) and valid in a range of populations, would be of considerable value.

We sought to examine whether weight and height could provide more accurate assessment of FM than that provided by BMI, particularly using prediction methods which have shown promise in estimating disease risks. 22-24

We report here on the development and validation of a prediction model that can estimate FM accurately in UK children and adolescents between 4 and 15 years of age of different ethnic origins, based on weight, height and routinely available basic demographic information.

**METHODS**

**Data sources and study population**

For this investigation we pooled data from four cross-sectional studies for the development of a prediction model, with a fifth study (not available at the time of model derivation) for external validation. All studies included data on weight, height and reference standard BF assessments based on the deuterium dilution (DD) method.

*Derivation Data*

Data from four separate cross-sectional studies17 25-27 (Supplementary Table 1), identified as the four available population-based UK based studies which contained DD measurements together with weight and height measurements in >200 participants in the 4-15 year age, were obtained and pooled for analysis (N=2375). Each of these studies used a similar protocol when conducting the DD method to measure total body water (TBW) (and indirectly FM), as described earlier.15 Three of the four studies included multi-ethnic populations; assessment of ethnicity was based on a combination of self-reported parental information on parental ethnicity,17 child ethnicity,17 26 27 with self-reported participant information on ethnicity for older children,25 26 using ethnic group categories based on the 2001 UK Census (Supplementary Table 1).

*External validation Data*

Data from a smaller separate cross-sectional study at the 11 year follow-up visit within the Avon Longitudinal Study of Parents and Children (ALSPAC)28 were obtained for external validation. ALSPAC is a birth cohort study containing detailed assessments from predominantly White children born in the Bristol area between April 1991 and December 1992 including; height, weight, sex, ethnicity and age. At the 11 year follow-up visit, a sub-sample of the cohort (stratified by sex and BMI to represent the whole cohort) were recruited to participate in a further study which involved FM assessment using the DD method alongside measures of height and weight taken simultaneously.29 Ethnicity was based on a combination of self-reported parental information on parental ethnicity.

**Defining the outcome of prediction models**

The primary aim of this investigation was to develop a model for predicting childhood FM, which could be estimated directly or indirectly (by predicting FFM from models and then subtracting resulting estimates from known weight) based on deuterium dilution measurements. We therefore first investigated the potential for modelling FM directly or indirectly by examining the distributions of FM and FFM in relation to height (one of the strongest predictors of body composition) in boys and girls separately. This showed that a regression model for FFM better met the assumptions of linear regression (more details in Appendix 1). The distribution of FFM (both in boys and girls separately and combined), was positively skewed (Supplementary Figure 1) and showed increased heterogeneity with increases in height and weight. FFM, transformed using natural logarithms (ln FFM), was therefore the outcome in the main analyses.

**Candidate predictors**

Weight, height, age, sex and ethnic group were considered as candidate predictors (variables) in the model development stage. Our derivation data, once restricted to those individuals with FFM or FM assessment, had no missing data on any of the candidate predictors mentioned. The large sample size of 2375 subjects meant that the number of candidate predictors being considered (along with non-linear terms), far exceeded both the minimum ‘10 subjects per candidate predictor’ rule of thumb,30 and the minimum sample size requirements for prediction models proposed by Riley et al. 31 Ethnicity was based on self-reported parental information on parental ethnicity. For the present analyses, child ethnicity was categorized as ‘White’ (of European origin), ‘Black’ (including black children of African and Caribbean descent), ‘South Asian’ (including children of Indian, Pakistani, Bangladeshi and Sri Lankan descent), `Other Asian’ (predominantly those of East Asian origins) and ‘Other’ (predominantly those of mixed ethnicity) groups.

**Statistical analysis for model development**

Stata v14 was used for all analyses. The TRIPOD guidance for development and reporting of multivariable prediction models was followed.32 All four available studies were used for model development, to avoid data splitting.33 A linear regression was used with ln FFM as the outcome and weight, height, age, sex and ethnic group considered as candidate predictors (variables). Using a stepwise approach through backwards elimination, beginning with a model that included all predictors, candidate predictors were excluded from the saturated model based on their statistical significance (pwald>0.05). Non-linear relationships between outcome and continuous predictors were considered by identifying, at each iterative step of the stepwise process, the best-fitting fractional polynomial terms34 35 (using Stata command mfp36). This model development process led to a final model for the prediction of ln FFM (and subsequently for FM = weight – exp[prediction of ln FFM]) based on the selected predictors along with their corresponding estimated beta coefficients and the associated intercept term. Although heterogeneity and clustering of patients across/ within studies was not considered for model development, the impact of this was checked via an internal-external validation approach (see below).

**Model performance and internal validation**

The performance of the final model was assessed using several approaches: i) R2 –proportion of the variance in ln FFM explained by the included predictors; ii) root mean square error (RMSE) –the average difference between the predicted and the observed value; iii) calibration slope – based on the model regressing observed on predicted values of ln FFM (with a slope of 1 being ideal); iv) calibration-in-the-large (CIL) – intercept term from the model regressing observed on predicted values of ln FFM (with an intercept of 0 being ideal); and v) comparing mean observed with mean predicted values of ln FFM. Calibration was also assessed graphically by displaying FFM and FM, on a calibration plot with a local regression (loess) smoother fitted across all individuals. RMSE of FM predictions was also assessed overall and within age-, ethnic- and sex- subgroups.

Internal validation was carried out to estimate optimism (the level of model overfitting),32 and correct measures of predictive performance (R2, calibration slope and CIL) for model overfitting by bootstrapping32 1000 samples of the derivation data (with replacement). The entire variable selection process, including the choosing of the fractional polynomial terms, was repeated within the model development for each of the 1000 bootstrap samples. This led to a set of 1000 bootstrap models that were derived using the same methods as in our original model development. Each of these bootstrap sample models were then applied within the original dataset to estimate optimism in the performance statistics (difference in test performance and apparent bootstrap performance) of R2, calibration slope and CIL (see Appendix 2 for further details), referred to as the R2adjusted, calibration slopeadjusted and CILadjusted. To adjust for optimism after model development, estimates of a uniform shrinkage factor were obtained (the average calibration slope from each of the bootstrap samples) and multiplied by the original beta coefficients to obtain optimism-adjusted coefficients.32 37 At this stage, the intercept of the model was re-estimated based on the adjusted coefficients to maintain overall model calibration,32 producing a final model.

**Internal-External Validation**

It is important to examine the generalisability of a prediction model developed using the above process. Due to the limited availability of appropriate external datasets, internal-external validation38 39 was conducted to further assess the performance of the derived model. This internal-external approach38 39 involved cross validation, omitting one of each of the four studies in turn from the development dataset and re-running the analyses described above. The following steps were undertaken; 1) Using the same model development strategy, a model was developed on three of the four studies and the beta coefficients from the model predicting ln FFM were obtained, 2) The predictive performance of the model from step 1. was assessed (overall and within sex and ethnic groups) within the fourth ‘external validation’ study data in terms of accuracy of predicted FM (the primary outcome) by means of the calibration slope, CIL and the R2 measures and finally 3) Steps 1-2 were repeated until ‘external validation’ was assessed for each of the four studies.

Overfitting was assessed in each round of the cross-validation, a uniform shrinkage factor was obtained37 and applied to the beta coefficients from step 1. Calibration slope, CIL and R2 measures derived from this procedure for each of the studies were then pooled and estimated via a random effects meta-analysis to assess the heterogeneity across studies (using with the τ2 statistic estimated using the Mantel-Haenszel method). The variance of R2 was estimated using the Wald-type method outlined by Tan 40 and used to pool the values.

**External Validation**

We applied our final prediction model to each individual in the external validation dataset based on their respective predictor values. In a very small number of children with missing ethnicity data, the child was reclassified as White in order to produce an estimate of FM. The performance of the model for predicting FM, by sex and overall, was assessed using the calibration slope, CIL, the R2, the RMSE and by comparing mean observed with mean predicted values. We also assessed the overall calibration of the model graphically in terms of FM by plotting agreement between predicted and observed values across tenths of predicted values. Finally, the intercept term from the final model was re-estimated for the external data, to maintain the calibration of the model, and the performance statistics were re-assessed.

**Patient involvement**

No patients were involved in setting the research question or the outcome measures, nor were they involved in the study design or implementation. No patients were involved in the interpretation or writing up of results. There are no plans to disseminate the results of the research to study participants or the relevant patient community.

**RESULTS**

**Study population**

A summary of the key participant characteristics from both datasets is presented in Table 1. The pooled derivation dataset (4 studies) included 2375 children, predominantly of White (37%), Black (23%) and South Asian (25%) origins, aged 4.0-15.9 years (median age 9.6 years, 48% male) had complete information on anthropometric, demographic and BF measurements. The external validation dataset included 176 children predominantly of White origins and on children aged 11-12 years (48% male), with complete data on anthropometric and BF measurements and missing ethnicity data on a very small number of children (<10%). For the pooled derivation dataset, the distribution of age within each of the four individual studies varied; one study contained children across the full age range (albeit in children of only White ethnic origin) while the other three studies each contained a restricted age range, but had marked ethnic diversity.

TABLE 1: CHARACTERISTICS OF INDIVIDUALS WITHIN THE DERIVATION AND VALIDATION DATASETS. VALUES PRESENTED ARE MEDIAN (25th – 75th CENTILE) UNLESS OTHERWISE STATED

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Derivation Dataset | | | | | Validation Dataset  (ALSPAC sub-sample)  N = 176 |
|  |  | Study | | | | |
|  |  | 1 (ABCC)  N=1,027 | 2 (ELBI)  N=382 | 3 (RC)  N=369 | 4 (SLIC)  N=597 | Overall  N=2,375 |
| Age (years) | | 9.3 (8.7-9.7) | 13.3 (12.4-14.3) | 11.1 (8.5-13.1) | 8.5 (7.1-10.1) | 9.6 (8.6-11.1) | 11.8 (11.8-12.0) |
| Height (metres) | | 1.4 (1.3-1.4) | 1.6 (1.5-1.6) | 1.5 (1.3-1.6) | 1.3 (1.2-1.4) | 1.4 (1.3-1.5) | 1.5 (1.5-1.6) |
| Weight (kg) | | 31.6 (27.3-38.1) | 47.3 (39.2-58.8) | 37.2 (27.4-48.5) | 30.3 (24.2-40.1) | 33.9 (27.4-43.8) | 43.3 (37.2-50.2) |
| Fat Mass (kg) | | 8.5 (6.2-12.8) | 10.4 (7.0-16.6) | 7.8 (4.9-12.5) | 7.4 (4.7-12.2) | 8.4 (5.8-13.2) | 9.5 (6.6-13.5) |
| Fat Free Mass (kg) | | 22.9 (20.4-25.9) | 35.4 (30.4-43.3) | 28.4 (21.9-36.3) | 22.7 (18.7-27.9) | 24.8 (20.8-30.6) | 33.8 (29.8-37.4) |
|  | |  |  |  |  |  |  |
| Male (N,%) | | 490 (48) | 182 (48) | 180 (49) | 284 (48) | 1136 (48) | 84 (48) |
|  | |  |  |  |  |  |  |
| Ethnic Group  (N,%) | White | 290 (28) | 91 (24) | 369 (100) | 135 (23) | 885 (37) | 161 (91) |
| Black | 252 (25) | 119 (31) | 0 (0) | 182 (30) | 553 (23) | - |
| South Asian | 325 (32) | 120 (31) | 0 (0) | 141 (24) | 586 (25) | - |
| Other Asian | 46 (4) | 44 (12) | 0 (0) | 22 (4) | 112 (5) | - |
| Other/ Missing | 114 (11) | 8 (2) | 0 (0) | 117 (20) | 239 (10) | 15 (9) |

FOOTNOTE: Fat mass and fat free mass assessments were made using the deuterium dilution method. The derivation dataset contained no missing data. Information on ethnic group was missing on a small number of individuals in the validation dataset. ABCC = The assessment of Body Composition in Children Study, ELBI = The East London Bioelectrical Impedance, RC = The Reference Child, SLIC = The Size and Lung function in Children Study, ALSPAC = The Avon Longitudinal Study of Parents and Children.

**Model development and apparent performance**

The final multivariable model included all of the five candidate predictors of height, weight, age, sex and ethnic group (i.e. none were excluded). Fractional polynomial terms for the continuous predictors (height, weight and age) were included in the final model to allow for non-linear relationships (Table 2). The model, demonstrated excellent apparent predictive performance for ln FFM (Table 3; R2 = 94.8%, RMSE = 0.068) and was, as expected, perfectly calibrated in the development data (apparent slope =1, apparent CIL = 0). This is confirmed by the calibration plot, assessing agreement between observed and predicted FFM and FM (Figure 1). The difference between the mean observed and the mean predicted values of ln FFM was zero. The RMSE values for FM were 2.0 kg and 1.9 kg in females and males respectively, and ranged between 0.9 kg and 3.3 kg within the one-year age groups. Within ethnic groups the RMSE ranged between 1.7 kg amongst South Asians and 2.4 kg amongst the Blacks.

TABLE 2: FINAL MULTIVARIABLE ANALYSIS MODEL IN DERIVATION DATASET AND OPTIMISM-ADJUSTED BETA COEFFICIENTS

|  |  |  |
| --- | --- | --- |
| **Variable** | **Developed Model: Coefficients (95% CI)** | **Final model coefficients after adjusting for overfitting:** |
| Height² | 0.308 (0.289 , 0.327) | 0.307 |
| (Weight/10)^-1 | -1.003 (-1.090 , -0.916) | -1.002 |
| Weight/10 | 0.046 (0.040 , 0.052) | 0.046 |
| Ethnicity (White) | Reference | Reference |
| Black | 0.014 (0.007 , 0.022) | 0.014 |
| South Asian | -0.065 (-0.072 , -0.058) | -0.065 |
| Other Asian | -0.026 (-0.040 , -0.013) | -0.026 |
| Other ethnicity | -0.017 (-0.027 , -0.008) | -0.017 |
| ln(age/10) | -0.919 (-1.086 , -0.753) | -0.918 |
| (age/10)^0.5 | 2.055 (1.708 , 2.401) | 2.052 |
| Sex (Female) | Reference | Reference |
| Male | 0.047 (0.042 , 0.053) | 0.047 |
|  |  |  |
| Constant | 0.692 (0.373 , 1.011) | 0.691 |

FOOTNOTE: Outcome of model was ln(fat free mass). Height in metres, Weight in kilograms and age in years.  
Ln refers to the natural logarithmic transformation. Constant term was re-estimated after adjustment for optimism (shrinkage factor = 0.99858) to uphold overall model calibration

**Model Validation**

*Internal validation*

Bootstrap internal validation showed little model overfitting, which was reflected in the extremely similar apparent and optimism-adjusted performance statistics (Table 3). After adjusting for overfitting, the final prediction model maintained a very high proportion of the variance in ln FFM with an R2adjusted value of 94.8%. The bootstrapping approach provided a shrinkage factor practically 1 (i.e. there was no important overfitting, with the mean calibration slope equal to 1 from the bootstrap models when tested in the original data). We also calculated the uniform shrinkage factors suggested by Van Houwelingen37 and this gave a value of 0.99858, again very close to 1. We chose the latter because it was slightly smaller than the bootstrap value, which was applied to the original beta coefficients from the model to obtain optimism-adjusted coefficients before re-estimation of the intercept term. Box 1 shows the prediction equation for the estimation of FM in children aged 4-15 years along with examples of how to calculate FM using the equation.

TABLE 3: MODEL PERFORMANCE STATISTICS BASED ON INTERNAL VALIDATION

|  |  |  |  |
| --- | --- | --- | --- |
| **Measure** | **Apparent performance** | **Average Optimism** | **Optimism Corrected** |
| R² | 94.83% (94.4 to 95.2) | 0.03% | 94.80% |
| Calibration Slope (95% CI) | 1.00 (0.99 to 1.01) | 1.63 x 10-9 | 1.00 (0.99 to 1.01) |
| Calibration-In-The-Large (95% CI) | 0.00 Kg (-0.03 to 0.03 Kg) | -3.54 x 10-9 Kg | 0.00 Kg (-0.03 to 0.03 Kg) |

Footnote: Outcome of model was ln(fat free mass).

*Internal-External validation*

Using the cross-validation approach mentioned, a model was developed in each of the three studies and applied within the fourth study. Assessments of model overfitting found extremely low levels of optimism at each round of cross-validation (shrinkage factor = 0.998 for each round). Within each of the studies being used as a validation dataset, after adjusting for optimism, the calibration slopes were very close to 1 and the CIL values were very close to 0, suggesting excellent model calibration in each of these four study populations (Figure 2, Supplementary Table 2).

The pooled calibration slopes and CIL values across the four studies were 1.00 [95%CI 0.95 to 1.04) and -0.29 (95%CI -0.83 to 0.25) for FM respectively, suggesting that – on average across the four populations – the model is likely to calibrate very well. The pooled R2 value for FM was 89.7% (95%CI: 87.8% to 91.7%) which indicates that the model, on average, explains a very high proportion of the variance in FM. The tau2 values for the calibration slope, CIL and R2 measures were 0.002, 0.267 and 0.0004 respectively, suggesting very little heterogeneity across the four populations. The calibration slopes and CIL values within sex and ethnic groups showed very good calibration for all subgroups upon each round of cross validation, suggesting that the final model is likely to calibrate well for children of both sexes and each ethnic group (Supplementary Figures 2&3).

*External Validation*

We applied our final prediction model (Box 1) to the independent population of 11-12 year-old children, reclassifying the very small number of children with missing ethnicity information as being from the White reference group. The resulting R2 value from the model was 90.0% (95%CI 87.2 to 92.8%), with a moderate RMSE of 2.6 kg and the model had average calibration in terms of FM (Figure 3); with a slope of 1.02 (95%CI 0.97 to 1.07) and CIL of -1.58 kg (95%CI -2.29 to -0.86 kg) (Table 4). The mean difference between observed and predicted FM was -1.29 kg (95%CI: -1.62 to -0.96 kg). The final model was observed to perform better in girls than in boys (Table 4). . After re-calibration of the intercept, the R2 value from the model was 90.0% (95%CI 87.1 to 92.8%), with a RMSE of 2.4 kg and the model had a calibration slope of 1.06 (95%CI 1.01 to 1.11) and CIL of 0.21 kg (95%CI -0.42 to 0.85 kg).

TABLE 4: EXTERNAL VALIDATION - MODEL PERFORMANCE STATISTICS BEFORE RECALIBRATION OF INTERCEPT

|  |  |  |  |
| --- | --- | --- | --- |
| **Measure** | **Boys** | **Girls** | **Overall** |
| R² (95% CI) | 87.9% (83.1 to 92.8) | 91.6% (88.4 to 94.9) | 90.0% (87.2 to 92.8) |
| Calibration Slope  (95% CI) | 1.05 (0.97 to 1.14) | 1.04 (0.97 to 1.10) | 1.02 (0.97 to 1.07) |
| Calibration-In-The-Large (95% CI) | -1.38 Kg (-2.41 to 0.36 Kg) | -2.25 Kg (-3.27 to -1.23 Kg) | -1.58 Kg (-2.29 to -0.86 Kg) |

Footnote: Performance in terms of fat mass predictions. Statistics presented before intercept term was re-estimated for the external data.

**Sensitivity analyses**

In our final model, we tested and found two-way interactions between sex and weight, and sex and age (along with their appropriate non-linear fractional polynomial terms) to be statistically significant at the 5% level. However, inclusion of additional terms for sex\*weight and sex\*age did not improve the apparent performance of the model (R2 = 94.9%, RMSE = 0.068) with very little difference between the Akaike’s Information Criterion from models including and excluding these terms. Therefore, these interaction terms were not added to the previously described prediction model.

We also investigated the use of the proposed model to estimate childhood FM where the ethnic origins of the child were unknown. Two approaches were taken, one omitting ethnic group as a predictor from the model and one treating children with unknown ethnic origins as being White (the reference group) for the purpose of FM predictions. Both approaches were carried out and compared, using an internal-external validation approach. FM predictions from both approaches had similar levels of bias when compared to observed FM values, suggesting that children of unknown ethnic origins can be treated as White with little effect on the predictive performance.

Finally, to investigate the direct approach of predicting FM, the model development strategy was repeated using ln FM as the primary outcome. The apparent performance of this model (R2 = 83.4%) was much less favourable than those of the main analyses using ln FFM as the primary outcome.

**DISCUSSION**

**Statement of principal findings**

We have developed a new prediction equation, based solely on readily available measures of height, weight, age, sex and ethnic group, to estimate FM levels (kg) for children and adolescents aged 4 to 15 years using a large representative sample of UK individuals. We then validated the model both internally and externally, firstly using a cross-validation approach within the derivation population and secondly in an independent dataset of 11-12 year old children. Both overall and within age-, sex- and ethnic subgroups, the developed model showed extremely high predictive ability with excellent calibration, low individual error with RMSE values less than 3.3 kg and useful R2 values greater than 88% from the derivation, cross-validation and external validation datasets. The average individual error associated with the predictions in the independent dataset was low with a RMSE of 2.6kg.

**Comparison with other studies**

There are few previous studies to our knowledge which have developed and validated prediction models based solely on weight, height and demographic factors to estimate FM in children and adolescents41. The majority of previously derived models for this purpose have focussed on older children and adolescents from the US using dual-energy X-ray absorptiometry as the method of assessing BF.41-46 Moreover, modelling has predominantly been based on the prediction of BF% and not absolute values of BF, making comparisons of the predictive ability of models difficult. The developed models, which have been shown to estimate %BF to a high level with R2 values >82%, have relied upon additional measurements including skinfold thickness, waist circumference and/or bioelectrical impedance in order to estimate BF.42-46 However, the model developed by Dugas et al.41 in US 12-20 year-olds included the same predictors as in our final model, of height and weight (in the form of a fractional polynomial non-linear term of BMI) as well as sex, age and ethnicity to estimate %BF. That model performed well, explaining a high proportion of the variance in %BF (R2 = 79.4%). However, the RMSE was not presented making direct comparisons between the accuracy of the models difficult.

**Strengths and limitations**

To our knowledge, this is the first model which has been developed to estimate childhood FM in UK children and adolescents. It has several strengths. The derivation dataset was sufficiently large, with complete information on all candidate predictors for all participants with FFM information, allowing us to test all of the candidate predictors along with their respective non-linear terms whilst adhering to the ‘10 subjects per candidate predictor’ rule of thumb.30 The wide age range of 4-15 years including a range of ethnic origins allowed derivation of a robust model applicable to a wider target population, with consistent performance of the model across the range of age, sex and ethnic groups. Data collection for all four derivation studies was completed within the last decade (2009-2013) and should have continuing relevance, with no indication that the associations between FFM and its predictors have changed. We were able to identify an additional independent dataset for external validation, though the range of age and ethnicity was narrower. The predictors on which the model are based are simple and already widely measured. The performance of the model is strong and allows discrimination between FM and FFM both in the whole study population and in specific ethnic groups, offering obvious potential advantages over the ethnic-specific BMI adjustments which we reported earlier,15 particularly if earlier reports suggesting that FM is more strongly associated with long term health outcomes than BMI.47 Although the inclusion of non-linear polynomial terms make the derived algorithm appear complicated for practical use, these have been integrated into a simple MS Excel calculator (Online Supplementary File). The derivation of the model was based on the reference standard DD method, which provides accurate, safe and minimally invasive measurements of total body water (and FFM) with an error of <1%.48 49 While potential differences may occur in the assessment of total body water and hydration between ethnic groups, previous studies have suggested that the hydration of lean body mass is remarkably consistent between individuals50 and that ethnic variations in the hydration of lean body mass are small.51 Moreover, the predictive ability of the final model is strong across the whole study population and does not differ appreciably between ethnic groups. The final prediction models should therefore be widely applicable within the UK population and may also be applicable in a range of other populations, though separate validation studies will be needed before such application.

**Implications for clinicians and policymakers**

The availability of a prediction model which can accurately assess FM in UK children has important potential implications for practice and policy. The model could be used to assess FM in individual children as a guide to clinical management, particularly when used as a height-standardized indicator. A simple Excel calculator would allow simple calculation of FM from the relevant predictor variables. An early application could be in the interpretation of routine surveillance of childhood adiposity, particularly in the National Child Measurement Programme, in which all the parameters needed for the prediction model are routinely measured. This would allow geographic, ethnic, socioeconomic and temporal variations in FM to be directly assessed, rather than relying on weight-based measures which do not distinguish between FM and FFM.

**Further Research**

Future research will seek to obtain clear evidence of the benefits of this approach when compared with conventional weight-for-height measures. It will also require the documentation of normal ranges for the relevant FM parameters in different age and sex groups. It will also explore whether childhood BF is more strongly associated than BMI with adult health outcomes, particularly type 2 diabetes and cardiovascular disease incidence. Finally, for international applications of the models, further validation in a range of different populations is needed.

**SUMMARY**

**What is already known on this topic**

Body mass index (BMI), the most widely used marker of body fat, has serious limitations particularly in children.

As a weight-based measure, it does not discriminate between lean and fat mass which can vary markedly in individuals with a given BMI and may relate differently to cardiometabolic disease risk.

More accurate simple methods, based on routinely available measurements, are needed to improve the assessment of boy fatness in childhood.

**What this study adds**

A prediction model to estimate fat mass levels in children and adolescents has been developed and validated.

This approach allows for accurate discrimination of lean and fat mass in UK boys and girls aged 4-15 years.

The equation is based on readily available markers of height, weight, age, sex and ethnic group (where available), without the need for more costly forms of assessment.

**Conflict of Interests:** All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) and declare: no support from any organisation for the submitted work; 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.

**Acknowledgements:** We are extremely grateful to children who took part in the DD studies and all the staff involved in recruitment and data collection. We are also grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The authors thank Professor JJ Reilly for helpful advice.

**Funding and ethical approval:** This research was supported by grants from the British Heart Foundation (PG/15/19/31336 and FS/17/76/33286). Diabetes prevention research at St George's, University of London, is supported by the National Institute of Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care South London (NIHR CLAHRC-2013-10022). Dr Nightingale is supported by the Wellcome Trust Institutional Strategic Support Fund (204809/Z/16/Z) awarded to St George's, University of London. Data collection in the ABCC, ELBI, RC and SLIC studies was funded by the British Heart Foundation (PG/11/42/28895), the BUPA Foundation (TBF-S09-019), Child Growth Foundation (GR 10/03), the Wellcome Trust (WT094129MA) and the Medical Research Council. The UK Medical Research Council and Wellcome (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. Ethical approval for the four studies from which data were used to derive the model was obtained from the relevant ethics committees and for the ALSPAC study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. The views expressed in this paper are those of the authors and not necessarily those of the funding agencies, the National Health Service, the NIHR or the Department of Health.

**Author Contributions:** Study design – MTH, JCKW, RDR, CGO, DGC, ARR, PHW, CMN. Data collection - PHW, CMN, CGO, SL, JEW, DH, MSF, JCKW, ARR, DGC. Data analysis – MTH, RDR, ARR, DGC, CMN. Data interpretation – MTH, RDR, PHW, ARR, CGO, DGC, CMN. Drafting manuscript – MTH, PHW, RDR, ARR, CMN. Critical evaluation and revision of manuscript – MTH, MSF, DH, SL, JEW, JCKW, RDR, CGO, DGC, ARR, PHW, CMN. The corresponding author, CMN, attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. This publication is the work of the authors who will serve as guarantors for the contents of this paper. 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 have been explained.

Data sharing: For access to data from the studies used to derive and validate the model please contact the study PIs.

Publishing Licence: This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0.

**REFERENCES**

1. World Health Organisation. Childhood overweight and obesity 2014 [Available from: [www.who.int/dietphysicalactivity/childhood/en/](http://www.who.int/dietphysicalactivity/childhood/en/) accessed Accessed August 2016.

2. Bridges S, Darton R, Evans-Lacko S, et al. Health Survey for England 2014. In: Craig R, Fuller E, Mindell J, eds., 2015.

3. Cole TJ, Bellizzi MC, Flegal KM, et al. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;320:1240-43.

4. Herman KM, Craig CL, Gauvin L, et al. Tracking of obesity and physical activity from childhood to adulthood: the Physical Activity Longitudinal Study. *Int J Pediatr Obes* 2009;4:281-88.

5. Whitlock G, Lewington S, Sherliker P, et al. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet* 2009;373:1083-96.

6. Shaper AG, Wannamethee SG, Walker M. Body weight: implications for the prevention of coronary heart disease, stroke, and diabetes mellitus in a cohort study of middle aged men. *BMJ* 1997;314:1311-7.

7. Staimez LR, Weber MB, Narayan KM, et al. A systematic review of overweight, obesity, and type 2 diabetes among Asian American subgroups. *Current diabetes reviews* 2013;9:312-31.

8. Reilly JJ, Kelly J, Wilson DC. Accuracy of simple clinical and epidemiological definitions of childhood obesity: systematic review and evidence appraisal. *Obesity reviews* 2010;11:645-55.

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

10. Wells JC. A Hattori chart analysis of body mass index in infants and children. *Int J Obes Relat Metab Disord* 2000;24:325-29.

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

12. Chinn S, Rona RJ, Gulliford MC, et al. Weight-for-height in children aged 4-12 years. A new index compared to the normalized body mass index. *Eur J Clin Nutr* 1992;46:489-500.

13. Fung KP, Lee J, Lau SP, et al. Properties and clinical implications of body mass indices. *Arch Dis Child* 1990;65:516-9.

14. Whincup PH, Cook DG, Adshead F, et al. Childhood size is more strongly related than size at birth to glucose and insulin levels in 10-11-year-old children. *Diabetologia* 1997;40:319-26.

15. 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.

16. Nightingale CM, Rudnicka AR, Owen CG, et al. Patterns of body size and adiposity among UK children of South Asian, black African-Caribbean and white European origin: Child Heart And health Study in England (CHASE Study). *International Journal of Epidemiology* 2011;40:33-44.

17. Nightingale CM, Rudnicka AR, Owen CG, et al. Are ethnic and gender specific equations needed to derive fat free mass from bioelectrical impedance in children of South asian, black african-Caribbean and white European origin? Results of the assessment of body composition in children study. *PLoS One* 2013;8:e76426.

18. 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;92:872-75.

19. Freedman DS, Wang J, Thornton JC, et al. Racial/ethnic differences in body fatness among children and adolescents. *Obesity (Silver Spring)* 2008;16:1105-11.

20. Duncan JS, Duncan EK, Schofield G. Ethnic-specific body mass index cut-off points for overweight and obesity in girls. *The New Zealand medical journal* 2010;123:22-9.

21. Prentice AM, Jebb SA. Beyond body mass index. *Obesity reviews* 2001;2:141-7.

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

23. Sultan AA, West J, Grainge MJ, et al. Development and validation of risk prediction model for venous thromboembolism in postpartum women: multinational cohort study. *BMJ* 2016;355:i6253.

24. Haybittle JL, Blamey RW, Elston CW, et al. A prognostic index in primary breast cancer. *British journal of cancer* 1982;45:361-6.

25. Haroun D, Taylor SJ, Viner RM, et al. Validation of bioelectrical impedance analysis in adolescents across different ethnic groups. *Obesity (Silver Spring)* 2010;18:1252-59.

26. 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.

27. Lum S, Bountziouka V, Sonnappa S, et al. Lung function in children in relation to ethnicity, physique and socioeconomic factors. *The European respiratory journal* 2015;46:1662-71.

28. Boyd A, Golding J, Macleod J, et al. Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2013;42:111-27.

29. 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. *Int J Pediatr Obes* 2010;5:111-15.

30. Harrell FE, Lee KL, Califf RM, et al. Regression modelling strategies for improved prognostic prediction. *Statistics in Medicine* 1984;3:143-52.

31. Riley RD, Snell KIE, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: Part I - Continuous outcomes. *Stat Med* 2018

32. Moons KM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): Explanation and elaboration. *Annals of Internal Medicine* 2015;162:W1-W73.

33. Steyerberg EW. Validation in prediction research: the waste by data splitting. *J Clin Epidemiol* 2018;103:131-33.

34. Sauerbrei W, Royston P. Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. *Journal of the Royal Statistical Society Series A* 1999;162:71-94.

35. Royston P, Altman DG. Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. *Journal of the Royal Statistical Society Series C (Applied Statistics)* 1994;43:429-67.

36. Royston PA, G. sg81: Multivariable fractional polynomials. *Stata Technical Bulletin* 1998;43:24-32.

37. Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. *Statistics in Medicine* 1990;9:1303-25.

38. Royston P, Parmar MK, Sylvester R. Construction and validation of a prognostic model across several studies, with an application in superficial bladder cancer. *Stat Med* 2004;23:907-26.

39. Steyerberg EW, Harrell FE. Prediction models need appropriate internal, internal-external, and external validation. *Journal of clinical epidemiology* 2016;69:245-47.

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

41. 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.

42. 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.

43. 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.

44. 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.

45. 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 (2005)* 2017;41:1669-75.

46. 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.

47. Zeng Q, Dong SY, Sun XN, et al. Percent body fat is a better predictor of cardiovascular risk factors than body mass index. *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas* 2012;45:591-600.

48. 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.

49. 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.

50. Wang Z, Deurenberg P, Wang W, et al. Hydration of fat-free body mass: review and critique of a classic body-composition constant. *Am J Clin Nutr* 1999;69:833-41.

51. Deurenberg P, Deurenberg-Yap M. Validity of body composition methods across ethnic population groups. *Acta diabetologica* 2003;40 Suppl 1:S246-9.

**Development and validation of a prediction model for fat mass in children and adolescents: an individual participant data (IPD) meta-analysis**

**PRINT ABSTRACT**

**Study question:** To develop and validate a model to predict fat mass in children and adolescents aged 4-15 years using routinely available risk factors.

**Methods:** Four multi-ethnic cross-sectional studies with data on anthropometry and deuterium dilution assessments of fat mass were pooled. Multivariable linear regression analysis, using backwards selection for inclusion of predictor variables and non-linear relationships, was used to develop a prediction model for fat free mass (and indirectly fat mass). Internal validation and internal-external cross-validation was used to examine overfitting and generalisability of the model’s predictive performance within the four development studies. Finally, external validation was conducted on a fifth independent dataset.

**Study answer and limitations:** The final model, based on a population of 2,375 children (48% boys), containing predictor variables of height, weight, age, sex and ethnicity had extremely high predictive ability (R2adjusted: 94.8% [95%CI: 94.4 to 95.2]) with excellent calibration of observed and predicted values. Internal validation demonstrated minimal overfitting and good model generalisability, with excellent calibration and predictive performance. External validation in 176 individuals aged 11-12 years demonstrated promising model generalisability (R2: 90.0% [95%CI: 87.2 to 92.8]) with good calibration of observed and predicted FM (slope: 1.02 [95%CI: 0.97 to 1.07]). The mean difference between observed and predicted FM was -1.29 kg (95%CI: -1.62 to -0.96 kg). Study limitations include the limited age-range of the external dataset and the lack of other independent datasets with deuterium dilution assessments in children.

**What this study adds:** The developed model accurately predicted FM in children and adolescents. The model, based on simple anthropometric measures without the need for more costly forms of assessment, could improve the accuracy of childhood body fatness assessments (compared to those provided by body mass index) for effective clinical and public health obesity surveillance, prevention and management.

**Funding:** British Heart Foundation (PG/15/19/31336; FS/17/76/33286)

Figure to be used: Figure 1 from manuscript figures