Development and validation of a prediction model for fat mass in children and adolescents: meta-analysis using individual participant data

Mohammed T Hudda,[1] Mary S Fewtrell,[2] Dalia Haroun,[3] Sooky Lum,[4] Jane E Williams,[2] Jonathan C K Wells,[2] Richard D Riley,[5] Christopher G Owen,[1] Derek G Cook,[1] Alicja R Rudnicka,[1] Peter H Whincup,[1] Claire M Nightingale[1]

[1]Population Health Research Institute, St George's, University of London, London SW17 0RE, 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 and 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:
C M Nightingale
cnightin@sgul.ac.uk
(or @mohammedhudda on Twitter ORCID 0000-0002-4803-7617)

ABSTRACT

OBJECTIVES
To develop and validate a prediction model for fat mass in children 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
Individual participant data meta-analysis.

SETTING
Four population based cross sectional studies and a fifth study for external validation, United Kingdom.

PARTICIPANTS
A pooled derivation dataset (four studies) of 2375 children and an external validation dataset of 176 children with complete data on anthropometric measurements and deuterium dilution assessments of fat mass.

MAIN OUTCOME MEASURE
Multivariable linear regression analysis, using backwards selection for inclusion of predictor variables and allowing non-linear relations, 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 were used to examine overfitting and generalisability of the model's predictive performance within the four development studies; external validation followed using the fifth dataset.

RESULTS
Model derivation was based on a multi-ethnic population of 2375 children (47.8% boys, n=1136) aged 4-15 years. The final model containing predictor variables of height, weight, age, sex, and ethnicity had extremely high predictive ability (optimism adjusted R^2: 94.8%, 95% confidence interval 94.4% to 95.2%) with excellent calibration of observed and predicted values. The internal validation showed minimal overfitting and good model generalisability, with excellent calibration and predictive performance. External validation in 176 children aged 11-12 years showed promising generalisability of the model (R^2: 90.0%, 95% confidence interval 87.2% to 92.8%) with good calibration of observed and predicted fat mass (slope: 1.02, 95% confidence interval 0.97 to 1.07). The mean difference between observed and predicted fat mass was −1.29 kg (95% confidence interval −1.62 to −0.96 kg).

CONCLUSION
The developed model accurately predicted levels of fat mass in children aged 4-15 years. The prediction model is based on simple anthropometric measures without the need for more complex forms of assessment and could improve the accuracy of assessments for body fatness in children (compared with those provided by body mass index) for effective surveillance, prevention, and management of clinical and public health obesity.

Introduction

With the increasing prevalence of obesity in children globally,[1] such as in the United Kingdom, where about one third of children aged 2-15 years are overweight or obese,[2] high body fatness in childhood represents a serious public health problem. High levels of body fatness in childhood have been associated with both overweight and obesity and increased risks of non-communicable diseases in adulthood—notably type 2 diabetes and cardiovascular diseases.[3-7]

Accurate and practical methods for quantifying body fatness in children are essential for effective monitoring, prevention, and management of high body fatness, overweight, and obesity in childhood.[8 9] Body mass index (BMI), the most widely used marker of childhood body fatness in clinical and public health practice, has serious limitations as a marker of body fatness in children.[9-11] Firstly, as a weight based measure, it does not discriminate between lean (fat-free mass) and fat mass, which can vary

WHAT IS ALREADY KNOWN ON THIS TOPIC

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

As a weight based measure, BMI does not discriminate between lean and fat mass, which can vary greatly in those with a given BMI and might relate differently to risk of cardiometabolic disease

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

WHAT THIS STUDY ADDS

A newly developed and validated prediction model to estimate fat mass levels in UK children aged 4-15 years allows for accurate discrimination of lean and fat mass

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

substantially in those with a given BMI.[10] Secondly, height squared provides poor height standardisation of weight in children—a higher power is needed to obtain height standardisation.[12-14] Finally, BMI in childhood is not a consistent marker of body fatness across different ethnic groups. In the UK and the United States, BMI has been shown to overestimate body fatness in black African children and underestimate body fatness in children of Asian origin.[15-19] Similar problems have been reported in other settings; BMI under estimates body fatness in South Asian girls and over estimates body fatness in Pacific Island girls in New Zealand.[20]

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

We examined whether weight and height as opposed to BMI could provide more accurate assessments of fat mass, particularly using prediction methods that have shown promise in estimating disease risks.[22-24]

We report on the development and validation of a prediction model to estimate fat mass accurately in UK children aged 4-15 years 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 body fatness assessments based on the deuterium dilution method.

Derivation data
Data from four separate cross sectional studies[17 25-27] (supplementary table 1), identified as the four available UK population based studies, which contained deuterium dilution measurements together with weight and height measurements in more than 200 children aged 4-15 years, were obtained and pooled for analysis (n=2375). Each of these studies used a similar protocol when conducting the deuterium dilution method to measure total body water (and indirectly fat mass), as described elsewhere.[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] and child ethnicity,[17 26 27] with self reported participant information on ethnicity for older children.[25 26] Ethnic group categories were based on the 2001 UK census (supplementary table 1).

External validation data
Data from a smaller separate UK 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 information on height, weight, sex, ethnicity, and age. At the 11 year follow-up visit, a subsample of the cohort (stratified by sex and BMI to represent the whole cohort) was recruited to participate in a further study that involved assessment of fat mass using the deuterium dilution 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
Our primary aim was to develop a model for predicting fat mass in childhood, which could be estimated directly or indirectly (by predicting fat-free mass from models and then subtracting resulting estimates from known weight) based on deuterium dilution measurements. Firstly, we investigated the potential for modelling fat mass directly or indirectly by examining the distributions of fat mass and fat-free mass in relation to height (one of the strongest predictors of body composition) in boys and girls separately. This showed that a regression model for fat-free mass better met the assumptions of linear regression (more details in Appendix 1). The distribution of fat-free mass (both in boys and girls separately and combined) was positively skewed (supplementary figure 1) and showed increased heterogeneity with increases in height and weight. Fat-free mass, transformed using natural logarithms, was therefore the outcome in the main analyses.

Candidate predictors
In the model development stage, we considered weight, height, age, sex, and ethnic group as candidate predictors (variables). Our derivation data, once restricted to those with fat-free mass or fat mass assessment, had no missing data on any of the candidate predictors. The sample size of 2375 participants meant that the number of candidate predictors being considered (along with non-linear terms) far exceeded both the minimum 10 people per candidate predictor rule of thumb[30] and the minimum sample size requirements for prediction models proposed elsewhere.[31] Ethnicity was based on self reported parental information on parental ethnicity. For the present analyses, we categorised child ethnicity as white (European origin), black (black children of African and Caribbean descent), South Asian (children of Indian, Pakistani, Bangladeshi, and Sri Lankan descent), other Asian (predominantly East Asian origins), and other (predominantly mixed ethnicity) groups.

Statistical analysis for model development

Stata v14 was used for all analyses. We followed the TRIPOD (transparent reporting of a multivariable model for individual prognosis or diagnosis) guidance for development and reporting of multivariable prediction models.[32] To avoid data splitting we used all four available studies for model development.[33] A linear regression was used with the natural logarithm of fat-free mass as the outcome, and weight, height, age, sex, and ethnic group as candidate predictors (variables). Using a stepwise approach through backwards elimination, beginning with a model that included all predictors, we excluded candidate predictors from the saturated model based on their statistical significance (Wald test P>0.05). Non-linear relations between outcome and continuous predictors were considered by identifying, at each iterative step of the stepwise process, the best fitting fractional polynomial terms[34 35] (using Stata command mfp[36]). This model development process led to a final model for the prediction of natural logarithm of fat-free mass (and subsequently for fat mass=weight−exp(prediction of natural logarithm of fat-free mass)) based on the selected predictors along with their corresponding estimated β coefficients and the associated intercept term. Although heterogeneity and clustering of patients across or within studies was not considered for model development, we checked the impact of this using an internal-external validation approach.

Model performance and internal validation

The performance of the final model was assessed using several approaches:

- R^2—proportion of the variance in natural logarithm of fat-free mass explained by the included predictors
- Root mean square error (RMSE)—the average difference between the predicted and observed values. The RMSE of fat mass predictions was also assessed overall and within subgroups for age, ethnicity, and sex
- Calibration slope—based on model regressing observed on predicted values of natural logarithm of fat-free mass (with a slope of 1 being ideal)
- Calibration-in-the-large—intercept term from the model regressing observed on predicted values of natural logarithm of fat-free mass (with an intercept of 0 being ideal)
- Comparing mean observed with mean predicted values of natural logarithm of fat-free mass.

Calibration was also assessed graphically by displaying fat-free mass and fat mass on a calibration plot with a local regression (loess) smoother fitted across all children

We carried out internal validation to estimate optimism (the level of model overfitting)[32] and correct measures of predictive performance (R^2, calibration slope, and calibration-in-the-large) for model overfitting by bootstrapping[32] 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. We then applied each of these bootstrap sample models within the original dataset to estimate optimism in the performance statistics (difference in test performance and apparent bootstrap performance) of R^2, calibration slope, and calibration-in-the-large (see Appendix 2 for further details), referred to as adjusted R^2, adjusted calibration slope, and adjusted calibration-in-the-large, respectively. To adjust for optimism after model development, we obtained estimates of a uniform shrinkage factor (the average calibration slope from each of the bootstrap samples) and multiplied these by the original β coefficients to obtain optimism adjusted coefficients.[32 37] At this stage, we re-estimated the intercept of the model 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 process discussed. Owing to the limited availability of appropriate external datasets, we conducted internal-external validation[38 39] to further assess the performance of the derived model. This internal-external approach[38 39] involved cross validation, omitting one of each of the four studies in turn from the development dataset, and developing a model within the remaining three datasets. The following three steps were undertaken: (1) using the same model development strategy, we developed a model on three of the four studies and obtained the β coefficients from the model predicting natural logarithm of fat-free mass; (2) the predictive performance of the model from the first step was then assessed (overall and within sex and ethnic groups) within the fourth external validation study data in terms of accuracy of predicted fat mass (the primary outcome) by means of the calibration slope, calibration-in-the-large, and the R^2 measures; and (3) we repeated the first two steps until we had assessed external validation for each of the four studies.

We assessed overfitting in each round of the cross validation and obtained a uniform shrinkage factor,[37] which was applied to the β coefficients from step 1. Calibration slope, calibration-in-the-large, and R^2 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 (with the τ^2 statistic estimated using the Mantel-Haenszel method). The variance of R^2 was estimated using the Wald type method outlined previously[40] and used to pool the values.

External validation

We applied our final prediction model to each participant in the external validation dataset based on

his or her respective predictor values. In a small number of children with missing ethnicity, we reclassified missing ethnicity data as white to produce an estimate of fat mass. The performance of the model for predicting fat mass, by sex and overall, was assessed using the calibration slope, calibration-in-the-large, R^2, and RMSE and by comparing mean observed values with mean predicted values. We also assessed the overall calibration of the model graphically in terms of fat mass by plotting agreement between predicted and observed values across 10ths of predicted values. Finally, we re-estimated the intercept term from the final model for the external data to maintain the calibration of the model and reassessed the performance statistics.

Patient and public 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

The pooled derivation dataset (four studies) included 2375 children, predominantly of white (37.3%, n=885), black (23.3%, n=553), and South Asian (24.7%, n=586) ethnicity, aged 4.0-15.9 years (median age 9.6 years, 47.8% (n=1136) boys) with complete information on anthropometric, demographic, and body fatness measurements (table 1). The external validation dataset included 176 children predominantly of white ethnicity and aged 11-12 years (47.7% boys, n=84), with complete data on anthropometric and body fatness measurements and missing data on ethnicity in a 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 children of only white ethnic origin), whereas

the other three studies each contained a restricted age range, but with noticeable ethnic diversity.

Model development and apparent performance

The final multivariable model included all five candidate predictors of height, weight, age, sex, and ethnic group (ie, 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 relations (table 2). The model showed excellent apparent predictive performance for natural logarithm of fat-free mass (table 3; R^2=94.8%, RMSE=0.068) and was perfectly calibrated in the development data (apparent slope=1, apparent calibration-in-the-large=0). This is confirmed by the calibration plot, assessing agreement between observed and predicted fat-free mass and fat mass (fig 1). The difference between the mean observed and mean predicted values of natural logarithm of fat-free mass was zero. The RMSE values for fat mass were 2.0 kg in girls and 1.9 kg in boys 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 among South Asian children and 2.4 kg among black children.

Model validation
Internal validation

Bootstrap internal validation showed little model overfitting, which was reflected in the similar apparent and optimism adjusted performance statistics (table 3). After we had adjusted for overfitting, the final prediction model maintained a high proportion of the variance in natural logarithm of fat-free mass with an adjusted R^2 value of 94.8%. The bootstrapping approach provided a shrinkage factor of practically 1 (ie, 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 previously,[37] and this gave a value of 0.99858, again close to 1. We chose to use this method because it was slightly smaller than the bootstrap value, which was applied

Table 1 | Characteristics of participants within derivation and validation datasets. Values are median (interquartile range) unless stated otherwise

Characteristics	Derivation dataset*					Validation dataset (ALSPAC subsample) (n=176)
	ABCC (n=1027)	ELBI (n=382)	RC (n=369)	SLIC (n=597)	Overall (n=2375)	
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 (m)	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)
Boys (No (%))	490 (47.7)	182 (48)	180 (49)	284 (47.6)	1136 (47.8)	84 (48)
Ethnic group (No (%))‡:						
White	290 (28.2)	91 (24)	369 (100)	135 (22.6)	885 (37.3)	161 (91)
Black	252 (24.5)	119 (31)	0 (0)	182 (30.5)	553 (23.3)	-
South Asian	325 (31.6)	120 (31)	0 (0)	141 (23.6)	586 (24.7)	-
Other Asian	46 (4.5)	44 (12)	0 (0)	22 (3.7)	112 (4.7)	-
Other/missing	114 (11.1)	8 (2)	0 (0)	117 (19.6)	239 (10.1)	15 (9)

ABCC=Assessment of Body Composition in Children study; ELBI=East London Bioelectrical Impedance; RC=Reference Child; SLIC=Size and Lung function in Children study; ALSPAC=Avon Longitudinal Study of Parents and Children.
*No missing data.
†Assessed using deuterium dilution method.
‡Information on ethnic group was missing on a small number of children in the validation dataset.

Table 2 | Final multivariable analysis model in derivation dataset and optimism adjusted β coefficients

Variable	Developed model: coefficients (95% CI)	Final model coefficients after adjusting for overfitting
Height^2 (m)	0.308 (0.289 to 0.327)	0.307
(Weight/10)$^{-1}$ (kg)	−1.003 (−1.090 to −0.916)	−1.002
Weight/10 (kg)	0.046 (0.040 to 0.052)	0.046
Ethnicity:		
White	Reference	Reference
Black	0.014 (0.007 to 0.022)	0.014
South Asian	−0.065 (−0.072 to −0.058)	−0.065
Other Asian	−0.026 (−0.040 to −0.013)	−0.026
Other	−0.017 (−0.027 to −0.008)	−0.017
ln(age/10) (years)	−0.919 (−1.086 to −0.753)	−0.918
(Age/10)$^{0.5}$ (years)	2.055 (1.708 to 2.401)	2.052
Sex:		
Girls	Reference	Reference
Boys	0.047 (0.042 to 0.053)	0.047
Constant*	0.692 (0.373 to 1.011)	0.691

ln=natural logarithmic transformation.
Outcome of model was ln(fat-free mass).
*Constant term was re-estimated after adjustment for optimism (shrinkage factor=0.99858) to uphold overall model calibration.

to the original β 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 fat mass in children aged 4-15 years, with examples of how to calculate fat mass using the equation.

Internal-external validation
Using the cross validation approach, we developed a model in each of the three studies and applied this within the fourth study. Assessments of model overfitting showed 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 close to 1 and the calibration-in-the-large values were close to 0, suggesting excellent model calibration in each of these four study populations (fig 2, supplementary table 2).

The pooled calibration slopes and calibration-in-the-large values across the four studies for fat mass were 1.00 (95% confidence interval 0.95 to 1.04) and −0.29 (−0.83 to 0.25), respectively, suggesting that, on average across the four populations, the model is likely to calibrate well. The pooled R^2 value for fat mass was 89.7% (95% confidence interval 87.8% to 91.7%), which indicates that the model, on average, explains a high proportion of the variance in fat mass. The τ^2 values for the calibration slope, calibration-in-the-large, and R^2 measures were 0.002, 0.267, and 0.0004, respectively, suggesting little heterogeneity across the four populations. The calibration slopes and calibration-in-the-large values within sex and ethnic groups showed good calibration for all subgroups during 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 and 3).

External validation
We applied our final prediction model (box 1) to the independent population of children aged 11-12 years, reclassifying the small number of children with

missing information on ethnicity as being from the white reference group. The resulting R^2 value from the model was 90.0% (95% confidence interval 87.2% to 92.8%), with a moderate RMSE of 2.6 kg, and the model had average calibration in terms of fat mass (fig 3); with a slope of 1.02 (95% confidence interval 0.97 to 1.07) and calibration-in-the-large of −1.58 kg (95% confidence interval −2.29 to −0.86 kg) (table 4). The mean difference between observed and predicted fat mass was −1.29 kg (95% confidence interval −1.62 to −0.96 kg). The final model was observed to perform better in girls than in boys (table 4). After recalibration of the intercept, the R^2 value from the model was 90.0% (95% confidence interval 87.1% to 92.8%), with a RMSE of 2.4 kg, and the model had a calibration slope of 1.06 (95% confidence interval 1.01 to 1.11) and calibration-in-the-large of 0.21 kg (95% confidence interval −0.42 to 0.85 kg).

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 (R^2=94.9%, RMSE=0.068), with little difference between the Akaike's Information Criterion (compares the relative quality of a set of statistical models for a given dataset) from models including and excluding these terms. Therefore, these interaction terms were not added to the previously described prediction model.

We also used two approaches to investigate the use of the proposed model to estimate fat mass in childhood when ethnic origins were unknown—omitting ethnic group as a predictor from the model, and treating children of unknown ethnic origin as being white (reference group) for fat mass predictions. Both approaches were carried out and compared using an internal-external validation approach. Fat mass predictions from both approaches had similar levels
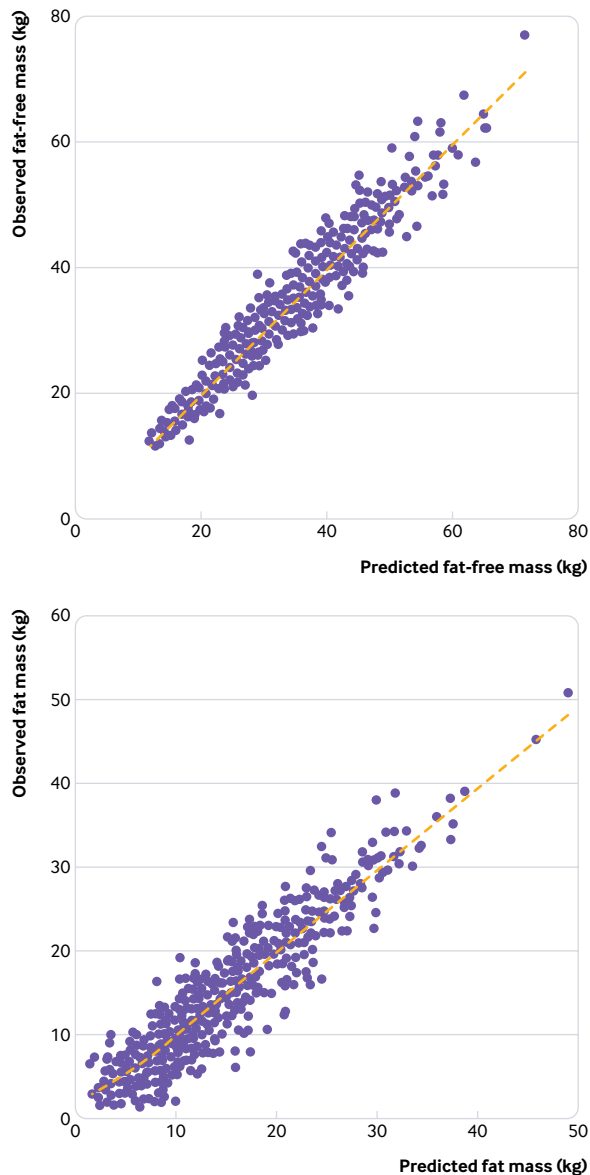
Fig 1 | Assessment of model calibration for fat-free mass and fat mass. The developed model for predicting natural logarithm of fat-free mass was used to derive estimates of fat-free mass and fat mass, which were each used to assess model calibration. Broken orange line represents a lowess smoother through the data points, showing a linear relation between observed and predicted values of both fat-free mass and fat mass

performance of this model (R^2=83.4%) was much less favourable than the performances of the main analyses using the natural logarithm of fat-free mass as the primary outcome.

Discussion

We developed a new prediction equation, based on readily available measures of height, weight, age, sex, and ethnic group, to estimate fat mass levels (kg) for children aged 4-15 years using a large representative sample from the UK. We then validated the model both internally and externally—firstly using a cross validation approach within the derivation population and then within an independent dataset of children aged 11-12 years. Both overall and within age, sex, and ethnic subgroups, the developed model showed high predictive ability, with excellent calibration; low individual error, with root mean square error (RMSE) values less than 3.3 kg; and useful R^2 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 2.6 kg.

Comparison with other studies

To our knowledge, few previous studies have developed and validated prediction models to estimate fat mass in children and adolescents based solely on weight, height, and demographic factors.[41] Most previously derived models for this purpose have focused on older children and adolescents from the United States with body fatness assessed using dual energy x ray absorptiometry.[41-46] Moreover, modelling has predominantly been based on the prediction of percentages and not absolute values of body fatness, making it difficult to compare the predictive ability of models. The developed models, which have been shown to estimate the percentage of body fatness to a high level, with R^2 values greater than 82%, have relied on additional measurements, including skinfold thickness, waist circumference, or bioelectrical impedance to estimate body fatness.[42-46] However, a previously developed model in 12-20 year-olds in the US included the same predictors as in our final model, of height and weight (in the form of a fractional polynomial non-linear term of body mass index, BMI) as well as sex, age, and ethnicity to estimate the percentage of body fatness.[41] That model performed well, explaining a high proportion of the variance in body fatness percentage (R^2=79.4%). The RMSE was not presented, however, making direct comparisons of accuracy between the models difficult.

of bias when compared with observed fat mass 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 fat mass, we repeated the model development strategy using the natural logarithm of fat-free mass as the primary outcome. The apparent

Table 3 | Model performance statistics based on internal validation

Measure	Apparent performance	Average optimism	Optimism corrected
R^2 (%)	94.83 (94.43 to 95.23)	0.03	94.80
Calibration slope (95% CI)	1.00 (0.99 to 1.01)	1.63×10^{-9}	1.00 (0.99 to 1.01)
Calibration-in-the-large (95% CI) (kg)	0.00 (−0.03 to 0.03)	-3.54×10^{-9}	0.00 (−0.03 to 0.03)

Outcome of model was natural logarithm of fat-free mass.

Box 1: Final equation for prediction of fat mass in children aged 4-15 years

Fat mass=weight−exp[0.3073×height^2−10.0155×weight^{-1}+0.004571×weight+0.01408×BA−0.06509×SA−0.02624×AO−0.01745×other−0.9180×ln(age)+0.6488×age$^{0.5}$+0.04723×male+2.8055]

- exp=exponential function, ln=natural logarithmic transformation
- Score 1 if child is of black (BA), south Asian (SA), other Asian (AO), or other (other) ethnic origins and score 0 if not
- If child is of unknown ethnic group, treat as of white ethnic origins
- Height is measured in metres, weight in kilograms, age in years, and fat mass in kilograms
- *Example 1*
- For a 6 year old white boy of height 1.4 m and weight 37 kg, fat mass would be estimated as:
- =37−exp[0.3073×1.4^2−10.0155×37^{-1}+0.004571×37+0.01408×0−0.06509×0−0.02624×0−0.01745×0−0.9180×ln(6)+0.6488×6$^{0.5}$+0.04723×1+2.8055=37−exp[3.2979]=37−27.0549=9.95 kg
- *Example 2*
- For a 12 year old black girl with a height of 1.6 m and a weight of 42 kg, fat mass would be estimated as:
- =42−exp[0.3073×1.6^2−10.0155×42^{-1}+0.004571×42+0.01408×1−0.06509×0−0.02624×0−0.01745×0−0.9180×ln(12)+0.6488×12$^{0.5}$+0.04723×0+2.8055
- =42−exp[3.5262]=42−33.9929=8.01 kg

Strengths and limitations of this study

This study has several strengths. The derivation dataset was sufficiently large, with complete information on candidate predictors for children with information on fat-free mass, allowing all of the candidate predictors to be tested along with their respective non-linear terms while adhering to the 10 people 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 during 2009-13 and should have continuing relevance, with no indication that the associations between fat-free mass and its predictors have changed. We were able to identify an additional independent dataset for external validation, although with a narrower range of age and ethnicity. The model is based on simple and already widely measured predictors. The performance of the model is strong and allows discrimination between fat mass and fat-free mass both in the whole study population and in specific ethnic groups, offering potential advantages over the ethnic specific BMI adjustments that we reported previously,[15] particularly if earlier reports suggested that fat mass is more strongly associated with long term health outcomes than is BMI.[47] Although the inclusion of non-linear polynomial terms makes the derived algorithm appear complicated for practical use, these terms have been integrated into a simple MS Excel calculator (supplementary file). The derivation of the model was based on the reference standard deuterium dilution method, which provides accurate, safe, and minimally invasive measurements of total body water (and fat-free mass) with an error of less than 1%.[48 49] Although potential differences might 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 highly consistent between people[50] 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 might also be
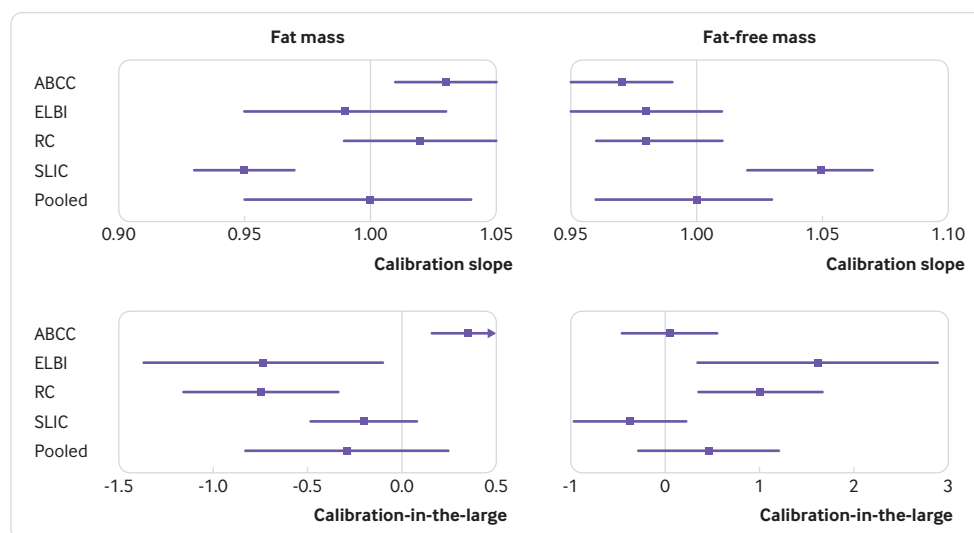

Fig 2 | Assessment of calibration slope and calibration-in-the-large for fat mass and fat-free mass from internal-external cross validation. Calibration slopes and calibration-in-the-large (and respective 95% confidence intervals) were obtained by fitting the final model in three studies and assessing the external validity in terms of the slope and intercept for fat-free mass and fat mass in the data from the fourth study. This was repeated until each of the four studies had been used as a validation dataset. A random effects meta-analysis was used to obtain the pooled estimates (95% confidence intervals) along with the τ^2 statistic for heterogeneity. Also see supplementary table 2 for data in tabular form. ABCC=Assessment of Body Composition in Children study; ELBI=East London Bioelectrical Impedance, RC=Reference Child, SLIC=Size and Lung function in Children study
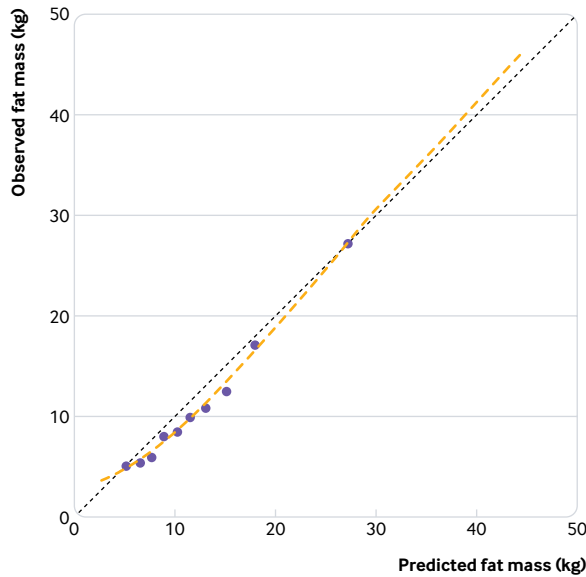
Fig 3 | Calibration plot of mean observed against mean predicted values, across 10ths of predicted fat mass, from external validation before recalibration of intercept. Data points are mean predicted against mean observed fat mass within 10ths of predicted fat mass. Individual level data points not shown for confidentiality reasons. Broken orange line represents a local regression smoother through individual level data points. Broken black line represents line of equality

applicable in a range of other populations, although separate validation studies will be needed before such application.

Implications for clinicians and policymakers

The availability of a prediction model that can accurately assess fat mass in UK children has important potential implications for practice and policy. The model could be used to assess fat mass in individual children as a guide to clinical management, particularly when used as a height standardised indicator. The Excel calculator (supplementary file) would allow simple calculation of fat mass from the relevant predictor variables. An early application could be in the interpretation of routine surveillance of adiposity in children, particularly in the National Child Measurement Programme, in which all the parameters needed for the prediction model are routinely measured. This would allow direct assessment of geographical, ethnic, socioeconomic, and temporal variations in fat mass rather than reliance on weight based measures, which do not distinguish between fat mass and fat-free mass.

Further research

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

Table 4 | External validation: model performance* statistics before recalibration of intercept in children aged 11-12 years

Measure	Boys	Girls	Overall
R^2 (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) (kg)	−1.38 (−2.41 to 0.36)	−2.25 (−3.27 to −1.23)	−1.58 (−2.29 to −0.86)

Statistics presented before intercept term were re-estimated for external data.
*Fat mass predictions.

1 World Health Organization. Childhood overweight and obesity 2014 www.who.int/dietphysicalactivity/childhood/en/ 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, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ 2000;320:1240-3. doi:10.1136/bmj.320.7244.1240
4 Herman KM, Craig CL, Gauvin L, Katzmarzyk PT. Tracking of obesity and physical activity from childhood to adulthood: the Physical Activity Longitudinal Study. Int J Pediatr Obes 2009;4:281-8. doi:10.3109/17477160802596171
5 Whitlock G, Lewington S, Sherliker PProspective Studies Collaboration. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. Lancet 2009;373:1083-96, et al. doi:10.1016/S0140-6736(09)60318-4
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. doi:10.1136/bmj.314.7090.1311
7 Staimez LR, Weber MB, Narayan KM, Oza-Frank R. A systematic review of overweight, obesity, and type 2 diabetes among Asian American subgroups. Curr Diabetes Rev 2013;9:312-31. doi:10.2174/15733998113099990061
8 Reilly JJ, Kelly J, Wilson DC. Accuracy of simple clinical and epidemiological definitions of childhood obesity: systematic review and evidence appraisal. Obes Rev 2010;11:645-55. doi:10.1111/j.1467-789X.2009.00709.x
9 Hall DM, Cole TJ. What use is the BMI?Arch Dis Child 2006;91:283-6. doi:10.1136/adc.2005.077339
10 Wells JC. A Hattori chart analysis of body mass index in infants and children. Int J Obes Relat Metab Disord 2000;24:325-9. doi:10.1038/sj.ijo.0801132
11 Wells JC, Fewtrell MS. Measuring body composition. Arch Dis Child 2006;91:612-7. doi:10.1136/adc.2005.085522
12 Chinn S, Rona RJ, Gulliford MC, Hammond J. 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, Chow OK, Wong TW, Davis DP. Properties and clinical implications of body mass indices. Arch Dis Child 1990;65:516-9. doi:10.1136/adc.65.5.516
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. doi:10.1007/s001250050681
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. Int J Obes (Lond) 2017;41:1048-55. doi:10.1038/ijo.2017.75
16 Nightingale CM, Rudnicka AR, Owen CG, Cook DG, Whincup PH. 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). Int J Epidemiol 2011;40:33-44. doi:10.1093/ije/dyq180
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. doi:10.1371/journal.pone.0076426
18 Shaw NJ, Crabtree NJ, Kibirige MS, Fordham JN. Ethnic and gender differences in body fat in British schoolchildren as measured by DXA. Arch Dis Child 2007;92:872-5. doi:10.1136/adc.2007.117911
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. doi:10.1038/oby.2008.30
20 Duncan JS, Duncan EK, Schofield G. Ethnic-specific body mass index cut-off points for overweight and obesity in girls. N Z Med J 2010;123:22-9.
21 Prentice AM, Jebb SA. Beyond body mass index. Obes Rev 2001;2:141-7. doi:10.1046/j.1467-789x.2001.00031.x
22 Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837-47. doi:10.1161/01.ClR.97.18.1837
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. doi:10.1136/bmj.i6253
24 Haybittle JL, Blamey RW, Elston CW, et al. A prognostic index in primary breast cancer. Br J Cancer 1982;45:361-6. doi:10.1038/bjc.1982.62
25 Haroun D, Taylor SJ, Viner RM, et al. Validation of bioelectrical impedance analysis in adolescents across different ethnic groups. Obesity (Silver Spring) 2010;18:1252-9. doi:10.1038/oby.2009.344
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. doi:10.3945/ajcn.112.036970
27 Lum S, Bountziouka V, Sonnappa S, et al. Lung function in children in relation to ethnicity, physique and socioeconomic factors. Eur Respir J 2015;46:1662-71. doi:10.1183/13993003.00415-2015
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. doi:10.1093/ije/dys064
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;5:111-5. doi:10.3109/17477160903060010
30 Harrell FEJr, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. Stat Med 1984;3:143-52. doi:10.1002/sim.4780030207
31 Riley RD, Snell KIE, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: Part I - Continuous outcomes. Stat Med 2019;38:1262-75. doi:10.1002/sim.7993
32 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. doi:10.7326/M14-0698
33 Steyerberg EW. Validation in prediction research: the waste by data splitting. J Clin Epidemiol 2018;103:131-3. doi:10.1016/j.jclinepi.2018.07.010
34 Sauerbrei W, Royston P. Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. J R Stat Soc [Ser A] 1999;162:71-94. doi:10.1111/1467-985X.00122 .
35 Royston P, Altman DG. Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. J R Stat Soc Ser C Appl Stat 1994;43:429-67.
36 Royston PA. G. sg81: Multivariable fractional polynomials. Stata Tech Bull 1998;43:24-32.
37 Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. Stat Med 1990;9:1303-25. doi:10.1002/sim.4780091109
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. doi:10.1002/sim.1691
39 Steyerberg EW, Harrell FEJr. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol 2016;69:245-7. doi:10.1016/j.jclinepi.2015.04.005
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, Durazo-Arvizu RA. Adiposity is not equal in a multi-race/ethnic adolescent population: NHANES 1999-2004. Obesity (Silver Spring) 2011;19:2099-101. doi:10.1038/oby.2011.52
42 Lohman TG, Caballero B, Himes JH, et al. Estimation of body fat from anthropometry and bioelectrical impedance in Native American children. Int J Obes Relat Metab Disord 2000;24:982-8. doi:10.1038/sj.ijo.0801318
43 Stevens J, Cai J, Truesdale KP, Cuttler L, Robinson TN, Roberts AL. Percent body fat prediction equations for 8- to 17-year-old American children. Pediatr Obes 2014;9:260-71. doi:10.1111/j.2047-6310.2013.00175.x
44 Stevens J, Ou FS, Cai J, Heymsfield SB, Truesdale KP. Prediction of percent body fat measurements in Americans 8 years and older. Int J Obes (Lond) 2016;40:587-94. doi:10.1038/ijo.2015.231
45 Stevens J, Truesdale KP, Cai J, Ou FS, Reynolds KR, Heymsfield SB. Nationally representative equations that include resistance and reactance for the prediction of percent body fat in Americans. Int J Obes (Lond) 2017;41:1669-75. doi:10.1038/ijo.2017.167
46 Truesdale KP, Roberts A, Cai J, Berge JM, Stevens J. Comparison of Eight Equations That Predict Percent Body Fat Using Skinfolds in American Youth. Child Obes 2016;12:314-23. doi:10.1089/chi.2015.0020

47   Zeng Q, Dong SY, Sun XN, Xie J, Cui Y. Percent body fat is a better predictor of cardiovascular risk factors than body mass index. *Braz J Med Biol Res* 2012;45:591-600. doi:10.1590/S0100-879X2012007500059

48   Wells JC, Fuller NJ, Dewit O, Fewtrell MS, Elia M, Cole TJ. 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. doi:10.1093/ajcn/69.5.904

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. doi:10.1053/beem.1999.0003

50   Wang Z, Deurenberg P, Wang W, Pietrobelli A, Baumgartner RN, Heymsfield SB. Hydration of fat-free body mass: review and critique of a classic body-composition constant. *Am J Clin Nutr* 1999;69:833-41. doi:10.1093/ajcn/69.5.833

51   Deurenberg P, Deurenberg-Yap M. Validity of body composition methods across ethnic population groups. *Acta Diabetol* 2003;40(Suppl 1):S246-9. doi:10.1007/s00592-003-0077-z