External validation of a prediction model for estimating fat mass in children and adolescents in 19 countries: individual participant data meta-analysis

Mohammed T Hudda,[1] Jonathan C K Wells,[2] Linda S Adair,[3] Jose R A Alvero-Cruz,[4] Maxine N Ashby-Thompson,[5] Martha N Ballesteros-Vásquez,[6] Jesus Barrera-Exposito,[7] Benjamin Caballero,[8] Elvis A Carnero,[9] Geoff J Cleghorn,[10] Peter S W Davies,[10] Malgorzata Desmond,[2] Delan Devakumar,[11] Dympna Gallagher,[12] Elvia V Guerrero-Alcocer,[13] Ferdinand Haschke,[14] Mary Horlick,[15] Houda Ben Jemaa,[16] Ashraful I Khan,[17] Amani Mankai,[16] Makama A Monyeki,[18] Hilde L Nashandi,[19] Luis Ortiz-Hernandez,[20] Guy Plasqui,[21] Felipe F Reichert,[22] Alma E Robles-Sardin,[6] Elaine Rush,[23] Roman J Shypailo,[24] Jakub G Sobiecki,[25] Gill A ten Hoor,[26] Jesús Valdés,[27] V Pujitha Wickramasinghe,[28] William W Wong,[24] Richard D Riley,[29] Christopher G Owen,[1] Peter H Whincup,[1] Claire M Nightingale[1]

For numbered affiliations see end of the article

Correspondence to: M T Hudda
mhudda@sgul.ac.uk
(@mohammedhudda on Twitter:
ORCID 0000-0001-7894-1159)
Additional material is published online only. To view please visit the journal online.

ABSTRACT

OBJECTIVE
To evaluate the performance of a UK based prediction model for estimating fat-free mass (and indirectly fat mass) in children and adolescents in non-UK settings.

DESIGN
Individual participant data meta-analysis.

SETTING
19 countries.

PARTICIPANTS
5693 children and adolescents (49.7% boys) aged 4 to 15 years with complete data on the predictors included in the UK based model (weight, height, age, sex, and ethnicity) and on the independently assessed outcome measure (fat-free mass determined by deuterium dilution assessment).

MAIN OUTCOME MEASURES
The outcome of the UK based prediction model was natural log transformed fat-free mass (lnFFM). Predictive performance statistics of R^2, calibration slope, calibration-in-the-large, and root mean square error were assessed in each of the 19 countries and then pooled through random effects meta-analysis. Calibration plots were also derived for each country, including flexible calibration curves.

RESULTS
The model showed good predictive ability in non-UK populations of children and adolescents, providing R^2 values of >75% in all countries and >90% in 11 of the 19 countries, and with good calibration (ie, agreement) of observed and predicted values. Root mean square error values (on fat-free mass scale) were <4 kg in 17 of the 19 settings. Pooled values (95% confidence intervals) of R^2, calibration slope, and calibration-in-the-large were 88.7% (85.9% to 91.4%), 0.98 (0.97 to 1.00), and 0.01 (−0.02 to 0.04), respectively. Heterogeneity was evident in the R^2 and calibration-in-the-large values across settings, but not in the calibration slope. Model performance did not vary markedly between boys and girls, age, ethnicity, and national income groups. To further improve the accuracy of the predictions, the model equation was recalibrated for the intercept in each setting so that country specific equations are available for future use.

CONCLUSION
The UK based prediction model, which is based on readily available measures, provides predictions of childhood fat-free mass, and hence fat mass, in a range of non-UK settings that explain a large proportion of the variability in observed fat-free mass, and exhibit good calibration performance, especially after recalibration of the intercept for each population. The model demonstrates good generalisability in both low-middle income and high income populations of healthy children and adolescents aged 4-15 years.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Improvements to the assessment of body fatness in children, currently based on body mass index (BMI), are required

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

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

WHAT THIS STUDY ADDS

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

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

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

Introduction

The high global prevalence of obesity in childhood poses a major global public health challenge. The World Health Organization recently estimated that about 18% of children and adolescents aged 5-19 years worldwide (>340 million individuals) are affected by overweight or obesity.[1] However, the most widely

used marker of high childhood adiposity, body mass index (BMI), has serious limitations.[2-4] BMI is poorly correlated with fat mass in childhood and, crucially, as a weight-for-height index, BMI is unable to discriminate between fat mass and fat-free mass, which can both vary markedly in individuals with a given BMI.[3] This is important, as fat mass in children has been shown to be more strongly associated with long term risk of type 2 diabetes than markers of body fatness based on overall weight (of which BMI is such a marker).[5] The availability of simple, accurate methods for assessment of fat mass could represent an important advance in the assessment of adiposity over the use of BMI.[4 6] In vivo techniques for the assessment of fat mass do exist, such as bioelectrical impedance analysis, dual energy x ray absorptiometry, and magnetic resonance imaging (MRI), but may lack accuracy,[4 7 8] and, in the case of dual energy x ray absorptiometry and MRI are inappropriate for general use. An alternative accurate method for assessment of fat mass, which has been developed and validated within the UK population of children and adolescents,[6 9] is based on the prediction of fat-free mass (and indirectly fat mass, as fat mass=weight–fat-free mass) using simple assessments of height, weight, sex, age, and ethnicity. Ethnicity was also included as a predictor within the developed model to allow for established differences in body fatness in children between ethnicities.[10-12] To maximise the accuracy of the predictions obtained from this approach, the model was developed using data on fat mass in children obtained from the deuterium dilution approach, a reference standard method of adiposity assessment, which provides accurate, safe, and minimally invasive measurements of total body water (and fat-free mass) with an error of less than 1%.[13 14] The developed model was also shown to predict fat mass levels as accurately as dual energy x ray absorptiometry and bioelectrical impedance analysis in UK children.[9] However, the predictive performance of the model has not so far been examined in children and adolescents outside the UK.[15] We therefore performed an external validation of the UK based fat mass prediction model to assess its predictive performance in a wide range of non-UK populations of children and adolescents aged 4-15 years with comparable fat mass assessment from the reference standard deuterium dilution method.

Methods

Data sources and study population

We carried out a Medline literature search through PubMed to identify all published studies that used the deuterium dilution method[13 14] to assess fat-free mass and fat mass in at least 100 healthy children or adolescents aged between 4 and 15 years and also included measurements of weight and height and basic personal information on age and sex (search dates: May 2020, search terms: "deuterium dilution", "study", "children" or "adolescents"). Twenty four relevant studies were identified; study authors were contacted and invited to contribute data to this investigation. Fourteen agreed to participate, four

were unable to participate owing to data sharing restrictions, and the remaining six did not respond. Study collaborators provided data from a further four studies (of which, one contained unpublished deuterium dilution data[16] and two were small[17 18]). Further information is available in the Preferred Reporting Items for Systematic Reviews and Meta-analyses individual participant data (PRISMA-IPD) flow diagram (see supplementary file). Each of the 18 studies provided data on weight, height, age, sex, ethnicity, and deuterium dilution assessed total body water, fat mass, and fat-free mass. The earliest study was conducted in 1981-82[19] and the latest in 2017-19[20]; most studies (83%) were carried out from 2000 (see supplementary table 1). In total, 5715 generally healthy children and adolescents aged 4-15 years were included from 18 studies spanning 19 countries: Australia,[17] Austria,[19] Bangladesh,[21] Brazil,[22] China,[23] Mexico,[24 25] Namibia,[20] Nepal,[26] the Netherlands,[27] New Zealand,[10] the Philippines,[23] Peru,[18] Poland,[16] Russia,[23] South Africa,[20 23] Spain,[28] Sri Lanka,[29] Tunisia,[30] and the US[31 32] (see supplementary table 1). Although five of these countries (Peru n=56, Spain n=92, China n=95, Philippines n=80, and Australia n=42) had smaller participant numbers than anticipated,[17 18 23 28] the data were included to avoid data wastage.

Outcome and predictor assessment

The outcome was natural log transformed fat-free mass (lnFFM), ascertained using the deuterium dilution reference method,[13 14] which was also used as the outcome for the previously developed UK based prediction model.[6] Five individuals with missing information on the outcome (lnFFM), one individual with an implausible weight value, and 16 individuals with implausible fat-free mass values (ie, fat-free mass greater than weight) were excluded. No other data were missing on the predictors of weight, height, age, sex, or ethnicity (classified as either white, black, South Asian, other Asian, or other). After these exclusions, 5693 children and adolescents aged 4-15 years were included in analyses.

Statistical analysis

All statistical analyses were conducted in Stata (version 17). We followed the Transparent Reporting of a multivariable model for Individual Prognosis Or Diagnosis (TRIPOD) guideline for the reporting of studies validating a multivariable prediction model.[15]

Evaluation of overall performance of UK based prediction model equation

We applied the UK based model equation[6] (fig 1) to all children and adolescents within this external validation dataset to obtain a prediction of lnFFM from weight, height, age, sex, and ethnicity.

The predictive performance of the model equation was assessed within each country by comparing the deuterium dilution assessed lnFFM value with the predicted lnFFM value obtained from the UK based

model. Model performance was assessed by examining established predictive performance measures of:

- R[2]—the percentage of variance in deuterium dilution observed lnFFM explained by predicted lnFFM estimated by the UK based prediction model.
- Root mean square error—the average difference between predicted lnFFM from the UK based model and the deuterium dilution observed lnFFM. Root mean square error indicates the absolute fit of the model to the data (ie, how close the model's predicted values of lnFFM are to the deuterium dilution observed values of lnFFM).
- Model calibration—assessed collectively by means of three measures: slope, calibration-in-the-large, and calibration plot.
  - *Slope*—obtained from the model regressing deuterium dilution observed values of lnFFM on model predicted values of lnFFM. The calibration slope reports on the accuracy of the predictions across the range of lnFFM values by evaluating the spread of the predicted values, and the slope has a target value of 1. A slope <1 suggests that predicted values are too high for children and adolescents with low observed values and too low for those with high observed values. A slope >1 suggests the opposite.
  - *Calibration-in-the-large*—intercept term obtained from a linear regression model of deuterium dilution observed values of lnFFM on model predicted values of lnFFM, where the slope is constrained to be the ideal of 1 (with a calibration-in-the-large of 0 being ideal). Calibration-in-the-large measures the overall agreement between average model predicted values of lnFFM and average deuterium dilution observed values of

lnFFM (ie, it tells you about the systematic bias of predicted lnFFM obtained from the model when compared with the deuterium dilution observed values of lnFFM).
  - *Calibration plot*—graph of deuterium dilution observed lnFFM plotted against model predicted values of lnFFM with a local regression (loess) smoother fitted across all individuals to produce a flexible calibration curve (created using pmcalplot on Stata).

To summarise results across all countries, we pooled the country specific performance measures of R[2], calibration slope, and calibration-in-the-large through a random effects meta-analysis to obtain an estimate of average performance and between country heterogeneity. The random effects meta-analysis model was fitted using the restricted maximum likelihood method[33] using the Hartung-Knapp[34] approach to adjust the standard error of the pooled performance measures. Heterogeneity was summarised for each of the performance measures of R[2], calibration slope, and calibration-in-the-large using the estimate of between study variance (τ^2) and a 95% prediction interval, which provides a range of values that the performance measure would be expected to fall within for a new individual setting.[35] For the random effects meta-analysis, the variance and confidence interval for R[2] was estimated using the Wald-type method outlined elsewhere,[36] and 95% prediction intervals were capped at the maximum values of 100%. The model performance was also assessed within subgroups of sex, three year age groups (4-6, 7-9, 10-12, and 13-15 years), and ethnic origin. Additionally, the model performance was assessed separately in low-middle income and high income populations of children and adolescents, categorised using the World Bank income classification.[37] The country specific classifications were obtained for the year of study commencement such that Australia, Austria, the Netherlands, New Zealand, Spain, the US, and Poland were classified as high income and the remaining 12 countries were classified as low-middle income.

Although the primary results focus on the model performance measures assessed comparing deuterium dilution observed values of lnFFM with model predicted lnFFM, as the UK based model was developed to predict fat-free mass on the log scale, root mean square error and calibration plots are also presented on the fat-free mass scale to improve the interpretability of the results.

Country specific recalibration of UK based prediction model equation

When systematic error in the UK based model's prediction of lnFFM was observed across countries, we recalibrated the intercept term of the UK based equation to provide updated country specific equations. Within each country, we first estimated the linear predictor portion of the UK based model for all individuals before fitting a linear regression model with deuterium dilution observed values of lnFFM

Fat mass = weight - exponential (0.3073 x height[2] - 10.0155 x weight[-1] + 0.004571 x weight + 0.01408 x BA - 0.06509 x SA - 0.02624 x AO - 0.01745 x other - 0.9180 x ln (age) + 0.6488 x age[0.5] + 0.04723 x male + country - specific constant term)

Score 1 if child is of black (BA), South Asian (SA), other Asian (AO), or 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, and age in years

Country	Constant term	Country	Constant term
North America		**North Africa**	
Mexico	2.8529	Tunisia	2.7858
US	2.8216	**Sub-Saharan Africa**	
South America		Namibia	2.7495
Brazil	2.8537	South Africa	2.7578
Peru	2.8447	**South Asia**	
Europe		Bangladesh	2.9009
Austria	2.8630	Nepal	2.8627
Netherlands	2.7787	Sri Lanka	2.7480
Poland	2.8450	**East Asia**	
Russia	2.6897	China	2.8867
Spain	2.8394	Philippines	2.7572
UK	2.8055	**Australasia**	
		Australia	2.8339
		New Zealand	2.8345

Fig 1 | Recalibrated country specific model equations for prediction of natural log transformed fat-free mass in children and adolescents. Country specific constant term for UK obtained from equation provided in Hudda et al[6]

Table 1 | Basic summary statistics of the analysis population, by country. Values are median (25th to 75th centile) unless stated otherwise

Region and country (No of participants)	Age (years)	Height (m)	Weight (kg)	DD fat mass (kg)*	DD fat-free mass (kg)*	No (%) boys	Ethnic group (No, %)				
							White	Black	South Asian	Other Asian	Other
North America											
Mexico (n=330)	8.4 (7.4-9.9)	1.30 (1.25-1.40)	31.6 (24.6-40.8)	7.6 (5.2-12.4)	22.7 (19.6-28.4)	170 (51.5)	0 (0)	0 (0)	0 (0)	0 (0)	330 (100)
US (n=1810)	10.6 (8.3-13.0)	1.44 (1.30-1.58)	39.2 (28.2-53.6)	9.1 (5.7-15.0)	28.5 (21.3-38.8)	867 (47.9)	571 (32)	457 (25)	2 (0.1)	283 (16)	497 (27)
South America											
Brazil (n=450)	13.3 (13.1-13.6)	1.58 (1.52-1.63)	49.5 (42.5-56.3)	10.1 (6.6-15.2)	38.1 (34.0-43.4)	236 (52.4)	0 (0)	0 (0)	0 (0)	0 (0)	450 (100)
Peru (n=56)	11.0 (8.5-13.5)	1.46 (1.32-1.55)	40.2 (32.4-50.6)	9.4 (5.8-14.6)	30.9 (23.9-35.5)	25 (44.6)	0 (0)	0 (0)	0 (0)	0 (0)	56 (100)
Europe											
Austria (n=107)	12.0 (11.1-13.6)	1.49 (1.42-1.60)	40.1 (34.0-48.0)	5.5 (4.2-8.5)	33.2 (28.5-40.6)	107 (100.0)	107 (100)	0 (0)	0 (0)	0 (0)	0 (0)
Netherlands (n=716)	13.5 (12.8-14.1)	1.62 (1.57-1.68)	51.3 (44.9-58.9)	11.9 (8.8-16.9)	38.6 (34.0-43.6)	342 (47.8)	716 (100)	0 (0)	0 (0)	0 (0)	0 (0)
Poland (n=174)	7.3 (6.1-8.7)	1.25 (1.18-1.34)	24.6 (21.0-28.0)	4.3 (3.3-5.8)	20.0 (17.5-22.8)	81 (46.6)	174 (100)	0 (0)	0 (0)	0 (0)	0 (0)
Russia (n=197)	10.8 (8.9-13.7)	1.47 (1.34-1.62)	37.9 (28.8-51.5)	11.3 (8.3-17.3)	25.8 (20.0-33.1)	97 (49.2)	0 (0)	0 (0)	0 (0)	0 (0)	197 (100)
Spain (n=92)	14.0 (13.0-15.2)	1.61 (1.55-1.68)	55.8 (48.1-63.0)	13.7 (8.8-17.9)	42.7 (37.0-49.2)	46 (50.0)	89 (97)	2 (2)	0 (0)	0 (0)	1 (1)
North Africa											
Tunisia (n=155)	9.0 (8.0-10.0)	1.38 (1.31-1.44)	32.0 (27.0-36.0)	8.0 (6.4-11.1)	23.1 (20.2-27.0)	80 (51.6)	0 (0)	0 (0)	0 (0)	0 (0)	155 (100)
Sub-Saharan Africa											
Namibia (n=151)	10.0 (9.0-11.0)	1.38 (1.33-1.46)	33.3 (27.2-43.4)	8.6 (6.1-14.5)	23.6 (20.6-28.1)	66 (43.7)	0 (0)	114 (76)	0 (0)	0 (0)	37 (25)
South Africa (n=411)	8.0 (7.0-8.8)	1.24 (1.17-1.31)	23.7 (20.4-28.3)	5.5 (4.4-7.6)	17.6 (15.3-21.2)	175 (42.6)	0 (0)	411 (100)	0 (0)	0 (0)	0 (0)
South Asia											
Bangladesh (n=187)	5.1 (5.0-7.1)	1.10 (1.02-1.17)	15.8 (14.1-18.4)	2.0 (1.4-2.8)	13.6 (12.2-16.2)	93 (49.7)	0 (0)	0 (0)	187 (100)	0 (0)	0 (0)
Nepal (n=100)	8.6 (8.3-9.0)	1.23 (1.16-1.29)	21.8 (18.0-25.8)	4.2 (3.3-5.8)	17.6 (14.6-20.2)	49 (49.0)	0 (0)	0 (0)	100 (100)	0 (0)	0 (0)
Sri Lanka (n=288)	10.0 (7.6-12.2)	1.37 (1.24-1.49)	31.5 (23.0-41.1)	9.1 (5.2-15.3)	21.5 (16.2-27.1)	162 (56.3)	0 (0)	0 (0)	288 (100)	0 (0)	0 (0)
East Asia											
China (n=95)	10.0 (9.4-10.6)	1.38 (1.33-1.43)	32.2 (27.4-37.0)	6.2 (3.9-8.7)	25.6 (22.8-28.2)	48 (50.5)	0 (0)	0 (0)	0 (0)	95 (100)	0 (0)
Philippines (n=80)	15.4 (15.1-15.7)	1.57 (1.52-1.64)	48.9 (43.4-54.5)	13.8 (9.5-17.2)	34.9 (31.3-40.8)	32 (40.0)	0 (0)	0 (0)	0 (0)	80 (100)	0 (0)
Australasia											
Australia (n=42)	8.2 (7.0-10.9)	1.33 (1.19-1.43)	27.2 (23.1-37.1)	6.6 (4.2-11.0)	21.9 (16.7-27.5)	27 (64.3)	0 (0)	0 (0)	42 (100)	0 (0)	0 (0)
New Zealand (n=252)	10.1 (7.4-12.4)	1.42 (1.27-1.54)	36.5 (28.6-50.0)	9.6 (6.3-15.6)	27.2 (21.7-37.1)	124 (49.2)	82 (33)	0 (0)	0 (0)	0 (0)	170 (67)

DD=deuterium dilution.

as the dependent variable and the linear predictor as the independent variable. The slope of this model was constrained to be 1, and the constant term from the model was used as the updated intercept term for that country. We re-examined the performance of the country specific recalibrated equations to assess the predictive performance of the updated model equations.

Patient and public involvement

For this external validation study based on secondary data analysis, no patients were directly involved in setting the research question, outcome measures, study design, or implementation. However, previous focus groups including members of the public had indicated the need to develop and validate new methods for body fatness assessment in children, which informed the development of this work. No patients were involved in the interpretation or writing up of results.

Results

Table 1 summarises the key characteristics of the 5693 children and adolescents included in the analysis, by country. The median age was 10.8 years—slightly higher than in the development population[6] (median age 9.6 years). All studies included both boys and girls, except for one study that included only boys.[19] Average levels of deuterium dilution assessed fat mass and fat-free mass were, as expected, higher in studies of older children, who were also generally taller and heavier than younger children (table 1). Median fat mass in this validation dataset (8.5 kg) was similar to that of the development population[6] (8.4 kg), although fat-free mass was higher in this external validation dataset (27.8 kg) than in the development population (24.8 kg).[6] The representation of different ethnic groups varied across the settings, with 15 countries including data from a single ethnic group (table 1).

Assessment of UK based model performance in non-UK settings

The UK based model equation, when applied to the external validation data, produced high R^2 values >75% in all countries and >90% in 11 of the 19 countries (table 2). When the country specific R^2 values were pooled through random effects meta-analysis, the overall pooled R^2 value was 88.66% (95% confidence interval 85.91% to 91.41%) (fig 2). The R^2 values showed some evidence of between country heterogeneity ($\tau^2 = 28.62$), and the 95% prediction interval for the R^2 value (ie, the expected range of R^2 values that would be expected from applying the model to a previously unstudied population of children and adolescents) was 77.04% to 100.00%. Root mean square error values were generally low (≤0.11 in all countries except for Russia and Sri Lanka; table 2). The model demonstrated high levels of calibration in terms of the slope in each of the countries. The observed calibration slope estimates ranged between 0.91 in Spain and 1.05 in Australia and South Africa (table 2), with the country specific

Table 2 | External validation predictive performance statistics based on natural log transformed fat-free mass, by country

Region and country	No	R^2 (%) (95% CI)	Calibration slope (95% CI)	Calibration-in-the-large (95% CI)	RMSE
North America					
Mexico	330	92.95 (91.49 to 94.42)	1.01 (0.98 to 1.04)	0.05 (0.04 to 0.05)	0.08
US	1810	93.32 (92.72 to 93.91)	1.00 (0.99 to 1.01)	0.02 (0.01 to 0.02)	0.10
South America					
Brazil	450	76.69 (72.92 to 80.46)	0.96 (0.91 to 1.01)	0.05 (0.04 to 0.06)	0.10
Peru	56	92.29 (88.41 to 96.17)	0.94 (0.87 to 1.01)	0.04 (0.02 to 0.06)	0.09
Europe					
Austria	107	91.47 (88.37 to 94.56)	0.96 (0.90 to 1.02)	0.06 (0.05 to 0.07)	0.09
Netherlands	716	81.53 (79.09 to 83.97)	1.00 (0.97 to 1.04)	−0.03 (−0.03 to −0.02)	0.09
Poland	174	93.28 (91.36 to 95.21)	0.96 (0.92 to 0.99)	0.04 (0.03 to 0.05)	0.07
Russia	197	91.30 (88.98 to 93.62)	0.93 (0.89 to 0.97)	−0.12 (−0.13 to −0.10)	0.15
Spain	92	80.85 (73.82 to 87.89)	0.91 (0.82 to 1.00)	0.03 (0.01 to 0.05)	0.10
North Africa					
Tunisia	155	80.98 (75.59 to 86.37)	1.02 (0.94 to 1.10)	−0.02 (−0.03 to −0.01)	0.08
Sub-Saharan Africa					
Namibia	151	90.14 (87.16 to 93.13)	0.93 (0.88 to 0.98)	−0.06 (−0.07 to −0.05)	0.09
South Africa	411	91.95 (90.46 to 93.44)	1.05 (1.02 to 1.08)	−0.05 (−0.05 to −0.04)	0.08
South Asia					
Bangladesh	187	89.50 (86.65 to 92.35)	0.99 (0.94 to 1.04)	0.10 (0.09 to 0.10)	0.11
Nepal	100	91.66 (88.53 to 94.79)	0.99 (0.93 to 1.04)	0.06 (0.05 to 0.07)	0.08
Sri Lanka	288	83.11 (79.56 to 86.67)	0.99 (0.94 to 1.04)	−0.06 (−0.07 to −0.04)	0.16
East Asia					
China	95	85.10 (79.57 to 90.63)	0.98 (0.90 to 1.07)	0.08 (0.07 to 0.10)	0.11
Philippines	80	81.03 (73.55 to 88.51)	1.04 (0.93 to 1.15)	−0.05 (−0.07 to −0.03)	0.09
Australasia					
Australia	42	96.64 (94.65 to 98.64)	1.05 (0.98 to 1.11)	0.03 (0.01 to 0.05)	0.07
New Zealand	252	92.96 (91.29 to 94.64)	0.98 (0.94 to 1.01)	0.03 (0.02 to 0.04)	0.10

RMSE=root mean square error.

95% confidence intervals around the respective slope term containing the ideal value of 1 in 15 of the 19 countries (table 2). The pooled calibration slope was 0.98 (95% confidence interval 0.97 to 1.00; fig 2), with no evidence of between country heterogeneity in the calibration slopes (τ^2=0), and a 95% prediction interval of 0.92 to 1.05 for the calibration slope expected to be observed in a new country. Although the country specific calibration-in-the-large values were close to the ideal value of 0 for most countries, ranging from −0.12 in Russia to 0.10 in Bangladesh, the associated 95% confidence intervals failed to contain the ideal value of 0 in any of the settings (table 2). When pooled through random effects meta-analysis, the overall pooled calibration-in-the-large value was almost equal to the ideal value of 0 (0.01, 95% confidence interval −0.02 to 0.04; fig 2). Some heterogeneity was evident in the calibration-in-the-large values (τ^2=0.0031), with a 95% prediction interval of −0.11 to 0.13 that would likely be observed in a new country. Graphically, the calibration plots also demonstrated the good levels of calibration across the range of lnFFM values within each of the countries (fig 3 and fig 4), including at the lower and upper ends of the distribution of lnFFM, with the flexible calibration curve close to the ideal 45 degree line of perfect calibration. However, the graphs further showed evidence of some systematic error in the prediction of lnFFM across settings (fig 3 and fig 4).

The generally high model performance was maintained when assessed across subgroups of sex (see supplementary figure 1), three year age groups (see supplementary figure 2), ethnic groups (see

supplementary figure 3), and national income level (table 3). The model showed low levels of heterogeneity in the performance statistics across the three year age groups (see supplementary figure 4) and ethnic groups (see supplementary figure 5).

On the fat-free mass scale, the country specific root mean square error values ranged between 1.32 kg and 4.83 kg, with a root mean square error value of <4 kg in 17 of the 19 countries (see supplementary table 2). Calibration plots showed good levels of agreement between deuterium dilution observed fat-free mass and model predicted fat-free mass within each of the countries, including at the lower and upper ends of the distribution of fat-free mass, with the flexible calibration curve close to the ideal 45 degree line of perfect calibration (see supplementary figures 6a and 6b).
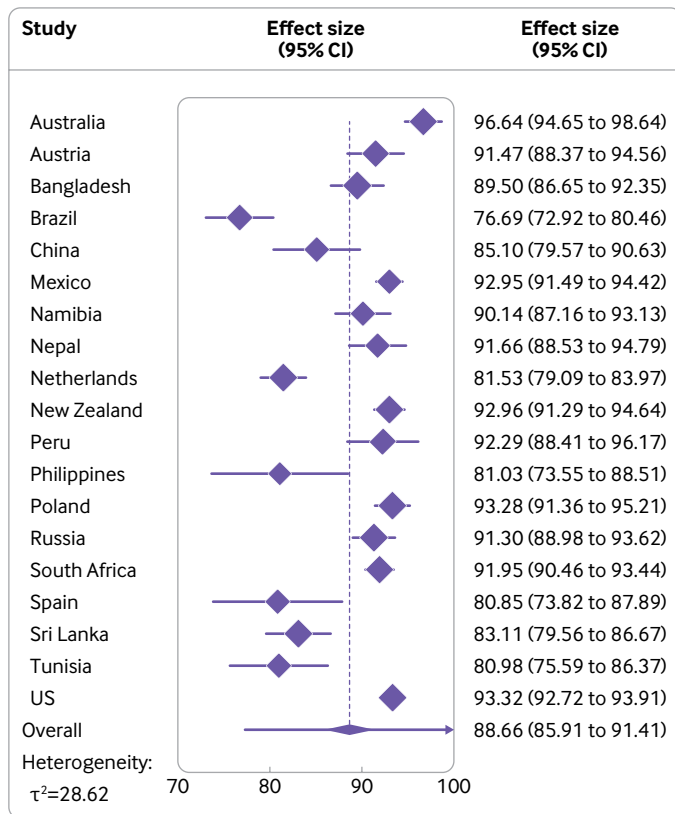
As a result of the observed systematic error in the UK based model's prediction of lnFFM, the equation was recalibrated for the intercept term to provide updated country specific equations (fig 1) and calibration plots (see supplementary figures 7a and 7b). After recalibration, the calibration-in-the-large and root mean square error values were closer to the ideal values of 0 in all settings, with all of the country specific 95% confidence intervals for the calibration-in-the-large now containing the ideal value (see supplementary table 3).
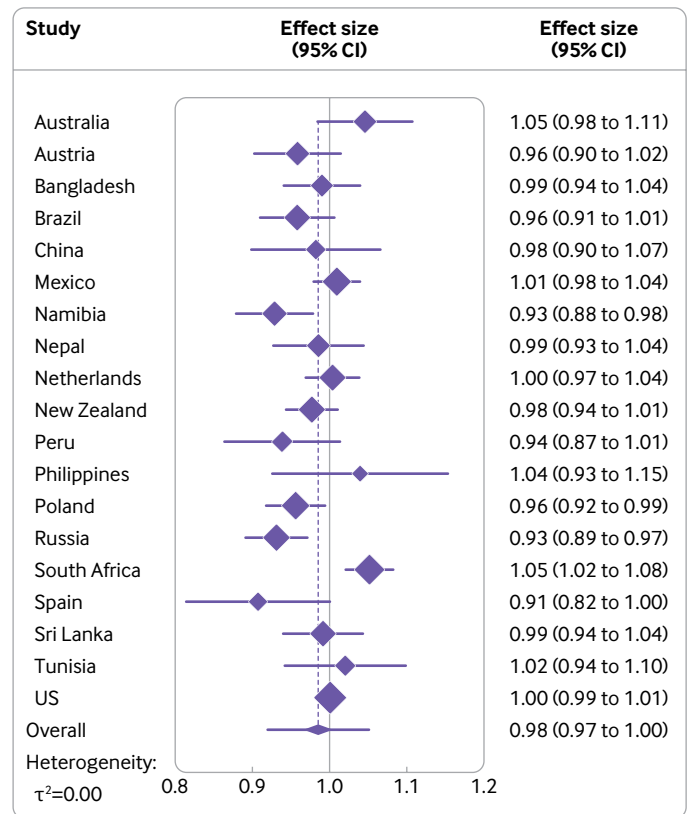
Discussion
Principal findings
We externally validated the predictive performance of a recently proposed model using weight, height,

R^2

Study	Effect size (95% CI)	Effect size (95% CI)
Australia		96.64 (94.65 to 98.64)
Austria		91.47 (88.37 to 94.56)
Bangladesh		89.50 (86.65 to 92.35)
Brazil		76.69 (72.92 to 80.46)
China		85.10 (79.57 to 90.63)
Mexico		92.95 (91.49 to 94.42)
Namibia		90.14 (87.16 to 93.13)
Nepal		91.66 (88.53 to 94.79)
Netherlands		81.53 (79.09 to 83.97)
New Zealand		92.96 (91.29 to 94.64)
Peru		92.29 (88.41 to 96.17)
Philippines		81.03 (73.55 to 88.51)
Poland		93.28 (91.36 to 95.21)
Russia		91.30 (88.98 to 93.62)
South Africa		91.95 (90.46 to 93.44)
Spain		80.85 (73.82 to 87.89)
Sri Lanka		83.11 (79.56 to 86.67)
Tunisia		80.98 (75.59 to 86.37)
US		93.32 (92.72 to 93.91)
Overall		88.66 (85.91 to 91.41)
Heterogeneity: τ^2=28.62	70 80 90 100	

Calibration slope

Study	Effect size (95% CI)	Effect size (95% CI)
Australia		1.05 (0.98 to 1.11)
Austria		0.96 (0.90 to 1.02)
Bangladesh		0.99 (0.94 to 1.04)
Brazil		0.96 (0.91 to 1.01)
China		0.98 (0.90 to 1.07)
Mexico		1.01 (0.98 to 1.04)
Namibia		0.93 (0.88 to 0.98)
Nepal		0.99 (0.93 to 1.04)
Netherlands		1.00 (0.97 to 1.04)
New Zealand		0.98 (0.94 to 1.01)
Peru		0.94 (0.87 to 1.01)
Philippines		1.04 (0.93 to 1.15)
Poland		0.96 (0.92 to 0.99)
Russia		0.93 (0.89 to 0.97)
South Africa		1.05 (1.02 to 1.08)
Spain		0.91 (0.82 to 1.00)
Sri Lanka		0.99 (0.94 to 1.04)
Tunisia		1.02 (0.94 to 1.10)
US		1.00 (0.99 to 1.01)
Overall		0.98 (0.97 to 1.00)
Heterogeneity: τ^2=0.00	0.8 0.9 1.0 1.1 1.2	

Calibration-in-the-large

Study	Effect size (95% CI)	Effect size (95% CI)
Australia		0.03 (0.01 to 0.05)
Austria		0.06 (0.05 to 0.07)
Bangladesh		0.10 (0.09 to 0.10)
Brazil		0.05 (0.04 to 0.06)
China		0.08 (0.07 to 0.10)
Mexico		0.05 (0.04 to 0.05)
Namibia		-0.06 (-0.07 to -0.05)
Nepal		0.06 (0.05 to 0.07)
Netherlands		-0.03 (-0.03 to -0.02)
New Zealand		0.03 (0.02 to 0.04)
Peru		0.04 (0.02 to 0.06)
Philippines		-0.05 (-0.07 to -0.03)
Poland		0.04 (0.03 to 0.05)
Russia		-0.12 (-0.13 to -0.10)
South Africa		-0.05 (-0.05 to -0.04)
Spain		0.03 (0.01 to 0.05)
Sri Lanka		-0.06 (-0.07 to -0.04)
Tunisia		-0.02 (-0.03 to -0.01)
US		0.02 (0.01 to 0.02)
Overall		0.01 (-0.02 to 0.04)
Heterogeneity: τ^2=0.00	-0.13 -0.08 -0.03 0.02 0.07 0.12	

Fig 2 | Assessment of R^2 values, calibration slope, and calibration-in-the-large based on natural log transformed fat-free mass, by country and overall. Overall estimates from random effect restricted maximum likelihood model with Hartung-Knapp standard errors. Purple line around the overall diamond indicates the 95% prediction intervals. Upper limit of the prediction interval for R^2 capped at 100%
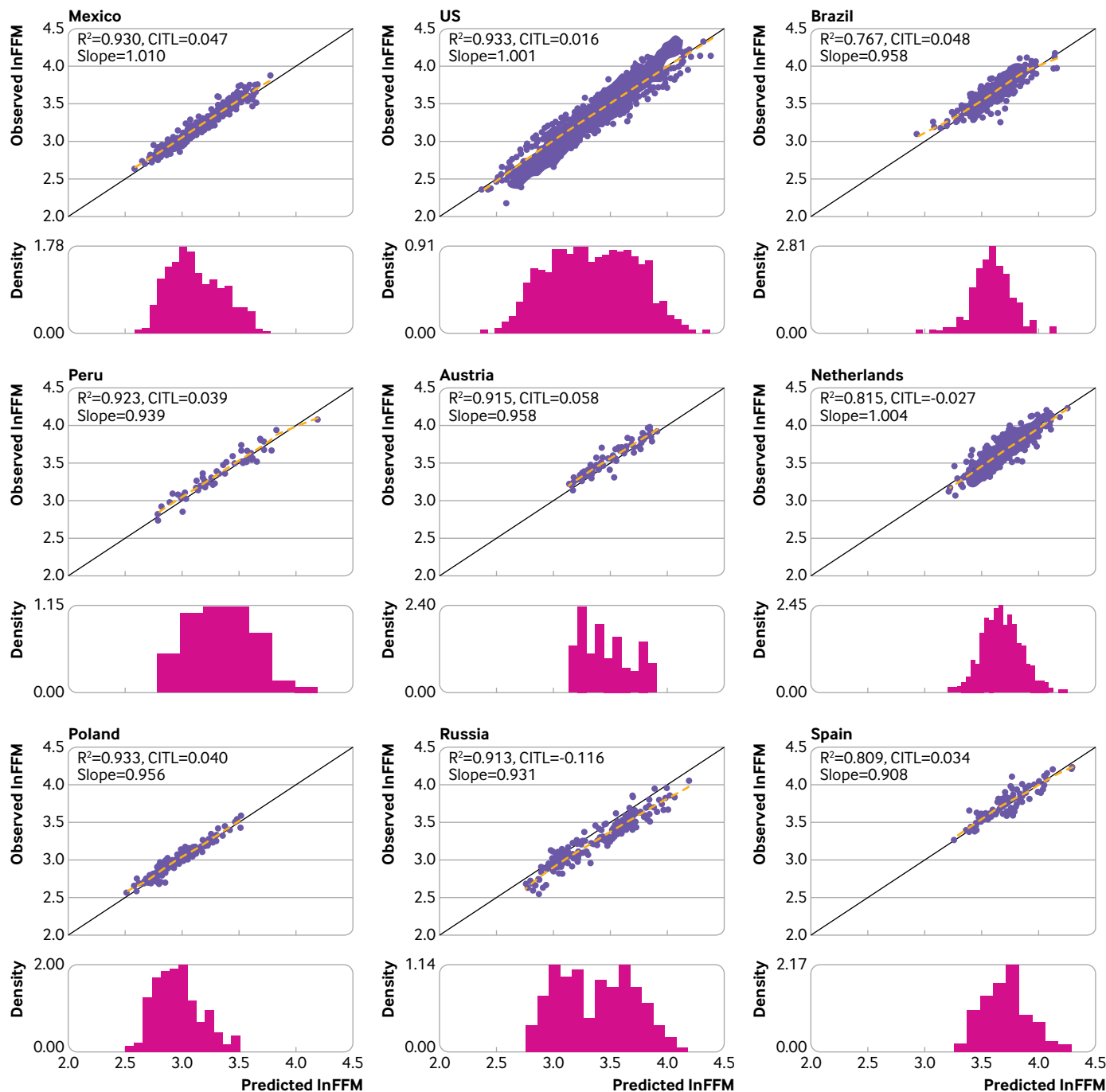
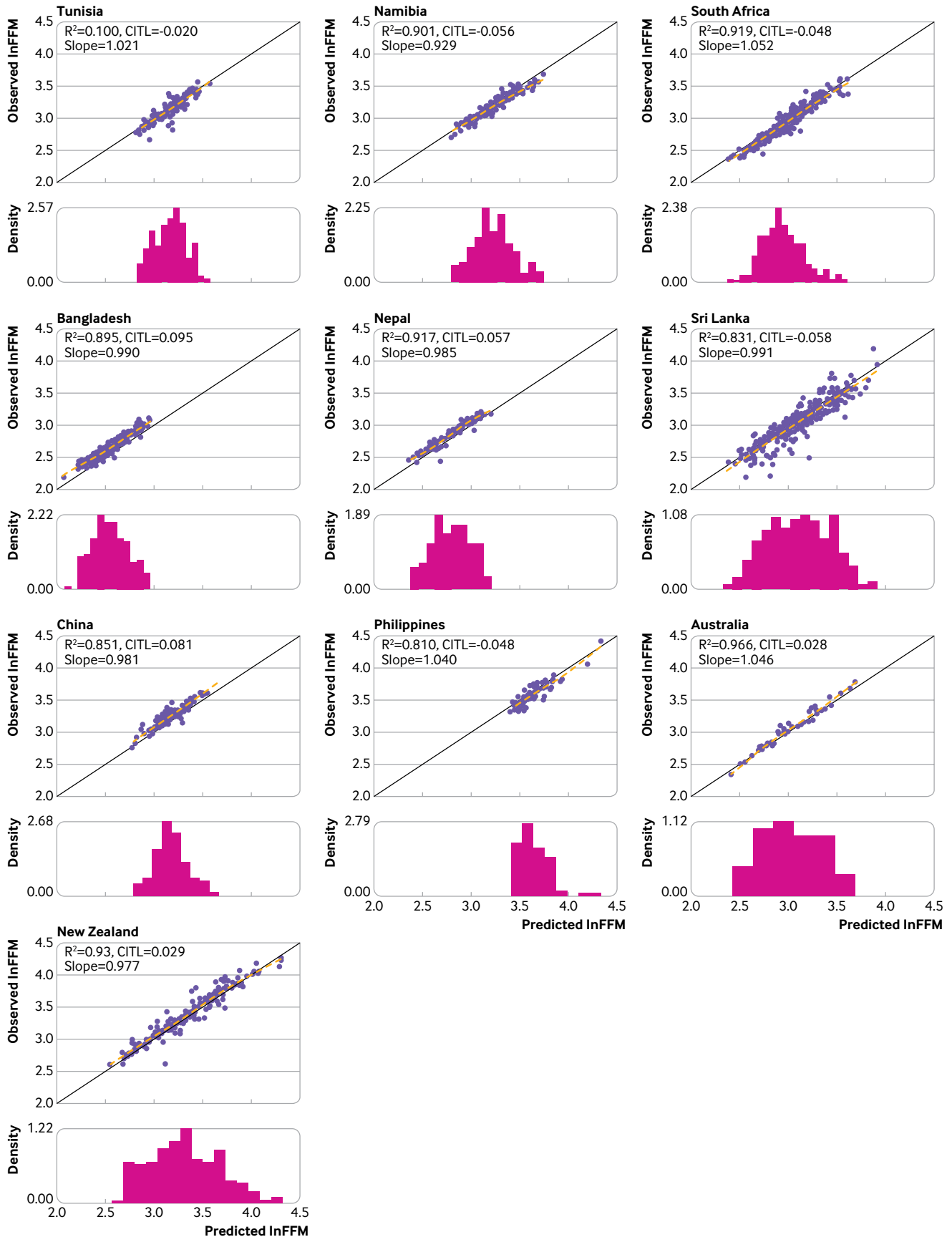**Fig 3 |** Calibration assessment of the model based on natural log transformed fat-free mass (lnFFM) in the Americas and European countries. Solid black line represents line of equality. Dashed line is a loess smoother through the individual data points. Histogram is the distribution of predicted ln(FFM). Slope=calibration slope; CITL=calibration-in-the-large

age, sex, and ethnicity to estimate lnFFM, developed and validated within a UK population of children and adolescents, in 19 other countries from several regions of the world. The developed model equation generally showed good predictive ability in these new settings, with good calibration of observed and predicted values, demonstrating the generalisability of the model in populations of children and adolescents outside the UK. The model equation produced high R^2 values >80% in all settings, with root mean square

error values (expressed in terms of fat-free mass for interpretability) <4 kg in most of the settings. The root mean square error of 0.1 for lnFFM indicates an error of 10% on predicted fat-free mass values, which takes into account that the absolute error in any individual child or adolescent will depend on the magnitude of his or her fat-free mass. Notably, the predictive performance was consistently high among both low-middle income and high income countries. Although the calibration slope statistics were close to

Fig 4 | Calibration assessment of the model based on natural log transformed fat-free mass (lnFFM) in African, Asian, and Australasian countries. Solid black line represents line of equality. Dashed line is a loess smoother through the individual data points. Histogram is the distribution of predicted ln(FFM). Slope=calibration slope; CITL=calibration-in-the-large

Table 3 | External validation predictive performance statistics based on natural log transformed fat-free mass, by World Bank income classifications*

Performance statistics	Low-middle income group† (n=2473)	High income group‡ (n=3193)
R^2 (%) (95% CI)	92.19 (91.60 to 92.78)	93.64 (93.21 to 94.07)
Calibration slope (95% CI)	0.98 (0.97 to 0.99)	0.97 (0.97 to 0.98)
Calibration-in-the-large (95% CI)	−0.00 (−0.01 to 0.00)	0.01 (0.01 to 0.01)
Root mean square error	0.11	0.10

*Ascertained for initial calendar year the study began.
†Comprises low, lower middle, and upper middle income groups. Includes Bangladesh, Brazil, China, Mexico, Namibia, Nepal, Peru, the Philippines, Russia, South Africa, Sri Lanka, and Tunisia.
‡Includes Australia, Austria, the Netherlands, New Zealand, Spain, US, and Poland.

the ideal value of 1 for all countries, the calibration-in-the-large values suggested a small systematic error in the prediction of lnFFM across settings. Therefore, the model equation was recalibrated for the intercept term to provide updated country specific prediction equations. After recalibration, the model performance showed improvement in the calibration-in-the-large and root mean square error values in all settings as expected.

Comparison with other studies

A small number of previous studies have developed models to estimate body fatness in populations of children and adolescents outside the UK,[38-43] producing R^2 values >80% comparable with those observed in the present study. However, direct comparisons of the performance of those models with the model being validated in this study are difficult for several reasons. Firstly, while those models also produced high R^2 values >80%, their outcomes (fat mass percentage) are different from the outcome of the current model (absolute fat-free mass values). Although fat mass needs to be standardised for height before interpretation or comparisons between individuals (typically expressed as a fat mass index or as fat mass percentage), the use of absolute values (of either fat mass or fat-free mass) as the outcome of the prediction model produces more accurate and precise predictions than estimating fat mass percentage, which is derived from fat mass and weight and is more variable than the absolute values. Furthermore, as discussed previously,[6] we chose to estimate absolute fat mass values indirectly (by predicting fat-free mass from the model and subtracting estimates from weight to obtain predicted fat mass) rather than directly (using fat mass as the outcome of the model), as the variability in fat-free mass with height (one of the strongest predictors of body composition) was more homogeneous than for fat mass and thus the indirect approach resulted in more precise fat mass predictions. We believe this modelling decision was one of the main reasons for the high predictive performance observed. Secondly, neither the calibration slopes nor the calibration-in-the-large values were assessed for these models. Thirdly, these models largely used dual energy x ray absorptiometry as the reference method for assessing body fatness (ie, as the outcome for model development). Estimation of fat mass using dual energy x ray absorptiometry is, however, prone to low levels of accuracy,[4 7 8 44 45]

which varies considerably by body shape, sex,[8 44] and different dual energy x ray absorptiometry devices and software used. Finally, most models used additional measurements, including skinfold thickness, waist circumference, or bioelectrical impedance to estimate fat mass, rather than being based on readily available anthropometric and demographic predictors.[39-43] One such study, which used the same data on Tunisian children used in our present study, also developed a prediction model to estimate fat-free mass.[30] However, the model was not based on readily available predictors as it relied on resistance from bioimpedance analysis and was not corrected for model optimism, which may explain why the model equation produced a R^2 value of 91.8% compared with the value of 81.0% obtained in the present study.

Strengths and limitations of this study

The current investigation had several strengths in its approach to validating the developed UK based prediction model. Crucially, from each of the 19 countries, body fatness assessments were made using the same reference standard deuterium dilution method on which the UK based model was based,[6] which provides accurate, safe, and minimally invasive measurements of total body water (and fat-free mass) with low error.[13 14] Most studies included in this external validation were conducted recently and were sufficiently large to provide accurate estimates of the country specific prediction performance statistics. The pooled data used for this external validation across the 19 settings spanned a wide age range of participants aged 4-15 years, allowing us to obtain an accurate picture of the model performance across childhood and adolescence in the new settings. The maintained high predictive performance of the model at both the lower and the upper ends of the distribution of fat-free mass indicates the model's potential utility for population based obesity surveillance (for example, in the English National Child Measurement Programme and the WHO Childhood Obesity Surveillance Initiative). The populations included about equal numbers of both high (56%) and low to middle income countries, which allowed for accurate subgroup assessment by income classification.

The study also had some limitations. Firstly, global representation was limited, with only a small number of participants included from East Asia and none from the Middle East. Secondly, five countries had fewer than 100 participants, which impacted the representativeness of the results and the precision of estimates obtained from these populations, although predictive performance statistics from these populations were consistent with those of the larger countries. Finally, deuterium dilution assessment involves estimating total body water, which is then converted into an estimate of fat-free mass using a chosen hydration constant.[46] The hydration constants used across the individual studies were those originally applied by individual study authors. The study

results were not, however, materially affected (results not presented) when using a consistent hydration constant[46] for all studies, and therefore the choice of hydration constants is unlikely to have affected the results appreciably.

Implications for clinicians and policy makers

Although BMI is widely used because of its simplicity and reliance solely on measures of height and weight, it has well recognised serious limitations as a marker of body fatness in children.[2-4] For example, BMI is unable to discriminate between fat mass and fat-free mass, and it has been shown that even at a given BMI among children of the same age and sex, both fat mass and fat-free mass can vary substantially,[3] which is particularly concerning in light of evidence that, independent of height, fat mass in childhood is strongly associated with emerging cardiometabolic risk[47-56] and is more strongly related with risk of type 2 diabetes in adulthood than overall weight (on which BMI is directly based).[5] Furthermore, it has been suggested that fat-free mass, independent of fat mass, also has implications on adverse health conditions. This emphasises the needed for a shift away from proxy weight-for-height indices such as BMI, which produce inaccurate body fatness assessment in children[12 57-59] and result in a large proportion being misdiagnosed with overweight or obesity, and towards more direct and accurate assessment of fat mass. The availability of the extensively validated prediction model provides a major advancement within this area of body fatness assessment in children. The accuracy and simplicity of the model, relying solely on readily available non-invasive measurements to assess fat-free mass and thus fat mass (fat mass=weight−exponential[predicted lnFFM]), has implications for its wider applicability both in routine healthcare practice and in population-wide obesity surveillance, monitoring, and prevention initiatives where more complex measurements of body fatness (such as waist circumference and skinfold thickness) are not so readily available. The consistency of the model performance across settings also strengthens the conclusions that the model, particularly after recalibration at local level, minimises the bias in fat-free mass predictions across a wide range of settings. However, the UK based model before local country specific recalibration of the intercept terms was shown to have good predictive ability across the settings, both in high and in low-middle income populations, and thus can be implemented in childhood settings where a recalibrated equation is not provided.

Body fatness markers require standardisation for height to minimise their correlation with height,[60] to provide a consistent measure across age and sex, allowing for efficient monitoring and tracking of body fatness in childhood. Two common approaches to provide appropriate standardisation of fat mass for height are to convert fat mass to fat mass percentage (fat mass/weight×100) or to fat mass index (fat mass/height^p, where p is the power of height needed

to obtain maximum height independence in the population of interest). Fat mass assessment (using either the updated country specific model equation provided when available or the original UK based model), after height standardisation (either as fat mass percentage or as fat mass index), can then be used in conjunction with respective established reference curves[61-63] for improved surveillance, management, and prevention of obesity in childhood. A MS Excel calculator has been developed (see supplementary file) to allow for simple calculation of fat-free mass, fat mass, and fat mass percentage from the relevant predictor variables. Assessments of fat mass can also be made readily available to clinicians by embedding the validated equation within existing computer software used by general practitioners and paediatricians. This approach would be consistent with the use of other prediction algorithms in clinical practice, such as the QRISK3 or Framingham risk score (cardiovascular disease).[64 65] It would also be possible and straightforward to apply the algorithms within existing obesity surveillance initiatives for children such as the English National Child Measurement Programme and the WHO Childhood Obesity Surveillance Initiative, to provide assessments of fat mass, fat-free mass, and fat mass index or fat mass percentage.

Further research

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

AUTHOR AFFILIATIONS
[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]Department of Nutrition, University of North Carolina Schools of Public Health and Medicine, NC, USA

[4]Faculty of Medicine, University of Malaga, Málaga, Spain

[5]Department of Pediatrics, New York Nutrition Obesity Research Center, Columbia University Medical Center, New York, NY, USA

[6]Coordinación de Nutrición, Centro de Investigación en Alimentación y Desarrollo, Hermosillo, Mexico

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

[8]Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

[9]Translational Research Institute, Adventhealth Orlando, Orlando, FL, USA

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

[11]UCL Institute for Global Health, London, UK

[12]Department of Medicine and Institute Human Nutrition, Division of Endocrinology, New York Nutrition Obesity Research Center, Columbia University Medical Center, New York, NY, USA

[13]Centro Universitario UAEM Amecameca, Universidad Autónoma del Estado de México, Amecameca de Juárez, Mexico

[14]Department of Pediatrics, PMU Salzburg, Salzburg, Austria

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

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

[17]International Centre for Diarrheal Disease Research, Dhaka 1212, Bangladesh

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

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

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

[21]Department of Nutrition and Movement Sciences, NUTRIM, Maastricht University, Maastricht, Netherlands

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

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

[24]Baylor College of Medicine, USDA/ARS Children's Nutrition Research Center, Houston, TX, USA

[25]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge, UK

[26]Department of Work and Social Psychology, Maastricht University, Maastricht, Netherlands

[27]Departamento de Bioquímica, Centro de Investigación y de Estudios Avanzados del IPN, Mexico City, Mexico

[28]Faculty of Medicine, University of Colombo, Sri Lanka

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

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

Dissemination to participants and related patient and public communities: We plan to disseminate these research findings to relevant stakeholders by presenting our findings at obesity related conferences, by sharing the findings with The Office for Health Improvement and Disparities (formerly Public Health England) who are responsible for the English National Child Measurement Programme and the WHO Childhood Obesity Surveillance Initiative Steering Group members, and by making plain language summaries available on social media.

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

1 World Health Organization. Obesity and overweight 2018. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight (accessed 09/12/2020).

2 Hall DM, Cole TJ. What use is the BMI? *Arch Dis Child* 2006;91:283-6. doi:10.1136/adc.2005.077339

3 Wells JC. A Hattori chart analysis of body mass index in infants and children. *Int J Obes Relat Metab Disord* 2000;24:325-9.

4 Wells JC, Fewtrell MS. Measuring body composition. *Arch Dis Child* 2006;91:612-7. doi:10.1136/adc.2005.085522

5 Hudda MT, Aarestrup J, Owen CG, et al. Association of Childhood Fat Mass and Weight With Adult-Onset Type 2 Diabetes in Denmark. *JAMA Netw Open* 2021;4:e218524. doi:10.1001/jamanetworkopen.2021.8524

6 Hudda MT, Fewtrell MS, Haroun D, et al. Development and validation of a prediction model for fat mass in children and adolescents: meta-analysis using individual participant data. *BMJ* 2019;366:l4293. doi:10.1136/bmj.l4293

7 Williams JE, Wells JC, Wilson CM, Haroun D, Lucas A, Fewtrell MS. Evaluation of Lunar Prodigy dual-energy X-ray absorptiometry for assessing body composition in healthy persons and patients by comparison with the criterion 4-component model. *Am J Clin Nutr* 2006;83:1047-54. doi:10.1093/ajcn/83.5.1047

8 Reilly JJ, Gerasimidis K, Paparacleous N, et al. Validation of dual-energy x-ray absorptiometry and foot-foot impedance against deuterium dilution measures of fatness in children. *Int J Pediatr Obes* 2010;5:111-5.

9 Hudda MT, Owen CG, Rudnicka AR, Cook DG, Whincup PH, Nightingale CM. Quantifying childhood fat mass: comparison of a novel height-and-weight-based prediction approach with DXA and bioelectrical impedance. *Int J Obes (Lond)* 2021;45:99-103. doi:10.1038/s41366-020-00661-w

10 Rush EC, Puniani K, Valencia ME, Davies PS, Plank LD. Estimation of body fatness from body mass index and bioelectrical impedance: comparison of New Zealand European, Maori and Pacific Island children. *Eur J Clin Nutr* 2003;57:1394-401. doi:10.1038/sj.ejcn.1601701

11 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

12 Hudda MT, Nightingale CM, Donin AS, et al. Body mass index adjustments to increase the validity of body fatness assessment in UK Black African and South Asian children. *Int J Obes (Lond)* 2017;41:1048-55. doi:10.1038/ijo.2017.75

13 Deurenberg P, Yap M. The assessment of obesity: methods for measuring body fat and global prevalence of obesity. *Baillieres Best Pract Res Clin Endocrinol Metab* 1999;13:1-11. doi:10.1053/beem.1999.0003

14 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

15 Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1-73. doi:10.7326/M14-0698

16 Desmond MA, Sobiecki JG, Jaworski M, et al. Growth, body composition, and cardiovascular and nutritional risk of 5- to 10-y-old children consuming vegetarian, vegan, or omnivore diets. *Am J Clin Nutr* 2021;113:1565-77. doi:10.1093/ajcn/nqaa445

17 Wickramasinghe VP, Cleghorn GJ, Edmiston KA, Murphy AJ, Abbott RA, Davies PS. Validity of BMI as a measure of obesity in Australian white Caucasian and Australian Sri Lankan children. *Ann Hum Biol* 2005;32:60-71. doi:10.1080/03014460400027805

18 Danysh HE, Gilman RH, Wells JC, et al. El Niño adversely affected childhood stature and lean mass in northern Peru. *Clim Change Responses* 2014;1:7. doi:10.1186/s40665-014-0007-z.

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

20 Diouf A, Adom T, Aouidet A, et al. Body mass index vs deuterium dilution method for establishing childhood obesity prevalence, Ghana, Kenya, Mauritius, Morocco, Namibia, Senegal, Tunisia and United Republic of Tanzania. *Bull World Health Organ* 2018;96:772-81. doi:10.2471/BLT.17.205948

21 Khan AI, Hawkesworth S, Hawlader MD, et al. Body composition of Bangladeshi children: comparison and development of leg-to-leg bioelectrical impedance equation. *J Health Popul Nutr* 2012;30:281-90. doi:10.3329/jhpn.v30i3.12291

22 Reichert FF, Wells JC, Ekelund U, Menezes AM, Victora CG, Hallal PC. Prospective Associations Between Physical Activity Level and Body Composition in Adolescence: 1993 Pelotas (Brazil) Birth Cohort. *J Phys Act Health* 2015;12:834-9. doi:10.1123/jpah.2013-0509

23 Eckhardt CL, Adair LS, Caballero B, et al. Estimating body fat from anthropometry and isotopic dilution: a four-country comparison. *Obes Res* 2003;11:1553-62. doi:10.1038/oby.2003.207

24 Ballesteros-Vásquez MN, Guerrero-Alcocer EV, Grijalva Haro MI, Ramírez-López E, Robles-Sardin AE. Effect of ingestion and excretion of fluids in determining body composition with deuterium dilution method in school children. *Nutr Hosp* 2015;32:1324-8.

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

26 Devakumar D, Grijalva-Eternod CS, Roberts S, et al. Body composition in Nepalese children using isotope dilution: the production of ethnic-specific calibration equations and an exploration of methodological issues. *PeerJ* 2015;3:e785. doi:10.7717/peerj.785

27 Ten Hoor GA, Rutten GM, Van Breukelen GJP, et al. Strength exercises during physical education classes in secondary schools improve body composition: a cluster randomized controlled trial. *Int J Behav Nutr Phys Act* 2018;15:92. doi:10.1186/s12966-018-0727-8

28 Alvero-Cruz JR, Alvarez Carnero E, Fernández-García JC, Barrera Expósito J, Carrillo de Albornoz Gil M, Sardinha LB. [Validity of body mass index and fat mass index as indicators of overweight status in Spanish adolescents: Esccola Study]. *Med Clin (Barc)* 2010;135:8-14. doi:10.1016/j.medcli.2010.01.017

29 Wickramasinghe VP, Lamabadusuriya SP, Cleghorn GJ, Davies PS. Validity of currently used cutoff values of body mass index as a measure of obesity in Sri Lankan children. *Ceylon Med J* 2009;54:114-9. doi:10.4038/cmj.v54i4.1451

30 Ben Jemaa H, Mankaï A, Khlifi S, et al. Development and validation of impedance-based equations for the prediction of total body water and fat-free mass in children aged 8-11 years. *Clin Nutr* 2019;38:227-33.

31 Horlick M, Arpadi SM, Bethel J, et al. Bioelectrical impedance analysis models for prediction of total body water and fat-free mass in healthy and HIV-infected children and adolescents. *Am J Clin Nutr* 2002;76:991-9. doi:10.1093/ajcn/76.5.991

32 Shypailo RJ, Wong WW. Fat and fat-free mass index references in children and young adults: assessments along racial and ethnic lines. *Am J Clin Nutr* 2020;112:566-75. doi:10.1093/ajcn/nqaa128

33 Viechtbauer W. Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *J Educ Behav Stat* 2005;30:261-93. doi:10.3102/10769986030003261.

34 Hartung J, Knapp G. On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Stat Med* 2001;20:1771-82. doi:10.1002/sim.791

35 Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;342:d549. doi:10.1136/bmj.d549

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

37 The World Bank. The world by income 2020. https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html (accessed 22/12/2021).

38 Dugas LR, Cao G, Luke AH, 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

39 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

40 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

41 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

42 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

43 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

44 Sinaiko AR, Donahue RP, Jacobs DRJr, Prineas RJ. Relation of weight and rate of increase in weight during childhood and adolescence to body size, blood pressure, fasting insulin, and lipids in young adults. The Minneapolis Children's Blood Pressure Study. *Circulation* 1999;99:1471-6. doi:10.1161/01.CIR.99.11.1471

45 Wong WW, Hergenroeder AC, Stuff JE, Butte NF, Smith EO, Ellis KJ. Evaluating body fat in girls and female adolescents: advantages and disadvantages of dual-energy X-ray absorptiometry. *Am J Clin Nutr* 2002;76:384-9. doi:10.1093/ajcn/76.2.384

46 Wells JC, Williams JE, Chomtho S, et al. Pediatric reference data for lean tissue properties: density and hydration from age 5 to 20 y. *Am J Clin Nutr* 2010;91:610-8. doi:10.3945/ajcn.2009.28428

47 Steinberger J, Jacobs DR, Raatz S, Moran A, Hong CP, Sinaiko AR. Comparison of body fatness measurements by BMI and skinfolds vs dual energy X-ray absorptiometry and their relation to cardiovascular risk factors in adolescents. *Int J Obes (Lond)* 2005;29:1346-52. doi:10.1038/sj.ijo.0803026

48 Lindsay RS, Hanson RL, Roumain J, Ravussin E, Knowler WC, Tataranni PA. Body mass index as a measure of adiposity in children and adolescents: relationship to adiposity by dual energy x-ray absorptiometry and to cardiovascular risk factors. *J Clin Endocrinol Metab* 2001;86:4061-7. doi:10.1210/jcem.86.9.7760

49 Falaschetti E, Hingorani AD, Jones A, et al. Adiposity and cardiovascular risk factors in a large contemporary population of pre-pubertal children. *Eur Heart J* 2010;31:3063-72. doi:10.1093/eurheartj/ehq355

50 Whincup PH, Gilg JA, Donald AE, et al. Arterial distensibility in adolescents: the influence of adiposity, the metabolic syndrome, and classic risk factors. *Circulation* 2005;112:1789-97. doi:10.1161/CIRCULATIONAHA.104.532663

51 Samouda H, de Beaufort C, Stranges S, et al. Adding anthropometric measures of regional adiposity to BMI improves prediction of cardiometabolic, inflammatory and adipokines profiles in youths: a cross-sectional study. *BMC Pediatr* 2015;15:168. doi:10.1186/s12887-015-0486-5

52 Gutin B, Islam S, Manos T, Cucuzzo N, Smith C, Stachura ME. Relation of percentage of body fat and maximal aerobic capacity to risk factors for atherosclerosis and diabetes in black and white seven- to eleven-year-old children. *J Pediatr* 1994;125:847-52. doi:10.1016/S0022-3476(05)81997-3

53 Wells JC, Cole TJ. Height, adiposity and hormonal cardiovascular risk markers in childhood: how to partition the associations?*Int J Obes (Lond)* 2014;38:930-5. doi:10.1038/ijo.2014.24

54 Weber DR, Leonard MB, Shults J, Zemel BS. A comparison of fat and lean body mass index to BMI for the identification of metabolic syndrome in children and adolescents. *J Clin Endocrinol Metab* 2014;99:3208-16. doi:10.1210/jc.2014-1684

55 Gishti O, Gaillard R, Durmus B, et al. BMI, total and abdominal fat distribution, and cardiovascular risk factors in school-age children. *Pediatr Res* 2015;77:710-8. doi:10.1038/pr.2015.29

56 Nightingale CM, Rudnicka AR, Owen CG, et al. Influence of adiposity on insulin resistance and glycemia markers among U.K. Children of South Asian, black African-Caribbean, and white European origin: child heart and health study in England. *Diabetes Care* 2013;36:1712-9. doi:10.2337/dc12-1726

57 L'Abée C, Visser GH, Liem ET, Kok DE, Sauer PJ, Stolk RP. Comparison of methods to assess body fat in non-obese six to seven-year-old children. *Clin Nutr* 2010;29:317-22. doi:10.1016/j.clnu.2009.11.009

58 Freedman DS, Sherry B. The validity of BMI as an indicator of body fatness and risk among children. *Pediatrics* 2009;124(Suppl 1):S23-34. doi:10.1542/peds.2008-3586E

59 Simmonds M, Llewellyn A, Owen CG, Woolacott N. Simple tests for the diagnosis of childhood obesity: a systematic review and meta-analysis. *Obes Rev* 2016;17:1301-15. doi:10.1111/obr.12462

60 Benn RT. Some mathematical properties of weight-for-height indices used as measures of adiposity. *Br J Prev Soc Med* 1971;25:42-50. doi:10.1136/jech.25.1.42

61 McCarthy HD, Cole TJ, Fry T, Jebb SA, Prentice AM. Body fat reference curves for children. *Int J Obes (Lond)* 2006;30:598-602. doi:10.1038/sj.ijo.0803232

62 Wells JC, Williams JE, Chomtho S, et al. Body-composition reference data for simple and reference techniques and a 4-component model: a new UK reference child. *Am J Clin Nutr* 2012;96:1316-26. doi:10.3945/ajcn.112.036970

63 Weber DR, Moore RH, Leonard MB, Zemel BS. Fat and lean BMI reference curves in children and adolescents and their utility in identifying excess adiposity compared with BMI and percentage body fat. *Am J Clin Nutr* 2013;98:49-56. doi:10.3945/ajcn.112.053611

64 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. CIR.97.18.1837

65 Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ* 2017;357:j2099. doi:10.1136/bmj.j2099

Supplementary information: supplementary tables 1-3, figures 1-7, and PRISMA flow diagram

Supplementary information: Excel spreadsheet showing fat mass calculator for children and adolescents