**SUPPLEMENTARY TABLES & FIGURES:**

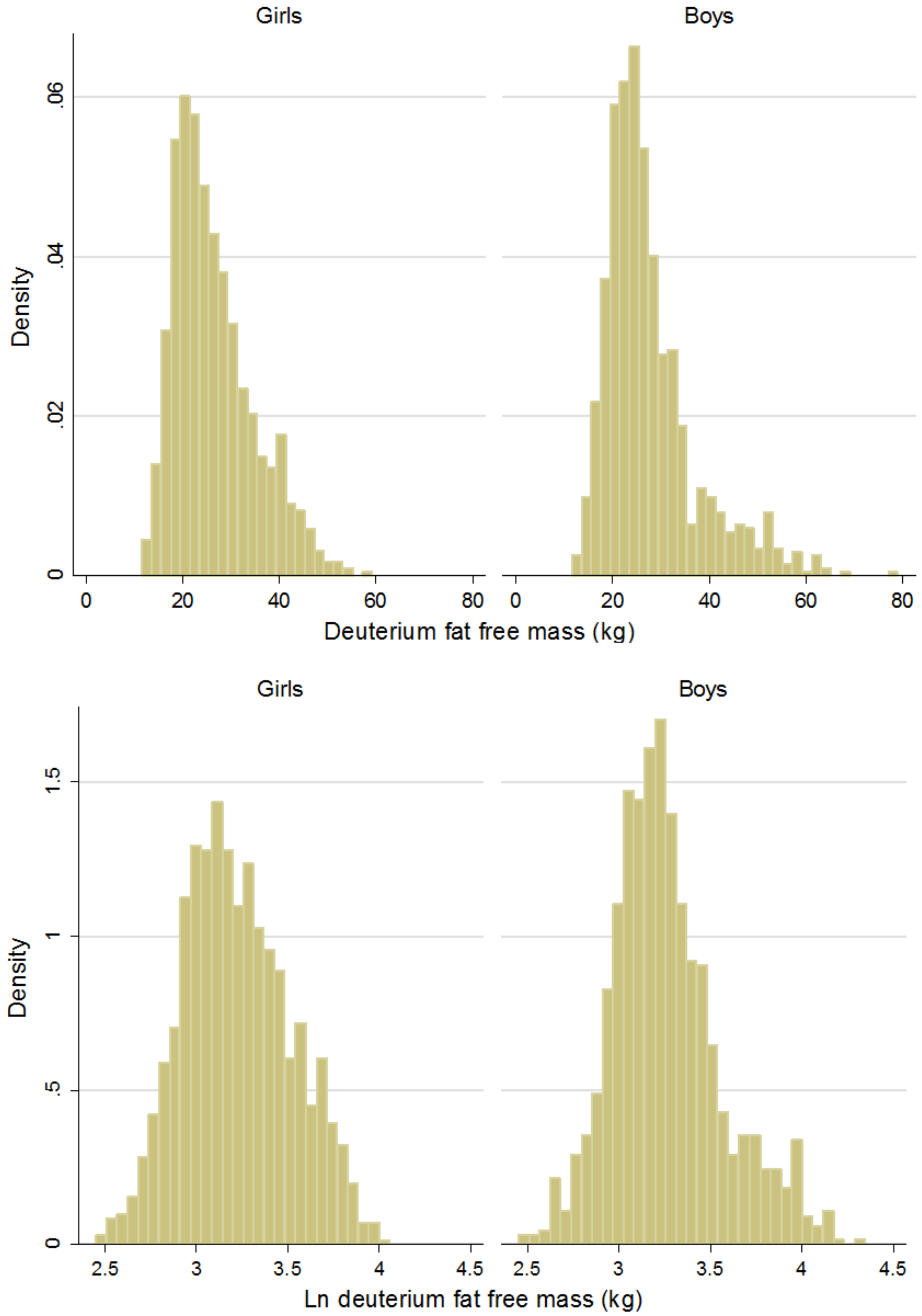
SUPPLEMENTARY TABLE 1: INFORMATION ON THE DERIVATION AND EXTERNAL VALIDATION STUDIES

| Study [ref] | Date of Study | Location | Setting | Age Group | Response Rate | Ethnicity Reporting |
|---|---|---|---|---|---|---|
| **Derivation Datasets:** | | | | | | |
| 1) The assessment of Body Composition in Children (ABCC) Study [17] | 2011-2012 | London | Primary schools | 8-10 years | 64% | Reported by parent |
| 2) The East London Bioelectrical Impedance (ELBI) [23] | 2008-2009 | London | Secondary schools, weight management clinic and volunteers recruited by advertisement | 11-15 years | Not Applicable | Self-reported |
| 3) The Reference Child (RC) [24] | 2001-2010 | London | Volunteers recruited by advertisement | 4-22 years | Not Applicable | Self-reported or reported by parent |
| 4) The Size and Lung function in Children (SLIC) Study [25] | 2012-2013 | London | Primary Schools | 5-11 years | 52% | Reported by parent |
| | | | | | | |
| **Validation Dataset:** | | | | | | |
| Avon Longitudinal Study of Parents and Children (ALSPAC) [26, 27] | 2002-2003 | Bristol | Population based advertising | 11-12 years | Not Applicable | Reported by parent |

SUPPLEMENTARY TABLE 2: ASSESSMENT OF CALIBRATION SLOPE AND CALIBRATION-IN-THE-LARGE IN TERMS OF FAT MASS AND FAT FREE MASS FROM INTERNAL-EXTERNAL CROSS-VALIDATION

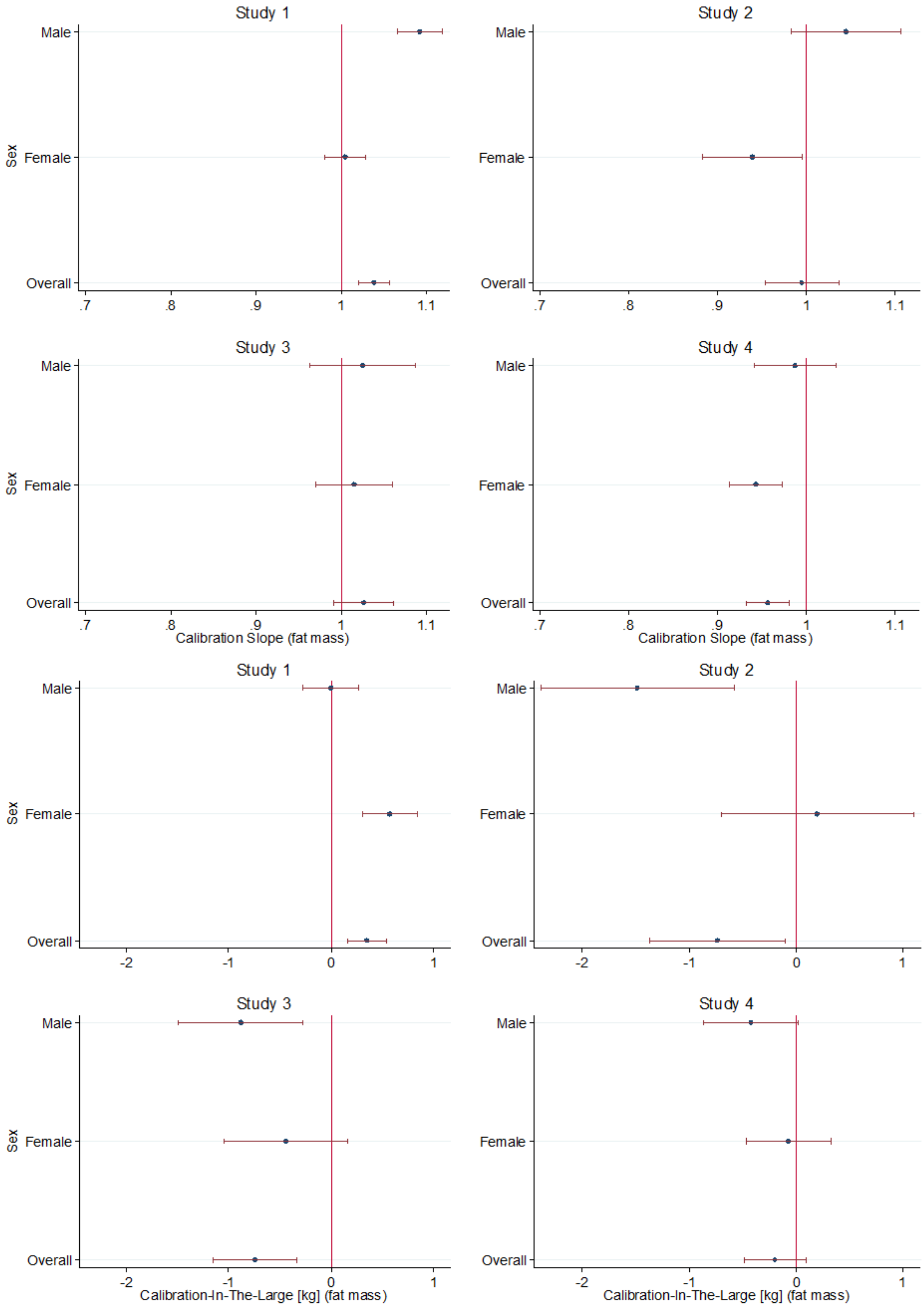
| **Study omitted for external validation** | **Fat Free Mass** | | **Fat Mass** | |
|---|---|---|---|---|
| | **Calibration-In-The-Large** | **Calibration Slope** | **Calibration-In-The-Large** | **Calibration Slope** |
| 1 | 0.05 (-0.45 to 0.55) | 0.97 (0.95 to 0.99) | 0.35 (0.16 to 0.54) | 1.03 (1.01 to 1.05) |
| 2 | 1.62 (0.34 to 2.90) | 0.98 (0.95 to 1.01) | -0.73 (-1.37 to -0.10) | 0.99 (0.95 to 1.03) |
| 3 | 1.02 (0.36 to 1.67) | 0.98 (0.96 to 1.01) | -0.74 (-1.15 to -0.33) | 1.02 (0.98 to 1.05) |
| 4 | -0.37 (-0.97 to 0.22) | 1.05 (1.02 to 1.07) | -0.20 (-0.48 to 0.09) | 0.95 (0.93 to 0.97) |
| Pooled | 0.46 (-0.30 to 1.21) | 1.00 (0.96 to 1.03) | -0.29 (-0.83 to 0.25) | 1.00 (0.95 to 1.04) |

SUPPLEMENTARY FIGURE 1: HISTOGRAM OF FAT FREE MASS (TOP) AND LN(FAT FREE MASS), BY SEX



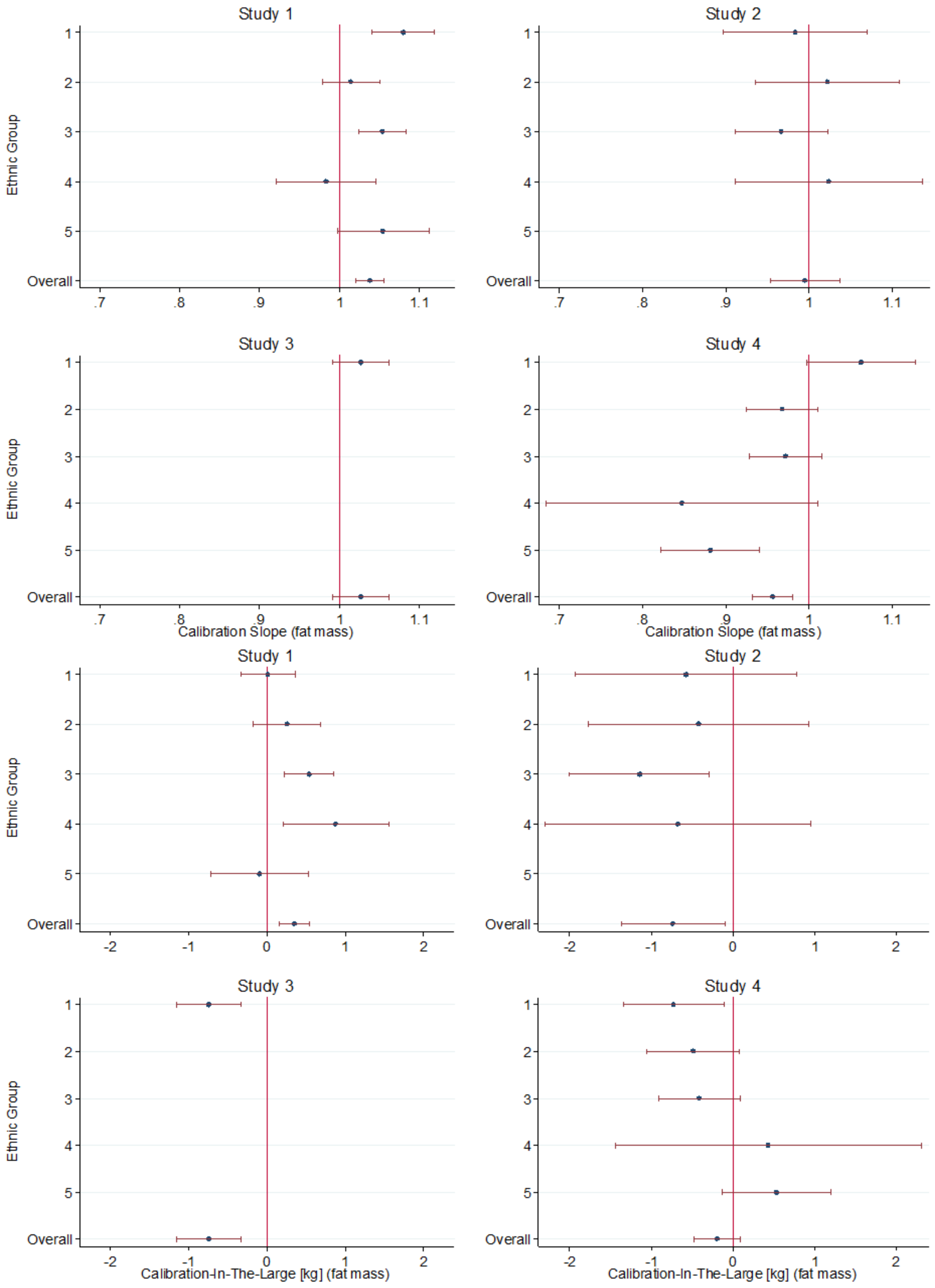

FOOTNOTE: LN = Natural logarithmic transformation

SUPPLEMENTARY FIGURE 2: CALIBRATION SLOPES AND CALIBRATION-IN-THE-LARGE (KG), BY SEX, FROM INTERNAL-EXTERNAL CROSS VALIDATION




FOOTNOTE: Study codes: 1 = The assessment of Body Composition in Children Study, 2 = The East London Bioelectrical Impedance, 3 = The Reference Child, 4 = The Size and Lung function in Children Study.

SUPPLEMENTARY FIGURE 3: CALIBRATION SLOPES AND CALIBRATION-IN-THE-LARGE (KG), BY ETHNIC GROUP, FROM INTERNAL-EXTERNAL CROSS VALIDATION




FOOTNOTE: Study codes: 1 = The assessment of Body Composition in Children Study, 2 = The East London Bioelectrical Impedance, 3 = The Reference Child, 4 = The Size and Lung function in Children Study. Ethnic group labels: 1 = White, 2 = Black, 3 = South Asian, 4 = Other Asian, 5 = Other.

**APPENDIX 1: COMPARISON OF ESTIMATING FM DIRECTLY OR INDIRECTLY**

The two approaches for estimating FM can be written as follows:

Equation 1: $FM_{direct}$ = f (height, weight, age, sex, ethnic group) + ε ; such that $\varepsilon \sim N(0, \sigma^2)$

Equation 2: $FM_{indirect}$ = weight - FFM = weight – [g(height, weight, age, sex, ethnic group) + μ] ; such that $\mu \sim N(0, \tau^2)$.

Therefore, possible estimates of FM are $\hat{f}$ or $weight - \hat{g}$ . It was clear from scatter plots of both FM and FFM, obtained from the DD method, plotted against height (Figure 1 below), that the variance in FFM with height was more homogeneous than for FM (i.e. $\tau^2 < \sigma^2$). Hence, a regression model for FFM (equation 2) should be estimated with greater precision than a model for FM (equation 1) because of the more homogenous relationship with height. As the variance of weight is ≈ 0 due to negligible measurement error, equation 2 ($weight - \hat{g}$ ) would be expected to provide more precise estimates of FM compared to estimates from $\hat{f}$ , making the indirect approach likely to be the preferred method for estimating FM.

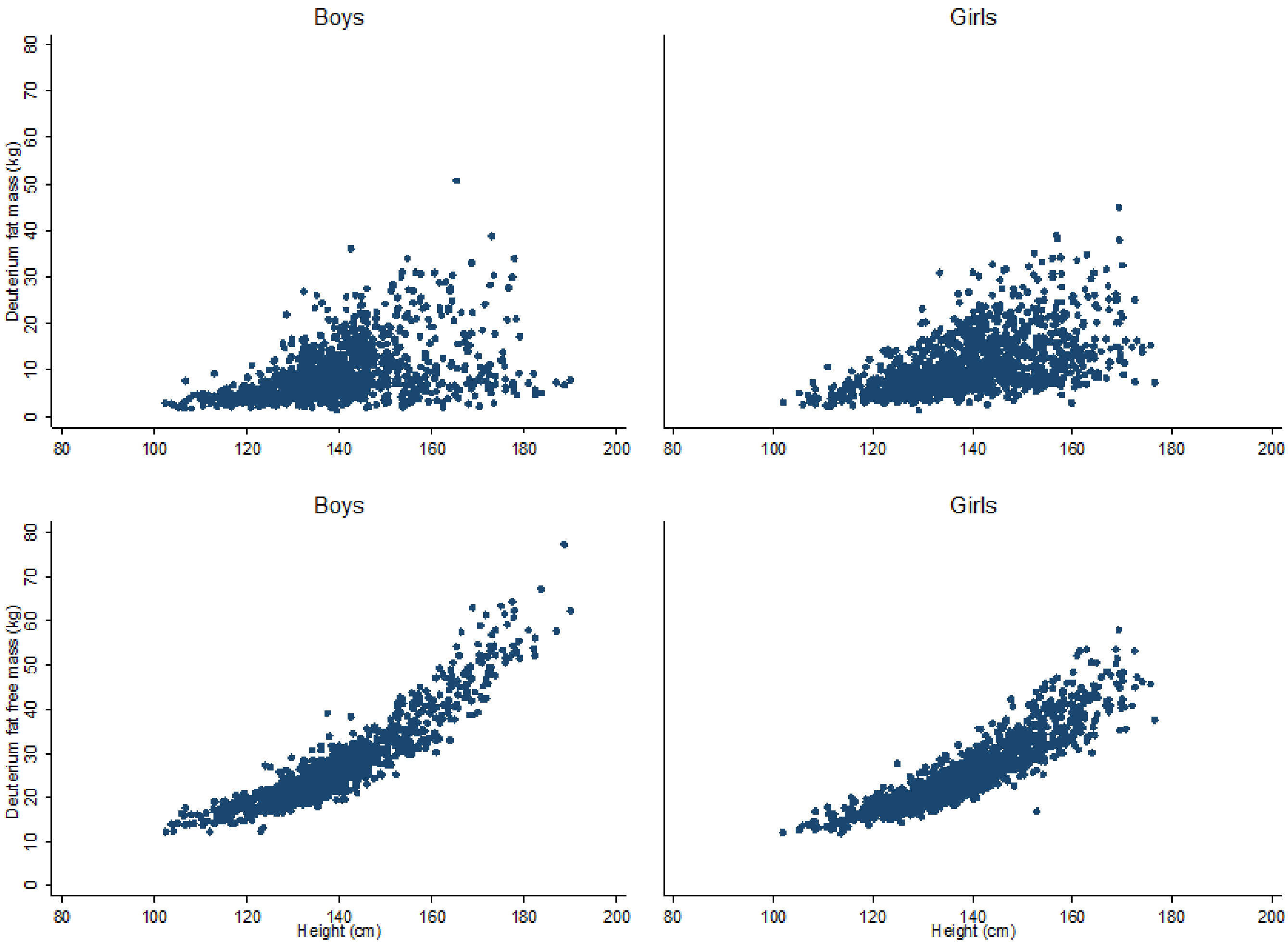



*Figure 1: Scatterplot of FM (top) and FFM (bottom) against height in boys (left) and girls (right)*

## APPENDIX 2: INTERNAL VALIDATION VIA BOOTSTRAPPING

The following steps were undertaken for the bootstrapping process[33]:

1. 1000 bootstrap samples were randomly selected from the entire DD derivation datasets, with replacement, such that the size of each bootstrap sample (N= 2375) was equal to that of the entire DD derivation dataset (N=2375). This selection process was stratified by sex, ethnic group and age to ensure that each bootstrap sample contained a representative sample of each of the subgroups.
2. The final developed prediction model was fitted within each bootstrap sample to obtain estimates of model performance based on $R^2$, the calibration slope and calibration-in-the-large (CIL).
3. Calculation of optimism-adjustments for $R^2$, the calibration slope and CIL:
    a. Values of $R^2$, the calibration slope and CIL were obtained, for each 1000 bootstrap sample
    b. The values from a. were subtracted from the original $R^2$, the calibration slope and CIL values from the original DD derivation datasets to obtain the level of optimism in each performance measure within each of the 1000 bootstrap samples
    c. The average of the differences in measures from b. across the 1000 bootstrap samples were determined
    d. Optimism-adjusted values of $R^2$ (denoted as $R^2_{adjusted}$), the calibration slope (denoted as Calibration $Slope_{adjusted}$) and CIL (denoted as $CIL_{adjusted}$) were obtained by subtracting the average of the differences (i.e. the value from c.) from the original $R^2$ and calibration slopes from the original DD derivation datasets