Sensitivity and specificity of body mass index and skinfold thicknesses in detecting excess adiposity in children aged 8–12 years

G. Bedogni†, L. Iughetti‡, M. Ferrari‡, M. Malavolti†, M. Poli†, S. Bernasconi§ and N. Battistini†

†Human Nutrition Chair, University of Modena and Reggio Emilia, Italy
‡Pediatrics Clinic, University of Modena and Reggio Emilia, Italy
§Pediatrics Chair, Parma University, Italy

Received 5 February 2002; in revised form 9 August 2002; accepted 14 August 2002

Summary. Primary objective: The study aimed to evaluate the sensitivity (SN) and specificity (SP) of body mass index (BMI) and skinfold thicknesses in detecting excess adiposity in children.

Research design: Cross-sectional.

Materials and methods: 986 children (500 females and 486 males) aged 10 ± 1 years (mean ± SD; range: 8–12 years) were studied. All underwent anthropometric measurements and bioelectrical impedance analysis (BIA). Dual-energy X-ray absorptiometry (DXA) was performed in 52 children to develop a population-specific algorithm for the assessment of fat-free mass (FFM) from BIA. The algorithm was applied to the remaining 934 children to estimate their FFM. Fat mass (FM) was obtained by subtracting FFM from weight (Wt). Values of FM:Wt were transformed in Z-scores and converted into 19 percentile categories (from 5 to 95 in steps of 5). The same procedure was performed with BMI and the log-transformed sum of four skinfold thicknesses (triceps, biceps, subscapular and suprailiac; lt-4SF). Excess adiposity was defined as a level of FM:Wt greater than the internally derived 85th percentile. SN and SP of each internally derived percentile of BMI and lt-4SF in detecting excess adiposity were calculated.

Results: In the pooled sample (n = 934), SN and SP were 0.39 and 0.99 for the 95th percentile of BMI, 0.65 and 0.95 for the 85th percentile of BMI, and 0.75 and 0.94 for the 85th percentile of lt-4SF.

Conclusions: BMI percentiles employed in the present study have a high SP but a low SN in detecting excess adiposity in 8–12-year-old children. The use of the sum of four skinfolds has the potential to increase the SN of a screening programme for excess adiposity in children of this age.

1. Introduction

Body mass index (BMI) has been proposed as an adiposity index in children because of its association with fat mass (FM) (Dietz and Bellizzi 1999, Maynard et al. 2001) and the possibility of using it to track adult BMI (Guo and Chumlea 1999). Another association that is being investigated is that between BMI and childhood metabolic disease (Bellizzi and Dietz 1999, Iughetti et al. 2000). As the association of BMI with FM is concerned, the traditional approach involves the use of regression models with FM as the dependent variable and BMI as the predictor variable (Pietrobelli et al. 1998, Bedogni et al. 2001). Another approach, which is more suitable for screening purposes, is based on the calculation of sensitivity (SN) and specificity (SP) of BMI in detecting excess adiposity (Lazarus et al. 1996). When applied to children, adolescents and young adults, BMI has generally a low SN and a high SP in detecting excess adiposity (Marshall et al. 1991, Lazarus et al. 1996, Warner et al. 1997, Reilly et al. 1999, Sardinha et al. 1999, Reilly et al. 2000, Sarria et al. 2001). The estimates of SN, however, are highly variable and there is some evidence that skinfolds may be more sensitive than BMI in detecting excess...
adiposity (Marshall et al. 1991, Sardinha et al. 1999, Sarria et al. 2001). This is to be expected because skinfolds are more directly associated than BMI with subcutaneous fat (Norgan 1991).

The available studies have employed different definitions of excess adiposity and study samples of different size and composition. As the technique employed to measure (or estimate) FM is concerned, this has been dual-energy X-ray absorptiometry (DXA) (Lazarus et al. 1996, Sardinha et al. 1999), body densitometry (Marshall et al. 1991, Sarria et al. 2001) and bioelectrical impedance analysis (Reilly et al. 1999). As discussed in detail by Lazarus et al. (1996), the variable of interest in this kind of studies is not the absolute value of FM but its ranking. Provided that the employed technique is able to produce an accurate ordering of FM%, the receiver–operator characteristic (ROC) curves obtained in different samples should be similar (Lazarus et al. 1996). However, even using a precise technique such as DXA, Lazarus et al. (1996) found wide 95% confidence intervals (95% CI) associated with the screening of excess adiposity from BMI and suggested that a larger number of children (~900 vs 230) were needed to produce an estimate of variability narrow enough for epidemiological applications. They also noticed that reference body composition techniques cannot be employed with large numbers of children because of formidable logistic challenges (Lazarus et al. 1996). In these conditions, indirect body composition techniques may be helpful, provided that population-specific equations are used and calibration has been performed against an accepted method (Guo et al. 1996). Bioelectrical impedance analysis (BIA) has a great potential for the assessment of body composition during epidemiological studies because it is non-invasive, rapid and portable (Deurenberg 1994).

Aiming at comparing the SN and SP of BMI and skinfolds in detecting excess adiposity in a large sample (>900) of children, we developed a population-specific algorithm for predicting DXA-measured fat-free mass (FFM) from BIA and obtained FM by subtracting FFM from weight (Wt).

2. Materials and methods

2.1. Study design

The study was performed in 986 apparently healthy children (500 females and 486 males) aged 8–12 years. They represented a convenience sample enrolled in primary and secondary schools of Modena and Parma (Italy). A subsample of 52 children underwent DXA at Parma University to develop a BIA population-specific algorithm for predicting FFM, with the informed consent of parents and under the approval of the local Ethical Committee. Informed consent to perform anthropometry and BIA was obtained from the parents of all children. Sample size was determined following the suggestion of Lazarus et al. (1996) that about 900 children were needed to produce an estimate of variability narrow enough for epidemiological applications. After the study of Reilly et al. (2000), which was however performed only in 7-year-old children (n = 4172), this is the largest study of this kind performed on children.

2.2. Anthropometry

Wt, height (Ht) and skinfolds (triceps, biceps, subscapular and suprailiac) were measured following the Anthropometric Standardization Reference Manual (Lohman, Roche and Martorell 1988). Body mass index was calculated as Wt (kg)/Ht (m)².
four measured skinfolds were summed to obtain a composite measure of subcutaneous fat (4SF) (Fiori et al. 2000).

2.3. BIA

Whole-body impedance (Z) was measured at a frequency of 50 kHz by using a four-polar impedance plethysmograph (Human-IM, Dietosystem, Milan, Italy) following standard procedures and after an overnight fasting (Deurenberg 1994). The impedance index (ZI), i.e. the Ht (cm)^2/Z (Ω) ratio, was employed as the predictor variable in the BIA algorithm (Bedogni et al. 1997). The reproducibility of measurements with this instrument, as determined by within-day test–retesting by us, is between 1 and 3 Ω. As the generation of the BIA algorithm is concerned, we calculated that a sample of 50 subjects has a power of 100% to detect a slope of 0.70 at an alpha level of 0.0001 under the assumption of a SD of 8 cm^2/Ω for ZI and a SD of 5 kg for FFM. The accuracy of the BIA algorithm was evaluated using the adjusted determination coefficient (R^2_adj) and the total and percent root mean square error (RMSE) (Guo et al. 1996).

2.4. DXA

DXA is increasingly used in children because it is more rapid and precise than other body composition methods (Goran et al. 1996, Lazarus et al. 1996, Pietrobelli et al. 1998, Sardinha et al. 1999). By measuring the differential attenuation of X-rays at two different energies, DXA allows the separation of body mass into FM, lean tissue mass (LTM) and bone mineral content (BMC). The sum of LTM and BMC gives FFM, which was the variable of interest in this study. DXA measurements were performed using a Lunar DPX-L densitometer (Lunar Corporation, Madison, WI, USA; paediatric software version 1.5). The precision of LTM and BMC measurements, as determined by 3 repeated measurements on two of the children, was ≤ 2.0 and ≤ 1.0%, respectively.

2.5. Statistical analysis

4SF was the only variable of interest that did not follow the normal distribution. Because log-transformation of 4SF obtained a normal distribution, the log-transformed value (lt-4SF) was used for analyses. Between-sex comparisons were performed by unpaired t-tests. FFM estimated from BIA was subtracted from Wt to obtain FM. Because age explained only a minimal portion of FM:Wt, BMI and lt-4SF variance (R^2_adj ≤ 0.05, p < 0.01 for all), we did not correct FM:Wt for age. Values of FM:Wt were transformed in Z-scores and converted into 19 percentile categories (from 5 to 95 in steps of 5) (Lazarus et al. 1996). The same procedure was performed with BMI and lt-4SF. Excess adiposity was defined as a value of FM:Wt corrected for age greater than the internally derived 85th percentile (Lazarus et al. 1996, Sarria et al. 2001). SN and SP of each percentile of BMI and lt-4SF in detecting excess adiposity were calculated (Newmann et al. 2001). 95% CI for SN and SP were calculated using Wilson’s method (Newcombe and Altman 2000). ROC curves were obtained by plotting SN versus (1 – SP) (Newmann et al. 2001). Statistical significance was set to a value of p < 0.05 for all tests. Statistical analysis was performed using Statview 5.0.1 (SAS Institute, Cary, NC, USA) and SPSS 10 (SPSS Inc., Chicago, IL, USA) on a MacOS computer.
3. Results

An algorithm for the prediction of FFM from BIA was developed in 52 of the 968 study children:

$$\text{FFM} \ (\text{kg}) = 4.8 + 0.7 \times \text{Ht} \ (\text{cm})^2 / Z(\Omega)$$

ZI explained 95% of FFM variance ($R^2_{\text{adj}} = 0.95$) and the RMSE was 1.5 kg (6%). Regression residuals were normally distributed (0.0 ± 1.4, mean ± SD) and uncorrelated with age, Wt, Ht, BMI and lt-4SF ($p = \text{NS}$ for all). No interaction was found between ZI and sex (ZI × sex, $p = \text{NS}$), showing that the generated algorithm could be employed independently of sex.

The measurements of the 934 study children to whom the BIA algorithm was applied are given in table 1. Age, Wt, Ht and BMI were similar in males and females ($p = \text{NS}$); however, 4SF ($p < 0.0001$) and $Z$ ($p < 0.0001$) were higher in females. FM:Wt was higher and FFM lower in females than males ($p < 0.0001$ for both).

Values of FM:Wt were transformed in Z-scores and classified into 19 percentile categories (from 5 to 95 in steps of 5). The same procedure was performed with BMI and It-4SF. SN and SP of BMI and It-4SF in detecting excess adiposity defined as a value of FM:Wt greater than the internally derived 85th percentile, were calculated for each sex. Because the areas under the ROC curves did not differ in males and females neither for BMI nor for It-4SF ($p = \text{ns}$ for both), ROC curves were plotted for the pooled sample (figure 1).

Although the overall accuracy of BMI (AUC = 0.94, 95% CI 0.92–0.95) was similar to that of It-4SF (AUC = 0.93, 95% CI 0.91–0.95), It-4SF had generally a greater SN than the corresponding percentile of BMI. SN and SP were 0.39 (95% CI 0.32–0.47) and 0.99 (95% CI 0.98–0.99) for the 95th percentile of BMI, 0.65 (95% CI 0.57–0.72) and 0.95 (95% CI 0.94–0.97) for the 85th percentile of BMI and 0.75 (95% CI 0.68–0.81) and 0.94 (95% CI 0.92–0.95) for the 85th percentile of It-4SF.

4. Discussion

The IOTF Committee has suggested to consider ‘overweight’ a child with a BMI greater than the 95th percentile for age and ‘at risk of overweight’ one with a BMI greater than the 85th and lower than the 95th percentile for age (Bellizzi and Dietz 1999). In our children, the 95th percentile of BMI showed a high SP (0.99; 95% CI
0.98–0.99) but a low SN (0.39; 95% CI 0.32–0.47) in detecting excess adiposity. Even if our SN is higher and less variable than that of Lazarus et al. (1996) (0.29; 95% CI 0.15–0.47) and Sarria et al. (2001) (0.19; 95% CI 0.07–0.39) (who used the same cut-point of FM% adopted by us), this value is clearly too low for BMI to be employed to screen positive cases of overweight in our children. The 85th percentile of BMI had a better SN (0.65; 95% CI 0.57–0.72) and a reasonable SP (0.95; 95% CI 0.94–0.97). The corresponding SN obtained by Lazarus et al. (1996) is higher but more variable (0.71; 95% CI 0.53–0.85) and that obtained by Sarria et al. (2001) is lower and even more variable (0.50; 0.30–0.70). However, the general pattern of our ROC curve mirror that observed by Lazarus et al. (1996), confirming their prediction that similar patterns of SN and SP should be obtained using different samples of children and body composition methods.

Even if the AUC under the ROC curves of BMI and lt-4SF were similar, suggesting a similar overall accuracy as the screening of excess adiposity is concerned, lt-4SF had generally a greater SN than the corresponding percentile of BMI. In particular, the use of the 85th percentile of lt-4SF in our children was associated with 10% less false negatives and only 1% more false positives than that of the 85th percentile of BMI. This confirms and extends the observations made in less numerous samples using single skinfolds (Sardinha et al. 1999, Sarria et al. 2001) or their sum (Marshall et al. 1991). It should nonetheless be noted that values of lt-4SF higher than the 90th percentile were no longer superior to the corresponding values of BMI, probably because measurements of skinfolds in overweight subjects are increasingly less accurate and reproducible.

The main limitation of this study is that it was not performed in a representative sample of the population. Among the available studies, only that of Reilly et al. (2000) was performed in such a sample. Thus, there is clearly the need for testing the accuracy of BMI and other anthropometric indicators in representative population samples. We chose to use internally derived centiles to set the cut-points of FM%,
BMI and lt-4SF, as done by Lazarus et al. (1996) and Sarria et al. (2001). Moreover, like them, we used a value greater than the internally derived 85th percentile to define excess FM%. Epidemiologically, this seems a reasonable choice but there is no reason why other values should not be employed and there is no doubt that some of the discrepancies in the literature arise from the choice of different cut-points. Use of externally derived centiles may have produced different results. Ideally, reference values of BMI for children should be determined on the basis of their associated risk of disease but this research area is still in its infancy (Bellizzi and Dietz 1999, Iughetti et al. 2000, Bedogni et al. 2002).

The ideal screening test should be both highly sensitive and highly specific. However, this combination is rarely attained and one has more commonly to weigh the relative importance of SN and SP and set the cut-off value accordingly (Newmann et al. 2001). Besides the prevalence of the condition being screened, one has to consider the practical implications of her or his choice. According to the majority of available studies on children, adolescents and young adults, BMI at conventional cut-off points is highly specific but has low sensitivity in detecting excess adiposity (Marshall et al. 1991, Lazarus et al. 1996, Warner et al. 1997, Reilly et al. 1999, Sardinha et al. 1999, Sarria et al. 2001). A notable exception is the study of Reilly et al. (2000), performed however only in 7-year-old children, where the 92nd percentile of BMI showed a SN = 0.92 and a SP = 0.92.

Thus, the available evidence suggests that some children with excess adiposity will be missed by a screening performed with BMI. On the other hand, there is a low risk of being wrongly labelled as overweight. As discussed in detail by Lazarus et al. (1996), this could presently be accepted because of the lack of good longitudinal data on which to estimate long-term health consequences and in view of the limited options for effective intervention. According to our study, however, the use of the 85th percentile of lt-4SF instead of the 85th percentile of BMI may increase the SN of screening programmes of excess adiposity in children without any relevant loss in SP.

In conclusion, this study of a large number of children of both sexes aged 8–12 years: (1) confirms that BMI at conventional cut-off points has an high SP but a low SN in detecting excess adiposity, and (2) suggests that the use of 4SF may increase the SN of a screening procedure for excess adiposity in children.

References
Bedogni, G., Bollea, M. R., Severi, S., Trunjfo, O., Manzieri, A. M., and Battistini, N., 1997, The prediction of total body water and extracellular water from bioelectric impedance in obese children. European Journal of Clinical Nutrition, 51, 129–133.
Bedogni, G., Iughetti, L., Ferrari, M., Malavolti, M., De Simone, G., Fiori, G., Battistini, N., and Bernasconi, S., 2002, Waist circumference vs body mass index as a predictor of fasting blood insulin in severely obese children. Diabetes, Nutrition and Metabolism, 15, 160–164.
Bedogni, G., Pietrobelli, A., Heymsfield, S. B., Borghi, A., Manzieri, A. M., Morini, P., Battistini, N., and Salvioli, G., 2001, Is body mass index a measure of adiposity in elderly women? Obesity Research, 9, 17–20.
Bellizzi, M. C., and Dietz, W. H., 1999, Workshop on childhood obesity: summary of the discussion. American Journal of Clinical Nutrition, 70, 1735S–1755S.
Deurenberg, P., 1994, International consensus conference on impedance in body composition. Age and Nutrition, 5, 142–145.
Dietz, W. H., and Bellizzi, M. C., 1999, Introduction: the use of body mass index to assess obesity in children. American Journal of Clinical Nutrition, 70, 1235S–1255S.
Fiori, G., Facchini, F., Pettener, D., Rimondi, A., Battistini, N., and Bedogni, G., 2000, Relationship between blood pressure, anthropometric characteristics and blood lipids in low and high altitude populations from Central Asia. Annals of Human Biology, 27, 19–28.
Zusammenfassung. Ziele: Das Ziel der Studie bestand darin, die Sensitivität (SN) und Spezifität (SP) des Body Mass Index (BMI) und der Hautfaltendicken bei der Feststellung extremer Adipositas bei Kindern zu ermitteln.

Design: Querschnittsuntersuchung

Material und Methoden: 986 Kinder (500 Mädchen und 486 Jungen) im Alter von 10 ± 1 Jahren (MW ± SD; Spannweite: 8–12 Jahre) wurden untersucht. Bei allen wurden anthropometrische Messungen und Bioelektrische Impedanzanalyse (BIA) durchgeführt. Dual-Energy-X-Ray Absorptiometrie (DXA) wurde bei 52 Kindern durchgeführt, um einen populationsspezifischen Algorithmus für die Bestimmung der fetten Masse (FFM) anhand von BIA zu entwickeln. Der Algorithmus wurde für die restlichen 934 Kinder angewendet, um ihr FFM zu schätzen. Die Fettmasse (FM) wurde durch Subtraktion der FFM der fetten Masse (FFM) und der log-transformierten Summe von vier Hautfalten (Trizeps, Bizeps, subscapular, suprailliacal; It-4SF) vorgegangen. Eine extreme Adipositas lag bei einer Überschreitung des intern bestimmen 85.
Perzentils für die Ratio FM:Wt vor. Für jedes intern abgeleitete BMI- und lt-4SF-Perzentil wurden die Sensitivität und Spezifität bestimmt.

**Ergebnisse:** In der gesamten Stichprobe \( n = 934 \) betrug SN und SP für das 95. BMI-Perzentil 0.39 und 0.99, 0.65 und 0.95 für das 85. BMI-Perzentil und 0.75 und 0.94 für das 85. Perzentil der lt-4SF.

**Zusammenfassungen:** Die BMI-Perzentile, welche in der vorliegenden Studie angewendet wurden, haben eine hohe Spezifität aber eine geringe Sensitivität bei der Ermittlung extremer Adipositas bei 8–12 jährigen Kindern. Durch die Verwendung der Summe von vier Hautfalten bietet sich die Möglichkeit, die Sensitivität eines Screeningprogrammes für extreme Adipositas bei Kindern dieser Altersgruppe zu erhöhen.

**Résumé.** *Objectif premier:* Cette étude a pour but d’évaluer la sensibilité (SN) et la spécificité (SP) de l’indice de masse corporelle (IMC) et des plis cutanés, pour détecter l’excès d’adiposité des enfants.

**Type de recherche:** Transversale

**Matériel et méthode:** On a étudié 986 enfants (500 filles et 486 garçons) âgés de 10 ± 1 ans (moyenne ± ET; étendue de variation: 8 à 12 ans). Tous ont subi des mensurations anthropométriques et une analyse d’impédance bioélectrique (AIB). On a pratiqué une absorptiométrie par rayons X double (AXD) sur 52 enfants afin de développer un algorithme spécifique pour la mesure de la masse maigre (MM) à partir de l’AIB. L’algorithme a été appliqué aux 934 enfants restants afin d’estimer leur MM. La masse grasse (MG) est obtenue en soustrayant la MM du poids (PDS). Les valeurs de MG:PDS sont transformées en Z scores et converties en 19 catégories de percentiles (de 5 à 95, de 5 en 5). La même procédure est appliquée pour l’IMC et le logarithme de la somme de quatre plis cutanés (biceps, triceps, sous scapulaire et supra iliaque; log4PC). L’excès d’adiposité est défini comme le niveau de MG:PDS plus grand que le 85ème percentile. On calcule la SN et la SP de chaque percentile d’IMC ainsi établi et du log4PC pour la détection de l’excès d’adiposité.

**Résultats:** Dans l’échantillon global \( n = 934 \), SN et SP sont de 0,39 et 0,99 pour le 95ème percentile d’IMC, 0,65 et 0,95 pour le 85ème percentile d’IMC et 0,75 et 0,94 pour le 85ème percentile de log4PC. L’excès d’adiposité est défini comme le niveau de MG:PDS plus grand que le 85ème percentile. On calcule la SN et la SP de chaque percentile d’IMC ainsi établi et du log4PC pour la détection de l’excès d’adiposité.

**Conclusions:** Les percentiles d’IMC employés dans cette étude ont une haute SP mais une basse SN pour détecter l’excès d’adiposité chez les enfants de 8 à 12 ans. L’utilisation de la somme des quatre plis cutanés a le potentiel d’accroître la SN du programme de détection.

**Resumen.**  *Objetivo principal:* El estudio pretende evaluar la sensibilidad (SN) y la especificidad (SP) del índice de masa corporal (BMI) y del espesor de los pliegues de grasa subcutánea en la detección del exceso de adiposidad en niños.

**Diseño de la investigación:** Transversal.

**Material y métodos:** Se estudiaron 986 niños (500 chicas y 486 chicos) de 10 ± 1 años (media ± SD; rango: 8–12 años). En todos los se tomaron medidas antropométricas y se les efectuó un análisis de impedancia bioeléctrica (BIA). A 52 niños se les realizó además una absorciometría dual de rayos X (DXA) con el fin de desarrollar un algoritmo poblacional específico para la estimación de la masa magra (FFM) por BIA. El algoritmo se aplicó a los 934 niños restantes para estimar su FFM. La masa grasa (FM) se obtuvo restando la FFM del peso (Wt). Los valores de FM:Wt se transformaron en puntuaciones Z y se convirtieron en 19 categorías percentilares (de 5 a 95, en intervalos de 5). El mismo procedimiento se siguió con el BMI y con el valor logarítmico de la suma de 4 pliegues (tríceps, bíceps, subescapular y supra iliaco; \( \text{lt-4SF} \)). El exceso de adiposidad se definió como un nivel de FM:Wt mayor que el percentil 85 derivado internamente. Para detectar el exceso de adiposidad, se calcularon los valores de SN y SP en cada percentil derivado internamente del BMI y del \( \text{lt-4SF} \).

**Resultados:** En el conjunto de la muestra \( n = 934 \), los valores de SN y SP fueron 0.39 y 0.99 para el percentil 95 del BMI, 0.65 y 0.95 para el percentil 85 del BMI, y 0.75 y 0.94 para el percentil 85 del \( \text{lt-4SF} \).

**Conclusiones:** Los percentiles del BMI empleados en el presente estudio poseen una elevada SP pero una SN baja para la detección del exceso de adiposidad en niños de 8–12 años de edad. El uso de la suma de 4 pliegues tiene la capacidad de incrementar la SN de un programa control del exceso de adiposidad en niños de esta edad.