Sensitivity and specificity of body mass index-based classification systems for overweight in children 7-10 years old

Danielle Biazzi Leal\textsuperscript{1,2}  
Filipe Ferreira da Costa\textsuperscript{1}  
Maria Alice Altenburg de Assis\textsuperscript{1,3}

Abstract – The objective of this study was to compare the sensitivity and specificity of BMI-based classification systems for detecting excess body fat in schoolchildren. A total of 2,795 schoolchildren aged 7 to 10 years were examined. Excess body fat was defined as the standardized residuals of sum of three skinfolds thickness ranking at or above the 90th percentile. The international BMI-based systems recommended by the International Obesity Task Force (IOTF) and the World Health Organization (WHO-2007) were evaluated on the basis of their sensitivity and specificity for detecting excess body fat and compared with a national BMI reference (Brazil-2006). The positive (LR+) and negative (LR-) likelihood ratios analysis was also used to compare the diagnostic accuracies of the three BMI criteria. The three classification systems presented moderately high sensitivity (78.4-98.6%) and specificity (75.9-91.6%) for both genders. Overall, the three classification systems showed both LR+ and LR- values consistent with a diagnosis of moderate evidence for overweight (LR+ above five and LR- below 0.2). The results showed that the three BMI classification systems can be used as screening instruments of excess body fat. However, the performance of the Brazil-2006 classification system was superior because it showed the best balance between the diagnostic accuracy indices.

Key words: Adiposity; Body mass index; Children; Sensitivity and Specificity.

Resumo – O objetivo deste estudo foi comparar a sensibilidade e especificidade de sistemas de classificação baseados no índice de massa corporal (IMC) na detecção do excesso de gordura corporal em escolares. Um total de 2.795 escolares com idade entre sete e dez anos foram examinados. O excesso de gordura corporal foi definido como os resíduos padronizados do somatório de três dobras cutâneas iguais ou superiores ao percentil 90. Os sistemas internacionais baseados no IMC recomendados pela International Obesity Task Force (IOTF) e Organização Mundial de Saúde (OMS-2007) foram avaliados com base em sua sensibilidade e especificidade para detecção do excesso de gordura corporal e comparado com uma referência do IMC nacional (Brasil-2006). Análise das razões de verossimilhança positiva (RV+) e negativa (RV-) também foi utilizada para comparar as precisões diagnósticas dos três critérios do IMC. Os três sistemas de classificação apresentaram sensibilidade (78,4-98,6%) e especificidade (75,9-91,6%) moderadamente alta para ambos os sexos. No geral, os três sistemas de classificação apresentaram valores de RV+ e RV-, condizentes com um diagnóstico de evidência moderada para o sobrepeso (RV+ acima de cinco e RV- abaixo de 0,2). Os resultados mostraram que os três sistemas de classificação do IMC podem ser usados como instrumentos de rastreio do excesso de gordura corporal. Entretanto, o desempenho do sistema de classificação Brasil-2006 foi superior porque mostrou o melhor equilíbrio entre os indices de acurácia diagnóstica.

Palavras-chave: Adiposidade; Crianças; Índice de massa corporal; Sensibilidade e Especificidade.
INTRODUCTION

Several nutritional status classification systems of children and adolescents have been used to estimate the frequency of individuals at nutritional risk for having elevated body mass. The main issues that underlie the use of the Body Mass Index (BMI) to assess the nutritional status of people under 20 years old refer to the use of international or national reference population and the determination of critical values that separate healthy individuals from the unhealthy ones.

The evidences of the first issue of using these references to assess the nutritional status of children and adolescents in each country indicate that the use of international references can cause distortions if the BMI does not have the same properties in both populations. Critical values of the classification systems based on BMI-for-age used in many countries have been confronted with reference measurements for body fat. In the absence of a “gold standard” to measure body fat content, such as multicomponent models, these studies have typically used indirect methods (e.g. bioimpedance, skinfolds) to define the proportion of individuals with excess body fat on the basis of these measurements.

Previous studies on the diagnostic accuracy of BMI-based classification systems when screening individuals with excess body fat have used statistical analyses such as sensitivity and specificity, area under the Receiver Operating Characteristic (ROC) curve, and likelihood ratio. In most studies, a significantly higher sensitivity was reported for national references when compared with the references of the International Obesity Task Force (IOTF) or of the World Health Organization (WHO-2007), and the opposite for specificity. Other studies have reported similar values of specificity between the IOTF reference and national reference data.

Guidelines for treating and preventing pediatric obesity published by experts in the area have recommended the use of the BMI national reference data to assess overweight and obesity in children and adolescents. Therefore, the objective of this study was to evaluate the diagnostic accuracy of BMI reference values proposed for Brazilian children and adolescents (Brazil-2006) in discriminating excess body fat in a representative sample of schoolchildren from Florianópolis-SC and compare it with the references internationally recommended by IOTF and WHO-2007.

METHODOLOGICAL PROCEDURES

Sample

This is a cross-sectional study with a probability sample of 7-10 year old schoolchildren from elementary schools in the city of Florianópolis (Santa Catarina). The study was conducted from September to November 2002. Sampling details can be found in another publication. In summary, a representative sample of schoolchildren engaged in first to fourth grade from elementary schools in the city was selected from a stratified sampling per
The information included anthropometric data (body mass, height, skinfolds, and arm, waist, and hip circumference) and socioeconomic status of the family. This article covers the measures of weight, height and three skinfolds (triceps, subscapular, and medial calf) of the children.

Of the 3,522 children selected in first to fourth grade from elementary schools, the data of 209 were excluded because they did not have the age range of the study (<7.0 and >10.0 years), and another 377 were excluded due to the lack of information (child was absent or ill on the day of anthropometric assessment or refused to participate in the study). Parents signed a term of informed consent for their children to participate in this study, which was approved by the Ethics and Human Research Committee of the Federal University of Santa Catarina (protocol No. 037/02).

**Anthropometric Measurements**

The administrative department of each school provided information on age and gender. The BMI and the sum of three skinfolds thickness (ΣSFT) were selected to assess the nutritional status of the participants. The measurements of weight, height, and skinfold thickness were taken by trained evaluators using standard techniques recommended by Lohman et al.\(^{17}\) Anthropometric measurements were done with the children barefoot, wearing light clothes. Weight was measured by a portable digital scale with a capacity of up to 180 kg (MARTE, model PP), and height was measured with a tape measure attached to a wall without a baseboard. The BMI was calculated using the body weight (in kg) divided by height (in meters) squared.

The skinfold thickness measurement was conducted on three points on the body: triceps, subscapular, and medial calf using skinfold calipers (Cescorf) on the right side of the body. Measurements were performed twice and a third measurement was performed if skinfolds differed more than 2 mm. The average of the readings at each point or close to two readings was used for the analyses.

Due to the importance of age in body fat variation, values of the ΣSFT were modeled against age polynomials (age in full years, age\(^2\), and age\(^3\)) through the linear regression model weighted by the inverse of the variance. Using the model with the highest Pearson coefficient (R\(^2\)), standardized residuals (skinfold variation independent of the linear age effect) were estimated for each gender. Values greater than ±4 standard deviations of the standardized residuals values were excluded (50 boys, 91 girls), resulting in a final sample of 2,795 children (52% boys). The unconventional value ±4 standard deviations has been chosen to preserve, as much as possible, the sample heterogeneity. The values of the ΣSFT (standardized residuals) ranking at or above the 90th percentile were adopted as the reference standard for the classification of excess body fat. Comparisons between the 90th percentile and the 85th/95th percentiles of the ΣSFT (standardized residuals) resulted in a variation in the number of misclassifications, as expected. However, the 90th percentile was chosen because the diagnostic accuracy of the IOTF, WHO-2007, and Brazil-2006 cut-offs showed the
same trend when considering the three critical values for excess body fat (data not shown). Although we are unsure whether 10% of the measure of body fat specifically identifies children with higher medical complications or health risks related to excess body fat, it is a statistical definition which is consistent with other practices in pediatrics, and has been employed in previous studies⁹,¹⁰,¹⁸,¹⁹.

**Statistical Analysis**

Individuals were classified as overweight (including obesity) according to three BMI-based classification systems: IOTF, WHO-2007, and Brazil-2006. Excess body fat was defined based on the value of the 90th percentile of the body fat reference measurement. Sensitivity was defined as the percentage of children with excess body fat (children in the top 10% of body fat distribution based on ΣSFT-standardized residuals) classified as overweight by BMI. Specificity was defined as the percentage of children without excess body fat (children not in the top 10% of body fat distribution based on ΣSFT-standardized residuals) classified as non-overweight by BMI.

Positive [sensitivity/(1-specificity)] and negative [(1-sensitivity)/specificity] likelihood ratio stratified by gender were calculated to express how many times more (or less) likely to find a result of a test in sick people compared to those who do not have the disease. In the context of this study, it indicated how many times more likely is the diagnosis (with or without overweight) according to the three criteria of the BMI in individuals with excess body fat compared to those without excess body fat. Positive likelihood ratio (LR+) above 10 and negative likelihood ratio (LR-) below 0.1 has provided diagnostics with convincing evidences, while values above five and below 0.2 indicate moderate evidences of diagnosis²⁰.

The level of significance was set at p<0.05. Statistical analyses were performed using SPSS 10.0 (Statistical Package for Social Sciences) and StATA version 10.0 (StataCorp, Lakeway Drive College Station).

**RESULTS**

The anthropometric characteristics of the individuals and the prevalence of overweight (including obesity) according to the references analyzed are shown in Table 1, stratified by gender. Compared to girls, boys were significantly taller, heavier, and had a higher mean BMI. Girls had a significantly higher mean of the ΣSFT than boys. The lower frequencies of overweight were obtained with the IOTF reference for both genders. According to the IOTF and WHO-2007 references, boys were significantly more likely than their female counterparts to be overweight. No statistical difference was observed between the genders for overweight prevalence according to the Brazil-2006 reference.

Table 2 shows the sensitivity, specificity, LR+ and LR- of BMI cut-offs for overweight of the IOTF, WHO-2007, and Brazil-2006 references. For both genders, the IOTF classification system showed the highest values of
specificity; however, its sensitivity was moderate for girls (78.4%). Classification systems of WHO-2007 and Brazil-2006 showed high sensitivity for both genders. The specificity of the WHO-2007 cut-offs was moderate for boys (75.9%). When comparing the genders, the BMI cut-offs, particularly those of the WHO-2007 and IOTF references, showed better sensitivity in boys and better specificity in girls, whereas the Brazil-2006 reference showed otherwise. The IOTF classification showed statistically significant differences between genders for diagnostic sensitivity and the classification of WHO-2007 for specificity. In contrast, no statistically significant difference between genders for sensitivity and specificity was found when the Brazil-2006 classification was used.

Table 1. Anthropometric measurements and frequencies of overweight (including obesity) using three references of BMI-for-age among 2795 schoolchildren, stratified by gender

<table>
<thead>
<tr>
<th></th>
<th>Girls (n=1341)</th>
<th>Boys (n=1454)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>8.5 (8.5; 8.6)</td>
<td>8.5 (8.5; 8.6)</td>
<td>0.426a</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>30.3 (29.9; 30.7)</td>
<td>31.3 (30.9; 31.7)</td>
<td>&lt;0.001a</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>133.2 (132.7; 133.7)</td>
<td>134.0 (133.6; 134.5)</td>
<td>0.010b</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>16.9 (16.8; 17.0)</td>
<td>17.2 (17.1; 17.4)</td>
<td>&lt;0.001a</td>
</tr>
<tr>
<td>∑SFT (mm)</td>
<td>31.5 (30.9; 32.1)</td>
<td>27.6 (27.0; 28.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Overweight (including obese)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOTF</td>
<td>15.4 (13.4; 17.3)</td>
<td>19.4 (17.4; 21.4)</td>
<td>0.005c</td>
</tr>
<tr>
<td>WHO-2007</td>
<td>22.7 (20.5; 25.0)</td>
<td>31.6 (29.2; 34.0)</td>
<td>&lt;0.001a</td>
</tr>
<tr>
<td>Brazil-2006</td>
<td>25.7 (23.3; 28.0)</td>
<td>23.6 (21.4; 25.8)</td>
<td>0.218a</td>
</tr>
</tbody>
</table>

BMI: Body mass index; ∑SFT: sum of three skinfolds thickness; IOTF: International Obesity Task Force; WHO: World Health Organization; 95%CI: 95% Confidence interval; a Student’s t test for equal variances; b Student’s t test for different variances; c Chi-square test.

The LR+ and LR- analysis suggested that a boy with excess body fat would be 4.1 to 8.3 times more likely to be classified as overweight than a boy without excess body fat depending on the classification system used, while a boy without excess body fat would just be 0.02 to 0.08 times more likely to be classified as overweight compared to a boy with excess body fat. Similarly, a girl with excess body fat would be 5.3 to 9.3 times more likely to be classified as overweight than a girl without excess body fat, while a girl without excess body fat would only be 0.06 to 0.24 times more likely to be classified as overweight compared to a girl with excess body fat (Table 2).

The IOTF classification system showed higher LR+ values than WHO-2007 and Brazil-2006 systems for both genders. However, they also showed greater LR- for girls, which would express lower accuracy in detecting children with truly excess body fat, causing higher rates of false negatives. In contrast, the WHO-2007 classification system had the lowest value of LR+, expressing less accuracy in detecting children with truly no excess body fat, causing higher rates of false positives.
DISCUSSION

The results of this study suggest that the three BMI classification systems have good performance in identifying excess body fat in children aged seven to ten years. However, the cut-offs of the national criterion showed the best balance between the diagnostic accuracy indices.

The classification systems based on BMI-for-age analyzed by previous studies showed differences in performance between genders. In general, better sensitivity was found for boys and better specificity for girls. Several studies compared the diagnostic accuracy between national and international BMI references, and found better sensitivity using the national references. The present study showed that the use of critical values for classifying overweight using the Brazil-2006 reference had better diagnostic sensitivity than the IOTF international reference (94.8% vs. 78.4% for girls and 94.5% vs. 93.1% for boys).

For specific comparisons with other observations we screened the literature to identify studies reporting sensitivity and specificity values of BMI cut-offs of at least one of the references used here and conducted in a population with a similar age to that in this study. In Brazil, the diagnostic accuracy of the same three classification systems for overweight used in the present study was assessed in 1570 schoolchildren aged seven to 12 years in the city of Paraíba. The sensitivity and specificity, using the body fat percentage as reference criteria (estimated by the sum of skinfolds) for IOTF were 86.3% and 94.7% for boys; 85.3% and 90.7% in girls, respectively. For WHO-2007, these figures were 64.7% and 97.9% in boys; 47.7% and 97.8% in girls. Using the Brazil-2006 classification, sensitivity and specificity values were 90.6% and 92.0% in boys; 97.2% and 84.8% in girls, respectively.

In Switzerland, a study done with a national sample of children aged six to 12 years showed similar sensitivity and specificity values for the IOTF criteria when compared with the results of the present study: using the body fat percentage, estimated from skinfolds, as reference standard, the IOTF sensitivity and specificity were respectively 78.8% and 94.4% in boys, 83.8% and 92.3% in girls. In a sample of British children seven years old, the sensitivity and specificity of the IOTF cut-offs for overweight, using the top 5% of the distribution of body fat percentage (estimated

Table 2. Values of sensitivity, specificity, and likelihood ratios for excess body fat for BMI-based references

<table>
<thead>
<tr>
<th></th>
<th>Boys</th>
<th></th>
<th></th>
<th>Girls</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sen</td>
<td>Spe (95%CI)</td>
<td>LR+</td>
<td>LR-</td>
<td>Sen (95%CI)</td>
<td>Spe (95%CI)</td>
<td>LR+</td>
<td>LR-</td>
</tr>
<tr>
<td>IOTF</td>
<td>93.1</td>
<td>(89.0; 97.2)</td>
<td>88.8</td>
<td>(87.1; 90.5)</td>
<td>8.31</td>
<td>0.08</td>
<td>78.4</td>
<td>(71.4; 85.3)</td>
</tr>
<tr>
<td>WHO-2007</td>
<td>98.6</td>
<td>(96.7; 100.0)</td>
<td>75.9</td>
<td>(73.5; 78.2)</td>
<td>4.09</td>
<td>0.02</td>
<td>92.5</td>
<td>(88.1; 97.0)</td>
</tr>
<tr>
<td>Brazil-2006</td>
<td>94.5</td>
<td>(90.8; 98.2)</td>
<td>84.3</td>
<td>(82.3; 86.2)</td>
<td>6.02</td>
<td>0.07</td>
<td>94.8</td>
<td>(91.0; 98.5)</td>
</tr>
</tbody>
</table>

IOTF: International Obesity Task Force; WHO: World Health Organization; Sen: Sensitivity; Spe: Specificity; LR+: Positive Likelihood Ratio; LR-: Negative Likelihood Ratio; 95%CI: 95% Confidence interval.
by hydrodensitometry) as reference standard, were respectively 90% and 92% for boys, 97% and 84% for girls\textsuperscript{25}. Differences in these values between studies may be explained by the children’s age, methods to measure body fat content, the definition of cut-offs for evaluation of excess body fat, and analytical approaches of the data.

The positive likelihood ratio had its highest values for the IOTF cut-offs. The higher the value, the stronger the association between having a positive outcome for overweight and having excess body fat. A likelihood ratio of 9.33 means that for the cut-off of the IOTF reference, the chance of a positive result being true is almost ten times greater than the chance of it being false. The negative likelihood ratio had its most significant values for WHO-2007 in the case of boys and for cut-off of the Brazil-2006 reference for girls. The lower the value, the stronger the association between having a negative outcome for overweight and not having excess body fat. Therefore, the three classification systems showed both LR+ and LR− values consistent with moderate diagnostic evidence (LR+ above five and LR− below 0.2). However, it is noteworthy that the LR+ of the WHO-2007 cut-offs for boys (4.09) and LR− of the IOTF cut-offs for girls (0.24) were borderline for what is considered ideal for public health interventions.

The final choice of which classification system should be adopted is conceptual, and includes issues related to ease of use, universal interpretation of results, and the ability of the cut-off values to predict the risk of future health problems. For international comparisons, the BMI references recommended by IOTF and WHO-2007 allows to identify prevalence rates globally acceptable and comparable. In Brazil, given the likely differences in health risks related to certain BMI values between populations, the national reference is likely to be more appropriate.

Some of the important aspects of this study include the large sample size and the inclusion of the analysis of the likelihood ratio, which goes beyond sensitivity and specificity. On the other hand, it should be noted that a reference measurement (gold standard) was not used for body fat due to the fact that this is a population-based study. The absolute measurement of skinfold thickness, despite the low reproducibility reported by other studies, is not based on predictive equations used to estimate the percentage of body fat, which are susceptible to errors\textsuperscript{26,27}. Furthermore, the use of standardized residuals of the measurement of skinfolds made this variable independent from the linear effect of age on variations in body fat. The skinfold thickness measurement has been used in other studies of diagnostic accuracy of the BMI as a reference measurement\textsuperscript{9,10,24,28,29}.

**CONCLUSION**

Data from this study indicate that the three BMI criteria have good ability to discriminate individuals with excess body fat from those without this condition. The international criteria established by the IOTF showed the lower diagnostic sensitivity of overweight for girls and the WHO-2007
criteria the lower specificity for boys. In addition, the performance of the Brazil-2006 classification system proved to be more suitable for national trials that aim to identify overweight children with similar characteristics in this study because it showed the best balance among the diagnostic accuracy indices.

REFERENCES


**Corresponding author**

Danielle Biazzi Leal  
Universidade Federal de Santa Catarina, Campus Universitário – Trindade.  
Programa de Pós-graduação em Educação Física, Centro de Desportos,  
88040-900 - Florianópolis, SC. Brasil  
E-mail: danibiazzi@yahoo.com.br