doi: 10.1111/j.1467-789X.2009.00709.x

Diagnostic in Obesity and Complications

Accuracy of simple clinical and epidemiological definitions of childhood obesity: systematic review and evidence appraisal

J. J. Reilly[1], J. Kelly[1,2] and D. C. Wilson[3]

[1]University of Glasgow Medical Faculty, Division of Developmental Medicine, Yorkhill Hospitals, Glasgow; [2]Scottish Intercollegiate Guidelines Network, 8–10 Hillside Crescent; [3]University of Edinburgh Medical Faculty, Section of Child Life and Health, Edinburgh Sick Children's Hospital, 20 Sylvan Place, Edinburgh, Scotland, UK

Received 5 August 2009; revised 9 October 2009; accepted 24 November 2009

Address for correspondence: Professor JJ Reilly, University of Glasgow Medical Faculty, Division of Developmental Medicine, Yorkhill Hospitals, Glasgow, Scotland G3 8SJ, UK. E-mail: jjr2y@clinmed.gla.ac.uk

Summary

The optimum means of defining obesity in children is unclear, creating variation in practice, and hindering obesity surveillance, prevention and treatment. This study aimed to review evidence on the use of body mass index (BMI) and waist circumference for diagnosis of high body fat content and adverse cardiometabolic risk factors in children and adolescents. A systematic literature review was carried out and evidence appraised using the Quality Assessment of Studies of Diagnostic Accuracy in Systematic Reviews method. Literature searching began following the last systematic review of this topic (end 2001) and collected evidence in MEDLINE and EMBASE in 0- to 18-year-olds that compared the accuracy of BMI vs. waist circumference and compared BMI interpreted relative to national reference data vs. BMI interpreted relative to Cole/International Obesity Task Force international reference data. Ten studies compared diagnostic accuracy of BMI vs. waist circumference: they reported no improved identification of adverse cardiometabolic risk profiles from waist circumference over that provided by high BMI. Eight studies compared BMI with national reference data vs. the international approach: 5/8 found significantly poorer accuracy (lower sensitivity) using BMI with the international approach; 3/8 found similar sensitivity; in 7/7 studies that compared specificity this was similar. In conclusion, the present review provides no compelling evidence for use of either high waist circumference or BMI interpreted using the International Obesity Task Force approach in preference to the use of national BMI percentiles for the identification of children and adolescents with excess fatness and adverse cardiometabolic risk profile.

Keywords: Body mass index, child, obesity diagnosis, waist circumference.

obesity reviews (2010) **11**, 645–655

Introduction

Surveillance, prevention and treatment of childhood and adolescent obesity require methods of defining obesity that are simple enough to be practical in most clinical and public health settings, but are also valid (1,2). The WHO defines obesity as a level of body fatness sufficiently high as to increase risk of morbidity or mortality (3). Valid methods of defining or diagnosing obesity in childhood and adolescence would therefore be able to identify the fattest individuals correctly and be able to identify those with greatest risk of morbidity (1). Systematic reviews have found that cardiometabolic risk factors are associated strongly with childhood obesity (4,5). As the cardiometabolic consequences of childhood obesity are of great clinical and public health concern (6), the optimal simple definition of child and adolescent obesity should be able to identify those with cardiometabolic comorbidity, which in paediatrics means those with adverse cardiometabolic risk factor profiles (1,5,6).

A systematic review published in 2003, which reviewed evidence published up to the end of 2001, concluded that defining obesity as a high BMI for age and sex using national population reference data (e.g. at or above the 95th percentile for BMI) identified the fattest children adequately, i.e. with moderately high sensitivity and high specificity (4). Absolute values of BMI are lower in children and adolescents than in adults and vary by age and sex, so have to be interpreted in an age and sex-specific manner, often using percentiles (1,2). After the publication of the 2003 review, two alternative simple approaches to defining obesity in youth have become popular: waist circumference (7) or indices based on waist circumference, such as waist : height ratio, BMI interpreted relative to international 'Cole/IOTF' criteria rather than national percentile cut-offs (8). These alternatives are being used widely in research, in clinical management of paediatric obesity and in population surveillance. However, at the time of the 2003 systematic review (4), no studies had determined the diagnostic accuracy of waist circumference as a measure of obesity in youth, and only a single study had compared the diagnostic accuracy of the BMI defined using the International Obesity Task Force (IOTF) approach vs. that using BMI with national reference data within the same sample (9). Furthermore, we are not aware of any systematic reviews of this topic that attempt to update the evidence since the end of 2001.

Recent guidelines on the management of obesity in youth (10–14) have recommended use of the BMI with national reference data and percentiles consistently, and have consistently recommended *not* using the newer alternatives for clinical applications because of a perceived lack of empirical evidence on their accuracy (10,11,13). However, preliminary literature searching for the present study showed that the amount of evidence on the use of the 'newer' definitions of paediatric obesity, BMI interpreted relative to international criteria using the IOTF method and use of waist circumference, has increased rapidly in recent years. The present study therefore aimed to identify the optimal simple definition of child and adolescent obesity by providing a systematic review and critical appraisal of recent evidence from studies that compared the accuracy of (i) BMI for age vs. waist for age and (ii) national vs. international definitions of excessive BMI for age.

Methods

Literature searching

The search strategy in MEDLINE is given below:

1. exp body mass index/;
2. bmi.tw.;
3. body mass index.tw.;
4. (waist circum$ or WC).tw.;
5. IOTF.tw.;
6. international obesity task force.tw.;
7. or/1–6;
8. exp Obesity/;
9. exp body weight/or exp overweight/;
10. exp Body Fat Distribution/;
11. (obes$ or overweight$).tw.;
12. or/8–11;
13. 7 and 12;
14. limit 13 to yr = '2002–08';
15. exp Adult/;
16. 14 not 15.

A similar search was run in EMBASE. The search results were not limited using a search filter, as these are generally considered inefficient in retrieving diagnostic studies (15,16).

We searched in MEDLINE and EMBASE from January 2002 to January 2008. We began the literature search from January 2002 because this is where searching ended in the previous systematic review of the optimal simple means of defining child and adolescent obesity (4,5). Reference lists of all eligible papers identified in the search and reference lists from recent guidelines on paediatric obesity (11,13,14) were all checked manually for additional potentially eligible studies. An attempt was also made to find additional eligible studies by cross-checking the results of the literature search against previous systematic reviews of population screening for obesity in youth (a related but distinct issue; (17,18)) and against three recent narrative reviews of diagnosis and definitions of obesity in youth (1,2,19).

Inclusion and exclusion criteria

Studies were only included if study participants were in the age range 0–18 years and if 'diagnostic accuracy' of BMI and/or waist was reported (with summary statistics, such as sensitivity and specificity, area under the curve, predictive values, receiver operator characteristic analysis). Abstracts were excluded because they failed to provide the necessary level of detail on study findings and methods. Studies were excluded if they reported only *associations* between BMI, waist and fatness or cardiometabolic risk factors as the present review addressed the specific issue of the accuracy of classification of obesity based on BMI and waist circumference.

The following evidence was regarded as being central to the primary aim of the present review

1. Studies that compared the ability of BMI for age vs. waist for age (or other indices based on waist, such as waist : height ratio) to identify high body fatness and/or cardiometabolic risk factors within the same sample.

2. Studies that compared the ability of BMI for age, interpreted using national BMI reference data vs. the Cole/IOTF international definition based on BMI (6) to identify high body fat content and/or cardiometabolic risk factors within the same sample.

The previous systematic review of this topic (4) had established the BMI for age with use of national percentiles as an adequate simple means of defining obesity in children. However, to provide completeness in the present review, studies published after January 2002 that only assessed the accuracy of BMI for age in the identification of the fattest children in the population were also included in the present review as a secondary aim.

Design and methods used in eligible studies

Studies typically measured body fat content of participants using a direct method (e.g. dual energy X-ray absorptiometry), defined a proportion of participants as excessively fat on the basis of these measures, and then calculated the extent to which the optimal cut-off points in the BMI or waist for age distribution (optimal cut-offs usually defined by receiver operator characteristic analysis) classified such children correctly. The only 'gold standards' for measurement of body fatness in children and adolescents are the multicomponent models, based on measurement of total body water plus body density (three component model) or total body water plus body density plus body mineral content (four component model; (20)).

Studies that compared the ability of BMI for age vs. waist for age to identify those with adverse cardiometabolic risk factor profiles generally used composite measures of cardiometabolic risk (e.g. presence of two or more widely accepted cardiometabolic risk factors, with analyses excluding waist and BMI in all of the eligible studies). In many studies individual cardiometabolic risk factors were also used. The predictive validity and clinical significance of cardiometabolic risk factors in youth is unclear at present (6). However, the risk factors used in the studies reviewed here represent widely accepted and readily available indices of cardiometabolic risk, and they tend to track from childhood and adolescence into adulthood (5,6).

Assessment of study validity and study data extraction

Two of the three authors (JR, DW) appraised the quality of each diagnostic study independently, using the Quality Assessment of Studies of Diagnostic Accuracy in Systematic Reviews (QUADAS) tool (21). In all cases differences of opinion over study quality ratings were resolved by discussion between the two reviewers. The QUADAS tool consists of a list of standardized criteria, and papers are rated on a scale from high (++), medium (+), to low (−) quality on the basis of the number of criteria satisfied when comparing 'index tests' (in this case the simple definitions of obesity; (21,22)) with the 'reference standards' (in this case the measures of body fatness and/or cardiometabolic risk factor profiles). In brief, the assessment criteria used were assessed using a standard checklist (http://www.sign.ac.uk; Methodology Checklist 5, Diagnostic Studies) and were as follows: study participants should be reasonably representative of those who would receive the test in practice; study participant selection criteria should be described clearly; the reference standard used should classify the condition adequately; the whole sample or a randomly selected subsample should be used in the study; the period between reference standard and index test being used should be short; study participants should receive the same reference standard regardless of the index test; reference standard and index test should be measured independently; index test results should be interpreted without knowledge of the results of the reference standard; reference standard results should be interpreted without knowledge of the results of the index test; the index test should be described in sufficient detail; the reference standard should be described in sufficient detail.

Results

Search results

The searches produced 437 citations (Fig. 1) of which 59 were identified as potentially eligible after initial sifting of paper titles and examination of the abstracts. After retrieval of these 59 full papers for further examination, 27 eligible papers remained for review and appraisal in the present study.

Body mass index vs. waist for classification of high body fat content and/or cardiometabolic comorbidities

Ten eligible studies were available for appraisal, all compared BMI vs. waist circumference within the same sample. Nine of the studies compared the ability of BMI vs. waist to diagnose cardiometabolic risk factors, while the other study compared the ability of BMI vs. waist circumference to identify high body fat content (Table 1). The study that compared waist vs. BMI for diagnosis of high fatness used dual energy X-ray absorptiometry to measure fatness. Table 1 shows a moderate amount of evidence that addresses the comparison of BMI vs. waist, from a wide range of populations. The eligible evidence ranged from low to high quality when appraised using the QUADAS tool (Table 1). The absence of a 'gold standard' measure of body fatness weakened the study that compared BMI vs. waist for assessment of high body fatness. In the studies of

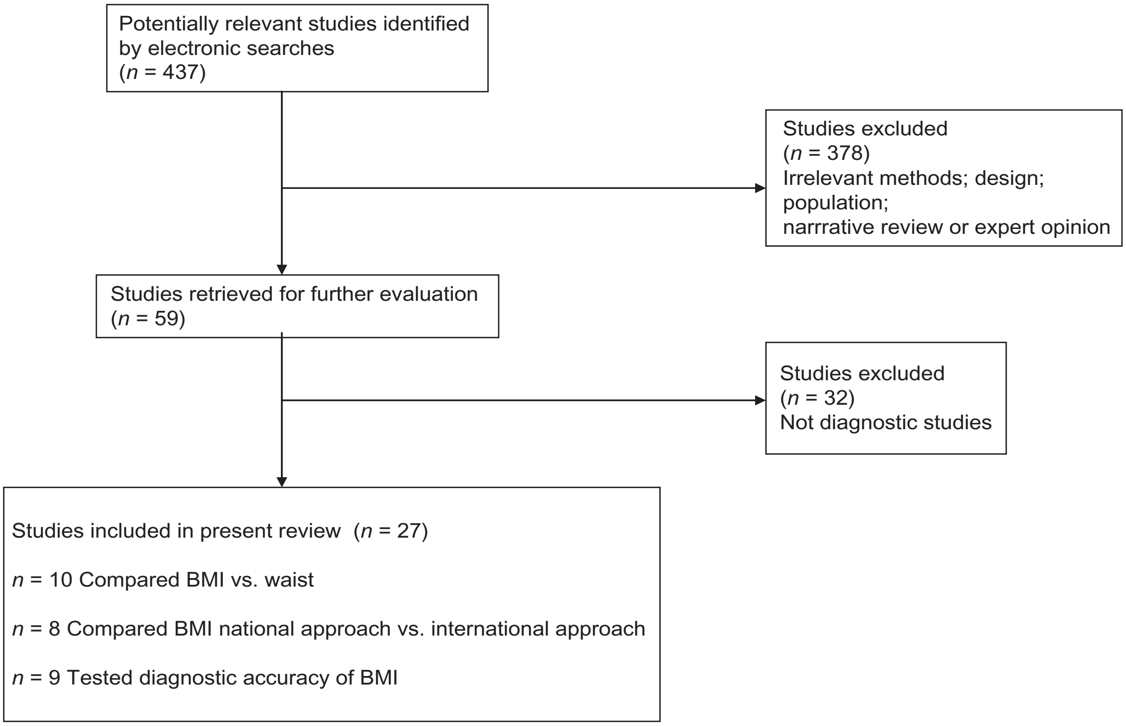

Figure 1 Study flow diagram. BMI, body mass index.

cardiometabolic risk factors the variables measured were generally appropriate (in the absence of a 'gold standard' for measurement of cardiometabolic risk in children and adolescents), and were independent of the 'index tests' of BMI and waist, but a number of weaknesses in study reporting (of recruitment, sampling and methodology) were common and these are summarized in Table 1. The consistency of the evidence was apparently very high: ability of waist and BMI for age to identify study participants with adverse cardiometabolic risk profiles was similar in all nine studies.

Body mass index using national reference data vs. International Obesity Task Force approach for classification of high body fat content and/or adverse cardiometabolic risk factors

Eight eligible studies were identified and appraised (Table 2) that compared, within the same samples, the accuracy of BMI to classify excessive fatness in children and adolescents using national reference data and percentiles vs. the IOTF approach. Seven of the eight eligible studies compared accuracy for excessive fatness only; one study compared accuracy of classification of both high body fatness and cardiometabolic risk. Studies were generally of low quality when appraised using the QUADAS tool (Table 2). The most common weakness was the absence of a 'gold standard' for the measurement of body fatness in study participants, but a number of other weaknesses were also apparent, particularly in study reporting (of recruitment, sampling and methodology), summarized in Table 2. Consistency of conclusions appeared to be relatively high: in 5/8 studies accuracy was reported to be significantly higher using national reference data and percentiles (significantly higher sensitivity) than when using the IOTF approach; in 3/8 studies accuracy was similar (and did not differ significantly) between the two approaches. In some studies significant differences in sensitivity were observed between the sexes for the IOTF approach that were not present when the national approach was used. In seven of the studies specificity of the two approaches was compared formally: specificity was similar and high for both approaches and in 7/7 studies did not differ significantly. The IOTF approach was not superior to the use of national percentiles for BMI in any of the eight eligible studies.

Evidence on the accuracy of body mass index for classification of excessive fatness using national reference data and percentiles

The secondary aim of the present review was to test whether the conclusions of the previous systematic review on the accuracy of high BMI for age using national reference data and percentiles (4) were supported by more recent evidence. Nine eligible studies, which addressed the issue of the accuracy of the national BMI approach only,

Table 1 Study and sample characteristics, results and quality rating: BMI vs. waist comparison for definition of excess fatness and cardiometabolic risk factors

Study	Participant characteristics	Sample size	Reference method	Index test	Comparison	Results	Conclusions	QUADAS rating comments*
Neovius *et al.* (2005) (23)	16- to 17-year-olds, Sweden	474	DXA	Fattest 5% of sample	BMI (kg m^{-2}) vs. waist (cm)	90–93% sensitivity both BMI & waist; 80–95% specificity both BMI & waist	Waist no improvement over BMI for high fatness	+
Moreno *et al.* (2002) (24)	7- to 15-year-olds, Spain	140	Cardiometabolic risk	Presence of 1 or more risk factors	BMI percentile vs. waist percentile	AUC for BMI 0.85; AUC for waist 0.87, NS	No improved diagnosis of cardiometabolic risk using waist	–, 1
Freedman *et al.* (2007) (25)	5- to 17-year-olds, USA	2498	Cardiometabolic risk	Presence of 1 or more risk factors	BMI *z*-score vs. waist : height ratio	Positive predictive values 0.13–0.38 for BMI *z*-score vs. 0.14–0.38 for waist *z*-score	No improved diagnosis of cardiometabolic risk using waist	+
Garnett *et al.* (2007) (26)	8-year-olds, Australia	172	Development of cardiometabolic risk 7 years after measurements of BMI and waist	Presence of 1 or more risk factors	BMI vs. waist *z*-score	Predictive validity risk NS different between BMI and waist	No improved diagnosis of cardiometabolic risk using waist	+
Kelishadi *et al.* (2006) (27)	6- to 18-year-olds, Iran	481	Cardiometabolic risk	Presence of single risk factors	BMI percentile vs. waist percentile	NS difference in sensitivity and specificity between waist and BMI	No improved diagnosis of cardiometabolic risk using waist	–, 1
Misra *et al.* (2006) (28)	14- to 18-year-olds, India	651	Cardiometabolic risk	Fasting insulin level	BMI (kg m^{-2}) vs. waist (cm) and waist–hip ratio	AUC risk NS different between the three indices	No improved diagnosis of cardiometabolic risk using waist	–, 1
Sung *et al.* (2007) (29)	6- to 12-year-olds, Hong Kong	1055	Cardiometabolic risk factors	Presence of 1 or more risk factors	BMI vs. waist percentiles	At optimal cut-points NS difference in sensitivity and specificity of waist vs. BMI	No improved cardiometabolic risk diagnosis using waist over BMI	–, 1
Ng *et al.* (2007) (30)	12- to 18-year-olds, China	2102	Cardiometabolic risk factors	Presence of ≥2 risk factors	BMI vs. waist percentiles	At optimal cut-points NS differences in AUC's between BMI vs. waist	No improved diagnosis using waist over BMI	+
Katzmaryk *et al.* (2004) (31)	5- to 8-year-olds, USA	2597	Cardiometabolic risk factors	Presence of ≥3 risk factors	BMI vs. waist percentiles	At optimal cut-points NS differences in AUC's between BMI vs. waist	No evidence of improved cardiometabolic risk using waist over BMI	+
Hirschler *et al.* (32)	4- to 13-year-olds, Argentina	530 subdivided by age and gender	Cardiometabolic risk factors	Presence of ≥3 risk factors	BMI vs. waist percentiles	AUC's between BMI vs. waist similar.	No improved diagnosis using waist over BMI	–, 1

*QUADAS rating comment 1 denotes limited description of study recruitment, sampling and/or limited description of study methodology.
AUC, area under the curve; BMI, body mass index; DXA, dual energy X-ray absorptiometry; NS, not significant; QUADAS, Quality Assessment of Studies of Diagnostic Accuracy in Systematic Reviews.

Table 2 Study and sample characteristics, results and quality rating: BMI used with national reference data vs. BMI using the international approach for definition of excessive fat mass and/or cardiometabolic risk factors

Study	Participant characteristics	Sample size	Reference method	Index test	Comparison	Results	Conclusions	QUADAS rating comments*
Neovius *et al.* (2004) (33)	16- to 17-year-olds, Sweden	474	Body density	High body fat percentage (morbidity related)	BMI percentile vs. international approach	Lower sensitivity of international approach, similar specificity	NS difference between definitions	–, 1
Moreno *et al.* (2006) (34)	13- to 17-year-olds, Spain	116 boys, 169 girls	DXA	High body fat percentage (top 15%)	BMI percentile vs. international approach	Similar sensitivity and specificity and specificity between methods	NS difference between the definitions	–, 1
Janssen *et al.* (2005) (35)	5- to 15-year-olds, USA	1709	(i) Cardiometabolic risk factors and (ii) later obesity	(i) Presence of 5 or more risk factors and (ii) obesity defined by BMI or waist after 13–24 years follow-up	BMI percentile vs. international approach	(i) AUC's identical for cardiometabolic risk factors and (ii) AUC's NS different for later obesity	NS difference between definitions	– for fatness; + for cardiometabolic risk diagnosis
Zimmerman *et al.* (2004) (36)	6- to 12-year-olds, Switzerland	2431	Skin-fold thickness	High body fat percentage (top 5% and 15% of fattest children)	BMI percentile vs. international approach	Sensitivity significantly higher with national BMI approach, specificity NS different	BMI national approach offers improved definition, high fat mass	–, 1
Yoo *et al.* (2006) (37)	8- to 12-year-olds, Korea	484 boys, 454 girls	Impedance	High body fat percentage (>35%)	BMI percentile vs. international approach	Sensitivity significantly higher with BMI national, similar specificity	Improved definition using national approach	–, 1
Wickramasinghe *et al.* (2005) (38)	5- to 15-year-olds, Australia	96 white Australian children; 42 Sri Lankan Australian children	Total body water	High body fat % (20% in boys, >30% in girls)	BMI percentile vs. international approach	Significantly higher sensitivity with national approach, identical and high specificity	Improved definition using national approach	–, 1
Oliveira *et al.* (2006) (39)	9- to 18-year-olds, Brazil	219 boys, 199 girls	DXA	High body fat % (top 5% of fattest children)	BMI percentile vs. international approach	Significantly higher sensitivity with national approach, NS difference in specificity	Improved definition using national approach	–, 1
Fu *et al.* (2003) (40)	6- to 11-year-olds, Singapore	321 boys	Impedance	High body fat % (fattest 5% of sample)	BMI national reference data vs. international approach	Significantly higher sensitivity with national approach, NS differences in specificity	Improved definition using national approach	–, 1

*QUADAS rating comment 1 denotes limited description of study recruitment, sampling and /or study methodology.

AUC, area under the cruve; BMI, body mass index; DXA, dual energy X-ray absorptiometry; NS, not significant; QUADAS, Quality Assessment of Studies of Diagnostic Accuracy in Systematic Reviews.

were identified and these are described in Table 3. The nine studies described in Table 3 were consistent in suggesting that high BMI for age provides an acceptable definition of excessive fatness in children and adolescents (high specificity; low to moderately high sensitivity depending on the precise cut-off point in the BMI distribution that is being used). The nine studies were of low quality when assessed using the QUADAS tool, and the most common weakness was the absence of a 'gold standard' for measurement of body fatness (see Table 3). In addition, the studies had a number of limitations in reporting of participant recruitment, sampling and methodology (Table 3).

Discussion

Main findings, implications and consistency with other evidence

The use of accurate yet simple means of defining obesity is central to the successful prevention, surveillance and treatment of obesity in childhood and adolescence. The present review found that the definition of obesity based on BMI for age with national reference data was superior to interpretation of BMI with the IOTF approach, with significantly higher sensitivity that did not differ significantly between the sexes and with similarly high specificity. In addition, the present review found that the use of obesity definitions based on high BMI for age was at least as accurate as definitions based on waist circumference for age, even for the diagnosis of adverse cardiometabolic risk factor profile. The present review therefore provides evidence that supports the use of the BMI for age with national reference data and percentiles and so supports recent expert committee recommendations that made tentative recommendations in favour of the use of BMI for age with national reference data to define paediatric obesity for clinical applications (11,13,14). The present review is also supportive of recent recommendations from expert committees that waist circumference should not replace BMI for clinical diagnostic purposes (11,13,14), and the IOTF approach should not be used in clinical practice (11,13,14).

One practical barrier to use of national reference data and percentiles for BMI for age in children and adolescents is availability: in many nations national reference data for BMI for age are not available. In the comparative studies included in the present review the studies from countries that did not have national BMI for age reference data tended to use reference data for BMI from the USA (Table 2) as the solution to this problem.

The decision as to which simple index of obesity is used should depend on the application, but users of simple obesity definitions should be aware of the implications of the choice of index that they make, and the present review should help inform such choices and illuminate the consequences of making particular choices. For clinical diagnostic applications, for example, the high specificity of the IOTF international approach to defining obesity is likely to be helpful, and the relatively low sensitivity may not be a problem, although evidence of differences in classification between the sexes may be more of a concern. Similarly, the IOTF approach is of great value in international comparisons of obesity prevalence (8,12). For within-nation surveillance of obesity prevalence, the low sensitivity of the IOTF approach is limited as it leads to an underestimation of the prevalence of excessive fatness in the population. Moreover, the significant difference in sensitivity between the sexes that has been reported when using the IOTF approach can also lead to spurious between-sex differences in childhood obesity prevalence, as in the UK for example (9,50). For many research applications in obesity the low sensitivity of the IOTF approach may be a concern because it can limit study power, in studies that depend on defining a proportion of study participants as obese this proportion will be much smaller when using the IOTF approach than when using the national approach, and these studies may become underpowered as a result.

The present review was consistent with the previous systematic review (4) in suggesting that a high BMI for age has low-moderately high sensitivity (depending on the cut-off point in the BMI distribution that is used), and high specificity. The evidence update (Table 3, 9 studies) was confirmatory of the evidence summary on this topic up to the end of 2001 (4).

The causes of the similarity in 'diagnostic accuracy' between BMI and waist circumference, and the differences in accuracy when using BMI with national percentiles vs. the IOTF approach are complex and go beyond the scope of the present study. Differences in the accuracy of classification of national and international BMI reference data may arise in part because of population differences in the pattern of BMI with age between nations. In addition, discrepancies in the ability of the IOTF obesity cut-offs to define obesity in boys and girls (large differences in sensitivity between the sexes) have been reported previously (1,9) and probably result from the decision to 'force' BMI curves through adult values of 25 and 30 kg m^{-2} at age 18 years, before growth has stopped. If the international BMI data are forced through adult values after growth has stopped the discrepancy in sensitivity between the sexes is abolished (51). The reasons why waist circumference does not appear to outperform BMI to define excessive fatness or cardiometabolic risk factors are perhaps more surprising, and in contrast to most evidence from adults (e.g. (52,53)). However, in children and adolescents there is evidence that total body fat and 'central' body fat are extremely highly correlated (54,55). The use of a simple index of central body fat, such as waist circumference, may therefore add little to the diagnostic accuracy that can be

Table 3 Study and sample characteristics, results and quality rating: diagnostic accuracy of BMI for age

Study, reference	Participant characteristics	Sample size	Body composition method, definition of excessive fatness	Index test	Results and conclusions concerning high BMI for age	QUADAS rating comments*
Bedogni *et al.* (2003) (41)	8- to 12-year-olds, Italy	986	Impedance	BMI percentile, internally generated	High specificity, low sensitivity	–, 1
Eto *et al.* (2004) (42)	3- to 5-year-olds, Japan	486	Impedance with >20% fat (boys) and >25% (girls)	BMI percentile	High specificity, moderate sensitivity	–, 1
Ghosh (2004) (43)	5- to 10-year-olds, India	450	Skin-fold thickness > 85th percentile, internally generated	BMI percentile	High specificity, moderate sensitivity	–, 1
Mast *et al.* (2002) (44)	5- to 7-year-olds, Germany	2286	Impedance and skin-fold thickness > 90th percentile for fatness, internally generated	BMI percentile, German reference data	High specificity and moderately high sensitivity	–, 1
Mei *et al.* (2007) (45)	5- to 18-year-olds , USA	1196	DXA-derived excess fatness, >95th percentile internally generated	BMI percentile, US reference data	High specificity and high sensitivity	–, 1
Mei *et al.* (2002) (46)	2- to 18-year olds, USA	190	DXA-derived 85th percentile for fatness, internally generated	BMI percentile, US reference data	High specificity and moderately high sensitivity	–,1
Taylor *et al.* (2003) (47)	8- to 15-year-olds, New Zealand	368	DXA-derived excess fatness: >25% in boys; >35% in girls	BMI percentiles, internally generated	High specificity and moderately high sensitivity	–,1
Field *et al.* (2003) (48)	5- to 18-year-olds, USA	596	DXA-derived excess fatness	BMI percentiles, US reference data	High specificity and moderately high sensitivity	–,1
Gaskin & Walker (2007) (49)	7- to 12-year olds, Jamaica	610	Skin-fold thickness > 85th percentile, US reference data	International definition	High specificity, sensitivity low and differed between the sexes	–,1

*QUADAS rating comment 1 denotes limited description of study recruitment, sampling and/or methodology.
BMI, body mass index; DXA, dual energy X-ray absorptiometry; QUADAS, Quality Assessment of Studies of Diagnostic Accuracy in Systematic Reviews.

obtained using a more global index of body fat, such as BMI. While speculative, there is also some recent evidence that suggests that relationships between intra-abdominal fat mass and cardiometabolic risk profile may differ between adults and children (56).

Strengths and weaknesses of the review and the evidence base

One consistent weakness in the diagnostic studies reviewed was the absence of a 'gold standard' to measure body fat content. None of the studies reviewed used a multicomponent method of measurement of body fat content and this contributed to their low-intermediate ratings in quality appraisal. Applying gold standard multicomponent methods in large studies that aim to validate simple methods of defining obesity is clearly difficult. For cardiometabolic risk in childhood and adolescence no single or composite index is recognized as a 'gold standard' at present, but the studies reviewed here all used standard and widely accepted measures of cardiometabolic risk profile that were independent of BMI and waist circumference (i.e. blood pressure and blood-based measures), and so independent of the simple obesity index being compared in each study.

The issue of the optimal definition of overweight, as distinct from obesity, is important but goes beyond the scope of the present study. The searches conducted for the present study suggested that there is a lack of evidence on this topic at present, and further original research is required in order to identify whether waist circumference, BMI with national reference data or BMI with the IOTF approach provides the optimal simple definition of overweight in particular circumstances, such as surveillance and clinical practice.

The QUADAS tool (21) that was used to appraise study quality in the present study is widely used and recommended, but it does not emphasize sample size and power, this aspect of a study does not form part of the QUADAS grading. As a result, issues of study power had a negligible contribution to the assessment of study quality in the present study. Many of the studies reviewed were relatively small, particularly when samples were subdivided by age, sex and/or ethnicity (all of which can influence diagnostic accuracy; (1)). The relatively small sample sizes may have compromised the ability of studies to identify differences in accuracy between BMI vs. waist or BMI national approach vs. BMI international approach. The apparent consistency between studies, in finding no significant differences between the accuracy of BMI and waist for example, could therefore in part reflect a lack of power to detect such differences. Simple methods for assessing power when designing diagnostic studies are available (57) and should be used more widely in future research on the validation of simple means of defining or diagnosing child and adolescent obesity. Very few of the studies identified in the present review referred to QUADAS or to the Standards for Reporting of Diagnostic Accuracy initiative (58), and greater use of these resources would improve future research on the optimal simple means of defining or diagnosing childhood and adolescent obesity.

Setting up (and funding) new studies that are sufficiently large to answer all the major questions on definitions of childhood obesity with greater confidence may be problematic, but one practical solution to this problem may be to make greater use of existing cohort studies in which both simple 'diagnostic' indices and sophisticated outcome measures have been measured in large samples. Large and comprehensive cohort studies, such as the Avon Longitudinal Study of Parents and Children (9,59) and the Bogalusa Heart Study (25) have been productive in methodological research on obesity definitions to date, and represent a potentially valuable existing resource for future diagnostic research in childhood obesity.

Conclusions: meaning and generalizability of the present review

The present review provides no compelling evidence that either waist circumference or BMI using the IOTF approach is superior to BMI percentile with national reference data and percentiles in identifying children and adolescents with excess fatness and adverse cardiometabolic risk factor profiles. A wide range of child and adolescent populations were studied in the eligible papers, and the apparent consistency of conclusions across these populations suggests that conclusions of the review might be widely applicable.

Conflict of Interest Statement

None declared.

References

1. Reilly JJ. Diagnostic accuracy of the BMI for age in pediatrics. *Int J Obes* 2006; **30**: 595–597.

2. Must A, Anderson SE. BMI in children and adolescents: considerations for population based applications. *Int J Obes* 2006; **30**: 590–594.

3. WHO Technical Consultation. *Obesity: preventing and managing the global epidemic*. WHO Technical Report Series, WHO, Geneva, 2000.

4. Reilly JJ, Wilson ML, Summerbell CD, Wilson DC. Obesity diagnosis, prevention, and treatment: evidence-based answers to common questions. *Arch Dis Child* 2002; **86**: 392–394.

5. Reilly JJ, Methven F, McDowell ZC, Hacking B, Alexander D, Stewart L, Kelnar CJ. Health consequences of obesity: systematic review. *Arch Dis Child* 2003; **88**: 748–752.

6. Daniels SR, Greer F, Committee on Nutrition. Lipid screening and cardiovascular health in childhood. *Pediatrics* 2008; **122**: 198–208.

7. McCarthy HD, Ellis SM, Cole TJ. Central overweight and obesity in British youth age 11–16 years: cross-sectional surveys of waist circumference. *Br Med J* 2003; **326**: 624–627.
8. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *Br Med J* 2000 **320**: 1240–1243.
9. Reilly JJ, Dorosty AR, Emmett PM, ALSPAC Study Team. Identification of the obese child: adequacy of the BMI for clinical practice and epidemiology. *Int J Obes* 2000; **24**: 1623–1627.
10. Scottish Intercollegiate Guideline Network. Management of obesity in children and young people: a national clinical guideline. Scottish Intercollegiate Guidelines Network Guideline Number 69, 2003. [WWW document]. URL http://www.sign.ac.uk (accessed April 2009).
11. National Institute for Health and Clinical Excellence. *Obesity: the prevention, identification, assessment and management of overweight and obesity in adults and children*. NIHCE Clinical Guideline 43, December 2006.
12. Reilly JJ, Wilson DC. ABC of obesity: childhood obesity. *Br Med J* 2006; **333**: 1207–1210.
13. Barlow S, the Expert Committee. Expert Committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: summary report. *Pediatrics* 2007; **120**: s124–s192.
14. August GP, Caprio S, Fennoy I, Freemark M, Kaufman FR, Lustig RH, Silverstein JH, Speiser PW, Styne DM, Montori VM. Prevention and treatment of pediatric obesity: endocrine society clinical practice guideline based on expert opinion. *J Clin Endocrinol Metab* 2009; **93**: 4576–4599.
15. Ritchie G, Glanville J, Lefebvre C. Do published search filters designed to identify diagnostic test accuracy studies perform adequately? *Health Info Libr J* 2007; **24**: 188–192.
16. Leeflang MM, Scholten R, Rutjes AW, Reitsma JB, Bossuyt PM. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. *J Clin Epidemiol* 2006; **59**: 234–240.
17. Westwood M, Fayter D, Hartley S, Rithalia A, Butler G, Glaszion P, Bland M, Nixon J, Stirk L, Rudolf M. Childhood obesity: should primary school children be routinely screened?: a systematic review and discussion of the evidence. *Arch Dis Child* 2007; **92**: 416–422.
18. Whitlock EP, Williams SB, Gold R, Smith PR, Shipman SA. Screening and interventions for childhood overweight: a summary of evidence for the US Preventive Services Task Force. *Pediatrics* 2005; **116**: e125–e151.
19. Freedman DS, Ogden CL, Berenson GS, Horlick M. BMI and body fatness in childhood. *Curr Opin Clin Nutr Metab Care* 2005; **8**: 618–623.
20. Wells JC, Fewtrell MS. Measuring body composition. *Arch Dis Child* 2006; **91**: 612–617.
21. Reitsma JB, Bossuyt PM, Kleijnen J. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess* 2004; **8**: 25–29.
22. Cote AM, Brown MA, Lam E, Von Dadelszen P, Firoz T, Liston RM, Magee LA. Diagnostic accuracy of urinary spot protein: creatinine ratio for proteinuria in hypertensive pregnant women: systematic review. *Br Med J* 2008; **336**: 1003–1006.
23. Neovius M, Linne, Y, Rossner S. BMI, waist circumference, and waist: hip ratio as diagnostic tests for fatness in adolescents. *Int J Obes* 2005; **29**: 163–169.
24. Moreno LA, Pireda I, Rodriquez G, Fleta J, Sarria A, Bueno M. Waist circumference for the screening of the metabolic syndrome in children. *Acta Pediatr* 2002; **91**: 1307–1312.
25. Freedman DS, Kaln HS, Mei Z, Grummer-Strawn LM, Dietz WH, Srinivasan SR, Berenson GS. Relation of BMI and waist: height ratio to cardiovascular disease risk factors in children and adolescents: the Bogalusa Heart Study. *Am J Clin Nutr* 2007; **86**: 33–40.
26. Garnett SP, Baur LA, Srinivasan S, Lee JW, Cowell CT. BMI and waist circumference in mid-childhood and adverse cardiovascular risk clustering in adolescence. *Am J Clin Nutr* 2007; **86**: 549–555.
27. Kelishadi R, Gheiratmand R, Ardalan G, CASPIAN Study Group. Association of anthropometric indices with cardiovascular disease risk factors among children and adolescents: CASPIAN Study. *Int J Cardiol* 2007; **117**: 340–348.
28. Misra A, Madhavan M, Vikram NK, Pardey RM, Dhingra V, Luthra K. Simple anthropometric measures identify fasting hyperinsulinaemia and clustering of cardiovascular risk factors in Asian Indian adolescents. *Metabolism* 2006; **55**: 1569–1573.
29. Sung RYT, Yu CCW, Choi KC, McManus A, Li AMC, Xu SLY, Chan D, Lo AFC, Chan JCN, Fok TF. Waist circumference and BMI in Chinese children: cut-off values for predicting cardiovascular risk factors. *Int J Obes* 2007; **31**: 550–558.
30. Ng VWS, Kong APS, Chow Choi K, Ozaki R, Wong GWK, Yee-so W, Tong PCY, Sung RYT, Xu LY, Chan MHM, Ho CS, Lau CWK, Chan JCN. BMI and waist circumference in predicting cardiovascular risk factor clustering in Chinese adolescents. *Obesity* 2007; **15**: 494–503.
31. Katzmaryk PT, Srinivasan SR, Chen W, Malira RM, Bouchard C, Berenson GS. BMI, waist circumference, and clustering of cardiovascular disease risk factors in a biracial sample of children and adolescents. *Pediatrics* 2004; **114**: e198–e205.
32. Hirschler V, Maccallini G, Calcagno M, Anuda C, Jadzinsky M. Waist circumference identifies primary school children with metabolic syndrome abnormalities. *Diabetes Technol Ther* 2007; **9**: 149–157.
33. Neovius M, Linne YM, Barkeling BS, Rossner SO. Sensitivity and specificity of classification systems for fatness in adolescdents. *Am J Clin Nutr* 2004; **80**: 597–603.
34. Moreno LA, Blay MG, Rodriguez G, Blay VA, Mesana MI, Olivares JL, Fleta J, Sarria A, Bueno M, AVENA Zaragoza Study Group. Screening performances of the International Obesity Task Force BMI cut-off values in adolescents. *J Am Coll Nutr* 2006; **5**: 403–408.
35. Janssen I, Katzmaryk PT, Srinivasan SR, Chen W, Malina RM, Bouchard C, Berenson GS. Utility of childhood BMI in the prediction of adulthood disease: comparison of national and international references. *Obesity Res* 2005; **13**: 1106–1115.
36. Zimmerman MB, Gubeli C, Punterer C, Molinari L. Detection of overweight and obesity in a national sample of 6–12 year old Swiss children: accuracy and validity of reference values for BMI from the US CDC and the IOTF. *Am J Clin Nutr* 2004; **79**: 838–843.
37. Yoo S, Lee SY, Kim KN, Sung E. Obesity in Korean preadolescent school children: comparison of various anthropometric measurements based on bio-electrical impedance analysis. *Int J Obes* 2006; **30**: 1086–1090.
38. Wickramasinghe VP, Cleghorn GJ, Edmiston KA, Murphy AJ, Abbott RA, Davies PSW. Validity of BMI as a measure of obesity in Australian white Caucasian and Australian Sri Lankan children. *Ann Hum Biol* 2005; **32**: 60–71.
39. Oliveira FLC, Taddei JAAC, Escrivao MAMS, Cobayashi EF, Barros ME, Vitdo MR, Colugnati FAB, Ancona-Lopez F. Accuracy of obesity diagnosis in Brazilian adolescents: comparison of Cole *et al.* and Must *el al.* criteria with DXA percentage of fat mass. *Nutr Hosp* 2006; **21**: 484–490.

40. Fu WPC, Lee HC, Ng CJ, Tay YKD, Seow CJ, Siak JK, Hong CY. Screening for childhood obesity: international vs population-specific definitions, which is more appropriate? *Int J Obes* 2003; **27**: 1121–1126.
41. Bedogni G, Lughetti L, Ferrari M, Malawalti M, Poli M, Bernasconi S, Battistini N. Sensitivity and specificity of BMI and skinfold thicknesses in detecting excess adiposity in children aged 8–12 years. *Ann Hum Biol* 2003; **30**: 132–139.
42. Eto C, Komiya S, Nakao T, Kikkawa K. Validity of the BMI and fat mass index as an indicator of obesity in children aged 3–5 years. *J Physiol Anthropol Appl Human Sci* 2004; **23**: 15–30.
43. Ghosh A. ROC curve analysis in 5–10 year old Bengalee girls from Calcutta, India. *Ann Hum Biol* 2004; **31**: 364–369.
44. Mast M, Langnase K, Labitzke K, Bruse U, Preub U, Muller MJ. Use of BMI as a measure of overweight and obesity in a field study on 5–7 year-old children. *Eur J Nutr* 2002; **41**: 61–67.
45. Mei Z, Grummer-Strawn L, Pietrobelli A, Goulding A, Goran MI, Dietz WH. Validity of BMI compared with other body composition screening indexes for the assessment of body fatness in children and adolescents. *Am J Clin Nutr* 2002; **75**: 975–978.
46. Mei Z, Grummer-Strawn L, Wang J, Thornton TC, Freedman DS, Pierson RN, Dietz WH, Horlick M. Do skinfold measurements provide additional information to BMI in the assessment of body fatness among children and adolescents? *Pediatrics* 2007; **119**: e1306–e1313.
47. Taylor RW, Falorni A, Jones IE, Goulding A. Identifying adolescents with high percentage body fat: a comparison of BMI cut-offs using age and stage of pubertal development compared with BMI cut-offs using using age alone. *Eur J Clin Nutr* 2003; **57**: 764–769.
48. Field AE, Laird N, Steinberg E, Fallon E, Semega-Janeh M, Yanovski JA. Which metric of relative weight best captures body fatness in children? *Obes Res* 2003; **11**: 1345–1352.
49. Gaskin PS, Walker SP. Obesity in a cohort of black Jamaican children as estimated by BMI and other indices of adiposity. *Eur J Clin Nutr* 2007; **57**: 420–426.
50. Reilly JJ. Assessment of childhood obesity: national reference data or international approach? *Obes Res* 2002; **10**: 838–840.
51. Chinn S, Rona RJ. International definitions of overweight and obesity in children: a lasting solution? *Ann Hum Biol* 2002; **24**: 306–313.
52. Lean ME, Han TS, Deurenberg P. Predicting body composition from simple anthropometric measurements. *Am J Clin Nutr* 1996; **63**: 4–14.
53. Yusuf S, Hawken S, Ounpuu S, INTERHEART Study Investigators. Obesity and risk of myocardial infarction in 27 000 participants from 52 countries: case control study. *Lancet* 2005; **366**: 1640–1649.
54. Toschke AM, Martin RM, von Kries R, Wells JC, Smith GD, Ness AR. Infant feeding method and obesity: BMI and DXA measurements at 9–10 years of age from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Am J Clin Nutr* 2007; **85**: 1578–1585.
55. Benfield LL, Fox KR, Peters DM, Blake H, Rogers I, Grant C, Ness A. Magnetic resonance imaging of abdominal adiposity in a large cohort of British children. *Int J Obes* 2008; **32**: 91–99.
56. Maffeis C, Manfredi R, Trombetta M, Sordelli S, Storti M, Bennuzzi T, Bonadonna RC. Insulin sensitivity is correlated with subcutaneous but not visceral body fat in overweight and obese pre-pubertal children. *J Clin Endocrinol Metab* 2008; **93**: 2122–2128.
57. Carley S, Dosman S, Jones SR, Harvon M. Simple nomograms to calculate sample size in diagnostic studies. *Emerg Med J* 2005; **22**: 180–181.
58. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou P, Irwig LM, Lijmer JG, Moher D, Rennie D, De Vet HCW, STARD Group. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Chem* 2003; **49**: 1–6.
59. Reilly JJ, Dorosty AR, Ghomazideh NM, Sheriff A, Wells JC, Ness AR. Comparison of waist circumference percentiles versus BMI percentiles for diagnosis of obesity in a large cohort of children. *Int J Pediatr Obes* 2009 [Epub ahead of print].