Article:

Simmonds, M. orcid.org/0000-0002-1999-8515, Llewellyn, A. orcid.org/0000-0003-4569-5136, Owen, C. G. et al. (1 more author) (2016) Simple tests for the diagnosis of childhood obesity:A systematic review and meta-analysis. Obesity reviews. pp. 46-60. ISSN: 1467-7881

https://doi.org/10.1111/obr.12462

Pediatric Obesity

Simple tests for the diagnosis of childhood obesity: a systematic review and meta-analysis

M. Simmonds,[1] A. Llewellyn,[1] C. G. Owen[2] and N. Woolacott[1]

[1]Centre for Reviews and Dissemination, University of York, York, UK, and [2]Population Health Research Institute, St George's, University of London, London, UK

Received 22 April 2016; revised 20 July 2016; accepted 20 July 2016

Address for correspondence: M. Simmonds, Centre for Reviews and Dissemination, University of York, York YO10 5DD, UK.
E-mail: mark.simmonds@york.ac.uk

Summary

There is a need to accurately quantify levels of adiposity in order to identify overweight and obesity in children. This systematic review aimed to identify all diagnostic accuracy studies evaluating simple tests for obesity and adiposity, including body mass index (BMI), skin-fold thickness and waist circumference, compared against high-quality reference tests. Twenty-four cohort studies including 25,807 children were included. BMI had good performance when diagnosing obesity: a sensitivity of 81.9% (95% confidence interval [CI]: 73.0 to 93.8) for a specificity of 96.0% (95% CI: 93.8 to 98.1). It was less effective at diagnosing overweight (sensitivity: 76.3%, 95% CI: 70.2 to 82.4; specificity: 92.1% 95% CI: 90.0 to 94.3). When diagnosing obesity, waist circumference had similar performance (sensitivity: 83.8%; specificity: 96.5%). Skin-fold thickness had slightly poorer performance (sensitivity: 72.5%; specificity: 93.7%). Few studies considered any other tests. There was no conclusive evidence that any test was generally superior to the others. BMI is a good simple diagnostic test for identifying childhood adiposity. It identifies most genuinely obese and adipose children while misclassifying only a small number as obese. There was no conclusive evidence that any test should be preferred to BMI, and the extra complexity of skin-fold thickness tests does not appear to improve diagnostic accuracy.

Keywords: BMI, Childhood obesity, diagnosis, meta-analysis.

Abbreviations: BMI, body mass index; CI, confidence interval; DXA, dual-energy X-ray absorptiometry; HSROC, hierarchical summary receiver operating characteristic (curve); QUADAS, quality assessment of diagnostic accuracy studies; RWt, relative weight; SD, standard deviation; SFT, skin-fold thickness; WC, waist circumference; WHtR, waist-to-height ratio; WHpR, waist-to-hip ratio.

obesity reviews (2016)

Introduction

Childhood obesity is an important public health issue (1). Childhood obesity can persist into adulthood (2–4) and so lead to an increased risk of many morbidities, including type II diabetes, cardiovascular disease and cancer (5–8). Identifying high adiposity in children (and hence overweight and obese individuals) is therefore important as these children are likely to go on to be obese adults at higher risk of morbidity (4,9).

Body mass index (BMI) is commonly used to measure adiposity, and hence to define obesity, but it has many problems. BMI does not measure the distribution of fat in the body and does not distinguish between adiposity and high muscularity. BMI does not perform well at the extremes of height (10). BMI may also be an imperfect measure to define ethnic differences in overweight or obesity in children: compared with children of white European ancestry, BMI underestimates adiposity among South Asian children (11,12) and overestimates adiposity in black African Caribbeans (12).

True adiposity may be measured using various methods. These include hydrostatic weighting, where the amount of water displaced by the body is measured; air displacement plethysmography, where air displacement is used instead of water; deuterium oxide dilution, to measure the amount of water and hence fat in the body; or dual-energy X-ray absorptiometry (DXA), which estimates fat composition based on the absorption patterns of X-rays (13,14). However, these methods are too complex, costly and time-consuming for regular use, and simple methods to estimate adiposity that are easy to perform are required. Many methods to measure obesity, other than BMI, are available, including waist circumference, skin-fold thickness, waist-to-hip ratio and waist-to-height ratio. This systematic review aimed to investigate the diagnostic accuracy of these tools to diagnose childhood obesity when compared with accurate reference standards such as densitometry.

Methods

This systematic review was conducted to comply with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidance (PRISMA). The protocol for the review is registered on PROSPERO (PROSPERO registration number: CRD42013005711). This review forms part of a broader Health Technology Assessment, which is reported in full elsewhere (15).

Search strategy

A range of databases were searched, including MEDLINE, EMBASE, PsycINFO and CINAHL, the Cochrane Library, DARE and Science Citation Index, up to June 2013. References of included studies and relevant systematic reviews were also checked. Searches were not restricted by language or publication status. A search strategy is reported in Table A1.

Study selection

Any population-based study of children/adolescents up to age 18 that compared the diagnostic performance of simple measures of adiposity to define overweight and obesity against reference standard measures was eligible for inclusion. Studies including only children who were not overweight or obese were excluded. Studies had to be of an index test that was a simple measure of adiposity (i.e. one that could be measured easily), such as BMI, skin-fold thickness, waist-to-height ratio, waist-to-hip ratio, Rohrer's Ponderal Index, Benn's Index, body adiposity index, fat mass index, bioelectrical impedance analysis or near-infrared interactance. The results of these index tests had to be presented so that children could be categorized as obese, overweight or normal weight. The performance of the index texts had to be compared with a reference standard that was one of hydrostatic weighting, air displacement plethysmography, DXA, deuterium dilution method (using deuterium oxide) or any multicompartment obesity measure. Studies had to report sensitivity and specificity of the index test(s) or data from which these could be calculated. Studies were selected by two reviewers independently. Disagreements were resolved through discussion or by another reviewer.

Data extraction and quality assessment

Data extraction was conducted by one reviewer and checked by a second reviewer. Sensitivity and specificity estimates were extracted, or data sufficient to calculate them. Also extracted were characteristics of the study (e.g. date and location) and demographic data (age, gender and ethnicity), and details of thresholds used to diagnose obesity and overweight (such as national or international standard definitions) and details of how index and reference standard tests were performed. The quality assessment of diagnostic accuracy studies (QUADAS-2) tool was used to assess quality of the included studies (16).

Statistical methods

Estimates of sensitivity and specificity of the index tests were calculated from presented data. Where two or more studies presented data on an index test, estimates of sensitivity and specificity were pooled using standard diagnostic meta-analysis techniques, namely, the bivariate model (17) to calculate summary sensitivity and specificity and the hierarchical summary receiver operating characteristic curve (HSROC) model (18) to generate summary receiver operating characteristic (ROC) curves. Separate analyses were conducted for each index test. Subgroup analyses were performed to identify differences between boys and girls and, for the bivariate model, to account for different index test thresholds (obese or overweight) and for differences in reference standards. All analyses were performed using the R software (19). Very few studies reported diagnostic accuracy in different age groups or in different ethnic populations, so the impact of these factors could not be assessed.

In studies that presented data on more than one simple index test, diagnostic odds ratios were calculated in order to compare the diagnostic accuracy of the different index tests (20). In order to aid comparison between tests, results are presented in terms of the estimated sensitivity at a 95% specificity based on the estimated diagnostic odds ratios, assuming that these ratios do not vary with specificity. This enabled the comparison of index tests within studies, where they were performed on the same children with the same reference standard. No meta-analyses or across-studies analysis of these comparative studies were performed because the studies were not consistent in which measures of obesity were compared.

Results

Searches identified a total of 10,269 unique references. After initial screening based on titles and abstracts, 794 papers were obtained. After further checks, 375 articles remained for further evaluation. Of these articles, 341 were excluded after detailed assessment, primarily because they did not present suitable diagnostic accuracy data. The remaining 34 unique studies met our inclusion criteria, but nine had insufficient sensitivity and specificity data to be included in the meta-analysis; hence, 25 papers representing 24 distinct cohorts were included in the meta-analysis (Fig. A1) (21–44).

A summary of the characteristics of the 24 included child cohorts is given in Table 1. BMI was the most widely used obesity measure (22 cohorts), but others considered were skin-fold thickness (seven cohorts), waist circumference (seven cohorts), waist-to-hip ratios (three cohorts) and waist-to-height ratios (two cohorts) and relative weight (two cohorts). The studies varied considerably in how obesity and overweight were defined from these index tests, with studies using different thresholds and different national or international standardizations of BMI (see Table A3 for full details). Skin-fold thickness was sometimes measured on the triceps, sometimes subscapular, or a combination of both.

Of the reference standards, only five studies used densitometry (hydrostatic weighting or air displacement plethysmography); one used deuterium dilution; and the rest used DXA. Studies generally reported results at the 85th centile of DXA, which we define as overweight, and the 95th centile for obesity, although there was some variation across studies (Table A3). These centiles appeared to be age-adjusted and sex-adjusted, although this was not always stated. There was more variation in the percentiles of body fat reported from densitometry and deuterium dilution reference standards, although defining obesity as above 30% body fat for girls and above 25% for boys was most common.

Most studies included any healthy children regardless of age, gender or ethnicity. One study (22) was in children referred to hospital, and one was in children with spinal muscular atrophy (38).

Study quality

The full results of the quality assessment are given in Table A2. The nature of the tests meant that all except one of the cohort studies avoided differential verification bias (where the results of the index test influence the reference standard) and incorporation bias (where the index test is a component of the reference standard). In one study (24), the results of DXA were imputed for some children, and thresholds of DXA used to define obesity appear to have been partly related to the results of the BMI analyses. It is unlikely that any time delay between conducting the index test and the reference standard would introduce bias, although no studies reported the timing of the tests. The description of the index tests was adequate in most studies; but little information on the reference standards was reported.

Body mass index

A total of 22 diagnostic accuracy studies evaluated BMI. Table 2 gives the results of the bivariate analysis of sensitivity and specificity, according to gender and whether the threshold was obesity (95th centile of BMI) or overweight (85th centile of BMI). Definitions of obesity varied across studies and included national BMI standardizations (including for the UK), International Obesity Task Force (45) and Centre for Disease Control and Prevention centiles (46). Figure 1 shows the sensitivity and specificity data from each study, according to gender and threshold (obese and overweight), and summary ROC curves from the HSROC model.

Overall BMI correctly detected 81.9% of obese (that is, highly adipose) children when compared with the reference standards with a false-positive rate of 4% (96% specificity – Table 2). So most obese children will be correctly identified and few non-obese children incorrectly classified as obese. BMI appears to perform less well at detecting overweight: detecting fewer overweight children (76.3% sensitivity) at a higher false-positive rate of 7.9% (Table 2).

Figure 1 shows that there was marked heterogeneity in the data across studies using BMI to detect overweight and obesity, both in sensitivity and specificity rates. The summary ROC curves suggest that BMI may be better at detecting overweight or obesity in girls than boys. At 95% specificity, the detection rate was around 75% for boys but 80% for girls. However, the wide 95% confidence intervals (CI) seen in Table 1 mean that this difference is not conclusive.

Other possible causes of heterogeneity are the varying thresholds and standardizations used to define obesity and overweight, although the HSROC model is designed to account for differences in thresholds, differences in populations and ethnicities and different reference standards. We

Table 1 Summary of the included studies

Study author	Year	Location	Sample size	Gender	Age at measurement	Index tests	Reference standard	Outcome threshold
Bartok ([21])	2011	USA/Canada	151	Girls	9 to 15	BMI	DXA	Obese and overweight
Dung ([48])	2006	Europe	393	Boys and girls	1 to 18	BMI	DXA	Overweight
Ellis ([23])	1999	USA/Canada	979	Boys and girls	3 to 18	BMI	DXA	Obese and overweight
Freedman ([24])	2013	USA/Canada	7,365	Boys and girls	9 to 18	BMI, SFT	DXA	Obese and overweight
Fujita ([49])	2011	UK	422	Boys and girls	10	BMI, WC, WHtR	DXA	Obese
Guntsche ([25])	2010	South America	108	Boys and girls	6 to 16	BMI, SFT, WC, WHpR, WHtR	DXA	Overweight
Harrington ([26])	2013	USA/Canada	423	Boys and girls	5 to 18	BMI	DXA	Obese
Himes ([27])	1989	USA/Canada	316	Boys and girls	8 to 18	BMI, SFT	HW	Obese
Johnston ([28])	1985	USA/Canada	235	Boys and girls	12 to 17	SFT, RWt	HW	Obese
Khadgawat ([29])	2013	Asia	1,640	Boys and girls	7 to 17	BMI	DXA	Obese and overweight
Marshall ([30])	1991	USA/Canada	540	Boys and girls	7 to 14	BMI, SFT, RWt	HW	Obese
Mei ([31])	2006	USA/Canada	1,196	Boys and girls	5 to 18	BMI, SFT	DXA	Obese
Moreno ([32])	2006	Europe	286	Boys and girls	13 to 17	BMI	DXA	Obese
Neovius ([33])	2004/2005	Europe	474	Boys and girls	15 to 18	BMI, WC, WHpR	ADP	Obese and overweight
Pandit ([34])	2009	Asia	586	Boys and girls	6 to 17	BMI	DXA	Obese and overweight
Reilly ([36])	2010	UK	7,722	Boys and girls	8 to 10	BMI, WC	DXA	Obese
Sarria ([37])	2001	Europe	175	Boys	7 to 16	BMI, SFT, WC	HW	Overweight
Sproule ([38])	2009	USA/Canada	25	Boys and girls	5 to 18	BMI	DXA	Obese and overweight
Taylor ([39])	2000	Australia/NZ	580	Boys and girls	3 to 19	WC, WHpR	DXA	Overweight
Telford ([40])	2008	Australia/NZ	741	Boys and girls	7 to 9	BMI	DXA	Obese and overweight
Vitolo ([41])	2007	South America	418	Boys and girls	10 to 19	BMI	DXA	Overweight
Warner ([42])	1997	UK	143	Boys and girls	6 to 18	BMI	DXA	Overweight
Wickramasinghe ([43])	2009	Australia/NZ	138	Boys and girls	5 to 15	BMI, WC	D2O	Obese
Zhang ([44])	2004	Asia	751	Boys and girls	9 to 14	BMI	DXA	Obese

ADP, air displacement plethysmography; BMI, body mass index; D_2O, deuterium dilution method; DXA, dual-energy X-ray absorptiometry; HW, hydrostatic weighting (densitometry); RWt, relative weight; SFT, skin-fold thickness; WC, waist circumference; WHpR, waist-to-hip ratio; WHtR, waist-to-height ratio.

Table 2 Results of bivariate analyses of sensitivity and specificity

Index test	Gender	Threshold	Sensitivity	95% confidence interval		Specificity	95% confidence interval	
Body mass index	Both	Obese	81.9	70.0	93.8	96.0	93.8	98.1
		Overweight	76.3	70.2	82.4	92.1	90.0	94.3
	Boys	Obese	75.2	52.2	98.3	96.3	93.6	99
		Overweight	80.1	73.5	86.7	91.4	89.2	93.5
	Girls	Obese	80.2	60.5	100	97.2	93.5	100
		Overweight	74.7	64.4	85.0	92.1	88.4	95.9
Skin-fold thickness	Both	Obese	72.5	58.7	86.3	93.7	90.2	97.2
		Overweight	78.0	69.2	86.9	90.3	88.0	92.5
	Boys	Obese	64.8	48.2	81.3	93.1	88.5	97.7
		Overweight	74.7	56.1	93.3	92.2	91.2	93.1
	Girls	Obese	67.5	39.4	95.6	99.1	73.9	100
Waist circumference	Both	Obese	83.8	61.2	100	96.5	92.1	100
		Overweight	73.4	58.6	88.1	94.7	91.1	98.4
	Boys	Obese	73.1	37.3	100	96.0	88.1	100
		Overweight	62.3	48.4	76.1	96.9	91.7	100
	Girls	Obese	77.7	45.5	100	96.6	88.4	100

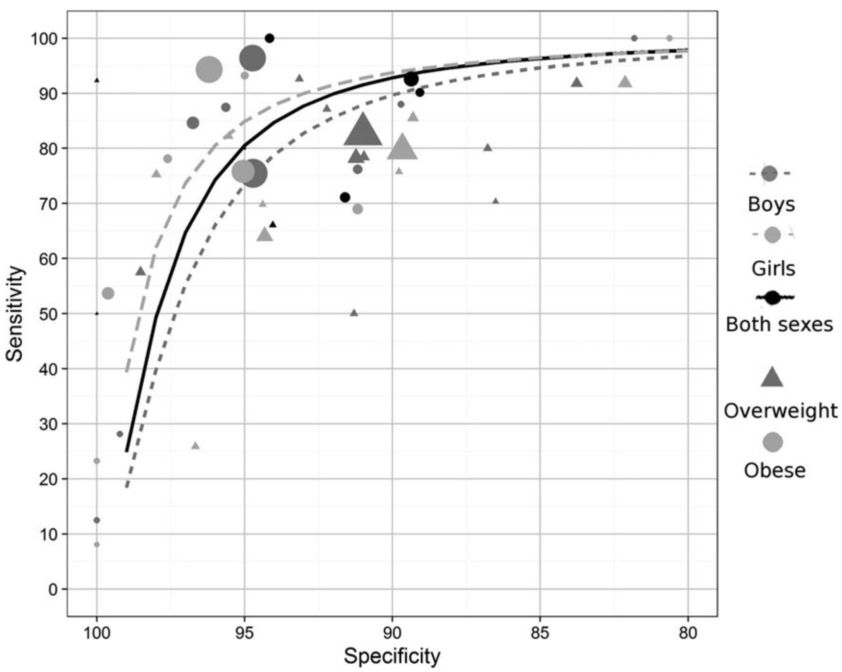

Figure 1 Sensitivity, specificity and summary hierarchical summary receiver operating characteristic curves when using body mass index.

performed a subgroup analysis comparing studies using DXA as a reference standard with those using other reference standards (Table 3). Results were broadly comparable between studies using DXA and non-DXA reference standards, except that sensitivity to detect obesity was lower for other reference standards (35.3%, 95% CI: 12.6 to 58.0) compared with using DXA (90.1% 95% CI: 84.8 to 96.5). This suggests that determination of obesity may be

Table 3 Subgroup analyses for diagnostic accuracy of BMI comparing studies using DXA as a reference standard with other standards

Reference standard	Threshold	Studies	Sensitivity	95% confidence interval		Specificity	95% confidence interval	
DXA	Obese	11	90.1	84.8	96.5	93.6	90.1	96.4
	Overweight	11	76.5	70.2	82.9	92.4	90.3	94.5
Not DXA	Obese	4	35.3	12.6	58.0	99.1	97.3	100
	Overweight	2	75.2	55.9	94.4	87.7	80.0	95.5

BMI, body mass index; CI, confidence interval; DXA, ?dual-energy? X-ray absorptiometry.

dependent on the choice of reference standard, although results should be interpreted with caution owing to the limited number of studies. In particular, the sensitivity was very low in the one study that used deuterium dilution (43). Results between DXA and other reference standards were more consistent for the diagnosis of overweight.

Skin-fold thickness

Seven studies reported data on skin-fold thickness. Studies reported data on both specific skin-fold locations (triceps or subscapular) and sums across locations. Where both were reported, sums of skin-fold thickness were used in this analysis. Table 2 gives the results of the bivariate analysis. There were no studies reporting data for predicting overweight in girls. Figure A2 shows the sensitivity and specificity data from each study and the summary ROC curve. There were too few studies to produce ROC curves by gender.

Skin-fold thickness correctly detected 72.5% of obese children when compared with the reference standards with a false-positive rate of 6.3% (93.7% specificity). So most obese children were correctly identified and few non-obese children incorrectly classified as obese, but using skin-fold thickness missed over one-quarter of obese children. Skin-fold thickness detected more overweight children (78% sensitivity) but had a higher 9.7% false-positive rate (90.3% specificity). There were too few studies of skin-fold thickness to reliably perform any subgroup analyses.

Waist circumference

Seven studies included data on waist circumference. Table 2 gives the results of the bivariate analysis, and Fig. 4b shows the sensitivity and specificity data from each study and the summary ROC curve.

Waist circumference had a similar performance to BMI, with waist circumference correctly identifying 83.8% of obese children when compared with the reference standards, with a false-positive rate of 3.5% (96.5% specificity). There was no conclusive evidence of any difference in effect between boys and girls. As with BMI, waist circumference appears to detect overweight less well: detecting fewer overweight children (73.4% sensitivity) at a higher false-positive rate of 5.3%. There were too few studies of waist circumference to reliably perform any subgroup analyses.

Other measures

Six studies presented data on three other measures: waist-to-height and waist-to-hip ratios and relative weight (that is, weight adjusted for age and gender). There were too little data to perform any meta-analyses, so the results of these studies are summarized in Table 3.

It is difficult to draw any conclusions from these limited data. Relative weight appears to have poor sensitivity of around 50% or less. Waist-to-hip ratio also has poor sensitivity of 45% or less in two of the three studies that used this test. Waist-to-height ratio has very high sensitivities of near 100% in the two studies including it, but in both studies, BMI also achieved near-100% sensitivity (Fig. 2).

Comparison of measures

Figure 2 shows the estimated sensitivity at 95% specificity for the 12 studies that included more than one index test in order to compare the performance of the index tests. Index tests are compared within each study here to give a fair comparison of tests because they were performed on the same children. There was little consistency in results across studies. For example, skin-fold thickness had lower sensitivity than BMI in the Himes (27) and Guntsche (25) studies, higher in the Marshall (30) study and similar in the Freedman (24), Mei (31) and Sarria (37) studies. Overall, particularly as the Freedman study is by far the largest (Table 1), the results suggest that skin-fold thickness has, at best, a marginally better diagnostic performance than BMI.

Waist circumference had a similar sensitivity to BMI in the six studies that included both tests. Relative weight had lower sensitivity than the alternative tests in the two studies including relative weight. Waist-to-hip ratio also had lower sensitivity than BMI or waist circumference in two of the three studies that included it (Table 4). These results suggest that relative weight and waist-to-hip ratio may be inferior to BMI, skin-fold thickness and waist circumference. Waist-to-height ratio was only included in two studies, with results similar to BMI and waist circumference.

Discussion

This systematic review has analysed the diagnostic accuracy of a number of tests for childhood obesity, including BMI and skin-fold thickness. Contrary to common opinion, we found that BMI is a good test for childhood obesity, identifying about 82% of genuinely obese, or highly adipose, children, while misclassifying only 4% of children. However, the 82% sensitivity does mean that 18% of obese children will not be identified as such using BMI. So, an appreciable minority of obesity cases will go undetected. BMI is slightly poorer at diagnosing overweight (or moderately elevated adiposity). This finding does not rule out the possibility that BMI is a poor test in some sub-populations, such as short or muscular children. None of the studies reported data on such sub-populations.

Results for skin-fold thickness were mixed. In bivariate models, skin-fold thickness had lower sensitivity than BMI, but in the largest study that compared them, skin-fold thickness had slightly higher sensitivity. These results

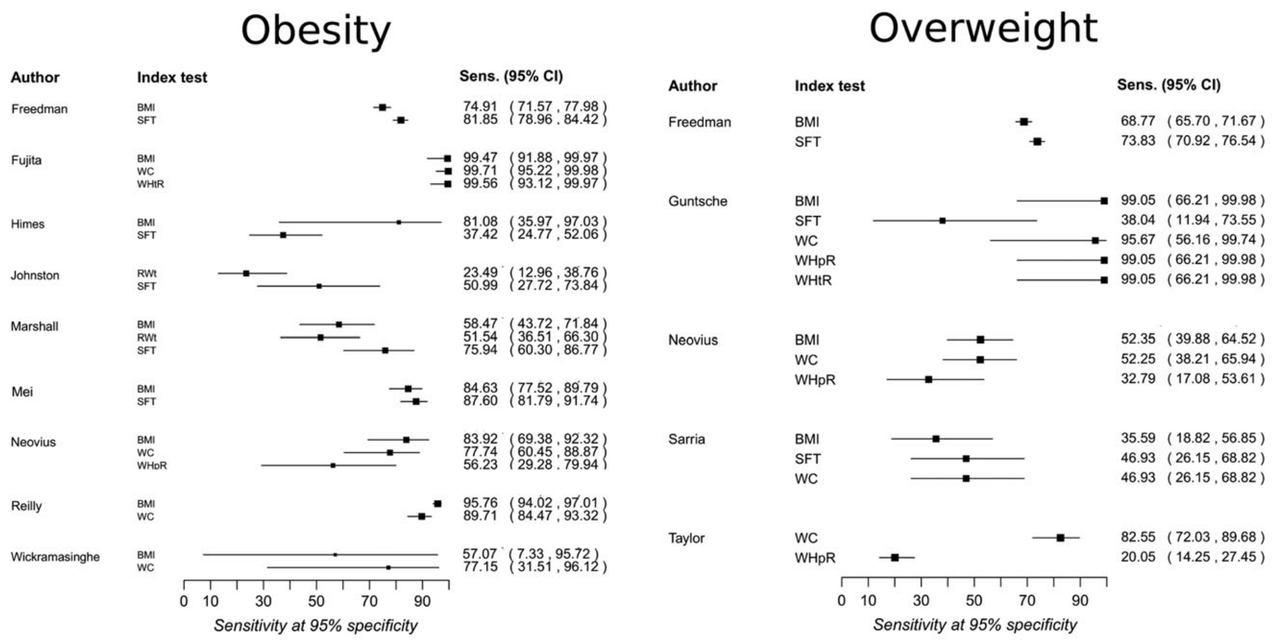

Figure 2 Sensitivity at 95% specificity in studies comparing index tests.

suggest that the extra complexity of performing a skin-fold thickness test, and the need for trained professionals to carry out the measurement, may outweigh any possible marginal improvements in diagnostic performance.

Data on other obesity tests were more limited, but there was no compelling evidence that any alternative test had better performance than BMI. Waist circumference appears to have a similar diagnostic performance to BMI, while the limited data on relative weight and waist-to-hip ratio suggest these perform less well.

There was considerable heterogeneity across studies, with differences in diagnostic accuracy according to gender and the reference standard used. Differences in thresholds used to classify obesity and differences in populations may also contribute to heterogeneity. Therefore, although BMI, skinfold thickness and waist circumference may perform well in general, diagnostic accuracy in practice may depend on which diagnostic thresholds are used and how well these apply to the population of interest.

Strengths and limitations

This systematic review used rigorous methods and followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. Extensive searches were performed to identify all relevant studies. Rigorous statistical methods were used to pool data across diagnostic accuracy studies.

A key limitation in this review was the diversity of the studies. Studies were in different populations at varying ages, and with different ethnicities (although diagnostic accuracy by subgroups were not routinely reported), and used differing definitions of obesity. While all studies used either obesity or overweight as their threshold, these thresholds were not

Table 4 Diagnostic accuracy results for relative weight, waist-to-hip and waist-to-height ratios

Author	Threshold	Gender	Sensitivity	(95% confidence interval)		Specificity	(95% confidence interval)	
Relative weight								
Johnston ([28])	Obese	Boys	51.6	34	69.2	86.2	80.5	92
		Girls	29.4	7.8	51.1	93.9	87.2	100
Marshall ([30])	Obese	Both	51.3	40.1	62.6	95	93.1	97
Waist-to-hip ratio								
Guntsche ([25])	Overweight	Both	96.4	86.7	100	98.6	94.9	100
Neovius ([33])	Overweight	Boys	24	7.3	40.7	97.7	95.5	99.9
		Girls	17.2	10.4	24.1	97.5	94.6	100
	Obese	Boys	40.7	27.6	53.8	97.3	94.6	99.9
Taylor ([39])	Overweight	Both	45.9	38.1	53.7	84.9	81.5	88.3
Waist-to-height ratio								
Fujita ([49])	Obese	Both	99.6	98.4	100	95	92.5	97.4
Guntsche ([25])	Overweight	Both	96.4	86.7	100	98.6	94.9	100

consistent across studies and so are unlikely to be consistent across different populations (47). Reporting on diagnostic performance by age or ethnicity was too limited to investigate the impact of these factors on obesity diagnosis. The studies also used several different reference standards, which may not be directly comparable, and may lead to differences in estimates of diagnostic accuracy. It was generally necessary to assume equivalence of these reference standards in the analyses, which is unlikely to be correct.

Another limitation was the small number of studies considering tests other than BMI, particularly other simple measures using different powered relationships between height and weight, such as the Ponderal Index. This restricted our ability to compare tests and draw any firm conclusions about their relative merits. Bioelectrical impedance may provide a routine measure of fat mass in the future, but no studies comparing these measures with reference standards were identified in the present review. This suggests that high-quality diagnostic test accuracy studies are needed for other tests, perhaps particularly for waist-to-height and waist-to-hip ratios. Such studies should use a high-quality reference standard for diagnosing obesity and measure BMI in order to compare the performance of different tests with BMI.

Conclusions

Perhaps contrary to popular opinion, this review found that BMI is a reasonably good, simple diagnostic test for identifying childhood obesity and adiposity. It identifies most adipose children correctly, but does fail to identify around 20% of obese or highly adipose children, while misclassifying only a small number as obese. The good diagnostic accuracy relies on selecting appropriate BMI thresholds to define obesity for the population of interest, which may vary according to age, gender and ethnicity. There were few studies of other simple diagnostic tests, and there was no conclusive evidence that any simple test should be preferred to BMI. In particular, the extra complexity involved in performing skin-fold thickness tests does not appear to result in any great improvement in diagnostic accuracy. While BMI is a good simple test for childhood obesity, it is not perfect, and some obese children will not be identified using BMI.

Conflict of interest statement

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; and no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

We would like to thank Dr Jane Burch and Dr Huiqin Yang for contributing to the design of the review and to the selection of the studies and for their comments and suggestions throughout the conduct of the review. We thank members of our advisory group, Professor Charlotte Wright and Dr Jason Halford, for their advice, comments and suggestions throughout the conduct of the review. We would also like to thank Professor Christine Power for her advice during the protocol development stage of the project. We thank Stephen Duffy for conducting the literature searches.

References

1. World Health Organization. Childhood overweight and obesity. http://www.who.int/dietphysicalactivity/childhood/en/ (accessed 6th June 2014).
2. Singh AS, Mulder C, Twisk JW, van Mechelen W, Chinapaw MJ. Tracking of childhood overweight into adulthood: a systematic review of the literature. *Obes Rev* 2008; **9**: 474–488.
3. Brisbois TD, Farmer AP, McCargar LJ. Early markers of adult obesity: a review. *Obes Rev* 2012; **13s**: 347–367.
4. Llewellyn A, Simmonds M, Owen CG, Woolacott N. Childhood obesity as a predictor of morbidity in adulthood: a systematic review and meta-analysis. *Obes Rev* 2016; **17**: 56–67.
5. Vucenik I, Stains JP. Obesity and cancer risk: evidence, mechanisms, and recommendations. *Ann N Y Acad Sci* 2012; **1271**: 37–43.
6. Shields M, Tremblay MS, Connor Gorber S, Janssen I. Abdominal obesity and cardiovascular disease risk factors within body mass index categories. *Health Rep* 2012; **23**: 7–15.
7. Ashwell M, Gunn P, Gibson S. Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: systematic review and meta-analysis. *Obes Rev* 2012; **13**: 275–286.
8. Prospective Studies Collaboration. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet* n.d; **373**: 1083–1096.
9. Simmonds M, Llewellyn A, Owen CG, Woolacott N. Predicting adult obesity from childhood obesity: a systematic review and meta-analysis. *Obes Rev* 2016; **17**: 95–107.
10. Trefethen N BMI (body mass index): calculate your "New BMI". 2013. http://people.maths.ox.ac.uk/trefethen/bmi.html (accessed Aug 29 2013).
11. Ehtisham S, Crabtree N, Clark P, Shaw N, Barrett T. Ethnic differences in insulin resistance and body composition in United Kingdom adolescents. *J Clin Endocrinol Metab* 2005; **90**: 3963–3969.
12. Nightingale CM, Rudnicka AR, Owen CG, Cook DG, Whincup PH. Patterns of body size and adiposity among UK children of South Asian, black African-Caribbean and white European origin: Child Heart and Health Study in England (CHASE study). *Int J Epidemiol* 2011; **40**: 33–44.
13. Wells JC, Fewtrell MS. Measuring body composition. *Arch Dis Child* 2006; **91**: 612–617.
14. Cornier MA, Després JP, Davis N *et al.* Assessing adiposity: a scientific statement from the American Heart Association. *Circulation* 2011; **124**: 1996–2019.
15. Simmonds M, Burch J, Llewellyn A, *et al.* The use of measures of obesity in childhood for predicting obesity and the development

of obesity-related diseases in adulthood; a systematic review and meta-analysis. *Health Technol Assess* 2015; **19**.
16. Whiting PF, Rutjes AW, Westwood ME *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155**: 529–536.
17. Reitsma JB, Glas AS, Rutjes AW, Scholten RJPM, Bossuyt PN, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; **58**: 982–990.
18. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001; **20**: 2865–2884.
19. Simmonds MC, Higgins JPT. A general framework for the use of logistic regression models in meta-analysis. *Stat Methods Med Res* 2014; (**Online first**): May 12. DOI: 10.1177/0962280214534409.
20. Deeks JJ. Systematic reviews of evaluations of diagnostic and screening tests. *BMJ*: *British Medical Journal* 2001; **323**: 157–162.
21. Bartok CJ, Marini ME, Birch LL. High body mass index percentile accurately reflects excess adiposity in white girls. *J Am Diet Assoc* 2011; **111**: 437–441.
22. Dung NQ. Body composition and nutritional status in neonates and sick children as accessed by dual energy x-ray absorptiometry, bioelectrical impedance analysis and anthropometric methods. In: Impact of Nutrition on Postnatal Growth. Ernst-Moritz-Arndt-Universität: Greifswald, 2006.
23. Ellis KJ, Abrams SA, Wong WW. Monitoring childhood obesity: assessment of the weight/height index. *Am J Epidemiol* 1999; **150**: 939–946.
24. Freedman DS, Ogden CL, Blanck HM, Borrud LG, Dietz WH. The abilities of body mass index and skinfold thicknesses to identify children with low or elevated levels of dual-energy X-ray absorptiometry-determined body fatness. *J Pediatr* 2013; **163**: 160–166.
25. Guntsche Z, Guntsche EM, Saravi FD *et al.* Umbilical waist-to-height ratio and trunk fat mass index (DXA) as markers of central adiposity and insulin resistance in Argentinean children with a family history of metabolic syndrome. *J Pediatr Endocrinol* 2010; **23**: 245–256.
26. Harrington DM, Staiano AE, Broyles ST, Gupta AK, Katzmarzyk PT. BMI percentiles for the identification of abdominal obesity and metabolic risk in children and adolescents: evidence in support of the CDC 95th percentile. *Eur J Clin Nutr* 2013; **67**: 218–222.
27. Himes JH, Bouchard C. Validity of anthropometry in classifying youths as obese. *Int J Obes (Lond)* 1989; **13**: 183–193.
28. Johnston FE. Validity of triceps skinfold and relative weight as measures of adolescent obesity. *J Adolesc Health Care* 1985; **6**: 185–190.
29. Khadgawat R, Marwaha RK, Tandon N, *et al.* Reference intervals of percentage body fat in apparently healthy North-Indian school children and adolescents. *Indian Pediatr* 2013; **50**: 859–866.
30. Marshall JD, Hazlett CB, Spady DW, Conger PR, Quinney HA. Validity of convenient indicators of obesity. *Hum Biol* 1991; **63**: 137–153.
31. Mei Z, Grummer-Strawn LM, Wang J *et al.* Do skinfold measurements provide additional information to body mass index in the assessment of body fatness among children and adolescents? *Pediatrics* 2007; **119**: e1306–e1313.
32. Moreno LA, Blay MG, Rodriguez G *et al.* Screening performances of the International Obesity Task Force body mass index cut-off values in adolescents. *J Am Coll Nutr* 2006; **25**: 403–408.
33. Neovius M Diagnostic tests for adiposity and metabolic risk factors in adolescence: results from the Stockholm Weight Development Study (SWEDES). Karolinska Institute; 2005.
34. Neovius MG, Linne YM, Barkeling BS, Rossner SO. Sensitivity and specificity of classification systems for fatness in adolescents. *Am J Clin Nutr* 2004; **80**: 597–603.
35. Pandit D, Chiplonkar S, Khadilkar A, Khadilkar V, Ekbote V. Body fat percentages by dual-energy X-ray absorptiometry corresponding to body mass index cutoffs for overweight and obesity in Indian children. *Clin Med Pediatr* 2009; **3**: 55–61.
36. Reilly JJ, Dorosty AR, Ghomizadeh NM, Sherriff A, Wells JC, Ness AR. Comparison of waist circumference percentiles versus body mass index percentiles for diagnosis of obesity in a large cohort of children. *Int J Pediatr Obes* 2010a; **5**: 151–156.
37. Sarria A, Moreno LA, Garcia-Llop LA, Fleta J, Morellon MP, Bueno M. Body mass index, triceps skinfold and waist circumference in screening for adiposity in male children and adolescents. *Acta Paediatr* 2001; **90**: 387–392.
38. Sproule DM, Montes J, Montgomery M *et al.* Increased fat mass and high incidence of overweight despite low body mass index in patients with spinal muscular atrophy. *Neuromuscul Disord* 2009; **19**: 391–396.
39. Taylor RW, Jones IE, Williams SM, Goulding A. Evaluation of waist circumference, waist-to-hip ratio, and the conicity index as screening tools for high trunk fat mass, as measured by dual-energy X-ray absorptiometry, in children aged 3–19 y. *Am J Clin Nutr* 2000; **72**: 490–495.
40. Telford RD, Cunningham RB, Daly RM *et al.* Discordance of international adiposity classifications in Australian boys and girls – the LOOK study. *Ann Hum Biol* 2008; **35**: 334–341.
41. Vitolo MR, Campagnolo PD, Barros ME, Gama CM, Ancona LF. Evaluation of two classifications for overweight among Brazilian adolescents. *Rev Saude Publica* 2007; **41**: 653–656.
42. Warner JT, Cowan FJ, Dunstan FD, Gregory JW. The validity of body mass index for the assessment of adiposity in children with disease states. *Ann Hum Biol* 1997; **24**: 209–215.
43. Wickramasinghe VP, Lamabadusuriya SP, Cleghorn GJ, Davies PS. Validity of currently used cutoff values of body mass index as a measure of obesity in Sri Lankan children. *Ceylon Med J* 2009; **54**: 114–119.
44. Zhang Q, Du WJ, Hu XQ, Liu AL, Pan H, Ma GS. The relation between body mass index and percentage body fat among Chinese adolescent living in urban Beijing. *Chung Hua Liu Hsing Ping Hsueh Tsa Chih* 2004; **25**: 113–116.
45. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000; **320**: 1240–1243.
46. Kuczmarski RJ, Ogden CL, Guo SS, *et al.* 2000 CDC Growth Charts for the United States: methods and development. *Vital and health statistics series 11*, *Data from the National Health Survey* 2002: 1–190.
47. Reilly JJ, Kelly J, Wilson DC. Accuracy of simple clinical and epidemiological definitions of childhood obesity: systematic review and evidence appraisal. *Obes Rev* 2010b; **11**: 645–655.
48. Dung NQ, Fusch G, Armbrust S, Jochum F, Fusch C. Body composition of preterm infants measured during the first months of life: bioelectrical impedance provides insignificant additional information compared to anthropometry alone. *Eur J Pediatr* 2007; **166**: 215–222.
49. Fujita Y, Kouda K, Nakamura H, Iki M. Cut-off values of body mass index, waist circumference, and waist-to-height ratio to identify excess abdominal fat: population-based screening of Japanese school children. *J Epidemiol* 2011; **21**: 191–196.

APPENDIX

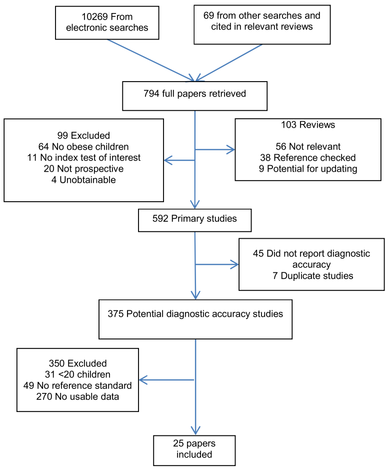

Figure A1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram for the systematic review.

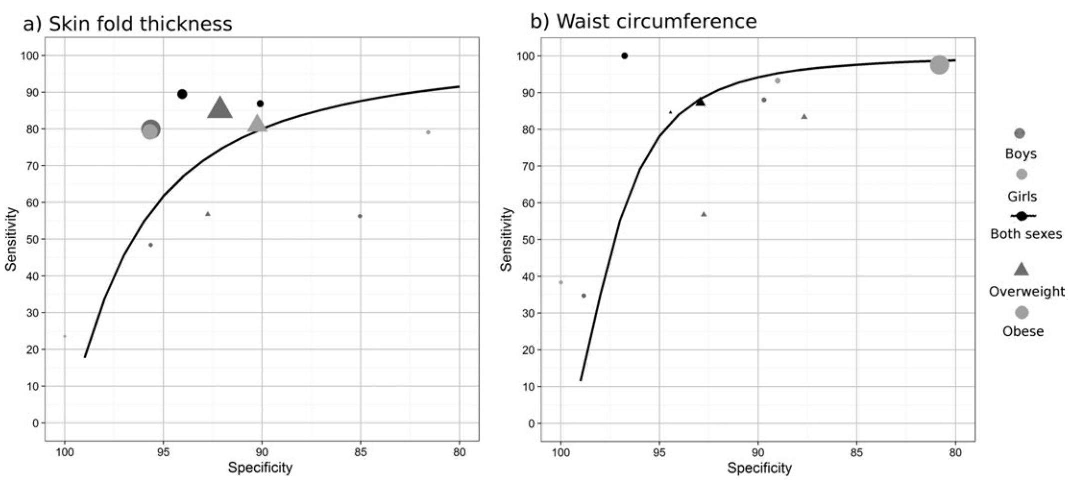

Figure A2 Sensitivity, specificity and summary hierarchical summary receiver operating characteristic curves when using (a) skin-fold thickness or (b) waist circumference.

Table A1 MEDLINE search strategy

1	exp Obesity/
2	Overweight/
3	Weight Gain/
4	Weight Loss/
5	obes$.ti,ab.
6	(overweight or over weight).ti,ab
7	(weight gain or weight loss).ti,ab.
8	or/1-7
9	Adiposity/ or Adipose Tissue/
10	exp Body Composition/
11	Body Weight/
12	(adiposity or adipose).ti,ab.
13	(body adj2 (composition or fat or weight)).ti,ab.
14	fatness.ti,ab.
15	or/8-14
16	Body Mass Index/
17	Skinfold Thickness/
18	Waist Circumference/
19	Waist-Hip Ratio/
20	Electric Impedance/
21	((body mass adj3 (index$ or indices)) or bmi or quetelet$).ti,ab.
22	((fat mass adj3 (index$ or indices)) or fmi).ti,ab
23	((fat free mass adj3 (index$ or indices)) or ffmi).ti,ab.
24	(body adipos$ adj3 (index$ or indices)).ti,ab.
25	(body fat adj2 percentage$).ti,ab.
26	((skinfold or skinfold) adj3 (thickness$ or test$ or measure$)).ti,ab.
27	((waist or hip or neck) adj3 circumference$).ti,ab.
28	((waist-to-hip or waist-hip) adj3 ratio$).ti,ab.
29	((waist-to-height or waist-height) adj3 ratio$).ti,ab.
30	(((bioelectric$ or electric$) adj3 (impedance or resistance)) or bia).ti,ab.
31	(near infrared interactance or NIR).ti,ab.
32	((benn$ or rohrer$ or ponderal or corpulence) adj3 (index$ or indices)).ti,ab.
33	(sagittal abdominal diameter$ or supine abdominal diameter$).ti,ab.
34	or/16-33
35	exp Densitometry/
36	exp Plethysmography/
37	Neutron Activation Analysis/
38	(body volume adj3 (index$ or indices)).ti,ab.
39	(densitometr$ or hydrodensitometr$).ti,ab
40	((hydrostatic or underwater or water) adj3 (weighing or analys$ or measure$)).ti,ab.
41	(absorptiometry or DXA or DEXA).ti,ab.
42	((water or air) adj3 displacement).ti,ab.
43	(air displacement plethysmograph$ or pea pod or peapod or infant body composition system$ or bodpod or bod pod).ti,ab.
44	(neutron$ adj3 activat$).ti,ab.
45	((multicomponent$ or multi component$ or multimodal$ or multi modal$ or composit$) adj3 model$).ti,ab
46	(deuterium adj3 dilut$).ti,ab.
47	or/35-46
48	exp child/
49	exp Infant/
50	Adolescent/
51	Young Adult/
52	(child$ or infant$ or pediat$ or paediat$ or schoolchild$ or school age$ or schoolage$).ti,ab.
53	(adolescen$ or juvenile$ or youth$ or teenage$ or youngster$).ti,ab
54	(girl or girls or boy or boys or kid or kids).ti,ab.
55	(young people or young person or young persons or young adult$).ti,ab.
56	or/48-55
57	15 and 34 and 47 and 56
58	exp Animals/ not Humans/
59	57 not 58

Table A2 Results of the quality assessment

Short title	1. Representative population	2. Progression bias	3. Partial verification bias	4. Differential verification bias	5. Incorporation bias	6. Description of selection criteria	7. Appropriateness of RS	8. Description of IT	9. Used validated IT	10. Description of RS	11. Uninterpretable/ intermediate results reported	12. Withdrawals explained	13. Measurement bias 13a. Training/ experience IT test personnel	13b. Number of IT assessors	13c. Training/ experience RS test personnel	13d. Number of RS assessors
Bartok (21)	No	UC	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	No	UC	UC	UC	UC
Dung (22)	No	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	None	UC	UC	UC	UC
Ellis (23)	Yes	Probably avoided	Avoided	Avoided	Avoided	Inadequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	None	Yes	Multi.	UC	UC
Freedman (24)	Yes	Probably avoided	Present	Avoided	Avoided	Adequate	Imperfect	Adequate *for BMI* Inadequate *for SFT*	Yes	Inadequate	Yes	No	UC	UC	UC	UC
Fujita (49)	Yes	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	Yes	UC	UC	UC	UC
Guntsche (25)	No	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Inadequate	Yes	Inadequate	Apparently none	None	UC	UC	UC	UC
Harrington (26)	No	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	Yes	UC	UC	UC	UC
Himes (27)	Yes	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Inadequate	Yes	Inadequate	Apparently none	None	UC	UC	UC	UC
Johnston (28)	No	Probably avoided	Avoided	Avoided	Avoided	Inadequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	None	UC	UC	UC	UC
Khadgawat (29)	Yes for India No for UK	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	Yes	UC	UC	UC	UC
Marshall (30)	Yes	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	Yes	UC	UC	UC	UC
Mei (31)	Yes	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	Yes	Yes	2	UC	UC
Moreno (32)	Yes	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	Yes	UC	UC	Yes	1
Neovius (34)	Yes	UC	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	Yes	UC	UC	UC	UC
Neovius (33)	Yes	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Inadequate	Yes	Inadequate	Apparently none	Yes	UC	UC	UC	UC
Pandit (35)	Yes	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	None	UC	UC	UC	1
Reilly (36)	Yes	UC	Avoided	Avoided	Avoided	Adequate	Imperfect	Inadequate	Yes	Inadequate	Apparently none	Yes	UC	UC	UC	UC
Sarria (37)	No	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Adequate	Yes	Yes	UC	UC	UC	UC
Sproule (38)	No	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	None	UC	UC	UC	UC
Taylor (39)	No*	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	None	UC	UC	UC	UC
Telford (40)	Yes		Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate		None	UC	UC	UC	UC

(*Continues*)

Table A2 (Continued)

													13. Measurement bias			
Short title	1. Representative population	2. Progression bias	3. Partial verification bias	4. Differential verification bias	5. Incorporation bias	6. Description of selection criteria	7. Appropriateness of RS	8. Description of IT	9. Used validated IT	10. Description of RS	11. Uninterpretable/ intermediate results reported	12. Withdrawals explained	13a. Training/ experience IT test personnel	13b. Number of IT assessors	13c. Training/ experience RS test personnel	13d. Number of RS assessors
		Probably avoided									Apparently none					
Vitolo (41)	Yes	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	Yes	UC	UC	UC	UC
Warner (42)	No	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	None	UC	UC	UC	UC
Wickramasinghe (43)	Yes	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	None	UC	UC	UC	UC
Zhang (44)	Yes	Probably avoided	Avoided	Avoided	Avoided	Adequate	Imperfect	Adequate	Yes	Inadequate	Apparently none	No	Yes	UC	UC	UC

IT, Index test; RS, Reference standard; UC, unclear.

Table A3 Thresholds for diagnosis of obesity and overweight for index tests and reference standards

Author	Index test reference population or measure	Index test cut-off (percentile)		Reference standard	Reference standard cut-off (percentile)	
		Obese	Overweight		Obese	Overweight
Body mass index						
Bartok (21)	CDC	85th	73rd	DXA	95th	85th
Dung (22)	German reference	—	90th	DXA	—	90th
Ellis (23)	Internal	95th	85th	DXA	95th	85th
Freedman (24)	CDC	95th	85th	DXA	≈82nd (to match centile obese according to BMI)	≈66th (to match centile overweight according to BMI)
Fujita (49)	Optimal (internal)	BMI 19.6 girls; BMI 20.8 boys	—	DXA	95th	—
Guntsche (25)	SD score (internal)	—	2.13	DXA	—	10 kg/m^2
Harrington (26)	CDC	96th	—	DXA	75th	—
Himes (27)	US national reference	85th	—	HW	90th	—
Khadgawat (29)	IOTF	95th	85th	DXA	95th	85th
Marshall (30)	Relative BMI	>120% of 'expected' BMI	—	HW	20%BF boys; 25%BF girls	—
Mei (31)	CDC	95th	—	DXA	95th	—
Moreno (32)	IOTF	—	≈85th Optimized for diag. accuracy	DXA	—	85th
Neovius (34)	IOTF	95th	85th	ADP	95th	25%BF boys; 30%BF girls
Pandit (35)	IOTF	95th	85th	DXA	95th	85th
Reilly (36)	UK90	95th	—	DXA	90th	—
Sarria (37)	Internal	—	85th	HW	—	85th
Sproule (38)	CDC	95th	85th	DXA	95th	85th
Telford (40)	IOTF	BMI 21.6	BMI 18.4	DXA	UK standard (McCarthy)	UK standard (McCarthy)
Vitolo (41)	IOTF		Not reported	DXA		25%BF boys; 30%BF girls
Warner (42)	CDC	—	*Z* score >1	DXA	—	USA 85th
Wickramasinghe (43)	CDC	95th	—	D_2O	25%BF boys; 30%BF girls	—
Zhang (44)	IOTF	BMI 30	—	DXA	25%BF boys; 35%BF girls	—
Skin-fold thickness						
Freedman (24)	Sum	≈82nd (to match centile obese according to BMI)	≈66th (to match centile overweight according to BMI)	DXA	≈82nd (to match centile obese according to BMI)	≈66th (to match centile overweight according to BMI)
Guntsche (25)	Skin-folds index	—	1.26	DXA	—	10 kg/m^2
Himes (27)	Triceps, subscapular, US reference	85th	—	HW	90th	—
Johnston (28)	Triceps, US reference	90th	—	HW	25%BF boys; 30%BF girls	—
Marshall (30)	Triceps + subscapular	85th	—	HW	20%BF boys; 25%BF girls	—
Mei (31)	Triceps	95th		DXA	95th	
Sarria (37)	Triceps + subscapular	—	85th	HW	—	85th
Waist circumference						

(*Continues*)

Table A3 (Continued)

Author	Index test reference population or measure	Index test cut-off (percentile)		Reference standard	Reference standard cut-off (percentile)	
		Obese	Overweight		Obese	Overweight
Fujita (49)	Umbilical optimal (internal)	76.5 boys; 73 girls	—	DXA	95th	—
Guntsche (25)	Umbilical	—	85 cm	DXA	—	10 kg/m^2
Neovius (34)	Smallest between ribs and iliac crest	95th boys; 85th girls	85.9 boys; 73.3 girls	ADP	95th	25%BF boys; 30%BF girls
Reilly (47)	UK 1988 reference	95th	—	DXA	90th	—
Sarria (37)	Smallest between ribs and iliac crest	—	85th	HW	—	85th
Taylor (39)	Smallest between ribs and iliac crest	—	80th	DXA	—	*Z* score >1
Wickramasinghe (43)	Smallest between ribs and iliac crest	98th	—	D_2O	25%BF boys; 30%BF girls	—
Waist-to-hip ratio						
Guntsche (25)	WC midpoint between ribs and iliac crest	—	0.91	DXA	—	10 kg/m^2
Neovius (34)	WC smallest between ribs and iliac crest	0.9 boys; 1.02 girls	0.9 boys; 0.84 girls	ADP	95th	25%BF boys; 30%BF girls
Taylor (39)	WC smallest between ribs and iliac crest	—	80th	DXA	—	*Z* score >1
Waist-to-height ratio						
Fujita (49)	WC umbilical optimal (internal)	0.519 boys; 0.499 girls	—	DXA	95th	—
Guntsche (25)	WC umbilical	—	0.54	DXA	—	10 kg/m^2
Relative weight						
Marshall (30)	—	120% of 'expected' weight	—	HW	20%BF boys; 25%BF girls	—
Johnston (28)	—	Not reported	—	HW	25%BF boys; 30%BF girls	—

ADP, air displacement plethysmography; BF, body fat; D_2O, deuterium dilution method; DXA, dual-energy X-ray absorptiometry; HW: hydrostatic (underwater) weighting; Internal, using study data only, no external reference given; IOTF: International Obesity Taskforce; Optimal: threshold giving optimal diagnostic accuracy; SD, standard deviation; UK90, the British 1990 growth reference; WC, waist circumference.