Check for updates

OPEN

The performance of anthropometric tools to determine obesity: a systematic review and meta-analysis

Isolde Sommer[1✉], Birgit Teufer[1], Monika Szelag[1], Barbara Nussbaumer-Streit[1], Viktoria Titscher[1], Irma Klerings[1] & Gerald Gartlehner[1,2]

The aim of this systematic review was to assess the performance of anthropometric tools to determine obesity in the general population (CRD42018086888). Our review included 32 studies. To detect obesity with body mass index (BMI), the meta-analyses rendered a sensitivity of 51.4% (95% CI 38.5–64.2%) and a specificity of 95.4% (95% CI 90.7–97.8%) in women, and 49.6% (95% CI 34.8–64.5%) and 97.3% (95% CI 92.1–99.1%), respectively, in men. For waist circumference (WC), the summary estimates for the sensitivity were 62.4% (95% CI 49.2–73.9%) and 88.1% for the specificity (95% CI 77.0–94.2%) in men, and 57.0% (95% CI 32.2–79.0%) and 94.8% (95% CI 85.8–98.2%), respectively, in women. The data were insufficient to pool the results for waist-to-hip ratio (WHR) and waist-to-height ratio (WHtR) but were similar to BMI and WC. In conclusion, BMI and WC have serious limitations for use as obesity screening tools in clinical practice despite their widespread use. No evidence supports that WHR and WHtR are more suitable than BMI or WC to assess body fat. However, due to the lack of more accurate and feasible alternatives, BMI and WC might still have a role as initial tools for assessing individuals for excess adiposity until new evidence emerges.

Obesity is widely recognised as a pandemic public health problem. According to the World Health Organization (WHO), in 2016 more than 650 million adults worldwide were obese[1]. These numbers have almost tripled since 1975[2]. Obesity increases the risk for many chronic diseases, such as diabetes mellitus, cardiovascular diseases and cancers[3], and is possibly associated with mental health disorders[4]. Associations have been shown to be strongest between obesity and the incidence of diabetes mellitus, particularly in women (risk ratio [RR] 12.41, 95% confidence interval [CI] 9.03–17.06).

Primary care is considered one of the main settings for the prevention, screening and management of obesity[5]. Individual studies indicate that patients are more likely to lose weight when they receive recommendations for lifestyle changes from their primary care physicians[6]. Because it can be difficult for physicians to accurately determine obesity solely through visually inspecting their patients[7], they need a reliable, efficient screening tool in order to ensure that those who need management and treatment receive it.

WHO conceptualises obesity as "abnormal or excessive fat accumulation that may impair health"[1]. It is most commonly assessed using body mass index (BMI), a simple and quick anthropometric tool that has a low cost. Adults with a BMI greater than or equal to 30 are classified as being obese[1] (Table 1). However, several researchers and professional associations[8–14] consider the use of BMI as the primary clinical index of obesity insufficient. They have called for a new definition that fully accounts for the complexity of the disease relating to the quantity, distribution and secretory function of adipose tissue.

A substantial body of evidence has shown that obesity (BMI ≥ 30) is associated with an increased risk of coronary heart disease[15] and mortality[16] relative to normal weight. For mortality, this association follows a J-shaped curve. Although a significantly higher mortality rate was found for all obesity grades combined (hazard ratio [HR] 1.18 [95% CI 1.12–1.25]), being overweight (BMI of 25–< 30) reduced the risk of all-cause mortality

[1]Department for Evidence-Based Medicine and Evaluation, Danube University Krems, Dr.-Karl-Dorrek-Straße 30, 3500 Krems, Austria. [2]RTI-UNC Evidence-Based Practice Center, Research Triangle Institute International, East Cornwallis Road, Post Office Box 12194, Research Triangle Park, NC 27709-2194, USA. ✉email: isolde.sommer@donau-uni.ac.at

1

Body mass index (BMI): person's weight in kilograms divided by the square of his/her height in meters (kg/m²)
Waist circumference (WC): waist circumference measured at approximate midpoint between the lowest rib and the top of the iliac crest
Waist-to -hip ratio (WHR): person's waist measurement divided by hip measurement taken around the widest portion of the buttocks
Waist-to-height ratio (WHtR): person's waist circumference divided by their height

Table 1. Definitions of anthropometric measurement tools. *Source*: World Health Organization[85].

(0.94 [95% CI 0.91–0.96]), and grade 1 obesity (BMI of 30–< 35) was not related with higher mortality (0.95 [95% CI 0.88–1.01])[16]. In older age, overweight and obesity as defined by BMI might even be protective against mortality[17–19].

Indeed, one of the main deficiencies of BMI is that it does not differentiate between fat mass and fat-free mass. Not all people with high levels of body fat have a BMI of 30 or greater, and some people with very high BMIs may have little fat mass. The proportion of body fat also differs across ethnic populations, sex, and age groups. For example, South Asian populations have a higher proportion of body fat than Caucasians for the same BMI[20]. Women have a significantly higher percentage of total and sub-cutaneous fat stores than their male counterparts[21]. The proportion of internal fat increases and muscle mass decreases with age, which can lead to sarcopenic obesity, the combination of obesity and muscle impairment[22]. In older populations, research even suggests that fat mass is associated with a decreased risk of morbidity and mortality[17,19,23,24], while a low fat-free mass might be a risk factor for mortality[25,26].

Another main deficiency of BMI is that it does not account for body fat distribution. The distribution of body fat is associated with the risk of metabolic syndrome and other cardiometabolic complications[10]. Longitudinal data have shown that the distribution of excess fat (resulting in a so-called apple or pear shape) has a greater influence on certain health risks, such as cardiovascular diseases or cancer, than total body fat[27,28]. Indices assessing the distribution of body fat include waist circumference (WC), waist-to-hip ratio (WHR) or waist-to-height ratio (WHtR) (Table 1). A growing body of evidence suggests that such indices are independently associated with cardiometabolic diseases and mortality[29–31]. They could thus provide additional value in determining obesity and the risk for associated comorbidities in clinical practice.

Imaging techniques allow for the measurement of body fat, its distribution, and body composition but are rarely used in clinical practice. They are generally considered more precise than anthropometric methods and continue to serve as "reference standards" in many research studies[14] until the concept of obesity is fully understood.

Despite the definitional problems with BMI, it remains the routine measurement to classify obesity in clinical practice. Within the last two decades, only two systematic reviews on the performance of anthropometric tools compared to that of body composition techniques have been published. The review by Okorodudu et al.[32] focused on the performance of BMI, and Mc Tigue et al.[33] reviewed the performance of BMI, WC and WHR in older adults. Both reviews are relatively old, with Okorodudu et al.[32] searching for studies until June 2008 and Mc Tigue et al.[33] until February 2003. Due to the emergence of new research evidence and the development of anthropometric tools other than BMI, we have aimed to provide an up-to-date systematic review using four anthropometric tools (BMI, WC, WHR and WHtR) for determining obesity in the adult population.

Methods

This systematic review was conducted following the Cochrane Methods for Systematic Reviews of Diagnostic Test Accuracy[34] and reported according to the Preferred Reporting Items for a Systematic Review and Meta-Analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA) statement[35]. The protocol is registered with the International Prospective Register of Systematic Reviews (PROSPERO), registration number CRD42018086888.

Information sources and searches. We searched the electronic databases Ovid MEDLINE, Embase.com (Elsevier), CINAHL (Ebsco) and PubMed (non-MEDLINE content) from 1 January 2000 to 16 January 2018, as well as the dissertation databases ProQuest Dissertations & Theses Global (ProQuest) and WorldCat dissertations from 1 January 2000 to 16 January 2018. In addition, we manually searched the reference lists of recent and relevant systematic reviews. Searches were limited to English and German language documents. An experienced information specialist developed a search strategy for Ovid/Medline MEDLINE, amended it to fit other electronic databases and performed all searches (see Supplementary file 1). In line with the peer review of the electronic search strategy (PRESS) statement[36], the Ovid MEDLINE search strategy was peer-reviewed by another information specialist.

Inclusion criteria. We included randomised controlled trials and prospective cohort or cross-sectional diagnostic studies assessing the performance of anthropometric tools (BMI, WC, WHR and WHtR) to determine obesity in adults (≥ 18 years) from any country. Our target population was adults aged 18 years from any country. We did not exclude studies with adults with diseases or disabilities that could have an impact on the body fat distribution. We used imaging techniques including computed tomography (CT), magnetic resonance imaging (MRI), dual energy X-ray absorptiometry (DXA) and ultrasound scanning (US) as reference standards because they are currently considered the most precise methods for assessing body composition[37]. We included studies that reported sensitivity, specificity, predictive values, likelihood ratios, diagnostic odds ratios, positivity

Characteristics	Inclusion	Exclusion
Population	Adults aged 18 years and above	Children, adolescents, and pregnant women
Index tests	Body mass index, waist circumference, waist-to-height ratio, waist-to-hip ratio	Any other anthropometric test, non-anthropometric tests
Target conditions	Overall and abdominal overweight, obesity, body fat distribution	Other conditions
Reference test	Imaging techniques including computed tomography, magnetic resonance imaging, dual energy X-ray absorptiometry, and ultrasound scanning	Any other tests measuring body composition including body volume/density tests, total body water or hydrometry tests, and impedance analyses
Outcomes	Sensitivity, specificity, predictive values, likelihood ratios, diagnostic odds ratios, positivity thresholds, ROC curves (including AUC)	Agreement measures
Study design	Randomised controlled trials (RCTs), prospective cohort studies, and cross sectional studies	All other study designs
Setting	Any setting	
Publication date	1/2000 onwards	Before 1/2000
Publication language	English, German	Other languages

Table 2. Eligibility criteria studies relevant to this systematic review.
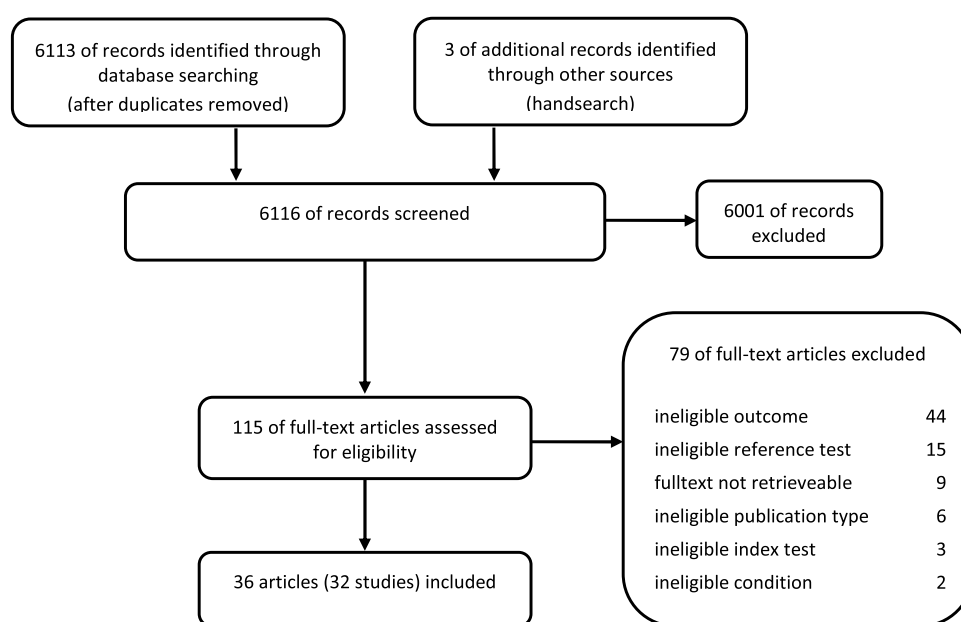

Figure 1. Flow chart of study selection process.

thresholds or receiver operating characteristics (ROC) curves (including area under the curve [AUC]) as outcomes. The eligibility criteria are described in more detail in Table 2.

Study selection. We developed and pilot-tested abstract and full-text review forms that reflected our inclusion and exclusion criteria. Two reviewers independently screened abstract and full-text articles and evaluated their eligibility for inclusion. Any discrepancies were resolved through discussion and consensus or by consultation with a third reviewer. The abstract and full-text reviews were carried out with Covidence (https://www.covidence.org/). Figure 1 summarises the flow of the literature review.

Data collection process and data items. We designed and pilot-tested a structured data abstraction form. One reviewer extracted data, and another checked for completeness and accuracy. For studies that met our inclusion criteria, we abstracted information related to (a) population; (b) index tests; (c) reference test; (d) obesity; (e) diagnostic values and f) funding source. We extracted or reconstructed the original classification data (2×2 table) at or close to WHO's recommended cut-offs (BMI: ≥ 30 kg/m^2, WC: ≥ 88 cm in women and ≥ 102 cm in men, WHR: 0.85 in women and 0.90 in men)[38] or utilised common definitions (body fat percentage: $> 35\%$ in women and $> 25\%$ in men) for further use in the meta-analyses. Otherwise, definitions of obesity as laid out in the articles were used. We contacted study authors via email if relevant data were not reported in an included publication.

Risk of bias and certainty of evidence assessment. Two independent reviewers assessed the risk of bias of diagnostic accuracy studies using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool[39]. We dually assessed the certainty of evidence for relevant outcomes using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach for diagnostic tests[40]. We resolved disagreements by discussion and consensus or by consulting a third reviewer.

Data synthesis. We conducted meta-analyses using the metandi command in STATA (version 15, Stata Corp.) when five or more studies were similar in terms of the index test, target condition and cut-offs used. The metandi command uses hierarchical logistic regression models to calculate meta-analyses of pairs of sensitivities and specificities. It displays the pooled estimates in both a bivariate and a hierarchical summary receiver operating characteristics (HSROC) model[41,42]. For each index test, we produced a paired forest plot of each study's sensitivity and specificity, as well as a plot of the sensitivities versus specificities in the ROC space. We assessed the heterogeneity by visually inspecting the CIs for sensitivity and specificity in the paired forest plots. For those index tests where we did not have sufficient studies to pool, we synthesised the data narratively.

Because of differences in the definition of the target condition between men and women, we conducted all analyses separately by sex. When information was available, we analysed the data by ethnicity. We further conducted sensitivity analyses to determine the impact of study quality on the robustness of the overall test performance measures. Subgroup analyses by age were not possible due to dissimilarities in the age categories in the studies.

Results

Our search yielded 6,116 records of which 32 studies (reported in 36 publications) met our a priori–defined eligibility criteria (Fig. 1)[43–78]. Twenty-seven studies (29 articles) assessed BMI[44,46–54,58–61,64–78], and 15 (19 articles) reported on waist measurements such as WC[43–47,49–53,57–60,62,63], WHR[43–47,49–51] and WHtR[51–56].

The majority of studies used DXA to evaluate anthropometric measurement tools[44,46,48,50–54,58–78], while four studies used CT[43,45,55–57,68] and three used MRI[46,47,49]. The cut-offs for obesity with DXA ranged from ≥ 30% to ≥ 43% body fat in women and from ≥ 20 to ≥ 34.6% in men. Studies using CT or MRI applied a cut-off of ≥ 100 cm^2 or ≥ 130 cm^2 of visceral adipose tissue area to provide reference data for both women and men. Fourteen studies were community-based[44,46,51,58,62–64,66–68,70,73,75,76,78], fourteen were primary care– or hospital-based[45,47,49,50,52–54,57,59–61,69,72,74,77], two were community- and hospital-based[43,48,55,56], one study was based in the army[65] and one did not report its setting[71]. Six studies included patients with various diseases or physical or cognitive disabilities[50,52–54,59,61,66,74]. Four studies stratified analyses by age groups[43,44,46,55,56,58,78]. The prevalence of obesity ranged widely from 5.7[51] to 95.8%[67]. Four studies did not report any prevalence data[45,47,60,73].

Of the included studies, we rated the risk of bias for six as low[52,53,58,60,63,67,70], for 16 as unclear[43–47,54–57,59,62,65,66,69,71,73–75,77] and for ten as high[48–51,61,64,68,72,76,78]. The reasons for the high risk of bias ratings included convenience sampling, inappropriate exclusion criteria for study participants and lack of predefined thresholds for index and reference tests. Eleven studies were conducted in Asia[44,46,47,49,51,60,63,68,70,71,75,78], ten studies in North America[50,52–54,58,65,66,72–74,77], eight in South America[43,45,55–57,59,61,62,67,69] and three[48,64,76] in Europe. Twenty-two studies were publicly funded, eight studies did not report their funding[48,49,64,65,68,69,71,76] and two received sponsoring from pharmaceutical companies[57,61]. Supplementary file 2 (Table S1) presents the characteristics of the included studies.

In the following sections, we first present the results of the four anthropometric measurement tools for women and men separately and if data allow, stratified by different ethnicities.

Body mass index (BMI). The 27 included studies[44,46–54,58–61,64–78] reported varying cut-offs for BMI to determine obesity. Thresholds ranged from 19.6 to 30 kg/m^2 for women and from 23.5 to 30 kg/m^2 for men. Studies often applied more than one cut-off to identify the threshold with the highest discriminative power. Cut-offs applied to Asian populations were generally lower than in other populations. Around 75% of studies from various countries, however, used the cut-off for obesity as suggested by WHO (≥ 30)[1] (see the characteristics of the included studies in Supplementary file 2, Table S1 and the results of all studies in Supplementary file 3, Table S1).

Based on a meta-analysis of 16 studies with data on 14,008 women of any race or ethnicity, the combined sensitivity of BMI (with thresholds from 25 to 30 kg/m^2) to detect obesity was 51.4% (95% CI 38.5–64.2%), with a corresponding specificity of 95.4% (95% CI 90.7–97.8%), as shown in the HSROC plot (Fig. 2). The HSROC plot shows the individual study estimates, a summary curve from the HSROC model, a summary estimate, a 95% confidence region for the summary estimate and a 95% prediction region. The confidence intervals of some studies failed to overlap for sensitivity, indicating considerable heterogeneity. The heterogeneity of the specificity was low (Fig. 3). A sensitivity analysis excluding studies with a high risk of bias had little impact on the results (sensitivity: 48.0% [95% CI 30.6–67.4%] and specificity: 96.1% [95% CI 88.3–98.8%]). When we excluded studies in Asian women from the meta-analyses, all the remaining studies (10 studies, n = 7,640) used a BMI cut-off of 30 kg/m^2. The results of the meta-analysis focusing on White, Latin or women of mixed ethnicity neither substantially altered the heterogeneity nor the summary estimates of the meta-analysis (sensitivity: 52.9% [95% CI 43.8–61.9%] and specificity: 97.0% [95% CI 90.8–99.1%]) (see Supplementary file 4, Figures S1 and S2). We rated the certainty of evidence of the pooled studies and considered it as very low for the sensitivity and as moderate for the specificity. The reasons for downgrading the certainty of evidence included the wide range and confidence intervals of the results for the sensitivity and the risk of bias for the specificity.

In men, the results of a meta-analysis including 12 studies with data on 11,320 men of any race or ethnicity show a combined sensitivity of 49.6% (95% CI 34.8–64.5%) and a specificity of 97.3% (95% CI 92.1–99.1%) for BMI cut-offs from 25 to 30 kg/m^2 (Fig. 2). The sensitivity varied considerably across studies (Fig. 4). A sensitivity
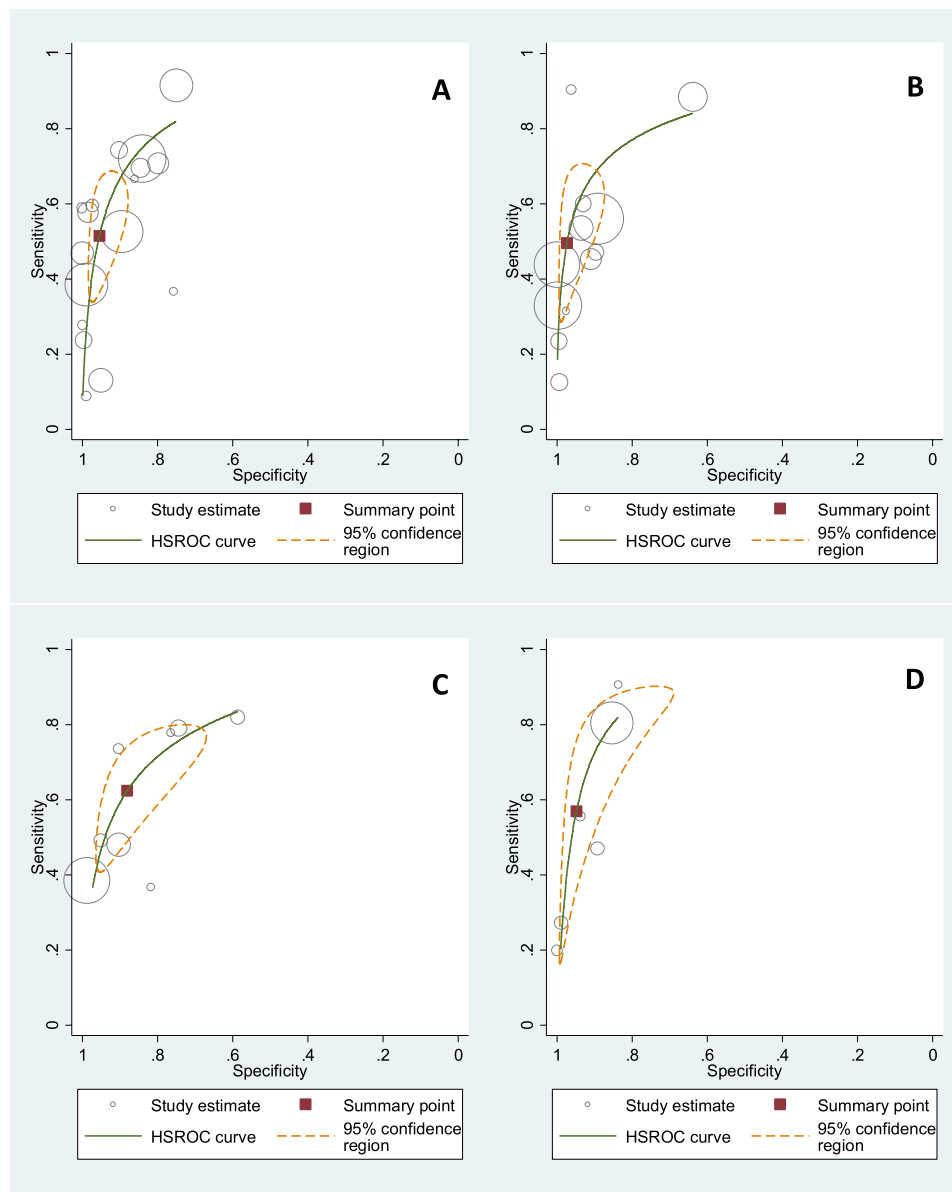
Figure 2. Summary ROC curve including summary point for Body Mass Index (BMI) in (**A**) Women (sensitivity: 51.4% [95%CI 38.5–64.2%]; specificity: 95.4% [95%CI 90.7–97.8%]) (**B**) Men (sensitivity: 49.6% [95% CI 34.8–64.5%]; specificity: 97.3% [95% CI 92.1–99.1%]) and for Waist Circumference (WC) in (**C**) Women (sensitivity: 62.4% [95% CI 49.2–73.9%] and for specificity 88.1% [95% CI 77.0–94.2%]) (**D**) men (sensitivity: 57.0% [95% CI 32.2–79.0] and for specificity 94.8% [95% CI 85.8–98.2%]).

analysis excluding studies with a high risk of bias had little impact on the results (52.4% [95% CI 28.6–75.1%] and specificity: 98.6% [95% CI 92.2–99.8%]). A subgroup analysis that excluded studies on Asian men and focused on men of White, Latin or mixed ethnicity (6 studies, n = 5,991, cuts-offs: 28.5–30 kg/m²) had little effect (sensitivity: 52.8% [95% CI 36.4–68.6%] and specificity: 98.9% [95% CI 93.8–99.8%]; see Supplementary file 4, Figures S1 and S2). We considered the certainty of evidence as very low for the sensitivity and moderate for the specificity. The reasons for downgrading the certainty of evidence included risk of bias as well as the wide range and confidence intervals of the sensitivity results.

Waist circumference (WC). For WC, the cut-offs to determine obesity in all 14 included studies[43–47,49–53,57–60,62,63] ranged from 65.8 to 107 cm in women and from 78.9 to 105 cm in men. Similar to studies assessing the performance of BMI, the analyses often applied more than one cut-off (see the characteristics of the included studies in Supplementary file 2, Table S1 and the results of all studies in Supplementary file 3, Table S2).

A meta-analysis including eight studies on 4,964 women rendered a sensitivity of 62.4% (95% CI 49.2–73.9%) and a specificity of 88.1% (95% CI 77.0–94.2%) for WC (80.5 to 92.3 cm) (Fig. 2). For both sensitivity and
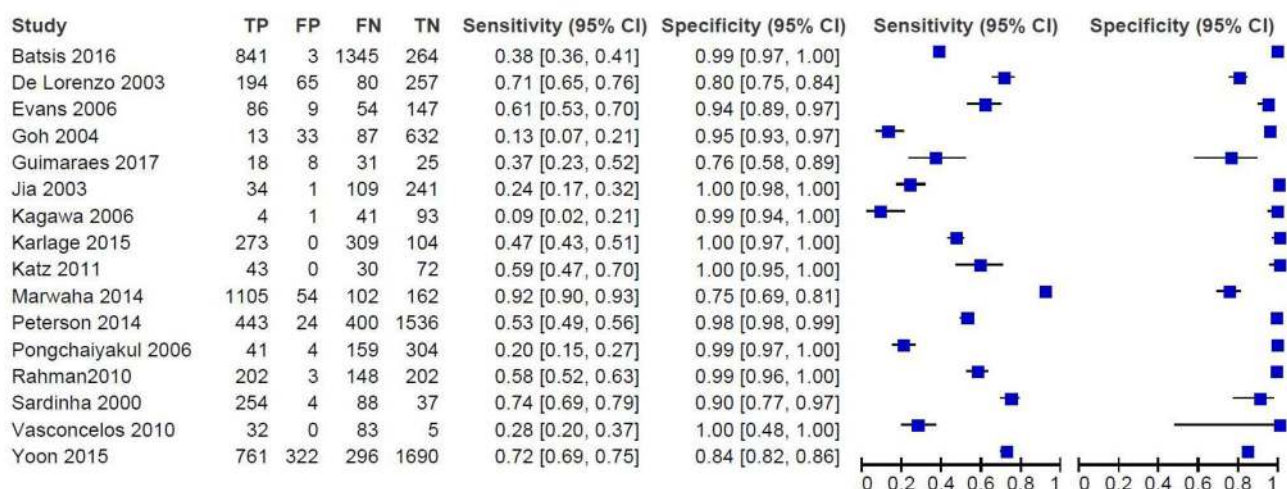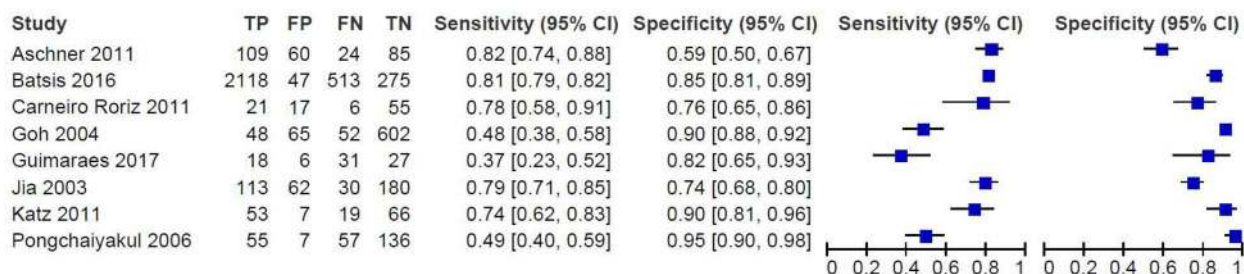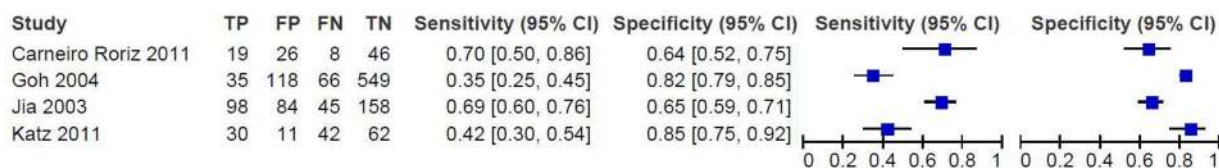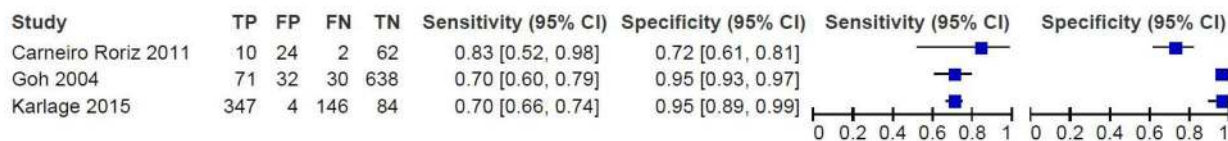
BMI Women

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Batsis 2016	841	3	1345	264	0.38 [0.36, 0.41]	0.99 [0.97, 1.00]
De Lorenzo 2003	194	65	80	257	0.71 [0.65, 0.76]	0.80 [0.75, 0.84]
Evans 2006	86	9	54	147	0.61 [0.53, 0.70]	0.94 [0.89, 0.97]
Goh 2004	13	33	87	632	0.13 [0.07, 0.21]	0.95 [0.93, 0.97]
Guimaraes 2017	18	8	31	25	0.37 [0.23, 0.52]	0.76 [0.58, 0.89]
Jia 2003	34	1	109	241	0.24 [0.17, 0.32]	1.00 [0.98, 1.00]
Kagawa 2006	4	1	41	93	0.09 [0.02, 0.21]	0.99 [0.94, 1.00]
Karlage 2015	273	0	309	104	0.47 [0.43, 0.51]	1.00 [0.97, 1.00]
Katz 2011	43	0	30	72	0.59 [0.47, 0.70]	1.00 [0.95, 1.00]
Marwaha 2014	1105	54	102	162	0.92 [0.90, 0.93]	0.75 [0.69, 0.81]
Peterson 2014	443	24	400	1536	0.53 [0.49, 0.56]	0.98 [0.98, 0.99]
Pongchaiyakul 2006	41	4	159	304	0.20 [0.15, 0.27]	0.99 [0.97, 1.00]
Rahman2010	202	3	148	202	0.58 [0.52, 0.63]	0.99 [0.96, 1.00]
Sardinha 2000	254	4	88	37	0.74 [0.69, 0.79]	0.90 [0.77, 0.97]
Vasconcelos 2010	32	0	83	5	0.28 [0.20, 0.37]	1.00 [0.48, 1.00]
Yoon 2015	761	322	296	1690	0.72 [0.69, 0.75]	0.84 [0.82, 0.86]

WC Women

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aschner 2011	109	60	24	85	0.82 [0.74, 0.88]	0.59 [0.50, 0.67]
Batsis 2016	2118	47	513	275	0.81 [0.79, 0.82]	0.85 [0.81, 0.89]
Carneiro Roriz 2011	21	17	6	55	0.78 [0.58, 0.91]	0.76 [0.65, 0.86]
Goh 2004	48	65	52	602	0.48 [0.38, 0.58]	0.90 [0.88, 0.92]
Guimaraes 2017	18	6	31	27	0.37 [0.23, 0.52]	0.82 [0.65, 0.93]
Jia 2003	113	62	30	180	0.79 [0.71, 0.85]	0.74 [0.68, 0.80]
Katz 2011	53	7	19	66	0.74 [0.62, 0.83]	0.90 [0.81, 0.96]
Pongchaiyakul 2006	55	7	57	136	0.49 [0.40, 0.59]	0.95 [0.90, 0.98]

WHR Women

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Carneiro Roriz 2011	19	26	8	46	0.70 [0.50, 0.86]	0.64 [0.52, 0.75]
Goh 2004	35	118	66	549	0.35 [0.25, 0.45]	0.82 [0.79, 0.85]
Jia 2003	98	84	45	158	0.69 [0.60, 0.76]	0.65 [0.59, 0.71]
Katz 2011	30	11	42	62	0.42 [0.30, 0.54]	0.85 [0.75, 0.92]

WHtR Women

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Carneiro Roriz 2011	10	24	2	62	0.83 [0.52, 0.98]	0.72 [0.61, 0.81]
Goh 2004	71	32	30	638	0.70 [0.60, 0.79]	0.95 [0.93, 0.97]
Karlage 2015	347	4	146	84	0.70 [0.66, 0.74]	0.95 [0.89, 0.99]

Figure 3. Paired forest-plots of sensitivity and specificity for body mass index (BMI), waist circumference (WC), waist-to hip ratio (WHR), waist to height ratio (WHtR) in women; data not weighted; TP, true positives; FP, false positives, FN, false negatives; TN, true negatives.

sensitivity, the heterogeneity of the included studies was high (Fig. 3). Excluding the study by Goh et al.[51] from the analysis because of its low cut-off (80.5 cm) did not substantially alter the results (cut-offs 86 cm to 92.3 cm; sensitivity: 64.4% [95% CI 50.1–76.6%] and specificity 88.0% [95% CI 74.5–94.8%]). Likewise, the results remained similar when excluding studies with a high risk of bias[49–51] (cut-offs 86–92.3 cm; sensitivity: 58.6% [95% CI 41.0–74.3%] and specificity 89.4% [95% CI 71.1–96.6]). Subgroup analysis on Latin women or women with mixed ethnicity (5 studies, n = 3,557, cuts-offs: 86–92.3 cm) reduced the heterogeneity and increased the sensitivity (73.4% [95% CI 52.5–87.4%]) but decreased the specificity (83.0% [95% CI 62.7–93.4%]; see Supplementary file 4, Figures S1 and S2). Because of methodological concerns and highly inconsistent and heterogeneous results, we rated the certainty of evidence as very low for both sensitivity and specificity.
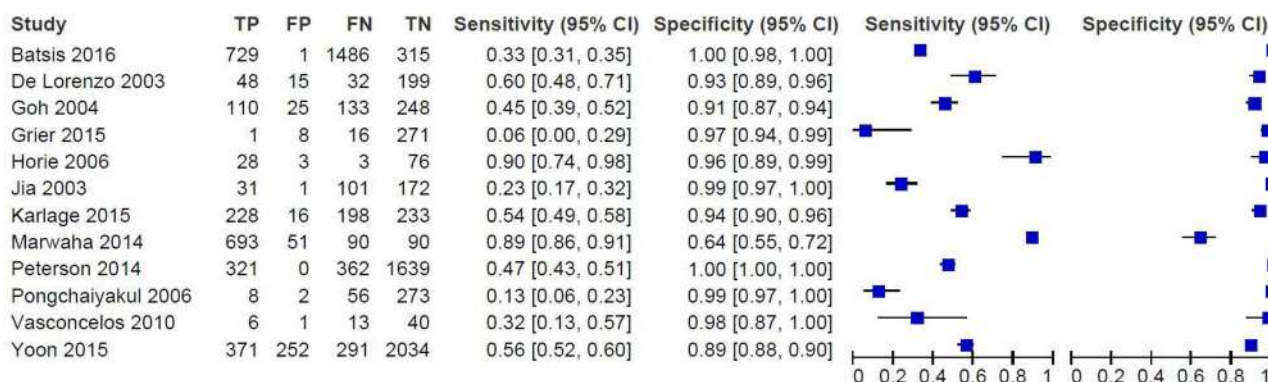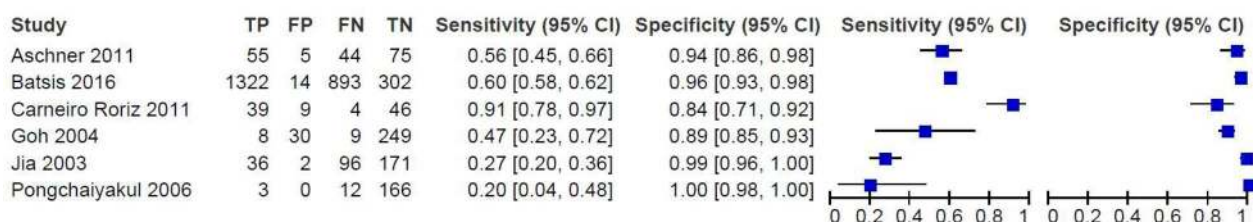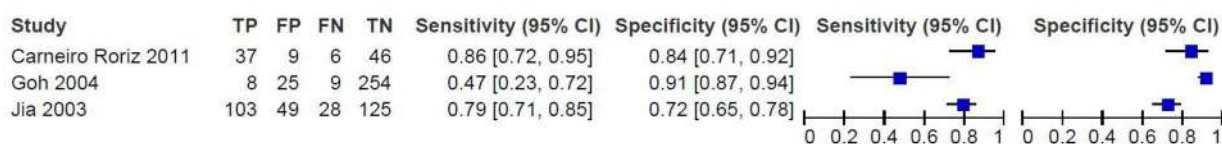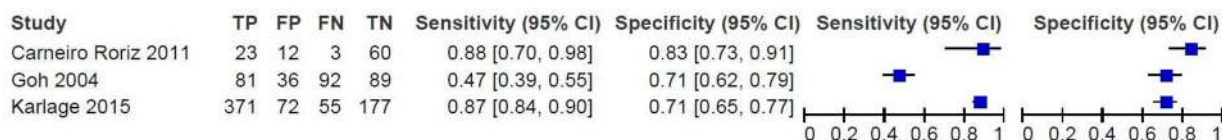
Figure 4. Paired forest-plots of sensitivity and specificity for body mass index (BMI), waist circumference (WC), waist-to hip ratio (WHR), waist to height ratio (WHtR) in men; data not weighted; TP, true positives; FP, false positives, FN, false negatives; TN, true negatives.

In men, the pooled estimates of six studies including 3,590 male participants were 57.0% (95% CI 32.2–79.0) for the sensitivity and 94.8% (95% CI 85.8–98.2%) for the specificity (Fig. 2). The cut-offs for WC ranged from 90.2 to 100.0 cm. The results of the included studies had a high heterogeneity for the sensitivity and a low heterogeneity for the specificity (Fig. 4). We were not able to perform sensitivity analyses or subgroup analysis due to the low number of studies included in the meta-analyses. Due to serious methodological concerns and highly inconsistent and heterogeneous results, we considered the certainty of evidence for the sensitivity as very low and for the specificity as low.

Waist-to-hip ratio (WHR). The cut-offs for determining obesity in the seven studies reporting on WHR ranged from 0.74 to 0.97 in women and from 0.85 to 0.96 in men[43–47,49–51] (see the characteristics of the included studies in Supplementary file 2, Table S1 and the results of all studies in Supplementary file 3, Table S3). We did not have enough data to conduct meta-analyses, as only four studies provided 2 × 2 tables[43,49–51]. In women, the sensitivities for WHR ranged from 34.4[51] to 92.3%[47], while the specificities ranged from 45.7[47] to 85.0%[51] (Fig. 3 and Supplementary file 3, Table S3). Two studies analysed the performance of WHR by age groups. Carneiro Roriz et al.[43] found sensitivity and specificity to be highest in young and middle-aged women (21–59 years)

in their study (n = 99) (sensitivity 0.83, specificity 0.72, no test for interaction). The study by Yang et al. and Li et al.[44,46] (n = 879) reported a higher sensitivity but a lower specificity in the 20–30 year-old age group compared to that in 31–45 year-old women (sensitivity 0.74 vs. 0.69, specificity 0.65 vs. 0.79, no test for interaction).

In men, the sensitivity ranged from 46.7[51] to 88.9%[44,46], and the specificity ranged from 25.0[47] to 90.9%[51] (Fig. 4). When stratifying the analysis by age groups, Carneiro Roriz et al.[43] found similar results in young and middle-aged men (21–59 years, n = 51) and elderly men (≥ 60 years, n = 47) (sensitivity 86.7% vs. 86.2%, specificity 83.3% vs. 83.3%, no test for interaction). Yang et al. and Li et al.[44,46] reported a similar sensitivity but a lower specificity in 31–45 year-old men (n = 185) compared to that in 20–30 year-old men (n = 694) (sensitivity 88.9% vs. 82.4%, specificity 64.1% vs. 78.4%, no test for interaction).

We rated the certainty of evidence for the sensitivity and specificity as very low in women and in men. The reasons for downgrading the certainty of the evidence related to methodological concerns, heterogeneous results and wide confidence intervals.

Waist-to-height ratio (WHtR).

We identified four studies[51–56] assessing WHtR. The data were insufficient to combine the results in a meta-analysis. Their cut-offs for defining obesity ranged from 0.50 to 0.59 in women and from 0.50 to 0.55 in men (see the characteristics of the included studies in Supplementary file 2, Table S1 and the results of all studies in Supplementary file 3, Table S4). In women, the sensitivity ranged from 51.0%[51] to 83.3%[55,56] and the specificity from 78.6%[55,56] to 95.2%[54] (Fig. 3 and Supplementary file 3, Table S4). The results in men were similar, from 46.7%[51] to 86.7%[54–56] for the sensitivity and from 71.0%[54] to 89.4%[51] for the specificity (Fig. 4 and Supplementary file 3, Table S4). Carneiro Roriz et al.[55,56] did not identify any differences in sensitivity and specificity between adults aged 20 to 59 years and adults aged 60 and older, irrespective of sex. Oreopoulos et al.[52,53] used slightly higher cut-offs for defining obesity (0.615 for women and 0.605 for men) in their study and reported a combined sensitivity of 77.4% and a specificity of 76.9%.

We rated the certainty of evidence for the sensitivity and specificity in both women and men as very low. The reasons for downgrading the certainty of evidence included methodological concerns, heterogeneous results and wide confidence intervals of the results.

Discussion

To the best of our knowledge, our work is the most recent and most comprehensive systematic review on the use of four anthropometric tools to determine obesity. Our findings, in general, indicate a lack of reliable scientific evidence on the performance of anthropometric tools to rule out or determine obesity as assessed by imaging techniques, which constitute the gold standard in obesity research until the concept of obesity is fully understood. Many of the included studies were fraught with methodological shortcomings. Consequently, we rated the certainty of evidence as low or very low, which indicates that we have little or very little confidence in the estimates of the effects.

The available studies focused mainly on BMI and WC to assess obesity. The pooled results of our meta-analyses consistently rendered low sensitivities and relatively high specificities for BMI and WC when compared to imaging techniques as reference standards. The sensitivities ranged from 49.6% (BMI for men) and 51.4% (BMI for women) to 57.0% (WC for men) and 62.4% (WC for women) in the pooled analyses. By contrast, the specificities ranged from 88.1% (WC for women) and 94.8% (WC for men) to 95.4% (BMI for women) and 97.3% (BMI for men).

These estimates are consistent with the findings from a previous systematic review by Okorodudu et al.[32], who reported a pooled sensitivity of 50% (95% CI 43–57%) and a pooled specificity of 90% (95% CI 86–94%) in their review of 25 studies. The studies included in this review went back to the 1990s and also used reference standards other than imaging techniques. For our systematic review, we employed more rigorous eligibility criteria than the Okordudu review[32] and included 17 additional studies that were published after the literature searches by Okordudu et al.

Our systematic review and the underlying evidence base have several notable limitations. A main limitation of the review is the substantial heterogeneity of the sensitivity estimates across studies. High heterogeneity is a common phenomenon in diagnostic test accuracy meta-analyses and is usually attributable to the spectrum effects and methodological shortcomings of the included studies. The subgroup analyses in our review that stratified meta-analyses by sex and by countries with predominantly White, Latin or mixed populations rendered similar estimates for BMI and WC as the overall analyses that also included Asian populations. We would have expected a difference, as WHO recommends lower cut-off values for Asian populations than for White populations[79]. Similarly, removing studies with a high risk of bias had little impact on the results. Nevertheless, many other factors, the impact of which we did not have sufficient data to explore, could have introduced heterogeneity. For example, the age of the participants, which varied widely among the studies, could have had an influence on the results. Without access to individual patient data, we were unable to assess the impact of age. Another potential source is the spectrum of prevalence rates among the studies (5.7[51] to 95.8%[67]). Studies with a higher disease prevalence most likely include more severely diseased patients, which ultimately leads to a better test performance in this population.

Heterogeneity could also stem from the use of different cut-offs both for determining obesity with the anthropometric measurement tools and for the reference tests in the primary studies. For BMI and WC, the majority of studies adhered to the cut-offs recommended by WHO (BMI: ≥ 30 kg/m², WC: ≥ 88 cm in women and ≥ 102 cm in men)[1,38]. The cut-offs for the reference tests ranged from ≥ 30% to ≥ 43% body fat in women and from ≥ 20% to ≥ 34.6% in men using DXA, with most studies referring to a body fat percentage > 35% in women and > 25% in men as the standards for defining obesity. Even though these cut-offs are widely applied and recommended, it is important to note that they were chosen arbitrarily and lack sound scientific basis[9,80,81]. For example, BMI

thresholds have only been based upon visual inspection of the relationship between BMI and mortality[82]. For body fat percentage, there is little evidence supporting the cut-offs due to the lack of studies investigating the relationship between a continuum of body fat percentage values and cardiometabolic disease and mortality[9,80]. In addition to the heterogeneity that is introduced by the application of various cut-off values, the cut-offs themselves remain an issue of debate and should be the focus of future research. However, their validity goes beyond the scope of this review.

The use of various imaging techniques, including DXA, CT, and MRI, could have led to differences in performance estimates. However, imaging techniques are currently considered to be the most accurate tools for body composition analysis because of their ability to accurately discriminate tissues[37,83]. We excluded all studies that used other reference standards, such as bioelectrical impedance analysis or dilution techniques, to increase homogeneity. Also limiting this review is the absence of a "gold standard" to diagnose obesity. Although imaging techniques are generally able to produce good-quality body composition data, they all have their shortcomings. For example, DXA does not differentiate between types of fat. Silver further argues that an accurate body composition analysis measuring excess body fat is insufficient for diagnosing obesity; it would rather need a tool that translates the interplay between body composition and metabolic risks into a new concept of obesity[84]. Nonetheless, until research has elucidated that interplay, obesity assessment relies on body composition data.

Finally, another major limitation of the underlying evidence base is the low methodological quality of the included studies that, together with the inconsistency and heterogeneity of the results, has contributed to the mostly low or very low confidence that we have in the evidence. We rated only six out of the 32 included studies as a low risk of bias. Many studies included convenience sampling, used inappropriate exclusion criteria for study participants, lacked predefined cut-offs for index and reference tests and failed to provide information about the numbers of participants included in the analysis.

The strengths of this systematic review include a comprehensive search strategy in four electronic databases combined with manual reference checking of pertinent research articles and a search for unpublished research studies. The search strategy was peer-reviewed by an additional information specialist. We contacted the authors of the included studies to receive the data of 2 × 2 tables when not reported. During the whole systematic review process, we followed Cochrane methods[34], which are known to be methodologically sound and rigorous. Despite these efforts, we cannot entirely rule out the possibility that we have overlooked a relevant research study.

The findings of our review should be interpreted cautiously within the context of clinical practice. Thresholds between normal weight, overweight, and obesity are arbitrary and not based on universally agreed upon standards. Our review emphasises the substantial uncertainties that obesity assessment with anthropometric tools bring with them. Methodologically sound studies with appropriate sampling strategies, predefined and valid cut-offs and complete analyses are needed for firm conclusions. Future research should focus on studies that differentiate between age groups, are conducted in a European setting and examine the combined use of anthropometric tools.

Conclusions
This systematic review shows that BMI and WC have serious limitations for use as obesity screening tools in clinical practice despite their widespread use, and no evidence supports that WHR and WHtR are more suitable than BMI or WC to access body fat. However, due to the lack of alternatives, BMI and WC might still have a role as initial tools for assessing individuals for excess adiposity until new evidence emerges. Nonetheless, one should be aware of the limitations of these tools when interpreting the results. In some clinical circumstances, particularly for BMI or WC results that are borderline between overweight and obesity, it might be useful to conduct further examinations of obesity-related risk factors or to confirm results with imaging techniques (e.g. DXA scans).

Data availability
2 × 2 tables that support the findings of this study are available from the first author (IS) upon reasonable request.

References
1. World Health Organization. *Obesity and Overweight—Fact Sheet*, https://www.who.int/mediacentre/factsheets/fs311/en/ (2017).
2. World Health Organization. *Global Health Observatory Data Repository*, https://apps.who.int/gho/data/node.main.A896?lang=en (2017).
3. Guh, D. P. *et al.* The incidence of co-morbidities related to obesity and overweight: A systematic review and meta-analysis. *BMC Public Health* **9**, 88. https://doi.org/10.1186/1471-2458-9-88 (2009).
4. Luppino, F. S. *et al.* Overweight, obesity, and depression: A systematic review and meta-analysis of longitudinal studies. *Arch. Gen. Psychiatry* **67**, 220–229. https://doi.org/10.1001/archgenpsychiatry.2010.2 (2010).
5. Campbell-Scherer, D. & Sharma, A. M. Improving obesity prevention and management in primary care in Canada. *Curr. Obes. Rep.* **5**, 327–332. https://doi.org/10.1007/s13679-016-0222-y (2016).
6. Rodondi, N. *et al.* Counselling overweight and obese patients in primary care: A prospective cohort study. *Eur. J. Cardiovas. Prev. Rehabil.* **13**, 222–228. https://doi.org/10.1097/01.hjr.0000209819.13196.a4 (2006).
7. Hite, A., Victorson, D., Elue, R. & Plunkett, B. A. An exploration of barriers facing physicians in diagnosing and treating obesity. *Am. J. Health Promot.* https://doi.org/10.1177/0890117118784227 (2018).
8. Hebert, J. R., Allison, D. B., Archer, E., Lavie, C. J. & Blair, S. N. Scientific decision making, policy decisions, and the obesity pandemic. *Mayo Clin. Proc.* **88**, 593–604. https://doi.org/10.1016/j.mayocp.2013.04.005 (2013).
9. Oliveros, E., Somers, V. K., Sochor, O., Goel, K. & Lopez-Jimenez, F. The concept of normal weight obesity. *Prog. Cardiovasc. Dis.* **56**, 426–433. https://doi.org/10.1016/j.pcad.2013.10.003 (2014).

10. Mechanick, J. I., Hurley, D. L. & Garvey, W. T. Adiposity-based chronic disease as a new diagnostic term: The American Association of Clinical Endocrinologists and American College of Endocrinology Position Statement. *Endocr. Pract.* **23**, 372–378. https://doi.org/10.4158/ep161688.Ps (2017).
11. Garvey, W. T. *et al.* American association of clinical endocrinologists and American College of endocrinology position statement on the 2014 advanced framework for a new diagnosis of obesity as a chronic disease. *Endocr. Pract.* **20**, 977–989. https://doi.org/10.4158/ep14280.Ps (2014).
12. Garvey, W. T. & Mechanick, J. I. Proposal for a scientifically correct and medically actionable disease classification system (ICD) for obesity. *Obesity* **28**, 484–492. https://doi.org/10.1002/oby.22727 (2020).
13. Frühbeck, G. *et al.* The ABCD of Obesity: An EASO position statement on a diagnostic term with clinical and scientific implications. *Obesity Facts* **12**, 131–136. https://doi.org/10.1159/000497124 (2019).
14. Cornier, M.-A. *et al.* Assessing adiposity. *Circulation* **124**, 1996–2019. https://doi.org/10.1161/CIR.0b013e318233bc6a (2011).
15. Mongraw-Chaffin, M. L., Peters, S. A. E., Huxley, R. R. & Woodward, M. The sex-specific relationship between body mass index and coronary heart disease: A systematic review and meta-analysis of 95 cohorts with 12 million participants. *Lancet Diabetes Endocrinol.* **3**, 437–449. https://doi.org/10.1016/S2213-8587(15)00086-8 (2015).
16. Flegal, K. M., Kit, B. K., Orpana, H. & Graubard, B. I. Association of all-cause mortality with overweight and obesity using standard body mass index categories: A systematic review and meta-analysis. *JAMA* **309**, 71–82. https://doi.org/10.1001/jama.2012.113905 (2013).
17. Auyeung, T. W. *et al.* Survival in older men may benefit from being slightly overweight and centrally obese—A 5-year follow-up study in 4,000 older adults using DXA. *J. Gerontol. A Biol. Sci. Med. Sci.* **65**, 99–104. https://doi.org/10.1093/gerona/glp099 (2010).
18. Lee, J. S. *et al.* Obesity can benefit survival-a 9-year prospective study in 1614 Chinese nursing home residents. *J. Am. Med. Dir. Assoc.* **15**, 342–348. https://doi.org/10.1016/j.jamda.2013.12.081 (2014).
19. Shil Hong, E. *et al.* Counterintuitive relationship between visceral fat and all-cause mortality in an elderly Asian population. *Obesity* **23**, 220–227. https://doi.org/10.1002/oby.20914 (2015).
20. Rush, E. C. *et al.* BMI, fat and muscle differences in urban women of five ethnicities from two countries. *Int. J. Obes. (Lond.)* **31**, 1232–1239. https://doi.org/10.1038/sj.ijo.0803576 (2007).
21. Thomas, E. L. *et al.* The missing risk: MRI and MRS phenotyping of abdominal adiposity and ectopic fat. *Obesity* **20**, 76–87. https://doi.org/10.1038/oby.2011.142 (2012).
22. Stenholm, S. *et al.* Sarcopenic obesity: Definition, cause and consequences. *Curr. Opin. Clin. Nutr. Metab. Care* **11**, 693–700. https://doi.org/10.1097/MCO.0b013e328312c37d (2008).
23. Bouillanne, O. *et al.* Fat mass protects hospitalized elderly persons against morbidity and mortality. *Am. J. Clin. Nutr.* **90**, 505–510. https://doi.org/10.3945/ajcn.2009.27819 (2009).
24. Lee, J. S. *et al.* Survival benefit of abdominal adiposity: A 6-year follow-up study with dual X-ray absorptiometry in 3,978 older adults. *Age* **34**, 597–608. https://doi.org/10.1007/s11357-011-9272-y (2012).
25. Han, S. S. *et al.* Lean mass index: A better predictor of mortality than body mass index in elderly Asians. *J. Am. Geriatr. Soc.* **58**, 312–317. https://doi.org/10.1111/j.1532-5415.2009.02672.x (2010).
26. Genton, L., Graf, C. E., Karsegard, V. L., Kyle, U. G. & Pichard, C. Low fat-free mass as a marker of mortality in community-dwelling healthy elderly subjects. *Age Ageing* **42**, 33–39. https://doi.org/10.1093/ageing/afs091 (2013).
27. Britton, K. A. *et al.* Body fat distribution, incident cardiovascular disease, cancer, and all-cause mortality. *J. Am. Coll. Cardiol.* **62**, 921–925. https://doi.org/10.1016/j.jacc.2013.06.027 (2013).
28. Ding, J. *et al.* The association of pericardial fat with incident coronary heart disease: The Multi-Ethnic Study of Atherosclerosis (MESA). *Am. J. Clin. Nutr.* **90**, 499–504. https://doi.org/10.3945/ajcn.2008.27358 (2009).
29. 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.* **13**, 275–286. https://doi.org/10.1111/j.1467-789X.2011.00952.x (2012).
30. Carmienke, S. *et al.* General and abdominal obesity parameters and their combination in relation to mortality: A systematic review and meta-regression analysis. *Eur. J. Clin. Nutr.* **67**, 573–585 (2013).
31. Corrêa, M. M., Thumé, E., De Oliveira, E. R. A. & Tomasi, E. Performance of the waist-to-height ratio in identifying obesity and predicting non-communicable diseases in the elderly population: A systematic literature review. *Arch. Gerontol. Geriatr.* **65**, 174–182 (2016).
32. Okorodudu, D. O. *et al.* Diagnostic performance of body mass index to identify obesity as defined by body adiposity: A systematic review and meta-analysis. *Int. J. Obes. (Lond.)* **34**, 791–799. https://doi.org/10.1038/ijo.2010.5 (2010).
33. McTigue, K. M., Hess, R. & Ziouras, J. Obesity in older adults: A systematic review of the evidence for diagnosis and treatment. *Obesity* **14**, 1485–1497. https://doi.org/10.1038/oby.2006.171 (2006).
34. Deeks, J., Wisniewski, S. & Davenport, C. in *Chapter 4: Guide to the contents of a Cochrane Diagnostic Test Accuracy Protocol* (eds JJ Deeks, PM Bossuyt, & C Gatsonis) (The Cochrane Collaboration, 2013).
35. McInnes, M. D. F. *et al.* Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA Statement. *JAMA* **319**, 388–396. https://doi.org/10.1001/jama.2017.19163 (2018).
36. McGowan, J. *et al.* PRESS peer review of electronic search strategies: 2015 Guideline statement. *J. Clin. Epidemiol.* **75**, 40–46. https://doi.org/10.1016/j.jclinepi.2016.01.021 (2016).
37. Prado, C. M. M. & Heymsfield, S. B. Lean tissue imaging: A new era for nutritional assessment and intervention. *J. Parent. Enteral Nutr.* **38**, 940–953. https://doi.org/10.1177/0148607114550189 (2014).
38. World Health Organization. *Waist Circumference and Waist-Hip Ratio: Report of a WHO Expert Consultation, Geneva, 8–11 December 2008* (World Health Organization, Geneva, 2008).
39. Whiting, P. F. *et al.* QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann. Intern. Med.* **155**, 529–536. https://doi.org/10.7326/0003-4819-155-8-201110180-00009 (2011).
40. Schünemann, H. J. *et al.* Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* **336**, 1106–1110 (2008).
41. Reitsma, J. B. *et al.* Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J. Clin. Epidemiol.* **58**, 982–990. https://doi.org/10.1016/j.jclinepi.2005.02.022 (2005).
42. Rutter, C. M. & Gatsonis, C. A. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat. Med.* **20**, 2865–2884 (2001).
43. Carneiro Roriz, A. K. *et al.* Methods of predicting visceral fat in Brazilian adults and older adults: A comparison between anthropometry and computerized tomography. *Arch. Latinoam Nutr* **61**, 5–12 (2011).
44. Yang, F. *et al.* Receiver-operating characteristic analyses of body mass index, waist circumference and waist-to-hip ratio for obesity: Screening in young adults in central south of China. *Clin. Nutr.* **25**, 1030–1039. https://doi.org/10.1016/j.clnu.2006.04.009 (2006).
45. Ribeiro-Filho, F. F., Faria, A. N., Azjen, S., Zanella, M. T. & Ferreira, S. R. Methods of estimation of visceral fat: Advantages of ultrasonography. *Obes. Res.* **11**, 1488–1494. https://doi.org/10.1038/oby.2003.199 (2003).
46. Li, L. M. *et al.* Anthropometric indices as the predictors of trunk obesity in Chinese young adults: Receiver operating characteristic analyses. *Ann. Hum. Biol.* **35**, 342–348. https://doi.org/10.1080/03014460802027049 (2008).
47. Gong, W. *et al.* A comparison of ultrasound and magnetic resonance imaging to assess visceral fat in the metabolic syndrome. *Asia Pac. J. Clin. Nutr.* **16**(Suppl 1), 339–345 (2007).

48. Donini, L. M. *et al.* How to estimate fat mass in overweight and obese subjects. *Int.J. Endocrinol. Print* **2013**, 285680. https://doi.org/10.1155/2013/285680 (2013).

49. Jia, W. P. *et al.* Prediction of abdominal visceral obesity from body mass index, waist circumference and waist-hip ratio in Chinese adults: receiver operating characteristic curves analysis. *Biomed. Environ. Sci* **16**, 206–211 (2003).

50. Katz, P. *et al.* Obesity and its measurement in a community-based sample of women with systemic lupus erythematosus. *Arthritis Care Res (Hoboken)* **63**, 261–268. https://doi.org/10.1002/acr.20343 (2011).

51. Goh, V. H., Tain, C. F., Tong, T. Y., Mok, H. P. & Wong, M. T. Are BMI and other anthropometric measures appropriate as indices for obesity? A study in an Asian population. *J. Lipid Res.* **45**, 1892–1898. https://doi.org/10.1194/jlr.M400159-JLR200 (2004).

52. Oreopoulos, A. *Exploring the Associations Between the Obesity Paradox, Body Composition and Prognostic Factors in Chronic Heart Failure* NR71203 thesis (University of Alberta, Canada, 2010).

53. Oreopoulos, A. *et al.* Do anthropometric indices accurately reflect directly measured body composition in men and women with chronic heart failure? *Congest. Heart Fail.* **17**, 90–92. https://doi.org/10.1111/j.1751-7133.2010.00204.x (2011).

54. Karlage, R. E. *et al.* Validity of anthropometric measurements for characterizing obesity among adult survivors of childhood cancer: A report from the St. Jude Lifetime Cohort Study. *Cancer* **121**, 2036–2043. https://doi.org/10.1002/cncr.29300 (2015).

55. Carneiro Roriz, A. K. *et al.* Discriminatory power of indicators predictors of visceral adiposity evaluated by computed tomography in adults and elderly individuals. *Nutr. Hosp.* **29**, 1401–1407, https://doi.org/10.3305/nh.2014.29.6.7185 (2014).

56. Carneiro Roriz, A. K. *et al.* Evaluation of the accuracy of anthropometric clinical indicators of visceral fat in adults and elderly. *PLoS ONE* **9**. https://doi.org/10.1371/journal.pone.0103499 (2014).

57. Aschner, P. *et al.* Determination of the cutoff point for waist circumference that establishes the presence of abdominal obesity in Latin American men and women. *Diabetes Res. Clin. Pract.* **93**, 243–247. https://doi.org/10.1016/j.diabres.2011.05.002 (2011).

58. Batsis, J. A. *et al.* Diagnostic accuracy of body mass index to identify obesity in older adults: NHANES 1999–2004. *Int. J. Obes.* **40**, 761–767. https://doi.org/10.1038/ijo.2015.243 (2016).

59. Guimaraes, M., Pinto, M., Raid, R., Andrade, M. V. M. & Kakehasi, A. M. Which is the best cutoff of body mass index to identify obesity in female patients with rheumatoid arthritis? A study using dual energy X-ray absorptiometry body composition. *Review* **57**, 279–285. https://doi.org/10.1016/j.rbre.2016.02.008 (2017).

60. Chen, Y. M., Ho, S. C., Lam, S. S. & Chan, S. S. Validity of body mass index and waist circumference in the classification of obesity as compared to percent body fat in Chinese middle-aged women. *Int. J. Obes.* **30**, 918–925. https://doi.org/10.1038/sj.ijo.0803220 (2006).

61. dos Santos Diniz, M., Couto Bavoso, N., Kakehasi, A. M., Weissheimer Lauria, M. S. S., M. M. & Machado-Pinto, J. Assessment of adiposity in psoriatic patients by dual energy X-ray absorptiometry compared to conventional methods. *Anais Brasileiros de Dermatologia* **91**, 150–155, https://doi.org/10.1590/abd1806-4841.20164082 (2016).

62. de Oliveira, A. *et al.* Waist circumference measures: Cutoff analyses to detect obesity and cardiometabolic risk factors in a South-east Brazilian middle-aged men population—A cross-sectional study. *Lip. Health Dis.* **13**, 141, https://doi.org/10.1186/1476-511X-13-141 (2014).

63. Pongchaiyakul, C., Pongchaiyakul, C., Wanothayaroj, E., Nguyen, T. V. & Rajatanavin, R. Association between waist circumference and percentage body fat among rural Thais. *J. Med. Assoc. Thai.* **89**, 1592–1600 (2006).

64. De Lorenzo, A. *et al.* How fat is obese?. *Acta Diabetol.* **40**, s254–s257. https://doi.org/10.1007/s00592-003-0079-x (2003).

65. Grier, T., Canham-Chervak, M., Sharp, M. & Jones, B. H. Does body mass index misclassify physically active young men. *Prev .Med.Rep.* **2**, 483–487. https://doi.org/10.1016/j.pmedr.2015.06.003 (2015).

66. Peterson, M. D., Al Snih, S., Stoddard, J., Shekar, A. & Hurvitz, E. A. Obesity misclassification and the metabolic syndrome in adults with functional mobility impairments: Nutrition Examination Survey 2003–2006. *Prevent. Med.* **60**, 71–76, https://doi.org/10.1016/j.ypmed.2013.12.014 (2014).

67. Vasconcelos Fde, A., Cordeiro, B. A., Rech, C. R. & Petroski, E. L. Sensitivity and specificity of the body mass index for the diagnosis of overweight/obesity in elderly. *Cad Saude Publica* **26**, 1519–1527 (2010).

68. Horie, N., Komiya, H., Mori, Y. & Tajima, N. New body mass index criteria of central obesity for male Japanese. *Tohoku J. Exp. Med.* **208**, 83–86 (2006).

69. Tello-Winniczuk, N. *et al.* Value of body mass index in the diagnosis of obesity according to DEXA in well-controlled RA patients. *Reumatol* **13**, 17–20. https://doi.org/10.1016/j.reuma.2016.02.003 (2017).

70. Pongchaiyakul, C. *et al.* Defining obesity by body mass index in the Thai population: an epidemiologic study. *Asia Pac. J. Clin. Nutr.* **15**, 293–299 (2006).

71. Kagawa, M., Uenishi, K., Kuroiwa, C., Mori, M. & Binns, C. W. Is the BMI cut-off level for Japanese females for obesity set too high? A consideration from a body composition perspective. *Asia Pac. J. Clin. Nutr.* **15**, 502–507 (2006).

72. Evans, E. M., Rowe, D. A., Racette, S. B., Ross, K. M. & McAuley, E. Is the current BMI obesity classification appropriate for black and white postmenopausal women? *Int. J. Obes.* **30**, 837–843. https://doi.org/10.1038/sj.ijo.0803208 (2006).

73. Blew, R. M. *et al.* Assessing the validity of body mass index standards in early postmenopausal women. *Obes. Res.* **10**, 799–808. https://doi.org/10.1038/oby.2002.108 (2002).

74. Temple, V. A., Walkley, J. W. & Greenway, K. Body mass index as an indicator of adiposity among adults with intellectual disability. *J. Intellect. Dev. Disabil.* **35**, 116–120. https://doi.org/10.3109/13668251003694598 (2010).

75. Marwaha, R. K. *et al.* Normative data of body fat mass and its distribution as assessed by DXA in Indian adult population. *J. Clin. Densitom.* **17**, 136–142. https://doi.org/10.1016/j.jocd.2013.01.002 (2014).

76. Sardinha, L. B. & Teixeira, P. J. Obesity screening in older women with the body mass index: A receiver operating characteristic (ROC) analysis. *Sci. Sports* **15**, 212–219. https://doi.org/10.1016/S0765-1597(00)80008-8 (2000).

77. Rahman, M. & Berenson, A. B. Accuracy of current body mass index obesity classification for white, black, and Hispanic reproductive-age women. *Obstet. Gynecol.* **115**, 982–988. https://doi.org/10.1097/AOG.0b013e3181da9423 (2010).

78. Yoon, J. L., Cho, J. J., Park, K. M., Noh, H. M. & Park, Y. S. Diagnostic performance of body mass index using the Western Pacific Regional Office of World Health Organization reference standards for body fat percentage. *J. Korean Med. Sci.* **30**, 162–166. https://doi.org/10.3346/jkms.2015.30.2.162 (2015).

79. Bassett, J. *The Asia-Pacific Perspective: Redefining Obesity and Its Treatment* (International Diabetes Institute, World Health Organization Regional Office for the Western Pacific, International Association for the Study of Obesity, International Obesity Task Force, Melbourne, 2000).

80. Snitker, S. Use of body fatness cutoff points. *Mayo Clinic Proc.* **85**, 1057; author reply 1057–1058. https://doi.org/10.4065/mcp.2010.0583 (2010).

81. Ho-Pham, L. T., Campbell, L. V. & Nguyen, T. V. More on body fat cutoff points. *Mayo Clinic Proc.* **86**, 584; author reply 584–585. https://doi.org/10.4065/mcp.2011.0097 (2011).

82. World Health Organization. *Physical status: The use and interpretation of anthropometry. Report of a WHO Expert Committee*. WHO Technical Report Series 854 (Geneva, 1995).

83. Smith, S. & Madden, A. M. Body composition and functional assessment of nutritional status in adults: A narrative review of imaging, impedance, strength and functional techniques. *J. Hum. Nutr. Diet.* **29**, 714–732. https://doi.org/10.1111/jhn.12372 (2016).

84. Silver, H. J., Welch, E. B., Avison, M. J. & Niswender, K. D. Imaging body composition in obesity and weight loss: Challenges and opportunities. *Diabetes Metab. Syndr. Obes.* **3**, 337–347. https://doi.org/10.2147/DMSOTT.S9454 (2010).

85. World Health Organization. *The WHO STEPwise approach to noncommunicable disease risk factor surveillance.* https://www.who.int/ncds/surveillance/steps/STEPS_Manual.pdf?ua=1 (2017).

Author contributions
IS, BT, BNS, and GG designed the study; IK performed the literature searches; IS, MS, BT, BNS, and VT conducted the review and extracted the data; GG performed the meta-analysis; and IS, BT, and BNS prepared the first draft of the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no financial or non-financial competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-69498-7.

Correspondence and requests for materials should be addressed to I.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.