BMJ Open

Comparison of different severe obesity definitions in predicting future cardiometabolic risk in a longitudinal cohort of children

Lisa Kakinami ,[1,2] Anna Smyrnova,[2] Gilles Paradis,[3] Angelo Tremblay,[4] Melanie Henderson [5,6]

To cite: Kakinami L, Smyrnova A, Paradis G, *et al*. Comparison of different severe obesity definitions in predicting future cardiometabolic risk in a longitudinal cohort of children. *BMJ Open* 2022;**12**:e058857. doi:10.1136/bmjopen-2021-058857

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-058857).

Received 01 November 2021
Accepted 06 May 2022

Check for updates

© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Dr Lisa Kakinami;
lisa.kakinami@concordia.ca

ABSTRACT

Objectives Severe obesity (SO) prevalence varies between reference curve-based definitions (WHO: ≥99th percentile, Centers for Disease Control and Prevention (CDC): >1.2×95th percentile). Whether SO definitions differentially predict cardiometabolic disease risk is critical for proper clinical care and management but is unknown.

Design Prospective cohort study

Setting SO definitions were applied at baseline (2005–2008, M_{age}=9.6 years, n=548), and outcomes (fasting lipids, glucose, homoeostatic model assessment (HOMA-IR) and blood pressure) were assessed at first follow-up (F1: 2008–2011, M_{age}=11.6 years) and second follow-up (2015–2017, M_{age}=16.8 years) of the Quebec Adipose and Lifestyle Investigation in Youth cohort in Montreal, Quebec.

Participants Respondents were youth who had at least one biological parent with obesity.

Primary outcome measures Unfavourable cardiometabolic levels of fasting blood glucose (≥6.1 mmol/L), insulin resistance (HOMA-IR index ≥2.0), high-density lipoprotein <1.03 mmol/L, low-density lipoprotein ≥2.6 mmol/L and triglycerides ≥1.24 mmol/L. Unfavourable blood pressure was defined as ≥90th percentile for age-adjusted, sex-adjusted and height-adjusted systolic or diastolic blood pressure.

Analysis Area under the receiver operating characteristic curve (AUC) and McFadden psuedo R^2 for predicting F1 or F2 unfavourable cardiometabolic levels from baseline SO definitions were calculated. Agreement was assessed with kappas.

Results Baseline SO prevalence differed (WHO: 18%, CDC: 6.7%). AUCs ranged from 0.52 to 0.77, with fair agreement (kappa=37%–55%). WHO-SO AUCs for detecting unfavourable HOMA-IR (AUC>0.67) and high-density lipoprotein (AUC>0.59) at F1 were statistically superior than CDC-SO (AUC>0.59 and 0.53, respectively; p<0.05). Only HOMA-IR and the presence of more than three risk factors had acceptable model fit. WHO-SO was not more predictive than WHO-obesity, but CDC-SO was statistically inferior to CDC-obesity.

Conclusion WHO-SO is statistically superior at predicting cardiometabolic risk than CDC-SO. However, as most AUCs were generally uninformative, and obesity definitions were the same if not better than SO, the improvement may not be clinically meaningful.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This is the first study to assess how well the Centers for Disease Control and Prevention or the WHO definitions of severe obesity could predict future cardiometabolic risk in youth.
⇒ Participants were from a large cohort sample of youth with baseline and two follow-up measures of fasting lipids and cardiometabolic data available.
⇒ The study used a non-random convenience cohort sample and may not be generalisable to all youth.

INTRODUCTION

Approximately 4%–6% of youth have severe obesity (SO).[1] Youth with SO are at greater future cardiometabolic risk compared with those with overweight and obesity,[2 3] and may thus be a distinctive sub-class of youth from those with obesity. In 2005, the American Medical Association, Health Resources and Service Administration, and the Centers for Disease Control and Prevention (CDC) convened an expert committee comprised 15 professional organisations to update recommendations for detecting and treating obesity during childhood and adolescence. Part of these updates included defining SO as ≥99th percentile for the CDC growth curves.[4] However, the method used for developing the CDC growth curve only allows for the calculation of percentiles between the 3rd and 97th percentiles. Values outside of this are extrapolated. Thus the definition of SO has since been modified to 1.2 times the 95th percentile[5]; ≥99th percentile is no longer recommended.[1 6 7]

In contrast, the WHO growth curve does not limit extrapolations above the 97th percentile.[8] Hence, the Canadian Paediatric Society (in collaboration with the Dietitians of Canada, the College of Family Physicians of Canada and Community Health Nurses of

Canada) recommends defining SO as ≥99th percentile (as a rounded percentile of the 99.9th) using the WHO growth curves.[9] However, the empirical evidence for these different recommendations is lacking. To the best of our knowledge, only a single study investigated the utility of these different SO definitions in identifying current cardiometabolic risk. This previous study concluded that the 1.2×95th percentile of either curve was superior in identifying children with cardiometabolic risk than using their respective ≥99th percentile definitions, and there were limited discriminatory differences between the CDC and WHO curves 1.2×95th percentile definitions.[10] However, the prediction of future risk was not possible in this cross-sectional sample.[10]

Thus although growth curves are designed to track growth and health risks, the discriminatory power of the SO definitions to detect future diabetes and cardiovascular disease risk is largely unknown. Improving our understanding of the predictive utility of SO definitions is critical for clinical care and management, as well as proper interpretation of research studies as there is no consensus on SO definition. Therefore the objective of this research was to determine whether the WHO or CDC growth curves SO definitions differ from one another in predicting diabetes and cardiovascular disease risks 2 years later. A secondary objective assessed prediction 7 years later, and assessed whether SO definitions were more predictive than obesity definitions.

MATERIALS AND METHODS

Participants

Participants for this study were from the QUALITY cohort (Quebec Adipose and Lifestyle Investigation in Youth), an ongoing longitudinal investigation of the natural history of obesity and cardiovascular risk in Quebec youth. A detailed description of the study design and methods is available.[11] Briefly, youth with at least one biological parent with obesity were eligible to participate. Data were collected at baseline (2005–2008: n=630), follow-up 1 (2008–2011, n=564) and follow-up 2 (2015–2017, n=377). For this study, the primary analyses were focused on baseline and follow-up 1 measurements (n=548). Secondary analyses were restricted to those with complete follow-up 1 and 2 data (n=356). Study participants or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Measures

Cardiometabolic

All cardiometabolic measures were assessed at baseline, follow-up 1 and follow-up 2. Participants fasted (no food or drink 12 hours prior to the visit). Trained nurses collected venous blood samples according to standardised protocols. Samples were immediately stored on ice. All samples were centrifuged, aliquotted and stored at −80°C until analysis. Fasting cardiometabolic measures included blood glucose, insulin, high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol and triglycerides. Insulin levels were measured using an electrochemiluminescence immunoassay method (Synchron LXi 725, Beckman Coulter). Homoeostatic model assessment (HOMA-IR) was calculated as (fasting insulin (pmol/L)/7.175×fasting glucose (mmol/L))/22.5.[12] Systolic and diastolic blood pressures were measured after a 5 min rest and at least 30 min after a meal using an oscillometric instrument (Dinamap XL, model CR9340, Critikon Company, Florida, USA).[13] Five consecutive readings at 1 min intervals were obtained. The mean of the last three measures were used in the analyses. The biochemical analyses were performed at the Department of Clinical Biochemistry at Centre Hospitalier Universitaire Sainte-Justine in accordance to the standardised guidelines of the International Federation of Clinical Chemistry.[14 15]

The cardiometabolic measures were categorised into normal and unfavourable, based on guidelines or as recommended in the literature.[16 17] Unfavourable glucose homeostasis was defined as fasting blood glucose ≥6.1 mmol/L, and insulin resistance was defined as HOMA-IR index ≥2.0.[18 19] Unfavourable lipid levels were defined as HDL cholesterol <1.03 mmol/L, LDL cholesterol ≥2.6 mmol/L and triglycerides ≥1.24 mmol/L.[16] Unfavourable blood pressure was defined as ≥90th percentile for age-adjusted, sex-adjusted and height-adjusted systolic or diastolic blood pressure.[20]

Anthropometric

Height and weight were measured using a stadiometer (height) and electronic scale (weight), with light indoor clothing and no shoes. Measurements were taken two times; a third measurement was taken when differences between the two initial measures were 0.1 cm (height) or 0.1 kg (weight) or more. The analysis used the average of the two closest measurements. Body mass index (BMI) was calculated (kg/m^2) and compared with age-specific and sex-specific CDC and WHO reference curves to calculate BMI centile and BMI z-score.[21 22] SO with the CDC reference curves was defined as BMI ≥1.2×95th percentile.[5] SO with the WHO reference curves was defined as ≥99th percentile.[9] For ease of readability, the SO definitions of ≥99th percentile (WHO) and 1.2×95th percentile (CDC) will be referred to as WHO-SO, and CDC-SO, respectively throughout the rest of this manuscript.

Other

Direct observation by trained nurses assessed sexual maturation (Tanner stages).[23 24] Youth were classified as pre-pubertal (Tanner stage 1) or pubertal (Tanner stages 2–5).[23 24] Relevant questionnaire data from parental report included the highest education obtained (by either parent), and the previous year's annual household income.

Table 1 Demographic and health characteristics for Quebec Adipose and Lifestyle Investigation in Youth participants at baseline and first follow-up (n=548)

	Baseline			1st follow-up		
	Boys (n=302)	Girls (n=246)	P value	Boys (n=302)	Girls (n=246)	P value
Characteristic, mean (SD)*						
Age, years	9.6 (0.9)	9.6 (0.9)	0.39	11.7 (0.9)	11.6 (0.9)	0.29
Height, cm	139.2 (7.9)	138.4 (8.1)	<0.0001	151.8 (9.8)	151.2 (8.6)	<0.0001
Weight, kg	38.1 (11.2)	38.3 (11.6)	0.85	49.4 (15.4)	49.3 (14.2)	0.92
BMI	19.4 (4.3)	19.5 (4.2)	0.73	21.1 (4.9)	21.2 (4.9)	0.75
Puberty initiated, (%)†	26 (9%)	88 (36%)	<0.0001	165 (55%)	204 (83%)	<0.0001
1 or 2 parents with university degree	167 (55%)	132 (54%)	0.46	136 (49%)	118 (51%)	0.60
Annual household income						
<$20 000	25 (8%)	22 (9%)	0.94	19 (6%)	18 (8%)	0.92
$20 000–39 999	116 (38%)	92 (38%)		92 (30%)	71 (29%)	
$40 000–59 999	97 (32%)	84 (33%)		109 (36%)	92 (37%)	
≥$60 000	62 (21%)	47 (19%)		80 (28%)	62 (26%)	
Unfavourable cardiometabolic factors‡						
HDL	73 (24%)	70 (28%)	0.24	89 (30%)	70 (28%)	0.79
LDL	78 (26%)	85 (35%)	0.24	78 (26%)	53 (22%)	0.24
Triglyceride	32 (11%)	31 (13%)	0.45	27 (9%)	37 (15%)	0.03
Blood glucose	3 (1%)	1 (0.4%)	0.42	2 (0.7%)	2 (0.8%)	0.84
HOMA-IR index	16 (5%)	24 (10%)	0.04	57 (19%)	67 (27%)	0.02
Blood pressure						
Systolic	0 (0%)	2 (0.8%)	0.12	2 (0.7%)	4 (2%)	0.28
Diastolic	0 (0%)	0 (0%)	N/A	0 (0%)	0 (0%)	N/A
Clustered risk factors§						
No cardiometabolic abnormality	158 (52%)	97 (39%)	0.003	136 (45%)	104 (42%)	0.52
1 cardiometabolic abnormality	101 (33%)	96 (39%)	0.16	102 (34%)	83 (34%)	0.99
2 cardiometabolic abnormalities	28 (9%)	37 (15%)	0.04	41 (14%)	35 (14%)	0.83
≥3 cardiometabolic abnormalities	14 (5%)	14 (6%)	0.57	23 (8%)	24 (10%)	0.37

*Mean (SD) unless noted otherwise; percentages may not add up to 100% due to rounding.
†Puberty defined as Tanner stage ≥2.
‡Unfavourablee cardiometabolic defined as: low HDL: <1.03 mmol/L; high LDL: ≥2.6 mmol/L; high triglycerides: ≥1.24 mmol/L; high glucose: ≥6.11 mmol/L; high HOMA-IR index: ≥2.0; high blood pressure ≥90th percentile for age, height and sex, as defined by the National High Blood Pressure Education Program.
§One or more of the cardiometabolic risk factors as defined in the previous superscript.
BMI, body mass index; HDL, high density lipoprotein cholesterol; HOMA-IR, homoeostatic model assessment; LDL, low density lipoprotein cholesterol.

Statistical analysis

All analyses were performed separately for boys and girls. The primary objective was assessed with WHO-SO and CDC-SO definitions at baseline with cardiometabolic risk factors at follow-up 1. Statistical comparisons between boys and girls were conducted with χ^2 for categorical variables and t-test for continuous variables. Publicly available SAS macros were used to calculate the BMI-for-age z-scores and percentiles in accordance to the CDC and WHO growth curves. Observations with missing data were excluded from analyses. All descriptive and statistical analyses were performed with SAS V.9.4 (SAS Institute).

The areas under the receiver operator characteristic (ROC) curves (AUC) for each definition of SO (WHO-SO, CDC-SO) and cardiometabolic risk factor of interest were calculated. The AUC represents the probability that a SO definition will detect an unfavourable level of cardiometabolic risk. An AUC of 0.50 is considered uninformative and detects cardiometabolic risk no better than chance; an AUC greater than 0.80 is considered to be good.[25] Statistical comparison between AUCs[26–28] used a macro available online (http://www.medicine.mcgill.ca/epidemiology/hanley/software/delong_sas.html). Although ROC curves combine sensitivity (in our

Table 2 Cross-sectional detection of cardiometabolic risk factors of Quebec Adipose and Lifestyle Investigation in Youth participants among those with severe obesity at baseline or first follow-up

	Baseline		First follow-up	
	CDC-SO	WHO-SO	CDC-SO	WHO-SO
Boys, n (%)	n=21*	n=59*	n=25*	n=59*
HDL <1.03 mmol/L	14 (67%)	26 (44%)	17 (68%)	35 (59%)
LDL ≥2.6 mmol/L	11 (52%)	22 (37%)	13 (52%)	38 (64%)
Triglyceride ≥1.24 mmol/L	11 (52%)	19 (32)	8 (32%)	12 (20%)
Blood glucose ≥6.11 mmol/L	1 (5%)	1 (2%)	0 (0%)	1 (2%)
HOMA-IR cut-off ≥2.0	6 (29%)	12 (20%)	18 (72%)	34 (58%)
Blood pressure ≥90th percentile†				
Systolic	0 (0%)	0 (0%)	0 (0%)	1 (2%)
Diastolic	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Clustered risk factors‡§				
None	2 (10%)	17 (29%)	2 (8%)	2 (15%)
1	5 (24%)	19 (32%)	2 (8%)	12 (20%)
2	6 (29%)	11 (19%)	11 (44%)	23 (39%)
≥3	8 (38%)	12 (20%)	10 (40%)	15 (25%)
Girls, n (%)	n=16*	n=40*	n=19*	n=36*
HDL <1.03 mmol/L	9 (56%)	23 (58%)	11 (58%)	18 (50%)
LDL ≥2.6 mmol/L	5 (31%)	14 (35%)	12 (63%)	23 (64%)
Triglyceride ≥1.24 mmol/L	8 (50%)	11 (28%)	8 (42%)	13 (36%)
Blood glucose ≥6.11 mmol/L	0 (0%)	0 (0%)	1 (5%)	1 (3%)
HOMA-IR cut-off ≥2.0	9 (56%)	16 (40%)	16 (84%)	26 (72%)
Blood pressure ≥90th percentile†				
Systolic	1 (6%)	1 (3%)	2 (11%)	3 (8%)
Diastolic	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Clustered risk‡§				
None	0 (0%)	4 (10%)	1 (5%)	5 (14%)
1	4 (25%)	15 (38%)	4 (21%)	8 (22%)
2	6 (38%)	11 (28%)	6 (32%)	10 (28%)
≥3	5 (31%)	9 (23%)	8 (42%)	13 (36%)

*Ns are for the boys and girls meeting severe obesity definitions at baseline or first follow-up from the full sample of 548 (302 boys, 246 girls).
†≥90th percentile for age, height and sex, as defined by the National High Blood Pressure Education Program.
‡Unfavourablee cardiometabolic defined as: low HDL: <1.03 mmol/L; high LDL: ≥2.6 mmol/L; high triglycerides: ≥1.24 mmol/L; high glucose: ≥6.11 mmol/L; high HOMA-IR index: ≥2.0; high blood pressure ≥90th percentile sex-matched and age-matched.
§One or more of the cardiometabolic risk factors as defined in the previous superscript.
CDC, Centers for Disease Control and Prevention; HDL, high density lipoprotein cholesterol; HOMA-IR, homoeostatic model assessment; LDL, low density lipoprotein cholesterol.

context, detecting cardiometabolic risk among those with SO) and specificity (no cardiometabolic risk among those without SO) into a single measure, these values were also separately calculated.

McFadden pseudo R^2 assessed the goodness of fit (0–0.2: poor, 0.2–0.4 good, 0.4+ excellent model fit).[29 30] Better model fit was assessed with the Akaike Information Criterion (AIC), with AIC at least two units lower deemed as statistically superior model fit. Kappa coefficients determined agreement between the curves (kappa of 81%–100%: high agreement, 61%–80%: substantial agreement, 41%–60%: moderate agreement, 21%–40%: fair, and 0%–20%: slight agreement).[31 32] Additionally, we assessed the ability of these definitions to detect any single, or multiple (two, and at least three) unfavourable risk factor clusters. Because there were too few cases of unfavourable blood glucose, systolic and diastolic blood pressure in the cohort, these risk factors were excluded from the risk factor clusters.

A sensitivity analysis calculated AUCs, sensitivity and specificity for each SO definition at follow-up 1 with cardiometabolic risk factors at follow-up 2. The AUCs and

Table 3 Sensitivity, AUC and kappa of CDC-defined and WHO-defined categories of severe obesity at baseline for predicting cardiometabolic risk at first follow-up for Quebec Adipose and Lifestyle Investigation in Youth participants (n=548), by sex

	Sensitivity		Specificity		AUC†		Pseudo R^2‡		
	WHO-SO¶	CDC-SO¶	WHO-SO¶	CDC-SO¶	WHO-SO¶	CDC-SO¶	WHO-SO¶	CDC-SO¶	Kappa§
Risk factor									
HDL									
Boys	39.3	15.7	88.7	96.7	0.64*	0.56	0.08††	0.04	44
Girls	30	11.4	89.2	95.4	0.59*	0.53	0.04††	0.01	51
LDL									
Boys	24.3	10.3	82.1	94.2	0.53	0.52	0.004	0.005	47
Girls	26.4	15.1	86.5	95.8	0.56	0.55	0.02	0.03††	53
Triglyceride									
Boys	48.1	25.9	80.7	93.3	0.66	0.6	0.07	0.06	45
Girls	37.8	18.9	87.6	95.7	0.63	0.57	0.06††	0.04	51
HOMA-IR									
Boys	61.4	26.3	90.2	97.5	0.76*	0.62	0.22††	0.1	37
Girls	41.8	19.4	93.3	98.3	0.67*	0.59	0.14††	0.08	47
Clustered risk‡‡									
0									
Boys	6.6	2.2	69.9	89.2	0.62*	0.54	0.07††	0.02	45
Girls	5.8	1	76.1	89.4	0.59*	0.55	0.05††	0.03	51
1									
Boys	11.8	0	76.5	89.5	0.55	0.55	0.02	0.05††	55
Girls	10.8	3.6	81	92	0.54	0.52	0.009	0.006	52
2									
Boys	53.7	24.4	85.8	95.8	0.70*	0.6	0.12††	0.07	43
Girls	28.6	14.3	85.8	94.8	0.57	0.54	0.02	0.02	52
≥3									
Boys	69.6	34.8	84.6	95.3	0.77*	0.65	0.19††	0.11	41
Girls	62.5	29.2	88.7	96	0.76*	0.63	0.19††	0.09	46

†Statistical comparison performed with an available macro online to calculate whether WHO significantly differed from CDC AUC: *p<0.05.
‡Statistical comparison performed with AIC to determine whether McFadden psuedo R^2 significantly differed between WHO and CDC.
§Kappa statistics calculated between WHO and CDC definitions as described in superscripts ¶,**.
¶Severe obesity is defined as ≥99th percentile for WHO (WHO-SO), and as ≥1.2×95th percentile for CDC (CDC-SO).
**Risk factor defined as: HDL <1.03 mmol/L; LDL ≥2.6 mmol/L; triglyceride ≥1.24 mmol/L; blood glucose ≥6.11 mmol/L; HOMA-IR cut-off ≥2.0 or blood pressure (systolic or diastolic) ≥90th percentile for age, height and sex, as defined by the National High Blood Pressure Education Program.
††AIC significantly lower.
‡‡Number of cardiometabolic risk factors (HDL, LDL, triglyceride, HOMA-IR) as defined in superscript **.
AUC, area under the receiver operating characteristic curve; CDC, Centers for Disease Control and Prevention; HDL, high density lipoprotein cholesterol; HOMA-IR, homoeostatic model assessment; LDL, low density lipoprotein cholesterol.

kappa were also calculated for obesity (CDC: BMI≥95th percentile, WHO: BMI≥97.7th percentile) as a point of reference. Comparisons between SO and obesity definitions, as well as between WHO-obesity and CDC-obesity definitions were conducted.

Patient and public involvement

Study participants or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

RESULTS

There were no significant differences in age, sex or BMI between those in the analytic sample (n=548) and those excluded due to missing data (n=82, data not shown). In the analytic sample of 548 participants, 55% were boys, mean age at baseline was 9.6 years (SD: 0.9) and 11.6 years (SD: 0.9) at follow-up 1 (table 1). Girls and boys significantly differed in pubertal stage at baseline and follow-up 1. Descriptive characteristics of cardiometabolic risk

Table 4 Comparison of AUC of CDC-defined and WHO-defined categories of obesity and severe obesity at baseline for predicting cardiometabolic risk at first follow-up for Quebec Adipose and Lifestyle Investigation in Youth participants (n=548), by sex

	WHO Obesity†			CDC Obesity†			WHO obesity vs CDC obesity	
Risk factor*	AUC	Pseudo R^2‡	vs WHO SO§ P value¶	AUC	Pseudo R^2‡	vs CDC SO§ P value¶	P value**	Kappa§§
HDL								
Boys	0.67	0.10††	0.05	0.64	0.08	0.001	0.05	86
Girls	0.59	0.03	0.83	0.6	0.04††	0.02	0.11	94
LDL								
Boys	0.57	0.02††	0.07	0.55	0.009	0.32	0.19	88
Girls	0.6	0.03	0.18	0.61	0.04	0.11	0.08	95
Triglyceride								
Boys	0.64	0.05	0.48	0.67	0.08††	0.15	<0.01	87
Girls	0.71	0.12	0.02	0.72	0.14††	<0.01	0.08	95
HOMA-IR								
Boys	0.77	0.22	0.45	0.76	0.21	<0.001	0.44	83
Girls	0.71	0.16††	0.08	0.7	0.14	<0.001	0.28	96
Clustered risk‡‡								
None								
Boys	0.65	0.10††	0.004	0.63	0.07	<0.001	<0.01	86
Girls	0.63	0.08	0.01	0.63	0.08	<0.001	0.78	97
1								
Boys	0.53	0.004	0.11	0.55	0.01††	0.89	0.23	89
Girls	0.53	0.003	0.43	0.53	0.005	0.59	0.32	98
2								
Boys	0.72	0.13††	0.42	0.69	0.1	0.02	0.23	86
Girls	0.59	0.03	0.47	0.6	0.03	0.16	0.08	96
≥3								
Boys	0.78	0.19	0.68	0.78	0.20††	0.01	0.97	85
Girls	0.81	0.24	0.21	0.81	0.25††	<0.001	0.76	95

*Risk factor defined as: HDL <1.03 mmol/L; LDL ≥2.6 mmol/L; triglyceride ≥1.24 mmol/L; blood glucose ≥6.11 mmol/L; HOMA-IR cut-off ≥2.0 or blood pressure (systolic or diastolic) ≥90th percentile for age, height and sex, as defined by the National High Blood Pressure Education Program.
†Obesity defined as ≥97.7th percentile according to WHO and as ≥95th percentile according to CDC.
‡Statistical comparison performed with AIC to determine whether McFadden psuedo R^2 significantly differed between WHO and CDC.
§Severe obesity is defined as ≥99th percentile for WHO (WHO-SO), and as ≥1.2×95th percentile for CDC (CDC-SO).
¶Statistical comparison performed with an available macro online to calculate whether obesity AUC significantly differed from severe obesity AUC.
**Statistical comparison performed with an available macro online to calculate whether WHO obesity AUC significantly differed from CDC obesity AUC.
††AIC significantly lower (better model fit) with this obesity definition compared with the other obesity definition.
‡‡One or more of the cardiometabolic risk factors (HDL, LDL, triglyceride, HOMA-IR) as defined in the superscript *.
§§ Kappa statistics calculated between WHO and CDC definitions of obesity
AUC, area under the receiver operating characteristic; CDC, Centers for Disease Control and Prevention; HDL, high density lipoprotein cholesterol; HOMA-IR, homoeostatic model assessment; LDL, low density lipoprotein cholesterol.

factors at baseline and follow-up among those meeting a SO definition are presented (table 2). At baseline, 99 and 37 youth were identified with SO using the WHO-SO or the CDC-SO criteria, respectively. Compared with baseline, more cases of SO were detected with the CDC-SO definition at follow-up (n=37 at baseline and n=44 at follow-up), but less were detected with the WHO-SO definition (n=99 at baseline and n=95 at follow-up). Although the overall prevalence of at least one cardiometabolic abnormality at follow-up 1 was 44% in the full sample, among those with SO based on the WHO-SO, or CDC-SO, the prevalence was 85% and 92%, respectively.

The diagnostic performance of SO definitions in predicting cardiometabolic risk at the first follow-up, as well as their agreement with one another is provided (table 3). Sensitivity of CDC-SO at baseline was lower than the WHO classification for all risk factors at the first follow-up (table 3). Based on kappas, agreement between the WHO-SO and CDC-SO was moderate (37%–55%). The AUCs from the SO definitions were largely uninformative for detecting cardiometabolic abnormalities. Similarly, good model fit was detected only for HOMA-IR and having more than three risk factors based on McFadden pseudo R^2s greater than 0.20; all others had poor model

fit. Results comparing follow-up 1 as a prediction of follow-up 2 were consistent with those from baseline and follow-up 1, although nearly all comparisons were not statistically significantly different (data not shown).

In contrast, the CDC and WHO obesity definitions had good agreement, with kappas all at least 83% (table 4). Notably, while the CDC-SO definitions demonstrated a significantly poorer AUC in comparison to the CDC obesity definition, the WHO-SO definition had approximately the same AUCs as the WHO obesity definition.

DISCUSSION

In this large cohort of youth, CDC-SO AUCs were generally inferior in comparison to the WHO-SO definition for detecting cardiometabolic abnormalities. However, as all SO definitions had relatively uninformative AUCs, it is unlikely that the statistical superiority of the WHO-SO and the CDC-SO definitions found in this study are of clinical importance.

To the best of our knowledge, only one previous cross-sectional study assessed the utility of SO definitions.[10] Notably, in a sample (n=3340, mean age: 11.2) comprised exclusively of children with overweight/obesity. SO AUCs[10] were similarly uninformative, emphasising the difference between clinical and statistical significance. Despite our study's relatively low prevalence of cardiometabolic markers, results were largely consistent with that of Valerio *et al*.[10] Indeed the prevalence of obesity in our study is likely more consistent with most populations than the SO prevalence of >50% in Valerio *et al* study.[10]

Although accurately classifying youth with SO has been recommended,[33] the necessity of doing so for the purposes of identifying cardiometabolic risks during childhood and adolescence is unclear. In fact, the AUCs from SO definitions in this study were similar to those from obesity definitions. Thus there is limited evidence that there is a discriminatory advantage of SO definitions over obesity definitions for identifying cardiometabolic risk in youth.

This study is not without limitations. Participants were not a representative sample and results may not be generalised to all youth. Although unmeasured confounding is possible, as this is a study comparing methodological definitions each participant served as their own comparator. Thus the growth curves are likely to perform similarly in another sample population in which prevalence of SO is approximately 5%–20%. Given that this study's sample size was relatively small, the prevalence of individual and clusters of cardiometabolic risk factors (1, 2, 3+) were assessed. Interpreting clusters is less straight-forward than the individual risk factors. However, as there were very few with high blood pressure or glucose, the majority of the risk factor clustering occurred with the lipids (HDL, LDL, triglycerides) and HOMA-IR.

Although this is the first cohort study to assess prediction of SO definitions of future cardiometabolic risk, there were only two follow-up visits (approximately 2 years and 5 years after baseline, respectively). Due to study attrition, identification of cardiometabolic risk at the second follow-up was likely underpowered. A repeated measures analysis may have more efficiently retained statistical power. However, as we were interested in determining whether the prediction models strengthened or worsened at the specific study visits, we analysed the data with separate models at the two follow-up visits. As an increase in HOMA-IR index is expected as youth enter puberty, the study should be reassessed in a sample in which adolescents have completed puberty.

ROCs may be less informative when datasets are imbalanced between diseased and non-diseased,[34] thus the McFadden pseudo R^2 were also calculated. Nevertheless, McFadden pseudo R^2 model fit and ROCs statistical comparisons indicated that WHO-SO performed better than CDC-SO. As this study was focused on the methodological and statistical differences between SO definitions, models were unadjusted.

Finally, the 99th percentile is a recommended rounded percentile of the 99.9th for defining SO in both research and clinical care. Using the exact percentile of 99.9th would decrease the number of SO youth in this cohort by two-thirds (from 95 to 28 youth) and would result in incalculable statistical comparisons due to insufficient sample sizes. However, as the AUCs for SO were largely uninformative, an increase in statistical precision is unlikely to be clinically useful.

CONCLUSION

In this large cohort sample of youth, the WHO definition of SO was consistently superior in detecting diabetes and cardiovascular disease risks at follow-up in comparison to the CDC definition of SO. From a clinical standpoint, the calculation of the WHO-SO is also simpler than the CDC-SO. Nevertheless, as the AUCs for SO using either CDC or WHO definition were generally uninformative, the improvement in sensitivity may not be clinically meaningful and should be used with caution.

Author affiliations
[1]PERFORM Centre, Concordia University, Montreal, Québec, Canada
[2]Department of Mathematics and Statistics, Concordia University, Montreal, Québec, Canada
[3]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Québec, Canada
[4]Département de kinésiologie, Université Laval, Quebec City, Quebec, Canada
[5]Department of Pediatrics, Université de Montréal, Montreal, Quebec, Canada
[6]Research Center of CHU Sainte Justine, Université de Montréal, Montreal, Quebec, Canada

Acknowledgements Dr Marie Lambert (July 1952–February 2012), paediatric geneticist and researcher, initiated the QUALITY cohort. Her leadership and devotion to QUALITY will always be remembered and appreciated. The cohort integrates members of TEAM PRODIGY, an inter-university research team including Université de Montréal, Concordia University, INRS-Institute-Armand Frappier, Université Laval, and McGill University. The research team is grateful to all the youth and their families who took part in this study, as well as the technicians, research assistants and coordinators involved in the QUALITY cohort project. Portions of

these data were presented in a poster at the Canadian Society for Epidemiology and Biostatistics 2019 National Conference, Ottawa, Ontario, 13–15 May 2019.

Contributors MH, GP and AT conceptualised and designed the study, coordinated and supervised data collection, and critically reviewed the manuscript for important intellectual content. LK conceptualised and designed the study, drafted the initial manuscript, and reviewed and revised the manuscript. AS conceptualised and designed the study, carried out the initial analyses, and reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work. LK is the guarantor.

Funding The QUALITY cohort is funded by the Canadian Institutes of Health Research (#OHF-69442, #NMD-94067, #MOP-97853, #MOP-119512), the Heart and Stroke Foundation of Canada (#PG-040291) and the Fonds de la Recherche du Québec - Santé. Mélanie Henderson holds a Diabetes Junior Investigator Award from the Canadian Society of Endocrinology and Metabolism - AstraZeneca and a Fonds de Recherche du Québec - Santé Junior 2 salary awards and Lisa Kakinami holds a Junior 1 salary award from the latter institution.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval The study was conducted ethically in accordance with the Tri Council policy statement on the ethical conduct for research involving humans. Ethics review boards of Centre Hospitalier Universitaire Sainte-Justine and the Quebec Heart and Lung Institute approved the study protocol (#MP-21-2005-79, 2040). The secondary data analysis for this project was approved by the Concordia University ethics board (#30016473). All parents provided informed consent, and study participants provided assent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. All data that support the findings of this study are available from the QUALITY research team upon reasonable request (www.etudequalitystudy.ca).

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Lisa Kakinami http://orcid.org/0000-0002-8496-7643
Melanie Henderson http://orcid.org/0000-0002-0102-2389

REFERENCES

1 Kelly AS, Barlow SE, Rao G, *et al*. Severe obesity in children and adolescents: identification, associated health risks, and treatment approaches: a scientific statement from the American Heart Association. *Circulation* 2013;128:1689–712.
2 Daniels SR, Kelly AS. Pediatric severe obesity: time to establish serious treatments for a serious disease. *Child Obes* 2014;10:283–4.
3 Li L, Pérez A, Wu L-T, *et al*. Cardiometabolic risk factors among severely obese children and adolescents in the United States, 1999-2012. *Child Obes* 2016;12:12–19.
4 Barlow SE, Expert Committee. Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: summary report. *Pediatrics* 2007;120:S164–92.
5 Flegal KM, Wei R, Ogden CL, *et al*. Characterizing extreme values of body mass index-for-age by using the 2000 centers for disease control and prevention growth charts. *Am J Clin Nutr* 2009;90:1314–20.
6 Freedman DS, Mei Z, Srinivasan SR, *et al*. Cardiovascular risk factors and excess adiposity among overweight children and adolescents: the Bogalusa heart study. *J Pediatr* 2007;150:12–17.
7 Koebnick C, Smith N, Coleman KJ, *et al*. Prevalence of extreme obesity in a multiethnic cohort of children and adolescents. *J Pediatr* 2010;157:26–31.
8 Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stat Med* 1992;11:1305–19.
9 Secker D. Promoting optimal monitoring of child growth in Canada: Using the new World Health Organization growth charts - Executive Summary. *Paediatr Child Health* 2010;15:77–9.
10 Valerio G, Maffeis C, Balsamo A, *et al*. Severe obesity and cardiometabolic risk in children: comparison from two international classification systems. *PLoS One* 2013;8:e83793.
11 Lambert M, Van Hulst A, O'Loughlin J, *et al*. Cohort profile: the Quebec adipose and lifestyle investigation in youth cohort. *Int J Epidemiol* 2012;41:1533–44.
12 Peplies J, Jiménez-Pavón D, Savva SC, *et al*. Percentiles of fasting serum insulin, glucose, HbA1c and HOMA-IR in pre-pubertal normal weight European children from the IDEFICS cohort. *Int J Obes* 2014;38:S39–47.
13 Paradis G, Lambert M, O'Loughlin J, *et al*. Blood pressure and adiposity in children and adolescents. *Circulation* 2004;110:1832–8.
14 Tate JR, Rifai N, Berg K, *et al*. International Federation of Clinical Chemistry standardization project for the measurement of lipoprotein(a). Phase I. Evaluation of the analytical performance of lipoprotein(a) assay systems and commercial calibrators. *Clin Chem* 1998;44:1629–40.
15 Tate JR, Berg K, Couderc R, *et al*. International federation of clinical chemistry and laboratory medicine (IFCC) standardization project for the measurement of lipoprotein(a). Phase 2: selection and properties of a proposed secondary reference material for lipoprotein(a). *Clin Chem Lab Med* 1999;37:949–58.
16 Cook S, Weitzman M, Auinger P, *et al*. Prevalence of a metabolic syndrome phenotype in adolescents: findings from the third National Health and Nutrition Examination Survey, 1988-1994. *Arch Pediatr Adolesc Med* 2003;157:821–7.
17 Expert Panel on Integrated Guidelines for Cardiovascular Health and Risk Reduction in Children and Adolescents, National Heart, Lung, and Blood Institute. Expert panel on integrated guidelines for cardiovascular health and risk reduction in children and adolescents: summary report. *Pediatrics* 2011;128:S213–56.
18 Shashaj B, Luciano R, Contoli B, *et al*. Reference ranges of HOMA-IR in normal-weight and obese young Caucasians. *Acta Diabetol* 2016;53:251–60.
19 Andrade MISde, Oliveira JS, Leal VS, *et al*. [Identification of cutoff points for Homeostatic Model Assessment for Insulin Resistance index in adolescents: systematic review]. *Rev Paul Pediatr* 2016;34:234–42.
20 National High Blood Pressure Education Program Working Group on High Blood Pressure in Children and Adolescents. The fourth report on the diagnosis, evaluation, and treatment of high blood pressure in children and adolescents. *Pediatrics* 2004;114:555–76.
21 Kuczmarski RJ, Ogden CL, Guo SS, *et al*. 2000 CDC growth charts for the United States: methods and development. *Vital Health Stat* 2002;246:1–190.
22 de Onis M, Onyango AW, Borghi E, *et al*. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 2007;85:660–7.
23 Marshall WA, Tanner JM. Variations in pattern of pubertal changes in girls. *Arch Dis Child* 1969;44:291–303.
24 Marshall WA, Tanner JM. Variations in the pattern of pubertal changes in boys. *Arch Dis Child* 1970;45:13–23.
25 Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978;8:283–98.
26 Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983;148:839–43.
27 Hanley JA, Hajian-Tilaki KO. Sampling variability of nonparametric estimates of the areas under receiver operating characteristic curves: an update. *Acad Radiol* 1997;4:49–58.
28 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36.
29 McFadden D. Conditional logit analysis of qualitative choice behaviour. In: *Economic theory and mathematical economics frontiers in econometrics*. New York, NY: Academic Press, 1974.
30 McFadden D. Quantitative methods for analysing travel behaviour of individuals: some recent developments. In: *Behavioural travel modelling*. London, UK: Croom Helm, 1979.
31 Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
32 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
33 Gulati AK, Kaplan DW, Daniels SR. Clinical tracking of severely obese children: a new growth chart. *Pediatrics* 2012;130:1136–40.
34 Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* 2015;10:e0118432.