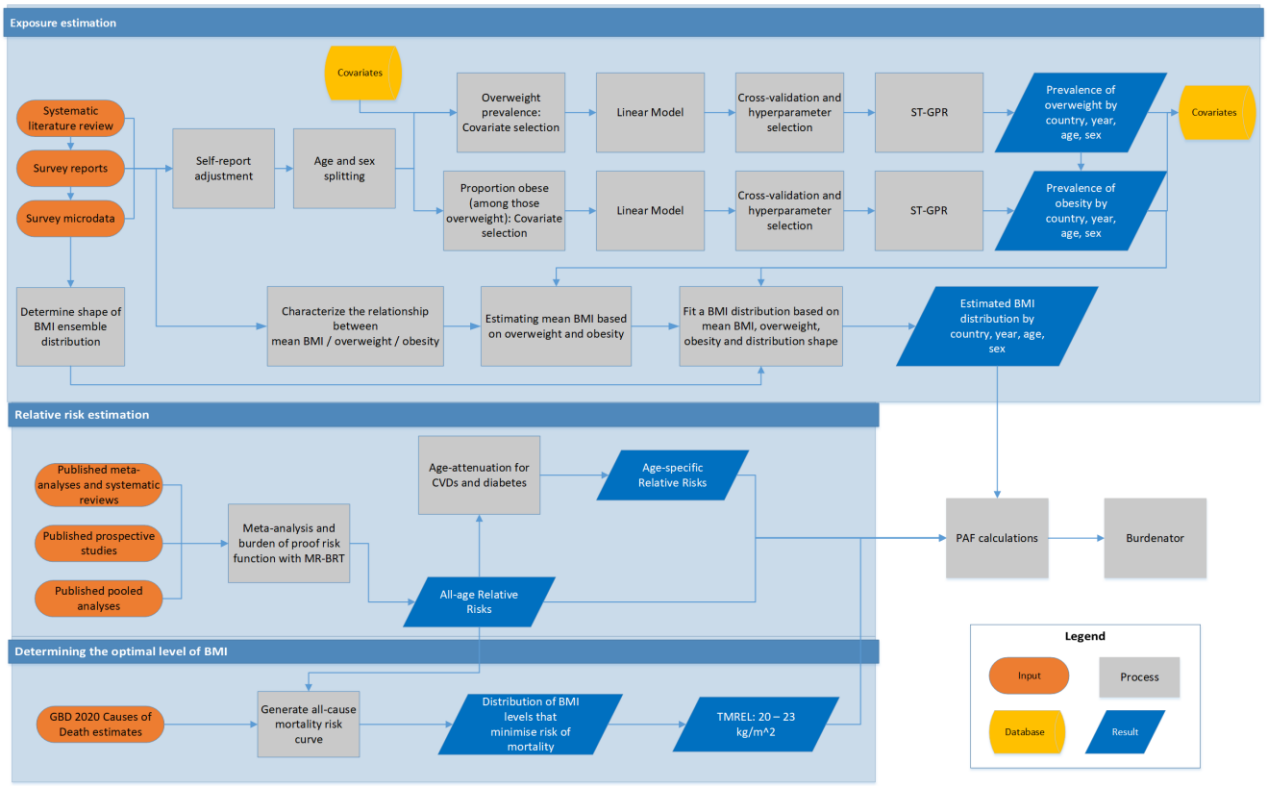
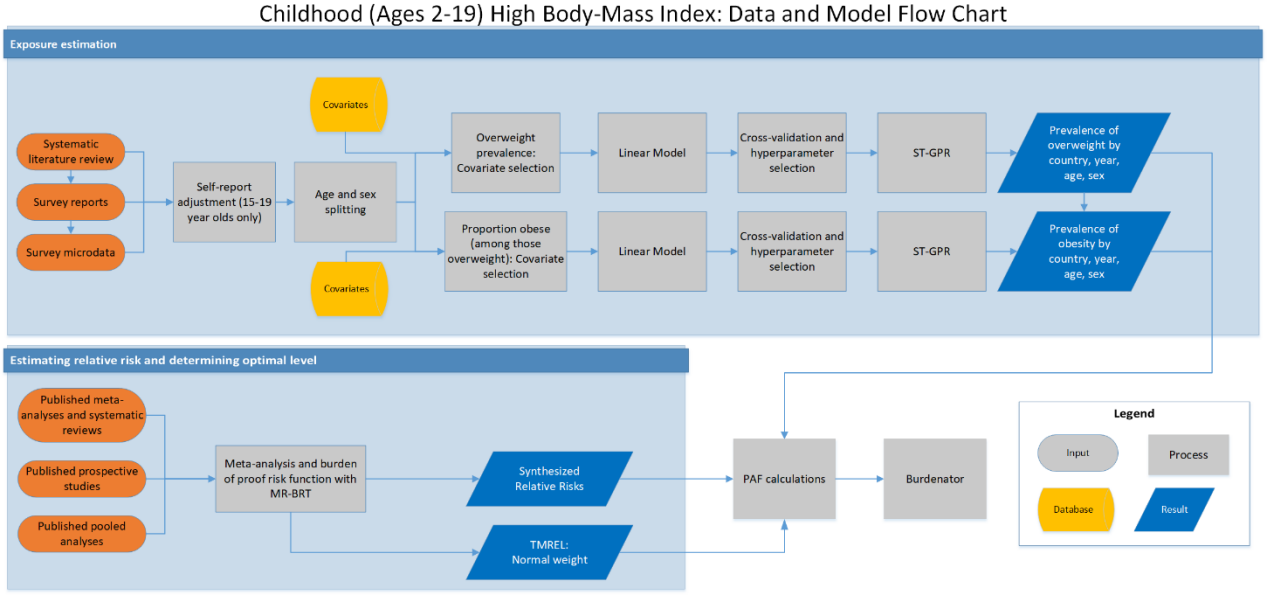
High body-mass index

Flowchart

Adult (Ages 20+) High Body-Mass Index: Data and Model Flow Chart



Childhood (Ages 2-19) High Body-Mass Index: Data and Model Flow Chart



Input data and methodological summary

Case definitions

Exposure

High body-mass index (BMI) for adults (ages 20+) is defined as BMI greater than 20 to 23 kg/m^2. High BMI for children and adolescents (ages 2–19) is defined as being overweight or obese based on International Obesity Task Force standards.[1]

Input data

Exposure

In GBD 2021, new data were added from sources included in the annual GHDx update of known survey series. We conducted a systematic review in GBD 2017 to identify studies providing nationally or subnationally representative estimates of overweight prevalence, obesity prevalence, or mean body-mass index (BMI). We limited the search to literature published between January 1, 2016, and December 31, 2016, to update the systematic literature search previously performed as part of GBD 2015.

The search for adults was conducted on 4 January 2017, using the following terms:

((("Body Mass Index"[Mesh] OR "Overweight"[Mesh] OR "Obesity"[Mesh]) AND ("Geographic Locations"[Mesh] NOT "United States"[Mesh]) AND ("humans"[Mesh] AND "adult"[MeSH]) AND ("Data Collection"[Mesh] OR "Health Services Research"[Mesh] OR "Population Surveillance"[Mesh] OR "Vital statistics"[Mesh] OR "Population"[Mesh] OR "Epidemiology"[Mesh] OR "surve*"[TiAb]) NOT (Comment[ptyp] OR Case Reports[ptyp] OR "hospital"[TiAb])) AND ("2016/01/01"[Date - Publication] : "2016/12/31"[Date - Publication]))

The search for children was conducted on 4 August 2016, using the following terms:
((("Body Mass Index"[Mesh] OR "Overweight"[Mesh] OR "Obesity"[Mesh]) AND ("Geographic Locations"[Mesh] NOT "United States"[Mesh]) AND ("humans"[Mesh] AND "child"[MeSH]) AND ("Data Collection"[Mesh] OR "Health Services Research"[Mesh] OR "Population Surveillance"[Mesh] OR "Vital statistics"[Mesh] OR "Population"[Mesh] OR "Epidemiology"[Mesh] OR "surve*"[TiAb]) NOT (Comment[ptyp] OR Case Reports[ptyp] OR "hospital"[TiAb])) AND ("2016/01/01"[Date - Publication] : "2016/12/31"[Date - Publication]))

Table 1: Data inputs for exposure for high body-mass index.

Input data	Exposure
Source count (total)	2016
Number of countries with data	194

Eligibility criteria

We included representative studies providing data on mean BMI or prevalence of overweight or obesity among adults or children. For adults, studies were included if they defined overweight as BMI ≥25 kg/m^2 and obesity as BMI ≥30 kg/m^2, or if estimates using those cutoffs could be back-calculated from reported categories. For children (children ages 2–19), studies were included if they used International Obesity Task Force (IOTF) standards to define overweight and obesity thresholds.[1] We only included studies reporting data collected after January 1, 1980. Studies were excluded if they used non-random samples (eg, case-control studies or convenience samples), conducted among specific subpopulations (eg, pregnant women,

racial or ethnic minorities, immigrants, or individuals with specific diseases), used alternative methods to assess adiposity (eg, waist circumference, skin-fold thickness, or hydrodensitometry), had sample sizes of less than 20 per age-sex group, or provided inadequate information on any of the inclusion criteria. We also excluded review articles and non-English-language articles.

Data collection process

Where individual-level survey data were available, we computed mean BMI using weight and height. We then used BMI to determine the prevalence of overweight and obesity. For individuals aged over 19 years, we considered them to be overweight if their BMI was greater than or equal to 25 kg/m^2, and obese if their BMI was greater than or equal to 30 kg/m^2. For individuals aged 2 to 19 years, we used monthly IOTF cutoffs[2] to determine overweight and obese status when age in months was available. When only age in years was available, we used the cutoff for the midpoint of that year. Obese individuals were also considered to be overweight. We excluded studies using the World Health Organization (WHO) standards or country-specific cutoffs to define childhood overweight and obesity. At the individual level, we considered BMI <10 kg/m^2 and BMI >70 kg/m^2 to be biologically implausible and excluded those observations.

The rationale for choosing to use the IOTF cutoffs over the WHO standards has been described elsewhere.[1] Briefly, the IOTF cutoffs provide consistent child-specific standards for ages 2–18 derived from surveys covering multiple countries. By contrast, the WHO growth standards apply to children under age 5, and the WHO growth reference applies to children ages 5–19. The WHO growth reference for children ages 5–19 was derived from United States data, which are less representative than the multinational data used by IOTF. Additionally, the switch between references at age 5 can produce artificial discontinuities. Given that we estimate global childhood overweight and obesity for ages 2–19 (with age 19 using standard adult cutoffs), the IOTF cutoffs were preferable. Additionally, we found that IOTF cutoffs were more commonly used in scientific literature covering childhood obesity.

From report and literature data, we extracted data on mean BMI, prevalence of overweight, and prevalence of obesity, measures of uncertainty for each, and sample size, by the most granular age and sex groups available. Additionally, we extracted the same study-level covariates as were extracted from microdata (measurement, urbanicity, and representativeness), as well as location and year.

In addition to the primary indicators described above, we extracted relevant survey-design variables, including primary sampling unit, strata, and survey weights, which were used to tabulate individual-level microdata and produce accurate measures of uncertainty. We extracted three study-level covariates: 1) whether height and weight data were measured or self-reported; 2) whether the study was predominantly conducted in an urban area, rural area, or both; and 3) the level of representativeness of the study (national or subnational).

Finally, we extracted relevant demographic indicators, including location, year, age, and sex. We estimated the standard error of the mean from individual-level data, where available, and used the reported standard error of the mean for published data. When multiple data sources were available for the same country, we included all of them in our analysis. If data from the same data source were available in multiple formats such as individual-level data and tabulated data, we used individual-level data.

Relative risk

In GBD 2021, we did not conduct an updated systematic review to identify new relative risk data sources. The last date of search in PubMed for evidence studying the health effects of high BMI on cardiovascular diseases and diabetes was 6 June 2019 using the following terms: ("Diabetes Mellitus"[Mesh] OR "diabetes"[title] OR "Stroke"[Mesh] OR "stroke[title]" OR "Heart Diseases"[Mesh] OR "Heart Diseases"[title] OR "Cardiovascular Diseases"[Mesh] OR "Cardiovascular Diseases"[title]) AND ("Obesity"[Mesh] OR "Obesity"[title] OR "Overweight"[Mesh] OR "Overweight"[title] OR "Body Mass Index"[Mesh] OR "Body Mass Index"[title]) AND ("cohort"[tiab]). For other risk-outcome pairs, we used existing meta-analyses and systematic reviews to identify and extract pooled cohorts and prospective studies for analysis.

Table 2: Data inputs for relative risks for high body-mass index.

Input data	Relative risk
Source count (total)	313
Number of countries with data	26

Data processing

Age and sex splitting

Any report or literature data provided in age groups wider than the standard five-year age groups or as both sexes combined were split using the approach used by Ng and colleagues.[2] We first modelled age-sex patterns with spatiotemporal Gaussian process regression (ST-GPR) using data sources reporting in sex-specific, standard five-year age units. To account for the large heterogeneity in overweight and obesity prevalence across geographical regions, we categorised each location into three categories of overweight and obesity prevalence. We then aggregated the modelled age and sex patterns into tertiles of overweight and obesity prevalence. Finally, the aggregated patterns were applied to split report and literature data based on the data source's location and its respective tertile of overweight or obesity prevalence. We did not propagate the uncertainty in the age pattern and sex pattern used to split the data as they seemed to have small effect.

Self-report bias adjustment

We included both measured and self-reported data. We tested for bias in self-report data compared to measured data, which is considered to be the gold standard. There was no clear direction of bias for children ages 2–14, so for these age groups we only included measured data. For individuals ages 15 and older, we adjusted self-reported data for overweight prevalence and obesity prevalence. We used MR-BRT to determine the level of self-report bias adjustment. For both overweight and obesity, we fit sex-specific MR-BRT models on the logit difference between measured and self-reported with a fixed effect on super-region. The bias coefficients derived from these two models are in Table 1 and 2.

A separate self-report bias adjustment was completed for the USA. Self-report data was compared to measured data from the NHANES survey series, which were selected as the gold standard for the USA. For individuals ages 2 and older, we adjusted self-reported data for overweight prevalence and obesity prevalence. We used MR-BRT to determine the level of self-report bias adjustment. For both overweight and obesity, we fit sex-specific MR-BRT models on the logit difference between NHANES measured and

self-reported with a fixed effect on 5-year age groups and decade when the data was collected. The bias coefficients derived from these two models are in Table 3.

Table 1: MR-BRT self-report crosswalk adjustment factors for overweight prevalence

Model	Data input	Reference or alternative case definition	Gamma	Beta coefficient, logit (95% CI)
Females	Measured data	Ref	0.26	---
	Self-reported data (southeast Asia, east Asia, and Oceania)	Alt		–0.53 (–1.03, –0.04)
	Self-reported data (central Europe, eastern Europe, and central Asia)	Alt		–0.20 (–0.69, 0.30)
	Self-reported data (high-income)	Alt		–0.25 (–0.75, 0.24)
	Self-reported data (Latin America and Caribbean)	Alt		–0.19 (–0.69, 0.31)
	Self-report data (north Africa and Middle East)	Alt		–0.38 (–0.89, 0.11)
	Self-report data (south Asia)	Alt		0.36 (–0.14, 0.85)
	Self-report data (sub-Saharan Africa)	Alt		–0.26 (–0.76, 0.24)
Males	Measured data	Ref	0.43	---
	Self-reported data (southeast Asia, east Asia, and Oceania)	Alt		–0.36 (–1.17, 0.50)
	Self-reported data (central Europe, eastern Europe, and central Asia)	Alt		–0.03 (–0.84, 0.82)
	Self-reported data (high-income)	Alt		0.05 (–0.77, 0.87)
	Self-reported data (Latin America and Caribbean)	Alt		–0.02 (–0.84, 0.81)
	Self-report data (north Africa and Middle East)	Alt		–0.21 (–1.04, 0.61)
	Self-report data (south Asia)	Alt		0.53 (–0.28, 1.37)
	Self-report data (sub-Saharan Africa)	Alt		–0.27 (–1.09, 0.55)

Table 2: MR-BRT self-report crosswalk adjustment factors for obesity prevalence

Model	Data input	Reference or alternative case definition	Gamma	Beta coefficient, logit (95% UI) *
Females	Measured data	Ref	0.38	---
	Self-reported data (southeast Asia, east Asia, and Oceania)	Alt		–0.11 (–0.86, 0.64)
	Self-reported data (central Europe, eastern Europe, and central Asia)	Alt		–0.95 (–1.70, –0.19)

	Self-reported data (high-income)	Alt		–0.42 (–1.16, 0.34)
	Self-reported data (Latin America and Caribbean)	Alt		–0.41 (–1.16, 0.34)
	Self-report data (north Africa and Middle East)	Alt		–0.48 (–1.23, 0.27)
	Self-report data (south Asia)	Alt		0.50 (–0.25, 1.26)
	Self-report data (sub-Saharan Africa)	Alt		–0.41 (–1.16, 0.34)
Males	Measured data	Ref	0.74	
	Self-reported data (southeast Asia, east Asia, and Oceania)	Alt		0.04 (–1.41, 1.53)
	Self-reported data (central Europe, eastern Europe, and central Asia)	Alt		–0.79 (–2.25, 0.71)
	Self-reported data (high-income)	Alt		–0.13 (–1.58, 1.40)
	Self-reported data (Latin America and Caribbean)	Alt		–0.26 (–1.70, 1.21)
	Self-report data (north Africa and Middle East)	Alt		–0.33 (–1.77, 1.16)
	Self-report data (south Asia)	Alt		0.66 (–0.78, 2.15)
	Self-report data (sub-Saharan Africa)	Alt		–0.41 (–1.86, 1.08)

**MR-BRT crosswalk adjustments can be interpreted as the factor the alternative case definition is adjusted by to reflect what it would have been had it been measured using the reference case definition. If the log/logit beta coefficient is negative, then the alternative is adjusted up to the reference. If the log/logit beta coefficient is positive, then the alternative is adjusted down to the reference.*

Table 3: MR-BRT self-report crosswalk adjustment factors for USA overweight and obesity prevalence

Model	Data input	Reference or alternative case definition	Gamma	Beta coefficient, logit (95% UI) *
Overweight prevalence				
Females	Measured NHANES data	Ref	0.0052	---
	Self-report (intercept)	Alt		0.08 (0.05, 0.11)
	Self-report (5-year age group)	Alt		–0.02 (–0.02, -0.01)
	Self-report data (decade)	Alt		0 (–0.02, 0.02)
Males	Measured NHANES data	Ref	0.016	---
	Self-report (intercept)	Alt		-0.42 (-0.46, -0.37)
	Self-report (5-year age group)	Alt		-0.003 (–0.005, -0.001)
	Self-report data (decade)	Alt		0 (–0.03, 0.03)
Obesity prevalence				
Females	Measured NHANES data	Ref	0.012	---
	Self-report (intercept)	Alt		-0.45 (-0.49, -0.41)
	Self-report (5-year age group)	Alt		0.003 (0.001, 0.004)
	Self-report data (decade)	Alt		0.01 (-0.02, 0.04)
Males	Measured NHANES data	Ref	0.018	---
	Self-report (intercept)	Alt		-0.46 (-0.50, -0.41)
	Self-report (5-year age group)	Alt		0 (–0.002, 0.001)

	Self-report data (decade)	Alt		0.01 (–0.02, 0.04)

**MR-BRT crosswalk adjustments can be interpreted as the factor the alternative case definition is adjusted by to reflect what it would have been had it been measured using the reference case definition. If the log/logit beta coefficient is negative, then the alternative is adjusted up to the reference. If the log/logit beta coefficient is positive, then the alternative is adjusted down to the reference.*

Modelling strategy

Exposure

Prevalence estimation for overweight and obesity

After adjusting for self-report bias and splitting aggregated data into five-year age-sex groups, we used ST-GPR to estimate the prevalence of overweight and obesity. This modelling approach has been described in detail elsewhere.

The linear model, which when added to the smoothed residuals forms the mean prior for GPR is as follows:

$$\text{logit}(\text{overweight})_{c,a,t} = \beta_0 + \beta_1 \text{educ}_{c,t} + \beta_2 \text{urban}_{c,t} + \beta_3 \text{agriculture}_{c,t} + \sum_{k=1}^{16} \beta_k I_{A[a]} + \alpha_s + \alpha_r + \alpha_c$$

$$\text{logit}(\text{obesity/overweight})_{c,a,t} = \beta_0 + \beta_1 \text{educ}_{c,t} + \beta_2 \text{urban}_{c,t} + \beta_3 \text{agriculture}_{c,t} + \sum_{k=1}^{16} \beta_k I_{A[a]} + \alpha_s + \alpha_r + \alpha_c$$

where educ is the age-standardised level of educational attainment; urban is the proportion of the population living in an urban area ; and agriculture is the proportion of the population working in agriculture. $I_{A[a]}$ is a dummy variable indicating specific age group A that the prevalence point captures, and $\alpha_s, \alpha_r, \text{and } \alpha_c$ are super-region, region, and country nested random intercepts, respectively. Random effects were used in model fitting but were not used in prediction.

We tested all combinations of the following covariates to see which performed best in terms of in-sample AIC for the overweight linear model and the obesity as a proportion of overweight linear model: ten-year lag-distributed energy per capita, proportion of the population living in urban areas, SDI, lag-distributed income per capita, educational attainment (years) per capita, proportion of the population working in agriculture, grams of sugar adjusted for energy per capita, grams of sugar not adjusted for energy per capita, and the number of two- or four-wheeled vehicles per capita. We selected these candidate covariates based on theory as well as reviewing covariates used in other publications. The final linear model was selected based on 1) if the direction of covariates matched what is expected from theory, 2) all the included covariates were significant, and 3) minimising in-sample AIC. The covariate selection process was performed using the dredge package in R.

Estimating mean BMI

To estimate the mean BMI for adults in each country, age, sex, and time period 1980–2021, we first used the following nested hierarchical mixed-effects model, fit using restricted maximum likelihood on data from sources containing estimates of all three indicators (prevalence of overweight, prevalence of obesity, and mean BMI), in order to characterise the relationship between overweight, obesity, and mean BMI:

$$\log(\mathrm{BMI}_{c,a,s,t}) = \beta_0 + \beta_1 \mathrm{ow}_{c,a,s,t} + \beta_2 \mathrm{ob}_{c,a,s,t} + \beta_3 \mathrm{sex} + \sum_{k=1}^{12} \beta_k I_{A[a]} + \alpha_s(1 + \mathrm{ow}_{c,a,s,t} + \mathrm{ob}_{c,a,s,t}) + \alpha_r(1 + \mathrm{ow}_{c,a,s,t} + \mathrm{ob}_{c,a,s,t}) + \alpha_c(1 + \mathrm{ow}_{c,a,s,t} + \mathrm{ob}_{c,a,s,t}) + \epsilon_{c,a,s,t}$$

where $\mathrm{ow}_{c,a,s,t}$ is the prevalence of overweight in country c, age a, sex s, and year t, $\mathrm{ob}_{c,a,s,t}$ is the prevalence of obesity in country c, age a, sex s, and year t, sex is a fixed effect on sex, $I_{A[a]}$ is an indicator variable for age, and α_s, α_r, and α_c are random effects at the super-region, region, and country level, respectively. The model was run in Stata 13.

We applied 1000 draws of the regression coefficients to the 1000 draws of overweight prevalence and obesity prevalence produced through ST-GPR to estimate 1000 draws of mean BMI for each country, year, age, and sex. This approach ensured that overweight prevalence, obesity prevalence, and mean BMI were correlated at the draw level and uncertainty was propagated.

Estimating BMI distribution

We used the ensemble distribution approach described in the manuscript. We fit ensemble weights by source and sex, with source- and sex-specific weights averaged across all sources included to produce the final global weights. The ensemble weights were fit on measured microdata. The final ensemble weights were exponential = 0.002, gamma = 0.028, inverse gamma = 0.085, log-logistic = 0.187, Gumbel = 0.220, Weibull = 0.011, log-normal = 0.058, normal = 0.012, beta = 0.136, mirror gamma = 0.008, and mirror Gumbel = 0.113.

1000 draws of BMI distributions for each location, year, age group, and sex estimated were produced by fitting an ensemble distribution using 1000 draws of estimated mean BMI, 1000 draws of estimated standard deviation, and the ensemble weights. Estimated standard deviation was produced by optimising a standard deviation to fit estimated overweight prevalence draws and estimated obesity prevalence draws.

Relative risk

In previous rounds of GBD, we reported the relative risk per five-unit change in BMI for disease endpoints using meta-analyses, and where available, pooled analyses of prospective observational studies. In GBD 2021, we assessed risk–outcome pairs included in previous rounds of the GBD based on the available evidence supporting a causal effect. We used MR-BRT to estimate the non-linear dose–response relationships between high BMI and risk for 26 disease endpoints. Specifically, we used the evidence score framework to systematically determine the risk function and evaluate the strength of evidence for each risk–outcome pair. Further details on the evidence score framework are available in the general methods of the Appendix.

The shape of dose–response relationships between BMI and risk for diseases has been well defined.[3,4] To best account for the various shapes (eg, J-shaped, increasing, and decreasing) of these relationships, we used the MR-BRT tool to estimate the log relative risk associated with each level of BMI on a continuous scale. Outcome-specific model characteristics are described in Table 4.

For each risk–outcome pair meta-regression, we considered study-level covariates that could potentially bias the study's reported effect size estimates. These study-level covariates included indication of whether the study used a washout period, whether the study population was randomly sampled from the

general population, whether the study measured or asked participants to self-report baseline BMI levels, whether the study determined outcomes based on administrative records or self-reports, and the level of adjustment for relevant confounders like age, sex, smoking, education, and income. We adjusted for these covariates in our meta-regression if they significantly biased our estimated relative risk function.

We implemented the Fisher scoring correction to the heterogeneity parameter, which corrects for data-sparse situations. In such cases, the between-study heterogeneity parameter estimate may be 0, simply from lack of data. The Fisher scoring correction uses a quantile of gamma, which is sensitive to the number of studies, study design, and reported uncertainty.

We also added methodology to detect and flag publication bias. The approach is based on the classic Egger's regression strategy, which is applied to the residuals in our model. In the current implementation, we do not correct for publication bias, but flag the risk–outcome pairs where the risk for publication bias is significant. We found no evidence of publication bias for the outcomes associated with high body-mass index.

There is a well-documented attenuation of the risk for cardiovascular disease and diabetes due to metabolic risks factors throughout one's life.[5] To incorporate this age trend in the relative risks, we first identified the median age-at-event across all cohorts and considered that as the reference age group. We then assigned our risk curves to this reference age group. Then, we derived attenuation factors by taking the ratio of excess risk between each age group and the reference. Finally, we applied 1000 draws of the age-specific attenuation factors to 1000 draws of the reference age group's risk curve to determine age-specific risk curves that propagated the uncertainty of both the risk function and age pattern.

For children and adolescent outcomes (ages 2–19), we computed dichotomous relative risks for overweight and obesity by modelling the log difference in relative risk between alternative groups (ie, overweight or obese) and reference groups (ie, normal weight) from prospective cohort studies.

Table 4: Model characteristics for outcomes related to high body-mass index in adults

Outcome	Non-linear specifications and constraints	Selected covariates	Mean gamma solution	Publication bias
Alzheimer's disease and other dementias	*	Reverse causality; representative population	0.332	No
Asthma	*		0.020	No
Atrial fibrillation and flutter	*		0.016	No
Breast cancer (in premenopausal women)	*		0.000	No
Breast cancer (in postmenopausal women)	*	Representative population	0.110	No
Cataract	**		0.157	No
Colon and rectum cancer	*		0.000	No

Diabetes mellitus type 2	*	Objective exposure measurement; objective outcome ascertainment	0.087	No
Gallbladder and biliary diseases	*	Objective outcome ascertainment; level of adjusted confounders	0.049	No
Gallbladder and biliary tract cancer	*		0.000	No
Gout	*		0.000	No
Intracerebral haemorrhage and Subarachnoid haemorrhage	*	Objective exposure measurement	0.118	No
Ischaemic heart disease	*	Objective exposure measurement	0.106	No
Ischaemic stroke	*		0.458	No
Kidney cancer	*		0.036	No
Leukaemia	*		0.000	No
Liver cancer	*		0.032	No
Low back pain	*		0.000	No
Multiple myeloma	**		0.000	No
Non-Hodgkin lymphoma	*		0.058	No
Osteoarthritis	*		0.045	No
Ovarian cancer	*		0.000	No
Pancreatic cancer	*		0.019	No
Thyroid cancer	*		0.000	No
Uterine cancer	*		0.008	No

** Cubic splines with 5 knots; left and right linear tails; Gaussian prior (0, 0.01) on max derivative of non-linear intervals.*

*** Cubic splines with 5 knots; left and right linear tails; Gaussian prior (0, 0.01) on max derivative of non-linear intervals.*

Theoretical minimum risk exposure level

For adults (ages 20+), the theoretical minimum risk exposure level (TMREL) of BMI (20–23 kg/m^2) was determined based on the BMI level that was associated with the lowest risk of all-cause mortality. Briefly, after estimating all-age, cause-specific dose–response risk curves, we generated 1000 draws of an all-cause mortality risk curve by taking weighted averages of 1000 draws of cause-specific risk curves. The weights were determined from the number of cause-specific global deaths from the GBD 2021 Causes of Death analysis. By generating the all-cause risk curve at the draw level, we were able to determine a distribution of the BMI levels that minimised all-cause mortality by assessing the level of BMI that minimised the risk for each of the 1000 draws.

For children and adolescents (ages 2–19), the TMREL is “normal weight,” that is, not overweight or obese, based on IOTF cutoffs.[1]

References

1.) Cole TJ, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatr Obes* 2012;**7**(4):284—94.

2.) Ng M, Fleming T, Robinson M, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2014; **384**: 766–81.

3.) Angelantonio ED, Bhupathiraju SN, Wormser D, et al. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet* 2016; **388**: 776—86. doi: 10.1016/S0140-6736(16)30175-1.

4.) Bhaskaran K, Dos-Santos-Silva I, Leon DA, Douglas IJ, Smeeth L. Association of BMI with overall and cause-specific mortality: a population-based cohort study of 3·6 million adults in the UK. *Lancet Diabetes Endocrinol* 2018; **6**(12): 944—53.

5.) Singh GM, Danaei G, Farzadfar F, et al. The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. *PLoS One* 2013; **8**(7): e65174.