nature portfolio

Corresponding author(s): Jean Joel Bigna

Last updated by author(s): Nov 13, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Using a pretested form, we extracted bibliometric information, country of recruitment, period of participants' inclusion, representativeness of the sample (national, subnational), mean or median age of the population, proportion of males, proportion of participants living in an urban area, diagnostic criteria for metabolic syndrome, number of participants included in the study and number with metabolic syndrome. When possible, we extracted data by age group, sex, setting, and year of data collection. Data were extracted from all included studies in this review.
Data analysis	We used a Bayesian multilevel logistic regression model to estimate the global, regional, and national prevalence of metabolic syndrome between 2000 and 2023, separately for men and women (Appendix: Supplementary Text 1). The modelling approach is aligned with previously published statistical and substantive frameworks. This hierarchical strategy accounted for the nested structure of the data, with prevalence estimates from individual studies clustered within countries, regions, and super-regions. The outcome was modelled as the number of individuals with metabolic syndrome out of the total number examined, assuming a binomial distribution with a logit link function. Several fixed effects were incorporated to account for population and study-level characteristics. Age was modelled as a continuous variable, centred at 50 years to improve model convergence and interpretability. To capture non-linear relationships between age and prevalence, we applied natural cubic splines. For men, knots were placed at 35, 65, and 75 years, with boundary knots at 20 and 80 years. For women, knots were located at 40, 60, and 65 years, with the same boundary knots. These placements were informed by epidemiological evidence suggesting that metabolic syndrome and its cardiometabolic components exhibit non-linear age trajectories, with key inflection points around midlife and older age. Additional fixed effects included the national representativeness of the study (national vs subnational), urbanicity (proportion of the population living in urban areas), gross domestic product (GDP) per capita (log-transformed), and the diagnostic definition of metabolic syndrome. For men, diagnostic categories included the JIS with waist circumference thresholds of 90 cm or 94 cm, as well as other definitions.

For women, definitions included the JIS threshold of 80 cm and others. An interaction between representativeness and log GDP was incorporated to assess whether the influence of economic context on prevalence varied by study coverage, capturing potential inconsistencies between national economic indicators and subnational data.

Random effects were introduced at multiple geographic levels to reflect spatial clustering and unmeasured heterogeneity. Random intercepts and age slopes were included at the super-regional level, allowing the age-prevalence association to vary across broad epidemiological zones. Nested random intercepts at subregional and national levels were also included to account for additional within-region and within-country variation not explained by fixed effects.

Priors were specified based on established epidemiological knowledge and assumptions about the distribution and determinants of metabolic syndrome. For the model intercept, we applied informative priors to reflect plausible global prevalence values under the JIS definition,8 anchoring the model without overly constraining it. For fixed effects, weakly informative priors were applied, centred at zero but allowing moderate variation, to reflect the well-documented associations of age, urbanicity, and GDP with metabolic syndrome while remaining conservative. Priors for the interaction term between GDP and representativeness were centred around a null effect. Priors on random effect standard deviations were weakly regularizing, helping to stabilize variance estimation without suppressing meaningful heterogeneity. Correlations between intercepts and age slopes at the super-regional level were also modelled, with structured priors permitting moderate correlation.

Models were implemented using the brm function from the brms package (version 2.22.0) in R (version 4.3.3), with Stan as the backend via cmdstanr. Estimation used four parallel Markov chains, each with 5,000 iterations and a 2,500-iteration warm-up. Posterior distributions were sampled using Hamiltonian Monte Carlo, yielding 10,000 post-warm-up draws.

Posterior predictions were generated to estimate the prevalence of metabolic syndrome across all countries and territories included in the study, including those without primary data. The predictions were informed by a set of time-varying covariates (age distribution, GDP per capita, and level of urbanicity) specific to each year between 2000 and 2023. This enabled the model to capture year-to-year changes in the prevalence of metabolic syndrome in relation to shifting demographic and socioeconomic conditions. In countries with available data, estimates incorporated both fixed effects and country-specific random effects. In countries without direct data but located in regions with available data, predictions were based on fixed effects and random effects at the regional and super-regional levels. For countries with no data at the country or regional level, predictions relied on fixed effects and super-regional random effects. This hierarchical modelling structure allowed estimates to borrow strength across geographic levels while maintaining internal coherence and accounting for data sparsity. Uncertainty was expressed through 95% credible intervals (95% CrI) derived from the posterior distributions, capturing both parameter uncertainty and structural variability across time and space.

The R code for modelling the prevalence of metabolic syndrome is available for download at : Github: https://doi.org/10.5281/zenodo.17599793.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data used in this study have been deposited in the Zenodo repository under accession code https://doi.org/10.5281/zenodo.15923329. The data are publicly available and unrestricted. No individual-level raw data were collected; therefore, no privacy restrictions apply. The processed data used to generate the figures and tables are available in the Source Data file. All authors had access to the data used in this study.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	In this study, we provide for the first time estimates of the prevalence and absolute burden of metabolic syndrome in women and men across 198 countries and territories from 2000 to 2023. When possible, we extracted data by age group, sex, setting, and year of data collection. According to sex, 1910 prevalence data were from women and 1310 from females. We used a Bayesian multilevel logistic regression model to estimate the global, regional, and national prevalence of metabolic syndrome between 2000 and 2023, separately for men and women.
Reporting on race, ethnicity, or other socially relevant groupings	Not applicable.
Population characteristics	Not applicable.
Recruitment	Not applicable.
Ethics oversight	Not applicable.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The number of datasets involved in the study was determined through a comprehensive systematic review of population-based observational studies, following the Joanna Briggs Institute methodology. The authors searched multiple international databases—including PubMed, EMBASE, Web of Knowledge, African Journals Online, and Global Index Medicus—up to September 20, 2022, using PRESS-compliant strategies and inclusive search terms such as "metabolic syndrome," "insulin resistance syndrome," and "cardiometabolic syndrome." From this process, 597 eligible reports were identified, yielding 3,236 prevalence data points derived from 45,549,151 adult participants. Studies were included regardless of diagnostic criteria, provided they used probability-based sampling and included at least 30 participants from the general adult population.
Data exclusions	Data exclusions in this study were based on predefined eligibility criteria established prior to data extraction. Specifically, studies were excluded if they were not conducted in the general adult population, lacked probability-based sampling methods, included fewer than 30 participants, or did not provide sufficient data to calculate crude prevalence. Duplicate reports or overlapping datasets were also excluded to avoid redundancy. These criteria were consistently applied during the systematic review process by independent reviewers using the Rayyan platform, with any disagreements resolved through consensus. The exclusions ensured methodological rigor and enhanced the reliability and generalizability of the findings.
Replication	Although a pre-registered protocol was not implemented for this study, several measures were taken to ensure the reproducibility of the findings. The reproducibility of the experimental findings was ensured through several methodological safeguards. First, the study adhered to established guidelines for systematic reviews and meta-analyses, including the Joanna Briggs Institute methodology and PRISMA standards. All data extraction and quality assessments were conducted independently by multiple reviewers, with discrepancies resolved by consensus, ensuring consistency and minimizing bias. The statistical modelling approach—Bayesian multilevel logistic regression—was implemented using open-source software (R version 4.3.3 and the brms package version 2.22.0), with all model specifications, priors, and covariates clearly documented. The code used for data analysis has been made publicly available via GitHub, allowing full transparency and enabling independent replication.
Randomization	Randomization was not applicable to this study. The research was based on a systematic review and modelling analysis of previously published, population-based observational studies. As there was no experimental intervention or assignment of participants to different groups by the investigators, randomization was not relevant to the study design or analysis.
Blinding	Blinding of investigators to group allocation during data collection and analysis was not applicable to this study. The research was based on a systematic review and modelling analysis of previously published, population-based observational studies. As such, the investigators did not conduct primary data collection or assign participants to experimental groups. Instead, they extracted and analyzed aggregated prevalence data from existing sources. Therefore, blinding was not relevant to the study design or methodology.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	Not applicable.
Novel plant genotypes	Not applicable.
Authentication	Not applicable.