

Small-area inequalities in weight excess and obesity among adults in Belo Horizonte, Brazil: a repeated cross-sectional ecological study using Vigitel and synthetic populations

Larissa F Araújo

larissafortunatoaraujo@gmail.com

Universidade Federal de Minas Gerais

Jussara M Almeida

Universidade Federal de Minas Gerais

Crizian S Gomes

Universidade Federal de Minas Gerais

Regina T I Bernal

Universidade Federal de Minas Gerais

Marcos A Gonçalves

Universidade Federal de Minas Gerais

Gregorio V Rodrigues

Universidade Federal de Minas Gerais

Thania MT R Faria

Universidade Federal de Minas Gerais

Celso França

Universidade Federal de Minas Gerais

Deborah C Malta

Universidade Federal de Minas Gerais

Research Article

Keywords: Weight excess, Obesity, Small-area analysis, Health vulnerability, Synthetic population, Vigitel, Sex differences, Urban health, Brazil

Posted Date: July 1st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9888682/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Public Health on September 16th, 2026. See the published version at <https://doi.org/10.1186/s12889-026-29377-8>.

Abstract

Background

Weight excess and obesity are increasing in Brazil and may be unevenly distributed within cities. However, city-level estimates can mask intra-urban inequalities and provide limited evidence for local public health planning. This study estimated the prevalence of weight excess and obesity among adults living in Belo Horizonte, Brazil, from 2006 to 2018, and examined temporal and territorial inequalities according to sex and Health Vulnerability Index (HVI)-based microareas.

Methods

This ecological study used repeated cross-sectional data from Vigitel, a national telephone-based surveillance system. Survey years were grouped into 2006–2009, 2010–2013 and 2014–2018. Weight excess (body mass index ≥ 25.0 kg/m²) and obesity (≥ 30.0 kg/m²) were analysed as binary outcomes. Estimates were stratified by sex and by nine HVI-based microarea clusters derived from socioeconomic, sanitation and housing indicators. Synthetic populations generated using conditional generative adversarial networks were incorporated in clusters with sparse observations to improve the stability of small-area estimates. Weighted analyses accounted for the complex survey design.

Results

The prevalence of weight excess and obesity increased over time in Belo Horizonte, with marked differences by sex and territorial vulnerability. Weight excess was generally more common among men, especially in lower-vulnerability clusters, whereas obesity was more frequent among women and showed a clearer social gradient in intermediate-, high- and very high-vulnerability clusters. The inclusion of synthetic populations increased the effective sample size in data-sparse microareas and improved the stability of local estimates.

Conclusions

Weight excess and obesity in Belo Horizonte were structured by the intersection of sex and territorial vulnerability. Small-area surveillance combining survey data, vulnerability-based territorial clusters and synthetic populations can make intra-urban inequalities more visible and support more targeted public health action.

Background

Weight excess and obesity are among the most important public health challenges of the twenty-first century because they contribute substantially to premature mortality, disability and the burden of non-communicable diseases worldwide. Their increase cannot be explained by individual behaviour alone. Excess body weight is shaped by food systems, labour organisation, transport patterns, housing conditions and the built environment, all of which structure daily opportunities for healthy eating and physical activity [1–3]. In cities, these conditions are distributed unequally, so the chance of protecting health varies according to where people live [4, 5].

Brazil has experienced this process intensely. National surveillance and population-based studies have shown a sustained increase in weight excess and obesity among adults, indicating that the nutritional transition has occurred alongside persistent social inequalities [6–8]. These inequalities are also territorial. Within the same city, residents may face very different food environments, unequal affordability of healthy foods, distinct levels of walkability and safety, and varying forms of work-related stress and time scarcity [5, 9–11].

Despite this complexity, weight excess and obesity indicators are usually produced for entire cities or broad sociodemographic groups, which may conceal important intra-urban heterogeneity. In Brazil, the Surveillance of Risk and Protective Factors for Chronic Diseases by Telephone Survey (Vigitel) has been essential for monitoring behavioural risk factors in the capitals and

documenting temporal trends in excess body weight [12, 13]. However, city-level averages may be insufficient for local planning because they can obscure priority areas and flatten social gradients within cities [14, 15].

Belo Horizonte is a relevant setting in which to address this gap. As a large Brazilian city marked by pronounced social and spatial inequalities, it combines a complex urban structure with a long-standing tradition of territorial planning in health. In this context, the Health Vulnerability Index (HVI) provides an analytically meaningful basis for identifying socially coherent microareas, since it incorporates socioeconomic, demographic and environmental dimensions that shape living conditions across the city [16]. Examining weight excess and obesity through HVI-based microareas may therefore reveal patterns hidden by municipal averages and strengthen the territorial intelligence needed for local surveillance and action [14].

Sex and gender-related differences further deepen these inequalities. In Brazil, weight excess has often been more frequent among men, whereas obesity tends to be more prevalent among women, especially among those living in socially disadvantaged contexts [6–8]. This distinction is epidemiologically important because weight excess and obesity, although related, may reflect different social processes. Gendered divisions of paid and unpaid labour, food insecurity, chronic stress and barriers to leisure-time physical activity may affect women and men differently, particularly in vulnerable territories [9, 11, 17].

A further challenge concerns the instability of local estimates derived from surveys when the number of observations is small. Socially vulnerable territories are often underrepresented in conventional samples, which reduces precision where detailed evidence is most needed. Recent advances in small-area estimation and synthetic population methods have expanded the possibilities for producing more stable local estimates while preserving the empirical basis of survey data [14, 15, 18–20].

From the perspective of food and nutrition public policies, this debate is especially relevant because obesity prevention depends on understanding how food environments and social vulnerability shape dietary practices within cities [5, 21]. Against this background, the present study estimated the prevalence of weight excess and obesity among adults living in Belo Horizonte between 2006 and 2018, analysing temporal trends and inequalities according to sex and HVI-based microareas, and explored the contribution of synthetic population strategies to improve the stability of local estimates.

Methods

Study design and setting

This ecological study analysed repeated cross-sectional data from the Surveillance of Risk and Protective Factors for Chronic Diseases by Telephone Survey (Vigitel), restricted to adult residents (≥ 18 years) of Belo Horizonte, Minas Gerais, Brazil, between 2006 and 2018. Belo Horizonte is a large metropolitan capital characterised by pronounced social and spatial inequalities, making it a relevant setting for the investigation of intra-urban differences in weight excess and obesity.

Vigitel is a national surveillance system implemented in 2006 by the Brazilian Ministry of Health to monitor behavioural and metabolic risk factors for non-communicable diseases in Brazilian state capitals using standardized annual surveys [12]. The system provides consistent time-series data and constitutes a key source for monitoring trends in overweight and obesity in Brazil [13]. The Vigitel methodology and questionnaire have been previously published [12,13].

Sampling design and weighting

The methodological framework of Vigitel has been described in detail elsewhere [12]. Briefly, the survey employs a two-stage probabilistic sampling design. In the first stage, landline telephone numbers are randomly selected from updated registries. In the second stage, one adult resident is randomly selected within each eligible household. To mitigate selection bias inherent to landline-based surveys and improve representativeness, post-stratification weights are applied. These weights adjust the sample distribution to match the adult population of Belo Horizonte by sex, age group and educational level, based on census projections [12,13]. All analyses accounted for sampling weights and the complex survey design.

Temporal stratification

To improve statistical stability and reduce random variability associated with annual estimates in small-area analyses, survey years were grouped into three consecutive periods: 2006-2009, 2010-2013 and 2014-2018. Temporal aggregation is recommended in small-area analyses based on survey data, as it increases effective sample size while preserving sensitivity to medium-term changes in prevalence [14].

Outcomes

The outcomes of interest were weight excess and obesity, defined using body mass index (BMI). In Vigitel, body weight and height are obtained through self-report during telephone interviews, in which respondents are asked to report their current weight in kilograms and height in meters [12,13]. BMI was calculated as weight (kg) divided by height squared (m^2). Weight excess was defined as $BMI \geq 25.0 \text{ kg/m}^2$ and obesity as $BMI \geq 30.0 \text{ kg/m}^2$, in accordance with World Health Organization recommendations [1]. Both outcomes were analyzed as binary variables. Although self-reported anthropometric measures may be subject to reporting bias, validation studies indicate acceptable agreement with measured values for population-level surveillance and trend analysis [22].

Sex stratification

All analyses were stratified by sex (men and women), as reported in the Vigitel questionnaire. Stratification was specified a priori, given robust evidence of sex-specific differences in the prevalence, determinants and temporal trajectories of overweight and obesity in Brazil and internationally [3,6].

Definition of microareas and HVI-based clusters

Microareas within Belo Horizonte were defined using clusters derived from the Health Vulnerability Index, a composite indicator integrating census-based socioeconomic, demographic and environmental variables and widely used by the municipality for health planning and resource allocation [16,23].

The original HVI classification includes four broad vulnerability categories (low, medium, high and very high). However, substantial internal heterogeneity within these categories, particularly in areas of higher vulnerability, may compromise the precision and interpretability of small-area health estimates [14]. To address this limitation, nine HVI-based clusters were employed as small areas: LA-0 (low vulnerability, subtype 0), LA-1 (low vulnerability, subtype 1), MD-0 (medium vulnerability, subtype 0), MD-1 (medium vulnerability, subtype 1), HI-1 (high vulnerability, subtype 0), HI-2 (high vulnerability, subtype 1), HI-3 (high vulnerability, subtype 2), VH-0 (very high vulnerability, subtype 0) and VH-1 (very high vulnerability, subtype 1).

These clusters were obtained using hierarchical clustering techniques applied to two key HVI components available per census tract: sanitation and housing conditions, and socioeconomic status, particularly income and education indicators. The best number of clusters was defined using the Silhouette Index [24]. The resulting groups were socially coherent, reduced internal heterogeneity and enhanced epidemiological relevance.

Each Vigitel respondent residing in Belo Horizonte was assigned to one of the nine HVI-based clusters according to the geocoded residential address, linked to census tracts using official municipal and national geographic databases [16,23].

Synthetic population generation

Sparse observations in socially vulnerable areas limit the reliability of small-area estimates derived from survey data. To address this issue, synthetic populations were incorporated as a form of statistical augmentation.

Synthetic populations consist of artificially generated datasets designed to reproduce the statistical properties of a real population while avoiding disclosure of individual-level information. In this work, synthetic data were generated using conditional generative adversarial networks (cGANs), which can model complex joint distributions in tabular health data while preserving privacy [18,19].

Synthetic samples were introduced only in clusters with fewer than 1,000 interviews over the aggregated period. The generation process preserved the empirical distribution of key variables by sampling proportionally to observed marginals. Generated

samples were then assigned to geographic small areas using supervised models, specifically an XGBoost classifier [25].

To ensure fidelity, similarity between real and synthetic data was evaluated using the KSComplement metric, where values closer to 1 indicate higher distributional alignment [26]. Additional distributional diagnostics included comparisons of marginal distributions and pairwise associations across key variables.

Models were trained separately for each temporal period using Vigitel data from Belo Horizonte. The feature space included age, educational level, BMI category and HVI-based cluster. Synthetic data generation targeted both marginal and joint distributions without replicating identifiable individuals, thereby preserving privacy. Importantly, synthetic observations served to augment, rather than replace, the original data, maintaining empirical grounding while improving robustness in data-sparse regions [18,27].

Statistical analysis

For each sex, temporal period and HVI-based cluster, prevalences of weight excess and obesity were estimated along with 95 % confidence intervals (95 % CI). All analyses accounted for the complex sampling design of Vigitel, including sampling weights and clustering.

Prevalence estimates were obtained using direct small-area estimation, consistent with recommendations for analysing survey data in geographically and socially stratified subgroups [14,15]. Temporal and spatial differences in prevalence were assessed descriptively through comparison of confidence intervals. These temporal increases were interpreted cautiously, with non-overlapping 95 % CI considered evidence of robust differences between periods. Differences between men and women were assessed using p values derived from weighted analyses, adopting a 5 % significance level.

Results

Table 1 presents the prevalence of weight excess and obesity according to sex, time period and HVI-based microarea cluster in Belo Horizonte. Overall, weight excess increased between 2006-2009 and 2014-2018 in both men and women across most microareas. These increases were evident in several clusters, particularly in areas of medium, high and very high vulnerability, as indicated by non-overlapping 95 % CI between periods.

Obesity also increased over time, although the pattern was less uniform than that observed for weight excess. Among women, the increase in obesity was more consistent across periods and was especially evident in clusters of intermediate, high and very high vulnerability, particularly between 2010-2013 and 2014-2018. Among men, increases in obesity were observed only in selected clusters, mainly in more vulnerable areas, whereas in less vulnerable clusters the 95 % CI overlapped across periods.

Sex differences also varied according to the outcome and the social profile of the microarea. Weight excess was generally more prevalent among men than among women in clusters of low and medium vulnerability and across most periods, with weighted p values indicating differences in several comparisons. However, this difference narrowed with increasing vulnerability, and in some clusters of high and very high vulnerability women showed similar or higher prevalence than men. In these more vulnerable clusters, sex differences were reduced or reversed depending on the period and microarea.

For obesity, women consistently showed higher prevalence than men in clusters of intermediate to very high vulnerability, with weighted p values indicating differences in most comparisons. These inequalities were observed across periods and became more marked in 2014-2018. In contrast, in the least vulnerable clusters, obesity prevalence was more similar between men and women.

Table 2 shows the sample size by period and data source. The distribution of respondents across HVI-based clusters remained broadly stable over time, although lower numbers were observed in the more vulnerable microareas, especially in the earliest period. The inclusion of synthetic populations increased the effective sample size in these clusters and contributed to greater stability of the prevalence estimates.

Discussion

The present study showed three main findings. First, weight excess increased over time in Belo Horizonte in both sexes and across most HVI-based microareas. Second, obesity displayed a steeper and more socially patterned distribution than weight excess, particularly among women living in intermediate-, high- and very high-vulnerability clusters. Third, the sex pattern varied according to territorial vulnerability, which men tended to show higher prevalence of weight excess in less vulnerable clusters, whereas women reached similar or higher prevalences in the most vulnerable clusters, a pattern that was consistent with the clearer female gradient observed for obesity. Together, these findings indicate that excess body weight in Belo Horizonte is not evenly distributed across urban space, but is structured by the interaction between sex, gendered social processes and territorial deprivation [6, 8, 14].

The temporal increase in weight excess and obesity is consistent with national and international evidence showing that the nutritional transition in middle-income settings has been accompanied by persistent growth in excess body weight, despite advances in health promotion discourse and regulation [2, 3, 6]. In Brazil, these trends have been linked to the expansion of ultra-processed foods, changing work routines, greater dependence on motorised transport and unequal access to environments that support healthy living [5, 10, 21]. The present findings add an important intra-urban perspective to this literature by showing that city-wide averages may conceal marked territorial gradients [4, 9]. In Belo Horizonte, the burden of obesity was not simply higher in the city as a whole over time; it was progressively concentrated in microareas marked by greater vulnerability, particularly among women.

The contrast between sexes across the HVI gradient suggests that weight excess and obesity are shaped by related, but not identical, social mechanisms. Among men, the higher prevalence of weight excess in lower-vulnerability clusters may reflect the social distribution of earlier stages of the nutritional transition, in which excess body weight becomes common among groups with greater purchasing power, more frequent eating outside the home and more sedentary occupational patterns [7, 17]. In these territories, men may be more exposed to desk-based work, long commutes by car and energy-dense diets without the same concentration of extreme material deprivation [2, 5].

By contrast, in high-vulnerability clusters, women showed higher prevalence of weight excess and a clearer gradient for obesity. This pattern is compatible with a cumulative process in which food insecurity, precarious employment, unpaid care work, time poverty, chronic stress and restricted access to safe spaces for physical activity disproportionately affect women and intensify obesogenic pressures in deprived territories [8, 9, 11]. The fact that the female burden is more pronounced for obesity than for weight excess suggests not only differential exposure, but also differential progression along the body-weight continuum under conditions of social adversity [7, 8].

These findings are coherent with previous Brazilian studies showing that obesity is more socially concentrated among women and that inequalities in body weight are strongly conditioned by socioeconomic disadvantage [7, 8]. They also reinforce the interpretation that HVI is more than a contextual label. Because the index synthesises socioeconomic, demographic and environmental constraints, the clusters derived from it provide socially coherent territories in which shared exposures are more likely to be captured than when analyses rely on conventional administrative units alone [14, 16].

This is a relevant analytical gain because administrative units often combine highly heterogeneous living conditions and may dilute local gradients, whereas HVI-based clusters are better suited to identify where the overlap between deprivation and nutritional risk is strongest [4, 9]. In this study, the use of clusters made it possible to distinguish areas where male predominance in weight excess persisted from those where female disadvantage became more evident, thereby improving the epidemiological intelligibility of the urban pattern.

Small-area analyses based on surveys are frequently weakened by sparse observations in the very territories where more precise evidence is needed, especially in vulnerable urban areas that are often underrepresented in telephone-based samples [14, 15]. By using synthetic observations to augment, rather than replace, empirical data, the study improved the effective sample size in selected clusters and increased the stability of local estimates. This does not eliminate the need for caution. Synthetic populations depend on the quality of the original data, on modelling assumptions and on the capacity of the algorithm to reproduce relevant joint distributions without generating implausible patterns [18, 19].

They may reduce random instability, but they do not correct structural bias in the original data source, nor do they eliminate uncertainty linked to undercoverage [15, 27]. Their value lies in making small-area comparisons more robust when used transparently, with clear diagnostics and without obscuring the empirical basis of inference. In this sense, the present approach contributes not only substantively, but also methodologically, by showing how artificial intelligence tools can be integrated into public health nutrition research to strengthen territorial surveillance and reduce the analytic invisibility of populations recurrently missed by conventional sampling [18, 20].

From a public policy perspective, the findings support the need for interventions that are both territorialised and gender-responsive. Broad city-level strategies remain important, but they are unlikely to reduce inequalities if they fail to prioritise microareas where the burden is greatest [4, 14]. In Belo Horizonte and in similar urban contexts, local health authorities could use small-area evidence to guide the allocation of community nutrition actions, women-centred obesity prevention programmes, culturally appropriate dietary counselling and partnerships with social assistance services in high-vulnerability territories. Improvements in food environments are also essential, including support for fresh-food retail, restrictions on the saturation of ultra-processed food outlets around vulnerable neighbourhoods and policies that make healthy food more affordable [5, 21]. At the same time, the stronger obesity gradient among women indicates that nutritional policies should be articulated with transport, urban safety, childcare, income protection and decent-work policies, since the ability to eat well and engage in physical activity is shaped by the organisation of daily life rather than by information alone [4, 10, 11].

This study has limitations. Weight and height were self-reported, which may introduce measurement error, although this strategy is widely accepted for surveillance purposes and has shown adequate validity for population monitoring [22]. The ecological design does not permit individual causal inference, and the temporal comparison based on non-overlapping confidence intervals is conservative [15]. In addition, the synthetic data strategy depends on model performance and should not be interpreted as a substitute for better empirical coverage in vulnerable areas [18, 19]. Nevertheless, the study also has important strengths, including long temporal coverage, use of a consolidated national surveillance system, a small-area analytical strategy anchored in socially meaningful clusters and an innovative integration of synthetic populations to improve estimate robustness [6, 14, 20]. These attributes make the results particularly useful for equity-oriented surveillance and local decision-making.

Conclusions

This study shows that weight excess and obesity in Belo Horizonte are shaped by the joint action of sex, gendered social processes and territorial vulnerability. Men concentrated higher prevalence of weight excess in less vulnerable clusters, whereas women bore the heaviest burden in high-vulnerability areas, especially for obesity. The use of HVI-based clusters and synthetic populations made these inequalities more visible and analytically tractable (14,16,20). For public health nutrition, the results reinforce that effective obesity prevention in unequal cities requires moving beyond municipal averages and addressing the territorial and gendered mechanisms through which obesogenic conditions are produced and reproduced (4,5).

Abbreviations

BMI, body mass index; cGAN, conditional generative adversarial network; CI, confidence interval; HVI, Health Vulnerability Index; KS, Kolmogorov-Smirnov; LA, low vulnerability; MD, medium vulnerability; HI, high vulnerability; VH, very high vulnerability; Vigitel, Surveillance of Risk and Protective Factors for Chronic Diseases by Telephone Survey.

Declarations

Ethics approval and consent to participate

This study was conducted according to the guidelines laid down in the Declaration of Helsinki, and all procedures involving human participants were approved by the Research Ethics Committee of the Universidade Federal de Minas Gerais (approval no. 6,538,883). Vigitel obtains informed consent verbally during telephone interviews; for this secondary analysis of anonymized surveillance data, individual written consent was not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The databases used for this study were provided by the Brazilian Ministry of Health and, due to the sensitive nature of some information, such as individuals' addresses, are stored in a data center with high levels of physical and virtual security under the responsibility of the Federal University of Minas Gerais. The data center adopts strict standards of governance and ethics to preserve confidentiality, maintain and update databases securely, and ensure that data processing prevents any possibility of individual identification. Therefore, the databases are not available for public access.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Center for Innovation and Artificial Intelligence for Health (CI-IA Saúde), with resources from the São Paulo Research Foundation (FAPESP, grant number 2020/09866-4), the Minas Gerais Research Foundation (FAPEMIG, grant number PPE-00030-21) and UNIMED Belo Horizonte; CNPq and Decit/SECTICS/MS call 21/2023 (grant number 444533/2023-0); the Ministry of Health/National Health Fund (TED 67/2023); FAPEMIG call B01/2025 (grant number APQ 02385-25); and the Ministry of Health/National Health Fund (TED No. 114/2024 - Artificial Intelligence Center for Health - NIAR-Saúde - UFMG). The funders had no role in the design, analysis or writing of this article.

Authors' contributions

L.F.A. conceived the study, formulated the research question, contributed to the study design, analyzed and interpreted the data, and drafted the manuscript. J.M.A., M.A.G. and C.F. contributed to the computational methods, synthetic population strategy and interpretation of artificial intelligence procedures. C.S.G. and R.T.I.B. contributed to the study design, survey methodology, statistical analysis and interpretation of findings. G.V.R. contributed to data organization, geocoding procedures and analytical support. T.M.T.R.F. contributed to the construction and interpretation of HVI-based territorial clusters and to the public health interpretation of the findings. D.C.M. supervised the study and contributed to study conception, interpretation and critical revision. All authors critically revised the manuscript and approved the final version submitted for publication.

Acknowledgements

Not applicable.

References

1. World Health Organization. (2000) Obesity: preventing and managing the global epidemic. WHO Technical Report Series no. 894. Geneva: WHO.
2. Swinburn BA, Sacks G, Hall KD, et al. The global obesity pandemic: shaped by global drivers and local environments. *Lancet*. 2011;378:804–14. 10.1016/S0140-6736(11)60813-1.
3. GBD 2021 Risk Factors Collaborators. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet*. 2024;403:2162–203. 10.1016/S0140-6736(24)00933-4.
4. Diez Roux AV. Neighborhoods and health: what do we know? What should we do? *Annu Rev Public Health*. 2016;37:15–30. 10.1146/annurev-publhealth-032315-021402.
5. Westbury S, Ghosh I, Jones HM, et al. The influence of the urban food environment on diet, nutrition and health outcomes in low-income and middle-income countries: a systematic review. *BMJ Glob Health*. 2021;6:e006358. 10.1136/bmjgh-2021-006358.

6. Silva LES, Oliveira MM, Stopa SR, et al. Temporal trend of overweight and obesity prevalence among adults in Brazilian state capitals and the Federal District, 2006–2019. *Epidemiol Serv Saude*. 2021;30:e2020294. 10.1590/S1679-49742021000100052.
7. Sousa APM, Pereira IC, Araujo LL, et al. Prevalence and factors associated with overweight in adults in Brazilian capitals and the Federal District, 2019. *Epidemiol Serv Saude*. 2021;30:e2020838. 10.1590/S1679-49742021000300014.
8. Flauzino PA, Malta DC, Souza Junior PRB, et al. Analysis of inequities between demographic and social determinants associated with overweight and obesity in Brazilian adults. *Rev Bras Epidemiol*. 2025;28:e250038. 10.1590/1980-549720250038.
9. Kong AY, Zhang X. The use of small-area estimates in place-based health research. *Am J Public Health*. 2020;110:829–32. 10.2105/AJPH.2020.305611.
10. Schmidt MI, Duncan BB, Azevedo e Silva G, et al. Chronic non-communicable diseases in Brazil: burden and current challenges. *Lancet*. 2011;377:1949–61. 10.1016/S0140-6736(11)60135-9.
11. Triaca LM, dos Santos AMA, Tejada CAO. Socioeconomic inequalities in obesity in Brazil. *Econ Hum Biol*. 2020;39:100906. 10.1016/j.ehb.2020.100906.
12. Malta DC, Bernal RTI, Oliveira M, et al. The implementation of the Surveillance System for Non-Communicable Diseases in Brazil, 2003 to 2015: scope and challenges. *Rev Bras Epidemiol*. 2017;20:661–75. 10.1590/1980-5497201700040009.
13. Brasil, Ministerio da Saude. Vigitel Brasil 2018: vigilancia de fatores de risco e protecao para doencas cronicas por inquerito telefonico. Brasilia: Ministerio da Saude; 2019.
14. Bernal RTI, Felisbino-Mendes MS, Carvalho QH, et al. A methodology for small area prevalence estimation based on survey data. *Int J Equity Health*. 2020;19:124. 10.1186/s12939-020-01220-5.
15. Rao JNK, Molina I. Small Area Estimation. 2nd ed. Hoboken, NJ: Wiley; 2015.
16. Horizonte PdeB, Secretaria Municipal de Saude. (2013) Indice de Vulnerabilidade da Saude de Belo Horizonte 2012: metodologia. Belo Horizonte: Prefeitura de Belo Horizonte. Available at https://bhgeo.pbh.gov.br/sites/geoportal.pbh.gov.br/files/BHGE0/RECURSO_METADADO/IVS/METODOLOGIA_IVS_2012.pdf (accessed April 2026).
17. Kant AK, Graubard BI. Secular trends in patterns of body mass index in adults. *Am J Clin Nutr*. 2015;102:698–705. 10.3945/ajcn.115.107615.
18. Chapuis K, Taillandier P, Drogoul A. Generation of synthetic populations in social simulations: a review of methods and practices. *J Artif Soc Soc Simul*. 2022;25:6. 10.18564/jasss.4762.
19. Murtaza H, Ahmed M, Khan NF, et al. Synthetic data generation: state of the art in health care domain. *Comput Sci Rev*. 2023;48:100546. 10.1016/j.cosrev.2023.100546.
20. Faria TMTR, Bernal RTI, Malta DC, et al. Improving the mapping of leisure-time physical activity inequities: the use of artificial intelligence to advance estimates of small areas in Brazil. *Public Health*. 2025;243:105727. 10.1016/j.puhe.2025.105727.
21. Monteiro CA, Cannon G, Levy RB, et al. Ultra-processed foods: what they are and how to identify them. *Public Health Nutr*. 2019;22:936–41. 10.1017/S1368980018003762.
22. Spencer EA, Appleby PN, Davey GK, et al. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. *Public Health Nutr*. 2002;5:561–5. 10.1079/PHN2001322.
23. Pitchon A, Malta DC, Bernal RTI, et al. Indice de vulnerabilidade da saude 2012. Belo Horizonte: Prefeitura Municipal de Belo Horizonte; 2013.
24. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math*. 1987;20:53–65. 10.1016/0377-0427(87)90125-7.
25. Davagdorj K, Pham VH, Theera-Umpon N, et al. XGBoost-based framework for smoking-induced noncommunicable disease prediction. *Int J Environ Res Public Health*. 2020;17:6513. 10.3390/ijerph17186513.
26. Patki N, Wedge R, Veeramachaneni K. (2016) The Synthetic Data Vault. In *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 399–410. Montreal: IEEE. 10.1109/DSAA.2016.49

27. El Emam K, Jonker E, Arbuckle L, et al. A systematic review of re-identification attacks on health data. PLoS ONE. 2011;6:e28071. 10.1371/journal.pone.0028071.

Tables

Table 1 and 2 are available in the Supplementary Files section.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table.docx](#)