The Generative AI Meta-Evaluation (GAME) Study Framework: Global, Regional, and Country-Specific Unequal Difficulty of High BMI Intervention

Chen Sun[1,3,†], Chuanwu Liu[1,3,†], Wenhua Lv[1,3,†], Wei She[1,†], Siyu Wei[1], Haiyan Chen[1,3], Junxian Tao[1], Jing Xu[2],

Tianyang Lei[1,3], Qiong Wu[1,3], Yuan Xu[1], Ning Wang[1], Yan Guo[1], Qinduo Ren[1], Chang Wang[1], Songlin Lu[1],

Zhenwei Shang[1,3], Chi Yan[1], Jingyang Hu[1], Tianyu Zhou[1], Qinghua Liu[1], GRD 2025 global high BMI and

intervention Consortium[3], Mingming Zhang[1,3,*], Hongchao Lyu[1,3,*], Yongshuai Jiang[1,3,*]

[1]College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China.
[2]Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, China.

[3]Global Riskfactor of Disease (GRD) 2025，DeepBurden，China

† These authors contributed equally to this work.
* Correspondence:
Yongshuai Jiang, College of Bioinformatics Science and Technology, Harbin Medical University, 194 Xuefu Road, Nangang District, Harbin, 150081, Heilongjiang Province, China. phone numbers: +86 13258669501; E-mail: jiangyongshuai@hrbmu.edu.cn;
Hongchao Lyu, College of Bioinformatics Science and Technology, Harbin Medical University, 194 Xuefu Road, Nangang District, Harbin, Heilongjiang Province, China. E-mail: hongchaolyu@hrbmu.edu.cn;
Mingming Zhang, College of Bioinformatics Science and Technology, Harbin Medical University, 194 Xuefu Road, Nangang District, Harbin, Heilongjiang Province, China. E-mail: zhangmingming@ems.hrbmu.edu.cn.

Abstract

Background

High body mass index (BMI) presents a serious and ongoing global health challenge. However, the difficulty of high BMI intervention has not yet been systematically evaluated.

Methods

We developed a Generative Artificial Intelligence Meta-Evaluation (GAME) framework, which integrated 18 indicators from 4 dimensions, including "Macro-System Level", "Socio-Cultural

Level", "Community-Family Level", and "Individual Level" to evaluate the difficulty of high BMI intervention across 226 locations. The GAME framework applies 8 leading AI models to generate intervention difficulty scores (IDS) of each indicator on a scale from 1 to 5, with higher scores indicating greater difficulty. Meta-analysis was conducted to derive combined scores, evaluate the heterogeneity and sensitivity. Final intervention difficulty scores were calculated as the weighted sum of all 18 indicators. Additionally, SHapley Additive exPlanation (SHAP) values were used to evaluate the importance of each indicator in determining the intervention difficulty.

Results

The global difficulty of high BMI intervention shows significant imbalance. Norway (IDS = 1.48) exhibited the easiest intervention, while Yemen (IDS = 4.56) faced the greatest challenge. Regions such as Western Europe, Australasia, and High-income Asia Pacific showed lower intervention difficulty, reflecting there are mature public health frameworks, supportive social-cultural environments for healthy lifestyles, and high levels of health awareness. On the contrary, countries in North Africa and Middle East, South Asia, Oceania, and Sub-Saharan Africa faced higher intervention challenges, suggesting the need for long-term, collaborative efforts from multiple sectors. Among the 18 indicators, "Cognition and Awareness" has the most significant impact on intervention difficulty, with the SHAP value of 31.03, followed by "Family life and cognitive patterns" (18.08) and "Health Care System" (11.7). Furthermore, the IDS for high BMI was significantly correlated with Socio-Demographic Index (SDI). Higher SDI values were associated with easier interventions. Finally, the independent external empirical verification demonstrated high consistency between intervention difficulty and increase in annual prevalence of obesity, population mean BMI, and national policies. It supported the GAME framework to characterize global heterogeneity in high BMI intervention challenge. Global results were freely available at http://www.deepburden.com/high-bmi.

Conclusion

The difficulty of high BMI intervention varies widely across countries and regions, highlighting the need for comprehensive strategies and governance to address the growing health issue effectively.

Introduction

Overweight and obesity is a major health crisis, presenting a significant and ongoing threat to global health progress [1-5]. In 2021, these conditions contributed to 3.71 million deaths and 129 million disability-adjusted life-years (DALYs) [6]. In the past two decades, the global age-standardised DALY rates attributable to overweight and obesity increased by more than 15% [6]. High body mass index (BMI) has been one of the top risk factors with the steepest increase in attributable burden [7-10].

A major challenge is implementing intervention policies to curb the rising rates of overweight and obesity [11-13]. However, the World Health Organization reports that in 2021, 40% of nations had implemented an operational policy, strategy, or action plan to address overweight and obesity. But the policy coverage dropped to below 10% in low-income countries [6]. For examples, the UK and the USA implemented the taxation of sugar-sweetened beverages and New York City's ban on artificial trans fats [6, 14, 15]. On the contrary, evidence for effective obesity intervention implementation remains scarce in low- and middle-income countries [6, 16]. The difficulty of high BMI intervention shows an inequality on a global scale.

Recently, generative Artificial Intelligence (AI) has made remarkable progress [17-21]. GPT, a generative transformer model, was developed based on the extensive and multiple dataset comprising diverse online texts, such as articles, books, and websites across various subjects and languages [22-24]. DeepSeek supports continuous learning by leveraging publicly available open-source datasets, improving its adaptability to rapidly evolving knowledge and scientific reasoning [25-27]. Such AI systems hold promise for various healthcare applications, including research acceleration, clinical decision support, medical education, and so on [28-30]. This raises the question of whether AI models could be employed to integrate data from various sources, thereby accurately assessing the difficulty of high BMI intervention.

In this study, we applied 8 large AI models to evaluate the difficulty of high BMI intervention in 226 locations. We developed a Generative AI Meta-Evaluation (GAME) framework (Fig. 1) to calculate the intervention difficulty scores according to 18 key indicators from 4 dimensions: "Macro-System Level", "Socio-Cultural Level", "Community-Family Level", and "Individual Level". The influence of Socio-Demographic Index (SDI) on the difficulty of high BMI intervention was evaluated. Furthermore, we verified the reliability of results based on both

independent external empirical baselines, including increase in annual age-standardised prevalence of obesity, population mean BMI, and national policies, and internal heterogeneity and sensitivity analysis.

Results

Global, regional, and national unequal difficulty of high BMI intervention

The results revealed significant regional and national differences in the difficulty of intervening in high BMI. At the country level, the mean intervention score of high BMI was 3.15, with Norway achieving the lowest score of 1.48, and Yemen representing the highest at 4.56. Regions such as Western Europe, Australasia, High-income Asia Pacific and High-income North America exhibited lower intervention difficulty, while countries in North Africa and Middle East, South Asia, Oceania, and Sub-Saharan Africa faced higher intervention challenges (Fig. 2, Supplementary Table 1).

In Western Europe, the mean intervention score of high BMI was 1.96, with countries like Norway (1.48), Sweden (1.51), Iceland (1.56), and Denmark (1.64) showing the lowest difficulty. These countries benefit from strong national public health frameworks, a social culture that supports healthy lifestyle, and the medical system focused on prevention. For example, Norway has implemented high sugar taxes, mandatory nutrition labels, and universal prevention-oriented primary care. In addition, there is a strong population awareness of obesity risks (Fig. 3B). Australasia, which includes Australia (1.95) and New Zealand (2.01) followed closely, with developed economies and rich health resources, though processed foods, such as dairy products and processed meats, are common in the diet (Fig. 3D). High-income North America had a mean intervention score of 2.55, including Canada (2.08) and United States of America (3.02). Despite high income and high health awareness, the United States of America faces challenges due to an uneven distribution of health resources, strong lobbying by unhealthy industries, contributing to widespread addiction to high sugar and high fat foods (Fig. 3C).

On the contrary, countries in North Africa and Middle East achieved the mean intervention score of 3.45, including Kuwait (2.83), Saudi Arabia (2.92), Jordan (3.25), Egypt (3.34), etc. While high-income oil-rich countries like Kuwait and Saudi Arabia offer good private medical

care and a high standard of family life, fast food culture and the popularity of traditional sweet treats dominate. Middle-income countries such as Jordan and Egypt, with diets high in oil and sugar, have accessible medical care but lack strong preventive measures (Fig. 3D). Oceania, which has the highest age-standardized prevalence of overweight [6], had a mean intervention score of 3.71, represented by countries like Nauru (4.12), Tonga (3.74), and Samoa (3.51). These nations have middle-income economies but weak medical systems and limited infrastructure, compounded by traditional high-starch diets and imported processed foods. Public awareness of the need for weight management is also low (Fig. 3D).

Sub-Saharan Africa had the highest mean intervention score of 3.88, with countries like Zambia (3.70) and Zimbabwe (3.92) reflecting this trend. These nations have low-to-middle incomes and healthy traditional diets, but there is limited awareness of obesity risks. In urban areas, the consumption of sweetened beverages is widespread, and although obesity is not yet a major public health issue, intervention efforts are hindered once it does occur (Fig. 3E).

Main factors influencing of unequal difficulty of high BMI intervention

We evaluated the difficulty of high BMI intervention based on 18 indicators from four dimensions, namely "Macro-System Level", "Socio-Cultural Level", "Community-Family Level", and "Individual Level" (Fig. 3, Supplementary part 5, Supplementary Figure 19-36). The SHapley Additive exPlanation (SHAP) method was applied to evaluate the impact of indicators on the unequal intervention difficulty (Fig. 4A). Among them, "Cognition and Awareness" showed the most significant impact (Fig. 4A, Fig. 4B), with the SHAP value of 31.03, followed by "Family life and cognitive patterns" (18.08), "Health Care System" (11.7), "Time and Energy Costs" (11.11), and "Implementation and Supervision" (10.92).

The factors at the "Individual Level" had the greatest influence on intervention difficulty, including "Cognition and Awareness", "Time and Energy Costs", "Economic Costs", and "Behavioral Habits and Reflexes". "Cognition and Awareness" evaluates an individual's health literacy, ability to actively learn and apply scientific knowledge, and awareness of their own health risks. Countries like Iceland, Norway, and Sweden had the lowest intervention scores for "Cognition and Awareness" (1.24), while Yemen and South Sudan had the highest scores (4.63), showing a 3.73-fold difference (Fig. 3). "Time and Energy Costs" also significantly impacted the

intervention difficulty, with a SHAP value greater than 10 (Fig. 4A). This factor assesses whether individuals have enough time and energy to plan and maintain a healthy lifestyle.

The second most important factors are at the "Macro-System Level", which encompass political, economic, technological, and public service systems. Among these indicators, the "Health Care System" had the greatest impact. It evaluates whether the system can provide general and preventive weight management and nutrition guidance, as well as treat obesity-related diseases. In cases where early prevention and weight management support are lacking, or where professional resources are scarce and medical services are expensive, the intervention difficulty is significantly higher. "Implementation and Supervision" also had a considerable impact, with a SHAP value of 10.92 (Fig. 4A). This indicator assesses whether regulations are strict, enforcement is efficient, and violations (such as false health claims) are severely punished. Norway had the lowest intervention difficulty for "Implementation and Supervision" (1.24), while countries such as Afghanistan, the Central African Republic, Chad, Somalia, South Sudan, and Yemen had the highest scores (5), showing a 4.03-fold difference (Fig. 3). Another significant indicator was "Information Environment Governance", which evaluates whether the public information environment is clean, whether unhealthy food advertisements are restricted, and whether scientific health information is easily accessible.

Factors at the "Community-Family Level" also played a significant role in the difficulty of high BMI intervention, including "Family Life and Cognitive Patterns", "Family Support Environment", and "Family Economic Status". These factors assess whether families practice a healthy lifestyle, support healthy health concepts and daily routines, whether family members understand and assist with an individual's health efforts, and whether a family can afford healthy food, fitness expenses, and professional consulting services. Norway and Denmark had the easiest intervention conditions at this level, while Yemen and Somalia experienced the greatest difficulty, with nearly 3 times the difference in scores (Fig. 3).

Lastly, the factors at the "Socio-Cultural Level" had a relatively smaller impact, including "Traditional Concepts and Misconceptions", "Gender Inequality", and "Social Habits", all with SHAP values less than 5 (Fig. 4A). These factors evaluate the mainstream cultural norms, values, and social atmosphere. "Traditional Concepts and Misconceptions" assesses whether mainstream views support health, whether scientific concepts are widely disseminated, and whether there are

common dietary misconceptions. "Gender Inequality" evaluates whether both genders have equal freedom and support in food choices, exercise time, and bodily autonomy. Finally, "Social Habits" reflects whether healthy eating and regular exercise are the societal norms and mainstream social activities.

Differences in difficulty of high BMI intervention across SDI levels

The difficulty of high BMI intervention is significantly correlated with SDI (R = -0.847, p-value = 1.1×10^{-56}, Fig. 5A). A higher the SDI value was associated with lower intervention difficulty.

In the 40 locations with a high SDI, the mean intervention difficulty score is 2.18, with Norway having the lowest score and the United States of America being the only country with a score higher than 3. In the 46 locations with a high-middle SDI, the mean intervention difficulty score is 2.98, ranging from the lowest in Israel (2.15) to the highest in Libya (3.99). Notably, 54.35% of countries in this level have scores above 3. In locations with a middle SDI, the mean intervention difficulty score increases to 3.27, with 82.5% of locations having scores above 3. Nauru (4.12) and the Syrian Arab Republic (4.37) have intervention difficulty scores above 4.

In the 42 locations with a low-middle SDI, the mean intervention difficulty score rises to 3.67, with only Bhutan scoring below 3 (2.95), and 6 locations having scores above 4. The highest intervention difficulty is observed in locations with a low SDI, where the mean intervention difficulty score is 4.06, and 64.16% of these locations have scores above 4.

Independent external empirical verification demonstrated high consistency between intervention difficulty and increase in annual prevalence of obesity, population mean BMI, and national policies

To verify the reliability of results, we compared the GAME intervention difficulty scores against the established baselines, including global increase in annual population mean BMI [31], prevalence of obesity [32], and national policies.

We observed a strong alignment between GAME intervention difficulty and real-world obesity trends (Fig. 6A). Specifically, the 10 countries ranked most difficult to intervene in by the GAME framework exhibited a upward trajectory in the annual increase of age-standardised prevalence of obesity from 1990 to 2022. In contrast, the 10 countries deemed easiest to intervene

in showed a declining trend in obesity growth over the same period. Notably, from 1990 to 2002, both groups experienced rising obesity prevalence. The annual increase was even higher in the easier-to-intervene countries during this early phase. However, a marked divergence emerged from 2003 onward. The annual increase in obesity prevalence decelerated in the easier-to-intervene countries, while it continued to rise in the more difficult-to-intervene countries. Under current trends, the annual increase in easier-to-intervene countries is gradually leveling off, whereas it remains on an upward trajectory in more difficult-to-intervene countries. The trends of annual increase of age-standardised prevalence of obesity are highly consistent with GAME intervention difficulty rating.

The similarly consistent pattern was observed in annual increase of population mean BMI. From 1985 to 2000, urban populations in both easier- and more difficult-to-intervene countries exhibited relatively stable trends in BMI growth (Fig. 6B). However, a clear divergence emerged after 2000. From 2000 to 2017, the annual increase in mean BMI in urban areas declined markedly in easier-to-intervene countries, whereas the decline in more difficult-to-intervene countries was limited. This divergence extended to rural populations as well (Fig. 6C). Moreover, the annual increase in population mean BMI remained consistently higher in more difficult-to-intervene countries than in easier-to-intervene ones from 1985 to 2017, particularly in rural areas. The findings further supported the validity of the GAME framework.

Additionally, we divided countries to two groups according to the presence or absence of national guidelines or standards for the management of high BMI. A highly significant difference was observed in the "Policy Integrity" intervention difficulty scores between the two groups (T-test, p-value = 2.3×10^{-10}, Fig. 6D). The scores in countries without such guidelines were higher than those with established guidelines. Similarly, the significant difference was also founded when comparing countries with and without national guidelines or standards for the management of physical inactivity (T-test, p-value = 3×10^{-11}, Fig. 6D). The results also demonstrated that the GAME framework is reliable in assessing the high BMI intervention of different countries.

Together, these findings provided robust and multi-dimensional evidence supporting the GAME framework to characterize global heterogeneity in high BMI intervention challenge.

Internal evaluation-low heterogeneity and sensitivity of 8 AI

The evaluations of the intervention difficulty across 226 locations, using eight distinct AI models based on 18 key indicators, demonstrated consistently low heterogeneity. For 99.9% indicators, the I² values were below 20% and Cochran's Q statistic were non-significant (all $p > 0.05$) (Supplementary Table 2-19). These results indicated a high concordance among the different AI methodologies in ranking global intervention challenges, supporting the reliability of the generated difficulty scores.

Furthermore, sensitivity analyses were performed by sequentially excluding the results from each AI model. The results revealed that the final conclusions remained essentially unchanged regardless of which model was removed, underscoring the robustness of the primary findings (Fig. 5B).

Methods

The GAME framework

We evaluated the difficulty of high BMI intervention using the Generative AI Meta-Evaluation (GAME) framework, which is structured across four levels: "Macro-System Level", "Socio-Cultural Level", "Community-Family Level", and "Individual Level". Each of these levels was further divided into two sub-levels to reflect both their potential impact and measurement difficulty. After conducting a comprehensive manual review of relevant literature in alignment with these dimensions, we identified and selected 18 key indicators that influence intervention difficulty. These indicators are "Policy integrity", "Implementation and supervision", "Information environment governance", "Regional economic development level", "Industry interest resistance", "Physical infrastructure", "Health care system", "Traditional concepts and misconceptions", "Social habits", "Gender inequality", "Family economic status", "Family life and cognitive patterns", "Family support environment", "Physiological addiction", "Cognition and awareness", "Behavioral habits and reflexes", "Economic costs", and "Time and energy costs" (Supplementary part 2). Each of these indicators is scored on a scale from 1 to 5, with higher scores representing greater difficulty in implementing interventions. We assigned a weight to each indicator such that the sum of all weights equals 1 (Table 1).

We utilized 8 AI models to assess the intervention difficulty of the 18 indicators for the 226 locations. The AI models used were GPT (https://chatgpt.com/), Deepseek (https://www.deepseek.com/), Doubao (https://www.doubao.com/chat), Grok (https://grok.com/), Qwen (https://www.tongyi.com/), Gemini (https://gemini.google.com/), Claude Sonnet 4 (https://claude.ai/), and Kimi (https://www.kimi.com/) (Supplementary part 1). For each location, we submitted identical questions to all eight AI models separately, asking them to score the 18 indicators on a scale from 1 to 5 (Supplementary parts 2 and 3). The scores from each model were then integrated to produce a combined score for each indicator, using meta-analysis combined with SuperCLUE (https://www.superclueai.com/), a comprehensive evaluation system for general AI. The combined score of each indicator for every country and region is calculated by:

$$T = \frac{\sum_{i=1}^{k} W_i S_i}{\sum_{i=1}^{k} W_i}$$

of which, T is the combined score, k is the number of AI, S_i is the score produced by AI_i, W_i is the weight of AI_i from SuperCLUE:

$$W_i = V_i$$

V_i represents the score of AI_i in SuperCLUE.

We tested the heterogeneity of AI models using I^2 and Cochran's Q (Supplementary part 4). In addition, we performed a clustering analysis on the scores of the 18 indicators generated by the eight AI models (Supplementary part 4, Supplementary Figure 1-18). Sensitivity analyses were also conducted by iteratively excluding the results from each AI model.

Finally, the intervention difficulty score for each country and region was calculated as the weighted sum of all 18 indicators.

Statistical analysis

The SDI values for each location were derived from the Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) 2021 (Supplementary Table 20). The SDI is a composite measure that reflects the social and economic conditions influencing health outcomes across different regions. It is calculated as the geometric mean of three indices: the total fertility rate (TFR) for individuals

under 25 years old (TFU25), the average education level for individuals aged 15 and older (EDU15+), and per capita lag-distributed income (LDI). In the GBD 2021 study, the SDI values were scaled by multiplying them by 100, resulting in a range from 0 to 100. To measure the correlation between the intervention difficulty scores for high BMI and SDI values, we used Pearson's correlation coefficient. A significant correlation was defined by an R^2 value greater than 0.5 and a p-value less than 0.05.

We also employed SHAP values to evaluate the overall importance of various indicators in determining the difficulty of intervening in high BMI. SHAP applies a game-theoretic approach to interpret model behavior by aggregating local feature contributions into global explanations [33, 34]. This methodology is considered more informative than most global approximation techniques. In addition to quantifying the overall indicator importance, SHAP helps to clarify how each individual variable influences specific predictions.

The age-standardised prevalence of obesity from 1990 to 2022 and the population mean BMI in both urban and rural from 1985 to 2017 were obtained in NCD Risk Factor Collaboration [31, 32]. A total of 197 countries were matched to the GAME framework. The annual increase of age-standardised prevalence of obesity was calculated by subtracting the prevalence of the previous year from the current year. The same method was applied to calculate the annual increase of population mean BMI. The existence of national guidelines or standards for the management of high BMI and physical inactivity in 194 countries were obtained from the World Obesity Atlas 2025 [35]. The difference between the countries with and without the guidelines was calculated using the two-sample student T-test.

Data analysis was conducted using the R statistical software package (v4.2.1). For statistical analysis, we used the following R packages: ggbeeswarm, tidyverse, stringr, iml, and randomForest.

Discussion

Despite global consensus on the severe health threats of overweight and obesity, the prevalence has persistently increased worldwide over the past three decades, affecting all age groups and sexes. High BMI intervention is a leading global health challenge. Here, we developed a GAME framework to evaluate the difficulty of high BMI intervention by integrating 18 indicators from

the broad societal structures to the individual's inner world. Our study revealed the imbalanced intervention difficulty across 226 locations. The results were freely available at http://www.deepburden.com/high-bmi (Supplementary part 6, Supplementary Figure 37-41).

The top 5 easiest locations to intervene in were Norway [1.48], Sweden [1.51], Iceland [1.56], Denmark [1.64], and Switzerland [1.67]. Almost 3 times difference was observed between the easiest locations to intervene in, Norway, and the most difficult locations, Yemen [4.56]. In the 57 locations with an intervention difficulty score lower than 2.67 (the 25th percentile value), 56.14% were classified as High SDI level, followed by 15.79% in the High-middle SDI level, and 1.75% in the Middle SDI level (the other 26.32% had no SDI values). The Middle SDI level locations that demonstrated lower difficulty of high BMI intervention was Costa Rica [2.37]. Notably, five Middle SDI level locations exhibited higher intervention difficulty scores than 3.68 (the 75th percentile value), including Tonga [3.74], Equatorial Guinea [3.88], Iraq [3.89], Nauru [4.12], and Syrian Arab Republic [4.37]. In addition, the findings highlighted that the difficulty of intervening in high BMI is influenced by a variety of factors across different levels. The most significant challenges arise from individual awareness and resources, followed by macro-level systemic factors, while socio-cultural factors play a comparatively smaller role.

We presented the first systematic attempt to evaluate the difficulty of high BMI intervention on a global scale. It offers a novel and comprehensive perspective for understanding multi-level determinants to address the obesity crisis. The GAME framework integrates a wide range of indicators spanning macro-system, socio-cultural, community-family, and individual levels, enabling a holistic view of the factors influencing unequal intervention difficulty. Furthermore, the use of eight leading AI models enhances the robustness of the estimates. Nevertheless, our study has several limitations that should be considered. First, the findings are dependent on the complex AI systems, the interpretability of specific scores for certain locations or indicators may lack transparency. Second, we applied the annual increase of obesity prevalence to verify the reliability of the results. Because the direct evidence of success interventions of high BMI for all 226 locations is constrained, particularly in settings with limited digital infrastructure or inconsistent reporting.

Overall, we provide the first large-scale, AI-driven evaluation of the difficulty of high BMI intervention across 226 locations, offering a crucial evidence to evaluate the effectiveness of

interventions and reduce global obesity.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Not applicable.

Competing interests

The authors declare no competing interests.

Funding declaration

This work was supported by the National Natural Science Foundation of China [Grant Nos. 31970651, 92286018]; the Excellent Youth Support plan of Education Department of Heilongjiang Province [Grant No.YQJH2023036]; Marshal Initiative Funding [Grant No. HMUMIF-22010]; XingLian Outstanding Talent Support Program 2024; the Joint Funds of the Zhejiang Provincial Natural Science Foundation of China [Grant No. LBY24H170001].

Author Contributions

C.S., C.W.L., W,H,L. and W.S. contributed equally to this work. Y.S.J. conceived and contributed the work. C.S., C.W.L., W,H,L., W.S., S.Y.W., H.Y.C., J.X.T., J.C.W., Y.P.Z., L.N.Y., R.L.L., J.X., T.Y.L., Q.W., Y.X., N.W., Y.G., Q.D.R., C.W., S.L.L., C.Y., J.Y.H., T.Y.Z., Q.H.L., Z.W.S, M.M.Z., H.C.L., and Y.S.J. drafted and modified the manuscript.

Acknowledgements

Not applicable.

REFERENCE:

1.   Bluher M: **Obesity: global epidemiology and pathogenesis**. *Nature reviews Endocrinology* 2019, **15**(5):288-298.

2.   Perdomo CM, Cohen RV, Sumithran P, Clement K, Fruhbeck G: **Contemporary medical, device, and surgical therapies for obesity in adults**. *Lancet* 2023, **401**(10382):1116-1130.

3.   Barbouni K, Jotautis V, Metallinou D, Diamanti A, Orovou E, Liepinaitiene A, Nikolaidis P, Karampas G, Sarantaki A: **When Weight Matters: How Obesity Impacts Reproductive Health and Pregnancy-A Systematic Review**. *Current obesity reports* 2025, **14**(1):37.

4.   Luo Y, Blackledge WC: **Microbiome-based mechanisms hypothesized to initiate obesity-associated rheumatoid arthritis**. *Obesity reviews : an official journal of the International Association for the Study of Obesity* 2018, **19**(6):786-797.

5.   Kulkarni A, Bowers LW: **The role of immune dysfunction in obesity-associated cancer risk, progression, and metastasis**. *Cellular and molecular life sciences : CMLS* 2021, **78**(7):3423-3442.

6.   Collaborators GBDAB: **Global, regional, and national prevalence of adult overweight and obesity, 1990-2021, with forecasts to 2050: a forecasting study for the Global Burden of Disease Study 2021**. *Lancet* 2025, **405**(10481):813-838.

7.   Dai H, Alsalhe TA, Chalghaf N, Ricco M, Bragazzi NL, Wu J: **The global burden of disease attributable to high body mass index in 195 countries and territories, 1990-2017: An analysis of the Global Burden of Disease Study**. *PLoS medicine* 2020, **17**(7):e1003198.

8.   Jin J: **Interventions for High BMI in Children and Teenagers**. *Jama* 2024, **332**(3):262.

9.   Liu C, Zhang Z, Wang B, Meng T, Li C, Zhang X: **Global health impacts of high BMI: A 30-Year analysis of trends and disparities across regions and Demographics**. *Diabetes research and clinical practice* 2024, **217**:111883.

10.   Ge C, Xiong J, Zhu R, Hong Z, He Y: **The global burden of high BMI among adolescents between 1990 and 2021**. *Communications medicine* 2025, **5**(1):125.

11.   Dong Y, Yuan C, Dang J, Song X, Cheng G, Chen Y, Wang H, Mi J, Xi B, Song Y: **Control of childhood obesity and implications for policy in China**. *The Lancet Public health* 2024, **9**(12):e1125-e1135.

12.   Damschroder LJ, Lowery JC: **Evaluation of a large-scale weight management program using the consolidated framework for implementation research (CFIR)**. *Implementation science : IS* 2013, **8**:51.

13.   Adom T, De Villiers A, Puoane T, Kengne AP: **A Scoping Review of Policies Related to the Prevention and Control of Overweight and Obesity in Africa**. *Nutrients* 2021, **13**(11).

14.   Collaborators GBDAB: **Global, regional, and national prevalence of child and adolescent overweight and obesity, 1990-2021, with forecasts to 2050: a forecasting study for the Global Burden of Disease Study 2021**. *Lancet* 2025, **405**(10481):785-812.

15.   Collaborators GUOF: **National-level and state-level prevalence of overweight and obesity among children, adolescents, and adults in the USA, 1990-2021, and forecasts up to 2050**. *Lancet* 2024, **404**(10469):2278-2298.

16.   Brown T, Moore TH, Hooper L, Gao Y, Zayegh A, Ijaz S, Elwenspoek M, Foxen SC, Magee L, O'Malley C *et al*: **Interventions for preventing obesity in children**. *The Cochrane database of systematic reviews* 2019, **7**(7):CD001871.

17. Xie WJ, Warshel A: **Harnessing generative AI to decode enzyme catalysis and evolution for enhanced engineering**. *National science review* 2023, **10**(12):nwad331.

18. Wang J, Wang YX, Zeng D, Zhu Z, Li D, Liu Y, Sheng B, Grzybowski A, Wong TY: **Artificial intelligence-enhanced retinal imaging as a biomarker for systemic diseases**. *Theranostics* 2025, **15**(8):3223-3233.

19. Reddy S: **Generative AI in healthcare: an implementation science informed translational path on application, integration and governance**. *Implementation science : IS* 2024, **19**(1):27.

20. Lu MY, Chen B, Williamson DFK, Chen RJ, Zhao M, Chow AK, Ikemura K, Kim A, Pouli D, Patel A *et al*: **A multimodal generative AI copilot for human pathology**. *Nature* 2024, **634**(8033):466-473.

21. Umer F, Adnan N: **Generative artificial intelligence: synthetic datasets in dentistry**. *BDJ open* 2024, **10**(1):13.

22. Hoppe JM, Auer MK, Struven A, Massberg S, Stremmel C: **ChatGPT With GPT-4 Outperforms Emergency Department Physicians in Diagnostic Accuracy: Retrospective Analysis**. *Journal of medical Internet research* 2024, **26**:e56110.

23. Rathje S, Mirea DM, Sucholutsky I, Marjieh R, Robertson CE, Van Bavel JJ: **GPT is an effective tool for multilingual psychological text analysis**. *Proceedings of the National Academy of Sciences of the United States of America* 2024, **121**(34):e2308950121.

24. Noda R, Tanabe K, Ichikawa D, Shibagaki Y: **GPT-4's performance in supporting physician decision-making in nephrology multiple-choice questions**. *Scientific reports* 2025, **15**(1):15439.

25. Temsah A, Alhasan K, Altamimi I, Jamal A, Al-Eyadhy A, Malki KH, Temsah MH: **DeepSeek in Healthcare: Revealing Opportunities and Steering Challenges of a New Open-Source Artificial Intelligence Frontier**. *Cureus* 2025, **17**(2):e79221.

26. Tordjman M, Liu Z, Yuce M, Fauveau V, Mei Y, Hadjadj J, Bolger I, Almansour H, Horst C, Parihar AS *et al*: **Comparative benchmarking of the DeepSeek large language model on medical tasks and clinical reasoning**. *Nature medicine* 2025, **31**(8):2550-2555.

27. Zeng D, Qin Y, Sheng B, Wong TY: **DeepSeek's "Low-Cost" Adoption Across China's Hospital Systems: Too Fast, Too Soon?** *Jama* 2025, **333**(21):1866-1869.

28. Teo ZL, Thirunavukarasu AJ, Elangovan K, Cheng H, Moova P, Soetikno B, Nielsen C, Pollreisz A, Ting DSJ, Morris RJT *et al*: **Generative artificial intelligence in medicine**. *Nature medicine* 2025, **31**(10):3270-3282.

29. Turk M, Zeydan E, Arslan SS, Turk Y: **The Future Role of Generative Artificial Intelligence (AI) in Medicine**. *Journal of clinical practice and research* 2024, **46**(3):211-213.

30. Preiksaitis C, Rose C: **Opportunities, Challenges, and Future Directions of Generative Artificial Intelligence in Medical Education: Scoping Review**. *JMIR medical education* 2023, **9**:e48785.

31. Collaboration NCDRF: **Rising rural body-mass index is the main driver of the global obesity epidemic in adults**. *Nature* 2019, **569**(7755):260-264.

32. Collaboration NCDRF: **Worldwide trends in underweight and obesity from 1990 to 2022: a pooled analysis of 3663 population-representative studies with 222 million children, adolescents, and adults**. *Lancet* 2024, **403**(10431):1027-1050.

33. Li J, Liu S, Hu Y, Zhu L, Mao Y, Liu J: **Predicting Mortality in Intensive Care Unit Patients With Heart Failure Using an Interpretable Machine Learning Model: Retrospective Cohort Study**. *Journal of medical Internet research* 2022, **24**(8):e38082.

34. Xiao Y, Zhao Z, Su CG, Li J, Liu J: **An interpretable machine learning model for predicting depression in middle-aged and elderly cancer patients in China: a study based on the CHARLS cohort**. *BMC psychiatry* 2025, **25**(1):610.

35. World Obesity Federation, **World Obesity Atlas 2025. Overweight, obesity and non-communicable diseases: New global, regional and national estimates of the prevalence of overweight and obesity in adults from 2000 to 2030 March**, 2025 Available from: https://s3-eu-west-1.amazonaws.com/wof-files/World_Obesity_Atlas_2025_rev1.pdf

Generative AI Meta-Evaluation (GAME) Study framework

Figure 1. Overview of the GAME framework.

Figure 2. Global difficulty of high BMI intervention.

B

Europe

	Macro-System							Socio-Cultural			Community-Family			Individual					Dimension
level	Policy Integrity	Implementation and Supervision	Information Environment Governance	Regional Economic Development Level	Industry Interest Resistance	Physical Infrastructure	Health Care System	Traditional Concepts and Misconceptions	Gender Inequality	Social Habits	Family Economic Status	Family Life and Cognitive Patterns	Family Support Environment	Physiological Addiction	Cognition and Awareness	Behavioral Habits and Reflexes	Economic Costs	Time and Energy Costs	Dimension
	1.88	1.88	1.88	1.50	2.01	1.62	1.63	2.15	1.50	2.15	1.50	1.88	1.64	2.25	1.75	2.02	1.50	1.88	Andorra
	1.64	1.76	1.76	1.39	2.65	1.49	1.89	2.00	1.65	2.53	1.77	2.12	2.13	2.64	1.76	2.43	1.64	2.28	Australia
	1.88	1.88	1.88	1.50	2.43	1.62	1.74	2.38	1.64	2.38	1.50	2.11	1.85	2.38	1.75	2.15	1.62	1.88	Austria
	2.01	1.88	1.88	1.50	2.66	1.62	1.74	2.25	1.50	2.15	1.50	1.97	1.85	2.39	1.75	2.15	1.62	1.88	Belgium
	2.24	2.24	2.38	2.01	2.52	2.01	2.15	2.65	2.01	2.78	2.28	2.51	2.24	2.78	2.24	2.78	2.28	2.38	Cyprus
	2.38	2.38	2.38	2.15	2.93	2.15	2.15	2.65	2.15	3.04	2.15	2.51	2.38	3.04	2.24	2.78	2.15	2.28	Czechia
	1.50	1.50	1.64	1.39	1.90	1.50	1.37	2.04	1.25	2.04	1.39	1.76	1.64	2.13	1.50	2.04	1.50	1.64	Denmark
	2.15	2.15	2.15	2.04	2.68	1.90	2.04	2.39	1.90	2.65	2.04	2.51	2.25	2.65	1.90	2.56	2.15	2.15	Estonia
	1.37	1.62	1.50	1.26	2.02	1.50	1.50	2.02	1.37	2.15	1.64	1.76	1.64	2.25	1.50	2.02	1.64	1.75	Finland
	1.25	1.25	1.39	1.13	2.27	1.37	1.25	2.00	1.53	2.31	1.65	2.26	1.88	2.54	1.50	2.31	1.65	1.65	France
	1.89	1.89	1.89	1.38	2.93	1.62	1.62	2.38	1.64	2.42	1.50	2.23	1.85	2.52	1.75	2.28	1.62	1.75	Germany
	1.51	1.49	1.38	1.26	1.91	1.38	1.24	1.78	1.13	1.91	1.65	1.77	1.51	2.01	1.24	1.91	1.78	1.51	Iceland
	1.89	2.01	1.89	1.50	2.65	1.88	1.88	2.38	1.76	2.42	1.50	2.23	1.97	2.38	1.88	2.28	1.62	1.88	Ireland
	1.90	2.02	1.77	1.89	2.93	1.75	2.02	2.53	2.15	2.42	1.89	2.23	2.11	2.52	1.89	2.42	1.89	2.02	Israel
	2.03	2.16	2.16	2.15	2.93	2.15	2.03	2.65	2.15	2.84	2.29	2.65	2.65	2.65	2.15	2.56	2.29	2.42	Italy
	2.41	2.54	2.41	2.68	2.78	2.41	2.54	2.78	2.41	2.78	2.68	2.64	2.50	2.78	2.54	2.68	2.68	2.54	Latvia
	2.41	2.41	2.41	2.27	2.78	2.27	2.37	2.78	2.27	2.78	2.27	2.50	2.37	2.65	2.27	2.55	2.27	2.41	Lithuania
	1.63	1.63	1.63	1.38	2.39	1.49	1.49	2.12	1.38	2.16	1.38	1.97	1.72	2.39	1.63	2.02	1.49	1.63	Luxembourg
	2.15	2.28	2.28	2.15	2.79	2.42	2.28	2.63	2.15	2.93	2.15	2.63	2.38	2.65	2.26	2.80	2.28	2.28	Malta
	1.75	1.62	1.62	1.24	2.23	1.49	1.50	2.25	1.24	2.28	1.24	1.97	1.73	2.38	1.62	2.15	1.37	1.62	Monaco
	1.64	1.75	1.75	1.26	2.42	1.50	1.50	2.25	1.38	2.28	1.38	2.10	1.85	2.39	1.75	2.28	1.49	1.75	Netherlands
	1.79	1.90	1.90	1.40	2.70	1.77	2.04	2.13	1.53	2.43	1.79	2.26	2.13	2.65	1.90	2.43	1.79	2.02	New Zealand
	1.24	1.24	1.38	1.26	1.78	1.38	1.24	1.78	1.13	1.78	1.39	1.73	1.61	1.87	1.24	1.78	1.51	1.51	Norway
	2.50	2.50	2.50	2.37	3.05	2.37	2.37	3.00	2.37	3.00	2.37	2.72	2.47	2.87	2.49	2.77	2.37	2.63	Poland
	1.90	2.15	2.15	2.15	2.65	2.15	2.15	2.52	2.15	2.69	2.15	2.38	2.38	2.79	2.15	2.56	2.15	2.15	Portugal
	2.10	2.10	2.10	1.63	2.39	1.88	1.75	2.38	1.63	2.42	1.63	2.23	1.86	2.38	1.98	2.38	1.63	2.10	San Marino
	2.04	2.15	2.15	2.02	2.52	1.89	2.02	2.38	2.02	2.42	2.02	2.38	2.25	2.38	2.02	2.28	2.02	2.15	Slovenia
	1.89	2.02	1.89	2.03	2.65	1.87	1.89	2.38	2.15	2.55	2.15	2.38	2.38	2.79	2.15	2.42	2.15	2.28	Spain
	1.38	1.38	1.38	1.26	1.78	1.38	1.24	1.78	1.13	1.78	1.51	1.73	1.61	1.87	1.24	1.78	1.51	1.51	Sweden
	1.49	1.36	1.49	1.24	2.39	1.49	1.36	2.12	1.24	2.16	1.38	1.97	1.72	2.25	1.36	2.02	1.49	1.63	Switzerland
	1.77	2.15	2.17	1.92	3.05	2.15	2.29	2.53	2.04	2.55	2.05	2.38	2.38	2.93	2.29	2.55	2.16	2.29	United Kingdom
	3.28	3.26	3.14	3.26	2.86	3.26	3.14	3.14	3.01	3.26	3.26	3.26	2.88	3.12	3.26	3.26	3.26	3.26	Albania
	3.14	3.26	3.14	3.00	3.13	2.87	3.14	3.13	2.77	3.13	3.00	3.00	2.88	2.87	2.87	3.00	3.00	3.00	Belarus
	3.18	3.31	3.18	3.04	3.04	3.04	3.17	3.28	3.05	3.28	3.04	3.28	3.15	3.00	3.18	3.18	3.04	3.18	Bosnia and Herzegovina
	3.05	3.05	3.03	3.03	2.89	3.05	3.05	3.26	2.91	3.39	3.16	3.26	3.02	3.12	3.16	3.16	3.16	3.16	Bulgaria
	2.64	2.77	2.77	2.64	3.04	2.49	2.63	3.00	2.64	3.13	2.64	3.00	2.74	3.00	2.64	2.90	2.64	2.77	Croatia
	2.25	2.64	2.65	2.65	3.18	2.50	2.64	2.88	2.50	3.15	2.78	2.88	2.88	3.15	2.78	2.92	2.78	2.78	Greece
	2.65	2.65	2.65	2.37	3.18	2.37	2.65	2.99	2.51	3.38	2.37	3.12	2.74	3.24	2.63	3.15	2.37	2.65	Hungary
	2.90	3.02	2.90	2.63	2.90	2.63	2.76	3.12	2.77	3.12	2.90	3.12	2.88	3.00	3.02	3.02	2.90	2.90	Montenegro
	3.04	3.02	2.90	2.90	3.04	2.90	2.90	3.12	2.77	3.25	2.90	3.12	2.88	3.00	3.02	3.02	2.90	2.90	North Macedonia
	3.39	3.64	3.53	3.64	3.24	3.64	3.64	3.39	3.30	3.52	3.64	3.64	3.40	3.11	3.64	3.64	3.64	3.64	Republic of Moldova
	3.05	3.05	3.05	2.77	3.04	2.77	3.05	3.14	2.91	3.28	2.90	3.14	2.88	3.00	3.05	3.05	3.05	3.05	Romania
	3.00	3.13	3.27	3.00	3.14	2.87	3.13	3.13	2.90	3.40	3.13	3.26	2.88	3.40	3.13	3.40	3.13	3.13	Russian Federation
	2.90	2.90	2.90	2.90	3.04	2.90	3.04	3.28	2.90	3.40	3.04	3.26	3.01	3.13	2.90	3.16	3.04	3.04	Serbia
	2.67	2.67	2.67	2.54	2.90	2.54	2.54	2.90	2.40	3.04	2.54	2.90	2.65	2.77	2.54	2.81	2.54	2.67	Slovakia
	3.80	3.92	3.68	3.78	3.52	3.78	3.94	3.39	3.30	3.39	3.91	3.52	3.40	3.00	3.65	3.52	3.78	3.78	Ukraine

level   ■ Medium Difficulty   ■ Low Difficulty

C Latin America, North America, and Caribbean

D   **North Africa and Middle East, Australasia, and Oceania**

level	Policy Integrity	Implementation and Supervision	Information Environment Governance	Regional Economic Development Level	Industry Interest Resistance	Physical Infrastructure	Health Care System	Traditional Concepts and Misconceptions	Gender Inequality	Social Habits	Family Economic Status	Family Life and Cognitive Patterns	Family Support Environment	Physiological Addiction	Cognition and Awareness	Behavioral Habits and Reflexes	Economic Costs	Time and Energy Costs	Dimension
	Micro-System							Socio-Cultural			Community-Family			Individual					Dimension
Low	1.64	1.76	1.76	1.39	2.65	1.49	1.89	2.00	1.65	2.53	1.77	2.12	2.13	2.64	1.76	2.43	1.64	2.28	Australia
Low	1.79	1.90	1.90	1.40	2.70	1.77	2.04	2.13	1.53	2.43	1.79	2.26	2.13	2.65	1.90	2.43	1.79	2.02	New Zealand
Low	2.37	2.50	2.63	1.48	3.17	2.49	1.87	3.04	2.90	3.40	1.73	3.03	2.52	3.27	2.75	3.31	1.73	2.37	United Arab Emirates
Medium	3.48	3.48	3.35	3.35	3.35	3.35	3.48	3.36	3.48	3.50	3.35	3.36	3.11	3.40	3.36	3.50	3.35	3.35	Algeria
Medium	2.72	2.86	2.99	1.96	3.26	2.89	2.49	3.12	2.87	3.62	2.21	3.11	2.73	3.25	2.97	3.38	2.21	2.71	Bahrain
Medium	3.28	3.14	3.27	3.26	3.26	3.26	3.14	3.40	3.53	3.64	3.39	3.26	3.14	3.40	3.26	3.55	3.39	3.26	Egypt
Medium	3.13	3.38	3.64	3.16	3.76	3.03	3.29	3.52	3.03	3.78	3.16	3.52	3.26	3.66	3.38	3.66	3.16	3.26	Fiji
Medium	3.26	3.38	3.39	3.51	3.26	3.38	3.39	3.52	3.53	3.75	3.51	3.51	2.87	3.38	3.51	3.65	3.51	3.51	Iran (Islamic Republic of)
Medium	2.74	3.14	3.01	3.13	3.27	3.26	3.01	3.40	3.53	3.63	3.26	3.13	3.54	3.26	3.26	3.54	3.26	3.26	Jordan
Medium	2.90	2.90	2.90	1.74	3.30	2.75	2.27	3.17	2.89	3.66	2.12	3.02	2.78	3.53	2.89	3.43	2.12	2.75	Kuwait
Medium	3.66	3.80	3.66	3.94	3.66	3.80	3.80	3.39	3.53	3.53	3.94	3.53	3.28	3.26	3.66	3.53	3.94	3.80	Lebanon
Medium	3.26	3.13	3.13	3.13	3.13	3.13	3.26	3.40	3.38	3.50	3.13	3.26	3.14	3.26	3.26	3.40	3.13	3.13	Morocco
Medium	3.01	3.40	3.14	3.14	3.15	3.00	3.26	3.54	2.52	3.66	3.54	3.14	3.40	3.02	3.54	3.13	3.54	3.40	Niue
Medium	2.63	2.76	2.90	1.83	3.05	2.63	2.23	3.02	3.03	3.53	1.99	3.01	2.77	3.39	2.88	3.43	1.99	2.61	Oman
Medium	3.12	3.12	3.25	3.15	3.25	3.28	3.40	3.77	2.90	3.89	3.28	3.63	3.13	3.77	3.50	3.77	3.40	3.50	Palau
Medium	2.49	2.76	2.90	1.61	3.17	2.75	1.99	3.24	2.90	3.78	1.86	3.26	2.64	3.38	3.00	3.42	1.86	2.74	Qatar
Medium	3.39	3.38	3.39	3.16	3.38	3.03	3.29	4.04	3.03	4.30	3.28	3.63	3.39	3.91	3.50	4.05	3.40	3.50	Samoa
Medium	2.62	2.75	3.02	1.86	3.29	2.87	2.12	3.42	3.54	3.65	2.11	3.28	3.04	3.52	3.00	3.56	2.11	2.88	Saudi Arabia
Medium	3.49	3.74	3.61	3.64	3.34	3.64	3.74	3.61	3.39	4.01	3.64	3.74	3.35	3.65	3.61	3.88	3.76	3.86	Solomon Islands
Medium	3.14	3.01	3.01	3.13	3.13	3.13	3.01	3.40	3.39	3.50	3.26	3.13	3.40	3.26	3.13	3.40	3.13	3.13	Tunisia
Medium	3.04	3.04	3.04	2.89	3.17	2.63	2.76	3.38	2.90	3.65	3.03	3.38	3.14	3.65	3.15	3.69	3.14	3.29	Turkey
Medium	3.39	3.77	3.77	3.67	3.64	3.53	3.67	3.64	3.42	3.78	3.67	3.77	3.38	3.91	3.64	3.91	3.79	3.89	Vanuatu
High	4.88	5.00	4.88	5.00	3.31	5.00	4.88	4.22	4.50	4.36	3.47	4.50	3.59	3.47	4.50	4.36	5.00	5.00	Afghanistan
High	3.99	4.12	4.01	4.03	3.74	4.16	4.01	3.74	3.78	3.88	4.03	4.00	3.50	3.38	3.87	4.00	4.03	4.26	Iraq
High	3.88	4.00	3.88	4.05	3.73	3.91	4.13	4.00	3.40	4.27	4.05	4.13	3.50	3.91	4.00	4.28	4.17	4.13	Kiribati
High	4.26	4.37	4.12	4.03	4.00	4.14	4.26	3.88	3.78	3.88	4.03	4.00	3.50	3.54	4.00	4.00	4.03	4.26	Libya
High	3.75	4.00	4.13	3.91	3.73	3.91	4.13	4.13	3.40	4.26	3.91	4.13	3.48	4.05	4.13	4.13	4.04	4.00	Marshall Islands
High	3.77	3.75	3.88	3.78	3.61	3.78	3.88	4.00	3.53	4.27	3.78	4.13	3.50	4.19	4.00	4.28	4.04	3.87	Micronesia (Federated States of)
High	4.00	4.12	4.27	4.16	3.86	4.16	4.38	4.38	3.26	4.53	4.04	4.38	3.50	4.44	4.25	4.53	4.16	4.25	Nauru
High	3.48	3.86	3.73	4.11	3.50	4.11	4.00	3.60	3.64	3.88	3.65	3.86	3.35	3.98	3.86	4.00	3.98	3.98	Papua New Guinea
High	4.62	4.87	4.62	4.76	3.72	4.87	4.76	3.97	4.25	3.97	4.76	4.10	3.86	3.39	4.23	4.10	4.76	4.76	Sudan
High	4.76	4.87	4.76	4.87	3.72	4.87	4.76	3.97	4.15	3.97	4.78	4.24	3.86	3.50	4.37	4.10	4.87	4.87	Syrian Arab Republic
High	3.39	3.51	3.64	3.54	3.64	3.54	3.55	4.29	4.02	3.52	3.54	4.02	3.76	4.16	3.75	4.30	3.79	3.88	Tonga
High	3.88	4.00	3.88	4.17	3.73	4.17	4.13	4.13	3.40	4.40	4.17	4.13	3.76	4.18	4.00	4.28	4.29	4.25	Tuvalu
High	4.88	5.00	4.88	5.00	3.67	5.00	4.88	4.10	4.63	4.38	5.00	4.63	3.71	3.76	4.63	4.50	5.00	5.00	Yemen

level ■ High Difficulty   ■ Medium Difficulty   ■ Low Difficulty

E Sub−Saharan Africa

level ■ High Difficulty ■ Medium Difficulty

F

Figure 3. Intervention difficulty scores of high BMI for 18 indicators of countries in (A) Asia, (B) Europe, (C) Latin America, North America, and Caribbean, (D) North Africa and Middle East, Australasia, and Oceania, (E) Sub−Saharan Africa, and (F) others.

Figure 4. The impact of 18 indicators on unequal difficulty of high BMI intervention . (A) SHAP summary plot of the 18 indicators. (B) The highest unequal indicators, Cognition and Awareness. (C)

The lowest unequal indicators, Industry Interest Resistance.

Figure 5. (A) The correlation between intervention difficulty scores of high BMI and SDI values. (B) The sensitivity analyses iteratively excluding each AI model and overall results.

Figure 6. Independent external empirical verification between intervention difficulty scores of high BMI and (A) increase in annual prevalence of obesity, (B) population mean BMI in urban, (C) population mean BMI in rural, and (D) national policies for the management of high BMI and physical inactivity.

Table 1. The 18 indicators in GAME framework.

Level 1	Level 2	Indicators	Weights
Macro-System Level	Policy and Governance	Policy integrity	0.0064
		Implementation and supervision	0.0064
		Information environment governance	0.0032
	Economic and Industry	Regional economic development level	0.084
		Industry interest resistance	0.056
	Infrastructure and Public Services	Physical infrastructure	0.04
		Health care system	0.06
Socio-Cultural Level	Culture and Beliefs	Traditional concepts and misconceptions	0.09
	Social customs and norms	Gender inequality	0.08
		Social habits	0.08
Community -Family Level		Family economic status	0.08
		Family life and cognitive patterns	0.06
		Family support environment	0.06
Individual Level	Biology and Cognition	Physiological addiction	0.045
		Cognition and awareness	0.045
	Behavior and Cost	Behavioral habits and reflexes	0.024
		Economic costs	0.018
		Time and energy costs	0.018