nature portfolio

Peer Review File

Epidemiology and future trend predictions of Ischemic stroke based on the Global Burden of Disease Study 1990-2021

Corresponding Author: Dr Jiayu Liu

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This report is among the most comprehensive and updated recent reports on stroke epidemiology world-wide. It is based on the 2021 GBD group data.

The title 'future trend predictions' could be shortened to 'trend predictions' or perhaps even 'trends'.

It is remarkable that the authors prefer to relate to the sociodemographic index rather than to the World Bank income group classification. Perhaps they can refer to any study showing that the SDI is more concise or informative?

It is important that the authors refer to the UN statement that shows stroke a priority target for reduction and that- following the trends- the SDG goals 2030 will not be reached by 2030. Perhaps they should put this already in the abstract.

While measures for prevention should be based on epidemiological data and their trends, the recommendations given are isolated and not systematic. The authors should refer to guideline papers from leading scientific organisations (eg Int J Stroke 2023; 18: 499-531)

A larger part of the discussion compares this paper to a recently published similar paper from the Shanghai group (Li et al. 2024). The arguments made are somewhat lengthy, even redundant. This part could be greatly reduced to perhaps a few sentences or one smaller paragraph.

Line 613: the authors recommend a more detailed reporting of stroke according to the etiology. They should also comment about the likely uncertainties that come with it.

Line 647: please explain 'chronic senile diseases'

Reviewer #2

(Remarks to the Author)

Reviewer #3

(Remarks to the Author)
This paper analyses GBD data in standard GBD methodology and provided useful observations.
1. It is important to discuss generally the quality of stroke incidence/prevalence data in different settings and the biases of using other country data where data are sparse in countries.
2. How does this manuscript relate to the Stroke GBD group who have produced stroke specific estimates and projections in the future. i.e what does the paper specifically add?
3. the paper is very data heavy and for clinicians, policy makers and stroke survivors does require attention to main questions and messages in less technical language.

4.Regarding future projections it would be useful to have more on the assumptions made regarding incidence, risk factors and outcomes and effects of acute care.
5. the Figures are quite detailed and their interpretation in the text could be more detailed. the use of abbreviations needs to be spelt out too.

Charles Wolfe

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The rebuttal letter is unusally long and contains a lot of information that has not been requested. The reviewer is overwhelmed by the the responses that have about the same length as the manuscript itself. I do not know how to respond to this response. I kindly ask the authors to respond to the requests in a short and concise way, possibly in one or two sentences. I can then make a decision based on clear and concise answers.

Reviewer #2

(Remarks to the Author)
Editorial Recommendation:
After a thorough review of the updated manuscript, we have observed that the authors have successfully integrated the previously requested modifications and clarifications, thereby significantly enhancing the robustness and transparency of the study. The following key points demonstrate compliance with our requirements:
1. Methodological Clarity and Data Handling:
The manuscript now provides a comprehensive explanation of the methods used for data imputation and managing missing data, particularly in resource-limited regions, which reinforces the robustness of the analysis.
2. Selection and Justification of Risk Factors: The discussion regarding the selection of primary risk factors has been expanded to include a detailed justification for excluding other potential factors, thereby situating the findings within a rigorous comparative framework.
3. Impact of COVID-19: The influence of the COVID-19 pandemic on the incidence and patterns of risk factors associated with cerebrovascular disease has been clearly addressed, adding valuable context to the study.
4. Compliance with Ethical and Publication Standards:
5. Statistical Analysis and Demographic Subgroup Analysis: The application of the GBD 2021 models (e.g., CODEm, DisMod-MR 2.1, ST-GPR) is presented transparently. Moreover, the enhanced stratification by sex, age, and sociodemographic index, along with comparative analyses across subpopulations, substantially strengthens the study's conclusions.

Considering that these modifications and improvements fully address the concerns raised in the previous review, we conclude that the updated manuscript meets the requirements specified by our review panel. This acceptance is based on the increased methodological rigor, the transparent presentation of results, and the added value provided by the discussion of critical aspects—such as the validation of predictive models and the impact of COVID-19.

Reviewer #3

(Remarks to the Author)
Thank you for revising the paper and addressing the questions comprehensively.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
None

Dear Editors and reviewers,

Thank you very much for taking time out of your busy schedule to review our manuscript and give us your very professional guidance, which helped our manuscript to be more perfect and rigorous. Based on the opinions of the editors and reviewers, we have carefully made supplementary modifications, hoping that the reviewers and editors can recognize our research, and we are looking forward to receiving the news that the article is accepted for publication as soon as possible.

We have refined our manuscript as suggested by you and the reviewer. First, as requested by your journal, we added the World Bank income group classification analysis and SDI analysis to more comprehensively assess the burden of ischemic stroke from the perspective of economic income and social development level. Second, as requested by your Journal, we have added systematic recommendations for the treatment and prevention of ischemic stroke from leading scientific organizations to integrate with our updated recommendations. Thirdly, we added ARIMA analysis in our study as a supplement to BAPC prediction analysis, and used different models to verify our conclusions and analyze the trend of ischemic stroke. Fourthly, stratified analysis of risk factors according to age subgroups was carried out in our paper to explain the distribution of high-risk risk factors for stroke in different age groups. Since there is no clear division and definition of rural and urban areas in the GBD database, and there is no precedent of classification analysis according to rural and urban areas in all previous GBD studies, this stratification is currently not feasible in terms of data sources and technology. We hope that the GBD database can add regional stratification for rural and urban areas in the future. Fifth, due to the space limitation of the main text, we have added detailed supplements about data sources, data analysis, data visualization and other contents in the supplementary materials section, so as to facilitate readers to better understand and repeat our results.

Reviewer 1#: We are very grateful to the expert for giving us very helpful and instructive valuable suggestions. We have carefully revised the manuscript according to your requirements! We hope that you will support the publication of our research and once again express our gratitude from the bottom of our hearts.

1. “The title ‘future trend predictions’ could be shortened to ‘trend predictions’ or perhaps even ‘trends’.”

Changing the future trend forecast to prediction will make our title more concise, which greatly improves the simplicity of our article, thank you very much for your suggestion!

Epidemiology and trends of Ischemic stroke based on the Global Burden of Disease Study 2021: a systematic analysis for the global, regional, and national burden of Ischemic stroke and its risk factors,1990-2021

2. “It is remarkable that the authors prefer to relate to the sociodemographic index rather than to the World Bank income group classification. Perhaps they can refer to any study showing that the SDI is more concise or informative?”

SDI classification and World Bank income classification have their own advantages, but at present

SDI is more used in GBD related research because of its comprehensivism. We have also supplemented the World Bank revenue disaggregation analysis to hopefully make our results more comprehensive.

The Sociodemographic Index (SDI) has the following advantages over the World Bank's income group classification that have made it more popular in some studies:

Simplicity

Composite indicator: The SDI is a composite indicator derived from a comprehensive assessment of data including the overall fertility rate of women under 25 years of age, the average education level of women 15 years of age and older, and per capita income. This comprehensive nature allows SDI to concisely reflect the state of socio-demographic development in a country or region without the need to consider multiple different economic indicators separately.

Easy to understand and apply: Since SDI integrates multiple related factors into a single indicator, researchers and policy makers can more intuitively understand and apply this indicator for comparison and analysis across countries or regions.

Richness of information

Multidimensional representation of development: SDI takes into account not only economic factors (such as per capita income), but also social factors such as demographics (such as fertility) and education levels. This multi-dimensional synthesis enables SDI to reflect the social and demographic development of a country or region more comprehensively, rather than just the economic level.

Association with health and disease burden: Studies have shown significant associations between SDI and health indicators such as healthy life expectancy (HALE) and disability-adjusted life year (DALY) rate. With the increase of SDI, the HALE increased, while the YLL and DALY rates decreased. This indicates that SDI can better reveal the impact of sociodemographic factors on health and disease burden, and provide richer information support for public health decision-making.

Revealing the socio-demographic characteristics of different diseases: The socio-demographic characteristics of different diseases (such as cardiovascular disease and cancer) are quite different. SDI can help researchers to analyze the relationship between these diseases and socio-demographic factors. For example, YLD rates increase as SDI increases in high SDI countries, which may be related to population aging and higher burden of chronic diseases in these countries.

In conclusion, the simplicity and information richness of SDI as a comprehensive socio-demographic indicator make it more advantageous in some studies, which can provide stronger support for the analysis and prediction of disease burden, health status, and the formulation of relevant policies.

World Bank Income Group Classification Analysis

The World Bank classifies countries into four income groups based on their Gross National Income (GNI) per capita: low income, lower-middle income, upper-middle income, and high income. This classification is updated annually on July 1, using the GNI per capita of the previous calendar year. The GNI is measured in US dollars using the Atlas method, which adjusts for exchange rate fluctuations over time.

Historical Trends

Low-Income Countries: In 1987, 30% of reporting countries were classified as low income. By 2023,

this proportion had decreased to 12%. For example, in South Asia, all countries were low income in 1987, but by 2023, only 13% remained in this category.
High-Income Countries: Conversely, the proportion of high-income countries has increased from 25% in 1987 to 40% in 2023. In Latin America and the Caribbean, the share of high-income countries rose from 9% in 1987 to 44% in 2023.
Regional Variations
Sub-Saharan Africa: The proportion of low-income countries in this region decreased from 74% in 1987 to 46% in 2022.
East Asia and Pacific: This region saw a significant shift, with the proportion of low-income countries dropping from 26% in 1987 to just 3% in 2022.
Middle East and North Africa: This region experienced an increase in low-income countries from 0% in 1987 to 10% in 2023.

SDI (Socioeconomic Development Index) Analysis

The SDI is a comprehensive measure that considers various socioeconomic factors beyond just income. It typically includes indicators such as education levels, healthcare access, and quality of life. While the World Bank's income classification primarily focuses on economic capacity, the SDI provides a broader perspective on development.
Importance of SDI
Holistic Development: The SDI helps in understanding the overall quality of life and the well-being of a country's population, which is not always reflected in income alone.
Policy Implications: By considering factors like education and healthcare, policymakers can identify areas that need improvement beyond economic growth, leading to more balanced development strategies.
In summary, while the World Bank's income group classification provides a useful framework for understanding economic development, the SDI offers a more comprehensive view by incorporating additional socioeconomic indicators. This dual approach allows for a more nuanced understanding of a country's development status and helps in formulating effective policies.

Socioeconomic Development Index (SDI) vs. World Bank Income Group Classification

World Bank Income Group Classification
Definition and Purpose: The World Bank classifies countries into four income groups: low income, lower-middle income, upper-middle income, and high income. This classification is based on Gross National Income (GNI) per capita, using the Atlas method to smooth exchange rate fluctuations.
Advantages:
Simplicity and Clarity: It provides a straightforward and easily understandable categorization that is useful for economic analysis and policy-making.
Economic Focus: It reflects a country's economic capacity and ability to generate income, which is crucial for assessing economic development and potential for growth.
Limitations:
Limited Scope: It focuses primarily on income levels and does not account for other important aspects of development such as education, healthcare, or environmental sustainability.
Potential Misinterpretation: High income does not necessarily equate to high quality of life or sustainable development, as it does not consider factors like income inequality or environmental degradation.

Socioeconomic Development Index (SDI)

Definition and Purpose: The SDI is a comprehensive measure that evaluates a country's development by considering various socioeconomic factors. It typically includes indicators related to economic growth, social well-being, and quality of life.

Advantages:

Holistic Approach: It considers a wide range of factors beyond income, such as education, healthcare, poverty reduction, and quality of life, providing a more comprehensive view of a country's development.

Long-term Perspective: By including indicators related to social well-being, the SDI encourages policies that not only promote current economic growth but also improve the overall quality of life.

Limitations:

Complexity: The inclusion of multiple indicators makes it more complex to measure and interpret compared to the income classification.

Data Availability and Consistency: Accurate and consistent data for all indicators can be challenging to obtain, especially for developing countries.

Which is Better?

Depends on the Context: The choice between the World Bank income group classification and the SDI depends on the specific goals and context of the analysis.

Economic Analysis: For purely economic analysis or when assessing a country's ability to finance its development, the World Bank income classification is more appropriate.

Comprehensive Development: For a more holistic understanding of a country's development, including social and economic aspects, the SDI is preferable.

Complementary Use: In many cases, both indicators can be used together to gain a more complete picture of a country's development status. The income classification can provide insights into economic capacity, while the SDI can highlight areas for improvement in terms of social well-being and quality of life.

In summary, while the World Bank income group classification offers a clear and focused view of economic development, the SDI provides a more comprehensive and holistic perspective. The choice between the two should be guided by the specific objectives of the analysis.

3. **“It is important that the authors refer to the UN statement that shows stroke a priority target for reduction and that- following the trends- the SDG goals 2030 will not be reached by 2030. Perhaps they should put this already in the abstract.”**

Conclusion

The disease burden of IS varies greatly among regions and populations, and DALYs in medium-high SDI countries are still high. The prevention and control measures taken so far for the reduction of IS are not enough to meet the Sustainable Development Goal targets by 2030. Countries should formulate prevention and control measures suitable for their national conditions based on the main risk factors, and learn from the experience of countries with good IS burden control.

4. **“While measures for prevention should be based on epidemiological data and their trends, the recommendations given are isolated and not systematic. The authors should refer to guideline papers from leading scientific organisations (eg Int J Stroke 2023; 18: 499-531)”**

According to your suggestions, we added in the abstract that the Sustainable Development Goals may not be achieved by 2030. This conclusion is indeed a very attractive one, and is expected to attract the attention of health management organizations.”

According to your request, we refer to the latest guidelines of major scientific organizations for the treatment and prevention of ischemic stroke.

We refer to guideline papers from leading scientific organisations (eg Int J Stroke 2023; 18: 499-531) for systematic suggestions.

believe that the existing work content has been able to assess the burden of IS in a more comprehensive and objective way, and we will make this study more detailed in the future.

By analyzing the temporal, spatial and population distribution, trend prediction and risk factor analysis of ischemic stroke worldwide based on GBD2021, we hope our study can provide reliable basis for national health policy makers to formulate targeted strategies.Our study provides targeted and innovative but fragmented recommendations for the prevention and treatment of stroke burden based on the latest GBD database. In order to make our guidance more comprehensive and systematic, we combined with the World Stroke Organization to summarize the recommendations of existing stroke guidelines for the management of stroke patients. The recommendations addressed pre-hospital, emergency, and acute hospital care. Strong recommendations were made for reperfusion therapies for acute ischemic stroke. For secondary prevention, strong recommendations included establishing etiological diagnosis; management of hypertension, weight, diabetes, lipids, and lifestyle modification; and for IS, management of atrial fibrillation, valvular heart disease, left ventricular and atrial thrombi, patent foramen ovale, atherosclerotic extracranial large vessel disease, intracranial atherosclerotic disease, and antithrombotics in non-cardioembolic stroke. For rehabilitation, there were strong recommendations for organized stroke unit care, multidisciplinary rehabilitation, task-specific training, fitness training, and specific interventions for post-stroke impairments.

Conclusion

This study covers countries and regions with different levels of social development around the world, covering a time span of more than 30 years. The epidemiological

Overview of Ischemic Stroke

Ischemic stroke is the most common type of stroke, accounting for about 87% of all strokes. It occurs when blood flow to a part of the brain is blocked, often by a blood clot. This can lead to brain tissue damage and various neurological deficits.

Leading Scientific Organizations and Guidelines

Several leading scientific organizations provide guidelines for the management of ischemic stroke, including the American Heart Association/American Stroke Association (AHA/ASA), the European Stroke Organization (ESO), and the World Stroke Organization (WSO). These guidelines are based on extensive research and clinical trials, ensuring that they reflect the best current evidence.

Systematic Recommendations
Prevention:
Lifestyle Modifications: Encourage patients to adopt a healthy lifestyle, including regular physical activity, a balanced diet, and smoking cessation. Lifestyle changes are crucial for reducing the risk of ischemic stroke.
Control of Risk Factors: Manage hypertension, diabetes, and hyperlipidemia through medication and lifestyle interventions. Regular monitoring and adjustment of treatment plans are essential to maintain optimal health.
Acute Management:
Thrombolysis: Administer intravenous alteplase within 4.5 hours of symptom onset for eligible patients. This treatment can help dissolve the clot and restore blood flow to the brain.
Endovascular Therapy: For patients with large vessel occlusions, endovascular therapy, such as mechanical thrombectomy, can be effective up to 24 hours after stroke onset.
Rehabilitation:
Multidisciplinary Approach: Utilize a team of healthcare professionals, including neurologists, physiotherapists, occupational therapists, and speech therapists, to provide comprehensive rehabilitation. Early mobilization and tailored rehabilitation programs can improve functional outcomes.
Ongoing Support: Provide ongoing support and follow-up care to address long-term needs and prevent recurrent strokes. This includes regular assessments and adjustments to the rehabilitation plan.
Long-term Management:
Secondary Prevention: Continue with medications such as antiplatelet agents, antihypertensive drugs, and statins to reduce the risk of recurrent stroke.
Education and Support: Educate patients and their families about stroke prevention and management. Provide resources and support groups to help them cope with the emotional and physical challenges of stroke recovery.
Implementation Strategies
Healthcare Professional Training: Ensure that healthcare providers are well-trained in the latest stroke guidelines and management techniques. Regular updates and continuing education programs can help maintain high standards of care.
Public Awareness Campaigns: Increase public awareness about stroke symptoms and the importance of seeking immediate medical attention. This can help reduce delays in treatment and improve outcomes.
Collaboration and Coordination: Foster collaboration between different healthcare providers and institutions to create a seamless system of care. This includes coordination between emergency services, hospitals, and rehabilitation centers.
By incorporating these systematic recommendations from leading scientific organizations, healthcare systems can improve the prevention, management, and rehabilitation of ischemic stroke, ultimately enhancing patient outcomes and quality of life.

5. **“A larger part of the discussion compares this paper to a recently published similar paper from the Shanghai group (Li et al. 2024). The arguments made are somewhat lengthy, even redundant. This part could be greatly reduced to perhaps a few sentences or one smaller**

paragraph."

Thank you for mentioning the problem of narrative redundancy in comparison with previous studies. We have corrected it.

6. "Line 613: the authors recommend a more detailed reporting of stroke according to the etiology. They should also comment about the likely uncertainties that come with it."

At your suggestion, we have added the results of the analysis that may bring uncertainty if the GBD database is supplemented with more detailed etiologic data.

Reporting stroke etiology in more detail may introduce some uncertainty in GBD studies, and the following are possible sources of uncertainty:

Data collection and diagnostic accuracy

Differences in diagnostic criteria: Different regions and medical institutions may use different criteria and methods for diagnosing stroke and its subtypes, which may lead to inconsistencies in data. For example, some areas may lack advanced imaging equipment to accurately distinguish ischemic stroke from hemorrhagic stroke.

Limitations of data sources: GBD studies rely on a variety of data sources, including hospital records, death certificates, and epidemiological surveys, among others. The completeness and accuracy of these data may vary by region, thus affecting the reporting of etiology.

Demographic and geographic differences

Distribution of causes in different populations: There may be significant differences in the distribution of causes of stroke in different regions and populations. For example, in some low-income countries, stroke may be more associated with hypertension and infectious diseases, whereas in high-income countries it may be more associated with atherosclerosis and lifestyle factors (e.g., smoking, diet). This difference makes for uncertainty when making etiological comparisons across regions.

Urban-rural differences: Medical resources and diagnostic capacity differ between urban and rural areas, potentially leading to bias in etiologic data. For example, urban areas may be more prone to report stroke causes related to lifestyle, whereas rural areas may report more causes related to inadequate medical access.

Temporal variations and trends

Trends in etiologic change: The etiology of stroke may change over time. For example, the incidence of ischemic stroke is likely to rise as the population ages and lifestyle changes occur. This changing trend increases the uncertainty of making etiological comparisons across time periods.

Lag in data update: There may be some lag in data update of GBD studies, which may result in the currently reported etiologic data not fully reflecting the latest changing trends.

Models and methodology

Model assumptions and parameter selection: Data need to be modeled and imputed when performing an etiologic analysis. Model assumptions and parameter choices may affect the accuracy of the results. For example, for some rare stroke subtypes, there may be large uncertainty in the estimation of the model due to the limited amount of data.

Attribution of risk factors: There may be attribution uncertainty when assessing the contribution of risk factors to stroke etiology. Multiple risk factors may coexist and interact, leading to inaccurate

attribution of a single risk factor.
These uncertainties need to be taken into account in the etiological analysis and reporting to ensure the reliability and scientific validity of the results.

Commentary on Likely Uncertainties
Data Limitations: We acknowledge that there are inherent uncertainties associated with reporting stroke by etiology. One major limitation is the variability in diagnostic accuracy and availability of specialized imaging techniques across different regions and healthcare settings. This can lead to misclassification or underreporting of certain stroke subtypes.
Etiological Ambiguity: In some cases, determining the exact etiology of stroke can be challenging due to overlapping clinical presentations or the presence of multiple contributing factors. The undetermined etiology category, in particular, reflects this ambiguity and the need for further research to elucidate the underlying causes.
Temporal and Geographical Variations: The etiological profile of stroke may change over time and across geographical locations due to shifts in risk factor prevalence, improvements in medical care, and other factors. This introduces uncertainties in projecting future trends and comparing historical data.
Impact on Interpretation: These uncertainties should be considered when interpreting the reported etiological data. While the detailed breakdown provides valuable insights, it is important to recognize that the numbers may not fully capture the true complexity and nuances of stroke etiology in all contexts.

By addressing the reviewer's concerns through more detailed reporting and transparent discussion of uncertainties, we hope GBD2023 will provide a more robust and informative analysis of stroke according to etiology, while also acknowledging the limitations and challenges associated with this approach.

To improve the accuracy of stroke etiology data, the following measures can be taken:
Improving diagnostic accuracy
Use of advanced diagnostic tools: use of advanced imaging equipment (such as magnetic resonance imaging MRI, computed tomography CT), cardiac monitoring equipment (such as electrocardiogram ECG, echocardiography), etc., to more accurately determine the cause of stroke.
Multidisciplinary collaboration: We should strengthen the collaboration of neurology, cardiology, endocrinology and other departments, and integrate the information of clinical history, physical examination, laboratory examination, imaging examination and other aspects to make a more comprehensive etiological diagnosis.
Professional training and education: Professional training of clinicians to improve their knowledge of stroke etiology and diagnostic skills, especially in the field of vascular neurology.
Data collection and management
Standardized data collection processes: Unified data collection standards and processes should be developed to ensure consistency among different medical institutions when collecting data on stroke etiology.
The use of electronic health records (Ehrs): The use of EHR systems to automatically collect and organize patients' medical records information to reduce manual entry errors and improve data

integrity and accuracy.
Data cleaning and quality control: The collected data were regularly cleaned to remove duplicate, erroneous or incomplete data records, and quality control was carried out to ensure the reliability of the data.
Applying artificial intelligence and machine learning
Develop automated classification tools: Using natural language processing and machine learning techniques, develop automated stroke etiology classification tools, such as StrokeClassifier, capable of extracting features from EHR text and predicting stroke etiology.
Ensemble multi-model prediction: Ensemble learning method is used to combine the prediction results of multiple machine learning models to improve the accuracy of stroke etiology classification.
Strengthen research and cooperation
Multi-center study: Multi-center cooperative study was conducted to collect larger scale and more representative data to improve the statistical power of etiological analysis and the extrapolation of results.
International cooperation and exchange: strengthen international academic exchanges and cooperation, share the experience and results of stroke etiology research, and promote the standardization and accuracy of stroke etiology data worldwide.
Through these measures, the accuracy of stroke etiology data can be effectively improved and more reliable information support can be provided for the prevention and treatment of stroke.

7. **Line 647: please explain ‘chronic senile diseases’**

The following is the definition of chronic senile diseases (chronic diseases of the elderly), which is a common expression in clinical work, and we add the following:
“Chronic senile diseases” or "Chronic diseases of the elderly" usually refer to chronic noncommunicable diseases that are common and persist for a long time in the elderly population. These diseases are usually characterized by:
Definition and Characteristics
Long-term: Chronic diseases of the elderly usually have a long course and require long-term management and treatment. For example, diseases such as hypertension, diabetes, and coronary heart disease often require ongoing medical interventions and lifestyle modifications once they occur.
Multifactorial disease: The etiology of these diseases is complex and usually related to multiple factors, including genetics, lifestyle, and environmental factors. For example, senile osteoporosis is associated with changes in hormone levels, malnutrition, lack of exercise, and other factors.
Atypical symptoms: In older adults, symptoms of chronic diseases may be less pronounced or typical than in younger adults. For example, pneumonia in the elderly may present with atypical symptoms and be more severe.
Common types
Cardiovascular diseases, such as hypertension, coronary heart disease and stroke, are among the most common chronic diseases in the elderly.
Metabolic diseases: such as diabetes, the efficiency of insulin secretion and action decreases with age, leading to an increased incidence of diabetes.
Respiratory diseases, such as chronic obstructive pulmonary disease (COPD), are closely linked to factors such as long-term smoking and air pollution.

Musculoskeletal diseases: such as osteoporosis and osteoarthritis, which cause weak bones and joint pain that affect mobility in older people.

Prevention and management

Healthy lifestyle: maintaining healthy dietary habits, moderate exercise, smoking cessation and alcohol consumption are essential for the prevention and management of chronic geriatric diseases.

Regular physical examination: Older adults should have regular physical examinations for early detection and intervention of chronic diseases.

Comprehensive management: The management of chronic geriatric diseases requires comprehensive consideration of drug therapy, rehabilitation training, psychological support and other factors.

Prevention and treatment of chronic diseases in the elderly is one of the important measures for elderly health care, which is of great significance for improving the quality of life and prolonging the healthy life span of the elderly.

Reviewer 2#: Thank you very much for your detailed and patient comprehensive analysis and evaluation of our study from different perspectives.

We will modify and reply as follows according to your suggestions. We hope that you can support the publication of our research and express our thanks from the bottom of our heart again.

1.Methodological clarity: The handling of data gaps in regions with limited resources and the extrapolation methods for missing data need more explanation.

Methods for dealing with data gaps

Use of alternative data sources: In areas with limited resources, there may be a lack of complete death registration systems or detailed health records. GBD studies use alternative data sources, such as censuses, population registries, sample registries, and oral autopsies. For example, in countries where complete death registration is not available, sample registries can be used to estimate causes of death.

Data standardization and integration: GBD research is committed to standardizing all available data sources so that comparisons can be made across regions and time periods. In this way, data from different sources can be integrated to fill data gaps.

Extrapolation methods for missing data

Statistical and epidemiological methods: GBD studies apply statistical and epidemiological methods to estimate missing data. For example, the spatio-temporal Gaussian process regression model (ST-GPR) was used to smooth the data and perform extrapolation. This approach enables the use of trends and patterns in existing data to predict missing data.

Model adjustments and assumptions: When making an extrapolation, the researcher makes adjustments and assumptions to the model based on available data and understanding of the epidemiology of the disease. This includes considering factors such as patterns of disease transmission in different regions, the distribution of risk factors, and time trends.

Measures to improve the accuracy and reliability of data

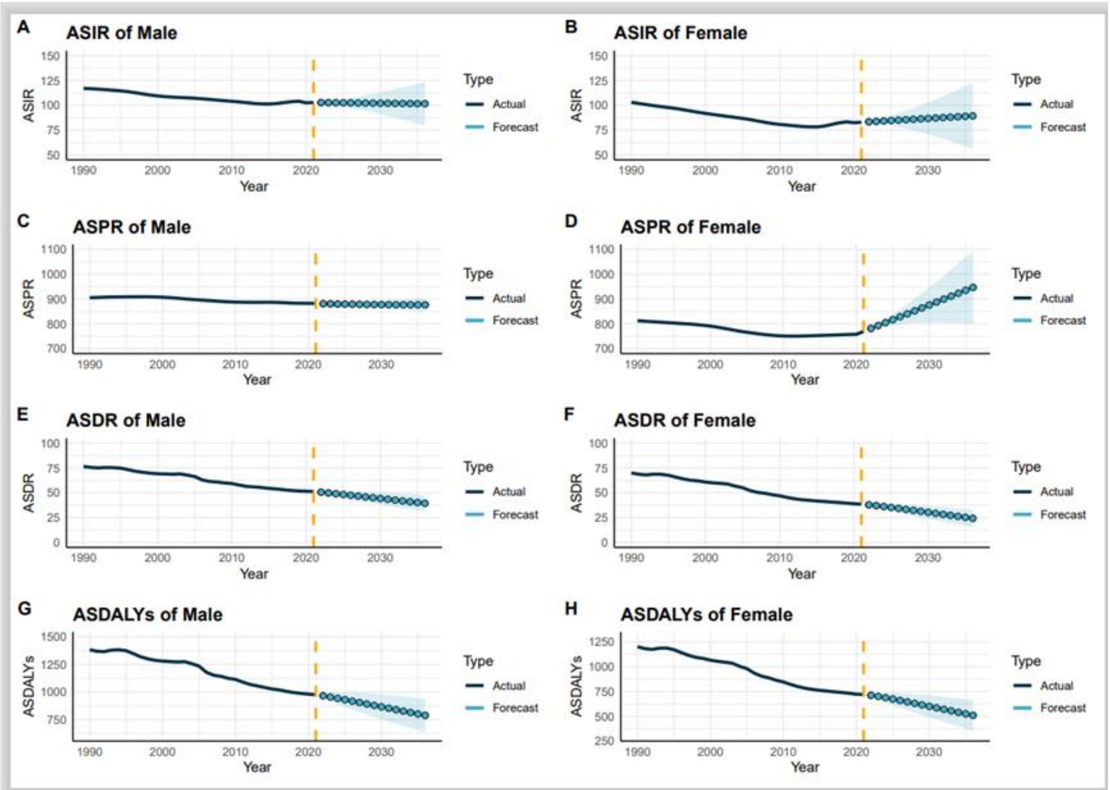
Multi-source data validation: Improve the accuracy and reliability of data by comparing and validating information from different data sources. For example, hospital data were contrasted with verbal autopsy data to ensure the accuracy of cause of death.

Continuous data update and improvement: GBD research is a dynamic, iterative process, and the accuracy and completeness of the data continue to improve as new data are continuously added and analytical methods are improved.
Through these methods and measures, the GBD study can effectively address data gaps and missing data in resource-limited Settings, and provide more accurate and reliable information for the assessment of global burden of disease.
Spatiotemporal Gaussian Process Regression (ST-GPR) model is a model for processing time-series data, which is mainly used to estimate risk factor exposure in the Global Burden of Disease (GBD) study. The following is a detailed explanation of the ST-GPR model:
Structure of model
Gaussian Process Regression Foundation: The ST-GPR model is based on Gaussian process regression (GPR), a nonparametric Bayesian method used for regression analysis. GPR models the distribution of the data by defining a Gaussian process, where each finite set of random variables follows a multivariate normal distribution.
Spatio-temporal modeling: ST-GPR extends GPR by adding modeling of temporal and spatial correlations. It is able to capture the correlation of data in time and space, so as to better deal with data with spatio-temporal characteristics.
Process of application
Data preprocessing: First, the data were preprocessed, including removal of systematic errors and normalization to ensure data accuracy and consistency.
Model fitting: The data were fitted using a linear mixed-effects model as the basis, combined with a spatio-temporal Gaussian process. The model takes into account the correlation of the data at different time and spatial locations.
Smoothing and forecasting: After fitting the model, ST-GPR smoothen the data to reduce the effect of random fluctuations and makes predictions of future values.
Advantages
Handling incomplete data: ST-GPR can effectively deal with missing values and incompleteness in the data, and fill the data gaps through the smoothing and prediction functions of the model.
Uncertainty quantification: The model is able to provide uncertainty quantification of the forecast outcome, such as giving confidence intervals for the predicted values, which is important for assessing the reliability of the forecast.
The challenge
High computational complexity: Due to the need to deal with spatio-temporal correlations, ST-GPR has high computational complexity, especially when dealing with large-scale datasets.
Kernel function selection: selecting the appropriate kernel function is the key of GPR model, and different kernel functions will affect the fitting effect and prediction accuracy of the model.
ST-GPR model plays an important role in GBD research, providing more accurate and reliable data support for the estimation of global burden of disease through its powerful modeling ability and the capture of spatiotemporal characteristics.
Potential problems regarding the availability and quality of raw data: GBD is a recognized global database with reliable data sources, and GBD database is the basis for all GBD research. Although it may have some limitations, its analytical methods and research framework are systematic and standardized, and its results are reliable. As in all previous published GBD studies.
In the Supplementary materials section, we elaborate on the methods to deal with the data gap and

the extrapolation of missing data in areas with limited resources, which is mainly limited by the GBD database itself. We mainly use the visual method to describe the GBD data more simply and easily, but do not change the GBD database itself. Therefore, all published GBD studies are faced with various problems faced by data in sparse areas.

2. Validation of future projections: The assumptions underlying the prediction models require further elaboration or cross-validation with alternative models.

Thank you for your insightful comment. In response to your suggestion for model validation, we have further validated our projections by employing the Autoregressive Integrated Moving Average (ARIMA) model, which is a widely accepted method for time series forecasting. We found that the trends predicted by the BAPC model were largely consistent with those generated by the ARIMA model for all outcomes (incidence, prevalence, mortality, and DALYs). This cross-validation reinforces the robustness of our initial projections. We have now included this validation approach and the corresponding results in the revised manuscript to provide further clarity and confidence in our findings.We supplemented the assumption basis of the BAPC prediction model and added the ARIMA model for additional cross validation. We hope that our modification can be approved by experts.



3.We supplemented the assumption basis of the BAPC prediction model and added the ARIMA model for additional cross validation. We hope that our modification can be approved by experts.

4.Selection of risk factors: While 12 major risk factors are addressed, justification for excluding others is not fully provided.

In GBD research, the selection and exclusion of risk factors is a complex process involving multiple considerations:

Selection of risk factors

Data-driven associations: GBD studies select risk factors based on data-driven risk-outcome associations. This means that a factor is included in the analysis only if there is sufficient evidence

of a statistically significant association between the factor and a particular health outcome.
Coverage at global, regional and country levels: Risk factors need to be selected considering their prevalence and importance at global, regional and country levels. For example, risk factors such as particulate air pollution, high systolic blood pressure, and smoking were identified as major contributors to the global burden of disease in 2021.
Contribution to health burden: Select those risk factors that contribute significantly to the burden of disease in order to better understand and address global health issues. For example, increased metabolic risk factors such as high BMI and high fasting glucose have a significant impact on the global burden of disease.
Exclusion of risk factors
Lack of data support: If a factor lacks sufficient data to support its association with health outcomes, or if the data quality is poor, then the factor may be excluded.
Nonsignificant associations: When the association between a risk factor and a health outcome is not significant or cannot be determined, the factor may be excluded to avoid introducing unnecessary complexity and uncertainty.
Adjustment for mediation effects: During the analysis, mediation effects are adjusted to account for the relationship between risk factors that act indirectly on outcomes through intermediate risks. A factor may be excluded if it does not have a significant effect on health outcomes after adjusting for mediating effects.
Through these selection and exclusion criteria and methods, GBD research can ensure that the risk factors included have scientific basis and practical significance, so as to provide accurate and reliable information support for the development and implementation of global health policies.
We selected the top 12 most important risk factors, whose selection was based on distribution weights from high to low, which were all risk factors at the third level. For factors with relatively low weight, they are not universally representative for guiding policy making at the macro level, and they occur more often in a small number of people, and the uncertainty of their results increases due to the decrease in population size. We hope that the story we tell can give readers and policy makers universal suggestions.
5. The conclusions drawn in the manuscript seem to be original, particularly in their comprehensive assessment of risk factors across different demographics and regions. However, if similar studies have been published recently, it would be beneficial to reference them to contextualize the findings

In fact, there were no GBD2021 studies related to ischemic stroke published when we wrote this article. However, due to the important role of ischemic stroke in the disease, three related studies have appeared in the Lancet and its sub-journals in recent months, and we included them in our article according to the requirements of the reviewer. However, their research Angle and content are different from ours, and our research has our unique innovation and value significance. Comparison with previous studies can also highlight our innovation points and make the GBD2021 IS-related research more comprehensive.
6. **Generally compliant, but missing explicit declarations of conflicts of interest and funding.**
We have added a clear statement of conflicts of interest and funding in the original article

Funding

This work was supported by National Major Science and Technology Projects for Cancer, Cardiovascular and Cerebrovascular Diseases, Respiratory and Metabolic Diseases (2023ZD0515900). The Bill and Melinda Gates Foundation, who were not involved in any way in the preparation of this manuscript, funded the Global Burden of Disease

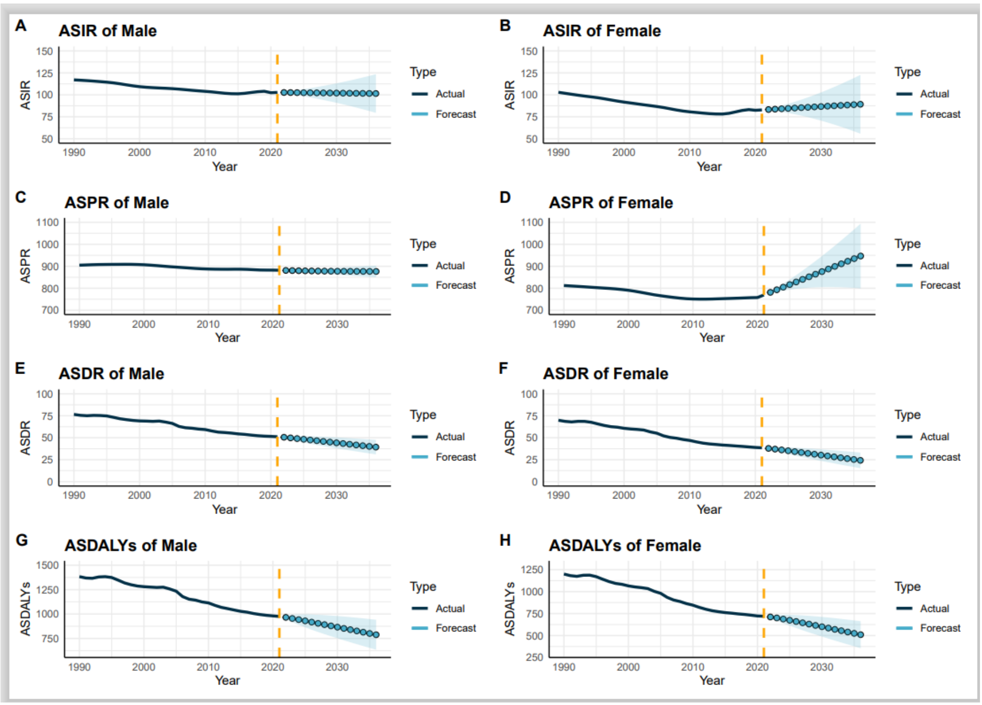
study.

Competing interests

The authors declare that they have no competing interests.

7. The methodology aligns with STROBE guidelines for observational studies but would benefit from clearer reporting of limitations and assumptions in the predictive models.

We add the ARIMA model for prediction and compare it with the BAPC model, and report the assumption basis and limitations of the two prediction models.



The ARIMA model was used to predict the same outcomes (incidence, prevalence, mortality, and DALYs) from 2021 to 2036, and the results were compared to the projections from the BAPC model.

Both models produced consistent trends, suggesting robustness in the long-term predictions.
The Bayesian age-period-cohort (BAPC) model relies on several key assumptions that are critical to the validity and accuracy of the model:
Separability of age, period, and cohort effects
Basic Assumptions: The BAPC model assumes that time-related changes in health outcomes can be separated into age, period, and cohort effects. The age effect refers to the changes in an individual's life course as he or she ages. Period effects reflect the impact of events occurring within a given year on health outcomes; Cohort effect is the variation due to the different characteristics of the population in different birth years.
The prior distribution of the parameters
Use of prior information: In a Bayesian framework, the BAPC model allows prior information to be incorporated into the parameter estimation process. This means that for an unknown parameter in the model, an appropriate prior distribution needs to be specified for it. For example, a highly concentrated Laplace distribution can be used as a prior for parameters with linear effects, whereas a diffusive normal distribution can be used for slopes of age and period.
Similarity of neighboring effects
Adjacent effect Assumption: The BAPC model assumes that the adjacent effects of age, period, and cohort are similar in time. This means that the model considers the effect changes to be smooth and continuous between adjacent time points or age groups.
Poisson distribution assumption of the data
Poisson distribution: In some cases, the BAPC model assumes that the data follow a Poisson distribution. This assumption applies to count data, such as disease incidence or mortality, etc., and it allows the model to use Poisson regression models when dealing with such data.
These assumptions provide the theoretical basis and computational framework for the BAPC model, but also limit the scope of model application and the interpretation of results. In practice, careful consideration needs to be given to whether these assumptions are applicable to specific research contexts and data characteristics.
The assumptions of the BAPC model have significant impacts on the analysis results. Here is a detailed explanation:
Separability of Age, Period, and Cohort Effects
Identification of Influencing Factors: The assumption that age, period, and cohort effects can be separated allows the model to identify and quantify the roles of these three effects in health outcomes. For example, if the incidence of a certain disease increases over time, the model can isolate the period effect to determine whether this is due to increased diagnoses from medical advancements or an actual rise in disease occurrence due to environmental changes.
Result Interpretation: This assumption enables clearer interpretation of the results by distinctly attributing changes in health outcomes to different factors. For instance, it can clearly indicate that a high disease incidence in a specific age group is due to physiological changes unique to that age (age effect) or due to particular environmental conditions when that age group was born (cohort effect).
Prior Distributions of Parameters
Robustness of Estimation Results: The choice of prior distributions affects the estimation of parameters. If the prior information is accurate and consistent with the actual data, it can enhance the robustness and accuracy of the estimates. For example, using highly concentrated prior

distributions can reduce the uncertainty in parameter estimation, bringing the results closer to the true values.
Subjectivity of Results: The setting of prior distributions may introduce subjectivity, as different priors can lead to different analysis results. This requires researchers to be cautious when setting priors and to provide sufficient justification for the choice of prior distributions.
Similarity of Adjacent Effects
Smoothness Constraint: Assuming similarity between adjacent effects essentially imposes a smoothness constraint on the changes in effects. This causes the model to tend to produce smooth curves in estimating effects, rather than abrupt fluctuations. For example, when estimating age effects, the model assumes that changes in incidence between adjacent age groups are gradual, not sudden jumps.
Reasonableness of Results: In many cases, the smoothness constraint is reasonable because changes in health outcomes are often continuous. However, in some special situations, if the actual effects do indeed change dramatically, this assumption may prevent the model from accurately capturing the true pattern of effects.
Poisson Distribution Assumption for Data
Applicability Limitations: The Poisson distribution assumption is suitable for count data, such as disease incidence or mortality rates. If the data do not conform to the characteristics of the Poisson distribution (such as overdispersion or zero inflation), this assumption can lead to poor model fit, thereby affecting the accuracy of the analysis results.
Bias in Results: When data do not meet the Poisson distribution assumption, it can cause bias in the estimation results. For example, if data exhibit overdispersion, using Poisson regression models may underestimate the variance, thus affecting the significance tests of parameters and the reliability of prediction results.
In summary, the assumptions of the BAPC model significantly impact the analysis results, helping to identify and interpret the roles of different factors while also potentially limiting the model's applicability and accuracy in certain situations. In practical applications, it is necessary to test and adjust these assumptions based on specific data and research contexts to ensure the reliability and validity of the analysis results.
There are some limitations in the application of Bayesian age-period-cohort (BAPC) prediction model, which are analyzed as follows:
Difficulty in parameter estimation
Identification problems caused by linear relationships: There is a linear relationship between the three factors of age, period and cohort, which makes parameter estimation more difficult. The trend of change estimated using the APC model may even be reversed without certain conditional constraints.
Data dependency
Impact of data quality: The predictive accuracy of the BAPC model depends to some extent on the quality of the input data. If the data is missing, wrong or incomplete, it will affect the fitting effect of the model and the prediction result.
Limitations of data sources: In some studies, data may come from GBD and other databases, but these data themselves are estimated based on system dynamics model and statistical model, which has certain limitations.
Limitations of model assumptions

No support for inclusion of covariates: BAPC model does not support inclusion of covariates in the analysis process, which means that other influencing factors (such as socioeconomic factors and environmental factors) other than age, period and cohort cannot be considered, which limits the ability of the model to explain the complex causes of changes in disease burden.

Computational complexity

High computational resource demand: BAPC model usually adopts Bayesian Markov chain Monte Carlo (MCMC) algorithm for parameter estimation, which has high computational complexity and requires a large amount of computing resources and time.

Limitations of the prediction range

Uncertainty in long-term projections: Although the BAPC model can be used to predict future trends in disease burden, the prediction results of the model may be more uncertain due to the greater uncertainty of future influencing factors when making long-term projections.

These limitations need to be considered and overcome when applying the BAPC model to improve the accuracy and reliability of the model.

To overcome the problem of data dependence of BAPC models, several strategies can be adopted:

Data quality improvement

Data cleaning and preprocessing: The original data were thoroughly cleaned to remove outliers, duplicate records, and inconsistent data to ensure the accuracy and integrity of the data.

Data calibration: Calibration of data to reduce systematic and random errors by comparison with other data sources or known facts.

Diversity of data sources

Integrating multiple data sources: In addition to traditional medical records and death registration data, multiple data sources such as censuses, socioeconomic surveys, and environmental monitoring can also be integrated to improve data comprehensivity and representativability.

Utilizing alternative data: In cases where some data is missing, alternative data can be sought to fill the gap. For example, sample survey data are used to estimate the overall picture.

Model refinement and validation

Model Hypothesis testing: The hypotheses in the model are rigorously tested to ensure that they are supported in the data. If some assumptions are found not to be true, the model structure should be adjusted in time.

Cross-validation: Cross-validation was used to validate the model and evaluate its performance on different datasets to ensure the stability and generalization ability of the model.

Advantages of the Bayesian approach

Utilization of prior information: The Bayesian approach allows prior information to be incorporated into the model, which can alleviate the problem of insufficient data to some extent. Additional information support for the model is provided by combining expert knowledge and historical data.

Uncertainty quantification: Bayesian methods are able to provide uncertainty quantification of parameter estimates, such as confidence intervals or probability distributions. This helps to assess the impact of data quality on model predictions.

Through these measures, the dependence of BAPC model on data can be effectively reduced, and the accuracy and reliability of model prediction can be improved.

The autoregressive integrated moving average (ARIMA) model is a statistical model widely used in time series analysis to predict future data points. Although the ARIMA model has been successful in many areas, it is based on a series of assumptions and suffers from some limitations. The following are the main assumptions and limitations of the ARIMA model:

The main assumptions of the ARIMA model

Stationarity:

Assumptions: Time series data are stationary, that is, their statistical properties (e.g., mean, variance) remain constant in time.

Limitations: Many real time series data are non-stationary, such as trends and seasonal variations in economic data. If the stationarity assumption is not met by the data, the predictive performance of the model may decrease significantly. Although it is possible to smooth the data through methods such as difference, excessive difference may lead to loss of information.

Linearity:

Assumption: The relationship in time series data is linear, that is, future values can be represented as linear combinations of past values.

Limitations: Many real time series data have nonlinear relationships, such as volatility clustering in financial markets. The inability of ARIMA models to capture these nonlinear features may lead to prediction bias.

Independent identically distributed (IID) error term:

Assumption: The error terms of the model are independently and identically distributed with zero mean and constant variance.

Limitations: In practice, the error term may not satisfy the assumption of independent identically distributed, such as the presence of autocorrelation or heteroscedasticity. This may lead to inaccurate estimation results of the model and affect the reliability of the prediction.

Model Order Selection:

Assumptions: The order of the model (p, d, q) is known or can be determined in some way.

Limitations: Selecting the appropriate model order is a complex process that usually relies on information criteria such as AIC, BIC, or methods such as cross validation. Inappropriate order selection may lead to overfitting or underfitting of the model, affecting the prediction performance.

Limitations of the ARIMA model

Limited capacity for non-stationary data:

The ARIMA model requires the data to be stationary, but many actual data are non-stationary. Although it is possible to smooth the data through methods such as difference, excessive difference may lead to loss of information and affect the predictive ability of the model.

Failure to capture nonlinear relationships:

Based on the assumption of linearity, ARIMA model cannot effectively capture the nonlinear relationship in time series. For data with nonlinear features, ARIMA models may not provide accurate predictions.

Sensitive to outliers:

The ARIMA model is sensitive to outliers in the data, which may lead to inaccurate model parameter estimation and thus affect the prediction results.

Model assumptions are strict:

The assumptions of the ARIMA model are strict, and it is often difficult for the actual data to fully meet these assumptions. For example, the assumption of independent identically distributed error terms may not hold in many cases, leading to inaccurate estimation results of the model.

The forecast is limited:

ARIMA model is mainly used for short-term prediction, but may not be effective for long-term prediction. This is because the model assumes that the statistical properties of the data remain constant over the forecast period, whereas the statistical properties of the actual data may change over time.

Model selection and parameter estimation are complex:

Selecting the appropriate ARIMA model order is a complex process that usually relies on methods such as information criteria or cross-validation. Inappropriate order selection may lead to overfitting or underfitting of the model, affecting the prediction performance.

Summary

The ARIMA model is a powerful tool for time series analysis, but its assumptions and limitations need to be carefully considered when applied. For data that are non-stationary, nonlinear, or have outliers, it may be necessary to combine other models, such as seasonal ARIMA models, GARCH models, or machine learning methods, to improve the accuracy and reliability of prediction.

Compared with other prediction models, ARIMA model has the following advantages and disadvantages:

Advantages of the ARIMA model

The model is simple, easy to understand and apply: The structure of ARIMA model is relatively simple, and it is mainly modeled by three parts: autoregression (AR), difference (I) and moving average (MA), without the help of other exogenous variables.

High prediction accuracy: For short-term prediction, ARIMA model has high prediction accuracy and reliability, and can better capture the trend and seasonality of time series.

High maturity: ARIMA model has a solid theoretical foundation, and after long-term development and application, it has formed a set of relatively complete theoretical system.

Suitable for non-seasonal data: ARIMA model is suitable for time series data with no obvious seasonal characteristics, and can smooth non-stationary time series by difference operation.

Disadvantages of the ARIMA model

The requirement for data stationarity is high: ARIMA model requires time series to be stationary, or to be stationary after difference operation. If there is an obvious trend or seasonality in the data, it needs to be smoothed first, which may affect the prediction effect of the model.

Parameter selection is complex: ARIMA model needs to determine three parameters p (autoregressive order), d (difference order) and q (moving average order). Improper parameter selection may lead to overfitting or underfitting of the model.

Only linear relationships can be captured: The ARIMA model can intrinsically only capture linear relationships, and the prediction effect may not be good for data with nonlinear relationships.

The long-term prediction effect is poor: the ARIMA model will gradually decrease the prediction accuracy due to the accumulation of errors in the long-term prediction.

The comparison between ARIMA model and BAPC model in time series prediction is as follows:

ARIMA model

Advantages:

The model is simple, easy to understand and apply: The structure of ARIMA model is relatively simple, and it is mainly modeled by three parts: autoregression (AR), difference (I) and moving average (MA), without the help of other exogenous variables.

High prediction accuracy: For short-term prediction, ARIMA model has high prediction accuracy

and reliability, and can better capture the trend of time series.
High maturity: ARIMA model has a solid theoretical foundation, and after long-term development and application, it has formed a set of relatively complete theoretical system.
Suitable for non-seasonal data: ARIMA model is suitable for time series data with no obvious seasonal characteristics, and can smooth non-stationary time series by difference operation.
Disadvantages:
The requirement for data stationarity is high: ARIMA model requires time series to be stationary, or to be stationary after difference operation. If there is an obvious trend or seasonality in the data, it needs to be smoothed first, which may affect the prediction effect of the model.
Parameter selection is complex: ARIMA model needs to determine three parameters p (autoregressive order), d (difference order) and q (moving average order). Improper parameter selection may lead to overfitting or underfitting of the model.
Only linear relationships can be captured: The ARIMA model can intrinsically only capture linear relationships, and the prediction effect may not be good for data with nonlinear relationships.
The long-term prediction effect is poor: the ARIMA model will gradually decrease the prediction accuracy due to the accumulation of errors in the long-term prediction.
The BAPC model
Advantages:
Comprehensive consideration of multiple factors: BAPC model took into account the influence of age, period and cohort on the distribution of NCDS, which could more comprehensively analyze the trends of incidence and mortality of NCDS.
Unique advantages of the Bayesian framework: In the Bayesian framework, BAPC model can use both sample information and prior information to obtain unique parameter estimates, which makes the results more robust and reliable.
Improved prediction accuracy: Compared with the traditional APC model, the BAPC model can obtain more stable and reliable results and improve the prediction accuracy.
Solving the difficulty in parameter estimation: There is a linear relationship among the three factors of age, period, and cohort in the classical APC model, which may lead to difficulty in parameter estimation. The BAPC model effectively solves this problem through the Bayesian approach.
Flexible model selection: The BAPC model allows the flexible selection of different model components and prior probability distributions, which improves the applicability and flexibility of the model.
Advanced computational methods: The BAPC model uses the ensemble nested Laplace approximation (INLA) for posterior marginal distribution approximation, which avoids mixing and convergence problems and shows a relatively low error rate.
Automatic and simplified calculation: The use of the BAPC package for parameter estimation not only does not require complex coding procedures and convergence checks, but also allows flexible selection of age-standardized ratio (ASR) parameters and prior probability distributions to simplify the calculation process.
Predicting future changes: BAPC models can not only describe trends in age, period, and birth cohort, but also predict future changes based on these trends.
Disadvantages:
High computational complexity: The computational process of BAPC model is relatively complex and requires high computational resources and time.

High data requirements: BAPC models require sufficient data to estimate model parameters, and for cases with a small amount of data, the accuracy and reliability of the model may be affected.
There are many model assumptions: the BAPC model is based on a Bayesian framework and requires assumptions about the prior distribution, which may affect the prediction results of the model.
Comparison
Applicable scenarios:
ARIMA model: Suitable for time series with no or very weak seasonality. The effect was poor when there was obvious seasonality in the data.
BAPC model: suitable for disease prediction that requires comprehensive consideration of age, period and cohort factors, especially in the field of public health and epidemiology.
Predictive power:
ARIMA model: performs well in short-term forecasting, but may not perform well for long-term forecasting and data with nonlinear relationships.
BAPC model: It can more accurately predict the change trend of diseases in different ages, periods and cohorts, and is suitable for long-term prediction and complex data structure.
Model complexity:
ARIMA model: The model structure is relatively simple, with few parameters, which is easy to understand and apply.
BAPC model: The model structure is complex and requires Bayesian estimation and complex calculations.
Data requirements:
ARIMA model: It has high requirements for the stationarity of the data and needs pre-processing such as difference.
BAPC model: Sufficient data are needed to estimate the model parameters, and high data volume and data quality requirements are required.

8. Impact of COVID-19: The manuscript mentions the impact of the COVID-19 pandemic but does not delve deeply into how it may have influenced ischemic stroke incidence and risk factors. A more detailed analysis of this impact would add valuable context.

The objective impact of COVID-19 on ischemic stroke was determined by comparing GBD2019 and GBD2021. Since GBD2021 was published after the COVID-19 epidemic in 2019, it can reflect the impact of COVID-19 on ischemic stroke in the first two years. The information of GBD2023 database published in the future will provide information on the impact of COVID-19 on ischemic stroke when COVID-19 quarantine measures are removed and COVID-19 is gradually fading out. We add a deeper look at how COVID-19 may affect the incidence and risk factors for ischemic stroke.
Impact on Ischemic Stroke Incidence
Decreased Admissions: Early observations in several countries reported a significant reduction in admissions for acute ischemic stroke during the pandemic. For example, some studies noted a decrease of up to 50% to 88% in stroke admissions. This could be attributed to patients avoiding healthcare facilities due to fear of COVID-19 infection, leading to underreporting of stroke cases.
Delayed Presentations: The pandemic led to delays in seeking medical care for stroke symptoms. Patients with mild strokes or transient ischemic attacks were less likely to present to hospitals, potentially increasing the risk of recurrent strokes and the need for secondary prevention.

Influence on Risk Factors

Increased Risk in COVID-19 Patients: Studies have shown that COVID-19 infection is an independent risk factor for ischemic stroke. Patients with COVID-19 have been found to have a higher risk of developing acute ischemic stroke, even after controlling for traditional vascular risk factors.

Exacerbation of Existing Risk Factors: The pandemic has exacerbated existing risk factors for stroke. Patients with pre-existing conditions such as hypertension, diabetes, and cardiovascular diseases were more susceptible to severe COVID-19 infections, which in turn increased their risk of stroke.

Lifestyle Changes: Changes in lifestyle during lockdowns, such as reduced physical activity, increased stress, and changes in diet, may have contributed to an increase in stroke risk factors like obesity and hypertension.

Clinical Outcomes and Treatment Delays

Worse Outcomes: Patients with ischemic stroke who also had COVID-19 infections tended to have more severe strokes and poorer functional outcomes. The mortality rate was also higher in these patients.

Treatment Delays: The pandemic led to delays in treatment for stroke patients. For instance, the time from hospital arrival to reperfusion treatment was significantly prolonged during the pandemic. This delay could have contributed to worse outcomes for stroke patients.

By incorporating these insights into the manuscript, it would provide a more comprehensive understanding of how the COVID-19 pandemic has influenced ischemic stroke incidence and risk factors, highlighting the need for continued vigilance and adapted healthcare strategies.

Delayed stroke presentations can have significant implications on long-term stroke risk and outcomes. Here are some key points to consider:

Increased Risk of Recurrent Stroke

Missed Treatment Window: When stroke symptoms are not promptly recognized or medical care is delayed, patients may miss the crucial treatment window for interventions like intravenous thrombolysis or mechanical thrombectomy. These time-sensitive treatments are most effective when administered within a few hours of symptom onset. Missing this window can leave patients with unresolved clot blockages, increasing the likelihood of recurrent strokes in the future.

Inadequate Secondary Prevention: Delayed presentations can lead to suboptimal initiation of secondary prevention measures. Early initiation of medications like antiplatelet agents, anticoagulants, and cholesterol-lowering drugs is crucial for reducing the risk of recurrent strokes. Patients who present late may not receive these preventive measures as soon as they should, allowing risk factors to persist and potentially leading to another stroke.

Worsened Post-Stroke Outcomes

Greater Brain Damage: The longer the delay in seeking medical care, the more time there is for brain tissue to be damaged due to lack of blood flow. This can result in larger infarct sizes and more severe neurological deficits. Patients with more extensive brain damage are at a higher risk of long-term complications and disability, which can indirectly increase their risk of subsequent strokes due to factors like immobility and increased dependency on caregivers.

Impaired Rehabilitation: Delayed presentations can disrupt the optimal timing for initiating rehabilitation therapies. Early rehabilitation is essential for maximizing functional recovery and independence. Patients who start rehabilitation later may face a slower and less effective recovery process, leaving them with lingering disabilities that can contribute to a higher risk of falls, injuries,

and subsequent strokes.
Impact on Risk Factor Management
Unaddressed Risk Factors: When stroke presentations are delayed, underlying risk factors like hypertension, diabetes, atrial fibrillation, and smoking may not be promptly identified and managed. Uncontrolled risk factors can continue to cause damage to blood vessels and the heart, increasing the likelihood of recurrent strokes over time.
Complications from Untreated Risk Factors: Delayed presentations can lead to complications from untreated risk factors. For example, uncontrolled hypertension can cause further damage to blood vessels, leading to conditions like vascular dementia or intracerebral hemorrhage, which can increase the risk of recurrent strokes.
Psychological and Social Factors
Stress and Anxiety: The experience of a delayed stroke presentation and the resulting uncertainty about one's health can cause significant stress and anxiety. These psychological factors can negatively impact a patient's adherence to medication regimens and lifestyle modifications, which are crucial for reducing long-term stroke risk.
Social Isolation: Patients who delay seeking care may face social isolation due to the prolonged recovery process and potential complications. Social isolation has been linked to increased stroke risk, as it can lead to poor mental health, lack of support for maintaining healthy behaviors, and reduced access to healthcare resources.
In summary, delayed stroke presentations can significantly increase long-term stroke risk by missing critical treatment windows, impairing rehabilitation, disrupting risk factor management, and introducing psychological and social challenges. It highlights the importance of prompt recognition of stroke symptoms, timely access to medical care, and comprehensive management of stroke patients to minimize the risk of recurrent strokes and improve long-term outcomes.

9. Comparative Analysis: The authors could conduct a comparative analysis of countries with reduced ischemic stroke burden to identify successful strategies that could be replicated in other regions.

We added pooled experience from countries with a lower ischemic burden. Hope to get the approval of the review experts.
1. Canada
Strategy: Canada promotes standardized Stroke management processes in emergency departments and acute stroke care by implementing the Canadian Stroke Best Practice Recommendations.
Effects: These measures significantly improved treatment outcomes for stroke patients and reduced mortality and disability from stroke.
2. The U.S.
Strategy: Endovascular treatment of acute ischemic stroke based on imaging assessment was promoted in the United States through the DEFUSE 3 trial.
Results: This treatment modality significantly increased patient recovery rates and reduced long-term disability.
3. Europe
Strategy: Europe assessed access to and delivery of acute ischemic Stroke treatment in 44 European countries through a "European Stroke Journal" survey.
Results: Establishing regional acute Stroke care networks and stroke units were found to be effective measures to reduce the stroke burden.

4. Japan
Strategy: Japan evaluated stroke incidence, mortality, and functional status in different regions through the SAMBA study, which analyzed stroke in multiple regions of Brazil.
Effect: These findings helped the Japanese government to develop targeted stroke prevention and treatment strategies, which significantly reduced the burden of stroke.
5. Chile
Strategy: Chile found through research that socioeconomic and cardiovascular variables are key factors explaining regional differences in stroke mortality.
Effect: By improving these factors, Chile succeeded in reducing stroke mortality.
6. Brazil
Strategy: Brazil analyzed the incidence and mortality of stroke in different regions through the SAMBA study.
Effect: These studies helped the Brazilian government to develop targeted stroke prevention and treatment strategies that significantly reduced the burden of stroke.
7. Global action
Strategies: The World Stroke Organization and the World Health Organization (WHO) have proposed comprehensive strategies to reduce the Burden of Stroke through the Global Burden of Stroke Report.
Effects: These strategies include enhanced stroke surveillance, improved preventive measures, and improved quality of acute care and rehabilitation services.
Summary
Successful experiences in these countries suggest that key strategies to reduce the burden of ischemic stroke include:
Strengthen stroke surveillance: Establish an effective stroke surveillance system that provides comprehensive stroke data.
Promoting standardized care: Promoting standardized stroke management processes in emergency departments and acute stroke care.
Raising public awareness: Improving public awareness of stroke and ability to cope with it through public education campaigns.
Lifestyle modification: Lifestyle modification by reducing salt intake and increasing physical activity.
Establishing a regional stroke care network: Improving treatment outcomes for stroke patients by establishing a regional stroke care network and a stroke unit.
To improve stroke care in our region, we can learn from successful strategies in other countries and regions. Here are some specific recommendations:
1. Establish a regional stroke care network
Success story: The IMPROVE Stroke Care program in the United States has successfully improved treatment outcomes for patients with acute stroke by establishing a regional stroke care network.
Specific measures:
Establish a network of comprehensive stroke centers (Hub Hospitals) and regional/community referral Hospitals (Spoke Hospitals).
Through regular meetings and performance reviews, best practices are promoted to improve the efficiency of treatment for stroke patients.
Strengthen emergency medical services (EMS) training to ensure rapid identification and transport

of stroke patients.
2. Promote standardized stroke management processes
Success story: Canada has significantly improved treatment outcomes for stroke patients by implementing the "Canadian Stroke Best Practice Recommendations" to promote standardized stroke management processes in emergency departments and acute stroke care.
Specific measures:
Develop and promote a guideline-based stroke management process to ensure it is followed by all healthcare facilities.
Regular training of healthcare staff to ensure they are familiar with the latest stroke treatment guidelines and techniques.
3. Strengthen public education
Success story: The American Heart Association/American Stroke Association (AHA/ASA) promotes healthy lifestyles and reduces stroke risk through the "Life's Essential 8" cardiovascular health indicators.
Specific measures:
Conduct public education campaigns to increase public awareness of stroke risk factors, such as hypertension, high cholesterol, and diabetes.
Promoting the Mediterranean diet pattern and encouraging lifestyle changes such as healthy diet, regular exercise, smoking cessation and alcohol reduction.
4. Improve your lifestyle
Success Story: The American Heart Association/American Stroke Association "Life's Essential 8" cardiovascular health indicators emphasize the importance of a healthy lifestyle.
Specific measures:
Encourage healthy diets with reduced intake of salt and saturated fats and increased intake of fruits and vegetables.
Promoting regular exercise, recommending that adults do at least 150 minutes of moderate-intensity aerobic activity or 75 minutes of vigorous-intensity aerobic activity per week.
Provide counselling on smoking cessation and alcohol restriction to help patients quit smoking and reduce alcohol intake.
5. Control medical conditions
Success Story: The American Heart Association/American Stroke Association guidelines emphasize the importance of controlling medical conditions.
Specific measures:
Check blood pressure, cholesterol, and blood sugar levels regularly to make sure they are in healthy ranges.
For patients with heart diseases such as atrial fibrillation, targeted treatment and management programs are provided.
6. Strengthen rehabilitation nursing
Success case: The StroCare program in Germany provided integrated rehabilitation care through cross-sectoral collaboration and significantly improved patient rehabilitation outcomes.
Specific measures:
To establish the continuity of rehabilitation care after discharge and reduce the waiting time of patients.
Regular follow-up and rehabilitation, including physical therapy, occupational therapy, and speech

therapy, were provided.

7. Leverage digital platforms

Success case: The StroCare project in Germany explored the possibility of using digital platforms to improve the care of stroke patients.

Specific measures:

Develop a digital patient platform to facilitate sharing of reports and information between health care providers and patients.

Leveraging electronic Health services (e-health) to improve communication and data sharing.

8. Focus on social determinants

Success Story: The American Heart Association/American Stroke Association guidelines highlight the influence of social determinants on stroke risk.

Specific measures:

Ensuring patients have access to affordable health services and medicines.

Provide resources to help patients address health-related social needs, such as food and housing insecurity.

By implementing these strategies, we can significantly improve stroke care in our region and reduce the burden of stroke.

10. Subgroup Analysis: Further exploration of specific subpopulations (e.g., rural vs. urban, young adults vs. elderly) could provide insights into how risk factors and outcomes differ across demographics.

Regarding subgroup analysis, we have conducted detailed subgroup analysis of risk factors and incidence rates according to age and gender. Since there is no clear division and definition of rural and urban areas in the GBD database, there is no precedent for the analysis of rural and urban areas in all previous GBD studies. Such stratification is currently not feasible in terms of data source and technology. We hope that the GBD database can add regional stratification for rural and urban areas in the future, but this is not feasible in terms of data source for our study at present, because there is no data for rural and urban areas.

Reviewer 3#:

Thank you very much for your professional comments and important guidance for us to improve the manuscript. We hope that you will support the publication of our research and express our gratitude from the bottom of our heart again.

1. "It is important to discuss generally the quality of stroke incidence/prevalence data in different settings and the biases of using other country data where data are sparse in countries."

The quality of data on stroke incidence and prevalence in different Settings and the bias of using data from other countries in countries with sparse data are discussed as follows:

Dealing with sparse data is a key challenge in GBD research. The following are some common approaches in GBD research to deal with sparse data problems:

Indirect methods and alternative data sources were used

Indirect methods: When direct data are scarce, GBD research will use indirect methods to estimate the distribution of health states. For example, mapping functions were developed to estimate GBD disability weights by analyzing self-reported disease prevalence and universal Health status assessment data from large population surveys.

Alternative data sources: In the absence of detailed disease data, GBD studies utilize other relevant data sources for estimation. For example, data such as censuses, surveys, surveillance systems, and cancer registries are used to supplement and calibrate estimates of disease burden.

Refinement of models and methods

Spatial-temporal Gaussian process regression (ST-GPR) : ST-GPR model can handle sparse data and fill data gaps through smoothing and interpolation methods, thereby improving the accuracy and reliability of estimation.

Causal Mortality Ensemble Model (CODEm) : This model improves its performance on sparse data by exploring multiple possible model combinations combined with optimal nonsample validation results.

Uncertainty assessment and transparency

Uncertainty intervals: The GBD study provides uncertainty intervals for each estimate to quantify the uncertainty introduced by sparse data through simulations and predictive validation tests.

Transparency and replicability: GBD research emphasizes transparency and replicability of methodology, ensuring that findings can be validated and understood by other investigators even when data are scarce.

Through these methods, GBD study can effectively deal with the problem of sparse data, and provide more accurate and reliable information support for the assessment of global burden of disease.

Quantifying uncertainty in disease burden is a critical step in GBD research to ensure reliability and transparency of findings. The following are the main methods used to quantify uncertainty in GBD research:

Calculation of uncertainty intervals

95% uncertainty interval: The GBD study provided a 95% uncertainty interval (UI) for each estimate. This is achieved through simulation and predictive validation tests, ensuring that a reasonable range of uncertainty is provided even when data are scarce.

Multiple sampling: In the estimation process, GBD uses 1000 or more samples to generate a posterior distribution for each quantity. This approach allows the researcher to capture uncertainty due to incomplete data and model assumptions.

Refinement of models and methods

Spatio-temporal Gaussian process regression (ST-GPR) : This model uses smoothing and interpolation methods to deal with sparse data and reduce the uncertainty caused by insufficient data.

Causal mortality Ensemble model (CODEm) : improves performance on sparse data by exploring multiple possible model combinations combined with optimal nonsample validation results.

Data integration and calibration

Multi-source data integration: GBD studies integrate data from different sources, including medical records, surveys, surveillance systems, and others. In this way, the bias and uncertainty introduced by a single data source can be reduced.

Data calibration: The data were calibrated to eliminate systematic and random errors, thereby improving the accuracy of the estimates.

Transparency and replicability

Transparency of methodology: GBD studies emphasize transparency of methodology to ensure that findings can be validated and understood by other investigators.

Disclosure of data: The GBD project makes raw data, model input data points, and final results freely available to the extent permitted by law for analysis and validation by other investigators.

Through these methods, GBD research can effectively quantify the uncertainty of disease burden and provide more accurate and reliable information for public health decision-making.

In GBD studies, the spatio-temporal Gaussian process regression (ST-GPR) model reduces uncertainty by:

Using the properties of Gaussian processes

Probability distribution modeling: The ST-GPR model treats the function as a Gaussian process, which means that it can provide a probability distribution for the predicted values. This probabilistic modeling approach allows the researcher to directly quantify prediction uncertainty, for example by calculating confidence intervals for predicted values.

Application of covariance function: The model captures the correlation between data points through the covariance function (kernel function). This correlation modeling facilitates more accurate interpolation and extrapolation in regions where data are sparse, thereby reducing uncertainty due to insufficient data.

Data smoothing and interpolation

Smoothing: The ST-GPR model can smooth the data and reduce the influence of random fluctuations. This smoothing effect helps to provide more stable and reliable estimates in the case of sparse data.

Interpolation ability: The model uses the interpolation property of Gaussian process to perform smooth interpolation between known data points to fill data gaps. This allows more accurate burden estimates to be generated even in regions or time periods where data are scarce.

Advantages of Bayesian inference

Integration of prior information: As one of the Bayesian methods, ST-GPR can combine prior information with observed data for Bayesian inference. Prior information can be derived from expert knowledge, historical data, or other relevant studies, which helps to provide additional information support and reduce uncertainty in cases of insufficient data.

Hyperparameter adaptation: The hyperparameters of the model, such as the parameters of the covariance function, can be automatically adjusted to fit the characteristics of the data by methods such as maximizing marginal likelihood. This adaptive ability allows the model to fit the data better and improve the accuracy of the estimation.

Multi-source data integration

Integrating multiple data sources: The ST-GPR model can integrate data from different sources, such as dynamic population registries, surveys, censuses, etc. By integrating these data, the model is able to capture the changing trend of disease burden from multiple perspectives, reducing the bias and uncertainty that may be introduced by a single data source.

Through these mechanisms, the ST-GPR model effectively reduces the challenges caused by data sparsity and uncertainty in GBD research, and provides strong support for accurate assessment of global burden of disease.

2. " How does this manuscript relate to the Stroke GBD group who have produced stroke specific estimates and projections in the future. i.e what does the paper specifically add?"

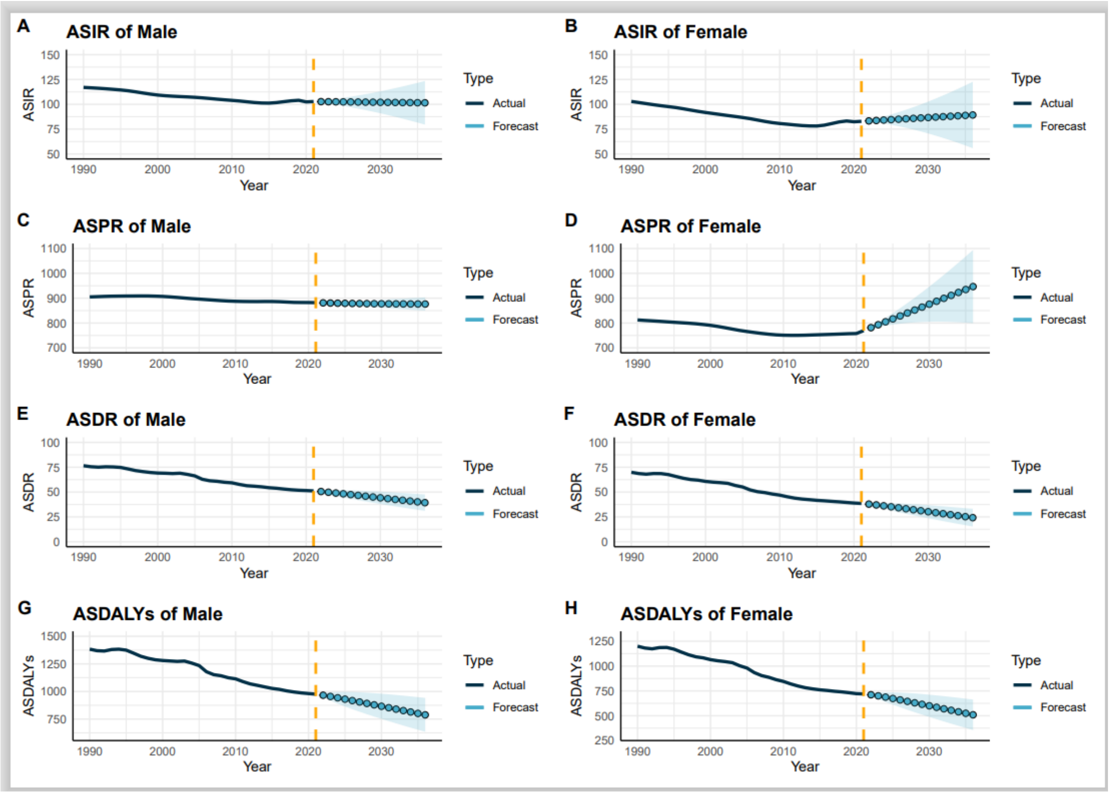
Different from the research of the stroke GBD research group, we analyzed the distribution of ischemic stroke in time, space, population and other aspects from a more specific perspective, and analyzed the distribution of major risk factors in different regions and different gender and age groups, and used frontier analysis to find countries with good and poor performance in the field of ischemic stroke prevention and treatment. We also used different prediction models to cross-validate the burden of ischemic stroke in different age groups and genders in the next 15 years, which would help to attract the attention of governments and guide them to take targeted measures.

3. "the paper is very data heavy and for clinicians, policy makers and stroke survivors does require attention to main questions and messages in less technical language."

We use less technical language as suggested by you to increase the readability of the article and make it more understandable to the audience.

4. "Regarding future projections it would be useful to have more on the assumptions made regarding incidence, risk factors and outcomes and effects of acute care."

For future projections, we added more assumptions about morbidity, risk factors, and outcomes and effects of acute care.



We add the ARIMA model for prediction and compare it with the BAPC model, and report the assumption basis and limitations of the two prediction models.

The ARIMA model was used to predict the same outcomes (incidence, prevalence, mortality, and DALYs) from 2021 to 2036, and the results were compared to the projections from the BAPC model. Both models produced consistent trends, suggesting robustness in the long-term predictions.

The Bayesian age-period-cohort (BAPC) model relies on several key assumptions that are critical to the validity and accuracy of the model:

Separability of age, period, and cohort effects

Basic Assumptions: The BAPC model assumes that time-related changes in health outcomes can be separated into age, period, and cohort effects. The age effect refers to the changes in an individual's life course as he or she ages. Period effects reflect the impact of events occurring within a given year on health outcomes; Cohort effect is the variation due to the different characteristics of the population in different birth years.

The prior distribution of the parameters

Use of prior information: In a Bayesian framework, the BAPC model allows prior information to be incorporated into the parameter estimation process. This means that for an unknown parameter in the model, an appropriate prior distribution needs to be specified for it. For example, a highly concentrated Laplace distribution can be used as a prior for parameters with linear effects, whereas a diffusive normal distribution can be used for slopes of age and period.
Similarity of neighboring effects
Adjacent effect Assumption: The BAPC model assumes that the adjacent effects of age, period, and cohort are similar in time. This means that the model considers the effect changes to be smooth and continuous between adjacent time points or age groups.
Poisson distribution assumption of the data
Poisson distribution: In some cases, the BAPC model assumes that the data follow a Poisson distribution. This assumption applies to count data, such as disease incidence or mortality, etc., and it allows the model to use Poisson regression models when dealing with such data.
These assumptions provide the theoretical basis and computational framework for the BAPC model, but also limit the scope of model application and the interpretation of results. In practice, careful consideration needs to be given to whether these assumptions are applicable to specific research contexts and data characteristics.
The assumptions of the BAPC model have significant impacts on the analysis results. Here is a detailed explanation:
Separability of Age, Period, and Cohort Effects
Identification of Influencing Factors: The assumption that age, period, and cohort effects can be separated allows the model to identify and quantify the roles of these three effects in health outcomes. For example, if the incidence of a certain disease increases over time, the model can isolate the period effect to determine whether this is due to increased diagnoses from medical advancements or an actual rise in disease occurrence due to environmental changes.
Result Interpretation: This assumption enables clearer interpretation of the results by distinctly attributing changes in health outcomes to different factors. For instance, it can clearly indicate that a high disease incidence in a specific age group is due to physiological changes unique to that age (age effect) or due to particular environmental conditions when that age group was born (cohort effect).
Prior Distributions of Parameters
Robustness of Estimation Results: The choice of prior distributions affects the estimation of parameters. If the prior information is accurate and consistent with the actual data, it can enhance the robustness and accuracy of the estimates. For example, using highly concentrated prior distributions can reduce the uncertainty in parameter estimation, bringing the results closer to the true values.
Subjectivity of Results: The setting of prior distributions may introduce subjectivity, as different priors can lead to different analysis results. This requires researchers to be cautious when setting priors and to provide sufficient justification for the choice of prior distributions.
Similarity of Adjacent Effects
Smoothness Constraint: Assuming similarity between adjacent effects essentially imposes a smoothness constraint on the changes in effects. This causes the model to tend to produce smooth curves in estimating effects, rather than abrupt fluctuations. For example, when estimating age effects, the model assumes that changes in incidence between adjacent age groups are gradual, not

sudden jumps.
Reasonableness of Results: In many cases, the smoothness constraint is reasonable because changes in health outcomes are often continuous. However, in some special situations, if the actual effects do indeed change dramatically, this assumption may prevent the model from accurately capturing the true pattern of effects.
Poisson Distribution Assumption for Data
Applicability Limitations: The Poisson distribution assumption is suitable for count data, such as disease incidence or mortality rates. If the data do not conform to the characteristics of the Poisson distribution (such as overdispersion or zero inflation), this assumption can lead to poor model fit, thereby affecting the accuracy of the analysis results.
Bias in Results: When data do not meet the Poisson distribution assumption, it can cause bias in the estimation results. For example, if data exhibit overdispersion, using Poisson regression models may underestimate the variance, thus affecting the significance tests of parameters and the reliability of prediction results.
In summary, the assumptions of the BAPC model significantly impact the analysis results, helping to identify and interpret the roles of different factors while also potentially limiting the model's applicability and accuracy in certain situations. In practical applications, it is necessary to test and adjust these assumptions based on specific data and research contexts to ensure the reliability and validity of the analysis results.
There are some limitations in the application of Bayesian age-period-cohort (BAPC) prediction model, which are analyzed as follows:
Difficulty in parameter estimation
Identification problems caused by linear relationships: There is a linear relationship between the three factors of age, period and cohort, which makes parameter estimation more difficult. The trend of change estimated using the APC model may even be reversed without certain conditional constraints.
Data dependency
Impact of data quality: The predictive accuracy of the BAPC model depends to some extent on the quality of the input data. If the data is missing, wrong or incomplete, it will affect the fitting effect of the model and the prediction result.
Limitations of data sources: In some studies, data may come from GBD and other databases, but these data themselves are estimated based on system dynamics model and statistical model, which has certain limitations.
Limitations of model assumptions
No support for inclusion of covariates: BAPC model does not support inclusion of covariates in the analysis process, which means that other influencing factors (such as socioeconomic factors and environmental factors) other than age, period and cohort cannot be considered, which limits the ability of the model to explain the complex causes of changes in disease burden.
Computational complexity
High computational resource demand: BAPC model usually adopts Bayesian Markov chain Monte Carlo (MCMC) algorithm for parameter estimation, which has high computational complexity and requires a large amount of computing resources and time.
Limitations of the prediction range
Uncertainty in long-term projections: Although the BAPC model can be used to predict future trends

in disease burden, the prediction results of the model may be more uncertain due to the greater uncertainty of future influencing factors when making long-term projections.

These limitations need to be considered and overcome when applying the BAPC model to improve the accuracy and reliability of the model.

To overcome the problem of data dependence of BAPC models, several strategies can be adopted:

Data quality improvement

Data cleaning and preprocessing: The original data were thoroughly cleaned to remove outliers, duplicate records, and inconsistent data to ensure the accuracy and integrity of the data.

Data calibration: Calibration of data to reduce systematic and random errors by comparison with other data sources or known facts.

Diversity of data sources

Integrating multiple data sources: In addition to traditional medical records and death registration data, multiple data sources such as censuses, socioeconomic surveys, and environmental monitoring can also be integrated to improve data comprehensivity and representativability.

Utilizing alternative data: In cases where some data is missing, alternative data can be sought to fill the gap. For example, sample survey data are used to estimate the overall picture.

Model refinement and validation

Model Hypothesis testing: The hypotheses in the model are rigorously tested to ensure that they are supported in the data. If some assumptions are found not to be true, the model structure should be adjusted in time.

Cross-validation: Cross-validation was used to validate the model and evaluate its performance on different datasets to ensure the stability and generalization ability of the model.

Advantages of the Bayesian approach

Utilization of prior information: The Bayesian approach allows prior information to be incorporated into the model, which can alleviate the problem of insufficient data to some extent. Additional information support for the model is provided by combining expert knowledge and historical data.

Uncertainty quantification: Bayesian methods are able to provide uncertainty quantification of parameter estimates, such as confidence intervals or probability distributions. This helps to assess the impact of data quality on model predictions.

Through these measures, the dependence of BAPC model on data can be effectively reduced, and the accuracy and reliability of model prediction can be improved.

The autoregressive integrated moving average (ARIMA) model is a statistical model widely used in time series analysis to predict future data points. Although the ARIMA model has been successful in many areas, it is based on a series of assumptions and suffers from some limitations. The following are the main assumptions and limitations of the ARIMA model:

The main assumptions of the ARIMA model

Stationarity:

Assumptions: Time series data are stationary, that is, their statistical properties (e.g., mean, variance) remain constant in time.

Limitations: Many real time series data are non-stationary, such as trends and seasonal variations in economic data. If the stationarity assumption is not met by the data, the predictive performance of the model may decrease significantly. Although it is possible to smooth the data through methods such as difference, excessive difference may lead to loss of information.

Linearity:
Assumption: The relationship in time series data is linear, that is, future values can be represented as linear combinations of past values.
Limitations: Many real time series data have nonlinear relationships, such as volatility clustering in financial markets. The inability of ARIMA models to capture these nonlinear features may lead to prediction bias.
Independent identically distributed (IID) error term:
Assumption: The error terms of the model are independently and identically distributed with zero mean and constant variance.
Limitations: In practice, the error term may not satisfy the assumption of independent identically distributed, such as the presence of autocorrelation or heteroscedasticity. This may lead to inaccurate estimation results of the model and affect the reliability of the prediction.
Model Order Selection:
Assumptions: The order of the model (p, d, q) is known or can be determined in some way.
Limitations: Selecting the appropriate model order is a complex process that usually relies on information criteria such as AIC, BIC, or methods such as cross validation. Inappropriate order selection may lead to overfitting or underfitting of the model, affecting the prediction performance.
Limitations of the ARIMA model
Limited capacity for non-stationary data:
The ARIMA model requires the data to be stationary, but many actual data are non-stationary. Although it is possible to smooth the data through methods such as difference, excessive difference may lead to loss of information and affect the predictive ability of the model.
Failure to capture nonlinear relationships:
Based on the assumption of linearity, ARIMA model cannot effectively capture the nonlinear relationship in time series. For data with nonlinear features, ARIMA models may not provide accurate predictions.
Sensitive to outliers:
The ARIMA model is sensitive to outliers in the data, which may lead to inaccurate model parameter estimation and thus affect the prediction results.
Model assumptions are strict:
The assumptions of the ARIMA model are strict, and it is often difficult for the actual data to fully meet these assumptions. For example, the assumption of independent identically distributed error terms may not hold in many cases, leading to inaccurate estimation results of the model.
The forecast is limited:
ARIMA model is mainly used for short-term prediction, but may not be effective for long-term prediction. This is because the model assumes that the statistical properties of the data remain constant over the forecast period, whereas the statistical properties of the actual data may change over time.
Model selection and parameter estimation are complex:
Selecting the appropriate ARIMA model order is a complex process that usually relies on methods such as information criteria or cross-validation. Inappropriate order selection may lead to overfitting or underfitting of the model, affecting the prediction performance.
Summary
The ARIMA model is a powerful tool for time series analysis, but its assumptions and limitations

need to be carefully considered when applied. For data that are non-stationary, nonlinear, or have outliers, it may be necessary to combine other models, such as seasonal ARIMA models, GARCH models, or machine learning methods, to improve the accuracy and reliability of prediction.

Compared with other prediction models, ARIMA model has the following advantages and disadvantages:

Advantages of the ARIMA model

The model is simple, easy to understand and apply: The structure of ARIMA model is relatively simple, and it is mainly modeled by three parts: autoregression (AR), difference (I) and moving average (MA), without the help of other exogenous variables.

High prediction accuracy: For short-term prediction, ARIMA model has high prediction accuracy and reliability, and can better capture the trend and seasonality of time series.

High maturity: ARIMA model has a solid theoretical foundation, and after long-term development and application, it has formed a set of relatively complete theoretical system.

Suitable for non-seasonal data: ARIMA model is suitable for time series data with no obvious seasonal characteristics, and can smooth non-stationary time series by difference operation.

Disadvantages of the ARIMA model

The requirement for data stationarity is high: ARIMA model requires time series to be stationary, or to be stationary after difference operation. If there is an obvious trend or seasonality in the data, it needs to be smoothed first, which may affect the prediction effect of the model.

Parameter selection is complex: ARIMA model needs to determine three parameters p (autoregressive order), d (difference order) and q (moving average order). Improper parameter selection may lead to overfitting or underfitting of the model.

Only linear relationships can be captured: The ARIMA model can intrinsically only capture linear relationships, and the prediction effect may not be good for data with nonlinear relationships.

The long-term prediction effect is poor: the ARIMA model will gradually decrease the prediction accuracy due to the accumulation of errors in the long-term prediction.

The comparison between ARIMA model and BAPC model in time series prediction is as follows:

ARIMA model

Advantages:

The model is simple, easy to understand and apply: The structure of ARIMA model is relatively simple, and it is mainly modeled by three parts: autoregression (AR), difference (I) and moving average (MA), without the help of other exogenous variables.

High prediction accuracy: For short-term prediction, ARIMA model has high prediction accuracy and reliability, and can better capture the trend of time series.

High maturity: ARIMA model has a solid theoretical foundation, and after long-term development and application, it has formed a set of relatively complete theoretical system.

Suitable for non-seasonal data: ARIMA model is suitable for time series data with no obvious seasonal characteristics, and can smooth non-stationary time series by difference operation.

Disadvantages:

The requirement for data stationarity is high: ARIMA model requires time series to be stationary, or to be stationary after difference operation. If there is an obvious trend or seasonality in the data, it needs to be smoothed first, which may affect the prediction effect of the model.

Parameter selection is complex: ARIMA model needs to determine three parameters p

(autoregressive order), d (difference order) and q (moving average order). Improper parameter selection may lead to overfitting or underfitting of the model.
Only linear relationships can be captured: The ARIMA model can intrinsically only capture linear relationships, and the prediction effect may not be good for data with nonlinear relationships.
The long-term prediction effect is poor: the ARIMA model will gradually decrease the prediction accuracy due to the accumulation of errors in the long-term prediction.
The BAPC model
Advantages:
Comprehensive consideration of multiple factors: BAPC model took into account the influence of age, period and cohort on the distribution of NCDS, which could more comprehensively analyze the trends of incidence and mortality of NCDS.
Unique advantages of the Bayesian framework: In the Bayesian framework, BAPC model can use both sample information and prior information to obtain unique parameter estimates, which makes the results more robust and reliable.
Improved prediction accuracy: Compared with the traditional APC model, the BAPC model can obtain more stable and reliable results and improve the prediction accuracy.
Solving the difficulty in parameter estimation: There is a linear relationship among the three factors of age, period, and cohort in the classical APC model, which may lead to difficulty in parameter estimation. The BAPC model effectively solves this problem through the Bayesian approach.
Flexible model selection: The BAPC model allows the flexible selection of different model components and prior probability distributions, which improves the applicability and flexibility of the model.
Advanced computational methods: The BAPC model uses the ensemble nested Laplace approximation (INLA) for posterior marginal distribution approximation, which avoids mixing and convergence problems and shows a relatively low error rate.
Automatic and simplified calculation: The use of the BAPC package for parameter estimation not only does not require complex coding procedures and convergence checks, but also allows flexible selection of age-standardized ratio (ASR) parameters and prior probability distributions to simplify the calculation process.
Predicting future changes: BAPC models can not only describe trends in age, period, and birth cohort, but also predict future changes based on these trends.
Disadvantages:
High computational complexity: The computational process of BAPC model is relatively complex and requires high computational resources and time.
High data requirements: BAPC models require sufficient data to estimate model parameters, and for cases with a small amount of data, the accuracy and reliability of the model may be affected.
There are many model assumptions: the BAPC model is based on a Bayesian framework and requires assumptions about the prior distribution, which may affect the prediction results of the model.
Comparison
Applicable scenarios:
ARIMA model: Suitable for time series with no or very weak seasonality. The effect was poor when there was obvious seasonality in the data.
BAPC model: suitable for disease prediction that requires comprehensive consideration of age,

period and cohort factors, especially in the field of public health and epidemiology.
Predictive power:
ARIMA model: performs well in short-term forecasting, but may not perform well for long-term forecasting and data with nonlinear relationships.
BAPC model: It can more accurately predict the change trend of diseases in different ages, periods and cohorts, and is suitable for long-term prediction and complex data structure.
Model complexity:
ARIMA model: The model structure is relatively simple, with few parameters, which is easy to understand and apply.
BAPC model: The model structure is complex and requires Bayesian estimation and complex calculations.
Data requirements:
ARIMA model: It has high requirements for the stationarity of the data and needs pre-processing such as difference.
BAPC model: Sufficient data are needed to estimate the model parameters, and high data volume and data quality requirements are required.

5. "the Figures are quite detailed and their interpretation in the text could be more detailed. the use of abbreviations needs to be spelt out too."

We would like to thank the experts for their recognition of our content richness. Due to space limitation, we have elaborated the table in more detail in the Supplementary materials section. As for abbreviations, we supplemented them in detail when they first appeared.

Thanks to the peer reviewers and editors for their patient and meticulous guidance, we have revised our manuscript according to the suggestions of the reviewers and provided two versions of the revised version. We hope editors and reviewers to see the important value of our research in guiding the formulation, organization and implementation of health policies and related comprehensive coordination work of national governments, and look forward to sharing our research results in such a high-level journal as your journal. We are also confident that it will become a highly cited article. We look forward to receiving your journal's acceptance and good news for the publication of our study.

Response to Reviewer#1 briefly:

We are extremely grateful for your valuable suggestions which have helped us improve our manuscript. We sincerely hope that our revisions and responses can be admired and approved by you, our expert. We also look forward to having our research published with your assistance. We wish you good health, smooth work and a happy Valentine's Day!

1. We abbreviated it as "Trend" following your instructions.
2. For the GBD-related research, most of them currently adopt SDI analysis because this indicator can comprehensively reflect indicators covering various dimensions such as economic income, education level, and female childbearing age to assess the development level of society. Compared with the information content of the World Bank's classification groups, it is more abundant. However, from a dialectical perspective, we cannot absolutely say which one is better. Therefore, we spent a lot of effort investigating and summarizing the advantages and disadvantages of both, and the suggestions from the reviewers have greatly inspired us. We not only added more literature explaining the richer information contained in SDI, but also added relevant analyses of the World Bank classification level, making our research the only one currently that simultaneously includes both types of analyses.
3. We included in the abstract the statement that the Sustainable Development Goals cannot be achieved by 2030 as per your instructions.
4. We referred to the guiding papers of major scientific organizations and added systematic suggestions as per your instructions.
5. We removed the lengthy discussion content following your instructions and added discussions on the latest literature.
6. We have commented on the uncertainty as follows following your instructions.
   More detailed etiological reports are helpful for guiding more precise treatment, but with the collection and collation of these data as well as their statistical analysis, some uncertainties inevitably arise. However, with a large amount of data and the GBD organization adopting more superior algorithms, these uncertainties will not cause fatal harm to the estimation of the results.

7.We have explained chronic geriatric diseases following your instructions.
“Chronic senile diseases” or "Chronic diseases of the elderly" usually refer to chronic noncommunicable diseases that are common and persist for a long time in the elderly population.

For detailed information, please refer to the following.

Reviewer #1 (Remarks to the Author):

This report is among the most comprehensive and updated recent reports on stroke

epidemiology world-wide. It is based on the 2021 GBD group data.

1.The title 'future trend predictions' could be shortened to 'trend predictions' or perhaps even 'trends'.

Epidemiology and trends of Ischemic stroke based on the Global Burden of Disease Study 2021: a systematic analysis for the global, regional, and national burden of Ischemic stroke and its risk factors,1990-2021

2.It is remarkable that the authors prefer to relate to the sociodemographic index rather than to the World Bank income group classification. Perhaps they can refer to any study showing that the SDI is more concise or informative?

Methods

The GBD 2021 analytical tools were applied to calculate ischemic stroke incidence, prevalence, deaths, disability-adjusted life-years (DALYs), and the population attributable fraction (PAF) of DALYs (with corresponding 95% uncertainty intervals [UIs]) associated with 12 risk factors, for 204 countries and territories from 1990 to 2021. We stratified epidemiological parameters by sex, age, the World Bank's classification of economies and sociodemographic index (SDI) levels and used the frontiers analysis to assess whether the burden of IS in each country matches its level of economic development.The Autoregressive Integrated Moving Average (ARIMA) and Bayesian age-period-cohort (BAPC) model was used to predict the burden of IS over the next 15 years.

Association with the standard World Bank classification of economies

In 2021, the lowest age-standardized DALYs were observed in high-income countries (385.13 per 100,000 population, [95% UI, 343.71 to 422.68]) and the highest in low-income countries(1053.94 per 100,000 population, [95% UI, 902.08 to 1254.22]). Similarly, the lowest ASDR were found in high-income countries (19.81 [95% UI, 16.78 to 21.39]), and the high ASDR were found in in low- income countries(55.1 [95% UI, 46.79 to 66.7]) and lower-middle-income countries(56.46 [95% UI, 50.69 to 61.91]) and upper-middle-income countries(56.39 [95% UI, 49.02 to 62.9]).The lowest ASIR were found in high-income countries (63.13 [95% UI, 56.04 to 70.56]), and the high ASIR were found in in low- income countries(104.96 [95% UI, 91.98 to 118.07]) and lower-middle-income countries(84.06 [95% UI, 72.86 to 95.63]) and upper-middle-income countries(118.35 [95% UI, 99.56 to 137.76]). The lowest ASPR were found in lower-middle-income countries (694.04 [95% UI, 635.91 to 750.41]), and the highest ASPR were found in in low- income countries(936.37 [95% UI, 888.99 to 980.65]).

To summarize, while more economically developed countries had the lowest burden of IS(measured in age-standardized DALYs ,ASDR and ASIR).

Table 2|Prevalence, Incidence, death rate, and disability-adjusted life years (DALYs) for ischemic stroke based on World Bank Income Group Classification in 2021 in age-standardized rates (ASRs) per 100 000, by Global Burden of Disease region, from 1990 to 2021 (Generated from data available at https://ghdx.healthdata.org/gbd-results-tool)

The sociodemographic index (SDI)

The Sociodemographic Index (SDI) is a composite indicator reflecting a country's development level based on lag-distributed income per capita, average educational attainment for individuals over 15 years, and the total fertility rate for those under 25 years.[6] The SDI ranges from 0 (minimal development) to 1 (maximal development).[53,54] Based on SDI values, the 204 countries and regions were categorized into five groups: high-SDI, medium-high-SDI, medium-SDI, medium-low-SDI, and low-SDI regions.[54]

World Bank Income Group Classification

Countries were classified by economic development according to the standard World Bank classification of economies (high-, upper-middle-, lower-middle-, and low-income countries).[55]

3.It is important that the authors refer to the UN statement that shows stroke a priority target for reduction and that- following the trends- the SDG goals 2030 will not be reached by 2030. Perhaps they should put this already in the abstract.

Conclusion

The disease burden of IS varies greatly among regions and populations, and DALYs in medium-high SDI countries are still high. The prevention and control measures taken so far for the reduction of IS are not enough to meet the Sustainable Development Goal targets by 2030.Countries should formulate prevention and control measures suitable for their national conditions based on the main risk factors, and learn from the experience of countries with good IS burden control.

4.While measures for prevention should be based on epidemiological data and their trends, the recommendations given are isolated and not systematic. The authors should refer to guideline papers from leading scientific organisations (eg Int J Stroke 2023; 18: 499-531)

been able to assess the burden of IS in a more comprehensive and objective way, and we will make this study more detailed in the future.

By analyzing the temporal, spatial and population distribution, trend prediction and risk factor analysis of ischemic stroke worldwide based on GBD2021, we hope our study can provide reliable basis for national health policy makers to formulate targeted strategies.Our study provides targeted and innovative but fragmented recommendations for the prevention and treatment of stroke burden based on the latest GBD database. In order to make our guidance more comprehensive and systematic, we combined with the World Stroke Organization to summarize the recommendations of existing stroke guidelines for the management of stroke patients. The recommendations addressed pre-hospital, emergency, and acute hospital care. Strong recommendations were made for reperfusion therapies for acute ischemic stroke. For secondary prevention, strong recommendations included establishing etiological diagnosis; management of hypertension, weight, diabetes, lipids, and lifestyle modification; and for IS, management of atrial fibrillation, valvular heart disease, left ventricular and atrial thrombi, patent foramen ovale, atherosclerotic extracranial large vessel disease, intracranial atherosclerotic disease, and antithrombotics in non-cardioembolic stroke. For rehabilitation, there were strong recommendations for organized stroke unit care, multidisciplinary rehabilitation, task-specific training, fitness training, and specific interventions for post-stroke impairments.[51]

51 Mead, G. E. *et al.* A systematic review and synthesis of global stroke guidelines on behalf of the World Stroke Organization. *Int J Stroke* **18**, 499-531 (2023). https://doi.org/10.1177/17474930231156753

5.A larger part of the discussion compares this paper to a recently published similar paper from the Shanghai group (Li et al. 2024). The arguments made are somewhat lengthy, even redundant. This part could be greatly reduced to perhaps a few sentences or one smaller paragraph.

Another study about global, regional, and national temporal trends of diet-related ischemic stroke mortality and disability from 1990 to 2019 evaluated six specific dietary factors, including sodium, red meat, fiber, vegetables, whole grains, and fruits to determine their individual and joint impact on ischemic stroke. They found high sodium diet was still a key driver of diet-related ischemic stroke. Besides, they found the developing countries with weak social and economic foundations were in the face of bigger challenges due to the ongoing burden of diet-related strokes compared with developed countries. In our study, the percentage of dietary risk in different SDI regions ranked from High to Low as High−middle SDI, Middle SDI, Low SDI, High SDI, and Low−middle SDI. This may be due to the fact that high-income countries pay more attention to healthy eating after food freedom, while developing countries are still in the stage of upgrading the level and quantity of diet, and do not pay attention to the quality and health impact of diet.[20]

6.Line 613: the authors recommend a more detailed reporting of stroke according to the etiology. They should also comment about the likely uncertainties that come with it.

More detailed etiological reports are helpful for guiding more precise treatment, but with the collection and collation of these data as well as their statistical analysis, some uncertainties inevitably arise. However, with a large amount of data and the GBD organization adopting more superior algorithms, these uncertainties will not cause fatal harm to the estimation of the results.

The uncertainties are manifested in the following aspects. Differences in diagnostic criteria: Different regions and medical institutions may adopt different diagnostic criteria and methods to diagnose stroke and its subtypes, which may lead to inconsistent data. For example, in some regions, there may be a lack of advanced imaging equipment to accurately distinguish ischemic stroke from hemorrhagic stroke. Limitations of data sources: The GBD study relies on multiple data sources, including hospital records, death certificates, and epidemiological investigations. The completeness and accuracy of these data may vary by region, thereby affecting the reporting of etiology. Demographic and geographical differences: The distribution of etiologies among different populations: The etiology distribution of stroke may be significantly different in different regions and populations. For example, in some low-income countries, stroke may be more associated with hypertension and infectious diseases, while in high-income countries, it may be more associated with atherosclerosis and lifestyle factors (such as smoking, diet). These differences lead to uncertainties when comparing etiology in different regions. Urban-rural differences: There are differences in medical resources and diagnostic capabilities between urban and rural areas, which may lead to deviations in etiological data. For example, urban areas may be more inclined to report stroke causes related to lifestyle, while rural areas may report more causes related to insufficient medical resources.

7.Line 647: please explain ‘chronic senile diseases’

“Chronic senile diseases” or "Chronic diseases of the elderly" usually refer to chronic noncommunicable diseases that are common and persist for a long time in the elderly population. These diseases are usually characterized by:

Definition and Characteristics

Long-term: Chronic diseases of the elderly usually have a long course and require long-term management and treatment. For example, diseases such as hypertension, diabetes, and coronary heart disease often require ongoing medical interventions and lifestyle modifications once they occur.

Multifactorial disease: The etiology of these diseases is complex and usually related to multiple factors, including genetics, lifestyle, and environmental factors. For example, senile osteoporosis is associated with changes in hormone levels, malnutrition, lack of exercise, and other factors.

Atypical symptoms: In older adults, symptoms of chronic diseases may be less pronounced or typical than in younger adults. For example, pneumonia in the elderly may present with atypical symptoms and be more severe.

Common types

Cardiovascular diseases, such as hypertension, coronary heart disease and stroke, are among the most common chronic diseases in the elderly.

Metabolic diseases: such as diabetes, the efficiency of insulin secretion and action decreases with age, leading to an increased incidence of diabetes.

Respiratory diseases, such as chronic obstructive pulmonary disease (COPD), are closely linked to factors such as long-term smoking and air pollution.

Musculoskeletal diseases: such as osteoporosis and osteoarthritis, which cause weak bones and joint pain that affect mobility in older people.

Prevention and management

Healthy lifestyle: maintaining healthy dietary habits, moderate exercise, smoking cessation and alcohol consumption are essential for the prevention and management of chronic geriatric diseases.

Regular physical examination: Older adults should have regular physical examinations for early detection and intervention of chronic diseases.

Comprehensive management: The management of chronic geriatric diseases requires comprehensive consideration of drug therapy, rehabilitation training, psychological support and other factors.

Prevention and treatment of chronic diseases in the elderly is one of the important measures for elderly health care, which is of great significance for improving the quality of life and prolonging the healthy life span of the elderly.

Dear Reviewer,

Thank you very much for your careful and patient review of our article and your valuable suggestions for improving our article. I wish you get higher achievements in scientific research and more amazing discoveries. Thank you very much for approving our opinions on revising the article and giving us the opportunity to publish. I wish you a smooth work, a sweet life like honey, every day can meet surprises and beauty!

Best wishes for you!