Geographical variation in dementia prevalence across China: a geospatial analysis

Yixuan Liu,[a,c] Xinyuan Gao,[a,c] Yongjin Zhang,[a,c] Minrui Zeng,[a,c] Yuyang Liu,[a,b] Yanjuan Wu,[a,c] Weihua Hu,[d] Yingsi Lai,[a,c,e,f,**] and Jing Liao[a,c,e,f,*]

[a]Department of Medical Statistics, School of Public Health, Sun Yat-sen University, Guangzhou, Guangdong 510000, China
[b]Shenzhen Health Development Research and Data Management Center, Shenzhen, China
[c]Sun Yat-sen Global Health Institute, Institute of State Governance, Sun Yat-sen University, Guangzhou, China
[d]Peking University Center for Public Health and Epidemic Preparedness & Response, Beijing 100871, China

The Lancet Regional Health - Western Pacific 2024;47: 101117

Published Online xxx

https://doi.org/10.1016/j.lanwpc.2024.101117

Summary

Background Dementia poses great health and social challenges in China. Dementia prevalence may vary across geographic areas, while comparable estimations on provincial level is lacking. This study aims to estimate dementia prevalence by provinces across China, taking into account risk factors of individual level and potential spatial correlation of provinces.

Methods In this study, 17,176 adults aged 50 years or older were included from the fourth wave of the China Health and Retirement Longitudinal Study (CHARLS 2018), covering 28 provinces, autonomous regions and municipalities. To improve provincial representativeness, we constructed provincial survey weights based on China 7th census (2020). The prevalence of dementia and 95% Bayesian credible intervals (BCIs) were estimated using a Bayesian conditional autoregressive (CAR) model with spatially varying coefficients of covariates.

Findings The weighted prevalence of dementia at provincial level in China in 2018 ranged from 2.62% (95%BCI: 1.70%, 3.91%) to 13.53% (95%BCI: 8.82%, 20.93%). High dementia prevalence was concentrated in North China, with a prominent high–high cluster, while provinces of low prevalence were concentrated on East and South China, characterized by a low–low cluster. Ordered by the median estimation of prevalence, the top 10% of provinces, include Xinjiang, Jilin, and Beijing. Meanwhile, Fujian, Zhejiang, and Guangdong rank among the last. The association between dementia prevalence and drinking, smoking, social isolation, physical inactivity, hearing impairment, hypertension, and diabetes exhibits provincial variation.

Interpretation Our study identifies a geospatial disparity in dementia prevalence and risk factor effects across China's provinces, with high–high and low–low clusters in some northern and southern provinces, respectively. The findings emphasize the need for targeted strategies, such as addressing hypertension and hearing impairment, in specific regions for more effective dementia prevention and treatment.

Funding National Science Foundation of China/the Economic and Social Research Council, UK Research and Innovation joint call: Understanding and Addressing Health and Social Challenges for Ageing in the UK and China. UK-China Health And Social Challenges Ageing Project (UKCHASCAP): present and future burden of dementia, and policy responses (grant number 72061137003, ES/T014377/1).

Keywords: Dementia; Prevalence; Geographical analysis; China

*Corresponding author. Department of Medical Statistics and Epidemiology, School of Public Health, Sun Yat-sen University, No.74 Zhongshan 2nd Road, Guangzhou, 510030, China.
**Corresponding author. Department of Medical Statistics and Epidemiology, School of Public Health, Sun Yat-sen University, No.74 Zhongshan 2nd Road, Guangzhou, 510030, China.
E-mail addresses: liaojing5@mail.sysu.edu.cn (J. Liao), laiys3@mail.sysu.edu.cn (Y. Lai).
[e]Senior authors.
[f]Contributed equally to this work and should be considered co-corresponding authors.

Research in context

Evidence before this study

We systematically reviewed studies published between 1st January 2020, and 31st December 2023 on geographical variation on dementia prevalence in China, searching PubMed Database using terms "China" or "Chinese", "dementia", "geographical", "variation", and "province", with no language restrictions. We searched previous meta-analyses have revealed a distinctive regional pattern, with higher prevalence observed in the north and west, and comparatively lower prevalence in the south and east. However, these findings heavily relied on studies from single province, introducing significant heterogeneity in research methods and dementia diagnostic criteria. There are a few multicenter large-scale studies on dementia prevalence in recent years, while their focus remained at the regional level, neglecting nuanced variations within regions. Moreover, previous studies did not use geospatial models to account for the spatial correlation of provinces when estimating dementia prevalence. Future study employing a standardized approach and a consistent dementia definition is needed to provide a more accurate and detailed estimation on geospatial variation in dementia prevalence across China.

Added value of this study

Our study estimated provincial dementia prevalence, considering the distribution and impact of risk factors across provinces. We employed Bayesian geospatial analyses incorporating both random and fixed effects of covariates to enhance the precision of our findings, and applied a prior validated dementia case algorithm to a nationwide ageing cohort, China Health and Retirement Longitudinal Study, with provincial sampling weights constructed from the latest China Population Census. We confirm the regional distribution pattern of weighted dementia prevalence in previous systematic reviews, with higher prevalence in the north and west. Regarding the geospatial distribution pattern of provincial dementia prevalence, the top 10% of provinces, ordered by the median estimation of prevalence, include Xinjiang, Jilin, and Beijing. Meanwhile, Fujian, Zhejiang, and Guangdong rank among the last. Furthermore, provinces exhibiting significantly high–high cluster dementia prevalence are concentrated in North China, while provinces with low–low cluster prevalence are clustered in East and South China.

Implications of all the available evidence

Our study provides estimation of provincial dementia prevalence by accounting for the spatial autocorrelation, and the distribution and effects of key established risk factors across provinces. Our findings underscore a crucial imperative for provinces in North China to intensify their focus on dementia prevention and treatment strategies.

Introduction

Dementia poses great health and social challenges in China.[1] It has been estimated that among the 51.6 million individuals living with dementia worldwide in 2019, over a quarter (13.1 million) are in China,[2] the burden of which is expected to escalate faster with the ageing of the Chinese population.[3]

Dementia burden is likely to vary across China. Previous meta-analyses[4–7] indicate a regional distribution pattern of dementia prevalence decreases from northern, central and southern China and with a particularly high estimate in western China. These estimations, however, were mainly based on studies conducted in a single province, greatly affected by heterogeneity in research methods and dementia diagnostic criteria.[8] Despite few multicentre large-scale conducted in recent years,[9–11] they only provided regional level estimates, overlooking internal heterogeneity and neglecting the exploration of geographical variation across provinces. Provincial dementia prevalence estimated with comparable research method and a consistent definition is still lacking.

Geographical variations on dementia prevalence may attribute to differences in established risk factors of dementia.[12] Different distributions of some risk factors have been found, such that diabetes,[13] hypertension[14] and obesity[15] tend to be more prevalent in north China than in south, while lifestyle-related factors (i.e. smoking and alcohol consumption) seem to be highest in northeastern China.[16] Less is known whether the strength of the associations between these risk factors and dementia prevalence would vary by provinces. There were significant socioeconomic disparities in the risk of dementia,[17] coupled with varying levels of socioeconomic development among provinces, indicate the possible existence of differences in dementia risk factors at the provincial level. Therefore, accounting for the provincial impact of risk factors on dementia and the spatial correlation of provinces is necessary to estimate provincial dementia prevalence.

To address the above-mentioned research gaps, we aimed to estimate the provincial dementia prevalence in China with a consistent dementia definition algorithm, accounting for the key established risk factors of individual level and potential spatial correlation of provinces. Specifically, we systematically considered the prevalence of the key risk factors and their effects on dementia at provincial level, ultimately exploring the spatial distribution of unmeasured factors' effects. Exploring provincial variations in dementia prevalence would assist policymakers and public health managers in assessing the disease burden across provinces and facilitates the allocation of medical resources.

Methods

Study population

We used data from the wave 4 in 2018 of the China Health and Retirement Longitudinal Study (CHARLS). CHARLS is a nationally representative longitudinal survey of Chinese aged 45 and older, whose sample was obtained via multistage stratified probability proportional to size (PPS) sampling. The CHARLS 2018 covered 28 provinces, municipalities and autonomous regions, 150 counties/districts, 450 communities/villages, involving 19,816 individuals. Due to the high prevalence of mid-life dementia-related risk factors (such as diabetes and hypertension, etc), we included individuals aged 50 and above in our analysis in line with prior studies[18] and to maintain an adequate sample size (n = 17,176). Variables with missing values are hypertension, diabetes, depressive symptoms, hearing impairment, smoking, drinking status, social isolation and physical inactivity; and the proportions of participants with missing values for all these variables are less than 1%. This study comprised 17,176 individuals aged 50 years old and above excluded the participants with all covariates missing (Appendix A). The provinces and municipalities in mainland China were categorized into seven geographical divisions based on the seven administrative geographical divisions: Northeast, North, Central, East, South, Northwest and Southwest (Appendix B). We followed the STROBE checklist for data extraction (Appendix C).

Ethics statement

This work was based on CHARLS, which was originally approved by the Ethical Review Committee of Peking University in June 2008 (IRB00001052-11015), and all participants signed informed consent at the time of participation. All data in this study were aggregated at areal level and did not contain any identifiable information at the individual or household levels. So, there are no specific ethical issues warranted.

Definition of dementia

Dementia was defined by an algorithmic case definition based on coexistence of cognitive impairment (CI) and functional impairment (FI), or a report of a doctor's diagnosis of dementia or memory-related disease by the participant or caregiver, in line with prior IMPACT-BAM studies.[19,20] The algorithmic case definition conforms to Diagnostic and statistical manual of mental disorders, Fifth Edition (DSM-5) and International Statistical Classification of Diseases and Related Health Problems (ICD-10) criteria in that it hinges on non-transient impairment in two or more cognitive domains, and has been validated against clinical diagnosis in COAST, a multisite cross-sectional survey.[21] If impaired domains of cognitive function were two or more, it was defined as CI. Common latent cognitive domains (i.e. orientation, memory, executive function and language) were identified via confirmatory factor analysis on cognitive tests of CHARLS and the Chinese Longitudinal Healthy Longevity Survey (CLHLS).[21] These cognitive tests are Telephone Interview of Cognitive Status (TICS) in CHARLS and Chinese Mini-Mental State Examination (C-MMSE) in CLHLS, which are widely used in many studies and have been modified to reduce literacy-related items.[22] Impairment in each cognitive domain was defined as a factor score of 1•5 standard deviations or lower below the mean of the population aged 50 years and above with the same level of education.[23] Based on the Katz-scale,[24] FI is defined as an inability to do one or more basic activities of daily living (BADL) requiring some form of caregiving for bathing, getting in or out of bed, changing clothes, using the toilet, cutting food and eating, controlling urination and bowel movements. If participants were incapable to complete the CI and FI tests, a doctor's diagnosis of dementia or memory-related disease reported by the participant or caregiver was used.

Demographic characteristics and key risk factors

We considered demographic characteristics, such as age and sex in the later analysis. Consistent with the Lancet Commission report,[25] 9 established dementia risk factors (less education, hypertension, diabetes, depressive symptoms, hearing impairment, smoking status, drinking status, social isolation and physical inactivity) were included. If the highest level of education attained by the respondent was below primary school, it was classified as "less education". Hypertension and diabetes were defined as having been diagnosed by a medical professional. Depressive symptoms were assessed using the 10-question version of the Center for Epidemiological Studies Depression Scale (CES-D), with scores ≥10 indicating depressive symptoms.[26] Hearing impairment was determined by participants self-reporting a hearing problem. Smoking was categorized as "never smoking", "former smoking", or "current smoking", based on participants' smoking history and current smoking behaviours (reference: never smoking). Drinking was classified as "never drinking", "drinking less than or equal to once a week", or "drinking more than once a week" (reference: drinking less than or equal to once per week), based on the frequency of alcohol consumption in the past year. Social isolation was defined as not participating in any social activities in the past month or not at all. Physical inactivity was defined as engaging in no vigorous, moderate, or mild physical activity for at least 10 min per week. Details can be found in Appendix D. Our analysis only included 9 out of 12 risk factors examined in the Lancet Commission[25] due to missing certain risk measures for obesity, Traumatic Brain Injury (TBI) and air pollution in the fourth wave of CHARLS. These 9 risk factors are at individual level. Specifically, physical examination was not conducted in this wave, thus height, weight measurements, blood pressure and fasting plasma glucose (FPG) values of participants were absent. There is no TBI-related question in CHARLS, indirect questions (e.g. do you feel head pain) as used in Chen (2023)[12] may nevertheless incur overestimation of TBI risk. Furthermore, the cohort did not include indicators related to air pollution at individual level, while substituting the provincial average air quality for individual exposures is not accurate.[27]

Statistical analysis

To estimate the provincial prevalence of dementia while accounting for risk factors and potential spatial correlation of provinces, we developed a Bayesian hierarchical regression model[28] with spatial-specific random effects.[29] We assumed outcome variable Y_{ik} (where $Y_{ik} = 1$ if individual k in province i suffered from dementia, otherwise $Y_{ik} = 0$) to follow a Bernoulli distribution with parameter π_{ik}, representing the probability of dementia of individual k in province i: $Y_{ik} \sim Bernoulli(1, \pi_{ik})$, $i = 1, \ldots, 28$.

We modelled covariates with the logistic hierarchical regression as $logit(\pi_{ik}) = \alpha + \boldsymbol{\beta X} + \sum_{k=1}^{K} f_s(\mu_{ik} X_{ik}) + u_i + v_i$. Here, covariates $\boldsymbol{X}$ is a matrix of explanatory spatial covariates and $\boldsymbol{\beta}$ is vector of fixed regression coefficients which are constant across all provinces. $f_s(\mu_{ik} X_{ik})$ is random slope of the covariates, subject to the standard normal distribution, representing the random effects of covariates in different provinces. Particularly, we assumed spatially structured random effects u_i followed a conditional autoregressive (CAR)[30] process as $u_i | u_{j, j \neq i} \sim N\left(\frac{1}{w_{ij}} \sum_{j=1}^{n_i} w_{ij} u_{ij}, \frac{\sigma_u^2}{n_i}\right)$. This represents the part of spatial correlation among provinces. w_{ij} is the element of spatial weights matrix of order 28*28 (28 is the number of provinces considered in this study). We set $w_{ij} = 1$ if province i and j are neighbouring, otherwise $w_{ij} = 0$, which is the most common and applicable setting in CAR model.[31] In addition, we conducted sensitivity analysis using a spatial weights matrix based on centroid distances of provinces, assuming that the effect diminishes with increasing centroid distance.[32] Results showed little impact of the spatial weight matrix selection on the findings, indicating the robustness of our setting. And $v_i \sim N(0, \sigma_v^2)$ is unstructured random effects following the independent identically distribution (IID) without spatial correlation. The sum of v_i and u_i is the random effects of unmeasured factors on dementia.

Based on the above, we formulated following three models. The model 1 does not conclude any covariates but considers the spatial correlation of provinces. The model 2 does not incorporate $\sum_{k=1}^{K} f_s(\mu_{ik} X_{ik})$, and model 3 extended this by additionally addressing the diverse provincial effects of risk factors on dementia. We chose the final model based on the deviance information criterion (DIC) value and pD (the number of effective parameters in model).

To identify geographic clusters of low-or high-prevalence provinces, Moran's test for global and local spatial autocorrelation analyses were used. The global indicator Moran's I statistic[33] reflects the similarity of attribute values of neighbouring spatial units in the entire space. The closer the absolute value is to 1, the closer the relationship between spatial units' attributes. The local Moran's I statistic[34] is a local indicator used to describe spatial relationships, where a positive value indicates a positive correlation between the attribute values of a spatial unit and its neighbouring units, while a negative value indicates a negative correlation. We identified provinces with high–high cluster, low–low cluster, high-low cluster and low-high cluster. Among them, the high–high and low–low clustering types indicate that the level of prevalence of dementia presents a strong positive correlation and spatial cluster, while the low-high and high-low clustering types indicate a strong negative correlation and spatial differentiation.

Bayesian regression analyses were implemented in the R-INLA (Integrated Nested Laplace Approximation) package.[35] The posterior distributions provide estimates of the parameters and their uncertainty (Bayesian credible interval, BCI), allowing for probabilistic inference. Other parameters and hyperparameters in the model are assigned weak informative prior distributions: $\log(\sigma_\alpha^2) \sim$ logGamma (1, 0.1), $\log(1/\sigma_\beta^2) \sim$ log Gamma (1, 0.1), $\log(1/\sigma_u^2) \sim$ logGamma (1, 0.05), $\log(1/\sigma_v^2) \sim$ logGamma (1, 0.05). The precision hyperparameter σ^2 controls the amount of variability. We used a stratified random forest method for multiple imputation of missing covariates in the data, and conducted sensitivity analyses to examine the robustness of our findings without imputation. The random forest method was employed for each of the missing covariates as the dependent variable, while including independent variables or predictors related to demographic characteristics and key risk factors within the dataset.[36] The imputation process iterated until reaching the maximum number of iterations(default value of 5).[37] These multiple imputed datasets were then pooled using Rubin's rules,[38] which involve combining estimates from each imputed dataset and adjusting standard errors to account for variability between imputations. Mean Squared Error (MSE) was calculated, which is closer to 0, the more effective and accurate the imputation method is. To obtain provincially representative prevalence, we constructed provincial sampling weight $weight_m$ (m = 1,2, ...,171,76) based on nationally representative sampling weight provided by CHARLS.[39] The formula of the weight is $weight_m = pweight_{ijkl} * ind\_weight_m$, where $pweight_{ijkl}$ is provincial post-stratification weight, ind_weight_m is the sampling weight adjusting age, sex and individual non-response corrections in CHARLS. Given CHARLS sampling weight is age specific, our study sample remained representative of the targeted age group (50 years and above) after excluding participants below age 50 (15%). The provincial post-stratification weight[40] based on population size by age group, sex, and type of residence in the sample and 7th China census (2020) from each province. The formula for calculating the provincial post-stratification weight is as $pweight_{ijkl} = \frac{N_{ijkl}/N_i}{n_{ijkl}/n_i}$, N_{ijkl} and n_{ijkl} are the population of age group j (j = 1,2, ...,11), the indicator for sex k (k = 1,2), and the residence l (l = 1,2)of province i (i = 1,2, ...,28) in the 7th census and CHARLS 2018 data, respectively. N_i and n_i are the total population of the province i in the 7th census and CHARLS 2018 data respectively. We integrated the individual probability of dementia in the posterior sample with the provincial sampling weight to derive provincial dementia prevalence.[40]

All data management and regression analyses were all based on the open-source R software.[41] (version 4.2.2).

Role of the funding source

The funders had no role in the design, data collection, data analysis and writing of this study.

Results

A total of 17,176 participants (51.69% women) aged 50 years or older were included in the analysis. We identified a total of 968 patients with dementia. Table 1 shows the demographic characteristics of the seven geographic regions. The spatial distribution of the nine risk factors can be found in the Appendix E. Comparison of the three models based on criteria (Appendix F), with model 1 having the poorest overall fit but the lowest complexity. Model 3, incorporating both fixed effects and spatial non-stationarity, demonstrates the best absolute fit despite a higher number of effective parameters. Maps and graphs presented reflect the outcomes of model 3 with minimal DIC.

The map displaying the estimated dementia prevalence corresponding to final model is shown in Fig. 1a. In 2018, the prevalence across the 28 provinces ranged from 2.62% (95%BCI: 1.70%, 3.91%) to 13.53% (95% BCI: 8.82%, 20.93%) (see Appendix G). Regarding the geospatial distribution pattern of provincial dementia prevalence, the top 10% of provinces, ordered by the median estimation of prevalence, include Xinjiang, Jilin, and Beijing. Meanwhile, Fujian, Zhejiang, and Guangdong rank among the last. Provinces with the highest prevalence of dementia were predominantly situated in Northwest, North, and Northeast China, while those with the lowest prevalence were concentrated in East and South China (Fig. 1a). Utilizing spatial autocorrelation analysis (local Moran analysis, $P < 0.05$), we identified provinces exhibiting positive correlation and spatial clustering (high–high cluster), as well as those with negative correlation and spatial clustering (low-high and high-low clusters). Regions with significantly high–high cluster dementia prevalence were primarily

Characteristic	Overall	Northeast	North	East	South	Central	Northwest	Southwest	P value
	(n = 17,176)	(n = 2005)	(n = 1475)	(n = 5300)	(n = 1460)	(n = 2633)	(n = 1294)	(n = 3009)	
Age, years (Mean ± SD)	63.6 ± 9.33	62.5 ± 8.83	63.1 ± 8.98	64.1 ± 9.48	64.0 ± 9.61	63.0 ± 9.10	62.6 ± 8.98	64.2 ± 9.60	<0.001
Women	8860 (51.6%)	1060 (52.9%)	742 (50.3%)	2747 (51.8%)	771 (52.8%)	1339 (50.9%)	664 (51.3%)	1537 (51.1%)	0.65
Type of residence									<0.001
Urban	6853 (39.9%)	1067 (53.2%)	460 (31.2%)	2174 (41.0%)	710 (48.6%)	975 (37.0%)	473 (36.6%)	994 (33.0%)	
Rural	10,323 (60.1%)	938 (46.8%)	1015 (68.8%)	3126 (59.0%)	750 (51.4%)	1658 (63.0%)	821 (63.4%)	2015 (67.0%)	
Less education	7236 (42.1%)	563 (28.1%)	421 (28.5%)	2556 (48.2%)	565 (38.7%)	943 (35.8%)	537 (41.5%)	1651 (54.9%)	<0.001
Smoking status									<0.001
Never smoking	9675 (56.3%)	1021 (50.9%)	774 (52.5%)	3052 (57.6%)	859 (58.8%)	1534 (58.3%)	726 (56.1%)	1709 (56.8%)	
Former smoking	2839 (16.5%)	372 (18.6%)	256 (17.4%)	927 (17.5%)	207 (14.2%)	427 (16.2%)	209 (16.2%)	441 (14.7%)	
Current smoking	4662 (27.1%)	612 (30.5%)	445 (30.2%)	1321 (24.9%)	394 (27.0%)	672 (25.5%)	359 (27.7%)	859 (28.5%)	
Drinking status									<0.001
Never drinking	11,470 (66.8%)	1294 (64.5%)	995 (67.5%)	3416 (64.5%)	984 (67.4%)	1788 (67.9%)	943 (72.9%)	2050 (68.1%)	
Drinking ≤1/week	2609 (15.2%)	315 (15.7%)	249 (16.9%)	731 (13.8%)	257 (17.6%)	404 (15.3%)	279 (21.6%)	374 (12.4%)	
Drinking >1/week	3097 (18.0%)	396 (19.8%)	231 (15.7%)	1153 (21.8%)	219 (15.0%)	441 (16.7%)	72 (5.6%)	585 (19.4%)	
Hearing impairment	2707 (15.8%)	345 (17.2%)	237 (16.1%)	762 (14.4%)	176 (12.1%)	419 (15.9%)	198 (15.3%)	570 (18.9%)	<0.001
Hypertension	7046 (41.0%)	924 (46.1%)	695 (47.1%)	2244 (42.3%)	450 (30.8%)	1015 (38.5%)	550 (42.5%)	1168 (38.8%)	<0.001
Diabetes	2364 (13.8%)	304 (15.2%)	248 (16.8%)	803 (15.2%)	150 (10.3%)	376 (14.3%)	171 (13.2%)	312 (10.4%)	<0.001
Social isolation	9624 (56.0%)	1085 (54.1%)	730 (49.5%)	3036 (57.3%)	756 (51.8%)	1316 (50.0%)	828 (64.0%)	1873 (62.2%)	<0.001
Physical inactivity	1809 (10.5%)	213 (10.6%)	192 (13.0%)	680 (12.8%)	120 (8.2%)	238 (9.0%)	116 (9.0%)	250 (8.3%)	<0.001
Depress symptoms	6449 (37.5%)	734 (36.6%)	566 (38.4%)	1686 (31.8%)	521 (35.7%)	979 (37.2%)	622 (48.1%)	1341 (44.6%)	<0.001
Dementia	968 (5.6%)	164 (8.2%)	103 (7.0%)	227 (4.3%)	59 (4.0%)	135 (5.1%)	99 (7.7%)	181 (6.0%)	<0.001

Table 1: Characteristics of study participants in CHARLS 2018.

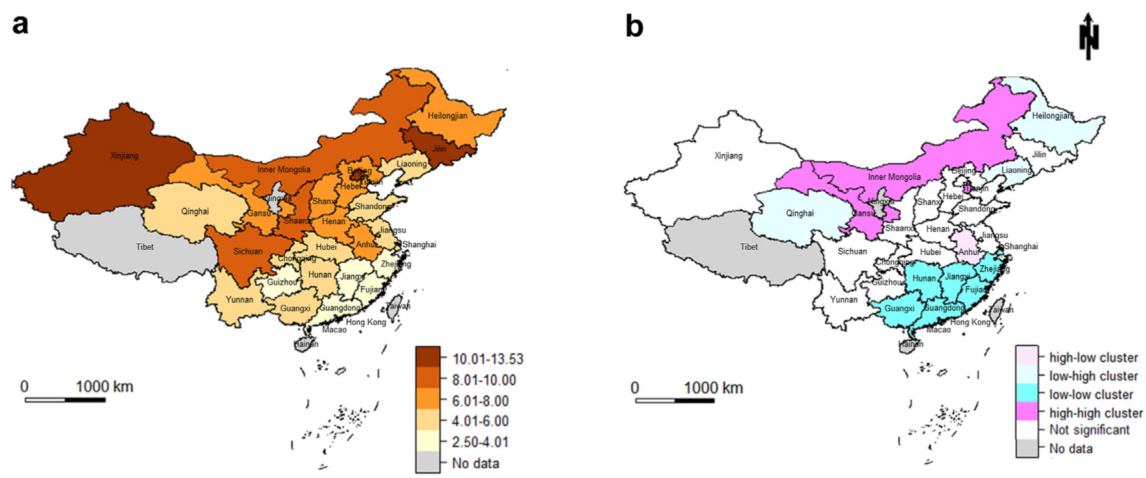

Fig. 1: Model-based weighted prevalence maps of dementia across provinces. Estimated prevalence was based on the median of the posterior estimated distribution in dementia prevalence (%) (a) and spatial autocorrelation analysis (b). Note: the provinces with only urban and rural residence are Beijing, Shanghai, Tianjin and Qinghai respectively.

concentrated in North China, while areas with low–low cluster were clustered in East and South China (Fig. 1b).

The risk factors with significant results of association included never drinking, drinking over once per week, former smoking, physical inactivity, social isolation, hearing impairment, diabetes, and hypertension. These effects showed a little variation across provinces, as illustrated in Fig. 2, with additional details provided in Appendix H. Compared to the status of "drinking less than or equal to once a week," never drinking exhibits a protective effect against dementia. The highest effect is observed in Fujian (0.61, 95% BCI: 0.41, 0.88), while the lowest is in Heilongjiang (0.79, 95% BCI: 0.55, 1.10). Provinces with a high-risk effect of "drinking more than once a week" are concentrated in the north. Moving on to smoking, in comparison to the status of "never smoking," the highest risk effect of the province for "former smoking" is in Fujian (2.02, 95% BCI: 1.32, 2.94), and the lowest is in Yunnan (1.64, 95% BCI: 1.15, 2.30). Transitioning to lifestyle factors, the risk effect of physical inactivity is significant in the south, ranging from 1.35 (95% BCI: 1.04, 1.80) in Anhui to 1.52 (95% BCI: 1.06, 2.12) in Guangdong. The highest risk effect of social isolation is in Jilin (2.07, 95% BCI: 1.44, 2.87), and the lowest is in Yunnan (1.71, 95% BCI: 1.26, 2.40). Examining health conditions, the range of the risk effect of diabetes is 1.43 (95% BCI: 0.89, 2.25) in Beijing to 1.89 (95% BCI: 1.17, 2.93) in Xinjiang. Hearing impairment and hypertension display greater risk effects in the north than in the south, with ranges of 1.23 (95% BCI: 0.90, 1.75) in Jiangxi to 2.64 (95% BCI: 1.49, 5.00) in Tianjin and 1.40 (95% BCI: 1.01, 2.04) in Guangdong to 2.26 (95% BCI: 1.49, 3.69) in Heilongjiang, respectively. Regarding to the unmeasured factors' effects, the distribution of random effects is shown in Fig. 3. The largest effects are observed in Henan and Shaanxi, while the smallest effects are in Zhejiang, Shandong, and Liaoning. Moreover, the box diagram of covariates' effects (Appendix H) serves as a visual representation summarizing the provincial effects. Detailed results of the sensitivity analysis in Appendix I show no statistical difference in model parameters before multiple imputation and main analysis, indicating the robustness of our findings. And the MSE of imputation was 0.13, which indicated the imputation method was effective and accurate.

Discussion

We employing a Bayesian regression spatial model to estimate provincial dementia prevalence in China, using a consistent dementia definition algorithm, and considering key established risk factors. Our findings reveal that dementia prevalence varied across China's provinces in 2018. High prevalence was concentrated in the provinces of North China, particularly marked by a high–high cluster, while low prevalence was concentrated in East and South China, characterized by a low–low cluster. Additionally, the association between different risk factors and dementia prevalence existed provincial variation.

Our study extends the current knowledge on provincial prevalence of dementia in China by a consistent dementia definition, providing uncertainty estimation and exploring the spatial clustering between providences. In line with previous systematic reviews,[4–7] we verified the regional pattern of higher dementia prevalence in the north and west, and lower prevalence in the south and east. Furthermore, we overcame the definition and methodological discrepancies in prior studies

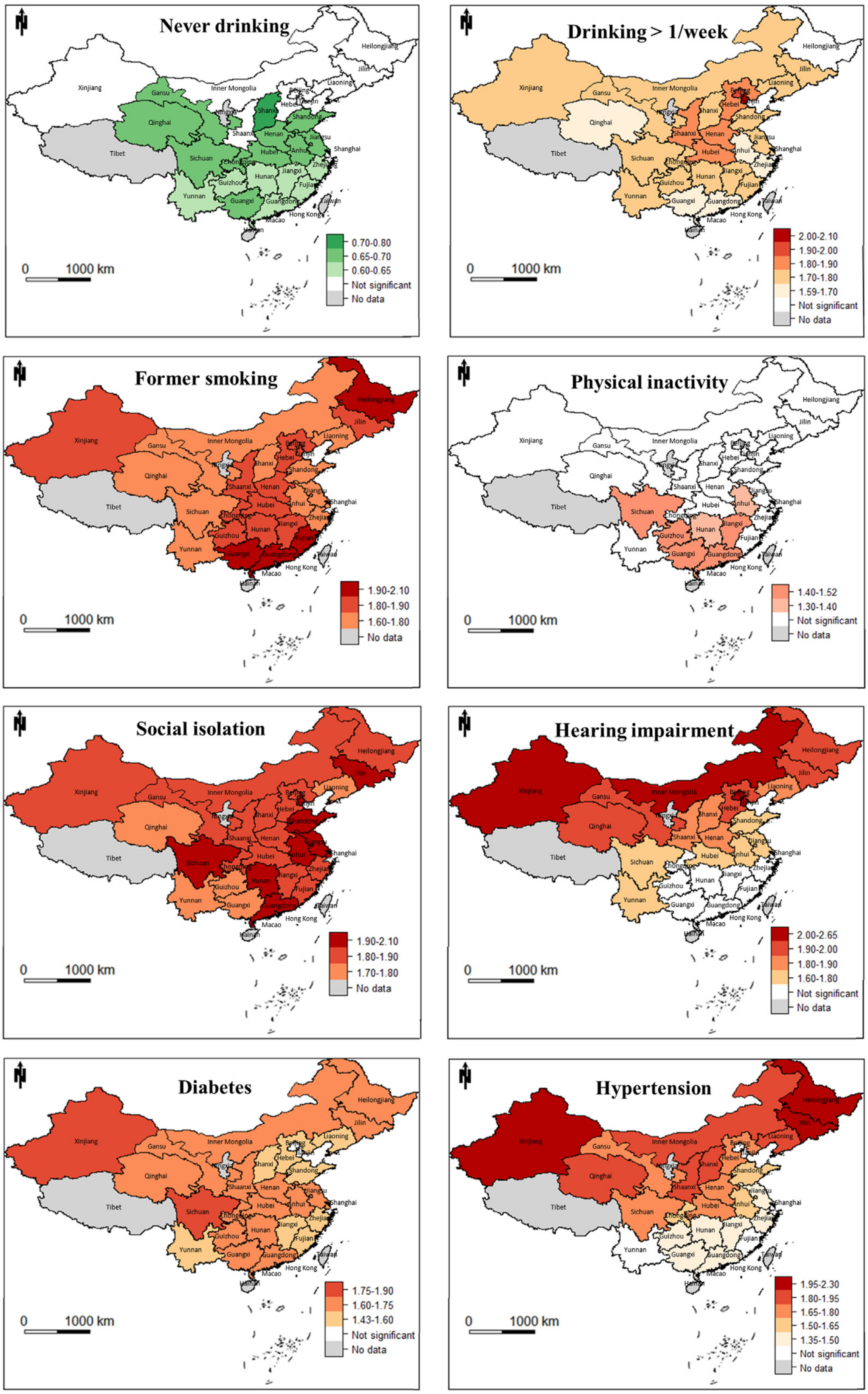

Fig. 2: Province-specific regression coefficients of dementia risk factors.

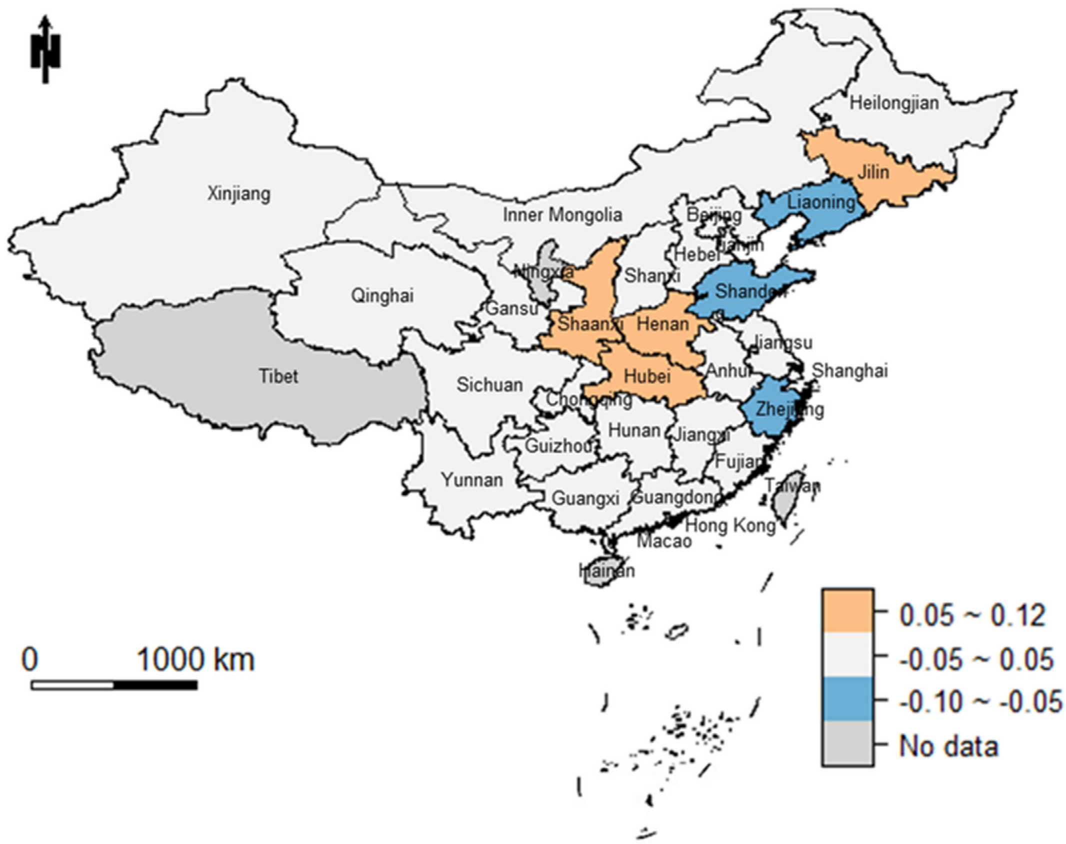

Fig. 3: Spatial distribution of random effects. The estimation was based on the median of the posterior distribution of risk in each province.

by employing a consistent and validated[21] definition of dementia to generated provincial-level rather than reginal-level prevalence estimates. Advancing the evidence from the China Alzheimer report (2022),[42] which only provided point estimation of provincial dementia prevalence, our estimation further provided Bayesian credible interval to quantify the uncertainty of prevalence and considered the spatial correlation between provinces. In addition, we examined the spatial clustering of provincial dementia prevalence, revealing a high–high clustering of dementia prevalence in three northern provinces (Inner Mongolia, Tianjin, and Gansu), while a low–low clustering in six southeast provinces (Zhejiang, Fujian, Jiangxi, Guangdong, Guangxi, and Hunan). This evidence has not been reported previously, and suggests that these three northern provinces with high–high clustering demand prioritized control interventions for dementia.

We delved into the provincial variations in dementia risk factors to estimate prevalence. The advantages of the model included considering spatially varying coefficients of different risk factors across provinces and highlighting the spatial patterns of these effects.[43] By considering the diverse effects of risk factors in different provinces, more accurate estimations of provincial dementia prevalence is possible. However, caution is warranted in interpreting the provincial effects of modifiable risk factors since the values showed a little difference across provinces but exhibited a certain spatial trend. The geographical variation in dementia prevalence might be partially explained by the relatively higher risk impacts of hypertension and hearing impairment in north China compared to the south as revealed by our analysis. We found that former smoking was a significant risk factor, while current smoking was not, which might be explained by health selection bias, as former smokers may have quit smoking due to health problems.[44] And protective effect of never drinking exhibits north-south differences. Due to the cold weather, alcohol drinking was common in the provinces of northern China.[16] Therefore, the varying impact of risk factors across provinces might be confounded by the other spatial-level factors. Besides, other unmeasured spatial-level risk factors of dementia as represented spatial random effects, also turned to show higher impact across central and northern regions. For instance, geographical differences in environmental factors(e.g. air populltion,[45] exposure to sunshine and Vitamin D intake[46]), regional lifestyle (such as dietary pattern,[47] behavioural habits[48]), climate (e.g. cold weather and low humidity in north etc.) and accessibility to health services[49] may all contribute to disparity in the identified dementia prevalence across China. Moreover, internal migration after retirement across provinces may also contribute to geospatial disparity in dementia prevalence. Previous studies suggested that the predominant internal migration patterns stem from the Central or Western provinces towards the Eastern China,[50] and as migrants tended to have better health status compared to non-migrants,[51] this may potentially contribute to decreased dementia prevalence of inflow provinces and increased prevalence of outflow provinces. We didn't identify any inter-provincial migrants among our study sample, while excluding these rural-to-urban migrants within province (13%) did not change the geospatial pattern of dementia identified in our main analysis. Future study could use data with rich information on inter-provincial migration to further explore this issue. Provinces with high risk should be paid more attention through the strategies including prevention and treatment of dementia. Targeted dementia prevention strategies could include controlling hypertension by reducing salt intake,[52] treating hearing impairment by using hearing aids,[53] and controlling smoking and drinking by reducing tobacco use[54] and alcohol consumption.[55] There is a need to prioritize the development of related health facilities in northern and western China for the prevention and treatment of dementia.

The results of our study must be interpreted in the context of several limitations. First, the prevalence interval estimates for Beijing, Tianjin, Shanghai, and Xinjiang exhibit wide ranges owing to small sample sizes, and participants from these provinces are exclusively from rural or urban areas, necessitating cautious interpretation. Nonetheless, we constructed provincial survey weights to enhance provincial result representation, potentially providing additional insights into provincial variations in dementia prevalence. Our study was confined to the provinces in the sample, and estimating dementia prevalence in other unsampled provinces requires further exploration. Second, majority of the risk factors examined in our study were self-reported, which may incur reporting bias particularly among participants with cognitive impairment.[56] Nevertheless, our findings on the association between risk factors and dementia are consistent with many other population-based dementia studies,[9–11] which also relied on self-reported data for risk factor assessment. Future research could incorporate objective measures for dementia risk factors, such as accelerator to quantify physical activity level, to enhance assessment reliability. Finally, the current study mainly examined the cross-sectional spatial distribution of provincial dementia prevalence in China, while its spatio-temporal variation is yet explored. Subsequent research can utilize data from multiple waves of CHARLS to investigate changes in geographical distribution of dementia prevalence over time, alongside time-changing impacts of dementia risk factors.

In conclusion, our study reveals a geospatial disparity in the prevalence of dementia and effects of risk factors among China's provinces. The prevalence of dementia in some provinces of northern and southern China showed high–high and low–low cluster, respectively. This study contributes evidence towards a comprehension of provincial dementia prevalence and risk factors in China. These findings provide new reference basis dementia prevention and treatment strategies tailored to specific provinces within China.

Contributors

YSL and JL contributed to the conception of the study and the study design, and obtained the research funding to support this study and provided continual supervision. YXL and XYG contributed to the data curation and the statistical analyses. YJZ, WHH and MRZ contributed to the code check. YXL, YYL and YJW contributed to the interpretation of the data. YXL, YSL and JL accessed and verified the data, wrote the initial version of the manuscript. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Data sharing statement

Citations for the data used in the study can be accessed in these official websites: https://charls.pku.edu.cn/and https://www.stats.gov.cn/. The raw data and code supporting the conclusions of this article will be made available by the authors, without undue reservation and these data will be available with publication. The detailed result of this study will be provided on a public website for downloading.

Editor note

The Lancet Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

Declaration of interests

The authors have no conflicts of interest to disclose. All authors confirm that they had full access to all the data in the study and accept responsibility to submit for publication.

Acknowledgements

This study was supported by the National Science Foundation of China/the Economic and Social Research Council, UK Research and Innovation joint call: Understanding and Addressing Health and Social Challenges for Ageing in the UK and China. UK-China Health and Social Challenges Ageing Project (UKCHASCAP): present and future burden of dementia, and policy responses. (Grant number 72061137003, ES/T014377/1). We would like to acknowledge the China Health and Retirement Longitudinal Study (CHARLS) team for providing high-quality, nationally representative data.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lanwpc.2024.101117.

References

1 Jia J, Wei C, Chen S, et al. The cost of Alzheimer's disease in China and re-estimation of costs worldwide. *Alzheimers Dement*. 2018;14(4):483–491.
2 Li X, Feng X, Sun X, Hou N, Han F, Liu Y. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2019. *Front Aging Neurosci*. 2022;14:937486.
3 Huang Y, Li X, Liu Z, et al. Projections of the economic burden of care for individuals with dementia in mainland China from 2010 to 2050. *PLoS One*. 2022;17(2):e0263077.
4 Wu YT, Ali GC, Guerchet M, et al. Prevalence of dementia in mainland China, Hong Kong and Taiwan: an updated systematic review and meta-analysis. *Int J Epidemiol*. 2018;47(3):709–719.

5 Li K, Wei S, Liu Z, et al. The prevalence of alzheimer's disease in China: a systematic review and meta-analysis. *Iran J Public Health*. 2018;47(11):1615–1626.
6 Zhang Y, Li Y, Ma L. Recent advances in research on Alzheimer's disease in China. *J Clin Neurosci*. 2020;81:43–46.
7 Zhu Y, Liu H, Lu XL, et al. Prevalence of dementia in the People's Republic of China from 1985 to 2015: a systematic review and meta-regression analysis. *BMC Publ Health*. 2019;19(1):578.
8 Jia L, Quan M, Fu Y, et al. Dementia in China: epidemiology, clinical management, and research advances. *Lancet Neurol*. 2020;19(1):81–92.
9 Jia L, Du Y, Chu L, et al. Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: a cross-sectional study. *Lancet Public Health*. 2020;5(12):e661–e671.
10 Huang Y, Wang Y, Wang H, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study. *Lancet Psychiatr*. 2019;6(3):211–224.
11 Qi S, Yin P, Zhang H, et al. Prevalence of dementia in China in 2015: a nationwide community-based study. *Front Public Health*. 2021;9:733314.
12 Chen H, Huang Y, Lv X, et al. Prevalence of dementia and the attributable contributions of modifiable risk factors in China. *Gen Psychiatry*. 2023;36(3):e101044.
13 Zhou M, Astell-Burt T, Bi Y, et al. Geographical variation in diabetes prevalence and detection in China: multilevel spatial analysis of 98,058 adults. *Diabetes Care*. 2015;38(1):72–81.
14 Li Y, Wang L, Feng X, et al. Geographical variations in hypertension prevalence, awareness, treatment and control in China: findings from a nationwide and provincially representative survey. *J Hypertens*. 2018;36(1):178–187.
15 Zhang X, Zhang M, Zhao Z, et al. Geographic variation in prevalence of adult obesity in China: results from the 2013-2014 national chronic disease and risk factor surveillance. *Ann Intern Med*. 2020;172(4):291–293.
16 Li R, Wang D, Chen J, Chai J, Tang M. Regional differences in smoking, drinking, and physical activities of Chinese residents. *Asia Pac J Public Health*. 2015;27(2):NP230–NP239.
17 Luo Y, Guo S, Wang Y, Pang L, Guo C, Zheng X. Socioeconomic disparities and dementia in China. *Psychiatry Res*. 2022;313:114611.
18 Kontari P, Fife-Schaw C, Smith K. Clustering of cardiometabolic risk factors and dementia incidence in older adults: a cross-country comparison in England, the United States, and China. *J Gerontol A Biol Sci Med Sci*. 2023;78(6):1035–1044.
19 Sara A, Maria G, Piotr B, et al. Temporal trend in dementia incidence since 2002 and projections for prevalence in England and Wales to 2040: modelling study. *BMJ*. 2017;358:j2856.
20 Chen Y, Araghi M, Bandosz P, et al. Impact of hypertension prevalence trend on mortality and burdens of de mentia and disability in England and Wales to 2060: a simulation model ling study. *Lancet Healthy Longev*. 2023;4(9):e470–e477.
21 Liu Y, Wu Y, Cai J, et al. Is there a common latent cognitive construct for dementia estimation across two Chinese cohorts? *Alzheimers Dement (Amst)*. 2022;14(1):e12356.
22 Yi Z. Introduction to the Chinese longitudinal Healthy longevity survey (CLHLS). In: Yi Z, Poston DL, Vlosky DA, Gu D, eds. *Healthy longevity in China: demographic, socioeconomic, and psychological dimensions*. Dordrecht: Springer Netherlands; 2008:23–38.
23 Chertkow H, Nasreddine Z, Joanette Y, et al. Mild cognitive impairment and cognitive impairment, no dementia: Part A, concept and diagnosis. *Alzheimers Dement*. 2007;3(4):266–282.
24 Katz S, Downs TD, Cash HR, Grotz RC. Progress in development of the index of ADL. *Gerontologist*. 1970;10(1):20–30.
25 Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet commission. *Lancet*. 2020;396(10248):413–446.
26 Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). *Am J Prev Med*. 1994;10(2):77–84.
27 Lu Y. Assessing air pollution exposure misclassification using high-resolution PM2.5 concentration model and human mobility data. *Air Qual Atmos Health*. 2023;16(11):2225–2238.
28 Julian B. Spatial interaction and the statistical analysis of lattice systems. *J Roy Stat Soc B*. 1974;36(2):192–225.
29 Gelfand A, Kim H, Sirmans C, Banerjee S. Spatial modeling with spatially varying coefficient processes. *J Am Stat Assoc*. 2003;98(462):387–396.
30 Julian B, Jeremy Y, Annie M. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Stat Math*. 1991;43(1):1–20.
31 Getis A. Spatial weights matrices. *Geogr Anal*. 2009;41(4):404–410.
32 Goodchild MF. Spatial data analysis. In: Liu L, Özsu MT, eds. *Encyclopedia of Database systems*. New York, NY: Springer New York; 2016:1–6.
33 George G. *Spatial autocorrelation. Spatial analysis methods and practice: describe – explore – explain through GIS*. Cambridge: Cambridge University Press; 2020:207–274.
34 Luc A. Local indicators of spatial association—lisa. *Geogr Anal*. 1995;27(2):93–115.
35 Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J Roy Stat Soc B*. 2009;71(2):319–392.
36 Shah AD, Bartlett JW, Carpenter J, Nicholas O, Hemingway H. Comparison of random forest and parametric imputation models for imputing missing data using mice: a caliber study. *Am J Epidemiol*. 2014;179(6):764–774.
37 Van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. *J Stat Software*. 2011;45(3):1–67.
38 Little RJ, Rubin DB. *Statistical analysis with missing data*. John Wiley & Sons; 2019.
39 Zhao Y, Strauss J, Yang G, et al. *China health and retirement longitudinal study 2011-2012 national baseline users' guide*; 2013. https://charls.charlsdata.com/Public/ashelf/public/uploads/document/2011-charls-wave1/application/CHARLS_nationalbaseline_users_guide.pdf.
40 Kulas J, Robinson D, Smith J, Kellar D. Post-stratification weighting in organizational surveys: a cross-disciplinary tutorial. *Hum Resour Manag*. 2018;57(2):419–436.
41 Marta B, Michela C, Gianluca B, Håvard R. Spatial and spatio-temporal models with R-INLA. *Spat Spatiotemporal Epidemiol*. 2013;7:39–55.
42 Ren R, Qi J, Lin S, et al. The China alzheimer report 2022. *Gen Psychiatr*. 2022;35(1):e100751.
43 Wan Q, Tang Z, Pan J, et al. Spatiotemporal heterogeneity in associations of national population ageing with socioeconomic and environmental factors at the global scale. *J Clean Prod*. 2022;373.
44 Inoue-Choi M, Ramirez Y, Fukunaga A, Matthews CE, Freedman ND. Association of adherence to Healthy lifestyle recommendations with all-cause and cause-specific mortality among former smokers. *JAMA Netw Open*. 2022;5(9):e2232778.
45 An Z, Huang RJ, Zhang R, et al. Severe haze in northern China: a synergy of anthropogenic emissions and atmospheric processes. *Proc Natl Acad Sci USA*. 2019;116(18):8657–8666.
46 Kang JH, Vyas CM, Okereke OI, et al. Effect of vitamin D on cognitive decline: results from two ancillary studies of the VITAL randomized trial. *Sci Rep*. 2021;11(1):23253.
47 Su X, Zhang J, Wang W, et al. Dietary patterns and risk of mild cognitive impairment among Chinese elderly: a cross-sectional study. *PLoS One*. 2020;15(7):e0235974.
48 Li X, Wu C, Lu J, et al. Cardiovascular risk factors in China: a nationwide population-based cohort study. *Lancet Public Health*. 2020;5(12):e672–e681.
49 Yip W, Fu H, Chen AT, et al. 10 years of health-care reform in China: progress and gaps in Universal Health Coverage. *Lancet*. 2019;394(10204):1192–1204.
50 Qi W, Abel GJ, Muttarak R, Liu S. Circular visualization of China's internal migration flows 2010–2015. *Environ Plann: Econ Space*. 2017;49(11):2432–2436.
51 Hou B, Nazroo J, Banks J, Marshall A. Impacts of migration on health and well-being in later life in China: evidence from the China health and retirement longitudinal study (CHARLS). *Health Place*. 2019;58:102073.
52 Liu Y, Chu H, Peng K, et al. Factors associated with the use of a salt substitute in rural China. *JAMA Netw Open*. 2021;4(12):e2137745.
53 Lin FR. Age-related hearing loss. *N Engl J Med*. 2024;390(16):1505–1512.
54 Otuyama LJ, Oliveira D, Locatelli D, et al. Tobacco smoking and risk for dementia: evidence from the 10/66 population-based longitudinal study. *Aging Ment Health*. 2020;24(11):1796–1806.
55 Jiang H, Room R, Hao W. Alcohol and related health issues in China: action needed. *Lancet Global Health*. 2015;3(4):e190–e191.
56 Newell SA, Girgis A, Sanson-Fisher RW, Savolainen NJ. The accuracy of self-reported health behaviors and risk factors relating to cancer and cardiovascular disease in the general population: a critical review. *Am J Prev Med*. 1999;17(3):211–229.