2024-01-01

Modelling the spatial risk pattern of dementia in Denmark using residential location data: A registry-based national cohort

Prince M. Amegbor *New York University*

Clive E. Sabel *School of Geography, Earth and Environmental Sciences*

Laust H. Mortensen *University of Copenhagen*

Amar J. Mehta *University of Copenhagen*

Spatial and Spatio-temporal Epidemiology

Modelling the spatial risk pattern of dementia in Denmark using residential location data: A registry-based national cohort

--Manuscript Draft--

Manuscript Number:	SSTE-D-23-00018R4
Article Type:	Research Paper
Keywords:	Dementia; socioeconomic factors; contextual factors; Bayesian spatial modelling; Stochastic partial differential equation (SPDE)
Corresponding Author:	Prince Michael Amegbor, PhD New York University New York, NY UNITED STATES
First Author:	Prince Michael Amegbor, PhD
Order of Authors:	Prince Michael Amegbor, PhD
	Clive E Sabel, PhD
	Amar J Mehta, PhD
	Laust H Mortensen, PhD
Abstract:	Dementia is a major global public health concern that is increasingly leading to morbidity and mortality among older adults. While studies have focused on the risk factors and care provision, there is currently limited knowledge about the spatial risk pattern of the disease. In this study, we employ Bayesian spatial modelling with a stochastic partial differential equation (SPDE) approach to model the spatial risk using complete residential history data from the Danish population and health registers. The study cohort consisted of 1.6 million people aged 65 years and above from 2005 to 2018. The results of the spatial risk map indicate high-risk areas in Copenhagen, southern Jutland and Funen. Individual socioeconomic factors and population density reduce the intensity of high-risk patterns across Denmark. The findings of this study call for the critical examination of the contribution of place of residence in the susceptibility of the global ageing population to dementia.

Modelling the spatial risk pattern of dementia in Denmark using residential location data: A registry-based national cohort

Abstract

Dementia is a major global public health concern that is increasingly leading to morbidity and mortality among older adults. While studies have focused on the risk factors and care provision, there is currently limited knowledge about the spatial risk pattern of the disease. In this study, we employ Bayesian spatial modelling with a stochastic partial differential equation (SPDE) approach to model the spatial risk using complete residential history data from the Danish population and health registers. The study cohort consisted of 1.6 million people aged 65 years and above from 2005 to 2018. The results of the spatial risk map indicate high-risk areas in Copenhagen, southern Jutland and Funen. Individual socioeconomic factors and population density reduce the intensity of high-risk patterns across Denmark. The findings of this study call for the critical examination of the contribution of place of residence in the susceptibility of the global ageing population to dementia.

Keywords: dementia, socioeconomic factors, contextual factors, Bayesian spatial modelling, Stochastic partial differential equation (SPDE).

1. Introduction

Dementia is an umbrella term for a range of neurological disorders usually characterised by the progressive loss of memory and cognitive abilities (Winblad et al., 2016; World Health Organization, 2017). The most common form of dementia is Alzheimer's disease (AD) which accounts for 60-70% of all dementia cases (World Health Organization, 2017). Other forms of dementia include vascular dementia, caused by impaired blood flow to the brain; Lewy body dementia, associated with an abnormal build-up of masses of proteins; and frontotemporal dementia (World Health Organization, 2017). While these forms of dementia together with AD are medically distinct, they often coexist and the boundaries between them are usually indistinct (World Health Organization, 2017).

Dementia has been recognised as a public health priority in both developed and developing countries. Dementia is currently the seventh leading cause of death among all diseases globally (GBD 2019 Collaborators, 2021) and in developed countries was ranked second in Australia in 2020 (Australian Bureau of Statistics, 2020), and sixth in the USA (Kochanek et al., 2020). Globally, an estimated 1.62 million people died from dementia in 2019 (GBD 2019 Collaborators, 2021). In Denmark, dementia is believed to be the fourth most common cause of death, after heart diseases, cancer and respiratory diseases (Taudorf et al., 2021b;

Vestergaard et al., 2020). In 2015, 47.47 million people were living with dementia worldwide (World Health Organization, 2015). Currently, over 50 million people are living with dementia and the figure is expected to triple to 152 million in the next three decades (Alzheimer's Disease International, 2019). Dementia cases in European Union member states account for nearly 18.2% (or 9.1 million cases) of global cases (OECD/EU, 2018). An estimate from a recent study shows 126,734 people were living with dementia in Denmark in 2013 representing 4.5% of individuals aged 50 years and above (Fann et al., 2018). The risk factors for dementia are multifaceted including socioeconomic status, environmental risk exposures, biological and lifestyle and behavioural factors (Chen et al., 2017; GBD 2016 Dementia Collaborators, 2019; Mukadam, Sommerlad, Huntley, & Livingston, 2019; Murayama et al., 2019; Wimo et al., 2017).

Current dementia research has paid little attention to how individual and contextual socioeconomic factors and demographic factors inform the spatial risk distribution patterns of dementia. Understanding factors that contribute to an elevated risk of dementia based on where one lives is crucial for addressing spatial inequalities in health outcomes as well as designing intervention measures for tackling dementia. The longitudinal Danish registries provide a rich resource, enabled via individual data linkage to link health records and individual-level data from the population registers across the whole life course (Pedersen, 2011; Pedersen et al., 2006). The insights from using geospatial analytic techniques, such as Bayesian spatial disease mapping, are vital for enhancing our understanding of geographical differences in the risk of dementia.

In this study, we explore the spatial risk patterns of dementia across Denmark using individual-level data and geocoded residential addresses. Specifically, the study sought to examine the spatial risk pattern of dementia in Denmark, accounting for individual and neighbourhood-level socioeconomic factors.

2. Data & Methods

2.1 Cohort

This register-based study consisted of all individuals aged 65 years and above who lived in Denmark for at least one year from 2005 through 2018 before the study entry – that is, before age 65 years (N = 1,757,168). This allowed us to capture a more comprehensive picture of individuals' residential history, including potential changes in address within a given year as well as exclude individuals who may have had dementia before coming or moving to Denmark, as their medical records may not be captured in the health registers. This helps to minimize potential biases in our analysis by focusing on individuals with a sufficient period of observation in the Danish healthcare system. Thus, the follow-up for this study starts at age 65 years. Individuals with dementia at the time of cohort entry (n=36,543) were excluded from the cohort. All individuals were followed from cohort entry until the incident date (lateonset dementia), date of death, date of emigration, loss-to-follow-up, or end of follow-up (31/12/2018), whichever came first. In Denmark, all residents are assigned a unique personal identification number and are registered in the Danish Civil Registration System (CRS) (Mainz et al., 2019; Pedersen, 2011; Schmidt et al., 2014). Information in the CRS includes historical complete residential addresses as well as the date of address change for all registered people in the country (Schmidt et al., 2014). This personal identification number enables linkage with the national population and health registers *at the individual level* and allows for follow-up of the population over time. The data used in this study included information on the place of residence at the time of entry into the cohort. Researchers are generally not given access to the residential address or information of the Danish population due to privacy concerns. However, this study was conducted under the auspices of Statistics Denmark under an initiative to see how researchers can use individual data, including geographic location data, in research without breaching Statistics Denmark's data protection, privacy and ethics rules. All analyses were done within Statistic Denmark.

In this study, we focused on non-institutionalised people living with dementia. Given the high prevalence of dementia among institutionalised older adults as well as their relatively frail health conditions to compared non-institutionalised individuals, including the former would invariably imply mapping the spatial distribution of long-term care homes; thus, skewing the true spatial effect.

2.2 Measures

2.2.1 Outcome

Consistent with a previous Danish study on the incidence of late-onset dementia (Taudorf et al., 2019), we defined diagnosis of dementia as a registered diagnosis of dementia (ICD-10 codes: F00.0-00.9, G30.0-30.9, F01.0-01.9, F02.0, F03.9, G31.9, G31.8) in the National Patient Register or the Psychiatric Central Research Register or as having filled at least one prescription for an anti-dementia drug (ATC codes: N06DX01, N06DA02, N06DA03, N06DA04) in the National Prescription Registry. All individuals who were diagnosed with dementia *before* their 65th birthday or cohort entry were excluded from the study.

2.2.2 Individual-level factors

Individual-level predictors used in the study were: age, sex, country of origin, the highest level of education, employment status and household wealth at the age 65 of years. Age as used in the study had two conception: the first is age defined as age at cohort entry (used in the descriptive analysis) and the second with age defined as a time varying covariate. The highest level of education obtained by people in the cohort was coded in the Danish education system with the following categories: basic or primary education, vocational educational training/qualifying educational program/upper secondary education (hereafter called Vocational training or Upper secondary education), Vocational bachelor's/short-cycle higher education, and bachelor's degree or higher. Given the compulsory basic education system in Denmark, individuals in the cohort who had missing information on the highest educational attainment were categorised as having basic or primary education (Hegelund et al., 2021). There were 147,440 older adults with missing education information. Missing education information was more common among elderly immigrants. Although we assumed all individuals with missing education information had basic education, there is the possibility their education level is higher than basic education. We conducted a crosstabulation analysis with chi-square and Cramer's V tests to compare the two groups (primary education and missing or unknown). The result of the tests shows that even though there were significant associations between SES and the groups, the associations were considerably weak with Cramer's V values of less than 0.2. A descriptive summary of the employment and sex distribution also shows the two categories are more similar compared to other educational groups.

Employment status had three categories: employed, pension and other. Household wealth in this data refers to the net difference of total assets minus total liabilities; obtained from Statistics Denmark’s family income register.

2.2.3 Neighbourhood-level contextual factors

The neighbourhood-level contextual variables used in this study were constructed by aggregating the individual and family-level socioeconomic characteristics into three kilometres square grids. The first contextual factor was population density which is a spatial covariate variable that depicts the density of people aged 65 years and above per grid cell (3km by 3km). The next contextual variables were the median age of the population per grid cell and the median household wealth per grid cell. We also computed the proportions of males, people born abroad, people with a secondary level of education and above and employed older adults per three-kilometre square grid as contextual variables.

2.3 Analysis

We utilised spatial survival analysis to explore the spatial inequalities in the risk of dementia among older adults in Denmark. In contrast to the standard survival analysis model, spatial survival analysis is a frailty model that recognises the effect of contextual factors and mechanisms on individual survival rate or hazard rate (Balan and Putter, 2020; Banerjee, 2016; Taylor, 2015). Spatial survival models are similar to standard survival models, however, they incorporate random effects (or frailties) whose joint distribution can be interpreted in relation to the spatial context (Balan and Putter, 2020; Taylor, 2015). Thus, by combining survival analysis and geostatistical methods, we can address key questions, such as where in the study area is the rate of dementia usually low or high. The spatial survival models were implemented using the integrated nested Laplace approximations (INLA) framework with the Stochastic Partial Differential Equations (SPDE) approach. The INLA framework is a computationally less-intensive (compared to Markov chain Monte Carlo (MCMC) simulative methods) deterministic algorithm for Bayesian inference based on the latent Gaussian model (van Niekerk et al., 2022, 2021).

2.3.1 SPDE modelling

The SPDE approach is a geostatistical method based on the assumption that the data can be modelled as a continuous spatial process using a discretely indexed Gaussian field. Thus, it uses an approximation solution via the finite element method (FEM) where the spatial area is divided into non-intersecting triangles resulting in a triangulated mesh or Delaunay triangulations. Details of the SPDE model can be found in (Krainski et al., 2019; Lindgren et al., 2021, 2011). We employed a Bayesian Weibull survival model with the stochastic partial differential equation (SPDE) approach. In this study, we adopted a Weibull survival model due to its flexibility via accommodating a wide range of survival patterns including both increasing and decreasing hazard rates over time. The Weibull distribution allows for different shapes of the hazard function, including exponential (constant hazard), increasing hazard (early failures), and decreasing hazard (late failures) (Carroll, 2003; Lamon, 2016;

Plana et al., 2022). This makes it suitable for capturing various survival patterns observed in real-world data. In addition, its parametric assumption makes it easier to estimate the model parameters and make predictions. It is also a versatile model compared to other survival models as it is known to provide results similar to Cox analysis even when the data does not follow an exact Weibull distribution (Carroll, 2003). It is also unique in that it simultaneously assumes both proportional hazards and accelerated time-failure properties. Thus, it can estimate both the relative rate of events and the relative extension in the survival time (Carroll, 2003). In addition, its parametric assumption makes it easier to estimate the model parameters and make predictions.

By incorporating the SPDE approach in our survival model, we account for the spatial effects of the risk of having dementia. The spatial effect here refers to the location-specific independent noise that cannot be accounted for by the study covariates (Lezama-Ochoa, Pennino, Hall, Lopez, & Murua, 2020; Muñoz, Pennino, Conesa, López-Quílez, & Bellido, 2013). That is, it corresponds to the spatial location of each measurement or observation with the assumption that the locations were sampled independently of the survival process. Residential address at cohort entry was the location data used in modelling the spatial risk distribution of dementia. That is, the hazard of a ith individual in location j having dementia, adjusting for individual socioeconomic and contextual factors, can be expressed as:

$$h(t_{ij}) = h_0(t_{ij})\exp(\mathbf{x}_{ij}\beta + \xi\_j) \qquad (1)$$

where t_{ij} represents the duration or elapsed time from entry into the cohort until the occurrence of dementia or end of follow-up for the ith individual in location j. h_0 is the baseline hazard function. $\mathbf{x}_{ij}$ is a matrix of the covariates for the fixed effect with regression coefficient β and ξ_j is the spatial latent Gaussian random field (GF) or Matérn GF. In INLA this can be written as:

ADRD = survival (time,censoring) ~ 1 + covariates + f(S, model=spde)

Where ADRD is the hazard of having dementia. 'S' is a spatial index for the Matérn GF (number of vertices in the SPDE model). This represents the spatial random effect that captures the spatial variation or clustering in the hazard of dementia that is not explained by the covariates. The 'model=spde' specifies that the spatial random effect is modelled using the SPDE approach with a Matérn covariance function. The number of vertices (from the Delaunay triangulation mesh) in the SPDE model influences the spatial resolution and complexity of the spatial random effect. We used a Weibull likelihood for the survival time with marginal variance and range for the spatial component and the shape parameter in the Weibull distribution as hyperparameters (Lindgren et al., 2011).

The Delaunay triangulation mesh used in defining the SPDE Matern had the geopolitical boundary of mainland Denmark and Bornholm as the domain boundary. We define mesh with a **cutoff** (minimum values for distance allowed between points) value of *2* and **max.edge** (the maximum allowed length of the triangle edges) value of *4*. The SPDE Matérn was defined with the mesh with a parameter alpha (the smoothness parameter of the process) of 2 – the default value in the inla.spde2.pcmatern() function. A penalised complexity (PC) prior was used with the spatial range parameter as 0.05 and 0.01; while the prior distribution for the spatial standard deviation (sigma) was specified as 1 and 0.01. This can be implemented in INLA with the code below:

```
inla.spde2.pcmatern(mesh = mesh, alpha = 2,  prior.range = c(0.05, 0.01), prior.sigma = c(1, 0.01))
```

Currently, there is no universally accepted guideline for the selection of prior parameters. For this study, we explored five different sets of prior parameters and based our final choice on the model diagnostic information, specifically opting for the one with the lowest WAIC value. The prior parameters used in this study were also influenced by similar parameters used in a SPDE survival analysis (Krainski et al., 2019). Nonetheless, this does not imply the parameters used are the definitive or optimal priors among a myriad of potential parameter values.

First, we modelled the baseline spatial distribution of the risk (log hazard rate ratio) of dementia in Denmark – Model 1; that is, a model without individual socioeconomic and neighbourhood-level contextual covariates. We stratified the baseline spatial risk by sex considering the hazards are different between males and females. Next, we modelled the spatial risk distribution of dementia using an incremental modelling approach. The first model (Model 2) adjusted for individual-level factors and population density. The second model (Model 3) adjusted for individual-level factors, population density and the mean age per grid cell, while the final model (Model 4) - a fully adjusted model –adjusts for all individual- and contextual-level factors. In the modelling, individual age, household wealth and neighbourhood-level contextual factors were included as continuous variables. The continuous variables were standardized to avoid computational challenges resulting from covariates of different units and with extreme values, such as wealth. We assumed linearity in the association between these factors and the risk of dementia. The assumption of a linear association with dementia was verified by fitting SPDE models with linear assumptions for the continuous variables and models with non-linear assumptions. The best fitting model – that is linear versus non-linear – was determined using the Watanabe–Akaike Information Criterion (WAIC) values. WAIC as model assessment criteria accounts for both the goodness-of-fit and the complexity of the model through the estimated effective number of parameters (Spiegelhalter et al., 2002). The WAIC values showed the linear model offers significant improvement in the fit than the non-linear model. The model diagnostics information can be found in Appendix E of the supplementary material. Age in Models 2 to 4 was an age as a time-varying covariate.

We also conducted simulation analyses to assess the performance of the SPDE model – that is, to determine if the SPDE models can recover the disease trend irrespective of population distribution patterns. A detailed description of the simulation study, analysis set-up, and results can be found in the supplementary material – Appendix B. We also conducted a sensitivity analysis to explore potential regional differences in diagnostic or coding of the outcome – supplementary material Appendix D. In the sensitivity analysis, we include a random term for region with an "independent and identically distributed" or “iid” assumption.

That is, in the sensitivity model, we assume the “region" effect to be independent of each other and follow an identical distribution. The regions are the five administrative regions of Denmark, namely: Hovedstaden (capital region), Midtjylland (Central Denmark region), Nordjylland (Northern Denmark region), Sjælland (Zealand region), and Syddanmark (Southern Denmark region). Each region is responsible for regional governance functions, including healthcare administration. The model was implemented in INLA as:

ADRD = survival (time,censoring) ~ 1 + covariates + f(region, model="iid") + f(S, model=spde)

Modelling was done using the R software (R Core Team 2020) and the R-INLA packages (Bakka et al. 2018; Lindgren and Rue 2011, 2015; Martins et al. 2013; Riebler et al. 2017; Rue et al. 2009). The posterior mean for the spatial effect was visualised at a 5km by 5km resolution using the basic plot function in R, after the results were transformed into raster data. The posterior mean for the spatial effect can be interpreted as the log hazard rate ratio which represents the baseline hazard rate for each spatial location compared to the average hazard rate. Positive values indicate locations with a higher baseline hazard rate compared to the average, while negative values indicate locations with a lower baseline hazard rate compared to the average. For the fixed effect parameters in the INLA-SPDE model, the results were considered statistically significant if the credible intervals indicated a positive or negative association with the outcome variable. In other words, if the credible intervals show consistent results in terms of the direction of the association (e.g., both lower and upper limits of the intervals are entirely above zero or entirely below zero), it indicates a significant relationship between the covariate and the outcome. The results of the fixed effects are displayed in Table C1 in Appendix C.

3. Results

3.1 Descriptive and cross-tabulation

The final cohort consisted of 1,658,575 individuals who were aged 65 years and above at any point during the period from 2005 to 2018 with valid geocoded addresses and complete sociodemographic information. Figure A1 in the supplementary material provides an overview of the cohort selection process. In this study, 110,038 individuals were diagnosed with dementia during the study period (from 2005 to 2018). The incidence rate of dementia among the cohort from 2005 to 2018 was 8.5 (95% CI = 8.4 – 8.5) cases per 1,000 person years at risk. Table 1 below shows the distribution of baseline characteristics of the study cohort and contextual factors. The table also shows the distribution of cases by sociodemographic groups and the incidence rate.

Table 1. Baseline characteristics of 1,658,575 persons aged 65 years and above who were followed for development of dementia from 2005 to 2018 where 110,038 developed dementia during 12,961,214 person years of follow-up.

	Number of People	Number of Cases	Incidence Rate (per 1,000 person years at risk)	Person (Each
Individual-Level Factors				
Age at Study Entry*				
65-69 years	1,154,340	33,264	4.0	8
70-74 years	181,517	22,755	11.6	1,
75-79 years	143,683	24,602	18.0	1,
80-84 years	102,560	18,518	23.2	7
85 years and above	76,475	10,899	26.8	4
Sex				
Female	885,703	66.121	9.3	5,
Male	772,872	43.917	7.5	7,
Place of Birth				
Denmark	1,578,533	105.767	8.5	12
Western country	46,112	2.885	8.4	3
Non-western country	33,930	1.386	6.2	2
Highest Level of Education				
Primary	744,368	65.766	10.9	6,
Vocational training/Upper Secondary Education	594,087	30,452	6.7	4,
Vocational Bachelors/Short-cycle Higher Education	239,810	10,288	5.8	1,
Bachelors Degree/Higher	80,310	3,532	6.1	5
Employment Status				
Employed	331,485	5.987	2.8	2,
Other	74,087	1,720	3.4	5
Pension	1,253,003	102.331	9.9	10
Wealth - Quintiles [a]				
Lowest - (-21,769)	331,715	20,876	9.2	2,
Lower (187,756)	331,715	26,549	10.6	2,
Medium (779,606)	331,715	23,040	8.6	2,
Richer (1,598,718)	331,715	20,979	7.6	2,
Richest (3,476,865)	331,715	18,594	6.7	2,
Contextual Variables (by 3x3km grid)				
Population Density - Quintiles [b]				
Lowest (0.1 - 37.6)	331,715	16,848	6.5	2,
Lower (37.6-132.0)	331,715	19,815	7.6	2,
Medium (132.0 - 323.0)	331,715	21,618	8.3	2,
Higher (323.0 - 649.0)	331,715	23,825	9.1	2,
Highest (649.0 - 2,852.0)	331,715	27,932	11.0	2,
Mean Age - Quintiles [b]				
Lowest (65 - 74)	331,715	16,141	6.4	2,

Lower (74 - 75)	331,715	20,409	7.9	2,
Medium (75 - 76)	331,715	22,793	8.7	2
Higher (76 - 77)	331,715	24,581	9.4	2
Highest (77 - 96)	331,715	26,114	9.9	2
Proportion of Male - Quintiles [b]				
Lowest (0.0 - 43.0)	331,715	27,372	10.8	2
Lower (43.0 - 44.8)	331,715	25,277	9.7	2
Medium (44.8 - 46.9)	331,715	22,006	8.4	2
Higher (46.9 - 49.6)	331,715	19,358	7.4	2
Highest (49.6 - 100.0)	331,715	16,025	6.2	2
Proportion Non-Danish Born - Quintiles [b]				
Lowest (0.0 - 2.1)	331,715	19,019	7.3	2
Lower (2.1 - 3.2)	331,715	20,010	7.7	2
Medium (3.2 - 4.6)	331,715	21,302	8.2	2
Higher (4.6 - 7.1)	331,715	23,624	9.0	2
Highest (7.1 - 100.0)	331,715	26,083	10.3	2
Proportion - Post-secondary Education - Quintiles [b]				
Lowest (0.0 -12.5)	331,715	21,391	8.3	2
Lower (12.5 - 15.6)	331,715	22,210	8.6	2
Medium (15.6 - 19.3)	331,715	21,981	8.5	2
Higher (19.3 - 25.4)	331,715	22,263	8.6	2
Highest (25.4 - 100.0)	331,715	22,193	8.5	2
Proportion Employed - Quintiles [b]				
Lowest (0.0 - 14.5)	331,715	23,440	9.0	2
Lower (14.5 - 17.3)	331,715	24,030	9.3	2
Medium (17.3 - 20.7)	331,715	22,011	8.5	2
Higher (20.7 - 24.4)	331,715	22,233	8.7	2
Highest (24.4 - 100.0)	331,715	18,324	7.0	2
Mean Wealth - Quintiles [a]				
Lowest (254,151)	331,715	23,054	10.7	2
Lower (553,532)	331,715	26,961	8.9	2
Medium (736,203)	331,715	21,757	8.4	2
Richer (974,479)	331,715	19,321	7.4	2
Richest (1,533,625)	331,715	18,945	7.1	2

*** = Age is defined here as age at cohort entry; a = median value for quintile; b = range for quintile**
NB: Age at end date (defined as the end of follow-up that is Age at event date, death, emigration or end of study which occurs first) was used in the Bayesian Spatial models

The results show that the incidence rate of dementia is higher among females, Danish-born, less-educated older adults and older adults receiving a pension. The incidence rate of dementia among female older adults was 9.27 per 1,000 person-years at risk, compared to 7.54 per 1,000 person-years at risk among males. Older adults with a primary level of education had an incidence rate of 10.9 per 1,000 person-years at risk, compared to 6.14 per 1,000 person-years at risk among those with a university degree or higher level of education.

The next sections present the results of the spatial risk pattern of dementia. Here, we present the results of the baseline model and the fully adjusted model. The results for the entire incremental model can be found in Appendix C of the supplementary material. First, we present the result of the spatial risk pattern for the baseline model – the unadjusted model; then compare this with the spatial risk patterns from the incremental models – Model 2, Model 3, and Model 4 – to determine if the risk patterns change after accounting for different individual and contextual factors.

3.2 Spatial risk pattern of dementia – location-specific log-hazard

The result of the simulation analyses shows that the SPDE model can recover the true spatial parameters irrespective of the situation – where we manipulate the spatial risk and the association with population density (Supplementary Material – Appendix B). Figure 1 below shows the spatial pattern of the location-specific effect on the risk (log-hazard rate ratio) of dementia in Denmark for the baseline model (Model 1) – unadjusted. The spatial effect observed in this model can be interpreted as the location-specific effect on the risk (loghazard rate ratio) of dementia in Denmark that is not explained by sex. The spatial risk pattern shows clusters of high-risk areas in the capital region (Copenhagen), Funen and southern Jutland. Please refer to Figure A2 in the supplementary material which displays the provinces or regions in Denmark.

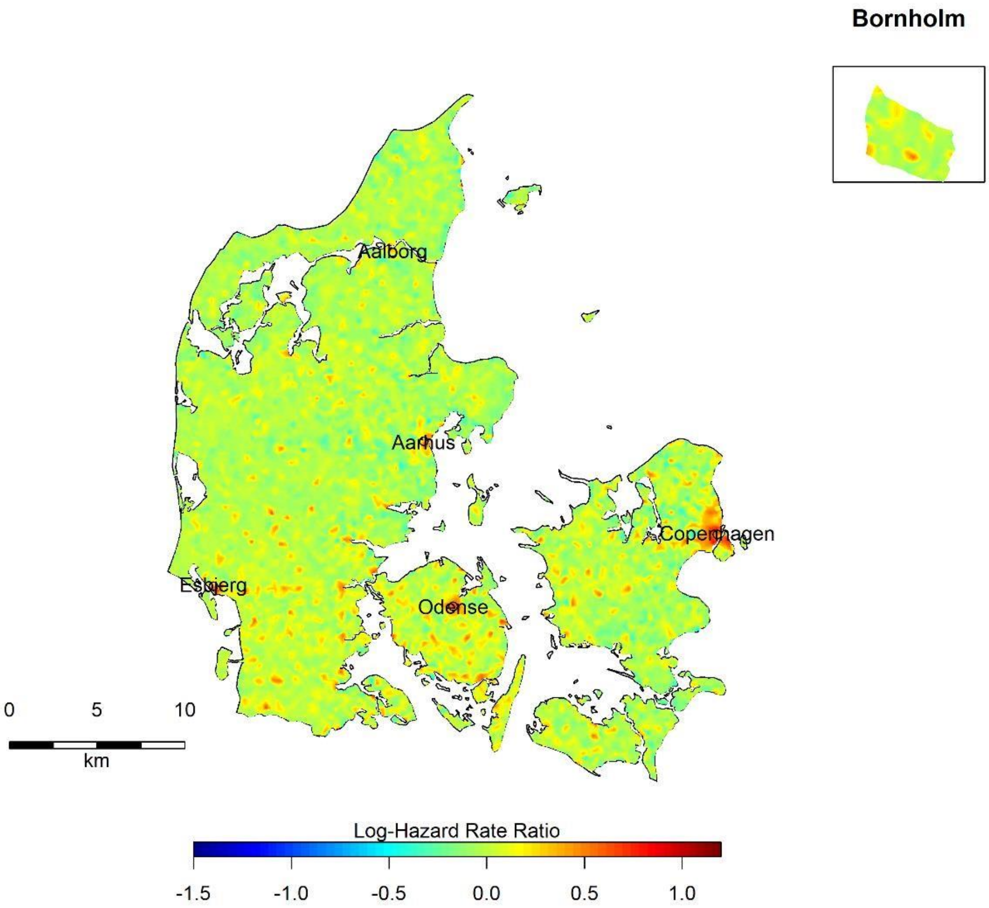

Figure 1. Map of the spatial effect (log-hazard rate ratio) for the baseline (Model 1) Bayesian Weibull survival model.

In the fully adjusted model (Figure 2 – below), the local intensity of the risk of dementia based on place of residential at age 65 years reduced across the country. For instance, the results show variations in location-specific effects on the risk of dementia in the Copenhagen region, Funen and southern Jutland compared to Figure 1. Some areas that were previously high-risk locations became lower-risk locations when we account for individual- and neighbourhood-level factors in the modelling process. However, some clusters of high-risk locations remain in the Copenhagen region, Odense and areas in southern Jutland.

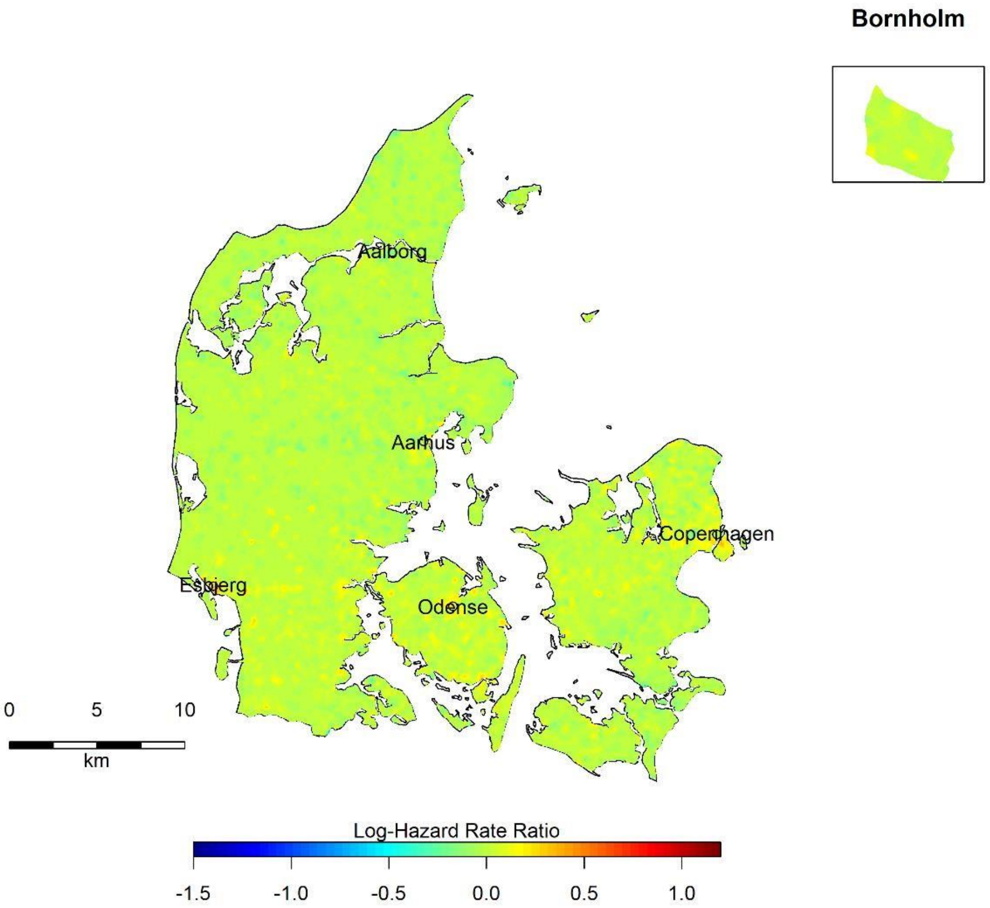

Figure 2. Map of the spatial effect (log-hazard rate ratio) for Model 4 Bayesian Weibull survival model – fully adjusted model.

Adjusting for various individual and neighbourhood-level factors generally reduced the spatial risk of dementia across Denmark (supplementary Figures C1_A to C4_A). The posterior standard deviation for the spatial effect is also presented in the supplementary material (Figures C1_B to C4_B). The pattern of the location-specific effect on the risk of dementia also changes depending on the factors adjusted for in the models. Appendix C of the supplementary material shows the results of the effect of individual and neighbourhoodlevel factors on dementia. The results of the sensitivity analysis for potential regional effect due to differences in diagnosis or coding of dementia are presented in Figure D1. The results for the fully adjusted model (Figure D1 - Panel B) show reductions in the high-risk intensity in the Copenhagen region, Odense and areas in southern Jutland – compared to the result in Figure 2.

4. Discussion

Dementia is a major public health issue and a leading cause of frailty among older adults (Prince et al., 2012; Wimo et al., 2017; World Health Organization, 2017). Current studies have enhanced our understanding of the risk factors and plausible pathways for its occurrence; albeit, there is limited knowledge on the risk distribution of the disease. While studies point to geographical differences in the prevalence and incidence of dementia (Russ et al., 2015, 2012), the scale of analysis is often too large for understanding the potential spatial variation of dementia. This is the first nationwide registry-based study to link individual-level geocoded residential location and demographic and health register information and to adopt Bayesian spatial survival analyses to explore the risk distribution of dementia in Denmark. The incidence rate (Table 1) and the results of Bayesian spatial survival analyses (Table C1 – supplementary material) show that people with lower socioeconomic status have a higher incidence rate and a greater risk of dementia.

4.1 Variations in Risk Patterns Across Denmark

The spatial models also reveal heterogeneity in the risk of dementia across Denmark with high-risk areas mainly in the capital region around Copenhagen, southern Jutland and the island of Funen. The spatial pattern in the unadjusted model suggests that the risk of dementia is higher in urban areas with notable clusters of high risk in the Copenhagen region, Odense and Aarhus, as well as other urban areas in southern Jutland, Funen and Zealand. Some of the high-risk areas identified in this study are also noted as high-risk locations for some cardiovascular disease outcomes, such as acute myocardial infarction (Kjærulff et al., 2016). Kjærulff et al. (2016) observed high-risk clusters of acute myocardial infarction in the Copenhagen region, Aarhus municipality and some regions in Funen and southern Denmark. The finding also corroborates the finding of an early study in Denmark that shows a higher incidence of psychiatric disorders in urban areas. However, the finding contradicts the results of studies from other geographical contexts that suggest a protective effect of urban living on the risk of dementia and other psychiatric disorders (Russ et al., 2015; Xiang et al., 2018). We postulate the differences in the findings on the effect of rural-urban residence may be explained by the geographical, demographic and socioeconomic composition of the populations, and in the differing geographic scale of analyses in the different study contexts. Risky behaviour, such as smoking and alcohol abuse, and poor socioeconomic status have been cited as potential contributors to the high incidence and risk of dementia among the rural population (Fors et al., 2009; Hegelund et al., 2021; Nguyen et al., 2008; O'Donovan et al., 2020; Tola-Arribas et al., 2013). The socioeconomic and environmental contexts of rural Denmark are different from the context of the other studies. The welfare system of Denmark coupled with other contextual factors means people in rural Denmark may not have the same socioeconomic vulnerabilities as observed in other countries. Thus rural Denmark may have a protective effect due to reduced exposure to these adverse lifestyles, socioeconomic vulnerabilities and environmental risks (Vassos et al., 2016).

We observed changes in the spatial risk patterns after adjusting for different individual- and contextual-level factors in the risk modelling. The intensity of high-risk areas reduces after accounting for individual- and contextual-level factors. For instance, in the Copenhagen region (Figure C2 and Figure C3) after adjusting for individual- and contextual socioeconomic factors, as well as, population density we observe a shrink in the high-risk area coverage. That is, places formerly identified as high-risk in the unadjusted model became low-risk in the adjusted models. This is not surprising considering the spatial effect (risk) is not independent (or orthogonal) to individual-level factors. This is because individual socioeconomic status and the contextual characteristics of their neighbourhoods may protect or increase their risk exposures. Thus, these micro and meso (contextual) level factors interact and may even modify or mediate how broader geographic factors contribute to the risk of dementia among older adults. Insight from previous studies suggests that there is an association between population density and the risk of psychiatric or mental health disorders (Colodro-Conde et al., 2018; Sariaslan et al., 2015; Xiang et al., 2018). They argue that stressors associated with densely populated areas may contribute to the increased risk among populations in these areas (Colodro-Conde et al., 2018; Sariaslan et al., 2015; Vassos et al., 2016). The findings of our study reveal that individual socioeconomic and contextual factors modify the risk distribution pattern of dementia across Denmark based on residential location at the age of 65 years. The results of the sensitivity analysis show that these observed spatial risk effects are an artefactual but true reflection of risk patterns based on the study cohort and defined measures used in the analysis.

4.2 Strengths and limitations

Our study has its strengths and limitations. A major strength of this research is the use of a registry-based design with large and comprehensive data on all older adults living in Denmark in the defined study period – 2005 to 2018. This research design approach eliminates potential selection bias associated with traditional forms of data collection, such as a survey. It also strengthens the generalisability of the findings of the study. High spatial resolution regarding the exact residential location of the individuals in the study cohort is also a strength of this present study. Last but not least, the stochastic partial differential equation (SPDE) approach used in modelling the spatial risk pattern proved to be robust as demonstrated in the sensitivity analyses. The sensitivity analyses show that the model was able to recover the actual disease risk pattern irrespective of the population distribution pattern. Our study has its limitations too. First, our methodology does not consider the potential effect of genetic predisposition on the risk distribution pattern of dementia in Denmark. Another limitation of this study is the spatial risk patterns of dementia are captured based on residential location at the time of entry into the cohort. A potential future study may want to explore prior residential locations in early life to compare the risk pattern to that of this study. Further, the modelling approach used in this study is based on a global model which is based on the assumption that the parameter effect or the association between socioeconomic factors, population density and dementia are constant across Denmark. Future studies could employ local statistical models, such as Bayesian spatial varying coefficient and geographically weighted regression models, to explore potential spatial variations in the association between socioeconomic factors, population density and dementia, as well as, environmental factors and dementia.

5. Conclusion

This is the first study to explore spatial variation in place-specific risk of late-onset dementia among older adults using nationwide administrative data from the Danish population and health registers. The study illustrates variation in the risk pattern of dementia across Denmark with high-risk areas mostly clustered in the Copenhagen region, southern Jutland and Funen. The risk pattern shows that the effect of place of residence on the risk of dementia generally tends to be low in the peripheral or suburban areas of major cities in Denmark, including

Copenhagen, and Aarhus. Another insight from this study was that individual and contextual socioeconomic factors modify the risk intensity pattern. The intensity of high-risk patterns reduces, after adjusting for individual and contextual factors, including population density. The findings of this study call for the critical examination of the contribution of place of residence in the susceptibility of the global ageing population to dementia. It also identifies hotspot locations in Denmark, thus paving the way for future research investigating the environmental, social and economic mechanisms in neighbourhoods that contribute to the high-risk burden of dementia.

Acknowledgements

This study was supported by BERTHA, the Danish Big Data Centre for Environment and Health funded by the Novo Nordisk Foundation Challenge Programme (grant number NNF17OC0027864).

References

Australian Bureau of Statistics. Causes of Death. Canberra: 2020.

Balan TA, Putter H. A tutorial on frailty models. Stat Methods Med Res 2020;29:3424–54. https://doi.org/10.1177/0962280220921889.

Banerjee S. Spatial Data Analysis. Annu Rev Public Health 2016;37:47–60. https://doi.org/10.1146/annurev-publhealth-032315-021711.

Carroll KJ. On the use and utility of the Weibull model in the analysis of survival data. Control Clin Trials 2003;24:682–701. https://doi.org/10.1016/S0197-2456(03)00072-2.

Colodro-Conde L, Couvy-Duchesne B, Whitfield JB, Streit F, Gordon S, Kemper KE, et al. Association between population density and genetic risk for schizophrenia. JAMA Psychiatry 2018;75:901–10. https://doi.org/10.1001/jamapsychiatry.2018.1581.

Fann JR, Ribe AR, Pedersen HS, Fenger-Grøn M, Christensen J, Benros ME, et al. Longterm risk of dementia among people with traumatic brain injury in Denmark: a populationbased observational cohort study. The Lancet Psychiatry 2018;5:424–31. https://doi.org/10.1016/S2215-0366(18)30065-8.

Fors S, Lennartsson C, Lundberg O. Childhood living conditions, socioeconomic position in adulthood, and cognition in later life: Exploring the associations. Journals Gerontol - Ser B Psychol Sci Soc Sci 2009;64:750–7. https://doi.org/10.1093/geronb/gbp029.

GBD 2019 Collaborators. Global mortality from dementia: Application of a newmethod and results from the global burden of disease study 2019. Alzheimer's Dement Transl Res Clin Interv 2021;7:1–28. https://doi.org/10.1002/trc2.12200.

Hegelund ER, Mehta AJ, Mortensen LH, J.Westendorp RG. The plasticity of late-onset dementia: A nationwide cohort study in Denmark. Alzheimer's Dement 2021. https://doi.org/10.1002/alz.12469.

Kjærulff TM, Ersbøll AK, Gislason G, Schipperijn J. Geographical clustering of incident acute myocardial infarction in Denmark: A spatial analysis approach. Spat Spatiotemporal Epidemiol 2016;19:46–59. https://doi.org/10.1016/j.sste.2016.05.001.

Kochanek KD, Xu J, Arias E. NCHS Data Brief No. 395, December 2020: Mortality in the United States, 2019. Hyattsville, MD: 2020.

Krainski ET, Gómez-Rubio V, Bakka H, Lenzi A, Castro-Camilo D, Simpson D, et al. Advanced Spatial Differential Equations Stochastic Partial Modeling with Using R and INLA. New York: Chapman and Hall/CRC; 2019. https://doi.org/https://doiorg.ez.statsbiblioteket.dk:12048/10.1201/9780429031892.

Lamon J. Statistical-Probabilistic Approaches to Brittle Fracture: The Weibull Model. In: Lamon J, editor. Brittle Fract. Damage Brittle Mater. Compos., Elsevier; 2016, p. 35–49. https://doi.org/10.1016/b978-1-78548-121-5.50002-8.

Lindgren F, Bolin D, Rue H. The SPDE approach for Gaussian and non-Gaussian fields: 10 years and still running 2021.

Lindgren F, Rue H, Lindström J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J R Stat Soc Ser B Stat Methodol 2011;73:423–98.

Mainz J, Hess MH, Johnsen SP. The Danish unique personal identifier and the Danish Civil Registration System as a tool for research and quality improvement. Int J Qual Heal Care 2019;31:717–20. https://doi.org/10.1093/intqhc/mzz008.

Nguyen CT, Couture MC, Alvarado BE, Zunzunegui, M. V. Life Course Socioeconomic Disadvantage and Cognitive Function Among the Elderly Population of Seven Capitals in Latin America and the Caribbean. J Aging Health 2008;20:347–62. https://doi.org/10.1177/0898264308315430.

van Niekerk J, Bakka H, Rue H, Schenk O. New Frontiers in Bayesian Modeling Using the INLA Package in R . J Stat Softw 2021;100. https://doi.org/10.18637/jss.v100.i02.

van Niekerk J, Krainski E, Rustand D, Rue H. A new avenue for Bayesian inference with INLA. J LATEX Templates 2022. https://doi.org/10.48550/arxiv.2204.04678.

O'Donovan G, Hamer M, Sarmiento OL, Hessel P. Education in early life markedly reduces the probability of cognitive impairment in later life in Colombia. Sci Rep 2020;10:1–8. https://doi.org/10.1038/s41598-020-74822-2.

Pedersen CB. The Danish civil registration system. Scand J Public Health 2011;39:22–5. https://doi.org/10.1177/1403494810387965.

Pedersen CB, Gøtzsche H, Møller JØ, Mortensen PB. The Danish Civil Registration System. A cohort of eight million persons. Dan Med Bull 2006;53:441–9.

Plana D, Fell G, Alexander BM, Palmer AC, Sorger PK. Cancer patient survival can be parametrized to improve trial precision and reveal time-dependent therapeutic effects. Nat Commun 2022;13:1–13. https://doi.org/10.1038/s41467-022-28410-9.

Prince M, Acosta D, Ferri CP, Guerra M, Huang Y, Rodriguez JJL, et al. Dementia incidence and mortality in middle-income countries, and associations with indicators of cognitive reserve: A 10/66 Dementia Research Group population-based cohort study. Lancet 2012;380:50–8. https://doi.org/10.1016/S0140-6736(12)60399-7.

Russ TC, Batty GD, Hearnshaw GF, Fenton C, Starr JM. Geographical variation in dementia: Systematic review with meta-analysis. Int J Epidemiol 2012;41:1012–32. https://doi.org/10.1093/ije/dys103.

Russ TC, Gatz M, Pedersen NL, Hannah J, Wyper G, Batty GD, et al. Geographical variation in dementia: Examining the role of environmental factors in Sweden and Scotland. Epidemiology 2015;26:263–70. https://doi.org/10.1097/EDE.0000000000000230.

Sariaslan A, Larsson H, D'Onofrio B, Långström N, Fazel S, Lichtenstein P. Does Population Density and Neighborhood Deprivation Predict Schizophrenia? A Nationwide Swedish Family-Based Study of 2.4 Million Individuals. Schizophr Bull 2015;41:494–502. https://doi.org/10.1093/schbul/sbu105.

Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. Eur J Epidemiol 2014;29:541–9. https://doi.org/10.1007/s10654-014-9930-3.

Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. J R Stat Soc Ser B Stat Methodol 2002;64:583–616. https://doi.org/10.1111/1467-9868.00353.

Stevenson M, Sergeant E, Nunes T, Heuer C, Marshall J, Sanchez J, et al. epiR: Tools for the Analysis of Epidemiological Data 2023.

Taudorf L, Nørgaard A, Brodaty H, Laursen TM, Waldemar G. Dementia increases mortality beyond effects of comorbid conditions: A national registry-based cohort study. Eur J Neurol 2021a;28:2174–84. https://doi.org/10.1111/ene.14875.

Taudorf L, Nørgaard A, Islamoska S, Jørgensen K, Laursen TM, Waldemar G. Declining incidence of dementia: A national registry-based study over 20 years. Alzheimer's Dement 2019;15:1383–91. https://doi.org/10.1016/j.jalz.2019.07.006.

Taudorf L, Nørgaard A, Waldemar G, Laursen TM, Wimo A. Mortality in Dementia from 1996 to 2015: A National Registry-Based Cohort Study. J Alzheimer's Dis 2021b;79:289–300. https://doi.org/10.3233/JAD-200823.

Taylor BM. Spatial Survival Analysis. In: Shekhar S, Xiong H, Zhou X, editors. Encycl. GIS, Cham, Switzerland: Springer Nature; 2015. https://doi.org/10.1007/978-3-319-23519-6.

Tola-Arribas MA, Yugueros MI, Garea MJ, Ortega-Valín F, Cerón-Fernández A, FernándezMalvido B, et al. Prevalence of Dementia and Subtypes in Valladolid, Northwestern Spain:
The DEMINVALL Study. PLoS One 2013;8:1–9. https://doi.org/10.1371/journal.pone.0077688.

Vassos E, Agerbo E, Mors O, Bøcker Pedersen C. Urban-rural differences in incidence rates of psychiatric disorders in Denmark. Br J Psychiatry 2016;208:435–40. https://doi.org/10.1192/bjp.bp.114.161091.

Vestergaard SV, Rasmussen TB, Stallknecht S, Olsen J, Skipper N, Sørensen HT, et al. Occurrence, mortality and cost of brain disorders in Denmark: A population-based cohort study. BMJ Open 2020;10:1–9. https://doi.org/10.1136/bmjopen-2020-037564.

Wimo A, Guerchet M, Ali GC, Wu YT, Prina AM, Winblad B, et al. The worldwide costs of dementia 2015 and comparisons with 2010. Alzheimer's Dement 2017;13:1–7. https://doi.org/10.1016/j.jalz.2016.07.150.

World Health Organization. Global action plan on the public health response to dementia 2017 - 2025. Geneva: 2017.

Xiang Y, Zare H, Guan C, Gaskin D. The impact of rural-urban community settings on cognitive decline: Results from a nationally-representative sample of seniors in China. BMC Geriatr 2018;18:1–9. https://doi.org/10.1186/s12877-018-1003-0.

Figure Captions

Figure 1. Map of the spatial effect (log-hazard rate ratio) for the baseline (Model 1) Bayesian Weibull survival model.

Figure 2. Map of the spatial effect (log-hazard rate ratio) for Model 4 Bayesian Weibull survival model – fully adjusted model.