Mapping Social Health and Dementia Risk: A Register-Based Study of Older Adults in Finland

Elisa Cisotto[1], Margherita Moretti[2 3], Margherita Silan[4], Joan Damiens[3], Pietro Belloni[4], Kaarina Korhonen[3], Pekka Martikainen[3]

[1] Department of Statistical Sciences, Università Cattolica del Sacro Cuore, Largo A. Gemelli 1, 20123 Milan, Italy
[2] Dondena Centre for Research on Social Dynamics, Bocconi University, Via Guglielmo Röntgen 1, 20136 Milan, Italy
[3] Helsinki Institute for Demography and Population Health, University of Helsinki & Max Planck – University of Helsinki Center for Social Inequalities in Population Health, Helsinki, Finland
[4] Department of Statistical Sciences, University of Padova, Via Cesare Battisti 241, 35121 Padova, Italy

Abstract

This study investigates the role of individual- and area-level social health factors in shaping geographic variation in dementia incidence among older adults in Finland. Using register data on all individuals born between 1935 and 1939 and residing in Finland in 2010 (N=161,909), we include several indicators of social health alongside key sociodemographic variables to estimate dementia incidence over a 10-year follow-up period. To address compositional confounding on the geographical comparison, we applied the Matching on Average Rank for Multiple Treatments (MARMoT), a matching-based method that equalizes the distribution of the key sociodemographic and social health characteristics across municipalities. We then used spatial scan statistics to detect clusters of excess dementia risk before and after this adjustment.

Results show that, prior to adjustment, several contiguous clusters of elevated dementia risk emerged, particularly in eastern and southern Finland. After balancing individual-level characteristics, the spatial clustering pattern changed substantially: some clusters disappeared, suggesting they were largely driven by population composition, while others persisted or newly emerged (central-western Finland), indicating the influence of contextual or unmeasured factors and revealing hidden geographic vulnerability.

Our findings highlight the relevance of social health for dementia prevention and demonstrate the importance of combining spatial and matching-based methods to disentangle compositional and

contextual drivers of health inequalities. This study also offers a novel application of register-based data for studying cognitive ageing, showing that spatially targeted policies may help reduce dementia risk in ageing societies.

Keywords: social health, dementia, Finland, MARMoT, spatial scan, older adults

Introduction

Maintaining cognitive functioning in later life is a fundamental condition for autonomy and quality of life, and dementia represents one of the most pressing public health challenges in ageing societies. Globally, over 57 million people live with dementia, with nearly 10 million new diagnoses each year, making it the seventh leading cause of death and a major contributor to disability among older adults (WHO, 2021).

While dementia is often perceived as an inevitable consequence of ageing, a growing body of research reveals substantial disparities in its incidence and prevalence across social groups and territories (Crimmins et al., 2018; Hudomiet et al., 2022). These disparities reflect both individual vulnerabilities and broader structural conditions. Socioeconomic status, education, family structure, and race/ethnicity are among the individual-level factors associated with dementia risk and patterns of diagnosis and classification (Brayne et al., 2006; Jones, 2017; Luth & Prigerson, 2022). Neighbourhood characteristics, service availability, and the quality of the social environment have been linked to contextual effects on cognitive decline (Clarke et al., 2015; Antonucci et al., 2024).

As a result, there is growing recognition that dementia should be examined not only through traditional biomedical and psychological models, but also within a broader framework that accounts for social conditions and environmental exposures (Winblad et al., 2016; Downs, 2000). Within this framework, the concept of *social health* has gained attention as a way to capture the interplay between individuals and their social environments. Originally introduced by the WHO in 1946 as one of the three dimensions of health, social health refers to the capacity to engage in meaningful relationships, participate in community life, and maintain autonomy (WHO, 1946; Vernooij-Dassen and Jeon, 2016; Vernooij-Dassen et al. 2018). In the context of dementia research, the concept has been operationalised to include both individual-level indicators (e.g. marital status, living arrangements, social participation) and contextual-level conditions (e.g. access to services, social infrastructure, neighbourhood safety).

The relationship between social resources and cognitive functioning is well-documented in epidemiological research (Bellou et al., 2017; Lenart-Bugla et al., 2022). Studies consistently shown that individuals with higher levels of social engagement, frequent interpersonal contact, and participation in cognitively stimulating leisure activities tend to experience more favourable cognitive outcomes and exhibit a lower risk of developing dementia (Duffner et al., 2022; Joyce et al., 2021; Kuiper et al., 2015; Maddock et al., 2023; Marseglia et al., 2023; Piolatto et al., 2022; Samtani et al., 2022; Sommerlad et al., 2019; Sutin et al., 2018; van der Velpen et al., 2022; Zhou et al., 2018). In contrast, social isolation have been linked to accelerated cognitive decline and increased dementia incidence, especially among older adults (Freak-Poli et al., 2022; Ma et al., 2018; Rafnsson et al., 2017; Sutin et al., 2018; Tsutsumimoto et al., 2017).

At the individual level, key determinants of social health, such as marital status, household composition, frequency of contact with family and friends, and active participation in social or leisure activities, have all been associated with the maintenance of cognitive functioning in later life. These associations may be mediated through mechanisms such as the preservation of cognitive reserve, which refers to the brain's resilience to neuropathological damage (Hahn et al., 2019; Livingston et al., 2020; Penninkilampi et al., 2018).

Beyond individual ties, contextual factors also play a critical role in shaping cognitive health. These include physical proximity to kin, availability of support networks, and broader neighbourhood characteristics such as access to social infrastructures and services (Clarke et al., 2012; Michael and Yen, 2014; Wu et al., 2015; Finlay et al., 2021). When operating jointly or intersecting, these conditions may generate cumulative exposures that shape individuals' cognitive resilience or vulnerability. Socially supportive environments can promote frequent social interactions and community engagement, thereby enhancing cognitive reserve by enabling individuals to access interpersonal and collective resources. In contrast, exclusionary conditions, such as limited access to support, social isolation, or environmental deprivation, may undermine cognitive resilience and contribute to increased vulnerability to cognitive decline.

A recent systematic review emphasizes that cumulative social disadvantages, including social isolation, economic hardship, and lack of meaningful roles in society, are strongly associated with higher dementia risk. Conversely, inclusionary conditions such as higher education, occupational complexity, and active social participation serve as protective buffers against cognitive decline (Vernooij-Dassen et al., 2021). These findings underscore the need for multidimensional frameworks that incorporate both individual and contextual dimensions of social health when examining the social determinants of dementia.

Growing attention has also been paid to the combinations of social health indicators. Studies show that the co-occurrence of multiple adverse social markers, such as living alone, limited social participation, and infrequent communication, is associated with increased dementia risk and broader declines in cognitive functioning (Ma et al., 2018; Tsutsumimoto et al., 2017; Yang et al., 2018). In contrast, combinations of protective factors, such as active engagement in social, cognitive, and physical leisure activities, are associated with a reduced likelihood of cognitive impairment (Opdebeeck et al., 2018; Wang et al., 2017; Zhu et al., 2017). For instance, Saito et al. (2018) found that among adults aged 65 and older, being married, having regular contact with friends and family, and being employed were jointly associated with a lower risk of developing dementia.

Despite these insights, a major analytical challenge remains: distinguishing between compositional effects (who lives in an area) and contextual effects (what the area is like). Many studies adjust for individual-level covariates through regression models but do not ensure comparability between geographic units in terms of population characteristics. As a result, elevated dementia incidence in a municipality may reflect both the concentration of high-risk individuals or the presence of place-specific environmental risk factors.

To address this, we propose a dual-level analytical framework that explicitly accounts for both individual and contextual contributions to dementia risk. This study focuses on Finland, a country with both a universalistic Nordic welfare regime and exceptionally high-quality administrative registers. These registers provide comprehensive data on individuals' social conditions (e.g. family status, proximity to kin), socio-economic background, and dementia diagnoses, enabling the analysis of population-level patterns of cognitive vulnerability. Moreover, Finland has socio-demographic characteristics similar to many other European or Nordic countries, and the results can thus be representative of many other contexts.

We restrict our sample to individuals born between 1935 and 1939, aged 76 to 80 at baseline in 2015, and dementia-free at the start of the observation period. We then observe new dementia diagnoses up to 2019. The outcome is measured using three register sources: health care records, medication purchases, and cause-of-death certificates. We distinguish between individual-level social health indicators, socio-demographic characteristics, and contextual characteristics at the municipal level

In line with recent conceptualizations of social health, we assess this construct using a set of indicators derived from population registers. At the individual level, the selected indicators include living with a partner, having children, having children residing in the same municipality, and having experienced widowhood. These variables reflect distinct dimensions of social health, including household composition, marital and parental status, and relational proximity, all of which represent potential sources of support or relational vulnerability in later life. Together, they capture complementary aspects of social connectedness and exclusion that are known to influence cognitive ageing.

Our selection was informed by a broader set of candidate variables derived from prior literature, including indicators of co-residence (e.g., living alone, living with children), partnership and parental history (e.g., widowhood, divorce, childlessness), adverse life events (e.g., recent child loss), and spatial proximity to kin. To ensure both conceptual clarity and statistical robustness, we conducted descriptive, correlational, and inferential analyses to examine collinearity and association with dementia risk. We retained those variables that contributed unique and interpretable information across multiple dimensions of social health, while minimizing redundancy and multicollinearity.

At the contextual level, we included four indicators capturing service provision and social infrastructure: the number of healthcare professionals, number of social workers, and the density of cultural venues and retail stores per 10,000 inhabitants. These environmental features reflect the broader capacity of local settings to support social inclusion, interaction, and access to care in later life.

To address confounding, we employ the Matching on Poset-based Average Rank for Multiple Treatments (MARMoT) approach (Silan et al., 2021), a non-parametric matching method designed to achieve covariate balancing across a large number of groups, in this case, 230 municipalities. This method enables the construction of a synthetic population in which the distribution of observed individual-level characteristics is similar across all areas, helping to more robustly compare areas, net of this composition

After balancing, we apply a spatial scan statistic to detect geographic clusters of elevated dementia incidence. Comparing the clustering patterns before and after adjustment allows us to identify areas where excess risk remains even if they had individual-level profiles similar to other areas. These clusters are interpreted as areas where unmeasured or contextual factors, beyond individual characteristics, may likely contribute to dementia risk, net of the composition of the population in terms of their social health.

Specifically, we ask to what extent geographic differences in dementia incidence across Finnish municipalities can be explained by differences in individual-level social health and contextual conditions. By comparing the spatial distribution of dementia risk before and after adjusting for these factors, we aim to estimate how much territorial inequalities could be reduced through interventions addressing social health and contextual disadvantages.

Our contribution is twofold. First, we provide an innovative application of the MARMoT approach in a highly granular spatial setting, exceeding the scale of most previous applications. Second, we offer new evidence on how the unequal distribution of dementia may be shaped by layered forms of social disadvantage, highlighting areas where either context-specific interventions or interventions aimed at improving social health conditions may be most needed to reduce dementia risks. The findings have relevance for public health policy in Finland and other ageing societies with spatially structured inequalities.

Comparing dementia clustering patterns before and after adjusting allows us to isolate residual risk and estimate the extent to which spatial inequalities might be reduced through targeted interventions on social health or on contextual factors.

Data

Study population

The study population comprises all individuals residing in Finland and born between 1935 and 1939, identified through the Finnish Population Register. These individuals were aged 76 to 80 in 2015 and were followed until 2019, reaching ages 79 to 84 during the observation period. We selected this specific cohort for two main reasons. First, dementia diagnoses prior to 2015 are considered less reliable due to changes in diagnostic practices and data coverage. Second, for individuals born before 1935, intergenerational links, such as those with adult children, are more difficult to trace in Finnish registers. This limitation has been acknowledged in previous research and was a criterion for the definition of the oldest cohorts in earlier register-based studies (Einiö et al., 2016).

The analytical sample was further restricted to individuals living outside institutional care at baseline and with no recorded diagnosis or treatment for dementia prior to 2015, which serves as the baseline year.

The study population identified through the Finnish Population Register was subsequently linked to national administrative datasets containing comprehensive longitudinal information on municipality of residence, health status, and key indicators of social health, both at the individual and contextual level.

Outcome

Dementia is the outcome of interest, and it was identified using three complementary sources:

1. The Care Register for Health Care (Finnish Institute for Health and Welfare), reporting dementia diagnosis based on ICD-10 codes (F00–F03, G30, F05.1) and ICD-9 codes (290, 291.2, 331.0, 333.1, 437.8A, 294.1A, 298.8C). ICD-9 codes were used for diagnoses recorded up to 1995 and served to identify individuals free of any dementia diagnosis between 1987 (the first year of data availability) and 1995. ICD-10 codes were used for diagnoses from 1996 onward.
2. The Drug Reimbursement Register (Social Insurance Institution of Finland), which tracks purchases of dementia-related medications based on Anatomical Therapeutic Chemical (ATC) code N06D.
3. The Cause of Death Register (Statistics Finland), which records dementia as either the underlying or a contributing cause of death, using the same ICD codes as in the Care Register. a primary or contributing cause of death, using the same ICD criteria described above[1].

Dementia status in 2015 was used to define the baseline sample, ensuring that all individuals were dementia-free at the start of the observation period. From 2016 to 2019, dementia was treated as the main outcome. Individuals were classified as having dementia from the date of their first recorded diagnosis or prescription of dementia-related medication. Once assigned, dementia status was considered irreversible and remained fixed for the remainder of the follow-up period.

[1] Except for ICD-10 code F02, which indicates dementia in other diseases classified elsewhere. This code is not assigned as an underlying cause of death in mortality records and was therefore excluded from the Cause of Death data.

For individuals whose death certificates indicated dementia but who had no prior recorded diagnosis or medication use, dementia status was retrospectively assigned up to two years prior to death. This approach is supported by evidence suggesting that clinical symptoms typically emerge several years before death (Joling et al., 2020).

To estimate the incidence of dementia, individuals were classified as either having or not having developed dementia by the end of the follow-up period. This classification includes both living individuals with a dementia diagnosis and deceased individuals for whom dementia was recorded as a primary or contributing cause of death. By design, all individuals included in the study were alive and dementia-free at baseline in 2015, aged 76 to 80.

Operationalizing Social Health

We conceptualize social health as a multidimensional construct encompassing both individual vulnerabilities and contextual resources that shape opportunities for social engagement, autonomy, and support in later life. Drawing on prior research (e.g., Vernooij-Dassen et al., 2021), we operationalize social health using administrative register indicators:

- Partnership status (currently living with a partner or not);
- Parental status (whether the individual has children);
- Geographic proximity to children (whether at least one child resides in the same municipality);
- History of widowhood (whether the individual has ever experienced widowhood during their lifetime);
- Disability status, capturing the most common diagnostic categories used in administrative records to approximate long-term functional, intellectual, developmental, or sensory impairments relevant to the concept of disability in later life (see Appendix A for ICD codes; Bister et al. 2025).

These variables reflect key dimensions of social support and relational vulnerability in later life. In line with our analytical strategy, they are treated as confounders, that is, individual-level characteristics that may influence both the exposure (contextual environment) and the outcome (dementia). Accordingly, they are included in the matching procedure to ensure covariate balance across municipalities. Moreover, we additionally adjust for two basic socio-demographic characteristics: gender and educational attainment.

At the contextual level, we include four area-level indicators reflecting local social infrastructure and service availability, measured per 10,000 residents in 2015:

- Number of healthcare professionals;
- Number of social workers;
- Number of cultural venues (e.g., theatres, libraries, museums);
- Number of retail stores, which serve as common sites for informal interaction.

Healthcare and social worker indicators are derived from employment records, assigning workers to municipalities based on their workplace location. Cultural and retail indicators are based on the number of establishments in each municipality, as these locations often facilitate informal social interaction and community engagement. Together, these indicators approximate the social infrastructure of local environments.

Contextual adjustment and aggregation

Population data from 2011 to 2018 served as the basis for computing per capita rates. These rates were calculated using denominators derived from the total municipal population, alongside sector-specific counts of employed individuals (in healthcare, social work, cultural, and retail sectors) and the number of establishments in the cultural and retail domains.

As of 2024, Finland comprises 308 municipalities with an average population size of 18,299 inhabitants. However, the distribution is highly skewed: the median municipal population is 5,977. The smallest municipality in mainland Finland is Lestijärvi (665 inhabitants), while the smallest in the Åland Islands is Sottunga (101 inhabitants). At the other end of the spectrum, Helsinki is by far the most populous municipality, with 684,018 residents in 2024.

To enhance the reliability of area-level comparisons, we aggregated municipalities with very small older populations. Municipalities with fewer than 130 residents aged 76–80 (below the first quartile) were merged based on spatial adjacency and demographic similarity. Adjacency was defined using a contiguity matrix (i.e., shared borders), and merging was guided by minimizing multidimensional Euclidean distances on sociodemographic characteristics relevant to aging and health: percentage living alone, childlessness, widowhood, proximity to children, child mortality, education, and presence of disability in the household. For geographically isolated areas (e.g., small islands, n = 8), aggregation was based on physical proximity. This procedure reduced the number of analytical units from 308 to 237, improving the stability and robustness of contextual analyses.

Method

Matching on Average Rank for Multiple Treatments (MARMoT)

To identify clusters of municipalities in Finland where older adults are exposed to a higher risk of dementia, we applied the Matching on Average Rank for Multiple Treatments (MARMoT) approach, a non-parametric matching method designed to balance covariates across more than two groups in observational settings (Silan et al., 2021). MARMoT builds on the theory of partially ordered sets (poset), particularly on methods developed for estimating and approximating average ranks in such structures (Brüggemann and Carlsen, 2011; De Loof et al., 2011; Caperna, 2019). These methods allow for the derivation of a scalar summary of each individual’s characteristics, the Average Rank (AR), which facilitates covariate balancing across multiple groups (i.e. Finnish municipalities), by generating a synthetic population in which the distribution of observed characteristics is similar across all geographic units.

By equalizing the composition of populations across municipalities, MARMoT allows for a credible interpretation of observed differences in dementia rates as reflecting contextual effects, rather than underlying compositional differences in individual-level risk factors.

The MARMoT procedure consists of the following steps:

1. Computation of the Average Rank (AR). Individual-level characteristics, both categorical and binary, are summarized into a single ordinal measure, the Average Rank (AR). This score reflects the degree to which each individual is exposed to characteristics potentially associated with dementia risk or selection into specific treatment groups.
2. Construction of a Frequency Table. A contingency matrix is generated, where rows represent distinct AR values and columns correspond to municipality. This table captures the distribution of individuals across municipalities by their AR profile.
3. Definition of a Reference Frequency. For each AR level, a reference frequency is defined to serve as a benchmark across all municipalities. This ensures that municipalities are equally represented for each AR profile, facilitating covariate balance.
4. Sampling to Create a Synthetic Balanced Population. Individuals are then randomly sampled within each municipality and AR level to match the reference frequency. This results in a synthetic population where the distribution of individual characteristics is balanced across municipalities.

The procedure is implemented using the R package `deloof`, which provides tools for computing average ranks and generating the synthetic balanced dataset.

The MARMoT approach offers several advantages for this study. First, it enables balancing across a large number of groups (i.e., hundreds of municipalities), without imposing parametric assumptions about treatment assignment. Second, it accommodates categorical and binary covariates in a principled and non-parametric manner. Third, it produces a synthetic population in which the distribution of observed characteristics is balanced across groups, thereby strengthening the credibility of inferences regarding contextual effects.

The goal of the method is to achieve comparability between groups with respect to observed confounders (i.e., demographic and social health characteristics), supporting the conditional independence (unconfoundedness) assumption required for causal inference. Unlike traditional resampling-based approaches, MARMoT applies a deterministic, rank-based matching strategy. It constructs a synthetic sample in which the distribution of observed covariates is balanced across multiple treatment groups. To assess the effectiveness of the matching procedure, we compute the Absolute Standardized Bias (ASB) for each covariate using the following formula:

$$ABS = \frac{\left|\underline{X}_t - \underline{X}\right|}{\sqrt{\frac{s_t^2 + s^2}{2}}} * 100$$

where $\underline{X}_t$ and $\underline{X}$ denote the mean of covariate in the treatment group and the overall sample, respectively, and s_t^2 and s^2 are the corresponding variances.

An ASB value below 10% is typically considered indicative of adequate balance. This diagnostic allows us to verify that the matching procedure has successfully reduced covariate imbalance prior to analysing contextual effects.

Detection of high-risk dementia clusters using spatial scan statistics

After balancing individual-level characteristics across municipalities using the MARMoT procedure, we proceeded to identify spatial clusters where the incidence of dementia remained significantly higher than expected. This step aimed to detect residual geographic variation that may reflect contextual risk factors, beyond population composition in terms of demographic and social health characteristics.

To this end, we applied spatial scan statistics (SSS), a class of methods designed to identify clusters of adjacent areas with significantly elevated or reduced risk of an outcome. SSS systematically scans the study area using moving windows, testing whether the incidence of dementia within each window significantly differs from that outside it (Naus, 1965; Kulldorff, 1999). The output consists of one or more statistically significant clusters, interpreted as areas with excess risk not explained by individual characteristics alone.

In this study, we employ a model-based scan statistic implemented in the R package `DClusterm` Gómez-Rubio et al., 2019), which builds on generalized linear models to identify spatial clusters. The method uses a series of dummy variables and performs likelihood ratio tests to compare nested models with and without the candidate cluster term. Previous studies have demonstrated how this method can be applied in different scenarios (Gómez-Rubio et al., 2018; Silan et al., 2023).

We first estimate a baseline Poisson regression model (null model), where the number of observed dementia cases per area is the dependent variable, and the number of individuals at risk is included as an offset term. This yields area-level dementia incidence rates. The model can also accommodate additional covariates, although the baseline model includes only an intercept.

Once the baseline model is fitted, the algorithm searches for spatial clusters by selecting a centroid (i.e., a starting area) and progressively adding neighbouring areas one by one. At each iteration, a new model is estimated that includes a binary term indicating whether an area belongs to the candidate cluster. A likelihood ratio test compares this extended model to the baseline. If the added area improves the model fit by indicating a significantly higher dementia incidence, it is retained in the cluster. The process continues until no further improvement is observed or no more neighbouring areas are available.

The algorithm can be customized using stopping rules, for example by setting thresholds on the maximum number of areas or total population allowed in a cluster. In this study, no population constraints were imposed during the spatial scan.

For each finalized cluster, `DClusterm` computes an approximate p-value based on a chi-squared distribution with one degree of freedom. Since no prior assumptions were made regarding the potential location of clusters, the algorithm was run from each area as a potential centroid. The final output includes only those clusters whose p-values fall below a pre-specified significance level, set at 0.10.

Results

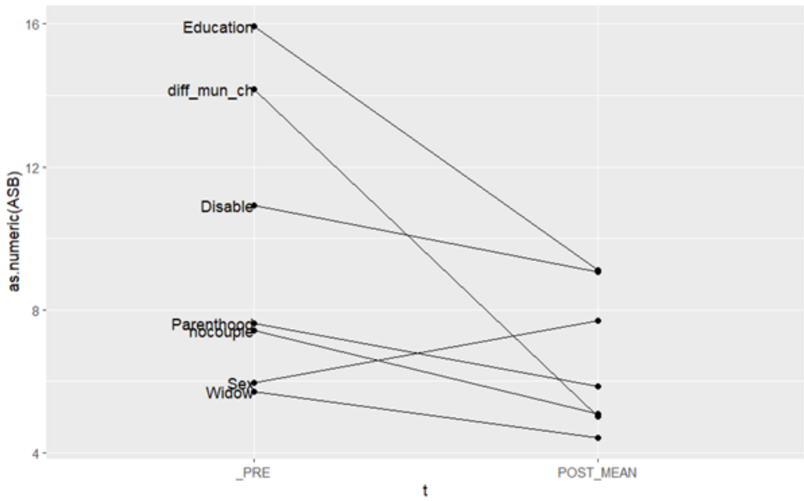
Social Health Covariate Balance and Dementia Incidence

To evaluate the effectiveness of the MARMoT procedure in balancing individual-level covariates across municipalities, we computed the Absolute Standardized Bias (ASB) for each variable, both on real data and on balanced data. Table 1 summarizes the distribution of ASB values across all covariates and areas, while Figure 2 illustrates the mean ASB per covariate before and after matching.

Table 1 Distribution of Absolute Standardized Bias (ASB) across municipalities, before and after MARMoT adjustment.

ASB	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Over 5%	Over 10%	Over 20%
Before	0.007	3.437	7.344	9.673	13.519	95.766	1,074	621	198
After	0.074	2.348	5.176	6.602	9.426	52.799	840	379	55

Figure 1 Mean Absolute Standardized Bias (ASB) by covariate, before and after MARMoT adjustment.



The results show a clear overall improvement in covariate balance following adjustment. In the unadjusted data, the median ASB was 7.34%, and the mean ASB 9.67%, suggesting notable heterogeneity in the distribution of individual-level characteristics across municipalities. Extreme values were also observed, with the maximum ASB reaching 95.8%, highlighting strong imbalance for certain covariates in specific areas.

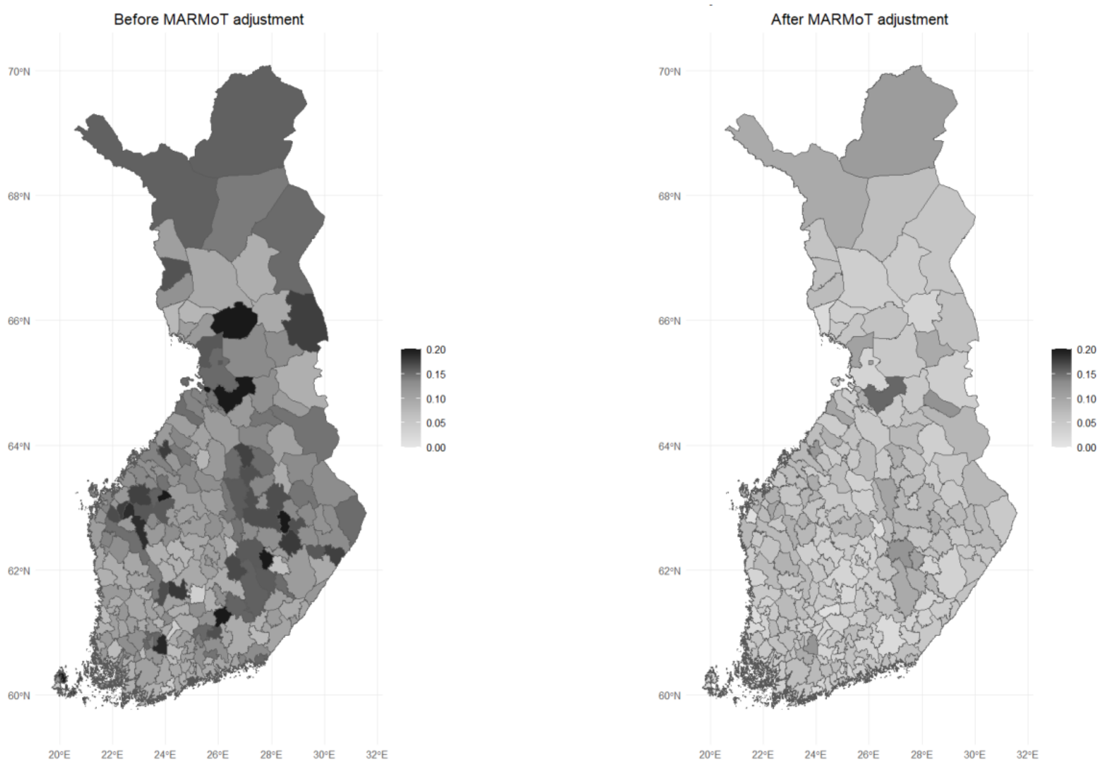
After MARMoT adjustment, the median ASB decreased to 5.18%, and the mean to 6.60%, indicating improved balance across the board. Although the maximum ASB remained high (52.8%), these cases were limited to a small number of outliers. Indeed, the share of municipalities with substantial imbalance, defined as ASB values exceeding commonly used thresholds, declined substantially after adjustment: the number of municipalities with ASB > 20% dropped from 198 to 55, those with ASB > 10% fell from 621 to 379, and those with ASB > 5% decreased from 1,074 to 840. These results confirm that the MARMoT matching procedure substantially reduced systematic differences in observed characteristics across municipalities, supporting the assumption of conditional independence for subsequent analyses.

As shown in Figure 2, education and municipality-level proximity to children (child proximity) were among the most imbalanced covariates before matching, with ASB values exceeding 14%. Both showed substantial reductions following the adjustment. Even for variables that were already relatively balanced, such as gender, partnership status, and widowhood, MARMoT further improved comparability across municipalities. While overall covariate balance improved, education and disability remained the most imbalanced variables after adjustment. Out of 237 municipalities, 99 and 89 still exhibited ASB values above 10% for education and disability, respectively; 40 and 45 exceeded 15%; and 21 and 19 exceeded 20%. Nonetheless, all variables met the commonly accepted criterion for adequate balance, with mean ASB below 10%, supporting the robustness of the matching procedure.

We also observed improvements in variables not directly included in the balancing procedure. For instance, the ASB for being divorced (yes/no) decreased from 11.5% to 8.4%, and for living alone (yes/no) from 8.1% to 7.3%. Although these covariates were not targeted during the matching process, the improvement in their balance underscores the broad stabilizing effect of the MARMoT algorithm.

Overall, the reduction in ASB values confirms that the matching procedure substantially improved balance on key socio-demographic and social health indicators, increasing the validity of subsequent comparisons of dementia incidence across municipalities.

Figure 2 illustrates the distribution of dementia incidence across Finnish municipalities, before (left panel)



In the unadjusted data, dementia incidence varied markedly across the country, with higher rates observed in northern and eastern Finland, and lower rates in central and southern regions. This spatial heterogeneity likely reflects underlying differences in population composition, particularly in terms of characteristics associated with social health.

Following adjustment for individual-level covariates using the MARMoT procedure, spatial disparities and overall dementia incidence were substantially reduced. The adjusted map reveals a more homogeneous distribution of dementia incidence across municipalities, suggesting that much of the observed variation in the raw data was attributable to compositional differences in socio-demographic and social health factors.

Importantly, under the assumption of unconfoundedness, any remaining differences in dementia incidence across municipalities after MARMoT adjustment can be attributed to contextual influences and other residual unmeasured individual-level factors. As shown in Figure 1 (right panel), only a few municipalities continue to exceed an incidence of 0.20 after adjustment.

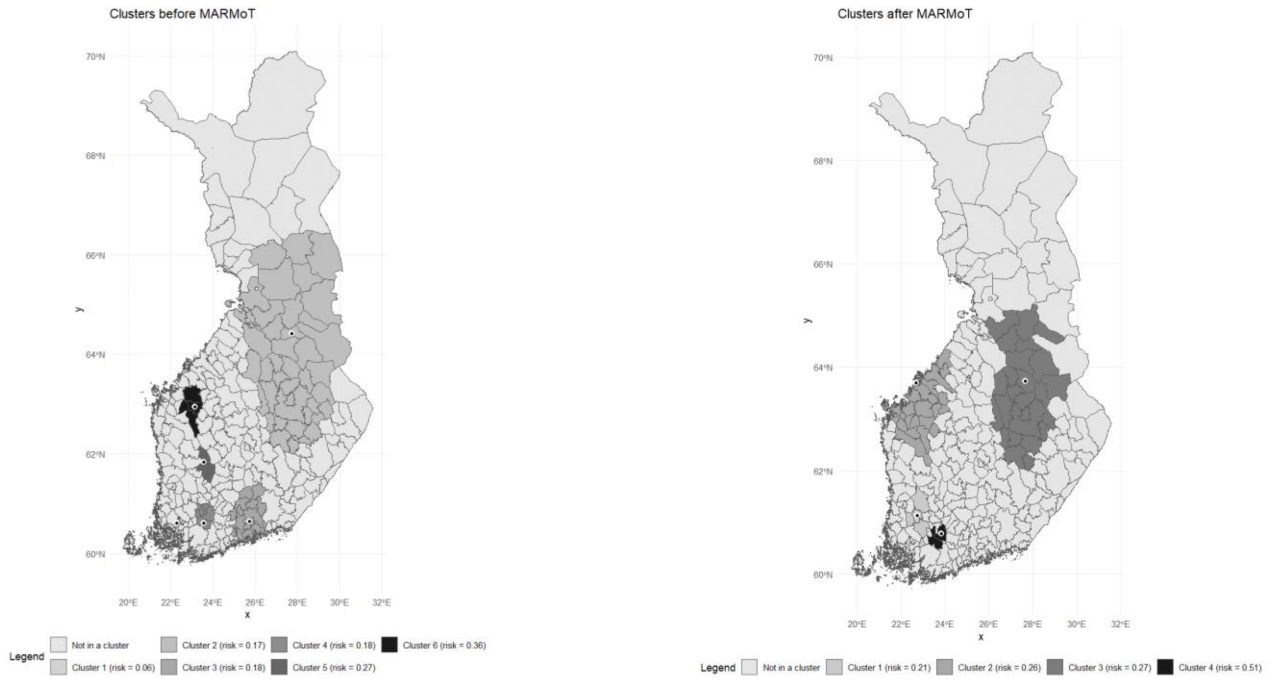
The post-balancing incidence can be interpreted as the expected incidence of dementia if all municipalities shared the same individual-level risk profile, in terms of the selected social health and sociodemographic characteristics. Therefore, the observed reduction in overall dementia risk and its spatial heterogeneity can be seen as an estimate of the extent to which dementia cases might be prevented by reducing inequalities in individual characteristics and their social health.

Spatial clustering of dementia incidence before and after adjustment

We applied spatial scan statistics to identify contiguous clusters of municipalities with elevated dementia incidence, both before and after adjusting for individual-level characteristics related to social health. Comparing the unadjusted and adjusted models allows us to assess whether high-risk clusters are primarily driven by the composition of the population in terms of individual sociodemographic and social health characteristics, or whether they reflect areas of intrinsically elevated dementia risk.

In the unadjusted model (Figure 3, left panel), six spatial clusters of high dementia incidence were identified. The largest cluster was located in the eastern-central region of the country and included several neighbouring municipalities with relative risks ranging from 0.18 to 0.36. Additional clusters were detected in southern Finland, including the Helsinki metropolitan area, and south-western and central-western parts of the country.

Figure 3 Spatial clusters of high dementia incidence before MARMoT (left) and after (right) adjustment.



After MARMoT adjustment, the spatial clustering pattern changed substantially. The model identified four clusters, with relative risks ranging from 0.21 to 0.51. First, the large central-eastern persisted, although its geographic extent was reduced.

This suggests that, while part of the elevated dementia risk in this area was attributable to compositional factors related to sociodemographic and social health characteristics, a residual excess persists. In other words, this area continues to exhibit an excess risk not explained by these individual-level factors, which may point to the influence of local environmental conditions or other structural and unmeasured individual factors. Second, the southern cluster around Helsinki, prominent before adjustment, disappeared after balancing. This indicates that the initially observed elevated risk in this urban area was largely due to differences in population composition, specifically a larger share of disadvantaged or socially vulnerable groups. Finally, a new central-west cluster emerged after adjustment, an area that was not identified in the unadjusted model. This pattern reveals a hidden elevated dementia risk that became visible only after accounting for measured compositional differences.

To further investigate the role of contextual factors, we replicated the spatial scan analysis including the four area-level indicators of social infrastructure (i.e., availability of healthcare professionals, social workers, cultural venues, and retail stores) in the MARMoT adjustment. The inclusion of these variables did not substantially alter the spatial clustering results, and we thus do not report these additional findings.

Discussions and conclusions

This study advances our understanding of the geography of dementia risk in ageing societies by using a spatial analytical framework to disentangle individual socio-demographic and social health characteristics and contextual factors. We applied a novel balancing method (MARMoT) to equalize the socio-demographic and social health composition of local populations, thereby isolating the role of these factors in shaping the patterns of dementia risk across Finnish municipalities. This was followed by spatial scan statistics to map high-risk clusters across Finnish municipalities.

Our findings underscore the importance of adjusting for individual-level confounders when interpreting spatial patterns of dementia. In fact, we found that, in unadjusted models, elevated risk in areas such as the Helsinki metropolitan one was fully explained by compositional factors, namely, a higher concentration of disadvantaged or socially vulnerable population subgroups. Once these differences were accounted for, the apparent excess risk in this area disappeared. Conversely, in other areas, particularly parts of eastern Finland, excess dementia risk persisted even after adjustment, suggesting the presence of residual effects. These may stem from unmeasured environmental exposures, differences in access to or quality of services, or the accumulation of disadvantages over the life course. While the specific mechanisms underlying these geographic disparities remain unclear, our findings indicate that they cannot be fully accounted for by adult socio-economic status or selected indicators of social health. This underscores the need for further interdisciplinary research that integrates environmental, institutional, and life-course perspectives.

Finally, a new high-risk cluster in central-western Finland emerged only after adjusting for population composition. This suggests that the area is, in fact, characterized by a comparatively favourable socio-demographic and social health profile. Once these compositional characteristics are accounted for, a previously hidden excess risk becomes visible, indicating that other unmeasured factors may be driving a potential elevated dementia incidence in this area.

Our aim was also to understand the role in shaping dementia risk given by the contextual factors, such as the characteristics of the municipality. Although the availability of local social infrastructure is commonly assumed to shape opportunities for social engagement and support in later life, the inclusion of area-level indicators in the adjustment did not substantially change the spatial clustering patterns of dementia risk in Finland. This suggests that, in this case, the observed geographic variation in dementia risk may be more strongly driven by the composition of individual characteristics than by the measured contextual ones. However, this result should be interpreted with caution, as it may also reflect limitations in the specificity or granularity of the available area-level indicators (e.g., we cannot measure service quality and satisfaction).

Beyond its substantive findings, this study also contributes methodologically by demonstrating the potential of MARMoT for spatial health research. Most studies rely on regression-based adjustment, which assumes that confounders are equally distributed among the different geographic units. In contrast, our matching-based approach generates a synthetic population balanced on key covariates. Moreover, the use of spatial scan statistics enables the detection of contiguous clusters rather than isolated high-risk units, providing more actionable insights for public health surveillance and territorial policy planning. This approach supports efforts to reduce area-level disparities and improve overall population health across the country.

One of the key insights of this study is that in areas where excess dementia risk is largely driven by selected individual-level compositional factors, targeted interventions to improve social health could contribute to reducing dementia incidence. In other words, addressing social vulnerability at the population level may yield measurable health gains. Conversely, in areas with a favourable population composition, elevated dementia risk may remain hidden and thus go undetected if one only considers individual-level profiles. This underscores the importance of combining compositional and contextual analyses to fully capture dementia risk and its spatial inequalities.

From a policy perspective, the results support the need for integrated approaches to dementia prevention that include targeted individual-level risk factors, promoting social health to further mitigate dementia risk at the population level and their geographical differences.

Identifying territorial clusters where populations experience poorer health outcomes is a key concern across multiple disciplines, including public health, social epidemiology, and health geography, as it offers critical insights into spatial and social health inequalities and informs evidence-based policy interventions. However, achieving this objective poses significant challenges due to the non-random nature of residential decisions and mobility. Individuals often self-select into specific areas based on personal characteristics or resources, leading to selection bias and complicating causal interpretation. In this study, we address this challenge by means of the MARMoT, improving the validity of comparisons between geographic contexts.

This study has some limitations. Although the Finnish register system provides rich and high-quality data, our operationalization of social health was constrained by the data availability. Important dimensions such as loneliness, informal caregiving dynamics, or perceived social support, typically included in survey-based measures of social health (e.g., Vernooij-Dassen and Jeon 2016; Vernooij-Dassen et al. 2021; Vernooij-Dassen et al., 2022) could not be directly captured.

In addition, although our matching strategy enhances causal interpretation, unmeasured confounding cannot be entirely ruled out. Future research could benefit from integrating further contextual indicators (e.g., air pollution, access to green spaces, or local service quality), as well as biological and health-related dementia risk factors, to better understand the determinants of cognitive ageing and its spatial inequalities.

Importantly, to our knowledge, this is the first study to apply the concept of social health to register data covering the whole population of a country. Despite the constraints of administrative sources, our findings demonstrate the potential of leveraging register-based indicators to study social vulnerability and its implications for dementia risk. This approach, combined with spatial methods, offers valuable tools for identifying social and place-based health disparities and informing targeted public health interventions in ageing societies.

References

Antonucci, T., Zahodne, L., Harrell, E., 2024. ADRD in context: Social, neighborhood, and national influences. *Innov. Aging* 8(Suppl_1), 14–15. https://doi.org/10.1093/geroni/igae098.0043

Bister, L., Balbo, N., Remes, H., Neri, E., & Martikainen, P. (2025) A novel framework to define child disability from a biopsychosocial perspective in population register data: the Degree-of-Limitation Index (DOLI) for child disability. *Manuscript in preparation. For the working paper, please contact the corresponding author Lara Bister:* lara.bister@wzb.eu

Bellou, V., Belbasis, L., Tzoulaki, I., Middleton, L., Ioannidis, J., Evangelou, E., 2017. Systematic evaluation of the associations between environmental risk factors and dementia: An umbrella review of systematic reviews and meta-analyses. *Alzheimers Dement.* 13, 406–418. https://doi.org/10.1016/j.jalz.2016.07.152

Brayne, C., Gao, L., Dewey, M., Matthews, F.E., MRC Cognitive Function and Ageing Study Investigators, 2006. Dementia before death in ageing societies—The promise of prevention and the reality. PLoS Med. 3(10), e397. https://doi.org/10.1371/journal.pmed.0030397

Brüggemann, R., & Carlsen, L. (2011). An improved estimation of averaged ranks of partial orders. MATCH Communications in Mathematical and in Computer Chemistry, 65, 383–414.

Caperna, G., 2019. Approximation of AverageRank by means of a formula. *Zenodo*. https://doi.org/10.5281/zenodo.2565699

Clarke, P.J., Ailshire, J.A., House, J.S., Morenoff, J.D., King, K., Melendez, R., Langa, K.M., 2012. Cognitive function in the community setting: The neighbourhood as a source of 'cognitive reserve'? *J. Epidemiol. Community Health* 66, 730–736. https://doi.org/10.1136/jech.2010.128116

Clarke, P.J., Weuve, J., Barnes, L., Evans, D.A., Mendes de Leon, C.F., 2015. Cognitive decline and the neighborhood environment. Ann. Epidemiol. 25(11), 849–854. https://doi.org/10.1016/j.annepidem.2015.07.001

Crimmins, E.M., Saito, Y., Kim, J.K., Zhang, Y.S., Sasson, I., Hayward, M.D., 2018. Educational differences in the prevalence of dementia and life expectancy with dementia: Changes from 2000 to 2010. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 73(Suppl_1), S20–S28. https://doi.org/10.1093/geronb/gbx135

De Loof, K., De Baets, B., De Meyer, H., 2011. Approximation of average ranks in posets. *MATCH Commun. Math. Comput. Chem.* 66, 219–229.

Downs, M., 2000. Dementia in a socio-cultural context: An idea whose time has come. *Ageing Soc.* 20, 369–375. https://doi.org/10.1017/S0144686X99007758

Duffner, L.A., Deckers, K., Cadar, D., Steptoe, A., de Vugt, M., Köhler, S., 2022. The role of cognitive and social leisure activities in dementia risk: Assessing longitudinal associations of modifiable and non-modifiable risk factors. *Epidemiol. Psychiatr. Sci.* 31, e5. https://doi.org/10.1017/S204579602100069X

Einiö, E., Nisen, J.T., Martikainen, P.T., 2016. Number of children and later-life mortality among Finns born 1938–50. *Popul. Stud.* 70, 217–238. https://doi.org/10.1080/00324728.2016.1162914

Finlay, J., Esposito, M., Li, M., Kobayashi, L.C., Khan, A.M., Gomez-Lopez, I., Melendez, R., Colabianchi, N., Judd, S., Clarke, P.J., 2021. Can neighborhood social infrastructure modify cognitive function? A mixed-methods study of urban-dwelling aging Americans. *J. Aging Health* 33, 772–785. https://doi.org/10.1177/08982643211008673

Freak-Poli, R., Wagemaker, N., Wang, R., Lysen, T.S., Ikram, M.A., Vernooij, M.W., et al., 2022. Loneliness, not social support, is associated with cognitive decline and dementia across two longitudinal population-based cohorts. *J. Alzheimers Dis.* 85, 295–308. https://doi.org/10.3233/JAD-210330

Gómez-Rubio, V., Moraga, P., Molitor, J., Rowlingson, B., 2019. DClusterm: Model-based detection of disease clusters. *J. Stat. Softw.* 90(14), 1–26. https://doi.org/10.18637/jss.v090.i14

Gómez-Rubio, V., Molitor, J., Moraga, P., 2018. Fast Bayesian classification for disease mapping and the detection of disease clusters, in: Cameletti, M., Finazzi, F. (Eds.), *Quantitative Methods in Environmental and Climate Research*. Springer, Cham, pp. 1–20. https://doi.org/10.1007/978-3-030-01584-8_1

Hahn, C., Lee, C., 2019. A brief review of paradigm shifts in prevention of Alzheimer's disease: From cognitive reserve to precision medicine. *Front. Psychiatry* 10, 786. https://doi.org/10.3389/fpsyt.2019.00786

Hudomiet, P., Hurd, M.D., Rohwedder, S., 2022. Trends in inequalities in the prevalence of dementia in the United States. *Proc. Natl. Acad. Sci. U.S.A.* 119(46), e2212205119. https://doi.org/10.1073/pnas.2212205119

Joling, K.J., Janssen, O., Francke, A.L., et al., 2020. Time from diagnosis to institutionalization and death in people with dementia. *Alzheimers Dement.* 16, 662–671. https://doi.org/10.1002/alz.12063

Jones, I.R., 2017. Social class, dementia and the fourth age. *Sociol. Health Illn.* 39, 303–317. https://doi.org/10.1111/1467-9566.12520

Joyce, J., Ryan, J., Owen, A., Hu, J., McHugh Power, J., Shah, R., et al., 2021. Social isolation, social support, and loneliness and their relationship with cognitive health and dementia. *Int. J. Geriatr. Psychiatry* 37. https://doi.org/10.1002/gps.5644

Kelly, M., Duff, H., Kelly, S., McHugh Power, J., Brennan, S., Lawlor, B., et al., 2017. The impact of social activities, social networks, social support and social relationships on the cognitive functioning of healthy older adults: A systematic review. *Syst. Rev.* 6, 259. https://doi.org/10.1186/s13643-017-0632-2

Kuiper, J., Zuidersma, M., Oude Voshaar, R., Zuidema, S., van den Heuvel, E., Stolk, R., et al., 2015. Social relationships and risk of dementia: A systematic review and meta-analysis of longitudinal cohort studies. *Ageing Res. Rev.* 22, 39–57. https://doi.org/10.1016/j.arr.2015.04.006

Kulldorff, M., 1999. Spatial scan statistics: Models, calculations, and applications, in: Glaz, J., Balakrishnan, N. (Eds.), *Scan Statistics and Applications*. Statistics for Industry and Technology. Birkhäuser, Boston, MA, pp. 303–322. https://doi.org/10.1007/978-1-4612-1578-3_14

Lenart-Bugla, M., Łuc, M., Pawłowski, M., Szcześniak, D., Seifert, I., Wiegelmann, H., et al., 2022. What do we know about social and non-social factors influencing the pathway from cognitive health to dementia? A systematic review of reviews. *Brain Sci.* 12, 1214. https://doi.org/10.3390/brainsci12091214

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S., Huntley, J., Ames, D., et al., 2017. Dementia prevention, intervention, and care. *Lancet* 390, 2673–2734. https://doi.org/10.1016/S0140-6736(17)31363-6

Luth, E.A., Prigerson, H.G., 2022. Socioeconomic status, race/ethnicity, and unexpected variation in dementia classification in longitudinal survey data. J. Gerontol. B Psychol. Sci. Soc. Sci. 77(12), e234–e246. https://doi.org/10.1093/geronb/gbac128

Ma, L., Sun, F., Tang, Z., 2018. Social frailty is associated with physical functioning, cognition, and depression, and predicts mortality. *J. Nutr. Health Aging* 22, 989–995. https://doi.org/10.1007/s12603-018-1054-0

Mackenzie, C. S., Smith, M. C., Hasher, L., Leach, L., & Behl, P. (2007). Cognitive functioning under stress: Evidence from informal caregivers of palliative patients. *Journal of Palliative Medicine*, *10*(3), 749–758. https://doi.org/10.1089/jpm.2006.0171

Maddock, J., Gallo, F., Wolters, F.J., Stafford, J., Marseglia, A., Dekhtyar, S., et al., 2023. Social health and change in cognitive capability among older adults: Findings from four European longitudinal studies. *Gerontology* 69(11), 1330–1346. https://doi.org/10.1101/2022.08.29.22279324

Marseglia, A., Kalpouzos, G., Laukka, E.J., Maddock, J., Patalay, P., Wang, H.X., et al., 2023. Social health and cognitive change in old age: The role of brain reserve. *Ann. Neurol.* 93(4), 844–855. https://doi.org/10.1002/ana.26591

Michael, YL., Yen, IH. 2014. Aging and place-neighborhoods and health in a world growing older. J Aging Health 26(8):1251–1260. https://doi.org/10.1177/0898264314562148

Naus, J.L., 1965. Clustering of random points in two dimensions. *Biometrika* 52(1–2), 263–266. https://doi.org/10.1093/biomet/52.1-2.263

Opdebeeck, C., Matthews, F., Wu, Y., Woods, R., Brayne, C., Clare, L., 2018. Cognitive reserve as a moderator of the negative association between mood and cognition: Evidence from a population-representative cohort. *Psychol. Med.* 48, 61–71. https://doi.org/10.1017/S003329171700126X

Penninkilampi, R., Casey, A., Singh, M., Brodaty, H., 2018. The association between social engagement, loneliness, and risk of dementia: A systematic review and meta-analysis. *J. Alzheimers Dis.* 66, 1619–1633. https://doi.org/10.3233/JAD-180439

Piolatto, M., Bianchi, F., Rota, M., Marengoni, A., Akbaritabar, A., Squazzoni, F., 2022. The effect of social relationships on cognitive decline in older adults: An updated systematic review and meta-analysis of longitudinal cohort studies. *BMC Public Health* 22, 278. https://doi.org/10.1186/s12889-022-12567-5

Saito, T., Murata, C., Saito, M., Takeda, T., Kondo, K., 2018. Influence of social relationship domains and their combinations on incident dementia: A prospective cohort study. *J. Epidemiol. Community Health* 72, 7–12. https://doi.org/10.1136/jech-2017-209811

Samtani, S., Mahalingam, G., Lam, B.C.P., Lipnicki, D.M., Lima-Costa, M.F., Blay, S.L., et al., 2022. Associations between social connections and cognition: A global collaborative individual participant data meta-analysis. *Lancet Healthy Longev.* 3, e740–e753. https://doi.org/10.1016/S2666-7568(22)00199-4

Schulz, R., Beach, S. R., Czaja, S. J., Martire, L. M., & Monin, J. K. (2020). Family caregiving for older adults. *Annual Review of Psychology*, *71*, 635–659. https://doi.org/10.1146/annurev-psych-010419-050754

Silan, M., Belloni, P., Boccuzzo, G., 2023. Identification of neighborhood clusters on data balanced by a poset-based approach. *Stat. Methods Appl.* 32, 1295–1316. https://doi.org/10.1007/s10260-023-00695-0

Silan, M., Boccuzzo, G., & Arpino, B. (2021). Matching on poset-based average rank for multiple treatments to compare many unbalanced groups. *Statistics in Medicine*, *40*(28), 6443-6458. https://doi.org/10.1002/sim.9192

Sommerlad, A., Sabia, S., Singh-Manoux, A., Lewis, G., Livingston, G., 2019. Association of social contact with dementia and cognition: 28-year follow-up of the Whitehall II cohort study. *PLOS Med.* 16, e1002862. https://doi.org/10.1371/journal.pmed.1002862

Sutin, A.R., Stephan, Y., Luchetti, M., Terracciano, A., 2018. Loneliness and risk of dementia. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 75, 1414–1422. https://doi.org/10.1093/geronb/gby112

Tsutsumimoto, K., Doi, T., Makizako, H., Hotta, R., Nakakubo, S., Makino, K., et al., 2017. Association of social frailty with both cognitive and physical deficits among older people. *J. Am. Med. Dir. Assoc.* 18, 603–607. https://doi.org/10.1016/j.jamda.2017.02.004

van der Velpen, I.F., Melis, R.J.F., Perry, M., Vernooij-Dassen, M.J.F., Ikram, M.A., Vernooij, M.W., 2022. Social health is associated with structural brain changes in older adults: The Rotterdam study. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 7, 659–668. https://doi.org/10.1016/j.bpsc.2021.01.009

Vernooij-Dassen, M., Jeon, Y.H., 2016. Social health and dementia: The power of human capabilities. Int. Psychogeriatr. 28(5), 701–703. https://doi.org/10.1017/S1041610216000260

Vernooij-Dassen, M., Moniz-Cook, E., Jeon, Y.-H., 2018. Social health in dementia care: harnessing an applied research agenda. *Int. Psychogeriatr.* 30(6), 775–778. https://doi.org/10.1017/S1041610217002769

Vernooij-Dassen, M., Moniz-Cook, E., Verhey, F., Chattat, R., Woods, B., Meiland, F., & de Vugt, M. (2021). Bridging the divide between biomedical and psychosocial approaches in dementia research: The 2019 INTERDEM manifesto. *Aging & Mental Health, 25*(2), 206–212. https://doi.org/10.1080/13607863.2019.1693968

Vernooij-Dassen, M., Verspoor, E., Perry, M., Wolf-Ostermann, K., 2022. Conceptual advancement for social health in dementia research. *Innov. Aging* 6(Suppl 1), 434–435. https://doi.org/10.1093/geroni/igac059.1705

Wang, H.X., MacDonald, S., Dekhtyar, S., Fratiglioni, L., 2017. Association of lifelong exposure to cognitive reserve-enhancing factors with dementia risk: A community-based cohort study. *PLOS Med.* 14, e1002251. https://doi.org/10.1371/journal.pmed.1002251

Winblad, B., Amouyel, P., Andrieu, S., Ballard, C., Brayne, C., Brodaty, H., et al., 2016. Defeating Alzheimer's disease and other dementias: A priority for European science and society. *Lancet Neurol.* 15, 455–532. https://doi.org/10.1016/S1474-4422(16)00062-4

World Health Organization (WHO), 1946. *Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference*. New York, NY: WHO.

World Health Organization (WHO), 2021. *Global Dementia Observatory data portal*. https://www.who.int/data/gho/data/themes/global-dementia-observatory-gdo

Wu, Y.T., Prina, A.M., Brayne, C., 2015. The association between community environment and cognitive function: A systematic review. *Soc. Psychiatry Psychiatr. Epidemiol.* 50, 351–362. https://doi.org/10.1007/s00127-014-0945-6

Yang, Y., Yeung, W., Feng, Q., 2018. Social exclusion and cognitive impairment – A triple jeopardy for Chinese rural elderly women. *Health Place* 53, 117–127. https://doi.org/10.1016/j.healthplace.2018.07.013