

Bayesian Spatial Risk Modelling of Suicide Mortality in South Korea

1. Background

South Korea consistently reports one of the highest suicide rates among developed nations, with 24 to 28 deaths per 100,000 people in recent years—more than twice the OECD average of 11 (Kim, 2024). This persistent trend positions South Korea as a global outlier and underscores the urgent need to investigate its underlying risk factors. Classical theories dating back to Durkheim (1951) emphasise the protective role of social integration, while recognising economic hardship and mental illness as key drivers of suicide. In South Korea, national crises such as the late-1990s Asian financial crash have been linked to suicide spikes, primarily through rising unemployment and social dislocation (Chang, 2009).

The analysis draws on a Bayesian Intrinsic Conditional Auto-Regressive (ICAR) model to estimate spatial variation in suicide risk across South Korea, using four key covariates: social isolation, psychological stress, unemployment, and unmet medical needs. Areas with more single-person households or divorced residents tend to report higher suicide rates, with higher rates observed in areas with weaker social ties (Jang, 2022). High levels of perceived stress, measured through the “stress awareness” indicator, are associated with depression, suicidal ideation, and increased regional suicide rates (Oh, 2020). Unemployment contributes to financial strain and psychological distress, and its association with higher suicide rates is particularly evident where social safety nets are weak (Park & Lester, 2006). Lastly, unmet medical needs—reflecting limited healthcare access—are correlated with higher suicide mortality, with only about 15% of affected individuals receiving treatment despite the high prevalence of mental health conditions (Kim, 2022).

To account for spatial variation in suicide risk, the Bayesian ICAR model incorporates information from adjacent regions to produce stable small-area risk estimates. Key outputs include both localised estimates of relative risk and global regression coefficients for the four covariates, providing an evidence base to inform public health interventions and guide regional resource allocation.

2. Data and Methods

2.1 Study Area

The study area encompasses all 229 municipalities in South Korea, including cities (*shi*), counties (*kun*), and districts (*ku*), collectively referred to as *shikunku*-level administrative units. This spatial resolution was selected as it represents the most granular level at which all covariates are consistently available and corresponds to the administrative scale at which suicide prevention policies are typically developed and implemented. Spatial boundaries were derived from official 2022 shapefiles, to which all statistical datasets were harmonised by resolving naming inconsistencies, aggregating sub-municipal divisions, and reversing backdated administrative updates. To ensure full connectivity in the Queen’s contiguity adjacency matrix, disconnected island polygons were linked to the mainland subgraph using nearest-neighbour centroids.

2.2 Data

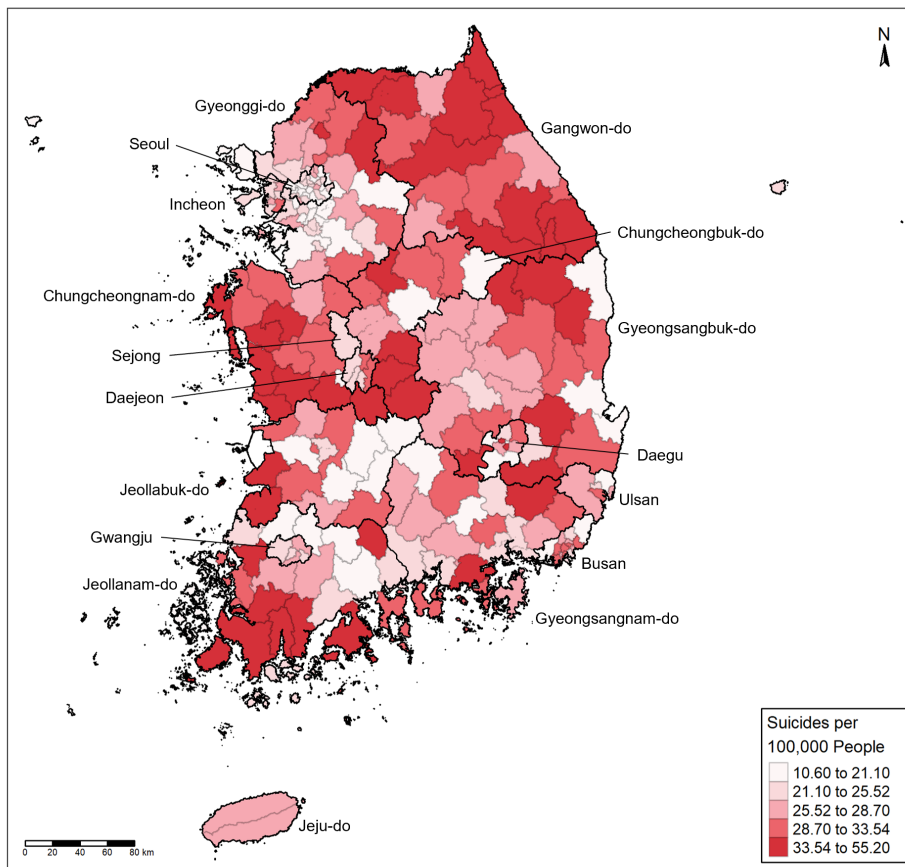
The primary outcome variable is the estimated number of suicide deaths in 2022 at the municipality (*shikunku*) level. This was calculated by multiplying the age-standardised suicide rate (per 100,000 people) by the mid-year population for each municipality. Suicide rates were obtained from Statistics Korea’s *Cause of Death Statistics*, while population estimates originate from the 2022 *Future Population Projections*. This indirect method was necessary due to the absence of disaggregated suicide counts at this spatial resolution.

Four predictor variables were selected based on their relevance in the literature: the single-person household ratio, stress awareness rate, unemployment rate, and unmet medical needs rate. All variables reflect conditions in 2022 and were obtained from national databases, as summarised in Table 1. Figure 1 provides a visual overview of suicide rates across the country.

Table 1: Summary of Datasets

Variable	Dataset	Usage	Source	Year
Suicide Rate	Cause of Death Statistics	Target variable	Statistics Korea (Population Trends Division)	2022
Single-Person Household Ratio	Population and Housing Census	Predictor variable	Statistics Korea (Population and Housing Census Division)	2022
Stress Awareness Rate	Community Health Survey	Predictor variable	Korea Disease Control and Prevention Agency (KDCA)	2022
Unemployment Rate	Local Employment Statistics	Predictor variable	Statistics Korea (Employment Statistics Division)	2022
Unmet Medical Needs Rate	Community Health Survey	Predictor variable	Korea Disease Control and Prevention Agency (KDCA)	2022
Estimated Population	Future Population Projections	Standardisation / Denominator	Statistics Korea (Regional Statistics Planning Division)	2022
Municipality Boundaries	Municipality Shapefiles	Spatial analysis	Statistical Geographic Information System (SGIS)	2022
Province Boundaries	Province Shapefiles	Mapping / Labelling	Statistical Geographic Information System (SGIS)	2022

Figure 1: Suicide Rates Across Municipalities



2.3 Methods

We employed a Bayesian Poisson regression model with spatial effects to estimate the relative risk (RR) of suicide mortality across South Korean municipalities. To account for spatial autocorrelation among neighbouring areas, a structured spatial random effect was included using an intrinsic conditional autoregressive (ICAR) prior (Besag et al., 1991), a standard approach in spatial epidemiology. An unstructured random effect was also added to capture residual heterogeneity not explained by spatial structure or covariates.

The model assumes a Poisson likelihood with a log link and a population-based offset, which is appropriate for modelling rare event counts in small-area epidemiological studies. Spatial adjacency was defined using Queen contiguity, where municipalities sharing either a boundary or vertex were considered neighbours.

$$Y_i \sim \text{Poisson}(\lambda_i), \quad \log(\lambda_i) = \log(E_i) + \alpha + \mathbf{X}_i\boldsymbol{\beta} + \sigma \cdot \phi_i^1 \quad (1)$$

Where Y_i denotes the observed number of suicide deaths, E_i the expected count based on population size, α the intercept, \mathbf{X}_i the vector of covariates for municipality i , $\boldsymbol{\beta}$ the corresponding coefficients, and ϕ_i the spatial random effect scaled by σ . Structured (ϕ_i) and unstructured (θ_i) effects are combined through a weighted average governed by ρ .

We specified the following priors:

- $\alpha \sim \mathcal{N}(0, 1)$ *(intercept)*
- $\boldsymbol{\beta} \sim \mathcal{N}(0, 1)$ *(covariate coefficients)*
- $\theta_i \sim \mathcal{N}(0, 1)$ *(unstructured random effects)*
- $\phi \sim \text{ICAR}(N, \text{node1}, \text{node2})$ *(structured spatial random effects)*
- $\sum_{i=1}^N \phi_i \sim \mathcal{N}(0, 0.001 \cdot N)$ *(sum-to-zero constraint on ϕ to ensure identifiability)*
- $\sigma \sim \mathcal{N}(0, 1)$ *(scaling factor for spatial effects)*
- $\rho \sim \text{Beta}(0.5, 0.5)$ *(mixing proportion between structured and unstructured effects)*

All priors were weakly informative, centred around zero to regularise estimates while allowing sufficient flexibility to reflect uncertainty without dominating the posterior. Model fitting was performed in *Stan* using the *rstan* interface in R. We ran six parallel chains with 20,000 iterations each (10,000 warm-up), using the No-U-Turn Sampler (NUTS) with a maximum tree depth of 12.

¹In the full model, we include both structured and unstructured random effects: $\log(\lambda_i) = \log(E_i) + \alpha + \mathbf{X}_i\boldsymbol{\beta} + \sigma \cdot (\sqrt{1-\rho}\theta_i + \sqrt{\rho}\phi_i)$, where $\theta_i \sim \mathcal{N}(0, 1)$ captures unstructured heterogeneity and ϕ_i follows an ICAR prior for spatial structure.

3. Results and Discussion

3.1 Global Parameter Estimates

Table 2: Posterior Estimates of Global Parameters

Parameter	Mean	SE Mean	2.5% CrI	97.5% CrI	n_{eff}	\hat{R}
α	-0.1956	0.0008	-0.4548	0.0685	25689.1	1.0005
β_1	0.0120	0.0000	0.0067	0.0174	27152.2	1.0003
β_2	-0.0052	0.0000	-0.0139	0.0035	26780.3	1.0001
β_3	-0.0385	0.0002	-0.0706	-0.0065	9776.5	1.0007
β_4	0.0089	0.0000	-0.0014	0.0190	28092.1	1.0001
σ	0.2118	0.0003	0.1611	0.2685	7507.3	1.0006

Note. SE Mean = Standard Error of Posterior Mean; CrI = Credible Interval; n_{eff} indicates the effective sample size. \hat{R} is the Gelman–Rubin convergence diagnostic, where values near 1.00 suggest good convergence.

The model converged successfully with all \hat{R} values below 1.05 and large effective sample sizes, indicating stable posterior estimates. The global intercept ($\alpha = -0.20$) corresponds to a relative risk (RR) of $\exp(-0.20) \approx 0.82$ (95% CrI: 0.63 to 1.07), suggesting slightly lower-than-expected baseline suicide risk. However, the result is not statistically significant, as the credible interval includes the null value of 1.00.

Among the covariates, the unemployment rate (β_3) showed the strongest and statistically significant association: a one-unit increase corresponds to a 4% decrease in suicide risk ($RR \approx 0.96$; 95% CrI: 0.93 to 0.99). The single-person household ratio (β_1) and unmet medical needs rate (β_4) had weak positive associations ($RR \approx 1.01$), but their credible intervals approached or included 1.00, indicating limited significance. Stress awareness (β_2) was negatively associated with suicide risk ($RR \approx 0.99$), though the effect was marginal.

The spatial standard deviation ($\sigma = 0.21$; 95% CrI: 0.16 to 0.27) indicates a meaningful level of spatial heterogeneity, supporting the inclusion of structured spatial effects in the model.

3.2 Spatial Distribution of Relative Risk

Figure 2: Map of Relative Risk Ratios

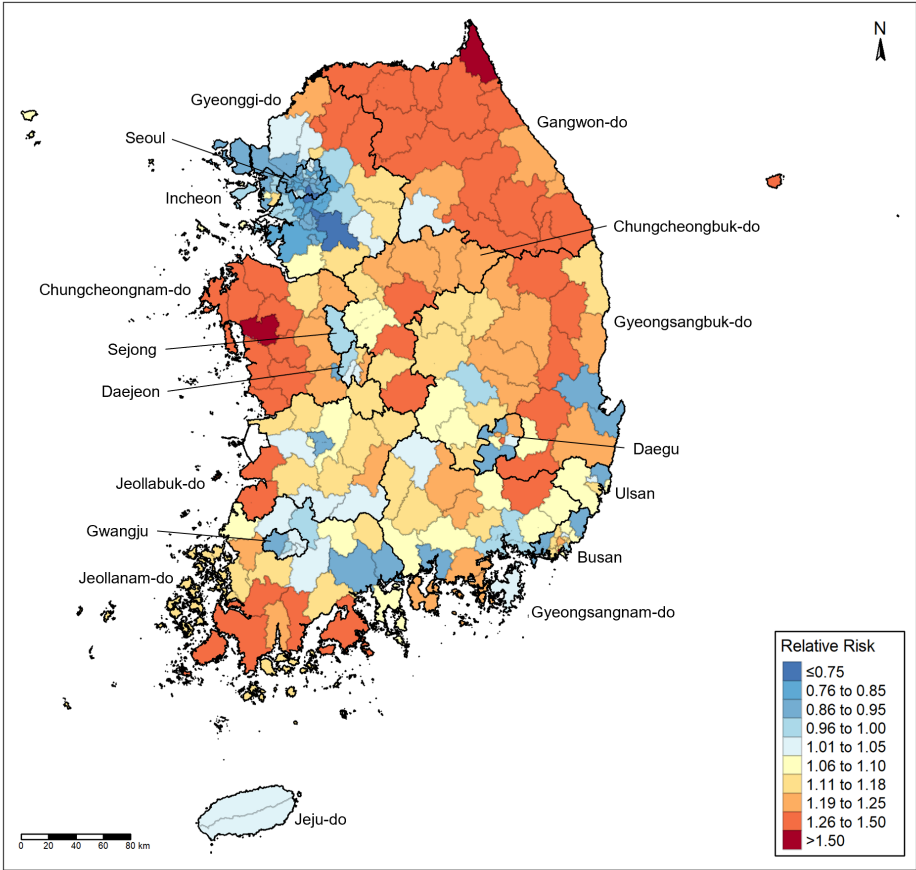


Figure 3: Map of Statistical Significance

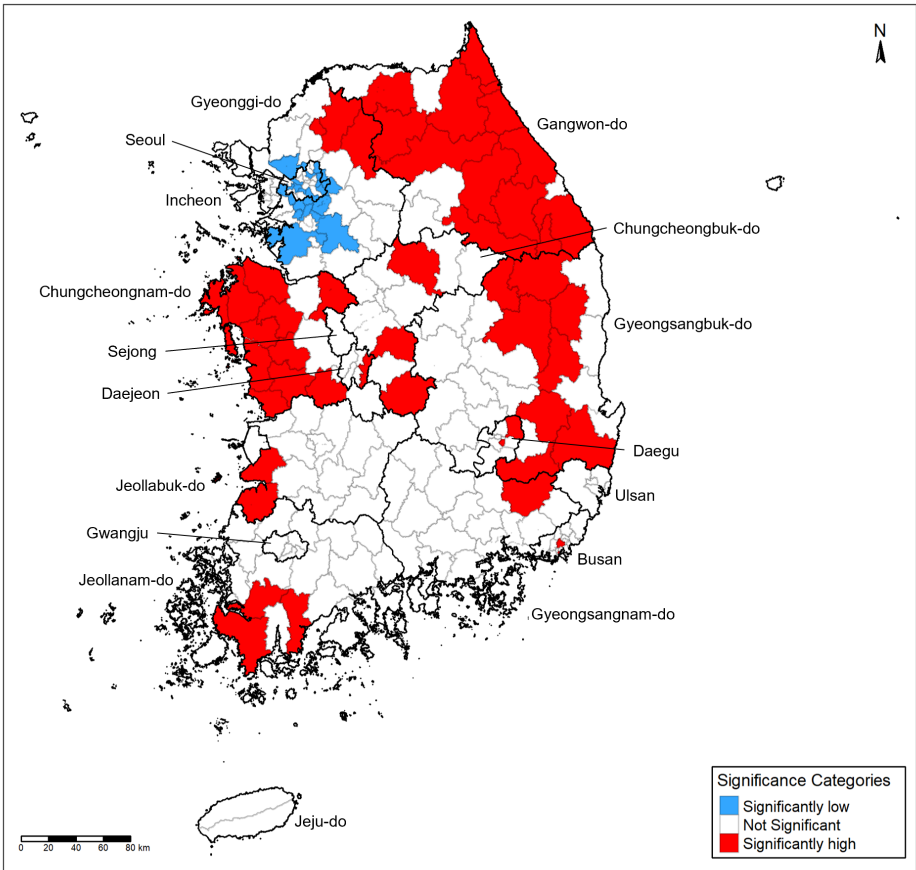
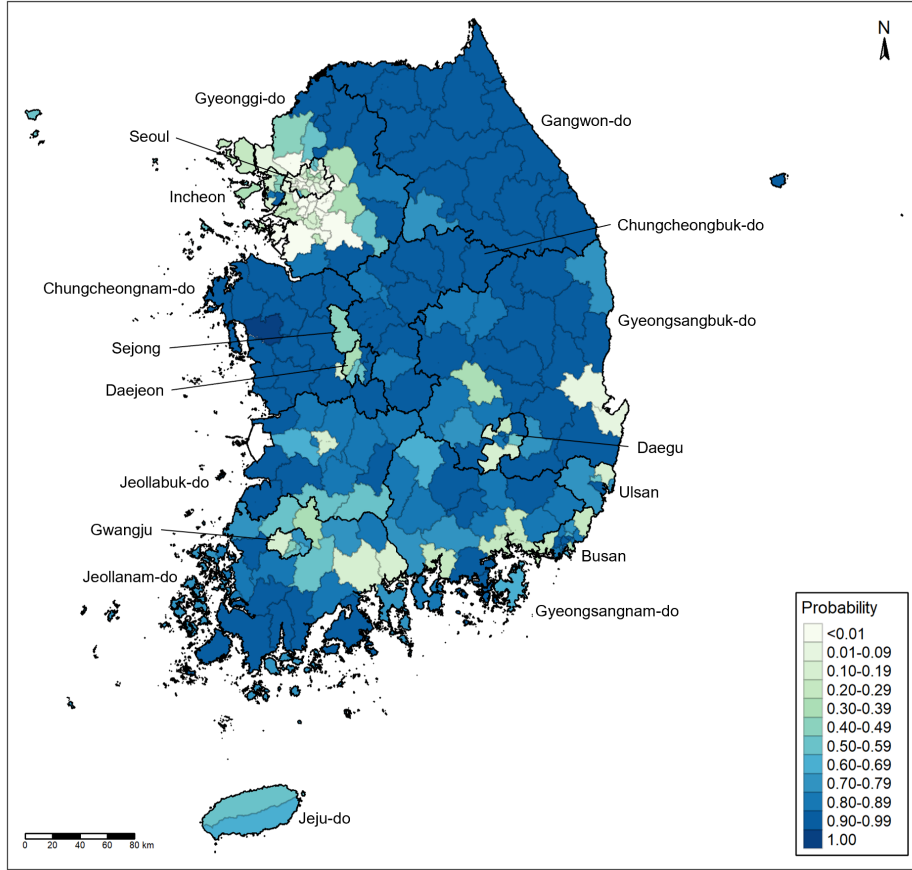


Figure 4: Map of Exceedance Probabilities



Figures 2, 3, and 4 collectively illustrate marked spatial disparities in suicide risk across South Korean municipalities. High-risk areas ($RR > 1.25$) are prominently clustered in rural regions, particularly in Chungcheongnam-do and Gangwon-do, with some municipalities exceeding an RR of 1.50. These areas are often characterised by limited access to healthcare, economic vulnerability, and higher proportions of older and male residents—demographics typically associated with elevated suicide risk.

In contrast, most low-risk areas ($RR < 1.00$) are concentrated in urban regions, particularly within the Seoul Metropolitan Area and major cities such as Busan and Daegu. These areas may benefit from stronger healthcare infrastructure, greater economic opportunity, and denser social networks. The clear clustering of high- and low-risk values indicates substantial spatial autocorrelation, underscoring a pronounced urban–rural divide in suicide risk.

The statistical significance of relative risk estimates further reinforces this pattern. Red municipalities ($95\% \text{ CrI} > 1.00$), indicating significantly elevated risk, are predominantly rural, while blue areas ($95\% \text{ CrI} < 1.00$), indicating significantly lower risk, appear almost exclusively in the capital region. Most municipalities, shown in white, do not exhibit statistically significant risk.

Exceedance probabilities ($P(RR > 1.00)$) provide an additional perspective. Municipalities with high probabilities (≥ 0.90) are widespread across non-capital regions, while the Seoul Metropolitan Area consistently shows low exceedance probabilities. This divide highlights substantial spatial inequities in suicide mortality and underscores the need for geographically targeted interventions beyond the capital region.

3.3 Discussion

These results carry important policy implications. Most notably, the unemployment rate showed the strongest association with suicide risk, with a one-unit increase linked to a 4% decrease in relative risk. Although counter-intuitive, this may reflect the buffering effect of South Korea’s unemployment benefits, which allow individuals to maintain economic stability while avoiding chronic workplace stressors. Robust safety nets may provide psychological protection, but they may also incentivise withdrawal from the highly stressful labour market, raising concerns about economic inactivity and long-term fiscal strain. These findings suggest that improving workplace mental health should be a priority, rather than blindly expanding welfare provisions.

Single-person household ratio and unmet medical needs rate show weak but positive associations with suicide risk, highlighting the relevance of social connection and access to healthcare. These findings align with existing literature and support continued investment in community formation, outreach to isolated individuals, and expansion of affordable mental health services.

Geographically, a sharp divide emerges between the Seoul Metropolitan Area and the rest of the country. Rural municipalities exhibit consistently higher relative risk and exceedance probabilities, while Seoul, Incheon, and Gyeonggi-do appear comparatively insulated. This suggests that national strategies risk neglecting the distinct vulnerabilities of peripheral regions, especially when policymakers themselves are based in the capital region. To address these disparities, policymakers should prioritise decentralised mental health services and pursue regional development strategies that foster economic opportunity and social cohesion outside the capital.

Several limitations should be acknowledged. Although preliminary diagnostics indicated modest overdispersion in suicide counts ($\theta \approx 1.28$), we opted for a Poisson model for its interpretability and established use in spatial epidemiology. The manual linking of disconnected subgraphs using centroid proximity ensured spatial contiguity but may not reflect true socioeconomic similarity, potentially biasing results in island geographies. The analysis is cross-sectional and based solely on 2022 data, limiting its ability to capture temporal dynamics. Finally, the observed relationship between unemployment and suicide risk may be confounded by unmeasured factors such as informal employment or labour market inactivity, which are not captured in standard unemployment statistics.

Despite these limitations, the study offers meaningful insight into the geography of suicide risk in South Korea. The findings underscore the urgent need for spatially targeted interventions and regionally responsive policy design.

References

- [1] Kim, E., & Kim, S. (2024). *Spatially clustered patterns of suicide mortality rates in South Korea: a geographically weighted regression analysis*. BMC Public Health, 24(1). doi:<https://doi.org/10.1186/s12889-024-19899-4>.
- [2] Lee, D. (2011). *A comparison of conditional autoregressive models used in Bayesian disease mapping*. Spatial and Spatio-temporal Epidemiology, 2(2), 79–89.
- [3] Rue, H., Martino, S., & Chopin, N. (2009). *Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (INLA)*. Journal of the Royal Statistical Society, Series B, 71(2), 319–392.
- [4] Statistics Korea. (2023). *Population and Health Indicators by Municipality*. Daejeon: KOSIS.
- [5] World Health Organization. (2023). *Suicide Worldwide in 2022: Global Health Estimates*. Geneva: WHO.
- [6] Zhu, L., Gorman, D. M., & Horel, S. (2006). *Hierarchical Bayesian spatial models for alcohol availability, drug "hot spots," and violent crime*. International Journal of Health Geographics, 5(1), 54.