Income and cancer mortality across Lombardy provinces

Francesco Danesi

February 13, 2026

1 Research Question

Cancer remains one of the leading causes of morbidity and mortality worldwide, posing a substantial burden on health systems and societies. Despite advances in prevention, diagnosis, and treatment, cancer outcomes continue to display marked inequalities across populations. A rich literature suggests that these disparities are closely linked to socioeconomic conditions, particularly income and income inequality, which can influence exposure to risk factors, access to healthcare, and timeliness of diagnosis and treatment (Eijkelboom et al. 2025; Ziehr et al. 2015; Tetzlaff et al. 2022; Santos Figueiredo and Adami 2019).

Large-scale analyses conducted in different national contexts show that populations with lower income levels tend to experience higher cancer incidence and increased mortality for several cancer types (Tetzlaff et al. 2022). Furthermore, these inequalities persist even in countries with universal healthcare systems, suggesting that socioeconomic determinants of cancer outcomes do not reside solely in access to treatment, but also influence a broader range of health determinants (Eijkelboom et al. 2025).

Beyond individual-level effects, income inequality at the regional level has been shown to play an important role in shaping cancer mortality patterns. For example, ecological studies conducted in Brazil have demonstrated that regions characterized by higher income inequality exhibit significantly higher age-standardized breast cancer mortality rates (Santos Figueiredo and Adami 2019). Evidence from the United States indicates that patients residing in lower-income areas are less likely to receive definitive or timely cancer treatment, resulting in worse survival outcomes, particularly for aggressive cancers such as high-risk prostate cancer (Ziehr et al. 2015). These findings suggest that unequal income distribution can exacerbate health disparities by limiting access to preventive healthcare services within specific territories. Income-related inequalities also affect cancer outcomes indirectly through disparities in treatment access and quality of care.

Additionally, these studies emphasize that cancer-related inequalities often display a significant spatial dimension (Goungounga et al. 2016). Such analyses frequently reveal substantial geographic heterogeneity in cancer incidence and mortality, reflecting differences in socioeconomic structure, environmental exposure, healthcare accessibility, and population composition. Spatial analytical approaches can therefore be useful to identify clusters, outliers, and spatial interaction processes that non-spatial methods may overlook.

Building on this background, the present study aims to investigate whether cancer mortality across Lombardy provinces exhibits spatial patterns and whether these patterns are associated with differences in income per capita. Although a considerable body of literature has examined cancer-related inequalities in Lombardy, most studies focus on specific cancer types (Giorgi Rossi et al. 2020; Stoppa et al. 2022), whereas this research adopts a broader perspective on overall cancer mortality. The choice of Lombardy was also driven by data availability, as relevant data—while not without limitations—were more accessible compared to many other regions. In addition, Lombardy represents a relatively wealthy and comparatively homogeneous region in terms of socioeconomic inequality, thus providing insight into how sensitive cancer outcomes may be to variations in material conditions.

The analysis combines measures of spatial autocorrelation with regression modeling. The objective is to contribute to the existing literature by assessing the extent to which socioeconomic disparities are reflected in the spatial distribution of cancer mortality at the provincial level. The research question is therefore formulated as follows:

> Do cancer mortality rates across Lombardy provinces exhibit spatial patterns? If so, can such patterns be associated with average income levels in those provinces?

2 Description of the Data

Cancer mortality data were obtained from the Italian National Institute of Statistics (ISTAT) ISTAT 2024a. The dataset provides information on mortality by cause of death, classified according to the European Short List of Causes of Death, and aggregated at the provincial level. The analysis focuses on mortality due to cancer, expressed as a mortality rate in order to normalize for differences in population size across provinces. The use of mortality data, rather than incidence data, was motivated by the fact that mortality represents a comprehensive indicator of disease burden, incorporating both incidence and survival outcomes. Moreover, mortality data were more readily accessible compared to incidence data.

Mortality rates were aggregated over a single year, 2019, selected as the most recent year not yet influenced by the COVID-19 pandemic, which may in some cases reduce the reliability and comparability of health statistics.

Income data were also retrieved from ISTAT ISTAT 2024b. The original income information was available at the municipal level; to ensure consistency with mortality data, municipal-level income data were aggregated to the provincial level. Total income values were then normalized by resident population to obtain income per capita for each province. This normalization allows for meaningful comparisons across provinces of different sizes. The choice of this indicator as a proxy for the average socioeconomic condition of the population was primarily driven by data availability, as more comprehensive measures of socioeconomic status were difficult to obtain. This represents a limitation of the present study.

Spatial data consist of a shapefile representing the administrative boundaries of Lombardy provinces, also obtained from ISTAT. The shapefile was used to construct spatial weights matrices based on contiguity relationships and to generate thematic maps illustrating the spatial distribution of cancer mortality and income per capita.

The final analytical dataset integrates cancer mortality rates, income per capita, population figures, and spatial geometries for the twelve Lombardy provinces.

3 Data Analysis

The analysis began with an Exploratory Spatial Data Analysis (ESDA) to examine the spatial distribution of cancer mortality rates and income per capita across Lombardy provinces. Choropleth maps were produced to provide an initial visual assessment of geographic heterogeneity and to identify potential spatial patterns. This exploratory phase aimed to guide the subsequent formal spatial analysis.

A spatial weights matrix was constructed based on Queen contiguity, whereby two provinces are considered neighbors if they share either a common boundary or a vertex. The matrix was row-standardized to ensure comparability across spatial units with different numbers of neighbors.

Global spatial autocorrelation was assessed using Moran's I, which measures the degree to which similar values cluster in space. Separate statistics were computed for cancer mortality rates and income per capita. Statistical significance was evaluated using permutation-based inference, allowing for a robust assessment of spatial dependence without relying on strong distributional assumptions. This step aimed to determine whether a systematic spatial structure was present at the regional scale.

To complement the global analysis, Local Indicators of Spatial Association (LISA) were computed using Local Moran's I to identify localized clusters and potential spatial outliers. Provinces were classified into standard cluster categories (high–high, low–low, high–low, and low–high) based on their values and those of neighboring provinces. The results were visualized through LISA cluster maps to highlight local spatial heterogeneity.

To examine the relationship between cancer mortality and socioeconomic conditions, Ordinary Least Squares (OLS) regression models were estimated with cancer mortality rate as the dependent variable and income per capita as the explanatory variable. Both linear and log-linear specifications of income were considered in order to account for potential non-linearities. Model diagnostics were examined to assess goodness of fit and the plausibility of the regression assumptions.

Finally, Moran's I was computed on the residuals of the OLS models to assess whether spatial dependence persisted after controlling for income. This diagnostic step allowed evaluation of whether the regression specification adequately captured the spatial structure of cancer mortality or whether additional spatial processes were likely present.

4 Conclusions

The choropleth maps of cancer mortality rates and income per capita show notable heterogeneity across Lombardy provinces. Cancer mortality varies considerably between provinces, although no clear contiguous areas of uniformly high or low values emerge. Similarly, income per capita displays marked differences, with higher values concentrated in major financial and industrial centers such as Milano and Monza, and lower values observed in more peripheral provinces. Visual inspection suggests the absence of large, homogeneous spatial clusters for either variable, indicating a potentially dispersed spatial structure. These preliminary observations motivated the use of formal spatial autocorrelation measures.

Global Moran's I statistics provide formal evidence of spatial dependence. For cancer mortality rates, Moran's I is equal to -0.381 and statistically significant ($p = 0.011$), indicating negative spatial autocorrelation. This result suggests spatial dispersion: provinces with relatively high mortality rates tend to be adjacent to provinces with lower rates, rather than forming contiguous clusters of similar values. This finding is consistent with the initial visual inspection.

In contrast, income per capita exhibits a negative but statistically non-significant Moran's I ($I = -0.316$, $p = 0.076$), indicating the absence of significant global spatial autocorrelation at the provincial level. While cancer mortality displays a structured spatial pattern, income per capita does not show a strong or consistent spatial organization across Lombardy provinces.

Local Moran's I statistics were computed to identify localized patterns of spatial association and potential spatial outliers. The LISA analysis revealed limited evidence of significant local clustering. Only the province of Cremona emerged as a statistically significant local outlier, while the remaining provinces did not exhibit meaningful local spatial associations. The absence of widespread significant local clusters aligns with the negative and dispersed pattern observed in the global Moran's I results. Rather than forming coherent high–high or low–low clusters, cancer mortality in Lombardy appears characterized by isolated deviations from neighboring values.

Ordinary Least Squares regression models were then estimated to assess the association between cancer mortality rates and income per capita. In the linear specification, income per capita displayed a negative coefficient; however, the relationship was not statistically significant ($p = 0.564$). The explanatory power of the model is limited, with an R^2 of 0.034, suggesting that income per capita alone explains only a small fraction of the variation in mortality rates across provinces. The log-linear specification yields similar results: the coefficient of log income per capita remains negative but statistically insignificant ($p = 0.780$), and model fit decreases further ($R^2 = 0.008$). These findings indicate that the absence of a significant association is robust across model specifications.

To assess whether spatial dependence persists after controlling for income, Moran's I was computed on the residuals of the OLS regression model. The residual Moran's I equals -0.392 and is statistically significant ($p = 0.012$), indicating that spatial structure remains present in cancer mortality beyond what is explained by income per capita. This result suggests that additional spatially structured factors contribute to the observed pattern.

Overall, the findings indicate that cancer mortality rates exhibit a statistically significant pattern of negative spatial autocorrelation, characterized by spatial dispersion rather than clustering. Income per capita does not display significant spatial autocorrelation at the provincial level, nor does it appear significantly associated with cancer mortality in regression analysis. Residual diagnostics further confirm the presence of unexplained spatial structure.

The results therefore suggest that socioeconomic conditions, as proxied by income per capita, are insufficient on their own to explain spatial differences in cancer mortality across Lombardy provinces. Other determinants, such as healthcare accessibility, environmental exposure, demographic composition, and spatial interaction processes, are likely to play a relevant role.

The answer to the research question appears to be negative. Cancer mortality rates across Lombardy provinces do exhibit spatial patterns, but these patterns cannot be meaningfully associated with differences in income per capita. This outcome may partly depend on the choice of study area and level of aggregation. Lombardy is relatively homogeneous in terms of income inequality, and provincial-level aggregation may obscure more localized socioeconomic disparities. The scale of analysis may therefore not be fully aligned with findings in the broader literature.

4.1 Limitations and Future Research

This analysis is subject to several limitations. First, the small number of spatial units (twelve provinces) limits statistical power and constrains the complexity of models that can be reliably estimated. Second, the use of aggregated province-level data may mask important within-province heterogeneity, raising the possibility of ecological fallacy.

Third, income per capita was the sole explanatory variable considered, whereas cancer mortality is influenced by multiple factors, including age structure, environmental conditions, lifestyle behaviors, and healthcare availability. The omission of such variables likely contributes to the limited explanatory power of the regression models.

Finally, although spatial autocorrelation was detected in cancer mortality, the small sample size discourages the use of more complex spatial econometric models. Future research could address these limitations by incorporating municipality-level data and additional economic, social, demographic, and environmental variables to better capture the spatial determinants of cancer mortality.

References

Eijkelboom, Anouk H. et al. (2025). "Income inequalities, tumor stage at diagnosis, and treatment in patients with non-small cell lung cancer: A population-based study in the Netherlands". In: *International Journal of Cancer*.

Giorgi Rossi, Paolo et al. (2020). "Geographic inequalities in breast cancer in Italy: Trend analysis of mortality and risk factors". In: *International Journal of Environmental Research and Public Health* 17.11, p. 4165.

Goungounga, Juste Aristide et al. (2016). "Impact of socioeconomic inequalities on geographic disparities in cancer incidence: comparison of methods for spatial disease mapping". In: *BMC Medical Research Methodology* 16.1, p. 136.

ISTAT (2024a). *Cause di morte per comune di residenza.* Accessed: 2026-02-13. URL: https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,Z0810HEA,1.0/HEA_DEATH/DCIS_CMORTE1_RES/IT1,39_494_DF_DCIS_CMORTE1_RES_1,1.0.

— (2024b). *Redditi IRPEF comunali.* Accessed: 2026-02-13. URL: https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,HOU,1.0/MEF_REDDITIIRPEF_COM/IT1,30_1008_DF_MEF_REDDITIIRPEF_COM_1,1.0.

Santos Figueiredo, Francisco Winter dos and Fernando Adami (2019). "Effects of the high-inequality of income on the breast cancer mortality in Brazil". In: *Scientific Reports* 9.1, p. 4173.

Stoppa, Giorgia et al. (2022). "Spatial analysis of shared risk factors between pleural and ovarian cancer mortality in Lombardy (Italy)". In: *International Journal of Environmental Research and Public Health* 19.6, p. 3467.

Tetzlaff, Fabian et al. (2022). "Time trends and income inequalities in cancer incidence and cancer-free life expectancy – A cancer site-specific analysis of German health insurance data". In: *Frontiers in Oncology* 12, p. 827028.

Ziehr, David R. et al. (2015). "Income inequality and treatment of African American men with high-risk prostate cancer". In: *Urologic Oncology: Seminars and Original Investigations* 33.1.