Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Regional Disparities in Cancer Incidence and Survival in Chile: A 20-Year Population-Based Analysis

Orietta Nicolis

orietta.nicolis@unab.cl

Andrés Bello National University

Diego Mellado

Andrés Bello National University

Denisse de los Angeles

Andrés Bello National University

Fabrizio Ruggeri

National Research Council

Carla Taramasco

Andrés Bello National University

Article

Keywords: Cancer Incidence, Survival Analysis, Spatial Analysis, Kaplan–Meier, Cox Model

Posted Date: July 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8434435/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Regional Disparities in Cancer Incidence and Survival in Chile: A 20-Year Population-Based Analysis

Orietta Nicolis[1,3,5*], Diego Mellado[1,2,3], Denisse de los Angeles[1,3], Fabrizio Ruggeri[4], Carla Taramasco[1,2,3]

[1*]Facultad de Ingeniería, Universidad Andrés Bello, Quillota 980, Viña del Mar, Chile.
[2]Instituto de Tecnología en Innovación en Salud y Bienestar (ITISB), Universidad Andrés Bello, Uno Oriente 1180, Viña del Mar, Chile.
[3]Centro para la Prevención y Control del Cancer (CECAN), Santiago, Chile.
[4]Consiglio Nazionale delle Ricerche, Istituto di Matematica Applicata e Tecnologie Informatiche, 20133, Milano, Italy.
[5]Department of Economics, Universitá degli Studi di Messina, Via dei Verdi, 75, 98122, Messina, Italy.

*Corresponding author(s). E-mail(s): orietta.nicolis@unab.cl;
Contributing authors: diego.mellado@unab.cl;
d.ramrezjara@uandresbello.edu ; fabrizio@mi.imati.cnr.it;
carla.taramasco@unab.cl;

Abstract

Purpose: Cancer is one of the leading causes of death in Chile, highlighting the need for regionally detailed and methodologically robust analyses to inform national public health policies. This study examines six major types of cancers: Breast, Lung, Colorectal, Stomach, Bladder, and Cervical. Across five regions with publicly available registry data over a 20-year period, we assessed the influence of geographic, demographic, and clinical factors on incidence and mortality.

Methods: We analyzed 26943 cancer cases from the Chilean Population Cancer Registry (1998-2019). After data cleaning and filtering, we conducted descriptive analysis, age-adjusted incidence estimation, Kaplan-Meier survival analysis, and

Cox proportional Hazards modeling to assess survival differences by cancer type, region, sex, and age.
Results: Marked geographic disparities were described. Northern regions, such as Antofagasta, showed a high incidence of lung and bladder cancers, whereas central-southern regions, such as Los Ríos, demonstrated elevated stomach cancer incidence. Survival outcomes varied widely across the different cancers: breast cancer exhibited the most favorable long-term survival, while lung and stomach cancers showed the poorest outcomes. The Cox model identified cancer type, age, sex, and region as significant predictors of mortality.
Conclusion: Chile faces substantial regional heterogeneity in cancer presence. Integrating spatial and survival analyses provides critical evidence to guide region-specific prevention, early detection, and treatment strategies.

Keywords: Cancer Incidence, Survival Analysis, Spatial Analysis, Kaplan–Meier, Cox Model

1 Introduction

Cancer is a leading cause of mortality in Chile, accounting for nearly one in four deaths nationwide. Over recent decades, it has emerged not only as a major public health priority but also as a significant socioeconomic challenge, particularly in regions exposed to elevated environmental and occupational risks. The rising incidence and uneven geographic distribution of cancer across Chile underscore the urgent need for reliable and systematic data to guide public health planning, allocate resources effectively, and reduce regional disparities in outcomes.

Despite its clinical and societal impact, Chile has historically lacked a unified, population-based cancer registry with national coverage. This absence of comprehensive data has limited the ability to monitor temporal trends, evaluate the effectiveness of screening programs, and design regionally tailored prevention strategies. In response to this gap, the Ministry of Health (*Ministerio de Salud*, MINSAL) established the *Plan Nacional de Cáncer*, which prioritizes the creation of an integrated cancer surveillance system. Accurate and consistent registration of cancer incidence, mortality, and survival is fundamental for evidence-based policymaking and for advancing equity in cancer care across the country.

This study contributes to these national efforts by analyzing six major cancer types—lung, bladder, breast, cervical, colorectal, and stomach—across five Chilean regions: Antofagasta, Arica y Parinacota, Biobío, Los Ríos, and Maule. These regions were selected because they are currently the only ones maintaining publicly available data from the *Registros Poblacionales de Cáncer* (RPC), and together they reflect Chile's geographic, demographic, and socioeconomic diversity. They also capture distinct environmental and lifestyle contexts, such as intensive mining activity in the arid north and unique dietary patterns in the southern agricultural zones. This heterogeneity provides a valuable foundation for exploratory and comparative analyses of cancer occurrence and outcomes.

In this study, we use publicly available data from the *Registros Poblacionales de Cáncer* (RPC, https://epi.minsal.cl/registros-poblacionales-de-cancer-bd/) provided by the Chilean Ministry of Health to conduct statistical analyses of cancer incidence, survival, and mortality. The dataset covers the period from January 1, 1998, to December 31, 2019, across the five aforementioned regions, each with varying registry coverage periods.

Our methodological framework combines Kaplan–Meier survival estimation with Cox proportional hazards modeling to identify both individual- and regional-level variations in cancer outcomes. By integrating spatial distribution with survival metrics, this study generates actionable insights that can inform cancer control policies and resource allocation. Consistent with the objectives of the *Centro de Prevención y Control del Cáncer* (CECAN), our findings aim to support the consolidation of a comprehensive national cancer registry, enhance the targeting of prevention and treatment strategies, and ultimately improve cancer outcomes and advance health equity across Chile.

1.1 Related works

In our literature review, we contextualized this research by focusing on survival, mortality, and geospatial patterns in cancer studies in Chile. These dimensions provide critical insight into the dynamics of cancer outcomes and their influencing factors. Below, we summarize key contributions from recent studies that inform our work.

Survival and mortality analyses in Chile often utilize Kaplan–Meier techniques and Cox models. For instance, [1] applied these methods to 622 breast cancer patients, showing that disease-free survival was significantly lower in younger women ($\leq$ 40 years: 72.3%) compared to older women ($\geq$ 70 years: 84.6%, $p = 0.007$). This finding highlights the importance of age-stratified analysis in understanding outcomes. Similarly, [2] reported a 45% reduction in cervical cancer mortality over 20 years, validating the utility of Kaplan–Meier and Cox methods for evaluating long-term trends. [3], in a study of 312 prostate cancer patients, showed that hormone therapy significantly improved survival rates ($p < 0.05$, Kaplan–Meier), underscoring the need to address disparities between urban and rural patient groups. Additionally, [4] analyzed 102,358 gastrointestinal cancer cases and reported 5-year survival rates ranging from 33% to 65% depending on tumor location. [5] examined 254 colorectal cancer patients, finding an overall mortality rate of 45% and significant sex-based differences in survival ($p < 0.05$, Kaplan–Meier). Collectively, these studies underscore the significance of demographic and geographic factors in shaping cancer outcomes.

In terms of mortality disparities, [6] identified a 72% mortality rate among 850 hematologic cancer patients from rural areas, illustrating the critical gap in treatment access across different regions. This observation reinforces the importance of addressing geographic inequities in cancer care. Similarly, [7] reported a five-year mortality rate of 64.7% among 948 gastric cancer patients, attributing this high mortality in part to geographic factors such as regional dietary practices. These findings underscore the importance of investigating geospatial variables and their impact on cancer outcomes in Chile. [8] provided a comprehensive evaluation of lung cancer epidemiology, mortality, risk factors, and clinical management in Chile. Notably, 92.4% of the 3,873 new

lung cancer cases in 2018 resulted in death, with Antofagasta identified as a high-risk region. Lung cancer mortality rates of 153 (men) and 50 (women) per 100,000 were linked to chronic arsenic exposure, reflecting long-term environmental risks in northern Chile. This geospatial finding aligns with our focus on regional risk patterns. In another study, [9] found that approximately 85% of the cancer burden occurred in lower-income communities, emphasizing the influence of socioeconomic factors on clinical outcomes. Together, these studies demonstrate the need to integrate geographic, demographic, and socioeconomic analyses to fully understand the factors influencing cancer incidence and survival in Chile. By doing so, our study aims to contribute to the development of more effective and equitable cancer prevention and treatment strategies.

2 Materials and Methods

The methodology of this study began with a comprehensive data cleaning process and exploratory data analysis (EDA). We first filtered the dataset to isolate records for the six selected cancer types and identified key variables for each cancer type. Descriptive statistics were generated, and the dataset was then spatially enriched by merging it with geographic information (using the `sf` library in `R`) to attach each case to its corresponding region's spatial geometry. This integration enabled the creation of tables and graphs to visualize the distribution of cancer cases across different regions.

Next, the distribution of cancer types was analyzed both regionally and comparatively. We performed Kaplan–Meier survival analysis, stratified by region, age group, and cancer type, to provide a detailed understanding of survival trends. We also conducted a mortality analysis using Cox proportional hazards modeling to evaluate risk factors associated with time-to-death across different patient groups.

2.1 Kaplan–Meier survival analysis

The Kaplan–Meier survival analysis is a non-parametric statistical method widely used to estimate the survival function from time-to-event data [10]. It is particularly valuable in medical research for evaluating the time until a specific event occurs (such as death, disease recurrence, or recovery). A key strength of this method is its ability to handle censored data, which arises when the event of interest has not occurred for specific individuals by the end of the study period or if they are lost to follow-up.

The Kaplan–Meier estimator estimates survival probabilities at specific time points and constructs a survival curve that represents the proportion of individuals surviving beyond each time interval. The survival curve is a step function, with each step corresponding to a decrease in survival probability resulting from an observed event at a specific time. The method adjusts for censored data by excluding individuals censored at a given time from the risk set in subsequent intervals, ensuring unbiased survival probability estimates.

Formally, the survival probability at a given time t, $S(t)$, is defined in equation 1:

$$S(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right) \tag{1}$$

Where t_i represents the time of the i-th event, d_i is the number of events (deaths) at time t_i, and n_i is the number of individuals at risk just before t_i. This formulation provides a cumulative estimate of survival probability, accounting for changes in the number of individuals at risk over time.

Kaplan–Meier analysis is often complemented by statistical tests such as the log-rank test to compare survival curves between groups (e.g., patients with different treatments, age ranges, or regional backgrounds). In cancer research, Kaplan–Meier curves play a pivotal role in understanding survival trends. For example, [1] used Kaplan–Meier methods to evaluate disease-free survival in breast cancer patients, revealing significant differences between age groups. The ability to visualize survival probabilities over time and identify disparities between groups makes Kaplan–Meier analysis a cornerstone of survival analysis in epidemiology.

2.2 The Cox Proportional Hazards Model

The Cox proportional hazards model, introduced by Cox in 1972 [11], is a semi-parametric regression model that has become fundamental in survival analysis. Its primary purpose is to evaluate the relationship between survival time and one or more predictor variables (covariates), such as patient demographics or clinical characteristics. Unlike fully parametric models, the Cox model does not assume a specific distribution for survival times, making it both flexible and robust for a wide range of data.

Mathematically, the Cox model expresses the hazard function $h(t|X)$, the instantaneous risk of the event at time t given covariates X, as shown in equation 2:

$$h(t|X) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p) \tag{2}$$

Where $h_0(t)$ is the baseline hazard function (the hazard when all covariates are zero) and $\beta_1, \beta_2, \ldots, \beta_p$ are regression coefficients for the covariates $X_1, \ldots, X_p$. One fundamental assumption of the Cox model is proportional hazards: the hazard ratios between any two individuals remain constant over time and depend only on their covariate values. Formally, for two individuals with covariates $X^{(1)}$ and $X^{(2)}$, the hazard ratio is estimated using equation 3 which is independent of time t.

$$\frac{h(t|X^{(1)})}{h(t|X^{(2)})} = \exp\{\beta \cdot (X^{(1)} - X^{(2)})\} \tag{3}$$

The Cox model is considered semi-parametric because it does not require specifying a functional form for $h_0(t)$. The model's coefficients β are typically estimated via partial likelihood, which sidesteps the need to estimate the baseline hazard directly. For example, in a model where age (X_1) and sex (X_2) are predictors, the hazard might be estimated as in equation 4:

$$h(t|\text{age}, \text{sex}) = h_0(t)\ \exp(\beta_1 \cdot \text{age} + \beta_2 \cdot \text{sex}) \tag{4}$$

In this case, $\exp(\beta_1)$ represents the hazard ratio for each one-year increase in age, and $\exp(\beta_2)$ represents the hazard ratio comparing males to females.

The Cox model is widely used in clinical trials, epidemiological studies, and public health research to identify significant risk factors, compare survival between groups, and quantify the impact of predictors on survival time. Its combination of flexibility (no need to assume a survival distribution) and interpretability (through hazard ratios) makes it an invaluable tool for analyzing time-to-event data in our study alongside the Kaplan–Meier analysis.

3 Methodology

3.1 Dataset description, preprocessing, and filtering

The study uses a dataset from the the *Population-Based Cancer Registries* (Registros Poblacionales de Cáncer, RPC), comprising 80,890 cancer case records and 19 variables. This dataset is publicly available for download within Chile via the Chilean Ministry of Health's epidemiological surveillance website and covers the period from January 1, 1998 to December 31, 2019. It includes key variables such as tumor morphology and behavior, patient age and sex, dates of diagnosis and death, and other clinical and demographic details. The data encompass the five regions of this study (Antofagasta, Arica y Parinacota, Biobío, Los Ríos, and Maule) and focus on six cancer types: lung, bladder, breast, cervical, colorectal (labeled as "colon" in the data records), and stomach. Registries contain descriptions using the Cancer Incidence in Five Continents, Volume XI (CI5-XI) [12], a key reference in global cancer epidemiology. CI5-XI defines specific 5-year age groupings (quinquennial age groups) for reporting cancer incidence, allowing consistent comparisons across populations and time periods. Additionally, two international coding systems are used for cancer data: the ICD-O3 (International Classification of Diseases for Oncology, 3rd Edition) [13] and the ICD-10 (10th Revision) [14]. The ICD-O3 is specialized for oncology, detailing tumor topography (site of origin) and morphology (cell type and behavior), and is widely used in cancer registries globally. The ICD-10 is a general disease classification system that assigns alphanumeric codes to all diagnoses, including cancer (e.g., C50 denotes a malignant breast neoplasm, C34 denotes lung cancer, etc.).

During data preprocessing, several issues were addressed to ensure data quality. First, a small number of records with problematic or invalid entries were removed. For example, one record had an outlier value for the *Behavior* variable, and three records contained municipality codes (*regMun*) that did not correspond to any recognized commune; these issues led to the removal of 2,694 entries, yielding a cleaned dataset of 78,195 records. We interpreted missing values in the *Death_cause* and *Death_date_up2020* fields as indicating that the patient was still alive at the last follow-up (i.e., no recorded death by the end of 2020). Two missing birth dates for patients older than 99 years were retained, as those patients' ages could be inferred to exceed our upper age cutoff. We also identified 2,229 duplicate *Id_person* entries (patients appearing more than once). Among these, 30 had inconsistent birth dates across entries and were removed from the database to avoid contradictions.

After cleaning, we filtered the dataset to focus on the six specific cancer types of interest, identified by their ICD-10 codes: C34 (lung cancer), C67 (bladder cancer), C50 (breast cancer), C53 (cervical cancer), C18 (colon/colorectal cancer), and C16

(stomach cancer). Notably, for cervical cancer, we specifically included ICD-10 C53 and deliberately excluded C54 (uterine cancer) to maintain a precise focus on cervical carcinoma. For the cancers that predominantly affect one sex (breast and cervical), we further filtered the records by sex; for example, the breast cancer dataset was restricted to female patients since male breast cancer cases were negligible (¡1% of breast cancer cases).

After all filtering steps, the refined dataset comprises 27,495 observations across the six targeted cancer types. Each variable was appropriately formatted for analysis (categorical variables converted to factors, dates converted to date types, etc.). Details of the final processed dataset, including data types and example entries, are shown in Table 1.

Table 1 Description of the processed dataset (selected variables and example values for cases in the Antofagasta region).

Variable Name	**Type and Example Values**
Id_person	(Integer) Identifier code per patient
Sex	(Categorical) Male or Female
Age	(Integer) Recorded age of patients ($0-105$ years, mean 64.20 ± 14.41)
Date of Birth	(Date) 1934-07-17, 1922-09-03, 1954-09-17, 1950-06-14, ...
Age Group (Quinquennial)	(Categorical) Age groups per 5 years: 0-4, 5-9, ..., 75-79, 80+
Age Group 1	(Categorical) Age groups: 0-14, 15-29, 30-49, 50-69, 70+
Diagnosis Date	(Date) between Jan. 1st, 1998 and Dec. 31, 2019
Year of Diagnosis	(Integer) Years between 1998-2019
ICD-10 Code	(Categorical) 412 listed ICD-10 Codes. Mode: C61: 7881 registries
Cancer Name	(Categorical) Code and Description of ICD-10 group names
Diagnosis Basis	(Categorical) Procedure of diagnosis: Clinical, Research, Citology, Histological, Death Certificate, etc.
Tumor Behavior	(Categorical) Benign, Malign, In-situ, Uncertain
Morphology	(Categorical) 460 ICD-O-3 Morphology codes listed. Mode: 8000: 15931 registries
Region Code (Commune)	(Integer) Code representing each commune.
RPC	(Categorical) Region of Population Registry of Cancer
Vital Status	(Categorical) Deceased or Alive
Date of Death (2022)	(Date) Recorded date of death. Between Apr. 1st, 1998 - Dec. 31th 2022
Cause of Death	(Categorical) Recorded cause of death: Cancer, other causes, Unknown.

3.2 Exploratory Data Analysis (EDA)

Exploratory data analysis (EDA) is a crucial first step in understanding the structure, patterns, and relationships within the dataset. In this study, EDA was used to examine the distributions of key variables and to investigate relationships among demographic factors, geographic location, and cancer characteristics. The final dataset analyzed in the EDA comprises 26943 cancer cases across six cancer types in the five Chilean regions under study. We computed summary statistics for variables such as

age, sex, cancer type, survival time, and region. A series of tables and visualizations was generated to capture the data's variability and patterns.

Table 2 shows the distribution of cases by cancer type and sex. Figure 1 provides a corresponding visualization of these sex-based differences in case counts for each cancer type.

Table 2 Number of cancer cases by sex for each cancer type.

Cancer Type	Male	Female	Total
Stomach	4869	2350	7219
Colorectal (Colon)	1720	2045	3765
Lung	3128	1708	4836
Breast	48	6712	6760
Cervical	0	2500	2500
Bladder	1224	639	1863
Total	10989	15954	26943

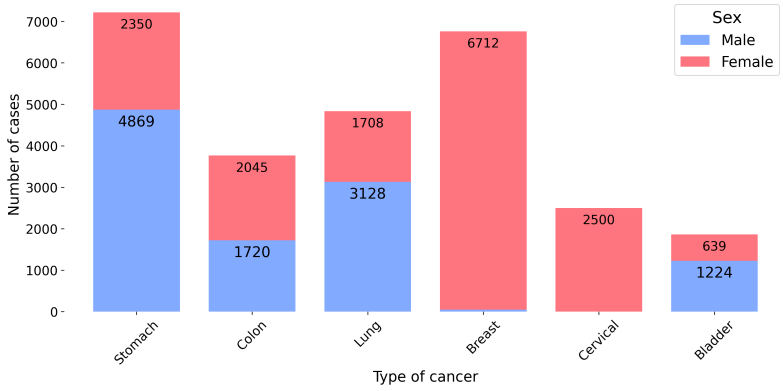


Fig. 1 Number of cases for each cancer type, by sex (Female vs. Male).

As shown in table 2 and figure 1, stomach cancer exhibits a substantially higher number of cases in men compared to women, a pattern similarly observed in lung and bladder cancers. This trend indicates a sex-specific difference in incidence that may be influenced by both biological factors and gender-associated environmental exposures (e.g., higher smoking rates or occupational hazards among men). On the other hand, colorectal cancer presents a more balanced distribution between women and men, suggesting a less pronounced sex bias in its incidence, potentially due to similar exposure to risk factors across sexes. Breast cancer cases are almost exclusively in women (male breast cancer accounted for $< 1\%$ of cases), which highlights the effectiveness of breast cancer screening programs in detecting a large number of cases (including early-stage tumors) among women. Similarly, cervical cancer, by definition,

affects only women; its considerable case count underlines the importance of regular Pap smear and HPV screening programs for early detection and intervention, which significantly improve treatment outcomes.

Table 3 Frequency of cancer cases by age group for each cancer type.

Cancer type	**0–14**	**15–29**	**30–49**	**50–69**	**≥70**
Stomach	0	29	636	3008	3546
Colorectal	2	51	400	1515	1797
Lung	2	7	256	2304	2267
Breast	0	55	1798	3230	1677
Cervical	1	116	1107	824	452
Bladder	1	4	142	770	946
All	6	262	4339	11651	10685

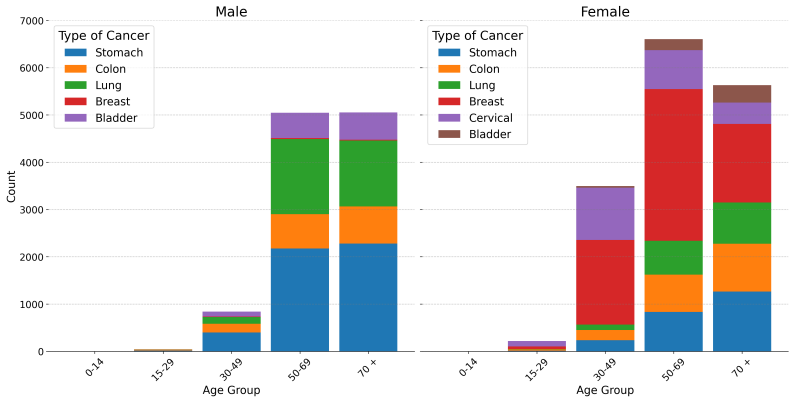


Fig. 2 Age distribution of cancer cases across the six cancer types, stratified by sex.

Table 3 summarizes the distribution of cancer cases by age group and cancer type, and Figure 2 depicts these patterns by sex. Overall, the results indicate distinct age-related trends across the cancer types analyzed. Stomach cancer is the most concentrated in older individuals, with the majority of cases occurring in the ≥ 70 age group. Breast cancer cases peak in the $50-69$ age group, which coincides with the age range of menopause (a known risk factor for breast cancer [15]); however, a substantial proportion of breast cancer cases also occur in the 30–49 group, underscoring that middle-aged women are significantly affected as well. Cervical cancer is predominantly diagnosed in women aged $30-49$, with a notable number of cases already appearing in the $15-29$ group, emphasizing the importance of early screening and HPV vaccination in younger women. Lung and bladder cancers show a marked increase in incidence among individuals aged $50-69$ and ≥ 70 (especially among men in these age groups), reflecting the cumulative effect of long-term exposures (such as smoking

or industrial carcinogens) that manifest later in life. Colorectal cancer appears to be more evenly distributed across middle and older ages, with significant case numbers in both the $50 - 69$ and ≥ 70 groups.

Figure 2 complements table 3 by illustrating these age distributions with an additional breakdown by sex. The plotted data highlight several gender-specific trends: for lung, stomach, and bladder cancers, men have a higher incidence at older ages compared to women, whereas colorectal cancer shows a relatively balanced pattern between sexes across age groups. These observations suggest that certain risk factors (like tobacco use or occupational exposures) that accumulate with age are more prevalent in men for some cancers. In contrast, other cancers are influenced by risk factors common to both men and women.

In addition to counts and frequencies, we computed summary statistics of patient age by cancer type, which are presented in table 4. This descriptive analysis includes the total number of cases, mean and median age, age standard deviation (SD), and the minimum and maximum ages observed for each cancer type.

Table 4 Descriptive statistics of patient age by cancer type.

Cancer Type	Count	Mean Age	Median Age	SD (Age)	Min Age	Max Age
Cervical	2,558	52.60	50.0	16.10	12	99
Colorectal	3,867	66.95	69.0	14.27	9	102
Stomach	7,442	67.85	69.0	12.95	17	101
Breast	6,760	59.23	58.0	14.02	19	105
Lung	4,900	68.14	69.0	11.28	0	98
Bladder	1,889	68.29	70.0	12.64	2	97

Table 4 indicates, for example, that cervical cancer patients have a mean age of approximately 52.6 years (median 50 years), whereas colorectal cancer patients are considerably older on average (mean $\sim$ 67 years, median 69). Stomach cancer cases (mean age 67.85) and lung cancer cases (mean 68.14) both reflect an older patient population. The largest age variability (SD 16.10) is seen in cervical cancer, which aligns with the fact that cervical cancer cases span a wide age range from the early 20s up to the 90s. In contrast, the ages of lung cancer patients are somewhat more tightly clustered (SD = 11.28). Notably, the minimum age observed for lung cancer was 0 (which likely indicates a very young patient or a potential data entry error), and the oldest patient in our data was 105 (with breast cancer).

Overall, the EDA highlights significant differences in how various cancers are distributed across sex and age groups, underlining the need to tailor public health interventions (like screening and education programs) to the demographic profiles most at risk for each cancer type.

4 Results

4.1 Temporal analysis

To understand the evolution of cancer incidence over time, we performed a temporal analysis examining trends in the annual number of diagnosed cases for each of the six major cancer types across the study period (1998–2019). This analysis helps detect temporal variations in incidence rates and assess the potential impact of public health initiatives (such as the introduction of screening programs or changes in environmental exposures) on cancer trends in Chile.

Figure 3 displays the number of new cases per year for each cancer type. It allows us to compare the temporal behavior of different cancers over the two-decade period.

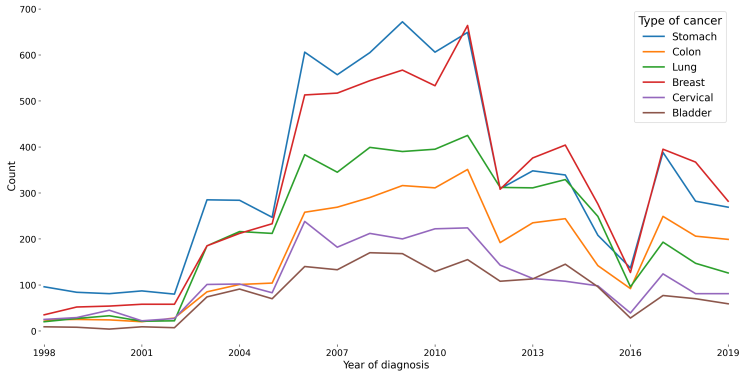


Fig. 3 Number of new cancer cases diagnosed each year (1998–2019), by cancer type.

The trends presented in Figure 3 show distinct temporal patterns across cancer types. Breast and cervical cancers exhibit a generally increasing trend over time, which may be partially attributed to improved screening (e.g., mammography and Pap smear programs implemented during the 2000s). In contrast, stomach cancer incidence appears relatively high throughout the late 1990s and early 2000s, with some fluctuations. Still, it does not show a strong upward trajectory, possibly reflecting long-standing risk factors (such as Helicobacter pylori infection rates) that have remained prevalent. Lung cancer shows an upward trend in the early 2000s, followed by variability, which may correlate with historical tobacco use rates and the delayed effects of anti-smoking policies. Colorectal cancer incidence has risen steadily, especially in the 2010s, potentially due to lifestyle changes and the aging population, as well as improved diagnostic capabilities. Bladder cancer has a lower overall incidence and a relatively flat trend with minor increases.

To further explore temporal changes, we highlight key health policy milestones in Chile, such as the introduction of the Explicit Health Guarantees program (GES) in 2005 [16] and the launch of the National Cancer Screening programs around the same time. Indeed, around 2005–2006, we observe an uptick or inflection in the incidence of breast and cervical cancers (coinciding with expanded screening), consistent with

the idea that increased screening initially raises case detection rates [17]. Similarly, improvements in diagnostic capabilities (e.g., broader availability of endoscopy and colonoscopy) might explain the noticeable rise in reported colorectal cancer cases in the late 2000s and 2010s.

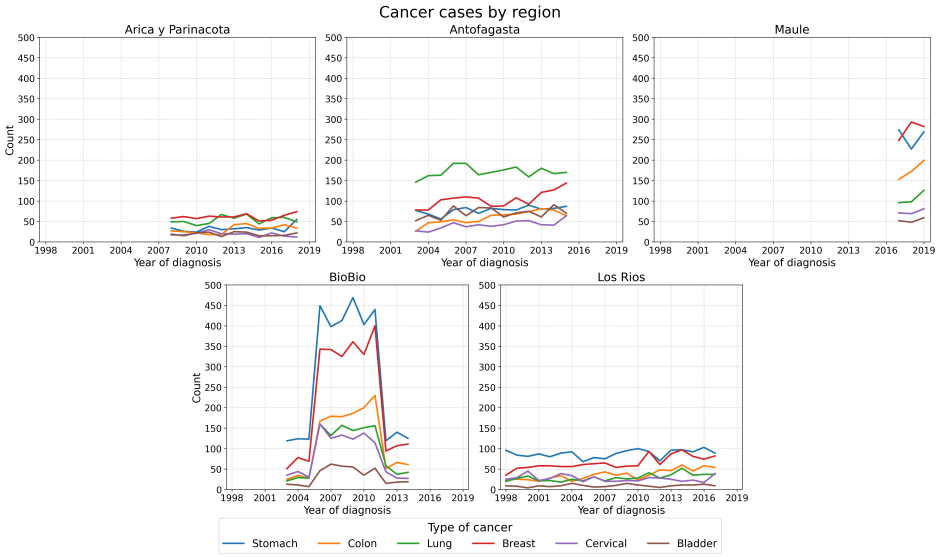


Fig. 4 Temporal trends of annual cancer cases (1998–2019) for each region, by cancer type. Each panel corresponds to one region (Antofagasta, Arica y Parinacota, Biobío, Los Ríos, Maule).

We conducted a temporal analysis disaggregated by region (Figure 4) to examine how cancer case trends vary across the five regions for each cancer type. The results reveal evident regional differences in temporal patterns. For instance, Antofagasta shows pronounced increases in lung and bladder cancer cases over time, aligning with the region's mining activities and known exposure to arsenic (as evidenced by historical studies on arsenic-related cancer risk in northern Chile[18]). Biobío, one of the most populous regions, shows rising trends in stomach and colorectal cancers during the 2010s, which may be linked to dietary habits and lifestyle factors, as well as improved reporting. Arica y Parinacota, a smaller region, exhibits more variable trends due to its smaller number of cases. Still, a notable spike in breast cancer detections in the mid-2010s is noteworthy, possibly reflecting targeted screening efforts. Los Ríos shows relatively steady or modestly increasing trends, though with high stomach cancer incidence persisting, reflecting that region's historically high gastric cancer burden [19]. In Maule, where registry data were available only for the last few years of the study period (2017–2019), the short time series precludes precise trend detection. As such, we have omitted Maule's curves from long-term trend interpretation due to the statistical insignificance of such a short interval. Overall, the temporal analysis indicates that while some improvements and increases in detection have occurred (especially for breast and cervical cancers with screening), other cancer types have not

seen declines, and some (like colorectal cancer) are on the rise. This underscores the importance of continuous monitoring and evaluating the effectiveness of public health interventions over time.

4.2 Spatial analysis

We next examined the spatial distribution of cancer cases to identify regional disparities in cancer frequency. Table 5 provides the total number of cases observed in each of the five regions for each cancer type. Figure 5 illustrates the geographic context of Chile and highlights the regions included in this study.

Table 5 Total number of cancer cases (1998–2019) by region and cancer type.

Cancer Type	Antofagasta	Arica y Parinacota	Biobío	Los Ríos	Maule
Cervical	541	203	1,000	593	221
Colorectal	769	340	1,406	828	524
Stomach	1,012	362	3,322	1,976	770
Breast	1,340	670	2,591	1,425	813
Lung	2,224	589	1,115	652	320
Bladder	918	210	390	212	159

Table 5 and the map in figure 5 help visualize the geographical distribution of cancer cases. We can identify regions where certain cancers have notably high frequencies. For example, Biobío (the most populous region in this study) has the highest counts for most cancer types, as expected given its larger population. However, raw case counts do not account for differences in population size across regions. To ensure comparability, we calculated cancer incidence rates per 100,000 inhabitants for each region using population data from the 2017 Chilean census.

Table 6 presents the age-adjusted incidence rates (per 100,000 population) for each cancer type by region (using 2017 population as a reference) [20]. These rates provide a clearer picture of the cancer burden relative to each region's population size.

Table 6 Age-adjusted cancer incidence rates (per 100,000 inhabitants) by region (based on 2017 census [20]).

Cancer Type	Antofagasta	Arica y Parinacota	Los Ríos	Biobío	Maule
Lung	332.09	266.08	185.72	76.57	29.00
Colorectal	135.18	153.59	230.15	116.10	48.85
Breast	419.08	598.35	737.52	418.45	145.45
Stomach	160.54	163.53	644.62	346.35	81.99
Cervical	274.43	181.29	361.17	170.42	42.85
Bladder	133.85	94.87	56.02	28.45	15.89

As shown in Table 6, once population size is accounted for, some patterns shift compared to raw case counts. For instance, Los Ríos, despite having fewer total stomach

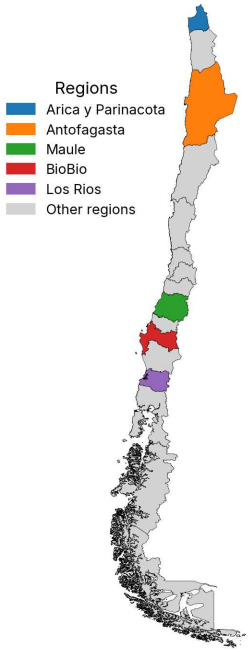


Fig. 5 Map of Chile indicating the five regions analyzed in this study (colored): Antofagasta, Arica y Parinacota, Biobío, Los Ríos, and Maule.

cancer cases than Biobío, shows the highest stomach cancer incidence rate (approximately 644.6 per 100,000), reflecting its smaller population and historically high gastric cancer risk in that region [19]. Similarly, Antofagasta, which had a moderate number of bladder cancer cases in absolute terms, exhibits the highest bladder cancer incidence rate among the five regions (133.9 per 100,000), underscoring the significant impact of environmental risk factors, such as arsenic in drinking water, which have historically affected northern Chile [18]. Lung cancer incidence is notably high in Antofagasta and Arica y Parinacota, consistent with the intense mining activity and related pollution in the north.

Breast cancer incidence is relatively high across all regions, with Los Ríos having the highest rate (over 737 per 100,000), suggesting that breast cancer is a widespread concern influenced by risk factors (genetic, reproductive, lifestyle) that are present nationwide [15]. Cervical cancer incidence is particularly high in Los Ríos and Antofagasta compared to Biobío and Maule, indicating regional differences that might relate to historical screening program reach or socio-cultural factors affecting HPV exposure and vaccination rates.

The spatial patterns observed suggest several important points. First, there is a clear north-south gradient for specific cancers: northern regions (Antofagasta, Arica) have higher incidence of cancers known to be linked with mining-related environmental

exposures (lung, bladder), whereas southern/central regions like Los Ríos and Biobío show higher incidence for cancers like stomach and colorectal, which might be influenced by local diet (such as high salt intake or red meat consumption) and possibly genetic predispositions in those populations. These regional disparities underscore a potential environmental health link; for example, the high lung cancer incidence in the north may reflect the legacy of arsenic contamination and air pollution, while the high stomach cancer incidence in the south-central zones may relate to localized dietary habits and the prevalence of H. pylori infection.

In summary, the spatial distribution of cancer incidence in Chile is not uniform: certain regions face a disproportionately high burden of specific cancers. This finding reinforces the need for geographically tailored public health strategies. Regions like Biobío and Los Ríos might benefit from targeted programs to prevent stomach cancer (such as dietary modifications and H. pylori eradication campaigns). At the same time, Antofagasta and Arica might require ongoing monitoring of environmental contaminants and aggressive smoking cessation programs to combat lung and bladder cancer.

4.3 Survival Analysis

We used the Kaplan–Meier survival analysis method to estimate and compare survival probabilities across patient groups. To ensure accurate survival estimates, the dataset for this analysis was further refined to include only cancer-related mortality cases: we filtered out any patients who were still alive or died from causes unrelated to cancer (these patients were treated as censored in the survival analysis). We then constructed survival curves for each cancer type and for subgroups by region. Before running the survival analysis, we prepared a subset of the data containing only the key variables necessary (patient ID, sex, age, cancer type, date of diagnosis, vital status, date of death, and cause of death). An excerpt of this survival analysis dataset (for illustrative purposes, using cases from Antofagasta) is shown in table 7.

Table 7 Subset of processed survival analysis data (example from Antofagasta region).

Variable Name	Type and Example Values
Id_persona	(Integer) 1102644, 1107469, ...
Sex	(Categorical) Male, Female
Age	(Integer) 77, 46, 74, 71, 58, ...
Type	(Categorical) Stomach, Breast, Colorectal ...
Diag_Date	(Date) 2006-07-04, 2008-05-06, 2006-01-30, 2003-10-20, ...
Vital Status (VM)	(Categorical) Alive, Deceased
Death_Date_2022	(Date) 2006-07-04, 2021-08-28, 2006-01-30, 2003-10-20, ...
Cause_of_Death	(Categorical) Cancer, other causes, unknown ...

The Kaplan–Meier survival curves for each cancer type (pooled across all regions) are shown in figure 6. In the same figure, the lower panel provides the number of patients at risk at various time intervals for each cancer type.

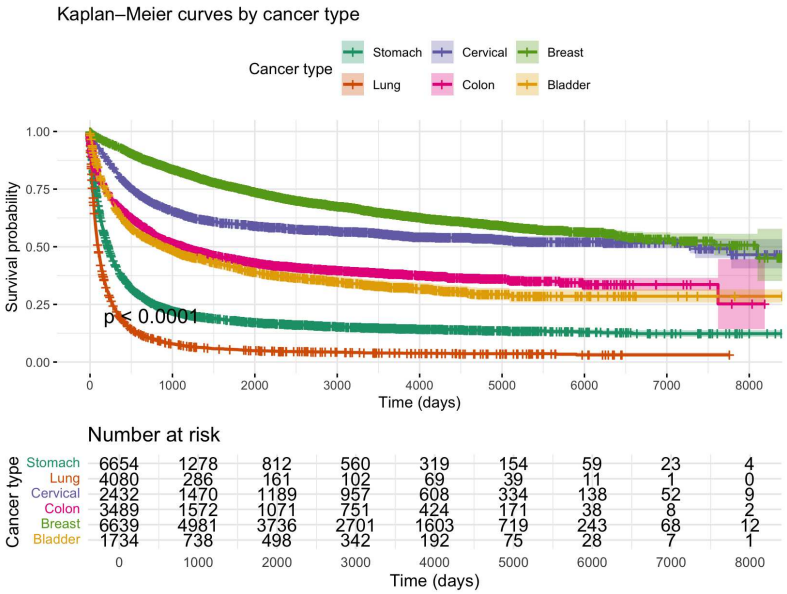


Fig. 6 Kaplan–Meier survival curves (top panel) and corresponding number-at-risk tables (bottom panel) for each cancer type.

The Kaplan–Meier curves in figure 6 illustrate stark differences in survival probabilities over time across cancer types. Among the cancers analyzed, breast cancer shows the highest long-term survival probability, indicating that patients diagnosed with breast cancer have a comparatively better prognosis, likely reflecting effective treatments and early detection efforts. In contrast, lung cancer shows the lowest survival probabilities, with a steep decline in the survival curve observed within the first couple of years after diagnosis. This highlights the aggressive nature of lung cancer and the high mortality risk shortly after diagnosis, which is consistent with global observations for lung malignancies.

Stomach cancer ranks next after lung cancer in terms of poor survival outlook, also demonstrating a steep decline in the early years. While its survival curve is slightly above that of lung cancer, it remains significantly lower than those of breast or cervical cancer. By approximately 1,000 days (about 2.7 years) post-diagnosis, less than 5% of lung cancer patients remain alive, whereas $10 - 30\%$ of patients with other cancers (such as breast or bladder cancer) are still surviving. At 2,000 days (~ 5.5 years), the survival probability for breast and bladder cancer patients is relatively high (over 36% alive for both), whereas lung and stomach cancer survival drops below 2%. By 4,000 days (~ 11 years), breast cancer still maintains a considerable proportion of survivors (though gradually declining), whereas lung cancer has virtually no survivors ($< 1\%$) and stomach cancer fewer than 2%. At 6,000 days (~ 16.5 years), survival for almost all cancer types has decreased significantly, but breast cancer continues to have the greatest residual survival.

These findings demonstrate a wide variability in disease aggressiveness and the effectiveness of interventions across different types of cancer. Breast and bladder cancers, for example, benefit from either early detection or relatively effective treatments, resulting in better survival outcomes. Lung and stomach cancers, on the other hand, are known for being detected at later stages and/or having limited treatment options in advanced stages, resulting in poor long-term survival.

We also stratified Kaplan–Meier survival curves by region for each cancer type, as shown in figure 7. Each panel in this figure corresponds to a specific region, and within each panel, the survival curves for different cancer types in that region are plotted.

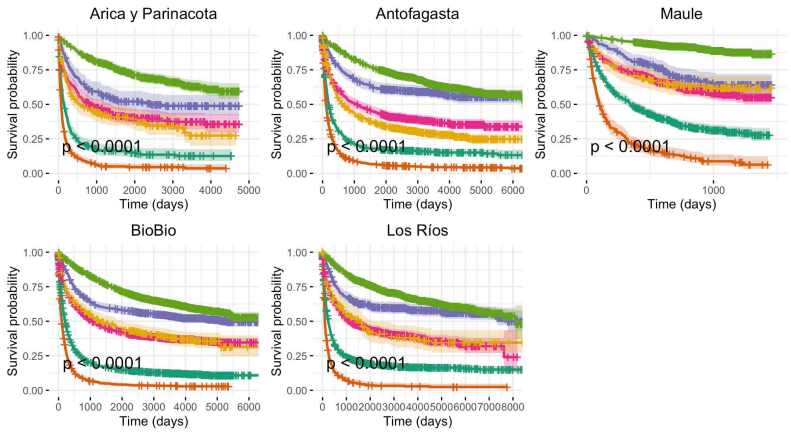


Fig. 7 Kaplan–Meier survival curves for each cancer type, stratified by region: Arica y Parinacota, Antofagasta, Biobío, Los Ríos, and Maule. The colors indicate the different types of cancer: green (breast), violet (cervical), magenta (colon and rectal), yellow (bladder), dark green (stomach), and brown (lung).

Each plot of Figure 7 reflects the overall survival trend: breast cancer shows the longest survival among the top cancers, followed by cervical cancer, whereas lung cancer exhibits the shortest survival. Across most regions, colon, rectal, and bladder cancer display very similar survival; the exception is Antofagasta, where colorectal cancer shows higher survival. In Maule, the curve spans a much shorter time frame because only two years of data are available; therefore, the results should be interpreted with caution.

In summary, figure 7 provides a compelling synthesis of how geographic context intersects with cancer type to shape survival outcomes. It reinforces the need for regionally tailored interventions: for example, improving cancer care infrastructure that shows consistently poorer survival, or investigating the factors that allow some regions to outperform others in specific cancers.

4.4 Survival analysis: unified cause-specific Cox model

We modelled time to *cancer-related death* using a unified cause-specific Cox proportional hazards model fitted to the entire cohort. Let T_i denote the follow-up time for individual i, measured from the date of diagnosis to the end of observation, and let δ_i be the event indicator. An event was defined as death due to cancer ($\delta_i = 1$). Individuals who were alive at the end of follow-up were administratively censored, and individuals who died from non-cancer causes were censored at their date of death ($\delta_i = 0$). Covariates included age at diagnosis (A_i, continuous), sex (S_i, coded as 1 for male and 0 for female), region of diagnosis (R_i, categorical with Antofagasta as the reference category), and cancer type (C_i, categorical with stomach cancer as the reference category). The model was specified as:

$$h(t \mid A_i, S_i, R_i, C_i) = h_0(t) \exp\left(\beta_{\text{age}} A_i + \beta_{\text{sex}} S_i + \sum_{r \neq r_0} \beta_r \, \mathbb{I}(R_i = r) + \sum_{c \neq c_0} \gamma_c \, \mathbb{I}(C_i = c) \right), \tag{5}$$

where $h(t \mid \cdot)$ is the cause-specific hazard of cancer death at time t, $h_0(t)$ is the baseline hazard, $\mathbb{I}(\cdot)$ denotes an indicator function, r_0 is the reference region (Antofagasta), and c_0 is the reference cancer type (stomach cancer). Results are reported as hazard ratios (HRs) with 95% confidence intervals (CIs).

Table 8 summarizes the estimates from the unified cause-specific Cox model for cancer-related mortality.

Table 8 Unified cause-specific Cox model estimates for cancer-related mortality.

Covariate	**HR**	**95% CI**	***p*-value**
Cancer type (ref: Stomach)			
Lung	1.80	(1.71–1.89)	$< 2 \times 10^{-16}$
Cervical	0.41	(0.37–0.44)	$< 2 \times 10^{-16}$
Colon	0.49	(0.46–0.52)	$< 2 \times 10^{-16}$
Breast	0.22	(0.20–0.23)	$< 2 \times 10^{-16}$
Bladder	0.51	(0.47–0.54)	$< 2 \times 10^{-16}$
Sex (ref: Female)			
Male	1.13	(1.08–1.18)	1.77×10^{-8}
Age (continuous)			
Age (per 1-year increase)	1.023	(1.022–1.025)	$< 2 \times 10^{-16}$
Region (ref: Antofagasta)			
Arica y Parinacota	1.12	(1.05–1.19)	0.000318
Biobío	1.07	(1.02–1.13)	0.0113
Los Ríos	0.92	(0.87–0.96)	0.000645
Maule	0.74	(0.69–0.79)	$< 2 \times 10^{-16}$

Notes: HR = hazard ratio from a unified cause-specific Cox model for cancer-related death. Non-cancer deaths were censored at the date of death; individuals alive at the end of follow-up were administratively censored. Reference categories are shown in parentheses. n = 19,679; cancer deaths (events) = 12,224.

Cancer type was the strongest predictor of cancer-specific death. Using stomach cancer as the reference category, lung cancer showed a substantially higher hazard of cancer-related death (HR=1.80), whereas cervical, colon, breast, and bladder cancers exhibited markedly lower hazards (HRs between 0.22 and 0.51). Age at diagnosis had a small but robust adverse association, with the hazard increasing by approximately 2.3% per additional year (HR per 1-year increase=1.023). Male sex was associated with a higher hazard compared with female sex (HR=1.13). Regional differences persisted after adjustment: compared with Antofagasta, Arica y Parinacota and Biobío showed modestly higher hazards, while Los Ríos and Maule showed lower hazards, with Maule exhibiting the largest regional reduction (HR=0.74). These results indicate substantial heterogeneity in cancer-specific mortality across malignancies and regions, even after accounting for age and sex.

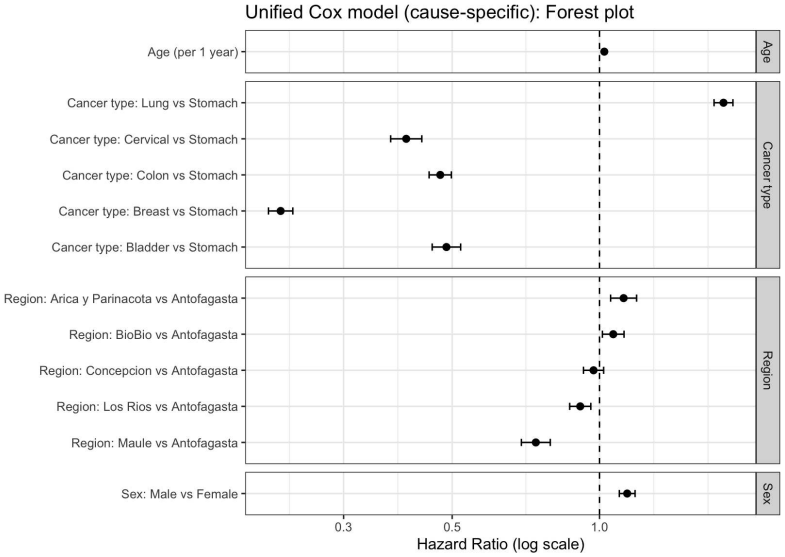


Fig. 8 Forest plot of hazard ratios (HR) from the Cox proportional hazards model. The horizontal bars indicate 95% confidence intervals. Variables include cancer types (relative to stomach cancer as reference), region of diagnosis (relative to Antofagasta), sex (Male vs. Female), and age (per year).

Figure 8 displays the forest plot of the unified cause-specific Cox model, reporting hazard ratios (HRs) and 95% confidence intervals on a logarithmic scale. Age at diagnosis showed a small but statistically robust association with cancer-related mortality, with HRs increasing per one-year increment. Using stomach cancer as the reference category, lung cancer presented substantially higher hazard, while cervical, colon, bladder and breast cancers had lower hazards. Region of diagnosis was also associated with survival, indicating geographically heterogeneous hazards relative to the reference region (Antofagasta). Finally, males exhibited a higher hazard of cancer-related death compared with females after adjustment for age, cancer type and region.

5 Discussion

This study examined the incidence and survival outcomes of six major cancer types across five Chilean regions (Antofagasta, Arica y Parinacota, Biobío, Los Ríos, and Maule) using a combination of exploratory data analysis, survival analysis, and spatial techniques. The findings reveal significant disparities in cancer distribution and outcomes across regions and cancer types. Here, we discuss the implications of these results in the context of Chile's public health landscape, compare them with the existing literature, and highlight the strengths and limitations of our approach.

Our exploratory analysis uncovered clear regional patterns in cancer incidence. For example, lung cancer showed a higher burden in Antofagasta (both in raw cases and incidence rate), whereas stomach cancer incidence was exceptionally high in Los Ríos. These patterns are consistent with known regional risk factors: northern Chile (Antofagasta and Arica) has a history of arsenic contamination and mining-related pollution, which has been linked to increased rates of lung and bladder cancer [21]. The elevated bladder cancer risk in Antofagasta aligns with [18], which reported persistently elevated cancer risks decades after reductions in arsenic exposure. In contrast, central-southern regions have historically exhibited high stomach cancer rates, often attributed to dietary practices and *Helicobacter pylori* prevalence. Our results reinforce this pattern, as Los Ríos exhibited the highest stomach cancer incidence rate despite its smaller population, consistent with high-risk populations described by [19].

The survival analysis highlighted large differences in outcomes by cancer type. Breast cancer showed the most favorable long-term survival, consistent with improvements in early detection and treatment and with Chile's implementation of mammography screening within the GES program since 2005 [17]. In contrast, lung cancer exhibited the poorest survival, with steep early declines in the Kaplan–Meier curves, underscoring the persistent challenge of late diagnosis and limited effectiveness of treatment at advanced stages. This aligns with [8], which emphasized the burden of lung cancer in Chile and the relevance of regional risk contexts such as arsenic exposure in Antofagasta. Stomach cancer also showed poor survival, consistent with reports of high gastric cancer mortality in Chile and ongoing concerns about late diagnosis [7]. Together, these patterns support strengthening prevention policies (e.g., tobacco control), improving timely diagnostic pathways, and considering targeted screening strategies in high-risk settings.

The unified cause-specific Cox model provided a multivariable perspective and confirmed that cancer type, age, sex, and region are significant predictors of cancer-specific mortality (Table 8). Cancer type emerged as the dominant determinant of hazard: relative to stomach cancer (reference), lung cancer was associated with a markedly higher hazard of cancer-related death, whereas breast, cervical, colon, and bladder cancers exhibited substantially lower hazards. Age at diagnosis showed a small but robust adverse association, with hazard increasing steadily per one-year increment. Male sex was associated with a higher hazard compared with female sex after adjustment for age, cancer type, and region.

Importantly, the updated adjusted regional effects do not support a simple pattern of uniformly worse outcomes in Antofagasta. Using Antofagasta as the reference, Arica y Parinacota and Biobío showed modestly higher adjusted hazards, whereas Los Ríos

and Maule showed lower hazards. These regional contrasts may reflect differences in case mix not captured in registry variables (e.g., stage at diagnosis, comorbidity), heterogeneity in access to specialized oncology services, treatment timeliness, and broader contextual factors (environmental, occupational, and socioeconomic). Because our registry data do not include stage or treatment, these regional associations should be interpreted as adjusted differences in hazard rather than causal effects, motivating further investigation with richer clinical data.

Sex differences also persisted after multivariable adjustment, with males experiencing higher cancer-specific hazard. Part of this gap may reflect differential exposure to risk factors (e.g., occupational hazards, smoking) and differences in health-seeking behavior and diagnostic delays. In addition, sex-specific case mix may contribute, given that cancers with comparatively more favorable survival (e.g., breast and cervical) occur predominantly in women, while cancers with poor survival (notably lung and stomach) are relatively more frequent in men. Future work incorporating stage at diagnosis and treatment would help disentangle these mechanisms.

Our findings must be viewed in light of certain limitations. First, the study was limited to the five regions with publicly available registry data, so results may not generalize to all of Chile. Second, registry completeness and outcome ascertainment may affect survival estimates; for example, assuming that missing death records imply survival to the end of follow-up could misclassify outcomes if reporting is incomplete. This underscores the need for continued investment in cancer registry infrastructure and data quality improvement [16]. Third, we lacked key clinical variables such as stage at diagnosis, tumor subtype, and treatment information, which likely explain part of the observed heterogeneity across cancers and regions.

Despite these limitations, this study provides a comprehensive assessment integrating incidence patterns with survival outcomes. The results have practical implications for Chile's cancer control strategy and support region-specific prevention and care planning. In particular, the consistently poor outcomes for lung cancer underscore the importance of intensified prevention and early detection efforts in high-risk populations, while the observed regional contrasts motivate targeted evaluations of access and care pathways to reduce inequities in cancer outcomes across the country.

6 Conclusion

This study demonstrates the value of integrating incidence, survival, and regional analyses to better characterize cancer patterns in Chile. Using Kaplan–Meier methods and a unified cause-specific Cox model, we identified substantial heterogeneity in cancer-specific mortality across cancer types, age, sex, and regions within the Population Cancer Registry (RPC) coverage.

Across cancer types, breast cancer showed the most favorable survival, consistent with the benefits of earlier detection and improved treatment, whereas lung and stomach cancers exhibited the poorest outcomes, highlighting the need for strengthened prevention and earlier diagnosis. The multivariable results further indicate that older age and male sex are associated with higher cancer-specific hazard, and that regional differences persist after adjustment for case mix across the six cancers considered.

These findings support prioritizing interventions for high-burden cancers—notably lung cancer through tobacco control and timely diagnostic pathways, and stomach cancer through targeted prevention and improved access to effective care in high-incidence areas.

From a public health perspective, the observed geographic variation underscores the importance of tailoring cancer control strategies to local contexts and ensuring equitable access to diagnostic and oncology services across regions. At the same time, our analysis highlights key data limitations of emerging registries, particularly incomplete information on stage at diagnosis and treatment, which are crucial for explaining survival differences. Continued investment in cancer registration and surveillance is therefore essential to improve data quality and to enable more detailed evaluations incorporating clinical factors, treatment pathways, and social determinants.

Overall, our results reinforce the need for a multifaceted, data-driven approach to cancer control in Chile. Future work should expand to additional regions as registry coverage grows, extend follow-up to assess temporal trends, and incorporate stage, treatment, comorbidities, and socioeconomic factors to better understand and reduce disparities in cancer outcomes nationwide.

References

[1] Acevedo F, Camus M, Sanchez C. Breast Cancer at Extreme Ages - a Comparative Analysis in Chile. Asian Pacific Journal of Cancer Prevention. 2015 Mar;16(4):1455–1461. https://doi.org/10.7314/APJCP.2015.16.4.1455.

[2] Navarrete Montalvo D, González M N, Montalvo V MT, Jiménez A A, Echiburú-Chau C, Calaf GM. Patterns of Recurrence and Survival in Breast Cancer. Oncology Reports. 2008 Sep;20(3):531–535. https://doi.org/10.3892/or_00000038.

[3] Sierra MS, Soerjomataram I, Antoni S, Laversanne M, Piñeros M, De Vries E, et al. Cancer Patterns and Trends in Central and South America. Cancer Epidemiology. 2016 Sep;44:S23–S42. https://doi.org/10.1016/j.canep.2016.07.013.

[4] Guerrero-Nancuante C, Eguiguren P, Bass C, Garmendia ML. Socio-Economic Factors Related to Premature Death from Colorectal Cancer in Santiago de Chile, 2014–2018: A Cross-Sectional Study. Public Health. 2024 Jun;231:1–6. https://doi.org/10.1016/j.puhe.2024.03.003.

[5] Lorenzo Bermejo J, Boekstegers F, González Silos R, Marcelain K, Baez Benavides P, Barahona Ponce C, et al. Subtypes of Native American Ancestry and Leading Causes of Death: Mapuche Ancestry-Specific Associations with Gallbladder Cancer Risk in Chile. PLOS Genetics. 2017 May;13(5):e1006756. https://doi.org/10.1371/journal.pgen.1006756.

[6] Smith AH, Marshall G, Roh T, Ferreccio C, Liaw J, Steinmaus C. Lung, Bladder, and Kidney Cancer Mortality 40 Years After Arsenic Exposure Reduction. JNCI: Journal of the National Cancer Institute. 2018 Mar;110(3):241–249. https://doi.org/10.1093/jnci/djx201.

[7] Manterola C, Vial M, Roa JC. Survival of a Cohort of Patients with Intermediate and Advanced Gall Bladder Cancer Treated with a Prospective Therapeutic Protocol. Acta Cirurgica Brasileira. 2010 Jun;25(3):225–230. https://doi.org/10.1590/S0102-86502010000300003.

[8] Burotto M, Aren O, Renner A, Samtani S, Jimenez De La Jara J. Lung Cancer in Chile. Journal of Thoracic Oncology. 2019 Sep;14(9):1504–1509. https://doi.org/10.1016/j.jtho.2019.02.024.

[9] Mondschein S, Subiabre F, Yankovic N, Estay C, Von Mühlenbrock C, Berger Z. Colorectal Cancer Trends in Chile: A Latin-American Country with Marked Socioeconomic Inequities. PLOS ONE. 2022 Nov;17(11):e0271929. https://doi.org/10.1371/journal.pone.0271929.

[10] Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association. 1958 Jun;53(282):457–481. https://doi.org/10.1080/01621459.1958.10501452.

[11] Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society Series B: Statistical Methodology. 1972 Jan;34(2):187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x.

[12] Bray F, Colombet M, Mery L, Piñeros M, Znaor A, Zanetti R, et al. Cancer Incidence in Five Continents Volume XI. No. 166 in IARC Scientific Publication. Lyon, France: International Agency for Research on Cancer; 2021.

[13] World Health Organization. International Classification of Diseases for Oncology (ICD-O): First Revision. 3rd ed. Nonserial Publications. Geneva: World Health Organization; 2013.

[14] World Health Organization, editor. International Statistical Classification of Diseases and Related Health Problems. 10th ed. Geneva: World Health Organization; 2004.

[15] Nicolis O, De Los Angeles D, Taramasco C. A Contemporary Review of Breast Cancer Risk Factors and the Role of Artificial Intelligence. Frontiers in Oncology. 2024 Apr;14:1356014. https://doi.org/10.3389/fonc.2024.1356014.

[16] Oliveira SCD, Machado CV, Hein AA, Almeida PFD. Políticas de Saúde No Chile (2000-2018): Trajetória e Condicionantes. Cadernos de Saúde Pública. 2020;36(11):e00002120. https://doi.org/10.1590/0102-311x00002120.

[17] Puschel K, Thompson B. Mammogram Screening in Chile: Using Mixed Methods to Implement Health Policy Planning at the Primary Care Level. The Breast. 2011 Apr;20:S40–S45. https://doi.org/10.1016/j.breast.2011.02.002.

[18] Steinmaus CM, Ferreccio C, Romo JA, Yuan Y, Cortes S, Marshall G, et al. Drinking Water Arsenic in Northern Chile: High Cancer Risks 40 Years after Exposure Cessation. Cancer Epidemiology, Biomarkers & Prevention. 2013 Apr;22(4):623–630. https://doi.org/10.1158/1055-9965.EPI-12-1190.

[19] Heise K, Bertran E, Andia ME, Ferreccio C. Incidence and Survival of Stomach Cancer in a High-Risk Population of Chile. World Journal of Gastroenterology. 2009;15(15):1854. https://doi.org/10.3748/wjg.15.1854.

[20] Instituto Nacional de Estadísticas.: Censo de Población y Vivienda. https://www.ine.gob.cl/estadisticas/sociales/censos-de-poblacion-y-vivienda/censo-de-poblacion-y-vivienda. Available from: http://www.censo2017.cl/microdatos/.

[21] García D, Tapia J, Aguilera J, Vega C, Zúñiga P, Lavin P, et al. Contamination of Urban Soils in a Historical Mining Town of Northern Chile. Environmental Earth Sciences. 2024 Mar;83(6):182. https://doi.org/10.1007/s12665-024-11445-0.

Statements and Declarations

Funding. This work was funded by the Agencia Nacional de Investigación y Desarrollo (ANID) FONDAP program, grant #152220002 (Center for Cancer Prevention and Control, CECAN).

Competing Interests. The authors declare they have no relevant financial or non-financial interests to disclose.

Author Contributions. Conceptualization, O.N. and C.T.; methodology, O.N.; software, O.N. and D.D.; validation, O.N. and D.D.; formal analysis, O.N. and D.D.; investigation, O.N.; data curation, D.D., D.M.; writing—original draft preparation, O.N.; writing—review and editing, O.N., F.R., D.M., and C.T.; visualization, O.N., D.M., and D.D.; supervision, O.N.; project administration, C.T.; funding acquisition, C.T. All authors have read and agreed to the published version of the manuscript.

Data Availability. Ministry of Health (Ministerio de Salud, MINSAL). The dataset can be accessed via the official MINSAL epidemiological surveillance portal at https://epi.minsal.cl/registros-poblacionales-de-cancer-bd/. However, access to the dataset may require a VPN connection from within Chile. If access is not available for any reason, both the original and the preprocessed data are available from the authors upon request.

Ethics Approval. This study uses publicly available data; therefore, no ethical approval is required.