JHGD
Journal of Human Growth and Development
www. jhgd.com.br

ORIGINAL ARTICLE

Trends of incompleteness of epidemiological variables in hospital cancer registries of women with malignant breast neoplasia in Espírito Santo, Brazil

Raphael Manhães Pessanha[a], Wesley Rocha Grippa[b], Luiz Cláudio Barreto Silva Neto[c], Larissa Soares Dell'Antonio[d], Cristiano Soares Silva Dell'Antonio[d], Luís Carlos Lopes-Júnior[a,b]

Open acess

[a]Programa de Pós-graduação de Saúde Coletiva (PPSCG), Centro de Ciências da Saúde, Universidade Federal do Espírito Santo (UFES), Vitória, ES, Brasil;

[b]Departamento de Matemática Aplicada (DMA), Centro Universitário Norte do Espírito Santo (CEUNES), Universidade Federal do Espírito Santo (UFES), São Mateus, ES, Brasil;

[c]Programa de Pós-graduação em Nutrição e Saúde (PPGNS), Centro de Ciências da Saúde, Universidade Federal do Espírito Santo (UFES), Vitória, ES, Brasil;

[d]Secretaria Estadual de Saúde do Estado do Espírito (SESA). Vitória, ES, Brasil.

Corresponding author
lopesjr.lc@gmail.com

Manuscript received: october 2024
Manuscript accepted: december 2024
Version of record online: april 2025

Abstract

Introduction: the World Health Organization classifies Non-Communicable Diseases as the leading cause of global mortality, with 71% of deaths associated with malignant neoplasms. In 2020, approximately 19.3 million new cancer cases and 10 million cancer-related deaths were reported worldwide, with the Americas accounting for 20.9% of global incidence. In Brazil, the National Cancer Institute estimates 704,000 new cancer cases for the 2023-2025 period, with breast and prostate cancers representing 15% of the new cases each. Age over 50 years, genetic factors, family history, and others are risk factors for breast cancer.

Objective: to evaluate the trend of incompleteness in the variables of the Hospital Cancer Registry for cases of malignant breast neoplasms across all hospitals in the Oncology Care Network of Espírito Santo.

Methods: this is a retrospective time-series study using secondary data from the Tumor Registration Forms of the Hospital Cancer Registries of the Oncology Care Network of Espírito Santo between 2000 and 2020. Data incompleteness was classified as excellent (<5%), good (5%-10%), regular (10%-20%), poor (20%-50%), and very poor (>50%), based on the percentage of missing information. The Mann–Kendall test was used to assess temporal trends, and the Friedman test was applied to evaluate the quality classifications of the historical series.

Results: a total of 16,587 breast cancer cases were recorded in the Hospital Cancer Registries of Espírito Santo between 2000 and 2020. The variables alcohol use ($p<0.001$), smoking ($p<0.001$), family history of cancer ($p<0.001$), and marital status ($p<0.001$) showed a statistically significant decreasing trend in incompleteness. On the other hand, variables related to diagnosis/treatment and tumor characteristics exhibited a lower trend of incompleteness over time, with an increase in completeness. These variables included the type of diagnostic exam performed ($p=0.03$), previous diagnoses and treatments ($p=0.008$), key diagnostic bases for the tumor ($p<0.001$), tumor laterality ($p=0.008$), and presence of more than one primary tumor ($p=0.01$).

Conclusion: the scores for the epidemiological variables in the Hospital Cancer Registries of women with breast cancer in Espírito Santo were predominantly classified as "excellent" in terms of completeness. However, clinically important variables, such as Tumor, Node, Metastasis staging, showed a decline in completeness throughout most of the historical series.

Keywords: breast cancer, oncology, hospital cancer registry, epidemiology, public health surveillance.

Suggested citation: Pessanha RM, Grippa WR, Neto LCBS, Dell'Antonio LS, Dell'Antonio CSS, Lopes-Júnior LC. Trends of incompleteness of epidemiological variables in hospital cancer registries of women with malignant breast neoplasia in Espírito Santo, Brazil. *J Hum Growth Dev. 2025; 35(1):111-123.* DOI: http://doi.org/10.36311/jhgd.v35.15861

Authors summary

Why was this study done?
The aim of this study was to assess the trend in the incompleteness of variables in the Hospital Cancer Registry (HCR) of breast cancer cases across all hospitals within the Oncology Care Network of Espírito Santo.

What did the researchers do and find?
The researchers conducted a retrospective observational study using tumor registry records from the HCRs of breast cancer, covering all hospitals in the Oncology Care Network of Espírito Santo state between 2000 and 2020. A total of 16,587 breast cancer cases were recorded in the historical series. Variables such as alcoholism, smoking, family history of cancer, and marital status showed a significant downward trend in completeness over time. On the other hand, variables related to diagnosis/treatment and tumor characteristics exhibited a smaller trend of incompleteness over time, with an increase in completeness, including variables like diagnostic exam type, main diagnostic bases for the tumor, tumor laterality, and the presence of more than one primary tumor.

What do these findings mean?
In summary, most epidemiological variables in the breast cancer HCRs from Espírito Santo, Brazil, were classified as having excellent completeness. However, important variables such as Tumor, Node, Metastasis (TNM) staging and clinical tumor staging by group exhibited high levels of incompleteness throughout the historical series. There is a pressing need for consistent, high-quality data in HCRs to improve the monitoring of epidemiological variables in tumor records. The study highlights the importance of the implementation, maintenance, updating, and availability of HCR data to better understand the cancer landscape and enhance monitoring and control efforts.

Highlights
The scores for the epidemiological variables in the hospital cancer registries of women with breast cancer in Espírito Santo were mostly classified as “excellent” in terms of completeness.
Clinically significant variables, such as Tumor, Node, Metastasis staging, showed a decline in completeness throughout most of the historical series.
Hospital Cancer Registries are systematic sources of information for evaluating the quality of hospital care provided to oncology patients.

INTRODUCTION

The World Health Organization classifies Non-Communicable Diseases (NCDs) as the leading cause of morbidity and mortality globally[1]. Malignant neoplasms account for 41 million deaths annually, representing 71% of all global deaths[2]. The incidence and mortality rates of cancer are rising rapidly worldwide, reflecting demographic, epidemiological, and nutritional transitions, as well as an increase in cancer risk factors[3,4].

The latest GLOBOCAN estimates for 2020 reported approximately 19.3 million new cancer cases (18.1 million excluding non-melanoma skin cancer) and 10 million cancer-related deaths (9.9 million excluding non-melanoma skin cancer)[1]. In the Americas, 20.9% of global cancer incidence and 14.2% of cancer-related mortality were reported, compared to other continents[1]. Among the highest expected incidence rates are found in Australia, New Zealand, Northern Europe, and Western Europe[1].

In Brazil, the National Cancer Institute (INCA) estimates 704,000 new cancer cases for the 2023-2025 period[5]. Excluding non-melanoma skin cancer, breast and prostate cancers each account for 15% of new cases[5]. In 2020, the global mortality rate for breast cancer, age-adjusted to the world population, was 11.84 deaths per 100,000 women. In the Southeast and South regions of Brazil, the highest rates were 12.64 and 12.79 deaths per 100,000 women, respectively[6]. Age over 50 years is a primary risk factor for breast cancer development; however, genetic factors (mutations in the BRCA1 and BRCA2 genes), family history of cancer, late menopause, obesity, and frequent exposure to ionizing radiation also represent significant risk factors[4,7].

Breast cancer is the most prevalent cancer among women (excluding non-melanoma skin cancer) across all regions in Brazil, with an incidence of 61.61 new cases per 100,000 women[6]. The Southeast region has the highest estimated risk, with 81.06 cases per 100,000 inhabitants[5]. In the state of Espírito Santo, 790 new cases of female breast cancer were estimated for 2022[5].

The Brazilian Hospital Cancer Registries (HCR) are a key data collection tool for cancer, helping to qualify healthcare services for planning purposes. These registries enable the evaluation of healthcare quality, the development of clinical research, and the creation of public health policies[8-10]. The National Cancer Institute (INCA) uses the HCR to monitor the epidemiological evolution of cancer every three years, offering nationwide training to improve hospital management, clinical treatments, health promotion, and disease prevention efforts. Additionally, these data are used for clinical and epidemiological research[8-10].

Advanced-stage breast cancer represents a barrier to early diagnosis and the initiation of treatment, making public health planning, especially at the primary care level, essential. Furthermore, public programs that provide epidemiological data on the social determinants of health, which shape cancer risk factors, are crucial[11].

A recent analysis of a time series up to 2020, encompassing the entire Oncology Care Network of Espírito Santo-comprising a CACON and seven High Complexity Oncology Assistance Units (UNACON) aimed at guiding cancer surveillance actions in the region, including the monitoring and evaluation of the HCRs of hospitals in the state oncology network, particularly regarding malignant breast neoplasm cases, remains unclear.

Thus, the objective of this study is to evaluate the trend of incompleteness in the variables of the HCR for malignant breast neoplasms in the hospitals of the Oncology Care Network of Espírito Santo, Brazil.

METHODS

Study design, location and period

This is a retrospective time series study using secondary data from the Oncology Care Network of Espírito Santo (ES), which consists of one High Complexity Oncology Center (CACON) and seven High Complexity Oncology Assistance Units (UNACON). The study covers the period from 2000 to 2020.

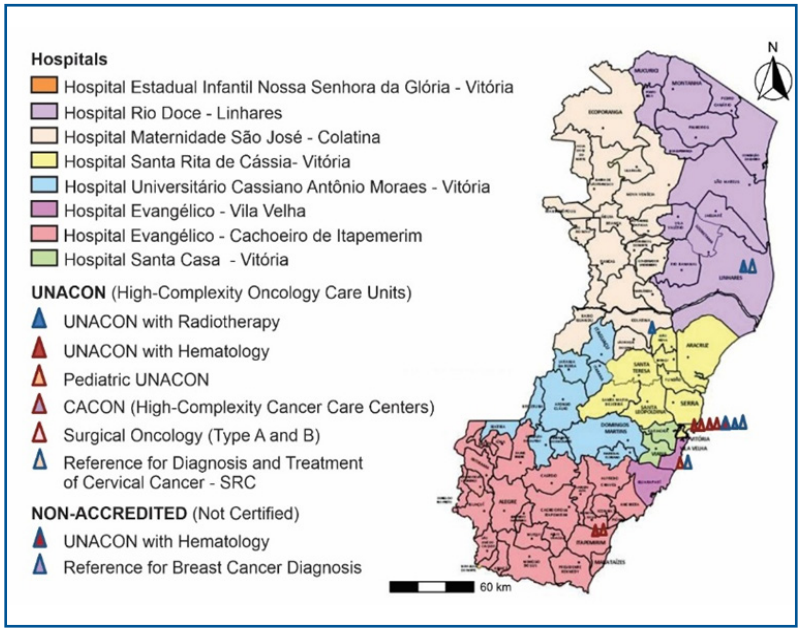

Figure 1: Spatial Distribution of the Oncology Care Network in Espírito Santo
Source: Adapted from (Espírito Santo, 2023).

The secondary data used were retrieved from the SIS-HCR, part of the INCA Integrated System, as well as from the Hospital Cancer Registries of the state of Espírito Santo, collected in collaboration with the State Health Department of Espírito Santo (SESA/ES). The state of Espírito Santo has an Oncology Care Network distributed across three health regions: Central-North, Metropolitan, and South (Figure 1)[9,10].

Population, eligibility criteria, and data collection

Data were collected from February to June 2023 in collaboration with the State Health Department of Espírito Santo (SESA/ES). All records of patients diagnosed with breast cancer, as per the International Statistical Classification of Diseases and Related Health Problems (ICD-10), C50: Malignant Neoplasm of the Breast, from the historical series studied (2000 to 2020), were extracted from the HCR database of the state of Espírito Santo via SESA/ES. Both analytical cases (where planning and treatment occur at the hospital where the record is made) and non-analytical cases (patients who arrive at the hospital already treated or who do not undergo the prescribed treatment) were included. It is important to note that all hospitals in the Oncology Care Network have up to two years to submit data for consolidation, which is why the complete historical series up to 2020 was chosen for this study. Regarding the epidemiological variables in this study, 44 variables from the tumor registration form provided by INCA were selected[12].

The quality dimensions defined by Lima *et al.*[13] were used, where completeness is determined by the proportion of fields containing non-null values. For the analysis of completeness, the classification described by Romero and Cunha[14] was applied. A variable's completeness was considered excellent when the filling percentage was >95%; good, when it ranged from 90.1% to 95%; regular, from 80.1% to 90%; poor, between 50.1% and 80%; and very poor, when ≤50%[14]. Completeness refers to the extent to which the analyzed fields are filled, measured by the proportion of notifications with a category different from those indicating missing data. In this study, a field marked as "ignored," zero, an unknown date, or a term indicating missing data were considered incomplete[9,14].

Data analysis

Statistical analyses were conducted using the free software RStudio (version 2022.07.2) and R (version 4.1.0). The completeness of the data was described based on the observed relative frequency and their respective completeness scores. The Friedman test was used to compare score classifications across different years. Additionally, the Mann-Kendall test[15] was employed to assess the presence of statistically significant temporal trends across the historical series evaluated. A significance level of 5% was adopted for all statistical analyses.

Ethical aspects

This research project was approved by the Research Ethics Committee of the Center for Health Sciences at the Federal University of Espírito Santo (UFES), under the opinion number 3.831.617, and Certificate for Ethical Appreciation: 25985219.3.0000.5060. We also obtained approval and authorization from the State Health Department of Espírito Santo (SESA) to collect secondary data and restricted data related to this research, in accordance with the recommendations of Resolution No. 466/12 of the National Health Council and the Guidelines and Regulatory Norms for Research Involving Human Beings in Brazil.

RESULTS

During the historical series from 2000 to 2020, a total of 16,587 cases of female breast cancer were recorded from the HCR across the entire Oncology Care Network of the state of Espírito Santo. Table 1 presents the percentage of incompleteness of the variables in the breast cancer tumor registration form for the evaluated historical series (Table1).

The Friedman test indicated no significant difference in the scores across the years (p=0.228), suggesting stability and consistency in the categories assigned to each variable. Table 2 presents the significant trends of either an increase or decrease in the incompleteness of the tumor registration form variables, as assessed using the Mann-Kendall test. The analysis results revealed a variety of incomplete patterns in the epidemiological variables examined.

Upon examining the sociodemographic variables of the patients, no significant trends of incompleteness were found for age (p=0.6077), race/ethnicity (p=0.9278), or educational level (p=0.8044). However, variables such as place of birth (p=0.0145), number of clinics attended (p=0.0218), clinical treatments (p<0.001), family history of cancer (p<0.001), alcohol consumption (p<0.001), and smoking (p<0.001) exhibited significant trends of incompleteness.

Regarding the variables related to diagnosis and treatment, patterns of increased incompleteness were observed. The variable "Federal Unit of residence" demonstrated significance (p=0.0112), suggesting an increase in data incompleteness. Similarly, the patients' place of origin also exhibited a significant trend of increased incompleteness (p=0.0112). Other breast cancer-related variables, such as previous diagnoses and treatments (p=0.0085), key diagnostic criteria (p<0.001), tumor laterality (p=0.0086), presence of multiple primary tumors (p=0.0136), other tumor clinical staging (p=0.0124), and the pTNM classification system (p=0.0037), also showed significance with an increase in incompleteness over time (Table 2).

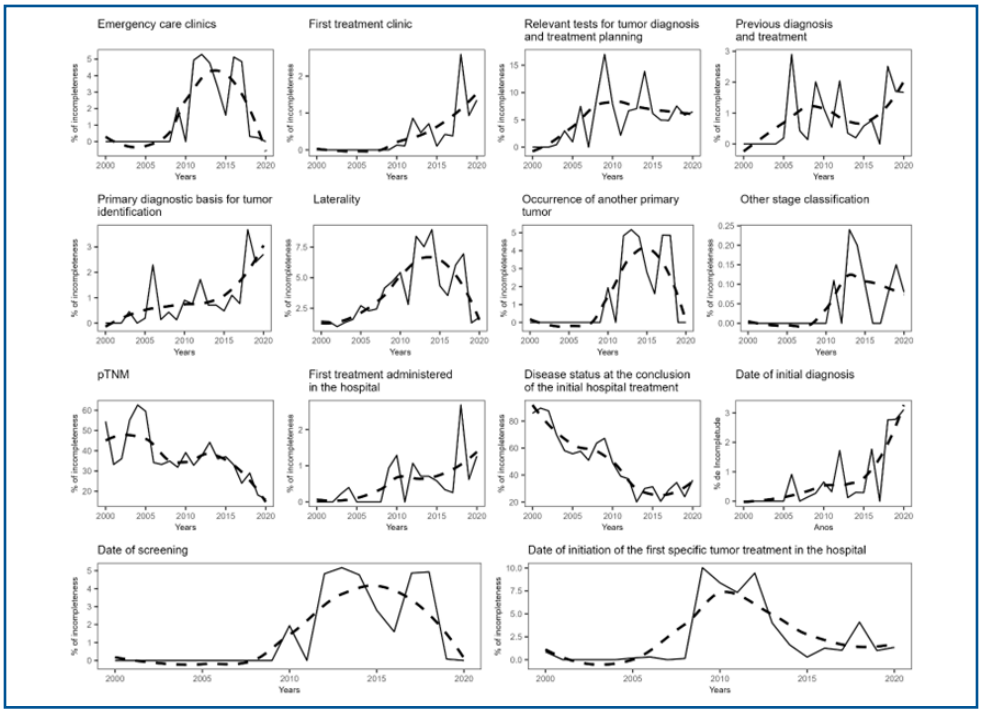

Figure 2: Trends in the incompleteness of sociodemographic variables in Hospital Cancer Records of women with breast cancer in the Oncology Care Network of Espírito Santo from 2000 to 2020 (N = 16,587)

Note: The dashed line represents the temporal trend, while the solid black line indicates the temporal evolution of incompleteness.

In figure 2, the graphs display the percentage of incompleteness for the sociodemographic, health, and lifestyle variables in the present historical series. An upward trend in incompleteness is observed for variables such as place of birth, marital status, and patient occupation, while the other variables showed fluctuations over time.

The trend curves for the incompleteness of clinical information in the tumor registration form variables of the HCR are shown in figure 3. All variables exhibited significant changes throughout the historical series. Variables such as key diagnostic criteria and disease status at the end of treatment had a high percentage of incompleteness in the year 2000. However, over the years of the temporal series, this percentage decreased, indicating an improvement in cancer records across all hospitals.

On the other hand, variables such as treatment clinics and most relevant exams showed a low percentage of incompleteness in the early years of the series but experienced an increase in incompleteness over time. Other variables followed a different trend pattern. For instance, the occurrence of multiple primary tumors, screening date, and treatment start date showed a low incompleteness early in the series but experienced a spike in incompleteness in the middle years, with a subsequent decrease in the following years.

Table 1: Percentage of incompleteness and classification of completeness for epidemiological variables in Hospital Cancer Records (HCRs) related to female breast cancer cases across all hospitals in the Oncology Care Network of Espírito Santo from 2000 to 2020 (N = 16,587)

Column	Incompleteness (%)	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Alcoholism	(%)	92.8	78.9	84.3	83.5	80.9	73.1	64.6	58.2	61.1	63.1	14.8	10.2	13.1	13.9	21	37.3	38.3	37.2	28
	Score	VB	VB	VB	VB	VB	VB	VB	VB	VB	VB	R	R	R	B	B	B	B	B	B
Year of initial diagnosis	(%)	0	0	0	0	0	0.91	7.88	0.14	0.27	0.65	0.32	1.72	0.12	0.3	0.29	1.77	0	2.76	2.78
	Score	E	E	E	E	E	E	G	E	E	E	E	E	E	E	E	E	E	E	E
Year of screening	(%)	0	0	0	0	0	0	0	0	0	1.94	0	4.83	5.17	4.77	2.8	1.6	4.87	4.93	0.08
	Score	E	E	E	E	E	E	E	E	E	E	E	E	G	E	E	E	E	E	E
Diagnostic basis for tumor identification	(%)	0	0	0	0.45	0	0.19	2.29	0.14	0.43	0.13	0.9	0.76	1.72	0	0.71	0.48	1.1	0.77	3.68
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Emergency care clinics	(%)	0	0	0	0	0	0	0	0	0	2.06	0	4.94	5.29	4.77	3.28	1.6	5.13	4.85	0.31
	Score	E	E	E	E	E	E	E	E	E	E	E	E	G	E	E	E	G	E	E
Treatment initiation clinic	(%)	0	0	0	0	0	0	0	0	0	0	0.13	0.11	0.86	0.47	0.71	0.1	0.42	0.38	2.59
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Year of initial specific cancer treatment	(%)	0.87	0	0	0	0	0.19	0.3	0	0.14	10	8.39	7.34	9.45	4	1.62	0.29	1.26	1.03	4.1
	Score	E	E	E	E	E	E	E	E	E	R	G	G	G	E	E	E	E	E	E
Date of death	(%)	0	0	0	0	0.2	0	0.15	0.43	0	0.13	0.26	0.11	0	0	0	0	0.08	0	0.08
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Previous diagnoses and treatments	(%)	0	0	0	0	0	0.19	2.9	0.43	0.14	2.01	1.16	0.54	2.04	0.35	0.2	0.58	0.76	0	2.51
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Date of initial diagnosis	(%)	0	0	0	0	0	0	0.91	0	0.14	0.27	0.65	0.32	1.72	0.12	0.3	0.29	1.77	0	2.76
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Year of initiation of first specific cancer treatment	(%)	3.48	4.79	3.48	3.12	1.79	3.88	1.83	2.29	0.57	10	8.39	7.34	9.45	9.17	8.32	5.21	6.82	7.31	10.9
	Score	E	E	E	E	E	E	E	E	E	R	G	G	G	G	G	G	G	G	R
Date of screening	(%)	0	0	0	0	0	0	0	0	0	0	1.94	0	4.83	5.17	4.77	2.8	1.6	4.87	4.93

Table 1: Percentage of incompleteness and classification of completeness for epidemiological variables in Hospital Cancer Records (HCRs) related to female breast cancer cases across all hospitals in the Oncology Care Network of Espírito Santo from 2000 to 2020 (N = 16,587)

	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	G	E	E	E	E	E
Clinical Tumor Staging (TNM Classification)	(%)	29.3	28.5	27.1	36.4	36.4	26.4	17.1	20.1	21.6	24.8	27.6	29.8	37.5	40.8	33	30.4	25	24.1	35.5
	Score	B	B	B	B	B	B	R	B	B	B	B	B	B	B	B	B	B	B	B
State of residence	(%)	0	0	0	0	0	0	0	0	0	0	0	0.22	0.21	0.24	0.3	0.19	0.67	0.38	0.17
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Current marital status	(%)	0	0	0.75	0	0.99	4.46	3.81	1.58	3.15	4.95	6.19	4.43	6.77	8.58	5.38	3.47	2.02	5.26	5.1
	Score	E	E	E	E	E	E	E	E	E	E	G	E	G	G	G	E	E	G	G
Disease status at the conclusion of initial hospital treatment	(%)	85.8	89.6	87.6	69.9	58.1	55.8	57.6	51	63.7	67.2	49.3	39.2	37.2	20.1	30	31.4	20.6	29	34.5
	Score	VB	VB	VB	VB	VB	VB	VB	VB	VB	VB	B	B	B	B	B	B	B	B	B
Relevant diagnostic and treatment planning tests	(%)	0	0	0	0.45	2.98	0.97	7.47	0	8.87	17	7.87	2.16	6.66	7.05	13.9	6.18	4.97	4.87	7.53
	Score	E	E	E	E	E	E	G	E	G	R	G	E	G	G	R	G	E	E	G
Family history of cancer	(%)	85.2	74.4	62.7	62.3	61.8	57	57.2	50.1	46.2	53.6	52	41	43.5	41.7	43	46.9	45.2	50.9	43.7
	Score	VB	VB	VB	VB	VB	VB	VB	VB	B	VB	VB	B	B	B	B	B	B	VB	B
Age	(%)	0.58	0.56	3.98	5.36	4.97	24.6	32.2	26.1	19.2	35.5	45.9	26	30	24.6	24.1	17.2	19.6	10.3	18.8
	Score	E	E	E	G	E	B	B	B	R	B	B	B	B	B	B	R	R	R	R
Education	(%)	0	0	0	0	0	0	0	0	0	0	0	0	0.11	0	0	0	0	0	0
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Laterality	(%)	1.45	1.41	1	1.34	1.59	2.71	2.29	2.44	4.15	4.69	5.42	2.81	8.38	7.52	8.92	4.34	3.54	6.03	6.94
	Score	E	E	E	E	E	E	E	E	E	E	G	E	G	G	G	E	E	G	G
Place of birth	(%)	8.7	0	3.23	9.82	8.55	2.71	3.66	0.29	0.72	3.21	2.97	3.67	3.76	5.41	6.29	4.54	5.14	6.41	9.78
	Score	G	E	E	G	G	E	E	E	E	E	E	E	E	G	G	E	G	G	G
Presence of another primary malignancy	(%)	0	0	0	0	0	0	0	0	0	0	1.94	0	4.83	5.17	4.77	2.8	1.6	4.87	4.85
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	G	E	E	E	E	E
Occupation	(%)	0	0.28	1.49	1.56	3.18	10.7	12.2	12.8	8.44	7.5	11.5	5.83	14.8	13.8	18.5	11.7	12.3	9.49	17.2
	Score	E	E	E	E	E	R	R	R	G	G	R	G	R	R	R	R	R	G	R

Table 1: Percentage of incompleteness and classification of completeness for epidemiological variables in Hospital Cancer Records (HCRs) related to female breast cancer cases across all hospitals in the Oncology Care Network of Espírito Santo from 2000 to 2020 (N = 16,587)

Referral source	(%)	15.9	20.3	21.4	7.81	8.55	10.5	25.3	23.2	21.5	13.9	16.7	10.9	16.5	18.6	21	17.1	13.5	20	13.6
	Score	R	B	B	G	G	R	B	B	B	R	R	R	R	R	B	R	R	B	R
Other stage classification	(%)	0	0	0	0	0	0	0	0	0	0	0	0.11	0	0.24	0.2	0.1	0	0	0.08
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Initial treatment received in hospital	(%)	0	0	0	0.22	0.4	0	0	0	0	0.94	1.29	0	1.07	0.71	0.71	0.58	0.34	0.26	2.68
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Municipality of residence	(%)	0	0	0	0	0	0	0	0	0	0	0	0.22	0.21	0.24	0.3	0.19	0.67	0.38	0.17
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
pTNM	(%)	54.5	33.2	36.1	55.1	62.6	59.5	34.2	33.2	34.9	31.9	39.1	32.8	37.5	44.2	36	37.1	33	24	29
	Score	VB	B	B	VB	VB	VB	B	B	B	B	B	B	B	B	B	B	B	B	B
Race/ethnicity	(%)	1.45	2.82	1.99	0.67	0.8	7.36	21.8	17.3	6.58	11.7	2.71	11.7	11.1	14.2	8.11	4.15	2.27	0.77	3.34
	Score	E	E	E	E	E	G	B	R	G	R	E	R	R	R	G	E	E	E	E
Primary reason for treatment non-compliance	%	0.58	1.13	1.49	0.45	1.59	1.74	77.1	3.3	72.1	67.9	64.9	2.81	3.11	3.06	11.1	2.51	4.38	5.9	12.4
	Score	E	E	E	E	E	E	VB	E	VB	VB	VB	E	E	E	R	E	E	G	R
Smoking	(%)	82.3	65.6	70.4	66.3	67.4	57.4	51.7	48.6	52.2	55.7	14.1	8.96	11.4	13	20.2	34.7	32.9	29.1	22.8
	Score	VB	VB	VB	VB	VB	VB	VB	B	VB	VB	R	G	R	R	B	B	B	B	B
TNM	(%)	38	39.7	37.3	47.1	45.9	35.5	27.9	38.8	32.9	32.9	48.1	58.6	58.5	56.6	43.2	43.9	43	44.9	46.3
	Score	B	B	B	B	B	B	B	B	B	B	B	VB	VB	VB	B	B	B	B	B
Hospital CNES (National Health Establishment Registry) number	(%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Date of first oncology consultation	(%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Year of 1st consultation	(%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E

Table 1: Percentage of incompleteness and classification of completeness for epidemiological variables in Hospital Cancer Records (HCRs) related to female breast cancer cases across all hospitals in the Oncology Care Network of Espírito Santo from 2000 to 2020 (N = 16,587)

Primary location	(%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Detailed primary location	(%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Municipality of the hospital	(%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Sex	(%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Histological type	(%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Case classification	(%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
State of the hospital	(%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Score	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Number of cases	n	345	355	402	448	503	516	656	698	699	747	775	926	931	851	986	1036	1187	780	1196
	%	2.08	2.14	2.42	2.7	3.03	3.11	3.95	4.21	4.21	4.5	4.67	5.58	5.61	5.13	5.94	6.25	7.16	4.7	7.21

Note: E: excellent; G: good; R: regular; B: bad; VB: very bad.

Table 2: Evaluation of the trend in incompleteness for epidemiological variables in Hospital Cancer Records (HCRs) related to female breast cancer cases across all hospitals in the Oncology Care Network of Espírito Santo from 2000 to 2020 (N = 16,587)

Variable	S*	p	Trend
Age	18	0.6077	Not significant
Place of birth	82	0.0145	Increase
Race/ethnicity	4	0.9278	Not significant
Education	4	0.8044	Not significant
Emergency care clinics	71	0.0218	Increase
First treatment clinic	131	< 0.001	Increase
Family history of cancer	-150	< 0.001	Decrease
Alcoholism	-120	< 0.001	Decrease
Smoking	-118	< 0.001	Decrease
State of residence	74	0.0112	Increase
Code of the Municipality of residence	74	0.0112	Increase
Year of diagnosis	75	0.0225	Increase
Referral source	-32	0.3492	Not significant
Relevant tests for tumor diagnosis and treatment planning	72	0.0314	Increase
Current marital status	111	< 0.001	Increase
Year of screening	66	0.0288	Increase
Previous diagnosis and treatment	87	0.0085	Increase
Primary diagnostic basis for tumor identification	127	< 0.001	Increase
Laterality	88	0.0086	Increase
Occurrence of another primary tumor	72	0.0136	Increase
TNM	40	0.2389	Not significant
Clinical tumor staging (TNM)	27	0.4322	Not significant
Other stage classification	70	0.0124	Increase
pTMN	-97	0.0037	Decrease
Primary reason for not receiving antineoplastic treatment in the hospital	62	0.06547	Not significant
Year of initiation of the first specific tumor treatment in the hospital	63	0.06106	Not significant
First treatment administered in the hospital	87	0.0074	Increase
Disease status at the conclusion of the initial hospital treatment	-146	< 0.001	Decrease
Primary occupation	132	< 0.001	Increase
Date of initial diagnosis	120	< 0.001	Increase
Date of screening	84	0.0052	Increase
Date of initiation of the first specific tumor treatment in the hospital	68	0.0415	Increase
Date of death	17	0.5992	Not significant

*Mann-Kendall test to analyze trend - For significance, p-value < 0.05

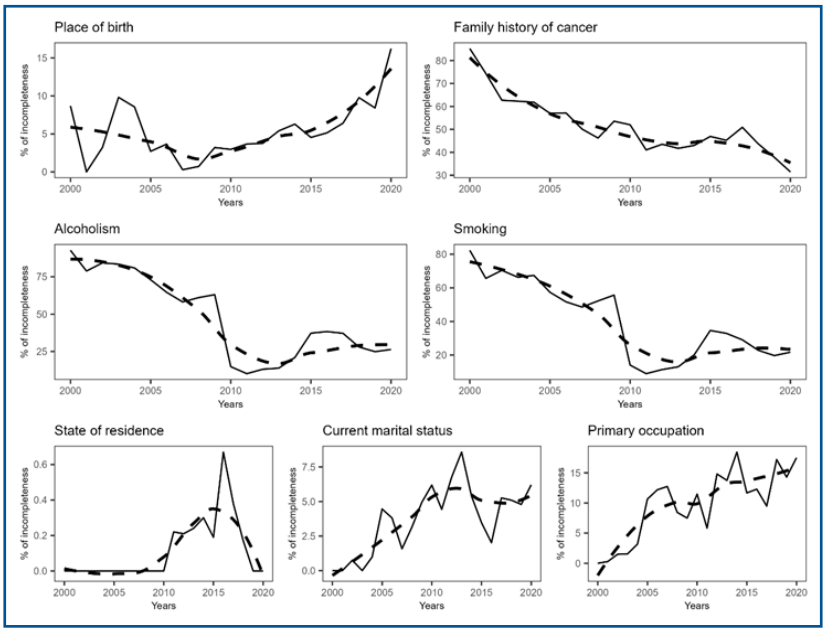

Figure 3: Trends in the incompleteness of clinical variables in Hospital Cancer Records of women with breast cancer in the Oncology Care Network of Espírito Santo from 2000 to 2020 (N = 16,587)

Note: The dashed line represents the temporal trend, while the solid black line indicates the temporal evolution of incompleteness.

DISCUSSION

Studies that evaluate data incompleteness provide a comprehensive view of the completeness of epidemiological variables in cancer registries, offering insights into data quality improvement within the Oncology Care Network of Espírito Santo. Our findings highlighted that certain variables, such as family history of cancer, smoking, alcoholism, pTNM, and disease status at the end of treatment, showed significant trends of decreased data completeness.

In Brazil, a study using a prostate cancer database found that sociodemographic variables, such as race/color, were classified as having regular incompleteness, while the place of birth showed a 14.25% incompleteness in 2000 (regular) but achieved excellent ratings in 13 years and good ratings in 3 subsequent years of the time series[10]. The education level variable was classified as poor in 2009 (24.14% of missing data) and 2010 (38.31% of missing data) but received excellent ratings in 8 years and good ratings in 7 years. In contrast, in our study, the education level variable consistently received an excellent score throughout the entire time series[10].

In another temporal series study conducted in Brazil, the authors evaluated the consistency of epidemiological variables in hospital records, such as origin and marital status, and demonstrated high consistency and stability in their records over the years, indicating a well-maintained data process[9]. However, for clinical variables like cancer staging by TNM, the classification was "very poor". Similarly, in our study, the TNM variable consistently received a "very poor" score throughout the entire time series[9].

In the Midwest region of Brazil, a study conducted in the state of Mato Grosso explored the quality of cancer hospital records by accessing data completeness and consistency. The study demonstrated that some variables, such as "marital status" and "sex," showed high consistency over time, indicating stable data quality[16]. The authors also emphasized that in that state, variables such as "race/skin color," "education," and "occupation" faced challenges in terms of completeness, with variations in the filling rates over the years[16].

In the state of São Paulo, a study on the completeness trends of medical records of elderly women with breast cancer showed that only the variable "family history of breast cancer" followed a decreasing trend of incompleteness, leading to an improvement in records over the years[17]. However, variables such as race/color, years of education, use of oral contraceptives, duration of oral contraceptive use, hormone replacement therapy, and breastfeeding showed worsening trends in records, with increasing incompleteness[17].

In the United States, a study[18] investigated racial and ethnic disparities in breast cancer survival by analyzing how tumor characteristics, sociodemographic factors, and treatment could mediate these disparities. The research found an increased risk of death from breast cancer for Black women compared to White women, especially when considering tumor subtype. In age-adjusted models, among women with estrogen receptor-positive tumors, Black women were three times more likely to die from breast cancer than White women in the first two years after diagnosis[18].

In a systematic review and meta-analysis[19] of 20 studies involving 14,103 Black women and 76,111 White women, it was found that Black women had a 19% higher likelihood of dying from breast cancer compared to White women, after adjusting for age, stage, and socioeconomic status[19]. The improvement in the quality of demographic data records contributes to the identification of risk factors for the development of breast cancer. A study conducted in São Paulo, Brazil, highlighted the importance of identifying risk factors and genetic polymorphisms for population screening and early breast cancer diagnosis, emphasizing that carcinogenesis is multifactorial and depends on specific tumor characteristics[20].

Regarding the TNM variable, our study observed a "poor" score throughout most of the observed period. Similarly, other studies also reported "poor" quality in this variable[16,21]. Na In research conducted in Cuiabá, Mato Grosso, the completeness quality was classified as "very poor" for the TNM variable and "poor" for staging[16]. In another study conducted nationwide in Brazil[22], a low percentage of inconsistency was found in staging for five cancers, including breast cancer. However, for the tumor laterality variable, the same nationwide study found a very high error rate in data filling, which contrasts with our study, where the completeness quality for this variable was "excellent" throughout almost the entire evaluated period.

Detecting significant trends in sociodemographic and clinical variables is essential for creating strategies for the prevention, diagnosis, treatment, and follow-up of cancer patients[23,24]. Identifying increases or decreases in variables related to risk factors, alcoholism, and smoking indicates the need to enhance the quality of HCRs. Another notable variable is family history of cancer, where individuals with a family history of cancer, particularly in first-degree relatives, are at higher risk of developing certain types of cancer due to genetic and environmental factors that may be shared within families[25-27]. The quality of cancer hospital records is crucial for reliable epidemiological analysis and for creating effective health policies[28,29]. Identifying variables with missing data or inconsistencies highlights the ongoing importance of improving data collection and monitoring.

The study has some limitations, such as being restricted to data from the HCR of a single Brazilian state, which limits the generalization of the findings to other states. Therefore, caution should be exercised when interpreting the external validity of the results. While HCRs provide valuable information about the quality of healthcare services, they do not fully represent the national epidemiology of cancer.

CONCLUSION

Based on the analysis of variables from the Hospital Cancer Registry (HCR) related to malignant breast cancer cases in the Oncological Care Network of Espírito Santo, a significant trend of incompleteness was observed in essential variables for monitoring and planning oncological care. This finding highlights the need to improve the quality of records and enhance the training of the teams responsible for data collection, aiming to obtain more robust and complete data. Improving the records can support more effective public policies and provide better targeted and more appropriate care for breast cancer patients.

Author Contributions

All authors contributed to the manuscript. Literature review: RMP, LCLJ. Data collection: RMP, LCBSN, LSDA, CSDA. Study design: RMP, LCLJ. Data analysis: WRG, LCLJ. Manuscript preparation: LCBSN, LSDA, CSDA. Manuscript review: LCLJ.

Funding

This research was funded by the Fundação de Amparo à Pesquisa e Inovação do Espírito Santo — FAPES, Public Notice FAPES/CNPq/Decit-SCTIE-MS/SESA Nº 09/2020 — PPSUS. Process Number: 2021-F0436.

Acknowledgments

We would like to thank the Cancer Surveillance Department of the State Health Secretariat of Espírito Santo (SESA/ES) for supporting this study.

Conflicts of Interest

The authors declare no conflicts of interest.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021 May;71(3):209-249. doi: 10.3322/caac.21660. Epub 2021 Feb 4. PMID: 33538338.
2. Sardinha L, Maheu-Giroux M, Stöckl H, Meyer SR, García-Moreno C. Global, regional, and national prevalence estimates of physical or sexual, or both, intimate partner violence against women in 2018. Lancet. 2022 Feb 26;399(10327):803-813. doi: 10.1016/S0140-6736(21)02664-7. Epub 2022 Feb 16. PMID: 35182472; PMCID: PMC8885817.
3. Omran AR. The epidemiologic transition. A theory of the epidemiology of population change. Milbank Mem Fund Q. 1971 Oct;49(4):509-38. PMID: 5155251.
4. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018 Nov;68(6):394-424. doi: 10.3322/caac.21492. Epub 2018 Sep 12. Erratum in: CA Cancer J Clin. 2020 Jul;70(4):313. PMID: 30207593.
5. Santos M de O, Lima FC da S de, Martins LFL, Oliveira JFP, Almeida LM de Cancela M de C. Estimation of Cancer Incidence in Brazil, 2023-2025. Rev. Bras. Cancerol. [Internet]. February 6, 2023 [cited August 21, 2023];69(1):e-213700. Available from: https://rbc.inca.gov.br/index.php/revista/article/view/3700
6. National Cancer Institute José Alencar Gomes da Silva. Atlas of Mortality. Rio de Janeiro: INCA; 2022. Database. Accessed on: August 21. Available from:https://www.inca.gov.br/app/mortalidade

7. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, Znaor A, Bray F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer. 2019 Apr 15;144(8):1941-1953. doi: 10.1002/ijc.31937.
8. National Cancer Institute José Alencar Gomes da Silva. Cancer Hospital Records: Planning and Management / National Cancer Institute, 2nd ed.; INCA; Rio de Janeiro; 2010 [accessed on August 14, 2023]. Available from: https://www.inca.gov.br/sites/ufu.sti.inca.local/files/media/document/registros-hospitalares-de-cancer-2010.pdf
9. Lopes-Júnior LC, Dell'Antonio LS, Pessanha RM, Dell'Antonio CS, da Silva MI, de Souza TM, Grassi J. Completeness and Consistency of Epidemiological Variables from Hospital-Based Cancer Registries in a Brazilian State. Int J Environ Res Public Health. 2022 Sep 22;19(19):12003. doi: 10.3390/ijerph191912003. PMID: 36231303; PMCID: PMC9566550.
10. Grippa WR, Dell'Antonio LS, Salaroli LB, Lopes-Júnior LC. Incompleteness trends of epidemiological variables in a Brazilian high complexity cancer registry: An ecological time series study. Medicine (Baltimore). 2023 Aug 4;102(31):e34369. doi: 10.1097/MD.0000000000034369.
11. Santos TB dos, Borges AK da M, Ferreira JD, Meira KC, Souza MC de, Guimarães RM, et al. Prevalence and factors associated to advanced stage breast cancer diagnosis. Ciênc Saúde Coletiva [Internet]. 2022 Feb;27(2):471–82. Available from: https://doi.org/10.1590/1413-81232022272.36462020
12. Brazil. National Cancer Institute. Cancer Hospital Records: Planning and Management / National Cancer Institute, 2nd ed. INCA: Rio de Janeiro, Brazil, 2010; 536 p.
13. Lima CR, Schramm JM, Coeli CM, da Silva ME. Review of data quality dimensions and applied methods in the evaluation of health information systems. Cad Saude Publica. 2009 Oct;25(10):2095-109. doi: 10.1590/s0102-311x2009001000002.
14. Romero DE, Cunha CB. Quality of socioeconomic and demographic data in relation to infant mortality in the Brazilian Mortality Information System (1996/2001). Cad Saude Publica. 2006 Mar;22(3):673-84. doi: 10.1590/s0102-311x2006000300022.
15. Mann HB. Nonparametric tests against trend. Econometrica. 1945;13:245–59.
16. Oliveira JCS, Azevedo EFS, Caló RS, et al. Hospital-based cancer registries of Mato Grosso, Brazil: analysis of completeness and consistency. Cad Saude Colet. 2021;29:330–43.
17. Brandão-Souza C, Amorim MHC, Zandonade E, Fustinoni SM, Schirmer J. Completeness of medical records of elderly women with breast cancer: a trend study. Acta Paul Enferm [Internet]. 2019Jul;32(4):416–24. Available from: https://doi.org/10.1590/1982-0194201900057
18. Warner ET, Tamimi RM, Hughes ME, Ottesen RA, Wong YN, Edge SB, Theriault RL, Blayney DW, Niland JC, Winer EP, Weeks JC, Partridge AH. Racial and Ethnic Differences in Breast Cancer Survival: Mediating Effect of Tumor Characteristics and Sociodemographic and Treatment Factors. J Clin Oncol. 2015 Jul 10;33(20):2254-61. doi: 10.1200/JCO.2014.57.1349.
19. Newman LA, Griffith KA, Jatoi I, Simon MS, Crowe JP, Colditz GA. Meta-analysis of survival in African American and white American patients with breast cancer: ethnicity compared with socioeconomic status. J Clin Oncol. 2006 Mar 20;24(9):1342-9. doi: 10.1200/JCO.2005.03.3472.
20. Pereira RGV, Souto RP, Longo PL, Fernandes CE, Filho IC, Felizi RT, Veiga MG, Oliveira E. Evaluation of the g473a polymorphism in the lysyl oxidase gene as a risk factor related to the occurrence of breast cancer in Brazilian women. J Hum Growth Dev. 2022; 32(2):242-247. DOI: http://doi.org/10.36311/jhgd.v32.13317
21. Pettersson A, Robinson D, Garmo H, Holmberg L, Stattin P. Age at diagnosis and prostate cancer treatment and prognosis: a population-based cohort study. Ann Oncol. 2018 Feb 1;29(2):377-385. doi: 10.1093/annonc/mdx742. PMID: 29161337.
22. Pinto IV, Ramos DN, Costa MCE, et al. Completeness and consistency of data in hospital-based cancer registries in Brazil. Cad Saude Colet. 2012;20:113–20.
23. D'Alessandro TAL, Antoniazzi BN, Abreu DMX. Hospital-based cancer registries of Minas Gerais, Brazil: analysis of the consistency of databases. Cad Saude Colet. 2010;18(3):410-7.
24. Velasco Yanez RJ, Carvalho Fernandes AF, de Freitas Corpes E, Moura Barbosa Castro RC, Sixsmith J, Lopes-Júnior LC. Palliative care in the treatment of women with breast cancer: A scoping review. Palliat Support Care. 2023 Dec 7:1-18. doi: 10.1017/S1478951523001840.
25. Alcantara SSA, Martinelli PM, Sousa LVA, Fonseca FLA. Epidemiological Profile of Prostate Cancer Mortality and Access to Hospital Care In Brazilian Regions - An Ecological Study.J Hum Growth Dev. 2021; 31(2):310-317. DOI: 10.36311/jhgd.v31.12227

26. Grippa WR, Pessanha RM, Dell'Antonio LS, Dell'Antonio CSDS, Salaroli LB, Lopes-Júnior LC. Completeness of variables in Hospital-Based Cancer Registries for prostatic malignant neoplasm. Rev Bras Enferm. 2024 Jul 29;77(3):e20230467. doi: 10.1590/0034-7167-2023-0467.
27. Pessanha RM, Grippa WR, Lopes-Júnior LC. Hospital-based registries of women with breast cancer from the oncology network of Espírito Santo, Brazil. Rev Enferm UFPI . 2024;13(1). Available from: https://periodicos.ufpi.br/index.php/reufpi/article/view/5562
28. Lima RAG, Lopes-Júnior LC, Maia EBS, Fuentes-Alabi S, Ponce MLV. Global Initiative for Childhood Cancer Control: Increasing access, improving quality, saving lives. Rev Lat Am Enfermagem. 2023 Oct 6;31:e3998. doi: 10.1590/1518-8345.0000.
29. Grassi J, Pessanha RM, Grippa WR, Dell'Antonio LS, Dell'Antonio CSDS, Faure L, Clavel J, Lopes-Júnior LC. Temporal Trends in the Completeness of Epidemiological Variables in a Hospital-Based Cancer Registry of a Pediatric Oncology Center in Brazil. Int J Environ Res Public Health. 2024 Feb 9;21(2):200. doi: 10.3390/ijerph21020200.

Resumo

Introdução: a Organização Mundial da Saúde classifica as Doenças Crônicas Não Transmissíveis como a principal causa de morte global, com 71% das mortes associadas a neoplasias malignas. Em 2020, ocorreram cerca de 19,3 milhões de novos casos de câncer e 10 milhões de mortes no mundo, sendo as Américas responsáveis por 20,9% da incidência global. No Brasil, o INCA estima 704 mil novos casos de câncer para 2023-2025, com destaque para os cânceres de mama e próstata, que representam 15% dos novos casos cada. A idade acima de 50 anos, fatores genéticos, histórico familiar e outros são fatores de risco para o câncer de mama.

Objetivo: avaliar a tendência de incompletude das variáveis do Registro Hospitalar de Câncer de casos de neoplasia maligna da mama de todos os hospitais que compõem a Rede de Atenção Oncológica do Espírito Santo.

Método: estudo retrospectivo de séries temporais utilizando-se dados secundários a partir da Ficha de Registro do tumor dos RHC da Rede de Atenção Oncológica do Espírito Santo entre 2000 e 2020. A incompletude dos dados foi classificada como excelente (<5%), boa (entre 5%-10%), regular (>10%<20%), ruim (>20%<50%) e muito ruim (>50%), de acordo com o percentual de ausência de informação. O teste de Mann–Kendall foi utilizado para avaliar as tendências temporais entre os anos, e o teste de Friedman para avaliar as classificações de qualidade da série histórica.

Resultados: foram registrados 16.587 casos de câncer de mama nos Registros Hospitalares de Câncer do Espírito Santo entre 2000 e 2020. As variáveis alcoolismo (p<0,001), tabagismo (p<0,001), história familiar de câncer (p<0,001), estado conjugal (p<0,001), tiveram uma incompletude com tendência de decréscimo estatisticamente significantes. Já as variáveis relacionadas ao diagnóstico/tratamento e características do tumor apresentaram uma tendência de incompletude menor ao longo do tempo, apresentando um acréscimo, tais como, tipo de exame diagnóstico realizado (p=0,03), diagnósticos e tratamentos anteriores (p=0,008), bases mais importantes para o diagnóstico do tumor (p<0,001), lateralidade do tumor (p=0,008), presença de mais de um tumor primário (p=0,01).

Conclusão: os escores das variáveis epidemiológicas dos Registros Hospitalares de Câncer de mulheres com câncer de mama no Espírito Santo tiveram em sua maioria completude classificada como "excelente". Porém, a variáveis de importância clínica como estadiamento TNM apresentou um decréscimo na completude durante a maior parte da série histórica.

Palavras-chave: câncer de mama, oncologia, registro hospitalar de câncer, epidemiologia, vigilância em saúde pública.