Quality of data from Hospital Cancer Registries: An analysis of Registered Cancer Cases in Brazil between 2000 and 2020

https://doi.org/10.32635/2176-9745.RBC.2024v70n1.4568

Qualidade dos dados dos Registros Hospitalares de Câncer: Uma Análise dos Casos Cadastrados de Câncer no Brasil entre 2000 e 2020

Calidad de los Datos de los Registros Hospitalarios de Cáncer: Un análisis de los Casos de Cáncer Registrados en el Brasil entre 2000 y 2020

Édnei César de Arruda Santos Junior[1]; Gulnar Azevedo e Silva[2]; Natalia Santana Paiva[3]

ABSTRACT

Introdução: Cancer has a significant impact on the Brazilian population, with a high incidence and mortality. Hospital Cancer Registries (HCR) are essential sources of information to evaluate cancer care provided by SUS, making it possible to conduct clinical and epidemiological researches and define strategic axes to plan health policies, cancer surveillance and control actions in the country. **Objective:** To evaluate the completeness and inconsistency of the HCR database in Brazil from 2000 to 2020. **Method:** Descriptive study to assess the quality of data of HCR in Brazil. The public available data were collected on *IntegradorRHC* at the National Cancer Institute. The Romero & Cunha score was utilized to classify the quality of the data according to completeness and inconsistency. **Results:** The worst completeness was observed for the following variables: occupation, Pathological Tumor-Node-Metastasis (pTNM) and disease status at the end of the first treatment in all studied years. With regard to inconsistencies, the combinations were substantially zero during the study period. **Conclusion:** The quality of the data of the *Sistema de Informação de Registros Hospitalares de Câncer* (SisRHC) has improved over the years, despite low completeness in some variables. The use of these data should be encouraged and can provide material for cancer surveillance and control in Brazil.

Key words: Data Accuracy; Neoplasms/epidemiology; Hospital Records; Public Health; Health Information Systems.

RESUMO

Introduction: O câncer tem um impacto significativo na população brasileira, apresentando elevada incidência e mortalidade no país. Os Registros Hospitalares de Câncer (RHC) são fontes de informações essenciais para a avaliação da assistência oncológica no SUS, possibilitando a condução de pesquisas clínicas-epidemiológicas e a definição de eixos estratégicos para o planejamento das políticas de saúde e ações de vigilância e controle do câncer no país. **Objetivo:** Avaliar a completude e inconsistência da base de RHC no Brasil, de 2000 a 2020. **Método:** Estudo descritivo de avaliação da qualidade dos dados de RHC no Brasil. Os dados foram coletados no IntegradorRHC e estão disponibilizados de forma pública pelo Instituto Nacional de Câncer. Utilizou-se o escore de Romero & Cunha para a classificação da qualidade dos dados segundo a completude e inconsistência. **Resultados:** As piores completudes foram observadas para as variáveis: ocupação, *Pathological Tumor-Node-Metastasis* (PTNM) e estado da doença ao final do primeiro tratamento em todos os anos de estudo. Em relação às inconsistências, as combinações foram substancialmente zero no período de estudo. **Conclusão:** A qualidade dos dados do Sistema de Informação de Registros Hospitalares de Câncer (SisRHC), apesar de baixo preenchimento em algumas variáveis, apresenta uma melhoria ao decorrer dos anos. A utilização desses dados deve ser estimulada e pode oferecer subsídios para a vigilância e controle do câncer no Brasil..

Palavras-chave: Confiabilidade dos Dados. Neoplasias/epidemiologia; Registros Hospitalares; Saúde pública; Sistemas de Informação em Saúde.

RESUMEN

Introducción: El cáncer tiene un impacto significativo en la población brasileña, con una alta incidencia y mortalidad en el país. Los Registros Hospitalarios de Cáncer (RHC) son fuentes esenciales de información para la evaluación de la atención del cáncer en el SUS, posibilitando la investigación clínica y epidemiológica y la definición de ejes estratégicos para la planificación de políticas de salud y acciones de vigilancia y control del cáncer en el país. **Objetivo:** Evaluar la completitud e inconsistencia de la base de datos del RHC en el Brasil, de 2000 a 2020. **Método:** Estudio descriptivo para evaluar la calidad de los datos de los RHC en el Brasil. Los datos fueron recogidos de IntegradorRHC y están puestos a disposición del público por el Instituto Nacional del Cáncer. Se utilizó la puntuación de Romero & Cunha para clasificar la calidad de los datos según su integridad e inconsistencia. **Resultados:** La peor completitud se observó para las siguientes variables: ocupación, *Pathological Tumor-Node-Metastasis (pTNM)* y estado de la enfermedad al final del primer tratamiento en todos los años del estudio. En cuanto a las incoherencias, las combinaciones fueron prácticamente nulas durante el periodo de estudio. **Conclusión:** La calidad de los datos del Sistema de Información de Registros Hospitalarios de Cáncer (SisRHC) ha mejorado con los años, a pesar de la baja completitud en algunas variables. El uso de estos datos debe ser estimulado y puede proporcionar ayudas para la vigilancia y control del cáncer en el Brasil.

Palabras clave: Exactitud de los Datos; Neoplasias/epidemiología. Registros Hospitalarios; Salud Pública; Sistemas de información en salud.

[1,2]Universidade do Estado do Rio de Janeiro (Uerj), Instituto de Medicina Social Hesio Cordeiro (IMS). Rio de Janeiro (RJ), Brasil. E-mails: edneicesar2@gmail.com; gulnar@ims.uerj.br. Orcid iD: https://orcid.org/0000-0002-6956-6139; Orcid iD: https://orcid.org/0000-0001-8734-2799
[3]Universidade Federal do Rio de Janeiro (UFRJ), Instituto de Estudos em Saúde Coletiva (Iesc). Rio de Janeiro (RJ), Brasil. E-mail: nataliapaiva@iesc.ufrj.br Orcid iD: https://orcid.org/0000-0003-0541-4686
Corresponding author: Édnei César de Arruda Santos Junior. IMS/Uerj. Rio de Janeiro (RJ), Brasil. CEP 20950-000. E-mail: edneicesar2@gmail.com

INTRODUCTION

Cancer has high incidence and mortality in Brazil, demanding great endeavor of public health authorities to control[1]. Both incidence and mortality are increasing significantly due to the sociodemographic and epidemiologic populational transformations and changes of the main risk factors[1,2].

The data produced by the Hospital Cancer Registries (HCR) are processed in the Hospital Cancer Registry Information System (SisRHC) and feed the *IntegradorRHC* which consolidates and offers hospital cancer registries from Brazil's HCRs[3].

HCR are essential to monitor cancer status in Brazil, which, in addition to evaluating the assistance provided by the cancer network, help the epidemiologic research, clinic and planning[4,5]. Therefore, the data produced contribute to cancer surveillance and creation of control actions in the country. However, many health surveillance systems have quality related issues with incomplete or insufficient data for correct populational monitoring[6,7].

The quality of health information depends on the analysis of the reliability of data to meet the objective it was created for. Therefore, the health information systems corroborate the prevention, promotion, planning and monitoring of harms to the population[8]. Its functioning and improvement of the population's health conditions are influenced by the flow of reliable information to the population with clear association with federal, state and municipal management[9-11].

Due to the importance of a public health information system, the evaluation of the quality of the data help to identify the fragilities and potentialities of the information produced within the scope of public health policies, identification of vulnerable individuals and impacts of the disease on the population and where they live[12,13].

The Centers for Disease Control and Prevention – CDC determines that the evaluation of a surveillance system comprehends quantitative and qualitative attributes, among them, the quality of the data is one of the most recommended for analysis, mainly regarding its reliability, validity, coverage and completeness[11,14].

The objective of this study is to evaluate the completeness and inconsistency of Brazil's HCR base between 2000 and 2020.

METHOD

Descriptive study of the evaluation of the quality of the data of HCR in Brazil which were extracted on July 29, 2023 from *IntegradorRHC*, a public deidentified database waiving an Institutional Review Board approval in compliance with Ordinance 510/2016[15] of the National Health Council (CNS). *IntegradorRHC* consolidates all the cases of patients with confirmed diagnosis of cancer who have been treated at a hospital approved by Brazil's National Health System (SUS).

All the reported analytical cases and without relapse were selected with the 1st consultation between 1st January 2000 and December 31st 2020.

The study variables were those whose completion was mandatory at the database, a standard procedure for all HCR in Brazil.

After the data were collected, two indicators – completeness and inconsistency – for evaluation of the quality of the data of *IntegradorRHC* were calculated in line with the Centers for Disease and Control – CDC guidelines for evaluation of Public Health Surveillance Systems[14].

COMPLETENESS

The calculation was performed from the proportion of complete data filled, subtracting the proportion of data filled as ignored or blank, i.e., 100% less the proportion of incompleteness for each year investigated. The score proposed by Romero & Cunha[16] was utilized to evaluate the completeness to classify the variables selected (Chart 1).

Chart 1. Score of Romero & Cunha matching completeness, inconsistency and classification

Romero & Cunha Score		
Proportion of incompleteness	**Matching completeness and inconsistency**	**Classification**
< 5%	100% to 96%	Excellent (E)
5% to 10%	95% to 90%	Good (G)
10% to 20%	90% to 80%	Regular (R)
20% to 50%	80% to 50%	Poor (P)
50%+	50% to 0%	Very poor (VP)

Fonte: Romero & Cunha[16].

INCONSISTENCY

The proportion of registries with inconsistencies (%) was calculated considering the number of time ranges of the dates recorded which had some inconsistency for all registries and multiplied by 100 for each year investigated. The score of Romero & Cunha[16] (Chart 2) was applied to classify the inconsistencies. Time ranges considered

for the analysis were: Date of the first visit < Date of the diagnosis in 12 months; Date of the first visit > Date of beginning of the treatment after 12 months; Date of the diagnosis > Date of beginning of the treatment after 12 months; Date of the first visit > Date of death after 12 months; Date of the diagnosis > Date of death after 12 months and Date of the beginning of the treatment after 12 months (Chart 2).

Considering the possibilities of patients referred without diagnosis, diagnosis after surgery and post-death diagnosis, only cases where the differences between the dates were higher or lower than 12 months were considered inconsistent, a period considered satisfactory for the diagnosis with more odds of being registrational errors depending on the inconsistency. The cutoff of 12 months was defined according to the study of Pinto et al.[17].

Chart 2. Variables selected and conditions to definition of inconsistency

Time range considered as inconsistency of the registries
Date of the first consultation < Date of the diagnosis
Date of the first consultation > Date of beginning of the treatment
Date of the diagnosis > Date of beginning of the treatment
Date of the first consultation > Date of death
Date of the diagnosis > Date of death
Date of beginning of the treatment > Date of death

The software R, version 4.3.0[18] was utilized for the analyses.

RESULTS

Between 2000 and 2020, 3,435,126 cases of cancer on *IntegradorRHC* were registered, a growth from 55,573 in 2000 to 172,889 in 2020, a percent increase of 210% based on cases of cancer registered at HCR in 20 years (Table 1).

The worst completeness were found for the variables occupation, PTNM[19] (Pathological Tumor-Node-Metastasis) and stage of the disease at the end of the first treatment for all years of the registries of the tumor.

For the variables of identification of the patient, the completeness was poor and very poor for race/skin color, education and occupation with completeness of 43.3%, 67.5% and 41.6%, respectively in 2000 and 67.2%, 77.1% and 54.1% in 2020 (Table 1).

The completeness of the characterization of the diagnosis was excellent for almost all of the variables, except the most important, the base, classified as poor, but it improved during the study, reaching completeness of 72.2% in 2020 (Table 1).

The variables of clinical staging of the tumor TNM[19] and PTNM[19] on the block of characterization of the tumor were classified as poor during the period. Completeness of tumor staging reached 80.7% in 2000 and 73.8% in 2019 and for the variable PTNM, completeness was 57.9% in 2000 and 51.3% in 2019 (Table 1), a drop in filling this variable.

For the characterization of the first treatment, the completeness of the variable stage of the disease at the end of the first treatment fluctuated between poor and very poor, with improvement in filling during the follow-up period, from very poor from 2000 (37.2%) to 2006 (49.2%) to poor from 2007 (51.6%) to 2020 (53.91%) (Table 1).

The combinations among the times of the variables were substantially zero during the study, showing that the inconsistencies of the dates on SisRHC are low, classified as excellent according to Romero & Cunha[16] (Table 2).

DISCUSSION

Improvement of completeness of the variables on the time span was found by the current investigation. In 2020, the completeness of the clinical variables increased in comparison with the last three years. Due to reduction of procedures of screening, diagnosis and treatment, except chemotherapy for cancer in 2020 when compared to 2019[20], the increase can be attributed to the COVID-19 pandemic which reduced the number of patients on HRC and better qualification of the tumors was possible.

The low volume of data of clinical variables filled is associated with failure of the health professional because as soon as the disease is registered at the HCR and its treatment begins, its clinical staging is known nd consequently a TNM and PTNM of the tumor.

The wrong filling of clinical variables was noticed in similar publications about the evaluation of SisRHC[12,17,21-25]. The study of D'Alessandro and Antoniazzi[12] revealed worst completeness for TNM and PTNM combined (35.0%) and staging (26.0%). Pinto et al.17 noticed poor completeness for the variables staging (29.7%) and TNM (49.8%).

In the study of Oliveira et al.[21], the completeness was very poor for TNM (77.2%) and poor for staging (42.2%). For Lopes-Júnior et al.[22] the completeness was poor for clinical staging of the tumor (45.7%) and very poor for TNM (67.8%). The mean completeness from 2017 and 2019 was poor (39.5%) for clinical

Chart 1. Proportion of fields filled and completeness of the variables of *IntegradorRHC* for all types of cancer, Brazil 2000-2020

Variables	Cancer*																					
	2000		2001		2002		2003		2004		2005		2006		2007		2008		2009		2010	
Identification of the patient	%		%		%		%		%		%		%		%		%		%		%	
Age	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E
Sex	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E
Race/Skin color	43.3	VP	50.5	P	53.8	P	52.2	P	54.2	P	55.1	P	56.6	P	60.5	P	61.0	P	61.4	P	63.7	P
Education	67.5	P	71.2	P	69.1	P	70.7	P	71.5	P	69.6	P	67.8	P	69.2	P	68.6	P	68.9	P	70.5	P
Occupation	41.6	VP	46.2	VP	47.8	VP	49.5	VP	53.9	P	56.2	P	55.0	P	59.7	P	57.7	P	56.5	P	55.4	P
State of origin	99.4	E	99.6	E	99.5	E	99.6	E	99.7	E	99.8	E	99.8	E	99.8	E	99.7	E	99.9	E	99.7	E
Characterization of the diagnosis	%		%		%		%		%		%		%		%		%		%		%	
Date of the 1st visit	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E
Date of the 1st diagnosis	99.4	E	98.8	E	98.8	E	98.9	E	99.2	E	99.2	E	99.2	E	98.6	E	98.6	E	98.2	E	97.8	E
Former diagnosis and treatment	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E
More important base for the diagnosis	48.5	VP	55.4	P	58.5	P	58.2	P	60.2	P	62.8	P	65.6	P	68.6	P	69.0	P	68.8	P	69.8	P
Characterization of the tumor	%		%		%		%		%		%		%		%		%		%		%	
Histology type	99.9	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	99.8	E	100.0	E	100.0	E
Clinical staging of the tumor (TNM)	80.7	R	79.7	P	79.4	P	79.1	P	78.6	P	79.9	P	78.9	P	79.2	P	78.2	P	78.9	P	77.2	P
TNM	73.6	P	72.3	P	67.9	P	68.7	P	69.5	P	71.4	P	69.1	P	70.4	P	70.1	P	71.1	P	68.7	P
PTNM	57.9	P	56.2	P	50.4	P	53.4	P	52.1	P	54.5	P	55.2	P	55.4	P	50.5	P	49.9	VP	50.5	P
Characterization of the 1st treatment	%		%		%		%		%		%		%		%		%		%		%	
Date of beginning of the 1st treatment	93.1	G	92.6	G	92.7	G	94.3	G	94.0	G	94.0	G	93.4	G	93.1	G	92.9	G	93.0	G	92.9	G
Main reason to not treat	98.0	E	98.1	E	98.1	E	98.6	E	98.4	E	97.2	E	91.7	G	95.4	G	91.9	G	90.2	G	85.6	R
1st treatment received	99.3	E	99.7	E	99.9	E	99.9	E	99.9	E	99.9	E	99.7	E	99.2	E	99.1	E	98.9	E	99.0	E
Stage of the disease at the end of the 1st treatment	37.2	VP	44.8	VP	43.5	VP	43.2	VP	45.1	VP	48.3	VP	49.2	VP	51.6	P	50.1	P	50.2	P	52.8	P
Date of Death	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E

. to be continued

Chart 1. continuation

Variables	Cancer*																			
	2011		2012		2013		2014		2015		2016		2017		2018		2019		2020	
Identification of the patient	%		%		%		%		%		%		%		%		%		%	
Age	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	99,9	E	100.0	E	100.0	E	100.0	E
Sex	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E
Race/Skin color	62.2	P	62.2	P	63.2	P	63.9	P	65.8	P	66.6	P	67.1	P	67.4	P	67.1	P	67.2	P
Education	72.0	P	75.2	P	76.2	P	76.5	P	76.2	P	77.4	P	77.3	P	77.6	P	76.1	P	77.1	P
Occupation	55.4	P	54.6	P	55.1	P	53.9	P	53.6	P	54.6	P	55.2	P	55.6	P	55.0	P	54.1	P
State of origin	99.5	E	99.2	E	99.3	E	99.4	E	99.3	E	99.6	E	98.9	E	99.5	E	99.4	E	99.6	E
Characterization of the diagnosis	%		%		%		%		%		%		%		%		%		%	
Date of the 1st visit	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E
Date of the 1st diagnosis	97.9	E	98.3	E	98.4	E	98.5	E	98.1	E	98.3	E	98.3	E	98.5	E	98.6	E	98.3	E
Former diagnosis and treatment	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E
More important base for the diagnosis	68.8	P	68.6	P	69.0	P	69.3	P	70.5	P	72.0	P	73.2	P	73.2	P	73.3	P	72.2	P
Characterization of the tumor	%		%		%		%		%		%		%		%		%		%	
Histology type	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E
Clinical staging of the tumor (TNM)	77.3	P	76.2	P	76.5	P	75.2	P	74.3	P	73.9	P	73.2	P	73.9	P	73.8	P	75.2	P
TNM	68.2	P	67.6	P	68.3	P	67.8	P	66.7	P	66.5	P	66.0	P	66.9	P	66.8	P	68.5	P
PTNM	49.8	VP	47.9	VP	49.3	VP	49.6	VP	50.5	P	51.9	P	52.2	P	51.8	P	51.3	P	52.2	P
Characterization of the 1st treatment	%		%		%		%		%		%		%		%		%		%	
Date of beginning of the 1st treatment	93.3	G	93.2	G	93.4	G	93.5	G	93.5	G	93.5	G	93.4	G	92.7	G	92.1	G	91.3	G
Main reason to not treat	94.6	G	97.2	E	97.9	E	97.9	E	97.7	E	97.7	E	97.3	E	97.0	E	97.1	E	97.1	E
1st treatment received	99.1	E	99.0	E	99.0	E	99.0	E	99.3	E	99.2	E	99.2	E	98.9	E	99.1	E	99.0	E
Stage of the disease at the end of the 1st treatment	52.4	P	51.8	P	51.0	P	52.1	P	53.0	P	53.9	P	54.3	P	55.0	P	55.0	P	53.9	P
Date of Death	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E	100.0	E

Captions: Classification: E = excellent; G = good; R = regular; P = poor; VP = very poor.

*All types of cancer

Chart 2. Proportion (%) of the inconsistency in variables combined of IntregradorRHC for all types of cancer, Brazil 2000-2020.

Inconsistencies %	Cancer*																					
	2000		2001		2002		2003		2004		2005		2006		2007		2008		2009		2010	
Date of the first consultation < Date of the diagnosis																						
	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E
Date of the first consultation > Date of beginning of treatment																						
	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E
Date of the diagnosis > Date of beginning of treatment																						
	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E
Date of the first consultation > Date of death																						
	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0	-	0	-
Date of the diagnosis > Date of death																						
	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0	-
Date of beginning of treatment > Date of death																						
	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0	-

. to be continued

Chart 2. continuation.

Inconsistencies %	Cancer*																					
	2011		2012		2013		2014		2015		2016		2017		2018		2019		2020			
Date of the first consultation < Date of the diagnosis																						
	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E		
Date of the first consultation > Date of beginning of treatment																						
	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E		
Date of the diagnosis > Date of beginning of treatment																						
	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E	0.0	E		
DDate of the first consultation > Date of death																						
	0	-	0	-	0	-	0	-	0	-	0	-	0	-	0.0	E	0	-	0	-		
Date of the diagnosis > Date of death																						
	0	-	0	-	0	-	0	-	0	-	0	-	0	-	0	-	0	-	0	-		
Date of beginning of treatment > Date of death																						
	0	-	0	-	0	-	0	-	0	-	0	-	0	-	0.0	E	0	-	0	-		

Legendas: Classification: E = excellent; G = good; R = regular; P = poor; VP = very poor.

*All types of cancer.

staging according to Keske[23], and very poor (62.9%) for TNM. Cardoso[24] concluded very poor completeness for the variables TNM (70.0%), PTNM (58.0%) and clinical staging (52.0%). Oliveira et al.[25] noticed missed information for TNM staging (36.6%).

The filling of the variable education is poor, however, it improved along the period, which is not restricted to HCR because national studies evaluating the Mortality Information System (SIM) and the National Diseases Reporting System (SINAN) reached similar results[26-30].

In view of the social inequities to access SUS, poor completeness of the variable race/skin color brings up a chronic problem of the health system. Race/skin color is a social marker of inequity[31] and whether the variable is missing, specific public policies to mitigate this scenario of accessing the health system will be difficult to implement.

The variable occupation had very poor and poor completeness along the years, corroborating what similar studies have concluded as well[22,24,32]. Low completeness of this variable creates barriers to achieve health surveillance actions targeted to occupational cancers affecting the workers, making difficult to obtain a profile of the professional occupations with high predominance of cancer.

The completeness of the variable stage of the disease at the end of the first treatment improved along the years, however, data filling is unsatisfactory yet. Former studies report low filling of this variable[17,21,23,25,33]. It is important to count with this variable to evaluate the treatments applied and standardize procedures and experience sharing among High Complexity Oncologic Hospitals (Unacon) and High Complexity Oncologic Assistance Centers (Cacon).

Low inconsistencies among time spans might be related to better qualification to fill in the dates to perform cancer surveillance, since this information can support several studies.

CONCLUSION

The quality of the data of Brazil's SisRHC, despite poor filling of some variables, improved along the years. If filling the variables rises, cancer surveillance and control expands as well from early diagnosis through hospital treatment.

Missing information in some fields may be associated with organizational issues at the HCR: flawed typing, difficulty to understand the system and inexistence of a flow of communication among the clinical body and those in charge of the registries. Filling the registration file of the tumor is based on the patient's chart, therefore, the missing information on the chart directly impacts the lack of information on the registration file of the tumor.

The limitations of this study consist in the impossibility to evaluate the trend of completeness and inconsistency and the identification of duplicity of registers. It was not possible to detect cases of individuals with more than one chart or when the name had been wrongly typed. In addition, further studies to evaluate the trend of completeness and inconsistency of the base of SisRHC in Brazil are encouraged.

Despite the limitations, the results found in the evaluation of the data of SisRHC show good filling and low inconsistency of the mandatory variables of the registration file of the tumor. Their utilization should be encouraged and improvement of the cancer information system is essential to ensure effectiveness of oncologic care in SUS services network.

CONTRIBUTIONS

All the authors contributed substantially to the study design, acquisition, analysis and interpretation of the data, wording and critical review. They approved the final version to be published.

DECLARATION OF CONFLICT OF INTERESTS

There is no conflict of interests to declare.

FUNDING SOURCES

Édnei César de Arruda Santos Junior was awarded funds from CAPES – Coordination for the Improvement of Higher Education Personnel, Code 001.

REFERENCES

1. Instituto Nacional de Câncer. Estimativa 2023: incidência de câncer no Brasil [Internet]. Rio de Janeiro: INCA; 2023. [acesso 2024 mar 17]. Disponível em: https://www.inca.gov.br/sites/ufu.sti.inca.local/files//media/document//estimativa-2023.pdf
2. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA A Cancer J Clin. 2021;71(3):209-49. doi: https://doi.org/10.3322/caac.21660
3. Instituto Nacional de Câncer José Alencar Gomes da Silva. IntegradorRHC: ferramenta para a vigilância hospitalar de câncer no Brasil [Internet]. Rio de Janeiro: INCA; 2011. [acesso 2024 mar 17]. Disponível em: https://www.inca.gov.br/publicacoes/informativos/integrador-rhc
4. Ministério da Saúde (BR), Departamento de Ciência e Tecnologia, Secretaria de Ciência e Tecnologia e Insumos Estratégicos. Integração de informações dos registros de câncer brasileiros. Rev Saúde Pública. 2007;41:865-8.
5. Instituto Nacional de Câncer. Registros Hospitalares de Câncer [Internet]. Rio de Janeiro: INCA; 2010. [acesso 2024 mar 17]. Disponível em: https://www.inca.gov.br/publicacoes/manuais/registros-hospitalares-de-cancer
6. World Health Organization. Health information systems and rehabilitation [Internet]. 2017. [acesso 2024 mar 19]. Disponível em: https://www.who.int/publications/m/item/WHO-NMH-NVI-17.2
7. Frias PG, Szwarcwald CL, Lira PIC. Avaliação dos sistemas de informações sobre nascidos vivos e óbitos no Brasil na década de 2000. Cad Saúde Pública [Internet]. 2014 [acesso 2024 mar 17];30(10):2068-280. Disponível em: https://www.scielo.br/j/csp/a/pgMLGPpm8pZHN3mPsf5LPLC/?lang=pt
8. Girianelli VR, Thuler LCS, Silva GA. Qualidade do sistema de informação do câncer do colo do útero no estado do Rio de Janeiro. Rev Saúde Pública [Internet]. 2009 [acesso 2024 mar 17];43(4):580-8. Disponível em: https://www.scielo.br/j/rsp/a/yM6Bmxn3ZTRrbBnCHDsq6Bc/
9. Agranonik M, Jung RO. Qualidade dos sistemas de informações sobre nascidos vivos e sobre mortalidade

no Rio Grande do Sul, Brasil, 2000 a 2014. Ciênc Saúde Colet. [Internet]. 2019 [acesso 2024 Mar 17];24(5):1945-58. Disponível em: https://www.scielo.br/j/csc/a/YtDvqL8GJjCVgXLtDxxbJ3N/?lang=pt

10. Braz RM, Oliveira PTR, Reis AT, et al. Avaliação da completude da variável raça/cor nos sistemas nacionais de informação em saúde para aferição da equidade étnico-racial em indicadores usados pelo índice de desempenho do sistema único de saúde. Saúde debate [Internet]. 2013 [acesso 2024 mar 17];37(99):554-62. Disponível em: https://www.scielo.br/j/sdeb/a/ZqDr6yqgFryL5zXqCyrLVLc/abstract/?lang=pt#
11. Correia LOS, Padilha BM, Vasconcelos SML. Métodos para avaliar a completitude dos dados dos sistemas de informação em saúde do Brasil: uma revisão sistemática. Ciênc Saúde Colet. 2014;19(11):4467-78. doi: https://doi.org/10.1590/1413-812320141911.02822013
12. D'Alessandro TAL, Antoniazzi BN, Abreu DMX. Registros hospitalares de câncer de Minas Gerais: análise de consistênciadas bases de dados. Cad saúde coletiva. 2021;29(3):330-43. doi: https://doi.org/10.1590/1414-462X202129030230
13. Oliveira MM, Andrade SSCA, Santiago G, et al. Avaliação do Sistema de Informações sobre Nascidos Vivos. Brasil, 2006 a 2010. Epidemiol Serv Saúde. 2015 [acesso 2024 mar 17];24(4):629-40. Disponível em: https://doi.org/10.5123/S1679-49742015000400005
14. German RR, Lee LM, Horan JM, et al. Updated Guidelines for Evaluating Public Health Surveillance Systems. MMWR [Internet]. 2001[acesso 2024 mar 17];50(RR13):1-35. Disponível em: https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm
15. Conselho Nacional de Saúde (BR). Resolução nº 510, de 7 de abril de 2016. Dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais cujos procedimentos metodológicos envolvam a utilização de dados diretamente obtidos com os participantes ou de informações identificáveis ou que possam acarretar riscos maiores do que os existentes na vida cotidiana, na forma definida nesta Resolução [Internet]. Diário Oficial da União, Brasília, DF. 2016 maio 24 [acesso 2023 mar 17]; Seção I:44. Disponível em: http://bvsms.saude.gov.br/bvs/saudelegis/cns/2016/res0510_07_04_2016.html
16. Romero DE, Cunha CB. Avaliação da qualidade das variáveis sócio-econômicas e demográficas dos óbitos de crianças menores de um ano registrados no sistema de informações sobre mortalidade do Brasil (1996/2001). Cad Saúde Pública. 2006;22(3):673-81. Disponível em: https://doi.org/10.1590/S0102-311X2006000300022
17. Pinto IV, Ramos DN, Costa MCE, et al. Completude e consistência dos dados dos registros hospitalares de câncer no Brasil. Cad Saúde Colet [Internet]. 2012[acesso 2024 mar 17];20(1):113-20. Disponível em: https://ninho.inca.gov.br/jspui/handle/123456789/10112
18. R: The R Project for Statistical Computing [Internet]. Version 4.3.0. [place unknown]: The R foundation. 2021 Fev 15 - [acesso 2023 mar 17]. Disponível em: https://www.R-project.org
19. Instituto Nacional de Câncer. TNM: Classificação de Tumores Malignos [Internet]. 6. ed. Rio de Janeiro: INCA; 2004. [2023 nov 9]. Disponível em: https://bvsms.saude.gov.br/bvs/publicacoes/inca/tnm2.pdf
20. Ribeiro CM, Correa FM, Migowski A. Efeitos de curto prazo da pandemia de COVID-19 na realização de procedimentos de rastreamento, investigação diagnóstica e tratamento do câncer no Brasil: estudo descritivo, 2019-2020. Epidemiol Serv Saúde. 2022 [acesso 2024 mar 19];31(1):e2021405. doi: https://doi.org/10.1590/S1679-49742022000100010
21. Oliveira JCS, Azevedo EFS, Caló RS, et al. Registros hospitalares de Câncer de Mato Grosso: análise da completitude e da consistência. Cad Saúde Coletiva. 2021 [acesso 2024 mar 17];29(3):330-43. doi: https://doi.org/10.1590/1414-462X202129030230
22. Lopes-Júnior LC, Dell'Antonio LS, Pessanha RM, et al. completeness and consistency of epidemiological variables from hospital-based cancer registries in a brazilian state. Internacional J Meio Ambiente Res Saúde Pública. 2022;19(19):e-12003. doi: https://doi.org/10.3390/ijerph191912003
23. Keske GL. Análise e proposta de gestão para a qualidade de informação do registro hospitalar de câncer do Hospital Universitário de Santa Maria - RS [dissertação na internet]. Santa Maria: Universidade Federal de Santa Maria; 2021 [acesso 2024 mar 17]; Disponível em: http://repositorio.ufsm.br/handle/1/23575
24. Cardoso AL. Avaliação da qualidade das informações do registro hospitalar de câncer do município de Santa Maria, Rio Grande do Sul, Brasil [dissertação na internet]. Santa Maria: Universidade Federal de Santa Maria; 2018 [acesso 2024 mar 17]; Disponível em: http://repositorio.ufsm.br/handle/1/20850
25. Oliveira AS, Vasconcelos MMN, Abath MB, et al. Registros hospitalares de câncer em Pernambuco: da gestão ao registro. Rev Bras Cancerol. 2019;63(1):21-8. doi: https://doi.org/10.32635/2176-9745.RBC.2017v63n1.152
26. Felix JD, Zandonade E, Amorim MHC, et al. Avaliação da completude das variáveis epidemiológicas do Sistema de Informação sobre mortalidade em mulheres com óbitos por câncer de mama na Região Sudeste: Brasil (1998 a 2007). Ciênc Saúde Colet. 2012;17(4):945-53. Disponível em: https://doi.org/10.1590/S1413-81232012000400016
27. Barbuscia DM, Rodrigues-Júnior AL. Completude da informação nas declarações de nascido vivo e nas declarações de óbito, neonatal precoce e fetal, da região de Ribeirão Preto, São Paulo, Brasil, 2000-2007. Cad

Saúde Pública. 2011;27(6):1192-200. doi: https://doi.org/10.1590/S0102-311X2011000600016

28. Marques CA, Siqueira MM, Portuga FB. Avaliação da não completude das notificações compulsórias de dengue registradas por município de pequeno porte no Brasil. Ciênc Saúde Colet. 2020;25(3):891-900. doi: https://doi.org/10.1590/1413-81232020253.16162018
29. Muzy J, Castanheira D, Romero D. Análise da qualidade da informação da mortalidade prematura por doenças crônicas não transmissíveis e sua utilização nos objetivos de desenvolvimento sustentável. Cad saúde colet. 2021;29(spe):152-64. doi: https://doi.org/10.1590/1414-462X202199010456
30. Sousa CMS, Mascarenhas MDM, Lima PVC, et al. Incompletude do preenchimento das notificações compulsórias de violência - Brasil, 2011-2014. Cad saúde colet. 2020;28(4):477-87. doi: https://doi.org/10.1590/1414-462X202028040139
31. Araújo EM, Costa MCN, Hogan VK, et al. A utilização da variável raça/cor em Saúde Pública: possibilidades e limites. Interface (Botucatu). 2009;13(31):383-94. Disponível em: https://doi.org/10.1590/S1414-32832009000400012
32. Grabois MF, Souza MC, Guimarães RM, et al. Completude da informação "ocupação" nos registros hospitalares de câncer do Brasil: bases para a vigilância do câncer relacionado ao trabalho. Rev Bras Cancerol. 2014;60(3):207-14. doi: https://doi.org/10.32635/2176-9745.RBC.2014v60n3.465
33. Chaffin VS. Análise dos registros hospitalares de câncer para o monitoramento do câncer infantojuvenil [dissertação na Internet]. Rio de Janeiro: Fundação Oswaldo Cruz. Instituto Nacional de Saúde da Mulher da Criança e do Adolescente Fernandes Figueira; 2016. [acesso 2024 mar 17]; Disponível em: https://www.arca.fiocruz.br/handle/icict/25245

Recebido em 19/2/2024
Aprovado em 27/3/2024

Associated-editor: Jeane Tomazelli. Orcid iD: https://orcid.org/0000-0002-2472-3444
Scientific-editor: Anke Bergmann. Orcid iD: https://orcid.org/0000-0002-1972-8777