

Cálculos de Estimaciones Regionales de Cáncer Utilizando la Razón Incidencia/Mortalidad: Tutorial con el Apoyo de Scripts del Software R

<https://doi.org/10.32635/2176-9745.RBC.2025v71n1.4836>

Cálculos de Estimativas de Câncer Regionais pela Razão Incidência/Mortalidade: Tutorial com Apoio de Scripts do Software R
Calculations of Regional Cancer Estimates Using the Incidence/Mortality Ratio: Tutorial with the Support of R Software Scripts

Gil Patrus Pena¹; Sara Oliveira Ribeiro²

RESUMEN

Introducción: Una de las funciones principales de la vigilancia del cáncer es brindar al gestor un escenario de la situación del cáncer, proyectando expectativas sobre la incidencia de los distintos tipos de cáncer en los territorios, posibilitando la planificación racional de las acciones de prevención, detección temprana y tratamiento dirigidos a los principales tipos de cáncer que ocurren en esa región. La elaboración de estimaciones de cáncer es una de las principales herramientas en este sentido, pero su elaboración detallada para diferentes tipos de cáncer y regiones es compleja y laboriosa. El *software* R es una poderosa herramienta de manipulación de datos y su uso aún no está extendido en las actividades de vigilancia de la salud. **Objetivo:** Describir la metodología para el cálculo de estimaciones de incidencia de cáncer y poner a disposición el conjunto de *scripts* del *software* R utilizados, con el objetivo de difundir y promover su uso en actividades de vigilancia del cáncer. **Método:** Se presenta la elaboración de las Estimaciones de Cáncer en Minas Gerais, para el año 2024, por macrorregiones de salud. Todo el proceso de manipulación y cálculo se estructuró en *scripts* R, que están disponibles como suplemento. **Resultados:** La elaboración de estimaciones se ve favorecida por el uso del *software* R, dada la facilidad para manejar la gran cantidad de datos y los cálculos necesarios, además de la posibilidad de presentar los resultados en gráficos y mapas. **Conclusión:** Con la disponibilidad de los *scripts*, se espera que otros Estados puedan adaptar todo el proceso para preparar sus propias estimaciones regionales. **Palabras clave:** Vigilancia en Salud Pública; Neoplasias/epidemiología; Incidencia.

RESUMO

Introdução: Um dos principais papéis da vigilância do câncer é fornecer ao gestor um cenário da situação do câncer, projetando expectativas sobre a incidência dos diferentes tipos de câncer nos territórios, possibilitando o planejamento racional das ações de prevenção, detecção precoce e tratamento direcionados aos principais tipos de câncer incidentes naquela Região. A elaboração de estimativas de câncer constitui uma das principais ferramentas nesse sentido, mas sua elaboração detalhada pelos diferentes tipos de câncer e regiões é complexa e trabalhosa. O *software* R é uma ferramenta poderosa de manipulação de dados e seu uso ainda é pouco difundido nas atividades de vigilância em saúde. **Objetivo:** Descrever a metodologia do cálculo das estimativas de incidência do câncer e disponibilizar o conjunto de *scripts* do *software* R utilizado, com o objetivo de difundir e fomentar o seu uso nas atividades de vigilância do câncer. **Método:** É apresentada a elaboração das Estimativas de Câncer em Minas Gerais, para o ano de 2024, por Macrorregiões de Saúde. Todo o processo de manipulação e cálculos foi estruturado em *scripts* do R, que são disponibilizados como um suplemento. **Resultados:** A elaboração de estimativas é potencializada pelo uso do *software* R, pela facilidade da manipulação da grande quantidade de dados e dos cálculos necessários, além da possibilidade de apresentação dos resultados em gráficos e mapas. **Conclusão:** Com a disponibilização dos *scripts*, espera-se que todo o processo possa ser adaptado por outros Estados, para a elaboração de suas próprias estimativas regionais. **Palavras-chave:** Vigilância em Saúde Pública; Neoplasias/epidemiologia; Incidência.

ABSTRACT

Introduction: One of the main roles of cancer surveillance is to provide the manager a scenario of the cancer status, predicting expectations about the incidence of different types of cancer in the territories, enabling rational planning of prevention, early detection and treatment actions for the main types of cancer occurring in that region. The preparation of cancer estimates is one of the main tools in this regard, but their detailed elaboration for different types of cancer and regions is complex and laborious. Software R is a powerful data manipulation tool and its use is still not widespread in health surveillance activities. **Objective:** To describe the methodology for calculating cancer incidence estimates and make available the set of software R scripts used, with the objective of disseminating and promoting its use in cancer surveillance activities. **Method:** Description of the preparation of 2024 Cancer Estimates in Minas Gerais by Health Macro-regions. The entire manipulation and calculation process was structured in R scripts, which are available as a supplement. **Results:** The preparation of estimates is enhanced by the use of software R, facilitating the handling of large amount of data and the necessary calculations, in addition to the possibility of presenting results in graphs and maps. **Conclusion:** With the availability of the scripts, it is expected that the entire process is adjusted by other states to prepare their own regional estimates. **Key words:** Public Health Surveillance; Neoplasms/epidemiology; Incidence.

¹Registro de Câncer de Base Populacional de Belo Horizonte, Coordenação de Vigilância das Doenças e Agravos Crônicos não Transmissíveis e Câncer. Superintendência de Vigilância Epidemiológica, Subsecretaria de Vigilância em Saúde, Secretaria de Estado da Saúde de Minas Gerais. Belo Horizonte (MG), Brasil. E-mail: gil.pena@saude.mg.gov.br. Orcid iD: <https://orcid.org/0000-0002-9395-6372>

²Superintendência de Vigilância Epidemiológica, Subsecretaria de Vigilância em Saúde, Secretaria de Estado da Saúde de Minas Gerais. Belo Horizonte (MG), Brasil. E-mail: sara.oribeiro@yahoo.com.br. Orcid iD: <https://orcid.org/0009-0008-1433-3056>

Dirección para correspondencia: Gil Patrus Pena. Secretaria de Estado da Saúde. Rodovia João Paulo II, 4143, Edifício Minas, 13º andar – Serra Verde. Belo Horizonte (MG), Brasil. CEP 31630-900. E-mail: gil.pena@saude.mg.gov.br



INTRODUCCIÓN

El cáncer es una enfermedad compleja que engloba numerosos tipos histológicos y diferentes topografías. Las acciones de vigilancia de la enfermedad incluyen el monitoreo de la incidencia y de la mortalidad. En las regiones no cubiertas por los Registros de Cáncer de Base Poblacional (RCBP), la incidencia se puede estimar usando modelado a partir de las informaciones de mortalidad, usando razones de incidencia/mortalidad derivadas de datos de RCBP específicos de la región.

La elaboración de estimaciones de cáncer, como las publicadas periódicamente por el Instituto Nacional del Cáncer (INCA), para las Unidades de la Federación (UF) y capitales del Brasil, para tipos seleccionados de cáncer y por sexo, es ardua. Para cada tipo de cáncer, sexo y región, se necesitan cálculos relativos a la población, al número de fallecidos, obteniéndose tasas de mortalidad bruta y ajustada. En el cálculo de la mortalidad ajustada, es necesario calcular las tasas de mortalidad específicas por edad y aplicar estas tasas a una población estándar. Además, es necesario obtener las tasas de incidencia por tipo de cáncer y sexo en los RCBP seleccionados¹.

Aunque los cálculos involucrados en la metodología se basen en la epidemiología clásica, con metodología relativamente accesible, la gran cantidad de datos y cálculos implicados, multiplicados para cada tipo de cáncer, región y sexo, ciertamente inhibe iniciativas de elaboración de estimaciones con mejor detalle regional.

En Minas Gerais, el Programa de Evaluación de Vigilancia del Cáncer y sus Factores de Riesgo (PAV-MG) publicó en 2013 la primera estimación para las macrorregiones de salud del estado². Una actualización de las estimaciones fue íntegramente calculada, pero, debido a una alteración de la regionalización administrativa de la salud³, no llegó a publicarse. Recientemente, una nueva alteración del Plan Rector de Regionalización en Salud fue instituida⁴, ejemplificando la necesidad de tener un análisis estructurado de las poblaciones y de la mortalidad que pueda rápidamente ser adaptado a cambios en la organización administrativa del territorio.

Es necesario el conocimiento del escenario del cáncer en Minas Gerais, detallado por las macrorregiones en salud, dada la extensión y la diversidad regional, con marcadas variaciones ambientales, culturales y económicas. Estas informaciones permiten al gestor definir prioridades y planear acciones de enfrentamiento de la enfermedad según las particularidades regionales.

Además de las estimaciones anteriores realizadas en el ámbito de la SES-MG², se realizaron estudios con estimaciones de cáncer para subregiones estatales, calculadas por el método de la razón de la incidencia

mortalidad, para las 17 Redes Regionales de Atención a la Salud (RRAS) del estado de São Paulo⁵.

Las estimaciones de incidencia pueden generarse con base en diferentes métodos, de acuerdo con la disponibilidad de informaciones en cada país. En las estimaciones mundiales por país, el *Global Cancer Observatory: Cancer Today*⁶ (Globocan) aplica diferentes métodos. El método utilizado prioritariamente involucra la proyección de tasas de incidencias de RCBP con gran cobertura nacional o subnacional. Ante la falta de disponibilidad de tasas de incidencia con la mayor cobertura, se utilizan estimaciones nacionales de mortalidad por modelado, usando razones de incidencia/mortalidad derivadas de datos de registro de cáncer específicos de la región o regiones vecinas.

Además del método directo para el cálculo de la razón incidencia/mortalidad, métodos estadísticos involucrando modelos de multiniveles de Poisson para la estimación de la razón de la incidencia/mortalidad fueron empleados por Jardim et al.⁷, permitiendo considerar la variación de los efectos aleatorios de los RCBP que no serían captados por los modelos tradicionales de efectos fijos.

El *software* R⁸ es una herramienta para la manipulación de datos y análisis estadísticos. Uno de los mejores recursos que R pone a disposición es la posibilidad de mantener en un *script* todas las manipulaciones de datos y análisis hechas. Incluso en los análisis más complejos, fue posible recomenzar y rehacer después de las correcciones, con mucha rapidez. Un *script* organizado permite tener la rutina del trabajo documentada, la automatización de los procesos y la reproducibilidad del trabajo⁹. Adicionalmente, R es un *software* libre, que puede ser instalado sin el pago de licencias. El uso de R se potencia usando los numerosos paquetes que agregan funciones específicas al programa, ampliando los recursos de manipulación de los datos, análisis y visualización (gráficos y publicaciones).

En este artículo, se describió detalladamente la metodología de la elaboración de las estimaciones de cáncer en el estado de Minas Gerais, para las 16 macrorregiones de salud, 21 tipos específicos de cáncer y para todas las neoplasias, para los sexos masculino y femenino. Todo el recorrido de la preparación de los datos y cálculos fue estructurado en *scripts* de R que están a disposición como un suplemento, pudiendo ser adaptados para la elaboración de estimaciones en otros territorios (Brasil, UF y divisiones regionales por UF).

MÉTODO

Toda la manipulación de los datos fue realizada en el *software* R, versión 4.3.2.⁸ Se usó un computador con procesador Intel® Core™ i5-4210U CPU 1.70GHz 2.40

GHz, con 8 GB de RAM y sistema operativo Windows 10 Pro de 64 bits. Se instalaron los siguientes paquetes con sus respectivas dependencias: *dplyr*¹⁰, *tidyr*¹¹, *readxl*¹², *ggplot2*¹³, *read.dbc*¹⁴, *car*¹⁵, *gmodels*¹⁶, *lmtest*¹⁷, *rmarkdown*¹⁸, *knitr*¹⁹ y *sf*²⁰. Con excepción de los *scripts* en *rmarkdown* (.Rmd), los análisis se realizaron en la interfaz *Rgui.exe*, un programa ejecutable disponible en la instalación de R base que abre una consola en una ventana de *Windows*. Los *scripts* del *rmarkdown* (.rmd) fueron ejecutados en la interfaz *Rstudio*²¹. Todos los *scripts* fueron íntegramente preparados para la realización de este estudio, con excepción del *script* para las regresiones lineales de las tasas de mortalidad que fue adaptado de una versión previa utilizada en la preparación de las estimaciones por el INCA¹. Una breve descripción de las bases utilizadas está presente en el Material Suplementario.

Las bases de datos referentes a la población fueron descargadas del Departamento de Informática del Sistema Único de Salud (DATASUS)²². Entre 1980 y 2012, fueron usadas informaciones de los censos (1980, 1991, 2000 y 2010), conteo (1996) y proyecciones intercensales (1981 a 2012), elaboradas por el Instituto Brasileño de Geografía y Estadística (IBGE) y puestos a disposición por el DATASUS, según grupo etario, sexo y situación de residencia. Se descargaron las bases de 1980 a 2012 referentes a todo el Brasil, desde la página del Tabwin²². Para los años de 2013 a 2021, se utilizaron las estimaciones preliminares elaboradas por el Ministerio de Salud/Secretaría de Vigilancia en Salud y Ambiente/Departamento de Análisis Epidemiológico y Vigilancia de Enfermedades no Transmisibles/Coordinación General de Informaciones y Análisis Epidemiológicos (MS/SVSA/DAENT/CGIAE)²³. Las informaciones fueron tabuladas año a año, para cada sexo, seleccionándose el estado de Minas Gerais y configurándose los municipios en la fila y el grupo etario en la columna. La información referente a 2022 corresponde a la población del censo y fue obtenida del IBGE²⁴. Con la asistencia del *script* de R, las informaciones de la población organizadas de los 853 municipios de Minas Gerais fueron agrupadas según las macrorregiones de salud, conforme con la planilla del Plan Rector de Regionalización de Minas Gerais (ver mapa en la Figura 1 del Material Suplementario). Además de las informaciones por macrorregión de salud, se prepararon las informaciones de la población de Belo Horizonte y de Poços de Caldas, que tienen cobertura por RCBP, con la finalidad de realizar los cálculos de incidencia del cáncer. Se emplearon grupos de rango etario de cinco en cinco años de 0 a 79 años y mayores de 80 años. Se creó un *data.frame* largo con las columnas macrorregión, año, sexo, grupo etario y población de la macro (disponible como suplemento). También con

la asistencia de los *scripts* y del *software* R se elaboraron gráficos de pirámides de las poblaciones para cada año y macrorregión.

Las informaciones de mortalidad de 1979 a 2022, provenientes del Sistema de Información sobre Mortalidad (SIM), fueron descargadas de la página del Tabwin²². Entre 1996 y 2022, las causas básicas están codificadas usando la décima revisión de la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados a la Salud (CIE-10)²⁵ y, entre 1979 y 1995, por la novena revisión (CIE-9)²⁶. No fueron incluidos en el análisis muertes sin información de sexo (n=373; 0,06%), edad (n=834; 0,14%) y residencia (n=546; 0,09%). Las muertes por cáncer fueron agrupadas de acuerdo con la Tabla 1.

Los datos de los casos clasificados como útero sin otra especificación (SOE) (C55) fueron incorporados en el análisis, con la intención de producir tasas más fidedignas para las ubicaciones cuello uterino (C53) y cuerpo del útero (C54). Esta reasignación es recomendada con la intención de producir tasas que permitan comparaciones significativas entre poblaciones. Los métodos de reasignación fueron basados en la edad y en las distribuciones específicas de los casos²⁷. En la serie completa, entre 1979 y 2022, la proporción de casos “útero SOE (C55)” correspondió al 35,5% del conjunto de casos “12-cuello uterino (C53)”, “13-cuerpo del útero (C54)” y “útero SOE (C55)”. Cuando la edad fue menor de 50 años, todos los casos “útero SOE (C55)” fueron reasignados para “12-cuello uterino (C53)”. Si la edad era mayor o igual a 50 años, los casos “útero SOE (C55)” fueron reasignados para “12-cuello uterino (C53)” en la proporción de casos:

$$\frac{\text{"12 – Cuello uterino (C53)"}}{\text{"12 – Cuello uterino (C53)" + "13 – Cuerpo del útero (C54)"}}$$

para cada macrorregión y año;

y para “13-cuerpo del útero (C54)”, en la proporción de:

$$1 - \frac{\text{"12 – cuello uterino (C53)"}}{\text{"12 – cuello uterino (C53)" + "13 – cuerpo del útero (C54)"}}$$

para cada macrorregión y año.

Con las informaciones de población y del número de muertos, fueron calculadas las tasas de mortalidad específicas por neoplasia, edad y sexo, las tasas brutas y las tasas ajustadas por el método directo, de acuerdo con la población mundial (Tabla 1 del Material Suplementario).

La tasa de mortalidad específica por grupo etario (i) fue calculada para cada macrorregión (m), neoplasia (n), sexo (s) y año (a).



Tabla 1. Grupos de neoplasia según la codificación usada por la CIE-9 (1979-1995) y CIE-10 (1996-2022)

Grupos de neoplasia	Códigos CIE-9 (1979-1995)	Códigos CIE-10 (1996-2022) Tres primeros dígitos
01-Cavidad oral (C00-C10)	140, 141, 142, 144, 145, 146	C00, C01, C02, C03, C04, C05, C06, C07, C08, C09, C10
02-Esófago (C15)	150	C15
03-Estómago (C16)	151	C16
04-Colon y recto (C18-C21)	153, 154	C18, C19, C20, C21
05-Hígado y vías biliares intrahepáticas (C22)	155	C22
06-Páncreas (C25)	157	C25
07-Laringe (C32)	161	C32
08-Tráquea, bronquios y pulmón (C33-C34)	162	C33, C34
09-Melanoma maligno de la piel (C43)	172	C43
10-Otras neoplasias malignas de la piel (C44)	173	C44
11-Mama femenina (C50)	174	C50 y sexo femenino
12-Cuello uterino (C53)	180	C53
13-Cuerpo del útero (C54)	182	C54
Útero SOE (C55)*	179	C55
14-Ovario (C56)	183	C56
15-Próstata (C61)	185	C61
16-Vejiga (C67)	188	C67
17-Sistema nervioso central (C70-C72)	191, 192	C70, C71, C72
18-Glándula tiroides (C73)	193	C73
19-Linfoma de Hodgkin (C81)	201	C81
20-Linfoma no Hodgkin (C82-C85; C96)	200, 202	C82, C83, C84, C85, C96
21-Leucemias (C91-C95)	204, 205, 206, 207	C91, C92, C93, C94, C95
22-Otras neoplasias (C00-C97; D46)	147, 148, 149, 152, 156, 158, 159, 160, 163, 164, 165, 170, 171, 175, 176, 177, 178, 181, 184, 186, 187, 189, 190, 194, 195, 196, 197, 198, 203, 238	C11, C12, C13, C14, C17, C23, C24, C26, C27, C28, C29, C30, C31, C35, C36, C37, C38, C39, C40, C41, C45, C46, C47, C48, C49, C51, C52, C57, C58, C59, C60, C62, C63, C64, C65, C66, C68, C69, C74, C75, C76, C77, C78, C79, C80, C86, C88, C90, C97, D46

Leyenda: (*) la localización en útero sin otra especificación (SOE) fue reasignada a las localizaciones en cuello uterino y cuerpo uterino²⁷.

$$Tasa\ de\ mortalidad\ específica_{imnsa} = \frac{número\ de\ muertes_{imnsa}}{población_{imnsa}}$$

$$Tasa\ de\ mortalidad\ ajustada_{mnsa} = \frac{\sum(Tasa\ de\ mortalidad\ específica_{imnsa} \times población\ estándar_i)}{\sum Población\ estándar_i}$$

Leyenda: i = grupo etario; m = macrorregión, n = grupo de neoplasia; s = sexo; a = año.

Esta tasa fue aplicada a la población estándar, resultando en el número de muertos esperados en cada grupo etario. La sumatoria del número de muertos esperados se dividió entonces entre el total de la población estándar, para obtener la tasa ajustada por edad.

La tasa de mortalidad bruta fue calculada para cada macrorregión (m), tipo de cáncer (n), sexo (s) y año (a):

$$Tasa\ de\ mortalidad\ bruta_{mnsa} = \frac{\sum número\ de\ muertes_{imnsa}}{\sum población_{imnsa}}$$

Las tasas de mortalidad bruta y ajustada, para cada tipo de cáncer, sexo y macrorregión, entre 1979 y 2022, fueron sometidas a una regresión lineal, colocando al

año como variable numérica independiente. El *script* para la realización de las regresiones fue desarrollado en *rmarkdown*, adaptado de un *script* previamente utilizado en la realización de las estimaciones por el INCA. Los resultados de las regresiones, con los respectivos coeficientes de determinación (r^2 , una medida de ajuste del modelo), fueron analizados en un documento *Word* (.docx), junto con gráficos de dispersión de la tasa bruta de mortalidad versus el año del muerte, de la distribución de los residuos con relación a los valores ajustados y a la tasa bruta, y un gráfico QQ normal (para cada neoplasia, sexo y macrorregión). Si la regresión se consideró adecuada (valor de $r^2 > 0,7$ y distribución aleatoria de los residuos), se utilizaron las tasas calculadas a partir de la ecuación obtenida con la regresión. En los otros casos, se optó por la tasa media (bruta o ajustada) de los últimos cinco años (2018-2022).

Para calcular la población proyectada para las macrorregiones, fueron utilizadas las proyecciones de población para el estado de Minas Gerais como un todo, por grupo etario y sexo (proyección de la población del Brasil y UF por sexo y edad para el período 2010-2060, edición 2018, producidas por el IBGE y disponibles en el DATASUS)²³. El cálculo de la población para cada macrorregión se hizo según la metodología del IBGE²⁸, considerando la tendencia de crecimiento de la población entre los censos de 2010 y 2022²⁸:

$$\begin{aligned} Población_m &= a_i \times Población_{MG} + b_i \\ a_i &= \frac{Población_{m2022} - Población_{m2010}}{Población_{MG2022} - Población_{MG2010}} \\ b_i &= Población_{m2010} - a_i \times Población_{MG2010} \end{aligned}$$

Donde:

$Población_m$ = Población proyectada de la macrorregión

$Población_{MG}$ = Población estimada para Minas Gerais

$Población_{m2022}$ = Población de la macrorregión en el censo de 2022

$Población_{MG2022}$ = Población de MG en el censo de 2022

$Población_{m2010}$ = Población de la macrorregión en el censo de 2010

$Población_{MG2010}$ = Población de MG en el censo de 2010

En el tabulador del INCA²⁹, fue tabulado el número de casos de cáncer por grupo etario, según la CIE-10, para cada sexo, en los RCBP de Belo Horizonte (informaciones de 2015 a 2019) y de Poços de Caldas (informaciones de 2010 a 2014). Las poblaciones cubiertas por los RCBP de Belo Horizonte (2 315 560 habitantes) y de Poços de Caldas (163 742 habitantes) representan el 12,1% de la población del estado (20 539 989 habitantes), considerando los resultados del censo de 2022. Los casos notificados como C55 (útero SOE) fueron reasignados con los códigos C53 y C54, siguiendo la misma metodología utilizada para las muertes, justificando la obtención de

las informaciones de los grupos etarios. La razón de la incidencia mortalidad fue calculada para Belo Horizonte y Poços de Caldas de acuerdo con la metodología usada por el INCA¹ para cada tipo de neoplasia (subíndice n en las ecuaciones) y sexo (subíndice s en las ecuaciones). No fueron previstos ajustes para los casos en que la razón de la incidencia/mortalidad resultase un valor menor de 1.

$$Razón \frac{incidencia}{mortalidad}_{ns} = \frac{\frac{Número \ de \ casos \ nuevos_{ns}}{\sqrt{Población}}}{\frac{Número \ de \ muertes_{ns}}{\sqrt{Población}}}$$

Las estimaciones de incidencia se realizaron utilizando la media de la razón incidencia/mortalidad calculada para los registros de Poços de Caldas y de Belo Horizonte. Esta razón fue aplicada a las tasas brutas y ajustadas de mortalidad estimadas por regresión para el año 2024, para cada sexo (indicado por el subíndice s), macrorregión (subíndice m) y neoplasia (subíndice n). Cuando la regresión no se mostró adecuada, fueron utilizadas las tasas de mortalidad medias (bruta y ajustada) entre 2018 y 2022 (cinco años).

$$Tasa \ incidencia \ estimada_{mns2024} = Tasa \ mortalidad_{mns2024} \times Razón \frac{incidencia}{mortalidad}_{ns}$$

Los valores absolutos para el número de casos nuevos se calcularon aplicando la tasa bruta estimada a la población proyectada para 2024, para cada macrorregión y sexo.

$$Número \ de \ casos \ nuevos_{mns2024} = \frac{Tasa \ incidencia \ estimada_{mns2024} \times Población_{ms2024}}{100 \ 000}$$

Siguiendo la metodología de las estimaciones del INCA¹, solo se presentan los resultados del número de casos nuevos si fuere mayor o igual a 20. Los valores absolutos fueron redondeados a múltiplos de 10.

Las tasas ajustadas de incidencia estimadas fueron utilizadas para la representación espacial con base en las distribuciones de las tasas por cuartil.

Como toda la metodología de cálculo fue estructurada en *scripts*, fue posible replicarla para ambos sexos, para todas las neoplasias (incluyendo o no piel no melanoma) y el estado como un todo. De este modo, el número de casos estimados para ambos sexos no necesariamente corresponde exactamente a la suma aritmética del número de casos estimados para los sexos masculino y femenino. De la misma manera, los números de casos estimados para todas las neoplasias y todas las neoplasias excepto piel no melanoma, y para todo el estado no necesariamente corresponden a las sumas aritméticas del número de casos estimados para cada neoplasia (incluyendo o no piel no melanoma) o para cada macrorregión, respectivamente.

La consolidación de los resultados para publicación se hizo en un *script* del *rmarkdown*, siendo presentados en texto los valores de los casos estimados para Minas



Gerais como un todo y de las cinco macrorregiones con las mayores tasas brutas, para cada sexo. Los resultados de las tasas estimadas y el número de casos nuevos para cada una de las macrorregiones del estado son presentados en una tabla (ilustrada en el Cuadro 2 del Material Suplementario). Además se elaboraron mapas con las tasas ajustadas de incidencia por sexo, con base en la distribución por cuartil. En este artículo, solo a título de demostración de la metodología utilizada, se presentan las informaciones referentes a todas las neoplasias, excepto piel no melanoma, también considerando la importancia de esa información en la programación de la atención en la alta complejidad estatal.

RESULTADOS

Las pirámides poblacionales para las macrorregiones del estado muestran una variación grande en el contingente poblacional entre las macrorregiones de salud (Figura 2 del Material Suplementario). En las diferentes macrorregiones del estado, se nota un envejecimiento de la población, con el cambio en la forma del gráfico entre 1980 y 2022.

Para ilustrar el funcionamiento del *script*, se presentan los resultados obtenidos para todas las neoplasias, excepto piel no melanoma, para el estado de Minas Gerais (sexo femenino, masculino y ambos sexos). En la Figura 1, se encuentran las tasas de mortalidad brutas y ajustadas, entre 1979 y 2022, para masculino, femenino y ambos sexos.

Los resultados de las tasas proyectadas por la regresión lineal, los valores del r^2 para cada regresión y la tasa media de los últimos cinco años se presentan en la Tabla 2.

Para los cálculos de incidencia y de la razón incidencia mortalidad, se trabajaron las informaciones de los últimos cinco años disponibles para los RCBP de Poços de Caldas (2010-2014) y de Belo Horizonte (2015-2019). Los resultados de los cálculos para todas las neoplasias, excepto piel no melanoma, se presentan en la Tabla 3.

La razón de la incidencia/mortalidad calculada para las poblaciones de Belo Horizonte y de Poços de Caldas fue aplicada sobre las tasas de mortalidad proyectadas para 2024, permitiendo estimar las tasas brutas y ajustadas de incidencia para los demás territorios. Las tasas brutas estimadas, a su vez, aplicadas sobre la población del territorio resultan en el número de casos nuevos estimados para aquella población. Un ejemplo de estos cálculos se presenta en la Tabla 4.

El último paso en la elaboración de las estimaciones es la publicación de los resultados. El paquete *rmarkdown* posibilita la conjugación de los resultados de los análisis con elementos de texto, gráficos, tablas y mapas, con la elaboración de un documento (en el formato *MS Word*, PDF o HTML). Como ejemplo del documento obtenido

con la ejecución del *script*, se presenta un trecho en el Cuadro 1 del Material Suplementario.

En el documento elaborado, se incluyeron además tablas y mapas. Un ejemplo de tabla producida en el *rmarkdown* se encuentra en el Cuadro 2 del Material Suplementario. La representación espacial de las tasas de incidencia ajustadas estimadas para las macrorregiones del estado, organizadas mediante la distribución en cuartiles, se ilustra en la Figura 2 del Material Suplementario.

DISCUSIÓN

En las acciones de la vigilancia en salud no se puede más abdicar de los recursos de manejo y análisis de informaciones, como los que pone a disposición el *software* R y sus paquetes. El dominio de estas tecnologías es esencial para mantener la actualidad de los análisis de situación, incorporar el mayor número de informaciones al presente estudio y utilizar métodos estadísticos adecuados para cada análisis. Además, el *software* R y sus paquetes contribuyen mucho en la presentación de los resultados, en la elaboración de gráficos, mapas, y hasta de texto, como se hizo en este estudio.

Al contrario de *softwares* con recursos de señalar y clicar (*point and click*), el aprendizaje del *software* R no es intuitivo, exigiendo estudio y práctica. Es una herramienta gratuita (*software* libre) capaz de soportar una cantidad de datos mucho más allá de las 1 048 576 filas disponibles en una hoja de *Excel*. Además, toda manipulación y análisis estructurados en *scripts* permiten que se rehaga todo el proceso siempre que sea necesario, con relativa simplicidad, posibilitando la actualización de las informaciones. Como ejemplo, la reestructuración de la regionalización puede ser rápidamente incorporada a los análisis.

Compartir los *scripts* es otra ventaja, permitiendo que el análisis desarrollado para Minas Gerais pueda ser adaptado para otros estados y para el Brasil como un todo. Los *scripts* utilizados en este estudio están disponibles (ver enlace en el Material Suplementario), con la esperanza de que sean utilizados y perfeccionados.

Una experiencia previa compartiendo *scripts* de R, para la preparación de las bases de datos de los RCBP para su remisión a un estudio internacional, facilitó la participación de registros brasileños en el estudio. La dirección del estudio consideró la estrategia útil y hará disponibles los *scripts* adaptados a registros de otros países³⁰. Con base en tales experiencias, se cree que las acciones de la vigilancia en salud puedan ser potenciadas con el uso del *software* R, mucho más allá de la elaboración de estimaciones.

Siempre que se instala una nueva versión actualizada de R, es necesario actualizar también los paquetes. Cuando

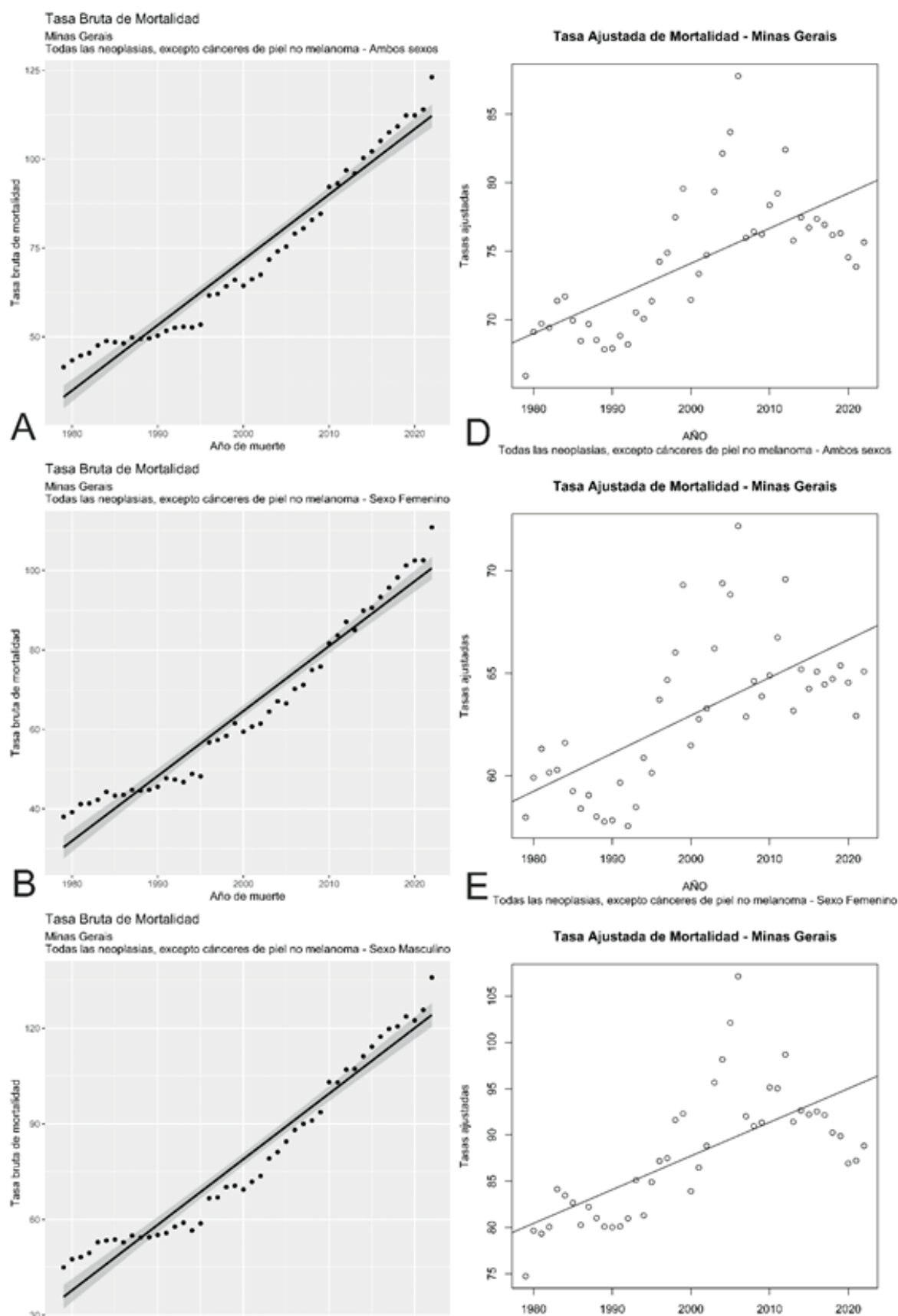


Figura 1. Tasas de mortalidad brutas (A, B, C) y ajustadas (D, E, F) para todas las neoplasias, excepto piel no melanoma en el estado de Minas Gerais, para ambos sexos (A, D), sexo femenino (B, E) y sexo masculino (C, F). Las rectas señalan la tendencia estimada por regresión lineal. Ejes y con escalas diferentes entre los gráficos



Tabla 2. Tasas proyectadas de mortalidad para el año 2024, para todas las neoplasias, excepto piel no melanoma usando la ecuación de la recta de regresión lineal, con la medida del r^2 . Tasa media de los últimos cinco años utilizada como alternativa, en los casos en que la regresión no fue considerada satisfactoria

Tipo de tasa	Sexo	Tasa proyectada por la regresión	r^2	Tasa media últimos cinco años	Tasa elegida para cálculo
Bruta	Ambos	115,85	0,95	114,12	115,85
Bruta	Femenino	103,82	0,95	103,04	103,82
Bruta	Masculino	128,29	0,95	125,59	128,29
Ajustada	Ambos	80,25	0,45	75,28	75,28
Ajustada	Femenino	67,37	0,42	64,52	64,52
Ajustada	Masculino	96,47	0,47	88,58	88,58

Tabla 3. Cálculo de la razón incidencia/mortalidad para todas las neoplasias, excepto piel no melanoma

Sexo		Belo Horizonte (2015-2019)			Poços de Caldas (2010-2014)		
		Muertes	Casos	Población	Muertes	Casos	Población
Ambos	Valores absolutos	16 550	36 613	12 461 005	928	2149	781 295
	Razón I/M	2,39			2,31		
	Mediana	2,35					
Femenino	Valores absolutos	8417	20 503	6 627 012	410	978	403 729
	Razón I/M	2,43			2,38		
	Mediana	2,41					
Masculino	Valores absolutos	8133	19 110	5 833 993	518	1171	377 566
	Razón I/M	2,35			2,26		
	Mediana	2,31					

Leyenda: I/M = incidencia/mortalidad.

Tabla 4. Cálculo de las tasas estimadas brutas y ajustadas para todas las neoplasias, excepto piel no melanoma en Minas Gerais, en el año 2024

Tipo de tasa	Sexo	Tasa de mortalidad proyectada (2024)	Razón incidencia/mortalidad	Tasa de Incidencia estimada (2024)	Población estimada (2024)	Casos nuevos estimados (2024)	Redondeo
Tasa bruta	Ambos	115,85	2,35	272,79	21 737 475	59 298,43	59 300
	Femenino	103,82	2,41	250,28	11 037 396	27 624,84	27 620
	Masculino	128,29	2,31	295,72	10 700 079	31 642,01	31 640
Tasa ajustada	Ambos	75,28	2,35	177,26			
	Femenino	64,52	2,41	155,54			
	Masculino	88,58	2,31	204,19			

un paquete no se actualiza a una nueva versión, este podrá salir del repositorio oficial en el *Comprehensive R Archive Network* (CRAN) y no estar más accesible fácilmente para los usuarios. Como estrategia para enfrentar esta situación, es posible mantener las versiones antiguas, con los paquetes previamente instalados.

La elaboración de los *scripts* de R involucra la creación de estrategias para la solución de problemas en el manejo de datos y análisis, en muchos aspectos limitados por el propio conocimiento de los múltiples potenciales que

pone a disposición la herramienta. En ese sentido, muchos aspectos pueden ser mejorados y pequeñas imperfecciones, presentes por casualidad, pueden ser corregidas.

El Sistema de Vigilancia en Salud tiene hoy en el DATASUS/Tabwin un gran aliado para poner a disposición las informaciones. Una mejor cohesión de las informaciones, sin embargo, sería deseable. En las informaciones de población, por ejemplo, fue posible descargar las bases completas en el Tabwin de 1980 a 2012, hacer tabulaciones en el DATASUS de estimaciones



poblacionales entre 2013 y 2021, y descargar las informaciones del censo de 2022 en el sitio web del IBGE. Aunque los archivos estén bien documentados y las fuentes de informaciones preservadas, sería deseable tener todos los datos disponibles en el mismo formato y en un mismo directorio. No está disponible, sin embargo, una retroproyección para 1979 (año en que inicia la serie del SIM), llevándonos a utilizar la población de 1980 también para 1979. No hay, hasta el momento, proyecciones de población a partir de 2023 distribuidas por municipio y grupos de edad, siendo necesario realizar los cálculos a partir de la población proyectada para el estado de Minas Gerais.

Las estimaciones de incidencia deben ser comprendidas como estimaciones, no como datos del mundo real. De este modo, alguna variabilidad en los resultados es esperada, no solo por la incorporación de nuevas bases, sino también por los resultados de los análisis de regresión y proyección utilizados. En los análisis realizados en este estudio, por ejemplo, están incorporadas las informaciones de mortalidad referentes al año 2022 y los datos poblacionales del censo de 2022, diferente de las proyecciones calculadas por el INCA para 2023¹, que utilizó datos hasta 2020. En comparación con las proyecciones elaboradas por el INCA¹, la presente estimación proyecta mayor número de casos, principalmente para el sexo masculino –52 090 casos en la proyección del INCA, 59 300 en esta estimación (Tabla 2 del Material Suplementario). La mayor parte de esa diferencia se puede atribuir al número de casos estimados para el cáncer de próstata –7970 casos nuevos estimados por el INCA y 11 800 casos nuevos en esta estimación. Específicamente en este caso, la metodología del INCA utiliza la mediana de la región geográfica, de modo que se evite la inflación de la tasa de incidencia producida por el aumento de la detección de la neoplasia, por la llegada de la prueba del antígeno específico de la próstata (PSA)¹.

Esta comparación solo busca demostrar la consistencia y la viabilidad de la metodología propuesta en el artículo, no estableciendo una comparación directa con otras estimaciones.

El cálculo de las estimaciones tiene limitaciones que pueden estar relacionadas a la calidad de las informaciones de incidencia, de la mortalidad y de la propia población. En lo que se refiere a la incidencia, las bases del RCBP de Poços de Caldas fueron analizadas y aprobadas para publicación en el *Cancer Incidence in Five Continents*, volúmenes XI y XII, respectivamente. Un criterio de calidad de las informaciones de los registros es la ocurrencia de una razón incidencia/mortalidad menor de 1, indicando posible subnotificación de casos, si las declaraciones de defunción son exactas y la incidencia y

la sobrevida de la neoplasia son constantes. Considerando las medianas de las razones utilizadas para el cálculo de las estimaciones, esta situación ocurrió para las neoplasias de hígado y vías biliares intrahepáticas (sexo masculino, 0,97; sexo femenino, 0,81 y ambos sexos, 0,93) y páncreas (sexo femenino, 0,89; y ambos sexos, 0,94), sabidamente neoplasias de alta letalidad y corta sobrevida. Con la decisión de no proponer cualquier ajuste en la razón incidencia/mortalidad, es posible que alguna subestimación ocurra para estas neoplasias. Tomando el ejemplo extremo de la neoplasia de hígado y vías biliares intrahepáticas, en el caso femenino, si la razón hubiese sido ajustada a 1, se esperarían 420 casos nuevos en Minas Gerais, al contrario de 340 casos nuevos calculados sin este ajuste. Es posible que parte de los fallecimientos notificados para esas neoplasias represente casos detectados ya en fase terminal, sin tiempo hábil para la confirmación del diagnóstico clínico, dificultando su inclusión en los Registros de Cáncer.

Se usó el modelo de regresión lineal para la proyección de las tasas de mortalidad, aplicado a una larga serie temporal (1979-2022). Este enfoque no capta cambios en la tendencia de crecimiento o disminución de las tasas a largo del tiempo. Esta dificultad fue esquivada empleando la tasa media de los últimos cinco años, en los casos en que el ajuste del modelo de regresión no fue considerado adecuado.

Las informaciones de la población también necesitan ser calificadas; los resultados del censo de 2022 revelaron una divergencia entre las proyecciones poblacionales anteriores y el conteo efectivo de la población, indicando la necesidad de conciliar estas informaciones. Para este estudio, específicamente, no se realizó ajuste alguno en las informaciones de la población puestas a disposición por el IBGE (censos de 1980, 1991, 2000, 2010 y 2022; conteo de 1996; y proyecciones intercensales de 1981 a 2012) o por el Ministerio de Salud (estimaciones preliminares de 2013 a 2021). Un posible efecto de enumeración por debajo de la población del censo de 2022 sería un aumento de la tasa de mortalidad para este año. Las proyecciones de población usadas para 2024, sin embargo, se basan en las informaciones hasta 2018 y no incorporan datos del censo de 2022.

En las informaciones sobre mortalidad, se hizo la reasignación de los casos notificados con útero SOE, siguiendo la metodología adoptada en las estimaciones elaboradas por el INCA. Diferentemente de lo realizado en el estudio de Jardim et al.⁷, no se hizo la reasignación de otros códigos de neoplasia maligna mal definidos (C26.0, C26.8, C26.9, C57.8, C57.9, C76.0, C76.2, C76.3, C78-C79 y C80) que, en el presente estudio, fueron incluidos en el grupo otras neoplasias. Considerando

solamente la mortalidad entre 1996 y 2022, el grupo de neoplasias mal definidas del tracto digestivo (C26.0, C26.8 y C26.9) representó el 6,9% del conjunto de muertes por neoplasias C15-C26, pudiendo acarrear una pequeña subestimación de casos atribuidos a estos específicos grupos objetivo.

Por el planeamiento estatal, las macrorregiones de salud constituyen la base territorial de planeamiento de la atención terciaria a la salud que engloba microrregiones de salud con población de unos 700 000 habitantes y que ofrece a su población servicios de salud hospitalarios de mayor densidad tecnológica⁴. Aun así, existen macrorregiones con población relativamente pequeña, como la macrorregión del Jequitinhonha, con 385 593 habitantes en 2022, llevando a la mayor fluctuación de las tasas, principalmente para los cánceres de menor frecuencia. Se trata de una limitación para el cálculo de las estimaciones de neoplasias con menor carga de enfermedad. No obstante, no hay otra forma de división en menor número de agrupamientos. Hubo un intento de superar esta limitación con la decisión de elaborar estimaciones para ambos sexos, lo que puede ser adecuado en el caso de tumores sin nítido predominio en uno de los sexos. También fue omitido el número de casos nuevos, cuando la estimación resultó un número menor de 20. A medida que la atención oncológica tiende cada vez más a la especialización, quizás sea adecuado agrupar neoplasias correlacionadas de acuerdo con la tendencia en la especialización en oncología (tumores hematológicos, tumores del tracto digestivo, del sistema genital femenino, etc.) como apoyo para el planeamiento de la organización de la red de atención. Por otro lado, neoplasias con especificidades propias en lo que se refiere a la prevención o detección temprana (por ejemplo, cáncer de cuello uterino, cáncer de mama, cáncer colorrectal) deben ser evaluadas individualmente.

El Brasil tiene una larga tradición en la elaboración de estimaciones de cáncer, publicadas regularmente desde mediados de la década de 1990. Dadas las enormes dimensiones del país, con apenas una parte de la población cubierta por los RCBP³¹, las estimaciones asumen un papel fundamental en el dimensionamiento de la magnitud y del impacto del cáncer en las diferentes regiones del país. La calidad de las estimaciones es altamente dependiente de la calidad de las informaciones utilizadas en su cálculo, incluyendo las informaciones de incidencia obtenidas de los RCBP, de la mortalidad (mediante el SIM) y de la propia población. El desafío de proporcionar informaciones de la incidencia de cáncer en subregiones estatales puede darle al gestor sustento para planear, monitorear y evaluar las acciones de control de cáncer y auxiliar en la gestión de las redes estatales de

atención oncológica tanto en el aspecto de la prevención como en el diagnóstico temprano y tratamiento como lo recomienda la Política Nacional de Prevención y Control del Cáncer en el ámbito del Sistema Único de Salud (SUS) y el Programa Nacional de Navegación de la Persona con Diagnóstico de Cáncer³².

CONCLUSIÓN

En este estudio se presentó todo el recorrido metodológico para la elaboración de estimaciones de cáncer para el estado de Minas Gerais y sus macrorregiones de salud, y fueron puestos a disposición como suplemento todos los *scripts* utilizados en la preparación de las informaciones, cálculos necesarios y presentación de los resultados. La expectativa es que este artículo pueda ser utilizado como un tutorial, junto con los *scripts* ahora disponibles, para la elaboración de estimaciones en otros territorios y poblaciones, permitiendo detallar la distribución de los diferentes tipos de cáncer a nivel intraestatal, donde estén disponibles informaciones adecuadas de incidencia, es decir, donde puedan usarse informaciones de RCBP existentes, con calidad de información y cobertura satisfactorias, específicos de la región o de regiones vecinas.

AGRADECIMENTOS

Al equipo de la División de Vigilancia y Análisis de Situación, de la Coordinación de Prevención y Vigilancia (Conprev) del INCA, Río de Janeiro (RJ), Brasil, en la persona de la entrañable Marcell de Oliveira Santos (*in memoriam*), quien contribuyó mucho para la realización de estas estimaciones regionales. A Fernanda Cristina da Silva de Lima, quien contribuyó con una versión de *script* del *rmarkdown* para la regresión de las tasas.

APORTES

Gil Patrus Pena y Sara Oliveira Ribeiro contribuyeron sustancialmente en la concepción y en el planeamiento del estudio; en la obtención, análisis e interpretación de los datos; en la redacción y revisión crítica. Gil Patrus Pena elaboró los *scripts* de R. Ambos autores aprobaron la versión final a publicarse.

DECLARACIÓN DE CONFLICTOS DE INTERÉS

Nada a declarar.

FUENTES DE FINANCIAMIENTO

No hay.

REFERENCIAS

1. Santos MO, Lima FCS, Martins LFL, et al. Estimativa de incidência de câncer no Brasil, 2023-2025. *Rev Bras Cancerol.* 2023;69(1):e-213700. doi: <https://doi.org/10.32635/2176-9745.RBC.2023v69n1.3700>
2. Secretaria de Estado da Saúde (MG). Programa de Avaliação e Vigilância do Câncer e seus fatores de risco (PAV-MG). Situação do câncer em Minas Gerais e suas macrorregiões de saúde: estimativas de incidência e mortalidade para o ano 2013, válidas para 2014: perfil da mortalidade: perfil da assistência na alta complexidade. Belo Horizonte: SES-MG; 2013. v. 1.
3. Secretaria de Estado de Saúde (MG). Subsecretaria de Gestão Regional. Ajuste do Plano Diretor de Regionalização de Saúde de Minas Gerais (PDR/MG) [Internet]. 1. ed. Belo Horizonte: SES-MG; 2020. [Acesso 2024 jul 7]. Disponível em: https://www.saude.mg.gov.br/images/1_noticias/06_2023/2-jul-ago-set/regionalizacao/1-PDR%202020.pdf
4. Secretaria de Estado de Saúde (MG) [Internet]. Belo Horizonte: SEM-MG; 2023. Plano Diretor de Regionalização, 2023 ago 17. [Acesso 2024 jul 7]. Disponível em: <https://www.saude.mg.gov.br/gestor/regionalizacao>
5. Luizaga CTM, Buchalla CM. Estimativa da incidência de câncer no Estado de São Paulo, Brasil, a partir de dados reais. *Cad Saúde Pública*, 2023;39(2):e00134222. doi: <https://doi.org/10.1590/0102-311XPT134222>
6. Ferlay J, Ervik M, Lam F, et al. Global Cancer Observatory: cancer today [Internet]. Lyon: International Agency for Research on Cancer; 2024 [acesso 2024 out 9]. Disponível em: <https://gco.iarc.who.int/today/en/data-sources-methods>
7. Jardim BC, Junger WL, Daumas RP, et al. Estimativa de incidência de câncer no Brasil e regiões em 2018: aspectos metodológicos. *Cad Saúde Pública*. 2024;40(6):e00131623. doi: <https://doi.org/10.1590/0102-311XPT131623>
8. R: The R Project for Statistical Computing [Internet]. Versão 4.3.2 Viena: The R foundation. 2021 nov 2 - [acesso 2022 set 6]. Disponível em: <https://www.r-project.org>
9. Lemes V, Baccaro FB. Introdução ao uso do programa R [Internet]. Manaus: Instituto Nacional de Pesquisas da Amazônia, 2011. [acesso 2024 out 9]. Disponível em: https://www.researchgate.net/profile/Victor-Landeiro/publication/275035302_Introducao_ao_uso_do_programa_R/links/553041070cf20ea0a06f67ca/Introducao-ao-uso-do-programa-R.pdf
10. Wickham H, François R, Henry L, et al. dplyr: a grammar of data manipulation [Internet]. Versão 1.1.3. [sem local]: The R foundation. 2023 nov 17. [acesso 2024 out 9]. Disponível em: <https://CRAN.R-project.org/package=dplyr>
11. Wickham H, Vaughan D, Girlich M. Tidy: Tidy Messy Data [Internet]. Versão 1.3.1, [sem local]: The R foundation. 2024 jan 1. [acesso 2024 out 9]. Disponível em: <https://CRAN.R-project.org/package=tidyr>
12. Wickham H, Bryan J. readxl: Read Excel Files [Internet]. Versão 1.4.3, [sem local]: The R foundation. 2023 nov 17. [acesso 2024 out 9]. Disponível em: <https://CRAN.R-project.org/package=readxl>
13. Wickham H. ggplot2: Elegant graphics for data analysis. New York: Springer-Verlag; 2016.
14. Petruzalek D. read.dbc: Read Data Stored in DBC (Compressed DBF) Files [Internet]. Versão 1.0.6. [sem local]: The R foundation. 2024 jul 6. [acesso 2024 out 9]. Disponível em: <https://CRAN.R-project.org/package=read.dbc>
15. Fox J, Weisberg S. An R companion to applied regression [Internet]. 3 ed. Thousand Oaks: Sage; 2023. [acesso 2024 out 9]. Disponível em: <https://www.john-fox.ca/Companion/>
16. Warnes GR, Bolker B, Lumley TL, et al. Package 'gmodels', 2024. Versão 1.0.6. [sem local]: The R foundation. 2024. doi: <https://doi.org/10.32614/CRAN.package.gmodels>
17. Zeileis A, Hothorn T (2002). "Diagnostic Checking in Regression Relationships" [Internet]. *R News* [Internet]. 2002 [acesso 2024 jul 17];2(3):7-10. Disponível em https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf
18. Allaire J, Xie Y, Dervieux C, et al. rmarkdown: Dynamic Documents for R [Internet]. Versão 2.25. São Francisco: GitHub, Inc.; 2023. [Acesso 2024 jul 17]. Disponível em <https://github.com/rstudio/rmarkdown>
19. Xie Y. knitr: A General-Purpose Package for Dynamic Report Generation in R. R package [Internet]. version 1.44. [Acesso 2024 jul 17]. Disponível em <https://yihui.org/knitr/>
20. Pebesma E. "Simple features for R: standardized support for spatial vector data." *R Journal*. 2018;10(1):439-46. doi: <https://doi.org/10.32614/RJ-2018-009>
21. RStudio [Internet]. Versão 2023.12.1+402. Boston: Posit Software, PBC. 2024 abr 1 - [acesso 2024 mar 1]. Disponível em: <http://www.rstudio.com/ide>
22. TABWIN [Internet]. Brasília (DF): DATASUS. c2008 - [acesso 2024 jul 28]. Disponível em: <https://datasus.saude.gov.br/transferecia-de-arquivos/>
23. TABNET [Internet]. Brasília (DF): DATASUS. c2008. [acesso 2024 jul 28]. Disponível em: <http://tabnet.datasus.gov.br/cgi/defhttm.exe?ibge/cnv/popsvsbr.def>
24. SIDRA: Banco de Tabelas Estatísticas [Internet]. Rio de Janeiro: IBGE; [sem data]. [acesso 2024 jul 28]. Disponível em: <https://sidra.ibge.gov.br/tabela/9514>
25. Organização Mundial da Saúde. CID-10: Classificação Estatística Internacional de Doenças e problemas relacionados à saúde. São Paulo: Edusp; 2008.



26. Organização Mundial da Saúde. CID-9: Classificação Estatística Internacional de Doenças: manual de lesões e causas de óbito. São Paulo: Centro Brasileiro para Classificação de Doenças em Português; 1979.
27. Loos AH, Bray F, McCarron P, et al. Sheep and goats: separating cervix and corpus uteri from imprecisely coded uterine cancer deaths, for studies of geographical and temporal variations in mortality. *Eur J Cancer*. 2004;40(18):2794-803. doi: <https://doi.org/10.1016/j.ejca.2004.09.007>
28. Instituto Brasileiro de Geografia e Estatística. Estimativas da população 2021 nota metodológica nº 1. Rio de Janeiro: IBGE; 2021
29. Registro de Câncer de Base Populacional [Internet]. Rio de Janeiro: INCA. [2012] – [acesso 2024 jul 20]. Disponível em: <https://www.inca.gov.br/BasePopIncidencias/Home.action>
30. Pena Gil P, Pongnikorn D, Khan Baloch F, et al. Sharing R scripts to re-code cancer registry data for international research. In: 12º Congresso Brasileiro de Epidemiologia; 2024 nov 23-27; Rio de Janeiro. Rio de Janeiro: Abrasco; 2024.
31. Pena GPM, Santos ATC, Pezzotti CM, et al. Registros de câncer de base populacional no Brasil: relevância, desafios e oportunidades. *Rev Bras Cancerol*. 2025;71(1):e-104878. doi: <https://doi.org/10.32635/2176-9745.RBC.2025v71n1.4878>
32. Presidência da República (BR). Lei nº. 14.758 de 19 de dezembro de 2023. Institui a Política Nacional de Prevenção e Controle do Câncer no âmbito do Sistema Único de Saúde (SUS) e o Programa Nacional de Navegação da Pessoa com Diagnóstico de Câncer; e altera a Lei nº 8.080, de 19 de setembro de 1990 (Lei Orgânica da Saúde). *Diário Oficial da União*, Brasília, DF. 2023 dez 20; Seção 1.

Recebido em 29/7/2024
Aprovado em 11/11/2024

