

Incremento no tempo de seguimento dos pacientes acompanhados no A.C.Camargo Cancer Center, por meio da vinculação entre as bases de pacientes e de óbitos

ISSN
2446-7537

Diego Rodrigues Mendonça e Silva
(diego.rodrigues@accamargo.org.br)
Doutor em Epidemiologia e coordenador do Registro Hospitalar de Câncer do A.C.Camargo Cancer Center

Bernadette Cunha Waldvogel
(bvogel@seade.gov.br)
Doutora em Saúde Pública e Gerente de Demografia da Fundação Seade

Magaly de Losso Perdigão
(mdellosso@seade.gov.br)
Analista de Projetos da Fundação Seade

Rosa Maria Vieira de Freitas
(rfreitas@seade.gov.br)
Analista de Projetos da Fundação Seade

Lilian Cristina Correia Morais
(lcorreia@seade.gov.br)
Analista de Projetos da Fundação Seade

Maria Paula Curado
(mp.curado@accamargo.org.br)
Doutora em Oncologia e chefe do Grupo de Epidemiologia e Estatística em Câncer do A.C.Camargo Cancer Center

RESUMO

A metodologia de vinculação determinística entre bases de dados representa alternativa eficaz para aprimorar a qualidade das informações sobre o fenômeno que se pretende estudar. No caso em que a vinculação é realizada entre registros de uma doença e os óbitos, é possível reduzir a perda de dados sobre o *status* vital, melhorar o tempo de seguimento e agregar a informação sobre a causa básica do óbito. O presente estudo analisou os resultados decorrentes da aplicação dessa metodologia entre as bases de dados do A.C.Camargo Cancer Center e as bases de registros civis de óbitos processadas para o Estado de São Paulo pela Fundação Seade.

Entre 2000 e 2021, 79.508 pacientes foram vinculados a 5.693.126 registros de óbitos ocorridos de 2002 a 2021, identificando 12.517 casos como falecidos. Destes, 53,0% eram do sexo masculino, 65,2% tinham mais de 60 anos e 40,3% apresentaram estadiamento clínico inicial. O tempo médio de acompanhamento aumentou significativamente após a vinculação, para a maioria dos cânceres analisados, sendo que para aqueles de maior letalidade, como pulmão e pâncreas, o impacto no tempo médio de seguimento foi menor.

A vinculação determinística realizada possibilitou relevante melhora na atualização do *status* vital e do tempo de seguimento, além do conhecimento da causa básica da morte, produzindo estimativas de sobrevida mais precisas e completas.

INTRODUÇÃO

A existência de diferentes bases de dados em saúde, independentes e criadas com objetivos específicos, pode revelar aspectos parciais ou fragmentados em relação a uma determinada doença quando analisadas separadamente. Todavia, a abordagem conjunta e integrada dessas bases, resultante da colaboração entre instituições parceiras, tem o potencial de ampliar a quantidade e a qualidade das informações disponíveis e permitir o conhecimento mais detalhado da questão de saúde analisada. A vinculação (*linkage*) determinística entre bases de dados representa relevante processo de tratamento conjunto nessa área.

Essa linha de pesquisa, que procura integrar diferentes bases de dados adotando técnicas de vinculação determinística, tem sido aprimorada na Fundação Seade por meio de inúmeras parcerias com diferentes instituições, centros de pesquisa e universidades. A aplicação dessa metodologia utilizando as bases de estatísticas do registro civil processadas no Seade tem sido responsável pela produção de teses, artigos científicos publicados e trabalhos apresentados em congressos nacionais e internacionais. Em estudo publicado pela Fundação Seade, Waldvogel *et al.* (2019) descrevem as diversas experiências com a

aplicação da metodologia de vinculação determinística entre diferentes bases de dados.

Um aspecto crítico para a realização de projetos que adotam o processo de vinculação determinística de bases de dados é a questão da confidencialidade. Na maioria das vezes, os pesquisadores não têm acesso a bases dessa natureza em função da existência de critérios de proteção à informação de natureza individual. De fato, garantir e controlar o compartilhamento e o tratamento de informações pessoais e sigilosas são exigências fundamentais contempladas na Lei Geral de Proteção de Dados (LGPD), com a qual o Seade está em conformidade, incluindo a produção de estatísticas demográficas.

A atuação da Fundação Seade na linha de pesquisa de vinculação entre bases de dados tem sido estratégica para esses estudos, pois, além de garantir arcabouço legal/jurídico entre as instituições, assegura também a não incorporação de variáveis de identificação nas bases resultantes da vinculação de seus registros, fornecendo aos pesquisadores acesso apenas às variáveis necessárias para a realização de seu estudo.

Para os estudos de sobrevida de pacientes com câncer, o seguimento atualizado com a última informação relativa ao *status* vital é essencial para a produção de estimativas de sobrevivência mais precisas, uma vez que altas taxas de perda de seguimento geram resultados equivocados, geralmente superestimados, reduzindo a confiabilidade das análises estatísticas e gerando potencial viés (Swaminathan *et al.*, 2008; Gil *et al.*, 2022; Pereira *et al.*, 2023).

A perda de seguimento pode ocorrer devido a término do tratamento, mudança de instituição de acompanhamento, retorno do paciente à cidade de origem, ou óbito. Assim, o relacionamento entre as bases de dados de pacientes com câncer e aquelas de registros de óbitos constitui alternativa eficiente para aprimorar a qualidade dos dados, possibilitando reduzir a perda de informação sobre o *status* vital do paciente, melhorar o tempo de seguimento, agregar a informação relativa à causa básica do óbito e, conseqüentemente, melhorar a precisão das estimativas de sobrevivência.

Firmou-se uma parceria entre o Registro Hospitalar de Câncer do A.C. Camargo Cancer Center e a Fundação Seade, com o intuito de recuperar informações sobre o *status* vital dos pacientes acompanhados neste hospital. Adotou-se o procedimento de vinculação determinística entre a base de pacientes do hospital e a base de óbitos processada na Fundação Seade, a partir das informações enviadas por todos os Cartórios de Registro Civil do Estado de São Paulo. Tal método tornou possível identificar os pacientes já falecidos, agregando a informação sobre a data do óbito e a respectiva causa básica de morte.

1. CONTEXTUALIZAÇÃO

O monitoramento dos casos incidentes de câncer é feito por meio do Registro de Câncer de Base Populacional (RCBP), em colaboração com os Registros Hospitalares de Câncer (RHC), fundamentais para compreender a magnitude dessa doença. Tais registros fornecem informações sobre incidência, mortalidade, sobrevida e fatores de risco, orientando estratégias de prevenção, detecção precoce e avaliação de novos tratamentos e métodos diagnósticos (Curado, 2019).

As diretrizes e a qualidade dos dados nos RCBP e RHC seguem o manual de práticas do Instituto Nacional do Câncer (Inca) recomendadas pela Agência Internacional de Pesquisa sobre o Câncer (IARC) (Brasil, Inca, 2010). No Estado de São Paulo, os RHCs são coordenados pela Fundação Oncocentro de São Paulo (Fosp), responsável pela identificação de todos os casos tratados e da qualidade dos casos registrados (Fosp, 2013).

A Fundação Seade realiza mensalmente uma pesquisa nos Cartórios de Registro Civil de todos os municípios do Estado de São Paulo e coleta informações sobre os eventos vitais: nascimentos, óbitos e casamentos. As bases de dados elaboradas integram o Sistema Nacional de Estatísticas do Registro Civil coordenado pelo Instituto Brasileiro de Geografia e Estatística – IBGE. As bases de mortalidade processadas pela Fundação Seade consideram dois tipos de instrumentos de coleta enviados pelos cartórios: arquivos com dados específicos dos registros legais e imagens das declarações de óbito. O processamento desses eventos gera bancos de mortalidade por sexo, idade, causas de morte e uma série de outras variáveis, que estão desagregadas para todos os municípios paulistas e distritos da capital, permitindo estudos minuciosos da população como, por exemplo, o acompanhamento contínuo da mortalidade por determinadas doenças (Waldvogel, 2020).

A disponibilidade de grandes bases de dados informatizadas sobre saúde tornou a técnica de relacionamento de registros uma alternativa viável para diferentes desenhos de estudo, proporcionando a criação de bancos de dados mais completos, com baixo custo operacional (Silva *et al.*, 2021; Fundação Seade, Gerência Demográfica, 2023).

Nesse contexto, o A.C. Camargo Cancer Center e a Fundação Seade firmaram Acordo de Cooperação Técnica-Científica visando a aplicação da técnica de vinculação determinística entre a base de pacientes com câncer do Registro Hospitalar de Câncer (RHC) da instituição e a base de óbitos registrados no Estado de São Paulo e processados pela Fundação Seade.

O presente estudo traz resultados da vinculação determinística adotada para relacionar a base do cadastro de pacientes tratados no A.C. Camargo Cancer Center e diagnosticados entre 2000 e 2021, com a base de óbitos processada no Seade entre 2002 e 2021, apontando os principais tipos de cânceres identificados e os correspondentes tempos de seguimento antes e após a realização desse procedimento.

2. MÉTODOS

O processo de vinculação determinística realizada entre as bases de dados dos parceiros teve início em setembro de 2008, cobrindo os registros de pacientes diagnosticados com câncer entre 2000 e 2021 e os óbitos ocorridos entre janeiro de 2002 e dezembro de 2021.¹ Os casos de óbitos já presentes nas bases do RHC tiveram a data da morte confirmada com a data existente nas bases do Seade, enquanto aqueles pacientes com perda de seguimento, que puderam ser identificados na base de óbitos, foram confirmados como óbito na base original. Nas duas situações – óbitos já identificados e perda

1. A Fundação Seade responde pelo Sistema de Estatísticas Vitais de São Paulo com base na Pesquisa Mensal de Eventos Vitais, que coleta informações, em todos os Cartórios de Registro Civil do Estado, sobre óbitos gerais, óbitos infantis, óbitos fetais, casamentos e nascidos vivos. Sobre a metodologia de produção das estatísticas, ver Waldvogel (2020).

de seguimento dos casos vinculados –, a causa básica da morte foi agregada às bases do RHC.

A metodologia de vinculação dos registros seguiu as etapas de seleção e padronização das variáveis e criação de variáveis derivadas. As variáveis selecionadas foram: nome do indivíduo; nome da mãe; data de nascimento; RG e/ou CPF; e endereço de residência. Na etapa de padronização, para as variáveis nome (paciente e mãe), suprimiram-se letras duplicadas, preposições, caracteres diferentes de A a Z, espaços em branco antes do primeiro nome e qualificações presentes nos campos (exemplo: brasileiro, já falecido, etc.). Os caracteres foram convertidos para maiúsculos e realizou-se a substituição das letras y por i, w por v, th por t, ph por f, oct por ot, etc. Nos campos de documentos foram suprimidos caracteres não numéricos e espaços em branco antes do número. O processo de padronização das variáveis é fundamental para obter maior eficácia na identificação dos casos comuns às duas bases. A criação de variáveis derivadas deu-se pela recodificação das variáveis nome do sujeito e nome da mãe. Foram criados novos campos contendo as primeiras letras do nome e do sobrenome, nome à esquerda, nome à direita e combinações (Fundação Seade, Gerência Demográfica, 2023).

Os critérios de comparação foram por igualdade, em que a combinação de variáveis selecionadas apresenta preenchimento idêntico nas duas bases, e por semelhança, que acontece quando as variáveis selecionadas possuem preenchimento igual ou superior a um nível predefinido. Se o par selecionado mostrar algum tipo de discordância em pelo menos um dos campos comparados, torna-se necessária uma análise visual, que é realizada caso a caso comparando-se as demais informações contidas nas bases. A seleção visual é feita com dupla verificação, ou seja, por duas pessoas que têm suas seleções comparadas e, na ocorrência de seleção discordante, nova análise conjunta é efetuada para a decisão final (Fundação Seade, Gerência Demográfica, 2023).

3. ANÁLISE ESTATÍSTICA

Os casos vinculados entre a base de dados do RHC e a base de mortalidade do Seade foram analisados quanto à frequência absoluta e relativa para as variáveis sexo, faixa etária, estadiamento clínico ao diagnóstico e tipo de câncer.

Em casos de pacientes com múltiplos cânceres primários, dois ou mais tumores primários na mesma região anatômica durante o período do estudo (tumor primário sincrônico ou metacrônico), apenas o primeiro tumor foi incluído na análise.

Análises estatísticas e descritivas foram realizadas para avaliar os casos diagnosticados e identificados com a vinculação determinística das bases, considerando o tempo de seguimento do paciente antes e após esse procedimento. Para tanto, foram calculados os seguintes indicadores:

- tempo de seguimento entre a data do diagnóstico e a data da última informação no RHC **antes da vinculação**;
- tempo de seguimento entre a data de diagnóstico e a data do óbito obtido **após a vinculação** com a base do Seade.

Calculou-se a média de tempo para os 15 tipos de câncer mais frequentes que foram identificados, realizando-se o teste de Wilcoxon (Massad *et al.*, 2004) para medir se havia diferença significativa entre a média de tempo de seguimento pré e pós-vinculação, a um nível de significância de 95%.

Os óbitos no período de 2002 a 2021, com informação das causas de morte, foram agrupados de acordo com a causa básica de óbito em capítulos de I a XXI da Classificação Internacional de Doenças (CID-10) e analisados quanto à frequência absoluta e relativa. Nem todos os óbitos de 2020 e 2021 contavam com a informação sobre a causa básica no momento da realização deste estudo, de modo que não foram analisados para este quesito em especial.

4. RESULTADOS

A vinculação determinística realizada, que contemplou 79.508 pacientes do Registro Hospitalar de Câncer e 5.693.126 registros de óbitos da Fundação Seade, possibilitou a identificação de 12.517 registros presentes nas duas bases de dados no período de 2002 a 2021, que correspondem a 15,7% dos pacientes considerados. Destes, 53% dos casos de câncer atualizados com informação de óbito eram do sexo masculino. A maioria correspondia a pacientes com idades acima de 60 anos (65,2%) e 40,3% dos casos apresentavam estadiamento² clínico inicial (*in situ*, I ou II) (Tabela 1).

Tabela 1 – Distribuição dos pacientes vinculados, segundo sexo, faixa etária e estadiamento clínico

Estado de São Paulo, 2002-2021

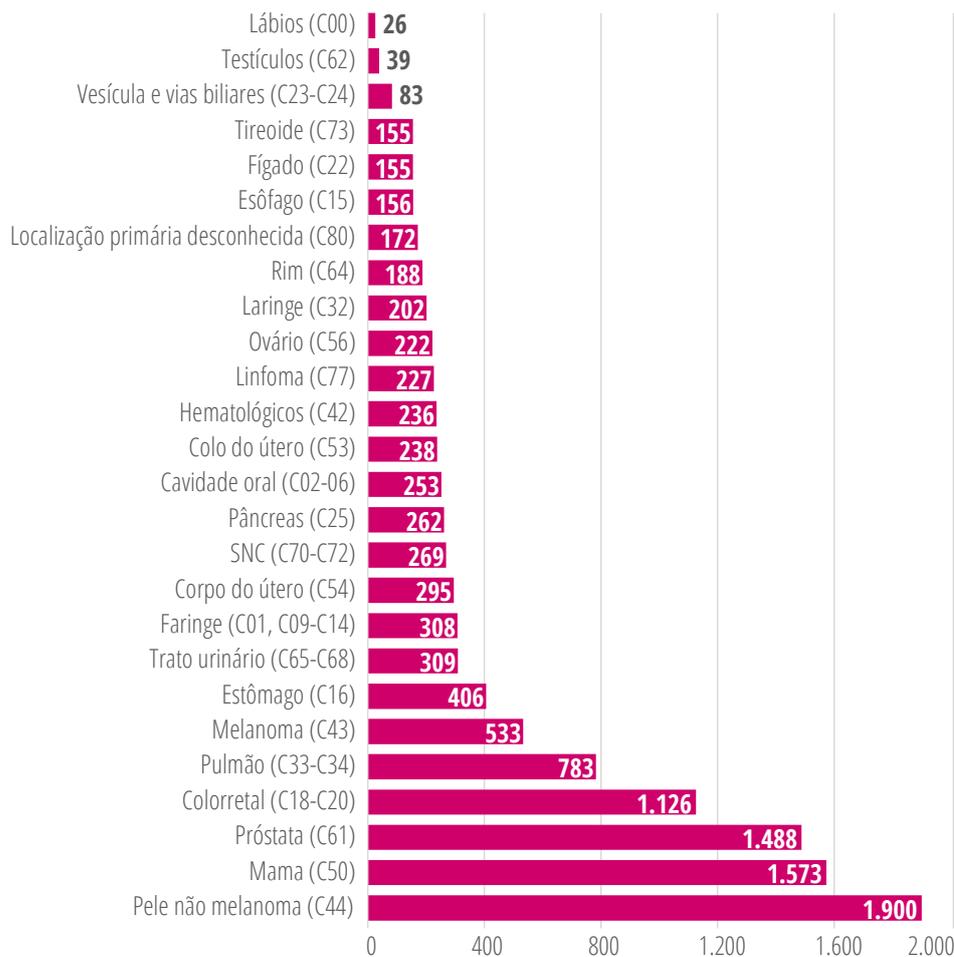
Variáveis	N. abs.	%
Total	12.517	100,0
Sexo		
Homens	6.637	53,0
Mulheres	5.880	47,0
Faixa etária		
Até 19 anos	166	1,3
De 20 a 39 anos	734	5,9
De 40 a 59 anos	3.455	27,6
60 anos ou mais	8.162	65,2
Estadiamento clínico		
0 (<i>in situ</i>)	434	3,5
I	2.412	19,3
II	2.185	17,5
III	1.671	13,3
IV	2.283	18,2
Sem dados	2.185	17,5
Não se aplica	1.347	10,8

Fonte: Parceria Fundação Seade/RHC/A.C.Camargo.

2. O sistema de estadiamento é utilizado para descrever a extensão da doença no corpo e envolve a avaliação do tamanho do tumor, a presença de linfonodos comprometidos e a disseminação para outros órgãos, podendo ser classificado em inicial (*EC in situ*, I ou II) ou estadiamento clínico avançado (EC III e IV). Constitui importante fator prognóstico e é utilizado na definição do plano terapêutico.

Os casos vinculados foram desagregados pelos tipos de câncer mais frequentes, procurando analisar quais tiveram maior frequência de informação adicionada sobre óbito. No período estudado, tiveram as informações atualizadas com o óbito 1.900 pacientes diagnosticados com câncer de pele não melanoma, um tumor de baixa letalidade e com alta chance de cura e sobrevida, 1.573 pacientes com diagnóstico de câncer de mama (C50) e 1.488 com câncer de próstata (C61). Os casos de tumores de testículos e lábios apresentaram o menor número identificado de mortes (Gráfico 1).

Gráfico 1 - Óbitos identificados pela vinculação determinística, segundo tipos de câncer
 Estado de São Paulo, 2002-2021



Fonte: Parceria Fundação Seade/RHC/A.C.Camargo.

Ao estratificar os resultados em tumores de baixa e alta letalidade, observou-se que 53% dos casos atualizados com informação de óbito estavam entre os cinco tipos de câncer mais frequentes com menor letalidade. Em contraste, os tumores de pulmão, estômago, pâncreas, sistema nervoso central e fígado representaram 15% dos casos atualizados com informação de óbito no grupo de alta letalidade (Tabela 2).

O tempo médio de acréscimo de seguimento para os casos de câncer na base de dados do RHC apresentou variação, com média de 47 meses para os casos de câncer de pele não melanoma, 36 meses para câncer de próstata e 28 meses para o câncer de mama e melanoma de pele (Gráfico 2). As menores médias foram observadas para os cânceres de pulmão (sete meses) e pâncreas (4,5 meses). Para os 15 cânceres analisados, observou-se melhora significativa no tempo médio de seguimento após a vinculação entre as bases de dados ($p < 0,05$).

Tabela 2 – Óbitos identificados na vinculação determinística, segundo nível de letalidade dos principais tipos de câncer

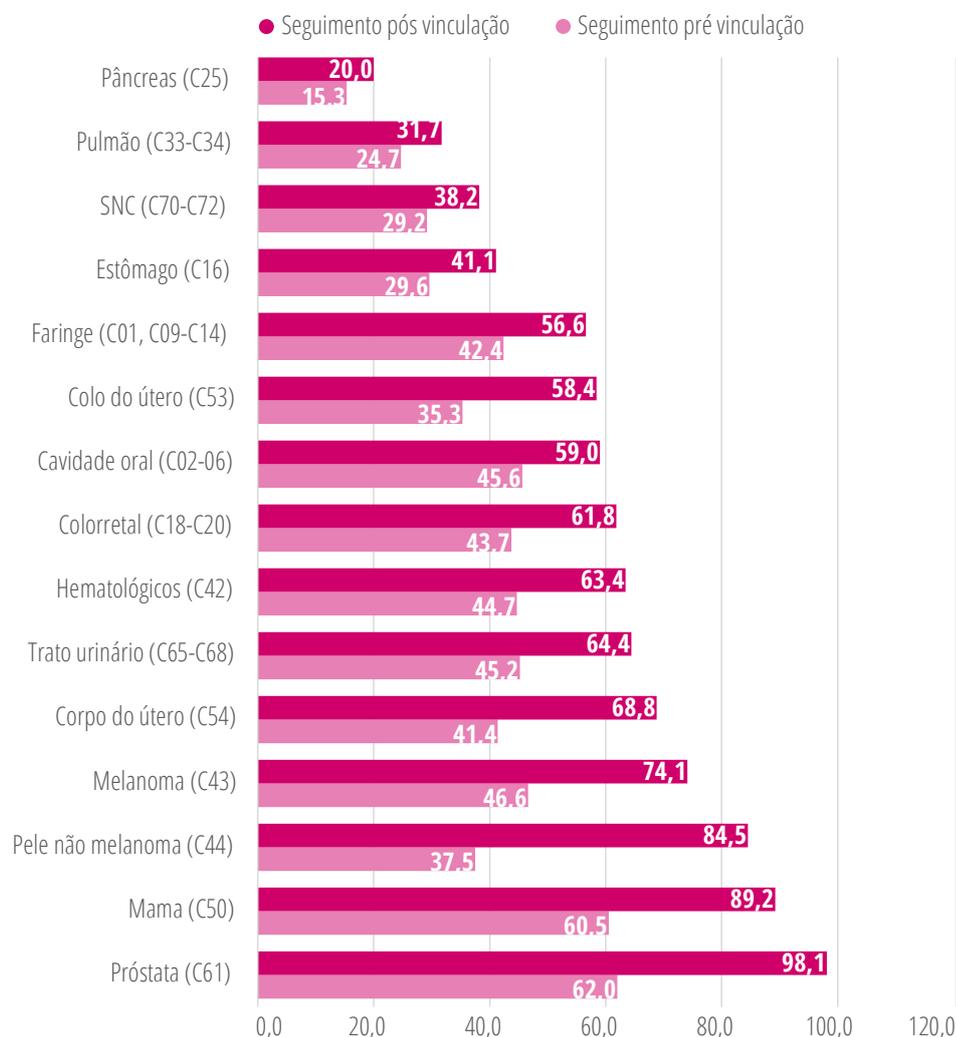
Estado de São Paulo, 2002-2021

Tipos de câncer	N. abs.	%
Baixa letalidade	6.620	53,0
Pele não melanoma (C44)	1.900	15,2
Mama (C50)	1.573	12,6
Próstata (C61)	1.488	11,9
Cólon e reto (C18-C20)	1.126	9,0
Melanoma (C43)	533	4,3
Alta letalidade	1.875	14,9
Pulmão (C33-C34)	783	6,3
Estômago (C16)	406	3,2
Sistema Nervoso Central (C70-C72)	269	2,1
Pâncreas (C25)	262	2,1
Fígado (C22)	155	1,2

Fonte: Parceria Fundação Seade/RHC/A.C.Camargo.

Gráfico 2 – Tempo médio de seguimento dos indivíduos pré e pós a realização da vinculação determinística, por tipo de câncer

Estado de São Paulo, 2002-2021, em meses



Fonte: Parceria Fundação Seade/RHC/A.C.Camargo.

A maior proporção dos casos identificados, entre 2002 e 2021, teve como causa básica do óbito as neoplasias (tumores) do capítulo II, com 6.571 óbitos (65%), seguidas pelas doenças do aparelho circulatório, capítulo IX, com 1.365 óbitos (13,4%), como infarto agudo do miocárdio, e por doenças do aparelho respiratório, com 755 óbitos (7,3%), incluindo broncopneumonia (Tabela 3).

Tabela 3 - Óbitos identificados na vinculação determinística, segundo causas básicas de morte categorizadas por capítulo da Classificação Internacional de Doenças (CID-10)

Estado de São Paulo, 2002-2021

Capítulo	Descrição	Códigos da CID-10	Óbitos n. abs.	%
I	Algumas doenças infecciosas e parasitárias	A00-B99	137	1,2
II	Neoplasias [tumores]	C00-D48	6.571	65,0
III	Doenças do sangue e dos órgãos hematopoiéticos e alguns transtornos imunitários	D50-D89	18	0,1
IV	Doenças endócrinas, nutricionais e metabólicas	E00-E90	145	1,3
V	Transtornos mentais e comportamentais	F00-F99	73	0,6
VI	Doenças do sistema nervoso	G00-G99	203	2,0
VII	Doenças do olho e anexos	H00-H59	0	0,0
VIII	Doenças do ouvido e da apófise mastóide	H60-H95	1	0,0
IX	Doenças do aparelho circulatório	I00-I99	1.365	13,4
X	Doenças do aparelho respiratório	J00-J99	755	7,3
XI	Doenças do aparelho digestivo	K00-K93	254	2,4
XII	Doenças da pele e do tecido subcutâneo	L00-L99	26	0,1
XIII	Doenças do sistema osteomuscular e do tecido conjuntivo	M00-M99	28	0,2
XIV	Doenças do aparelho geniturinário	N00-N99	239	2,3
XIX	Lesões, envenenamentos e algumas outras consequências de causas externas	S00-T98	0	0,0
XV	Gravidez, parto e puerpério	O00-O99	1	0,0
XVI	Algumas afecções originadas no período perinatal	P00-P96	0	0,0
XVII	Malformações congênitas, deformidades e anomalias cromossômicas	Q00-Q99	1	0,0
XVIII	Sintomas, sinais e achados anormais de exames clínicos e de laboratório, não classificados em outra parte	R00-R99	88	0,8
XX	Causas externas de morbidade e de mortalidade	V01-Y98	176	1,5
XXI	Fatores que influenciam o estado de saúde e o contato com os serviços de saúde	Z00-Z99	0	0,0

Fonte: Parceria Fundação Seade/RHC/A.C.Camargo.

5. CONSIDERAÇÕES FINAIS

Houve incremento relevante no tempo de seguimento dos pacientes com câncer tratados no A.C.Camargo Cancer Center após a realização da vinculação determinística com as bases de óbitos processadas pela Fundação Seade. Foi observada maior completude dessa informação para os tipos de câncer de menor letalidade, tendo sido identificadas causas básicas de morte distintas de neoplasias, como doenças do aparelho circulatório e broncopneumonia. Para as neoplasias de maior letalidade, o incremento foi menor, pois a informação sobre o óbito esteve mais presente na base original antes da vinculação realizada.

A colaboração entre o A.C.Camargo Cancer Center e a Fundação Seade ressalta a importância de parcerias institucionais para otimizar a qualidade e a abrangência das informações no monitoramento dos pacientes com câncer. Essa cooperação permitiu superar desafios

inerentes à perda de seguimento, agregar a informação sobre a causa básica do óbito e melhorar a precisão das estimativas futuras de sobrevida.

No presente estudo foram identificados 15,7% de casos do RHC A.C.Camargo Cancer Center vinculados à base de óbitos do Seade, representando melhora significativa no tempo de seguimento dos pacientes acompanhados para os 15 tumores mais frequentes. Os resultados encontrados destacam a importância do procedimento de vinculação determinística adotado para a complementação de informações, reduzindo a perda do tempo de seguimento.

Os cânceres com maior número de óbitos identificados neste estudo foram aqueles com melhor prognóstico, como câncer de pele não melanoma, de mama e de próstata, sendo que 55% das causas básicas de morte agregadas não eram neoplasias. Já os cânceres mais letais, como de pulmão, estômago, pâncreas, sistema nervoso central e fígado, tiveram menor impacto na atualização do *status* vital, em que 78% das causas básicas de morte foram de fato câncer e apenas 12% dos óbitos não foram. Esse resultado deve-se à maior agressividade do tumor e, provavelmente, o maior número de óbitos ocorreu no período de acompanhamento hospitalar, com melhor atualização desse evento no registro hospitalar.

O incremento no tempo de seguimento foi superior à média de 36 meses para câncer de próstata e de pele não melanoma. Tal aumento, após o processo de vinculação com a base de óbitos, deve-se ao maior número de mortes que foram constatadas e à retificação do último contato e das datas do óbito presentes no RHC.

Neste estudo, 65% dos óbitos relacionados entre 2002 e 2021 tiveram como causa básica o câncer. Dentre os pontos fortes deste estudo, destaca-se que o procedimento adotado, vinculação determinística, possibilita a preservação da individualidade dos casos e a maior atualização do *status* vital e do tempo de seguimento dos pacientes. A metodologia contribui para a precisão das análises epidemiológicas e das estimativas de sobrevida.

Assim, a vinculação determinística realizada favoreceu maior acurácia das causas de óbito dos pacientes com neoplasias malignas. Além de identificar relevante incremento no tempo do seguimento após a adoção da vinculação entre as bases de dados, fornecendo uma visão abrangente do panorama dos pacientes com câncer ao longo de quase duas décadas, entre diagnósticos e óbitos. Os resultados destacam o incremento das informações sobre o seguimento dos pacientes e sobre a causa básica da morte, otimizando a qualidade e a abrangência das informações para o monitoramento dos pacientes com câncer.

6. REFERÊNCIAS

BEHERA, P.; PATRO, B. K. Population based cancer registry of India – The challenges and opportunities. *Asian Pac. J. Cancer Prev.*, v. 19, p. 2.885-2.889. APJC, 2018. Disponível em: <https://doi.org/10.22034/apjcp.2018.19.10.2885>.

BRASIL. Ministério da Saúde. Instituto Nacional de Câncer José Alencar Gomes da Silva. *Registros hospitalares de câncer: planejamento e gestão*. 2. ed. Rio de Janeiro: Inca, 2010.

CURADO, M. P. Importance of hospital cancer registries in Africa. *Ecancermedicalscience*, v. 13, Jul. 2019.

ESTADO DE SÃO PAULO. Secretaria da Saúde/CRT-DST/Aids; FUNDAÇÃO SEADE. *Dados para repensar a Aids no Estado de São Paulo: resultados da parceria entre Programa Estadual DST/Aids e Fundação Seade*. São Paulo, 2010.

FUNDAÇÃO ONCOCENTRO DE SÃO PAULO. Fosp. *Registro Hospitalar de Câncer: conceitos, rotinas e instruções de preenchimento*. 2. ed. São Paulo: Fundação Oncocentro de São Paulo, 2013. Disponível em: http://www.fosp.saude.sp.gov.br:443/epidemiologia/docs/ManualRHC_2013.pdf. Acesso em: 4 out. 2020.

FUNDAÇÃO SEADE. *Sigilo e proteção de dados*. Institucional. São Paulo: Fundação Seade, s/d. Disponível em: <https://www.seade.gov.br/institucional/sigilo-protexcao-de-dados/>.

FUNDAÇÃO SEADE. *Painel Estatísticas Vitais – Óbitos*. Disponível em: <https://estatisticasvitais.seade.gov.br/>.

FUNDAÇÃO SEADE. Gerência Demográfica. *Vinculação determinística entre bases de dados*. São Paulo: Fundação Seade, 2023. Disponível em: <https://metodologia.seade.gov.br/wp-content/uploads/sites/4/2023/12/seade-metodologia-vinculacao-deterministica-entre-bases-dados.pdf>

GIL, F. *et al.* Impact of the management and proportion of lost to follow-up cases on cancer survival estimates for small population based cancer registries. *Journal of Cancer Epidemiology*, article 9068214, 2022.

BRENNER, H.; HAKULINEN, T. Implications of incomplete registration of deaths on long-term survival estimates from population-based cancer registries. *International Journal of Cancer*, v. 125, n. 2, p. 432-437, 2009.

HILL, T. *et al.* Data linkage reduces loss to follow-up in an observational HIV cohort study. *Journal of Clinical Epidemiology*, v. 63, n. 10, p. 1.101-1.109, 2010. Disponível em: <https://doi.org/10.1016/j.jclin.epi.2009.12.007>.

LAURENTI, R.; MELLO JORGE, M. H. P.; GOTLIEB, S. L. D. Mortalidade segundo causas: considerações sobre a fidedignidade dos dados. *Rev. Panam. Salud Publica*, v. 23, n. 5, p. 349-356, 2008.

MASSAD, E; MENEZES, R.X.; SILVEIRA, P.S.P.; ORTEGA, N.R.S. *Métodos quantitativos em medicina*. São Paulo: Manole, 2004.

MORAIS, L. C. C. *et al.* Vinculação de casos de Aids com óbitos no Estado de São Paulo. *In: Seminário Internacional de Linkage de Bases de Dados. Anais [...]*. Rio de Janeiro: Instituto Nacional de Cardiologia, 2010.

OKUYAMA, A.; BARCLAY, M.; CHEN, C.; HIGASHI, T. Impact of loss-to-follow-up on cancer survival estimates for small populations: a simulation study using Hospital-Based Cancer Registries in Japan. *BMJ Open*, v. 10, n. 1, article e033510, 2020.

PEREIRA, A. P. E.; DA GAMA, S. G. N.; DO CARMO LEAL, M. Mortalidade infantil em uma amostra de nascimentos do município do Rio de Janeiro, 1999-2001: "linkage" com o Sistema de Informação de Mortalidade. *Revista Brasileira de Saúde Materno Infantil*, v. 7, n. 1, p. 83-88, 2007. Disponível em: <https://doi.org/10.1590/S1519-3829200700100010> (2007).



Governador do Estado
Tarcísio de Freitas

Vice-Governador do Estado
Felício Ramuth

Secretário da Fazenda e Planejamento
Samuel Kinoshita

SEADE

Presidente do Conselho Curador
Carlos Antonio Luque

Diretor Executivo
Bruno Caetano

**Diretor-adjunto de
Produção e Análise de Dados**
Carlos Eduardo Torres Freire

**Diretor-adjunto de
Comunicação e Informação**
Marcelo Moreira

**Diretor-adjunto Administrativo
e Financeiro**
Luiz Ricardo Santoro

Chefe de Gabinete
Sérgio Meirelles Carvalho

SP DEMOGRÁFICO

A série SP Demográfico, iniciada em 1998, procura veicular os principais indicadores demográficos do Estado de São Paulo, de suas regiões, municípios e distritos da capital, com ênfase na análise das projeções populacionais e das Estatísticas do Registro Civil, produzidas pela Fundação Seade.

Coordenação e edição
Bernadette Cunha Waldvogel

Corpo editorial

Bernadette Cunha Waldvogel; Carlos Eduardo Torres Freire; Carlos Eugênio de Carvalho Ferreira e Valmir José Aranha

Autores deste número

Diego Rodrigues Mendonça e Silva, Bernadette Cunha Waldvogel, Magaly de Losso Perdigão, Rosa Maria Vieira de Freitas, Lillian Cristina Correia Morais e Maria Paula Curado

Assessoria de Editoração e Arte

Responsável técnico: Paulo Emirandetti Junior
Equipe técnica: Cristiane de Rosa Meira, Elisabeth Magalhães Erharter, Maria Aparecida Batista de Andrade, Rita Bonizzi e Vania Regina Fontanesi

Endereço para correspondência

Av. Professor Lineu Prestes, 913 - Cidade Universitária
05508-000 - São Paulo - SP
Fone (11) 3324.7200
seade.gov.br
sicseade@seade.gov.br
ouvidoria@seade.gov.br

PEREIRA, T. F.; ARANHA, V. J.; WALDVOGEL, B. C.; DA COSTA, A. M.; TAVARES GUERREIRO FREGNANI, J. H. Deterministic linkage for improving follow-up time in a Brazilian population-based cancer registry. *Scientific Reports*, v. 13, Article 4816, Mar. 2023. Disponível em: <https://doi.org/10.1038/s41598-023-31303-6>.

PERES, S. V. *et al.* Melhora na qualidade e completude da base de dados do Registro de Câncer de Base Populacional do município de São Paulo: uso de técnicas de linkage. *Revista Brasileira de Epidemiologia*, v. 19, n. 4, p. 753-765, São Paulo: Assoc. Brasileira de Saúde Coletiva, 2016. Disponível em: <https://doi.org/10.1590/1980-5497201600040006>.

SILVA, D. R. M. E.; LUIZAGA, C. T. M.; TOPORCOV, T. N.; ALGRANTI, E. Agreement and validity of asbestos-related cancers in the hospital information system of the Brazilian Unified Health System. *Revista Brasileira de Epidemiologia*, v. 24, e210044. São Paulo: Assoc. Brasileira de Saúde Coletiva, 2021.

SUZUKI, K. M. F. *O uso de método de relacionamento de dados (record linkage) para integração de informação em sistemas heterogêneos de saúde: estudo de aplicabilidade entre níveis primário e terciário.* Tese (Doutorado) – Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto, Ribeirão Preto, 2012.

SWAMINATHAN, R.; RAMA, R.; SHANTA, V. Lack of active follow-up of cancer patients in Chennai, India: implications for population-based survival estimates. *Bull World Health Organ*, v. 86, n. 7, p. 509-515, 2008.

WALDVOGEL, B. C.; FERREIRA, C. E. de C. Estatísticas da vida. *São Paulo em Perspectiva*, v. 17, n. 3-4, p. 55-66. São Paulo: Fundação Seade, 2003. Disponível em: <https://doi.org/10.1590/S0102-88392003000300007>.

WALDVOGEL, B. C. *Produção das estatísticas do Registro Civil no Estado de São Paulo.* São Paulo: Fundação Seade, nov.2020. (Seade Metodologia). Disponível em: https://metodologia.seade.gov.br/wpcontent/uploads/sites/4/2021/05/Metodologia_Estatisticas_Registro_Civil.pdf.

WALDVOGEL, B. C. *et al.* Pesquisa pioneira recupera casos de Aids no Estado de São Paulo: integração das bases do Sinan-Aids e do Seade. *Boletim Epidemiológico*, n. 1, out. 2004.

WALDVOGEL, B. C. Fundação Seade aprimora linha de pesquisa sobre relacionamento de registros administrativos adotando metodologia de vinculação determinística de bases de dados. *Boletim de Estatísticas Públicas*, n. 12, p. 63-65, dez. 2016.

WALDVOGEL B. C. *et al.* Experiência da Fundação Seade com a aplicação da metodologia de vinculação determinística de bases de dados. *Ensaio & Conjuntura.* São Paulo: Fundação Seade, mar.2019. Disponível em: https://produtos2.seade.gov.br/produtos/midia/2019/04/Ensaio_conjuntura_Vinculacao.pdf.