

Estado da publicação: Não informado pelo autor submissor

O que aprendemos diante dos dados que não temos? O caso da variável profissão/ocupação nos bancos de síndrome gripal, síndrome respiratória aguda grave e mortalidade no Brasil

Cleber Santos, Vanessa Melo Ferreira, João Roberto Cavalcante, Patricia Canto Ribeiro, Hermano Albuquerque de Castro, Adriana Coser Gutierrez, Ingrid D'avilla Freire Pereira, Marcio Fernandes Nehab, Maria Martha Duque de Moura, André Reynaldo Santos Perissé

<https://doi.org/10.1590/SciELOPreprints.3560>

Submetido em: 2022-02-03

Postado em: 2022-02-09 (versão 1)

(AAAA-MM-DD)

Nota de pesquisa

Título em português:

O que aprendemos diante dos dados que não temos? O caso da variável profissão/ocupação nos bancos de síndrome gripal, síndrome respiratória aguda grave e mortalidade no Brasil

Título em inglês:

What can we learn with the data we didn't have? The case of the variable profession/occupation on flu-like syndrome, severe acute respiratory syndrome and mortality databases in Brazil

Título em espanhol:

¿Qué podemos aprender con los datos que no tenemos? El caso de la variable profesión /ocupación sobre base de datos de la síndrome gripal, síndrome respiratorio agudo severo y de mortalidad en Brasil

Título resumido em português:

O que podemos aprender com os dados que não temos? O caso da variável profissão/ocupação

Autores:

Cleber Vinicius Brito dos Santos^{1,2†} (e-mail: cleber.vini@yahoo.com.br ; ORCID: 0000-0001-5710-2866)

Vanessa de Melo Ferreira^{1,2†} (e-mail: vanessademelo.ferreira@gmail.com ; ORCID: 0000-0002-7308-0676)

João Roberto Cavalcante^{1,2} (e-mail: joao.rcs@hotmail.com ; ORCID: 0000-0003-2070-3822)

Patricia Canto Ribeiro^{1,3} (e-mail: patriacanto.ribeiro@gmail.com; ORCID: 0000-0002-2731-331X)

Hermano Albuquerque de Castro^{1,3} (e-mail: hermanocastro10@gmail.com ; ORCID: 000-0003-1191-5671)

Adriana Coser Gutierrez^{1,3} (e-mail: adriana.coser@fiocruz.br; ORCID: 0000-0002-7305-5841)

Ingrid D'avilla Freire Pereira^{1,3} (e-mail: ingrid.davilla@fiocruz.br; ORCID: 0000-0002-2042-2871)

Marcio Fernandes Nehab^{1,3} (e-mail: marcio.nehab@gmail.com; ORCID: 0000-0002-4659-6247)

Maria Martha Duque de Moura^{1,3} (e-mail: martha@iff.fiocruz.br; ORCID: 0000-0002-4853-6491)

André Reynaldo Santos Perissé^{1,3} (e-mail: aperisse41@gmail.com ; ORCID: 0000-0002-5253-5774)

Afiliação:

1. Grupo de Trabalho de Retorno Escolar, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

2. Departamento de Epidemiologia, Instituto de Medicina Social, Universidade Estadual do Rio de Janeiro, Rio de Janeiro, Brasil.

3. Escola Nacional de Saúde Pública Sergio Arouca. Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

† Os autores contribuíram igualmente

Correspondência:

Rua São Francisco Xavier 524, bloco E, 7o andar, Rio de Janeiro, RJ 20550-013, Brasil.

E-mail: cleber.vini@yahoo.com.br; vanessademelo.ferreira@gmail.com

Telefone: (21) 96926-8835; (21) 98209-4486

Fonte de financiamento: Nenhum.

Conflito de Interesses: Nenhum.

RESUMO

Objetivo: descrever o preenchimento da variável profissão/ocupação nos bancos de dados do Sistema de Informação de Mortalidade (SIM), de Síndrome Respiratória Aguda Grave (SRAG) e Síndrome Gripal (SG) entre 2020 e 2021 no Brasil.

Métodos: foram calculadas as frequências absolutas e relativas do preenchimento da variável profissão/ocupação por Unidade da Federação para cada base de dados durante o período de estudo.

Resultados: Para o período de estudo, verificou-se 94,83% de incompletude para as notificações de SG, 97,73% para casos notificados de SRAG e 17,06% para óbitos por todas as causas.

Conclusões: Recomendamos uma articulação entre o Ministério da Saúde, Ministério do Trabalho e Previdência e representações de trabalhadores, como conselhos e sindicatos para formulação de estratégias para contornar o problema da falta de dados de ocupação/profissão nos bancos de dados públicos.

Palavras-chave: Dados faltantes, Vigilância, Saúde do trabalhador, Pandemia.

ABSTRACT

Objective: To describe the completeness of the profession/ occupation on flu-like syndrome, severe acute respiratory syndrome and mortality databases in Brazil between 2020 and 2021

Methods: the absolute and relative frequencies of the completeness of the variable occupation were calculated for each state for the between 2020 and 2021.

Results: For the study period, there was 94.83% incompleteness for notifications of Flu-like syndrome, 97.73% for notified cases of SARS and 17.06% for deaths from all causes.

Conclusions: We recommend an articulation between the Ministry of Health, Ministry of Labor and Welfare and workers' representations, such as councils and unions, to solve the problem of lack of data on occupation/profession in public databases.

Keywords: Missing data, Surveillance, Worker's health, Pandemic.

RESUMEN

Objetivo: describir la completitud de la variable profesión/ocupación en las bases de datos de la síndrome gripal, síndrome respiratorio agudo severo y de mortalidad, entre 2020 y 2021, en Brasil.

Métodos: las frecuencias relacionadas con el llenado de la variable profesión / ocupación por Unidad de Federación se calcularon para cada base de datos durante el período de investigación.

Resultados: Para el período de estudio, hubo 94,83% de incompletitud para las notificaciones de síndrome gripal, 97,73% para los casos notificados de SRAS y 17,06% para mortalidad por todas las causas.

Conclusiones: Recomendamos una articulación entre el Ministerio de Salud, el Ministerio de Trabajo y Bienestar y las representaciones de los trabajadores, como consejos y sindicatos, para resolver el problema de la falta de datos sobre ocupación / profesión en bases de datos públicas.

Palabras-clave: Datos faltantes, Vigilancia, Salud del trabajador, Pandemia.

INTRODUÇÃO

A ocorrência de dados faltantes (*missing data*) é um problema comum em pesquisas científicas, especialmente na epidemiologia (1). A perda de dados é um grande desafio para o planejamento e análise de estudos epidemiológicos, onde, comumente o objetivo é investigar preditores que contribuam para prever a ocorrência ou não de um agravo numa população. Perda de informações, tanto nos preditores quanto no desfecho, pode levar a problemas sérios na análise dos dados, podendo-se destacar viés de colisão (*collider bias*) (também chamado de viés de seleção, viés de amostragem ou paradoxo de *Berkson*) (2).

O impacto da ausência da notificação de óbitos, impõe barreiras que comprometem a qualidade e a cobertura das bases de dados oriundas dos sistemas de informação do país. Ao se considerar a saúde de trabalhadores, tais fragilidades impactam diretamente na completude das informações que poderiam ser extraídas para auxiliar estratégias eficientes e continuadas de prevenção de agravos ocupacionais (3).

Durante Emergências em Saúde Pública essa situação se complica ainda mais. Caso não tivéssemos um preenchimento adequado dos nascimentos de crianças com microcefalia e/ou alterações no sistema nervoso central no Sistema de Informações sobre Nascidos Vivos (SINASC), não teríamos descoberto o aumento desses casos em 2015 e 2016 e, posteriormente, a associação desses casos com o Zika vírus, por exemplo (4).

Durante a pandemia de COVID-19, por sua vez, informações sobre o perfil sociodemográfico e socioeconômico das populações mais atingidas são fundamentais para o planejamento e gestão da emergência. Sobretudo em um cenário de recrudescimento das desigualdades socioeconômicas, no qual as atividades laborais tem mobilizado intenso fluxo de pessoas — a despeito das recomendações de distanciamento

social — e cada indivíduo se torna um potencial agente transmissor do vírus SARS-CoV-2, seja nos transportes públicos, no ambiente de trabalho ou até mesmo dentro do próprio domicílio (5,6).

Portanto, não saber a profissão/ocupação dos casos leves, casos graves e óbitos, se torna um problema para atuação e intervenção do setor saúde, principalmente no que diz respeito à saúde do trabalhador dado o impacto da COVID-19 enquanto doença relacionada ao trabalho (6,7).

O presente estudo teve por objetivo descrever a incompletude da variável profissão/ocupação, nas fichas de notificação de Síndrome Gripal (SG), de síndrome respiratória aguda grave (SRAG) e de mortalidade por todas as causas do Sistema de Informação de Mortalidade (SIM), durante os anos de 2020 e 2021.

MÉTODOS

Foram coletados os dados de notificação de SG do eSUS VE Notifica e SRAG do Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe) e de óbitos por todas as causas do Sistema de Informações sobre Mortalidade (SIM) em todo território brasileiro para os anos de 2020 e 2021 (8–12).

Cada conjunto de dados foi agrupado por Unidade da Federação e pelas variáveis correspondentes a profissão/ocupação. Posteriormente, foram calculadas as frequências absolutas e relativas de dados faltantes da variável profissão/ocupação para cada estado dentro do período de estudo.

A presente investigação foi realizada a partir de dados secundários de domínio público disponibilizados de maneira agregada, sendo assim, não houve necessidade de aprovação por um Comitê de Ética em Pesquisa.

RESULTADOS

Para o período de janeiro de 2020 a dezembro de 2021, pode-se verificar que em todo o Brasil houve altas porcentagens de não preenchimento da variável profissão/ocupação (Tabela 1). Observando-se 94,83% de incompletude para as notificações de SG e 97,73% para casos notificados de SRAG. Como exceção, destacou-se as notificações de mortalidade por todas as causas, com 17,06%.

Observando separadamente os bancos de dados nas unidades federativas, para o banco de dados de SG o estado com menor porcentagem de incompletude foi o Espírito Santo (89,94% de dados faltantes) e o estado com o pior preenchimento foi o Paraná (99,48% de dados ausentes). Para o banco de SRAG o estado com melhor preenchimento foi Roraima (78,17% de dados faltantes) e com o menor preenchimento foi o estado do Rio Grande do Sul (99,37%). Por último, quanto aos dados provenientes do SIM, o estado com a menor porcentagem de incompletude foi o Paraná (7,37%) e com a maior incompletude foi o estado de Alagoas (45,32%).

Tabela 1 — Porcentagem de incompletude da variável profissão/ocupação nos bancos de dados de Síndrome Gripal, Síndrome Respiratória Aguda Grave e mortalidade por todas as causas segundo unidades federativas, 2020-2021, Brasil (continua)

UF	SÍNDROME GRIPAL			SÍNDROME RESPIRATÓRIA AGUDA GRAVE			ÓBITOS POR TODAS AS CAUSAS		
	Casos com CBO	Total de casos	Incompletude (%)	Casos com CBO	Total de casos	Incompletude (%)	Óbitos com CBO	Total de óbitos	Incompletude (%)
AC	25294	304795	91,70	106	6777	98,43	5170	8731	40,78
AL	36714	815418	95,50	294	32695	99,10	22300	40785	45,32
AM	74418	1217509	93,89	1887	49142	96,16	39974	46316	13,69
AP	5081	263943	98,07	1120	6628	83,10	5330	8017	33,51
BA	305536	4029150	92,42	2498	104891	97,61	139917	180220	22,36
CE	145512	2453507	94,07	3202	115489	97,22	103097	118532	13,02
DF	61797	851904	92,75	544	59988	99,09	25876	29588	12,54
ES	92581	920445	89,94	505	23917	97,88	29583	49388	40,10
GO	86830	2760417	96,85	3417	93664	96,35	76727	91739	16,36
MA	65757	1229813	94,65	1410	34936	95,96	59103	73094	19,14
MG	401897	7593232	94,71	9418	311153	96,97	246860	279210	11,58
MS	123814	1428931	91,34	404	48882	99,17	33485	37322	10,28
MT	37030	935981	96,04	2752	51808	94,68	38606	44559	13,35
PA	66544	1431014	95,35	2261	69926	96,76	68699	88776	22,61
PB	97149	1590155	93,89	1121	44935	97,50	30801	53934	42,89

PE	174404	2623736	93,35	1636	110636	98,52	114253	131117	12,86
PI	46090	973935	95,27	698	31027	97,75	32935	41071	19,80
PR	9469	1813397	99,48	3102	202565	98,46	151260	163301	7,37
RJ	265077	5450840	95,14	3391	254899	98,66	237740	298117	20,25
RN	118202	1491007	92,07	1158	27165	95,73	30812	41855	26,38
RO	56092	938699	94,02	806	21894	96,31	16443	20425	19,49
RR	18278	409500	95,54	1179	5403	78,17	5254	6649	20,98
RS	401319	6958595	94,23	963	154234	99,37	133776	172288	22,35
SC	158935	7164836	97,78	3177	100368	96,83	74526	86252	13,59
SE	36291	596033	93,91	663	27863	97,62	23761	27281	12,90
SP	1074651	20809101	94,84	15457	833760	98,14	564429	648319	12,93
TO	30869	554377	94,43	1260	20686	93,90	14777	17235	14,26
Brasil	4015631	77610270	94,83	64429	2845331	97,73	2325494	2804121	17,06

DISCUSSÃO

No presente estudo, foi observada uma alta proporção de falta de preenchimento da variável profissão/ocupação em duas importantes bases de dados pesquisadas, SG e SRAG. Foi também observado que os altos índices de dados faltantes ocorrem de maneira generalizada em todas as unidades federativas.

Apesar da abundância de métodos para lidar com a ausência de informações em bancos de dados, os níveis elevados de ausência em todas as camadas impossibilitam a inferência de quaisquer relações entre a ocupação e o risco à saúde (13–16).

Dados precisos e confiáveis sobre doenças ocupacionais são considerados o ponto de partida essencial para o desenvolvimento de programas de prevenção de agravos ocupacionais (17). Esses dados nos informam sobre quem ficou doente e também como, quando e onde ocorreu o agravo. Esses dados, portanto, têm o potencial de identificar alvos para medidas preventivas e de controle. Adicionalmente, tais informações também têm um potencial considerável como ferramentas na avaliação do progresso no cumprimento das metas de prevenção já estabelecidas por lei.

Ao observarmos os dados individualmente, verifica-se que o banco de mortalidade proveniente do SIM, talvez por se basear no registro de óbito e ser a principal fonte de dados sobre mortalidade no âmbito das estatísticas vitais no Brasil, apresenta menores valores de ausência de preenchimento da variável ocupação. Óbitos, por se tratar de um desfecho duro, ou definitivo (do inglês *hard endpoint*), são eventos bem definidos e pouco passíveis de subjetividade e viés de aferição. Dessa forma, naturalmente se é esperado um melhor preenchimento dos dados individuais da ficha de notificação (18).

Especificamente no cenário da pandemia da COVID-19, o Brasil apresenta uma das menores taxas de testagem do mundo (19). Alguns trabalhos utilizam dados de taxa mortalidade geral como forma de quantificar a carga da doença a nível nacional com uso

de dados de séries históricas de mortalidade, com o cálculo do excesso de mortalidade, porém tais métodos se tornam impraticáveis para a quase totalidade das classes profissionais (20,21).

Outra limitação para utilização de dados relacionados à saúde do trabalhador é a impossibilidade do cálculo da taxa de ocorrência de agravos para cada ocupação. Como previamente exposto, para a quantidade de óbitos é possível obter dados precisos sobre o número de casos, ou seja, o numerador da taxa. No entanto, a interpretação desses dados costuma ser difícil na ausência de dados confiáveis sobre o número de trabalhadores sob risco, quantitativo essencial para compor seu denominador. Normalmente, utiliza-se dados de censo ou pesquisas sobre a força de trabalho para fornecer estimativas do número de trabalhadores em grupos ocupacionais, industriais e sociodemográficos. Na prática, tendemos a nos concentrar nas dificuldades de obter dados precisos do numerador e perdemos de vista o fato de que os dados do denominador, advindos do censo, por exemplo, são frequentemente desatualizados — e também inadequados para alguns grupos ocupacionais, como trabalhadores sazonais, trabalhadores informais ou temporários. De tal forma, estimativas podem não ser confiáveis (22,23).

Em algumas situações, dados de várias fontes podem ser combinados para desenvolver denominadores mais apropriados para uma análise particular (24,25). No entanto, deve-se reconhecer que, embora seja possível obter dados precisos para alguns resultados, as limitações nos dados do denominador disponíveis podem introduzir vies e erros aleatórios no cálculo das taxas.

Embora cada uma das fontes de dados utilizadas forneça informações valiosas, cada uma compreende apenas um subconjunto de casos e, portanto, só pode descrever as taxas de doenças ocupacionais de sua perspectiva particular.

Em princípio, a ligação (do termo inglês *linkage*) entre arquivos de dados fornece um meio de combinar dados de duas ou mais fontes e, assim, extrair mais informações valiosas. Ao vincular os registros de um indivíduo em conjuntos de dados, as informações dessa pessoa podem ser reunidas em um único arquivo. Isso permitiria que as pessoas representadas em vários conjuntos de dados fossem identificadas, assim como forneceria informações mais abrangentes sobre esses casos. Essa melhoria já foi observada para doenças como tuberculose (26). Esta abordagem fornece um mecanismo para fazer o uso mais eficaz dos dados disponíveis. Na prática, no entanto, as oportunidades de vincular arquivos são severamente limitadas pela falta de um identificador pessoal exclusivo nos arquivos de dados, além de fragilidades em relação à privacidade dos usuários (27,28).

Em cenário pandêmico, como o atual da COVID-19, onde o desafio está nas medidas de distanciamento social com restrições de atividades produtivas, a ausência de informações sobre a saúde dos trabalhadores e o impacto da doença em grupos específicos, como profissionais de saúde e da educação, dificultam as orientações de políticas públicas, tanto para restrição, quanto para retorno das atividades.

Por fim, ressaltamos que é essencial a união de esforços por parte do Ministério do Trabalho e Previdência, Ministério da Saúde e representações de grupos ocupacionais, como sindicatos para articular, em conjunto, na formulação de soluções para o problema de falta de dados de saúde do trabalhador.

CONTRIBUIÇÃO DOS AUTORES

ACG: contribuiu para concepção do estudo, investigação, redação do manuscrito (revisão).

ARSP: contribuiu para concepção do estudo, investigação, metodologia, supervisão e redação do manuscrito (revisão e edição).

CVBS: contribuiu para concepção do estudo, construção dos bancos de dados, análise dos dados, investigação, metodologia e redação do manuscrito (rascunho original, revisão e edição).

HAC: contribuiu para concepção do estudo, investigação e redação do manuscrito (revisão e edição).

IDFP: contribuiu para concepção do estudo, investigação, redação do manuscrito (revisão).

JRC: contribuiu para concepção do estudo, construção dos bancos de dados e redação do manuscrito (revisão e edição)

MFN: contribuiu para concepção do estudo, investigação, redação do manuscrito (revisão).

MMDM: contribuiu para concepção do estudo, investigação, redação do manuscrito (revisão).

PCR: contribuiu para concepção do estudo, investigação e redação do manuscrito (revisão e edição).

VMF: contribuiu para concepção do estudo, construção dos bancos de dados, análise dos dados, investigação, metodologia e redação do manuscrito (rascunho original, revisão e edição).

Todos os autores participaram da redação do manuscrito, aprovaram sua versão final e afirmam serem responsáveis por todos os aspectos do trabalho, garantindo sua precisão e integridade.

REFERÊNCIAS

1. Rubin DB. Multiple Imputation After 18+ Years. *Journal of the American Statistical Association*. 1996;91(434):473–89.
2. Heijden GJMG van der, Donders ART, Stijnen T, Moons KGM. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: A clinical example. *Journal of Clinical Epidemiology*. 1o de outubro de 2006;59(10):1102–9.
3. Batista AG, Santana VS, Ferrite S. Registro de dados sobre acidentes de trabalho fatais em sistemas de informação no Brasil. *Ciênc saúde coletiva*. março de 2019;24(3):693–704.
4. Marinho F, Araújo VEM de, Porto DL, Ferreira HL, Coelho MRS, Lecca RCR, et al. Microcefalia no Brasil: prevalência e caracterização dos casos a partir do Sistema de Informações sobre Nascidos Vivos (Sinasc), 2000-2015. *Epidemiol Serv Saúde*. 26 de setembro de 2016;25:701–12.
5. Demenech LM, Dumith S de C, Vieira MECD, Neiva-Silva L. Income inequality and risk of infection and death by COVID-19 in Brazil. *Rev bras epidemiol*. 5 de outubro de 2020;23:e200095.
6. Maeno M. COVID-19 como uma doença relacionada ao trabalho. *Rev bras saúde ocup* [Internet]. 13 de dezembro de 2021 [citado 1o de fevereiro de 2022];46. Disponível em: <http://www.scielo.br/j/rbso/a/YfgLv9nQNCSTqRdNjXVQnPJ/?lang=pt>
7. Medeiros LCA de, Borges M da CAL, Gonsalves MP, Teodózio GC, Santos WM dos. Perfil sócio demográfico dos pacientes acometidos pela COVID- 19. *Environmental Smoke*. 31 de agosto de 2021;4(2):42–8.
8. Ministério da Saúde. Notificações de Síndrome Gripal - Open DataSUS [Internet]. 2022 [citado 1o de fevereiro de 2022]. Disponível em: <https://opendatasus.saude.gov.br/dataset/notificacoes-de-sindrome-gripal>

9. Ministério da Saúde. SRAG 2020 - Banco de Dados de Síndrome Respiratória Aguda Grave - incluindo dados da COVID-19 - Open DataSUS [Internet]. 2021 [citado 10 de setembro de 2021]. Disponível em: <https://opendatasus.saude.gov.br/dataset/bd-srag-2020>
10. Ministério da Saúde. SRAG 2021 e 2022 - Banco de Dados de Síndrome Respiratória Aguda Grave - incluindo dados da COVID-19 - Open DataSUS [Internet]. 2022 [citado 1o de fevereiro de 2022]. Disponível em: <https://opendatasus.saude.gov.br/dataset/srag-2021-e-2022>
11. Ministério da Saúde. Sistema de Informação sobre Mortalidade – SIM - Open DataSUS [Internet]. 2021 [citado 31 de agosto de 2021]. Disponível em: <https://opendatasus.saude.gov.br/dataset/sistema-de-informacao-sobre-mortalidade>
12. Ministério da Saúde. Sistema de Informação sobre Mortalidade (SIM) - Dados Abertos - Centrais de Conteúdos - DASNT - SVS/MS [Internet]. 2022 [citado 1o de fevereiro de 2022]. Disponível em: <http://svs.aids.gov.br/dantps/centrais-de-conteudos/dados-abertos/sim/>
13. Santin G, Geoffroy B, Bénézet L, Delézire P, Chatelot J, Sitta R, et al. In an occupational health surveillance study, auxiliary data from administrative health and occupational databases effectively corrected for nonresponse. *Journal of Clinical Epidemiology*. 1o de junho de 2014;67(6):722–30.
14. Bang H, Robins JM. Doubly Robust Estimation in Missing Data and Causal Inference Models. *Biometrics*. 2005;61(4):962–73.
15. Baumgarten M, Oseasohn R. Studies on occupational health: a critique. *J Occup Med*. março de 1980;22(3):171–6.
16. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Stat Methods Med Res*. 1o de junho de 2013;22(3):278–95.

17. Feyer A-M, Williamson AM. The Role of Work Practices in Occupational Accidents. Proceedings of the Human Factors Society Annual Meeting. 1o de setembro de 1991;35(15):1100–4.
18. Jorge MHP de M, Laurenti R, Gotlieb SLD. Análise da qualidade das estatísticas vitais brasileiras: a experiência de implantação do SIM e do SINASC. Ciênc saúde coletiva. junho de 2007;12(3):643–54.
19. Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Giattino C, Ortiz-Ospina E, et al. Coronavirus Pandemic (COVID-19). Our World in Data [Internet]. 5 de março de 2020 [citado 12 de novembro de 2021]; Disponível em: <https://ourworldindata.org/coronavirus>
20. Silva GA e, Jardim BC, Santos CVB dos. Excesso de mortalidade no Brasil em tempos de COVID-19. Ciênc saúde coletiva. setembro de 2020;25(9):3345–54.
21. Orellana JDY, Cunha GM da, Marrero L, Moreira RI, Leite I da C, Horta BL. Excesso de mortes durante a pandemia de COVID-19: subnotificação e desigualdades regionais no Brasil. Cad Saúde Pública. 2021;37(1):e00259120.
22. Macaskill P, Driscoll TR. National occupational injury statistics: what can the data tell us? In: Occupational Injury. CRC Press; 1998.
23. Macaskill P, Mandryk JA, Leigh J. Improved coverage and relevance of national occupational health and safety surveillance in Australia. Epidemiology and Surveillance Unit. Worksafe Australia de 1995;
24. Corvalan CF, Driscoll TR, Harrison JE. Role of migrant factors in work-related fatalities in Australia. Scandinavian Journal of Work, Environment & Health. 1994;20(5):364–70.
25. Driscoll TR, Ansari G, Harrison JE, Frommer MS, Ruck EA. Traumatic work related fatalities in commercial fishermen in Australia. Occup Environ Med. setembro de 1994;51(9):612–6.

26. Bartholomay P, Oliveira GP de, Pinheiro RS, Vasconcelos AMN. Melhoria da qualidade das informações sobre tuberculose a partir do relacionamento entre bases de dados. *Cad Saúde Pública*. novembro de 2014;30(11):2459–70.
27. Sibthorpe DB, Kliewer E, Smith L. Record linkage in Australian epidemiological research: health benefits, privacy safeguards and future potential - Sibthorpe - 1995 - *Australian Journal of Public Health* - Wiley Online Library. *Australian Journal of Public Health*. junho de 1995;19(3):250–6.
28. Pool J, Fatehi F, Hassandoust F, Akhlaghpour S. Health Data Privacy: Research Fronts, Hot Topics and Future Directions. *Integrated Citizen Centered Digital Health and Social Care*. 2020;167–71.

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença [Creative Commons CC-BY](#).
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.