

Tuberculose e diabetes: relacionamento probabilístico de bases de dados para o estudo da associação entre ambas doenças*

doi: 10.5123/S1679-49742017000200013

Tuberculosis and diabetes: probabilistic linkage of databases to study the association between both diseases

Ricardo Gadelha de Abreu¹

Artur Iuri Alves de Sousa²

Maria Regina Fernandes de Oliveira³

Mauro Niskier Sanchez⁴

¹Ministério da Saúde, Secretaria de Vigilância em Saúde, Brasília-DF, Brasil

²Ministério da Saúde, Agência Nacional de Vigilância Sanitária, Brasília-DF, Brasil

³Universidade de Brasília, Núcleo de Medicina Tropical, Brasília-DF, Brasil

⁴Universidade de Brasília, Departamento de Saúde Coletiva, Brasília-DF, Brasil

Resumo

Objetivo: descrever o perfil dos casos de comorbidade tuberculose-diabetes no Brasil. **Métodos:** estudo descritivo com dados do Sistema de Informação de Agravos de Notificação – tuberculose (Sinan-tuberculose) e do Sistema de Gestão Clínica de Hipertensão Arterial e Diabetes *Mellitus* da Atenção Básica (Hiperdia), de 2007 a 2011; o relacionamento probabilístico foi realizado com auxílio do *software* Reclink. **Resultados:** foram encontrados 24.443 casos de comorbidade, incluindo 3.181 não registrados no Sinan-tuberculose; do total de casos recuperados, predominaram indivíduos do sexo masculino (57,2%), faixa etária de 40 a 59 anos (52,3%), raça/cor da pele preta/parda (68,4%), cinco a oito anos de estudo (78,4%), uso não regular de álcool (86,5%) e sorologia negativa para o vírus da imunodeficiência humana – HIV (91,8%). **Conclusão:** os casos captados tiveram perfil semelhante àqueles registrados no Sinan-tuberculose e o relacionamento dos dados de diferentes sistemas de informações possibilitou a detecção de casos não captados pela vigilância.

Palavras-chave: Tuberculose; Diabetes *Mellitus*; Sistemas de Informação; Epidemiologia Descritiva.

Abstract

Objective: to describe the profile of cases of tuberculosis and diabetes comorbidity in Brazil. **Methods:** this is a descriptive study with data from the Brazilian Information System for Notifiable Diseases – tuberculosis (Sinan-TB) and from the System of Registration and Monitoring of Hypertension and Diabetes *Mellitus* (Hiperdia), from 2007 to 2011; probabilistic linkage was carried out with Reclink software. **Results:** 24,443 cases of comorbidity were found, including 3,181 cases not registered on Sinan-TB; of the total number of recovered cases, mostly were males (57.2%), aged 40-59 years (52.3%), black/brown-skinned (68.4%), with five to eight years of schooling (78.4%), with no regular use of alcohol (86.5%) and negative serology for the HIV virus (91.8%). **Conclusion:** the cases found had similar profile to those registered on Sinan-TB and the probabilistic linkage of data from different information systems enabled the detection of cases not captured by surveillance.

Keywords: Tuberculosis; Diabetes *Mellitus*; Information Systems; Epidemiology, Descriptive.

*Este artigo compõe a tese de Doutorado em Medicina Tropical de Ricardo Gadelha de Abreu, a ser defendida junto à Universidade de Brasília em 2017.

Endereço para correspondência:

Ricardo Gadelha de Abreu - CA 09, Lotes 1, 3 e 5, Torre 3, Apto. 206, Lago Norte, Brasília-DF, Brasil. CEP: 71503-509
E-mail: ricardo.gadelha@gmail.com

Introdução

Tuberculose e diabetes são problemas de Saúde Pública de grande relevância. O risco de uma pessoa com diabetes desenvolver tuberculose pode representar de 2,44 a 8,33 vezes o mesmo risco para uma pessoa sem diabetes.¹ Todavia não foi encontrada explicação sobre a maior ocorrência da infecção pelo bacilo de Koch, causador da tuberculose, entre diabéticos. A comorbidade tuberculose-diabetes requer atenção e cuidados mais complexos, uma vez que a diabetes pode interferir no metabolismo dos fármacos antituberculose.²

A tuberculose é um agravo de notificação compulsória. As principais medidas para o controle da doença e interrupção de sua cadeia de transmissão são o diagnóstico e o tratamento correto e oportuno dos casos de tuberculose pulmonar.^{3,4}

Em 2014, em todo o mundo, foram estimadas 9,6 milhões de pessoas com tuberculose: 5,4 milhões de homens, 3,2 milhões de mulheres e 1 milhão de crianças. Entretanto, cerca de 6 milhões de casos novos da doença – menos de dois terços (62,5%) – foram notificados à Organização Mundial da Saúde (OMS), indicando que 37,5% dos casos não foram diagnosticados ou não foram registrados. Para o mesmo ano de 2014, foram estimadas 1,5 milhões de mortes por tuberculose.⁵

Tuberculose e diabetes são problemas de Saúde Pública de grande relevância. O risco de uma pessoa com diabetes desenvolver tuberculose pode representar de 2,44 a 8,33 vezes o mesmo risco para uma pessoa sem diabetes.

De acordo com a OMS, em 2015, o Brasil foi um dos 22 países que concentraram 80% da carga mundial de tuberculose, ocupando a 18ª posição em número absoluto de casos novos e a 22ª em relação ao coeficiente de incidência.⁵ Segundo o Ministério da Saúde, de 2003 a 2012, a taxa de incidência da tuberculose apresentou declínio de 38,7%; e a taxa de mortalidade, redução de 33,6%.^{6,7}

O crescimento no número de pessoas com diabetes no mundo, de 171 milhões de casos em 2000 para uma estimativa de 440 milhões em 2030, pode representar um fator adicional a dificultar o controle da tuberculose.²

Em 2012, entre os países da América Central e do Sul, a prevalência de diabetes foi estimada em 26,4 milhões de pessoas, sendo projetada para 40 milhões em 2030. No Brasil, a prevalência de diabetes foi estimada em 4,6% no ano 2000 (8ª posição entre os referidos países). A estimativa para 2030 aponta uma prevalência de 11,3% (6ª posição) em 2030.^{8,9}

No Brasil, os casos de tuberculose são registrados no Sistema de Informação de Agravos de Notificação (Sinan). O cadastramento e o acompanhamento das pessoas com diabetes no Sistema Único de Saúde (SUS) eram feitos no Sistema de Gestão Clínica de Hipertensão Arterial e Diabetes Mellitus da Atenção Básica (Hiperdia), até 2013. A partir de então, o Hiperdia foi substituído pelo e-SUS Atenção Básica (AB), ou e-SUS AB: um processo de informatização qualificada, estratégia do Departamento de Atenção Básica, da Secretaria de Atenção à Saúde do Ministério da Saúde, para reestruturar as informações da Atenção Básica oferecida pelo Sistema Único de Saúde (SUS).¹⁰⁻¹⁴

A crescente disponibilidade de sistemas informatizados capazes de armazenar dados de abrangência nacional requer a utilização de metodologias que relacionem esses dados de forma a obter bases mais robustas para análises epidemiológicas.^{15,16} A ausência de um identificador unívoco nas bases de dados para associar registros dos diferentes sistemas de informações em saúde existentes representa um desafio adicional.

Visando garantir a qualidade das fontes de registro que auxiliam a vigilância da tuberculose, o entre bases de dados justifica-se pela qualificação das informações de forma não onerosa e com baixo custo operacional, possibilitando detectar casos identificados por outros sistemas, embora não captados pela vigilância da doença.¹⁷⁻¹⁹

Este estudo objetivou descrever o perfil dos casos de comorbidade tuberculose-diabetes no Brasil, no período de 2007 a 2011.

Métodos

Foi realizado estudo descritivo com dados dos casos de comorbidade tuberculose-diabetes no Brasil, obtidos a partir do relacionamento probabilístico das bases de dados do Sinan referentes às notificações por tuberculose (Sinan-tuberculose), e do Hiperdia, referente aos registros de casos por diabetes. A base de dados do Sinan-tuberculose considerou os casos

notificados de 1º/1/2001 até 31/12/2012, e a base de dados do Hiperdia, de 1º/1/2007 a 31/12/2011.

O Sinan-tuberculose é alimentado pela Ficha de Notificação/Investigação de Tuberculose, que contempla informações obtidas do caso, seu local de residência, a situação clínica e a classificação de acordo com seu tipo de entrada, entre outras. Para fins de vigilância epidemiológica, o Programa Nacional de Controle da Tuberculose definiu como caso notificado no Sinan-tuberculose somente o caso confirmado da doença.²⁰

O Hiperdia era utilizado para cadastramento e acompanhamento das pessoas com diabetes e/ou hipertensão atendidas na rede ambulatorial do SUS. Criado em 2002, o sistema permitia gerar informações que subsidiassem a aquisição, dispensação e distribuição de medicamentos para os cadastrados, além de contribuir para o planejamento eficaz de prevenção e controle de complicações nos indivíduos já diagnosticados com a doença.¹⁰⁻¹³ O registro com entrada única da pessoa diagnosticada com diabetes no Hiperdia possibilitava sua identificação e vinculação às unidades de saúde ou equipes da Atenção Básica do SUS, que registravam os dados, permitindo a geração de informações sobre o desempenho e os resultados clínicos durante o acompanhamento.²¹

De acordo com a Portaria Conjunta nº 2 das Secretarias de Vigilância em Saúde e Executiva do Ministério da Saúde, de 5 de março de 2002, em seu art. 5º, o gestor municipal, para integrar o Programa Nacional de Assistência Farmacêutica para Hipertensão Arterial e Diabetes *Mellitus*, deveria assinar um Termo de Adesão responsabilizando-se por integrar o programa de assistência farmacêutica, cadastrar e acompanhar as pessoas com essas doenças no Hiperdia, de forma a assegurar aos cadastrados o recebimento dos medicamentos padronizados prescritos.²²

A alimentação da base de dados nacional do Hiperdia, a cargo dos municípios que aderiram ao Programa de Assistência Farmacêutica à Hipertensão Arterial e à Diabetes *Mellitus*, era feita via cadastramento dos usuários no Subsistema Centralizador Municipal do sistema ou mediante a exportação de dados de aplicativo próprio.^{13,23}

Na fase de pareamento, para reduzir possíveis erros, realizou-se uma preparação prévia das bases com a exclusão de registros duplicados e a padronização (aplicação da mesma grafia para as primeiras sílabas

de nomes com a mesma fonética) e codificação dos campos. Para melhorar a qualidade de campos de 'nome' utilizados no processo, foi necessária a remoção de informações ou caracteres, tais como &, #, @, entre outros equivocadamente adicionados ao nome da pessoa.

Visando obter uma base conjunta, contemplando os casos de tuberculose e diabetes, e diminuir o número de comparações entre os registros, foram utilizados, em ambas as bases, procedimentos probabilísticos estabelecidos pelo programa Reclink versão 3.1.6.3160, empregando-se campos comuns para verificar se os registros pareados pertenciam ao mesmo indivíduo. Foram adotadas as seguintes rotinas para a padronização dos campos comuns: (i) comparação de 'nome da pessoa', 'nome da mãe' e 'data de nascimento', e aplicação de algoritmos para comparação aproximada de cadeias de caracteres, com o objetivo de contornar erros fonéticos e de grafia; (ii) blocagem, dividindo-se as bases em blocos mutuamente exclusivos, com comparações restritas aos registros de um mesmo bloco e otimização da comparação entre registros; (iii) cálculo de escores, para indicar o grau de concordância entre registros de um mesmo par; e (iv) definição de limiares para classificação dos pares de registros e sua identificação como verdadeiros ou não pares.^{24,25}

Os procedimentos probabilísticos também foram aplicados na identificação de possíveis duplicidades na base de dados do Hiperdia. A rotina de remoção de duplicidade não foi aplicada na base do Sinan-tuberculose, porque foi feita pelo Programa Nacional de Controle da Tuberculose previamente à cessão dos dados.

Para agilizar o processamento dos bancos, utilizou-se a blocagem em múltiplos passos, uma vez que os bancos somavam mais de 1.000.000 (1 milhão) de registros e os computadores utilizados não dispunham de memória e velocidade suficientes para o processamento de tantos dados. Assim, foi realizada uma divisão do banco do Hiperdia em seis blocos – considerando-se a primeira letra do nome da pessoa –, para serem relacionados com o banco Sinan-tuberculose:

1º Nomes iniciados por A até C

2º Nomes iniciados por D até I

3º Nomes iniciados por J até L

4º Nomes iniciados por M

5º Nomes iniciados por N até R

6º Nomes iniciados por S até Z

A comparação dos registros dentro de cada bloco considerou três passos:

- 1) Comparação do nome completo do paciente, nome completo da mãe do paciente e data de nascimento utilizando-se a chave de blocagem PBLOCO+UBLOCO+SEXO
- 2) Comparação do nome completo do paciente, nome completo da mãe do paciente e data de nascimento utilizando-se a chave de blocagem PBLOCO+UBLOCO
- 3) Comparação do nome completo do paciente, nome completo da mãe do paciente e data de nascimento utilizando-se a chave de blocagem PBLOCO+SEXO

Os relacionamentos probabilísticos requerem maior atenção aos pares duvidosos, situados na fronteira de incerteza, conhecida como zona cinza.²⁶ Para maximizar a capacidade da ferramenta em identificar os pares verdadeiros, os parâmetros de relacionamento foram estimados com a aplicação da ferramenta disponível do ReLink – procedimentos ‘gera matriz’ e ‘calcula parâmetros’ – sobre as informações das próprias bases de dados, gerando a distribuição teórica do escore: -1,6843 a 30,0763.

Como as bases de dados eram muito extensas, utilizou-se a fração amostral de 1% com o propósito de escolher os melhores parâmetros no relacionamento, ou seja, a capacidade de melhor discriminar os pares dos não pares. Realizou-se uma inspeção na base, com o objetivo de identificar o melhor limiar para classificação dos pares: os registros com valores de escore maiores ou iguais a 24,9 foram considerados pares verdadeiros e os com escores inferiores a 24,9 foram considerados não pares ou pares duvidosos. Com os parâmetros estimados, alcançou-se um bom desempenho na identificação dos pares verdadeiros, permitindo supor que a revisão manual da chamada zona cinza não acrescentaria um volume substancial de registros ao conjunto dos pareados. Assim, optou-se por não revisar a zona cinza e chamar os pares duvidosos de não pares.²⁷

Após o processamento dos seis blocos relacionados ao banco do Sinan-TB, os respectivos blocos foram juntados novamente, resultando em um único banco, cujas denominações para as origens dos dados foram codificadas por:

01 = banco Sinan-tuberculose

10 = banco Hiperdia

11 = banco comorbidade tuberculose-diabetes (relacionado)

Para o banco final, foram incluídos no estudo os códigos 01 e 11 (respectivamente: casos de tuberculose sem a comorbidade tuberculose-diabetes notificados no Sinan de 1º/1/2007 até 30 de junho de 2011; e casos com comorbidade tuberculose-diabetes – casos notificados no Sinan-tuberculose com a informação da associação com a diabetes e casos recuperados após o relacionamento dos bancos do Sinan-tuberculose e do Hiperdia). Foram considerados sem comorbidade os casos de tuberculose notificados no Sinan-tuberculose sem diabetes associada, ou seja, casos de tuberculose cuja variável ‘doenças e agravos associados – diabetes’ da ficha do Sinan apresentasse o código ‘2’ (Não); foram excluídos os casos com código ‘9’ (ignorado) ou sem preenchimento (em branco) para a mesma variável. Para os casos com a comorbidade tuberculose-diabetes, foram considerados:

- caso de tuberculose notificado no Sinan em que a variável ‘doenças e agravos associados – diabetes’ apresentava o código ‘1’ (Sim);
- caso de diabetes registrado no Hiperdia antes de o caso de tuberculose ser notificado no Sinan; e
- casos de diabetes e tuberculose registrados e notificados no mesmo ano em ambos os sistemas – Hiperdia e Sinan.

Foram excluídos os casos de tuberculose registrados no Sinan após junho de 2011, uma vez que os indivíduos diagnosticados no segundo semestre de 2011 deveriam concluir o tratamento recomendado (de seis meses) somente em 2012, período fora do estudo e sem informações dos registros de diabetes.

Para descrever o perfil sociodemográfico das pessoas com a comorbidade tuberculose-diabetes, foram consideradas as seguintes variáveis da ficha de notificação do Sinan-tuberculose:

- sexo (masculino; feminino);
- faixa etária (em anos: até 19; 20 a 39; 40 a 59; 60 ou mais);
- raça/cor de pele (branca; preta/parda; outras);
- escolaridade (em anos de estudo: até 4; 5 a 8; 8 ou mais);
- uso regular de álcool (sim; não); e
- sorologia para o vírus da imunodeficiência humana, o HIV (positivo; negativo).

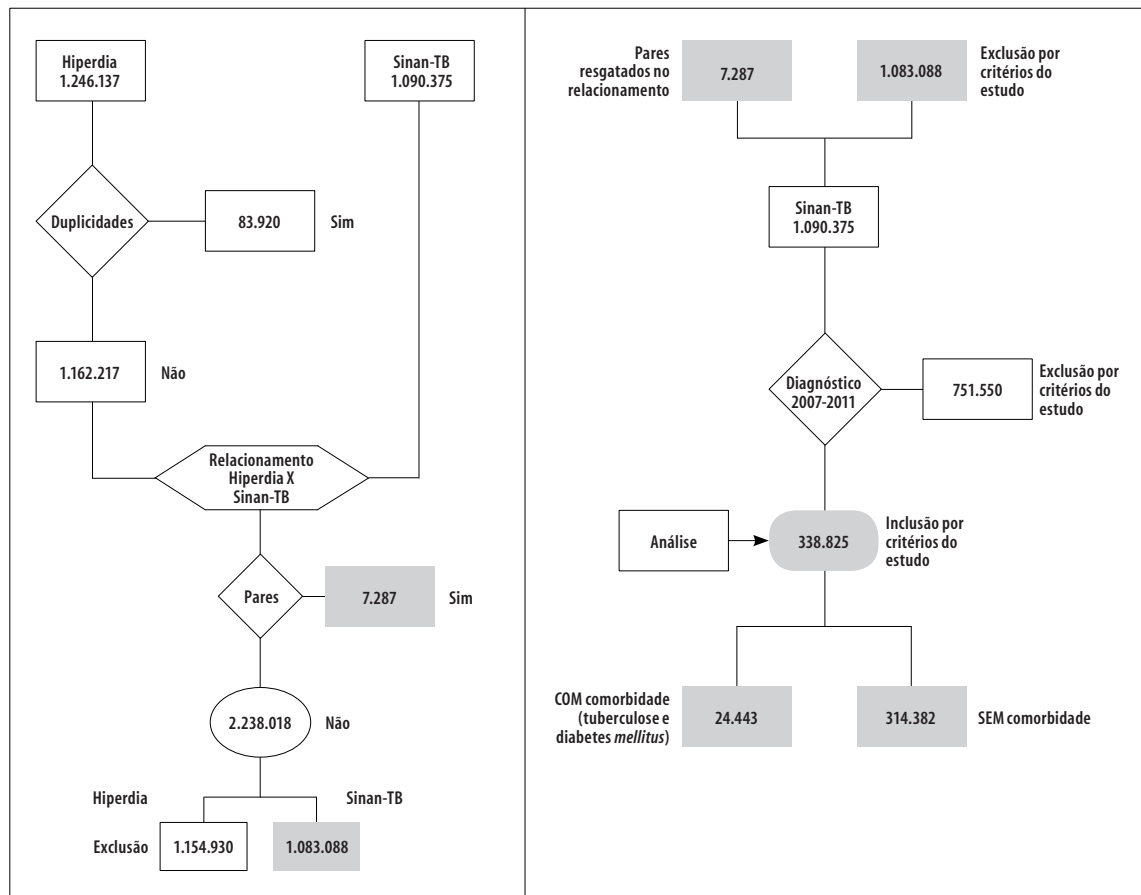
As análises foram feitas pelo teste de qui-quadrado de Pearson, para comparação das proporções entre o grupo do relacionamento e o do Sinan-tuberculose, com nível de significância estatística de 5%.

As bases de dados nominais do Sinan e do Hiperdia foram disponibilizadas pelo Ministério da Saúde, por meio do Programa Nacional de Controle da Tuberculose e do Departamento de Informática do SUS (Datasus), mediante assinatura de Termo de Responsabilidade pelo pesquisador. O projeto da pesquisa foi aprovado pelo Comitê de Ética em Pesquisa da Faculdade de Ciências da Saúde da Universidade de Brasília: Parecer nº 552.561, de 11 de março de 2014.

Resultados

Após organização das bases, foram contabilizados 1.090.375 registros de casos de tuberculose e 1.246.137 cadastros de diabetes relativos ao período estudado. (Figura 1).

Do primeiro bloco, de A a C do banco Hiperdia, de 209.370 cadastros, foram removidas 13.347 duplicidades. Do segundo bloco, de D a I do banco Hiperdia, de 225.089 cadastros, foram removidas 15.497 duplicidades. Do terceiro bloco, de J a L do banco Hiperdia, de 238.660 cadastros, foram removidas 14.795 duplicidades. Do quarto bloco, letra inicial M do banco Hiperdia, de 284.337 cadastros, foram removidas 20.387 duplicidades. Do quinto bloco, de N a R do banco Hiperdia, de 146.261 cadastros, foram removidas 10.746 duplicidades. Por fim, do total do sexto bloco, de S a Z do banco Hiperdia, de 142.412 cadastros, foram removidas 9.148 duplicidades. Ao todo, foram removidas 83.920 duplicidades. Outrossim, do total de cadastros de diabetes (n=1.246.137),



Legenda:
 Sinan-TB: Sistema de Informação de Agravos de Notificação – Tuberculose
 Hiperdia: Sistema de Gestão Clínica de Hipertensão Arterial e Diabetes Mellitus da Atenção Básica

Figura 1 – Fluxograma da base de dados das pessoas com tuberculose, com e sem a comorbidade tuberculose-diabetes mellitus

oito continham informação não válida no campo 'nome do indivíduo' para criação do bloco, sendo esses excluídos da análise.

Após o relacionamento entre os bancos do Sinan-tuberculose e do Hiperdia, foram encontrados 7.287 registros de pessoas em ambos os bancos, ou seja, com a comorbidade tuberculose-diabetes (Tabela 1):

- bloco 1 [A a C], 1.057 pares;
- bloco 2 [D a I], 987 pares;
- bloco 3 [J a L], 2.097 pares;
- bloco 4 [M], 1.765 pares;
- bloco 5 [N a R], 734 pares; e
- bloco 6 [S a Z], 647 pares.

Afinal, conforme a origem do banco (Sinan-tuberculose; relacionamento Sinan-Hiperdia) e considerando-se os critérios de inclusão de casos para o estudo, resultaram para análise 335.644 casos de tuberculose com ou sem comorbidade no Sinan e 3.181 casos da comorbidade diabetes-tuberculose recuperados no relacionamento, ou seja, que não estavam com o campo da informação da associação tuberculose-diabetes assinalado no Sinan-tuberculose, totalizando 338.825 casos. Desse total, 314.382 não apresentavam a comorbidade e 24.443 tinham tuberculose e diabetes (21.262 provenientes do

Sinan-tuberculose e 3.181 do banco relacionado) (Figuras 1 e 2).

Entre os casos de comorbidade tuberculose-diabetes identificados após o relacionamento, observou-se predominância do sexo masculino (57,2%), faixa etária de 40 a 59 anos (52,3%), raça/cor da pele preta/parda (68,4%), cinco a oito anos de estudo (78,4%), uso não regular de álcool (86,5%) e sorologia negativa para HIV (91,8%). Entre os casos de comorbidade registrados no banco Sinan-tuberculose, predominaram indivíduos do sexo masculino (63,1%), faixa etária de 40 a 59 anos (53,1%), raça/cor da pele preta/parda (56,1%), cinco a oito anos de estudo (87,0%), uso não regular de álcool (84,9%) e sorologia negativa para HIV (91,4%). Não houve diferença estatisticamente significativa ($p > 0,05$) entre os casos de comorbidade identificados após o relacionamento e os casos identificados exclusivamente no Sinan, no que concerne a sexo, raça/cor da pele e anos de estudo (Tabela 2).

Discussão

No presente estudo, o relacionamento probabilístico entre as bases de dados possibilitou a melhoria da captação da informação sobre a comorbidade

Tabela 1 – Número de pares encontrados após relacionamento entre os blocos dos bancos padronizados do Sinan-tuberculose^a e do Hiperdia.^b Brasil, 2007-2011

Blocos (letras iniciais dos nomes dos pacientes)	Total de casos nas bases de dados	Número de pares encontrados
A-C	Hiperdia = 195.172 Sinan-tuberculose = 206.645	1.057
D-I	Hiperdia = 208.749 Sinan-tuberculose = 206.840	987
J-L	Hiperdia = 222.301 Sinan-tuberculose = 248.411	2.097
M	Hiperdia = 262.531 Sinan-tuberculose = 140.690	1.765
N-R	Hiperdia = 134.891 Sinan-tuberculose = 145.100	734
S-Z	Hiperdia = 132.736 Sinan-tuberculose = 135.498	647
Total de pares		7.287

a) Sinan-tuberculose: Sistema de Informação de Agravos de Notificação – Tuberculose

b) Hiperdia: Sistema de Gestão Clínica de Hipertensão Arterial e Diabetes *Mellitus* da Atenção Básica

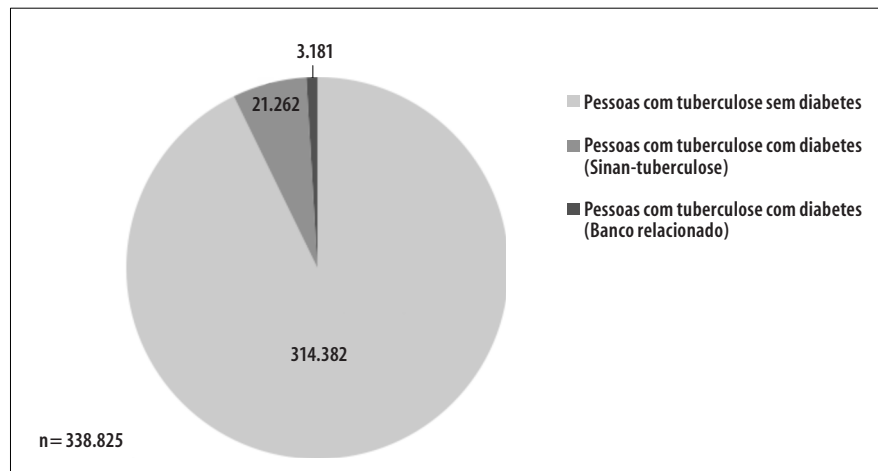


Figura 2 – Número de indivíduos com tuberculose, com e sem diabetes, segundo base de dados de identificação da comorbidade, Brasil, 2007-2011

Tabela 2 – Distribuição dos pacientes com comorbidade tuberculose-diabetes *mellitus* segundo origem do banco de dados e variáveis sociodemográficas, Brasil, 2007-2011 (n= 24.443)^a

Variáveis sociodemográficas		Comorbidade tuberculose-diabetes <i>mellitus</i>		
		Relacionamento Sinan-tuberculose ^b e Hipertensão ^c		Valor de p ^d
		n (%)	Sinan-tuberculose ^b n (%)	
Sexo	Masculino	1.820 (57,2)	13.410 (63,1)	<0,001
	Feminino	1.361 (42,8)	7.850 (36,9)	
Faixa etária (em anos)	≤19	34 (1,1)	297 (1,4)	0,002
	20-39	385 (12,3)	2.918 (13,9)	
	40-59	1.643 (52,3)	11.120 (53,1)	
	≥60	1.078 (34,3)	6.593 (31,5)	
Raça/cor da pele	Branca	876 (30,3)	7.822 (42,0)	<0,001
	Preta/parda	1.978 (68,4)	10.446 (56,1)	
	Outras	39 (1,3)	349 (1,9)	
Escolaridade (em anos de estudo)	≤4	428 (19,7)	1.424 (10,6)	<0,001
	5-8	1.705 (78,4)	11.713 (87,0)	
	≥8	43 (2,0)	319 (2,4)	
Uso regular de álcool	Sim	369 (13,5)	2.908 (15,1)	0,029
	Não	2.356 (86,5)	16.303 (84,9)	
HIV ^e	Positivo	120 (8,2)	966 (8,6)	0,583
	Negativo	1.347 (91,8)	10.258 (91,4)	

a) Número de observações válidas

b) Sinan-tuberculose: Sistema de Informação de Agravos de Notificação – Tuberculose

c) Hipertensão: Sistema de Gestão Clínica de Hipertensão Arterial e Diabetes *Mellitus* da Atenção Básica

d) Teste de qui-quadrado de Pearson

e) HIV: vírus da imunodeficiência humana

tuberculose-diabetes não registrada na base nacional do Sinan-tuberculose. Seus resultados evidenciam a importância do trabalho de integração do Sinan-tuberculose com outros bancos de dados, principalmente com aqueles que detêm informações dos agravos associados à tuberculose, caso da diabetes.²⁹ Mesmo sendo o Hiperdia o sistema utilizado para o cadastramento das pessoas com diabetes no período de 2002 a 2013 e existindo, no Sinan, a diabetes como campo-agravo associado à tuberculose, não foram encontrados estudos na literatura, tampouco publicações do Ministério da Saúde que abordassem o relacionamento entre esses bancos.

O relacionamento nominal de bases de dados é normalmente realizado com a aplicação de métodos probabilísticos que utilizam processos de pareamento de duas ou mais bases, com probabilidades de concordância – ou não – entre um conjunto de variáveis comuns a essas bases, buscando pares de registros referentes à mesma pessoa, ademais da padronização e qualificação dessas informações.²⁸

Uma vez que na ficha do Sinan, o campo ‘agravo associado à tuberculose’ não é de preenchimento obrigatório, esse registro apresentou-se com grande percentual em branco. Assim, é possível que esta análise esteja prejudicada e por conseguinte, o desconhecimento das pessoas sobre a comorbidade tenha dificultado uma melhor assistência, além de prejudicar a investigação intradomiciliar de novos casos. Ressalta-se que o registro do referido campo é de extrema importância: a ocorrência dessa comorbidade vem crescendo e as condutas clínicas e o monitoramento dos casos tornam-se cada dia mais necessários.

A discordância de registros encontrada entre os dois sistemas demonstra a necessidade de melhorar a integração e o fluxo de dados entre sistemas de informações em saúde, também observada em estudos com outros sistemas oficiais.^{18,30}

Mesmo diante da indisponibilidade, neste estudo, de computadores com memória suficiente para processar bancos de dados com grandes quantidades de registros, a divisão dos bancos em seis blocos possibilitou o relacionamento entre o Sinan-tuberculose e o Hiperdia sem causar prejuízos aos resultados encontrados. Não obstante, para estudos futuros, recomenda-se utilizar uma estratégia mais eficiente de relacionamento das bases, por meio do emprego,

em todos os passos, de uma chave de blocagem (indexação) formada de, no mínimo, quatro campos: soundex do primeiro nome, do último nome, do primeiro nome da mãe; ano de nascimento; sexo; e município de residência. A aplicação dessa estratégia no relacionamento de bases de dados permitiria a formação de diferentes chaves, com pelo menos quatro campos.

Como limitações deste estudo, destacam-se a falta de completude de variáveis utilizadas para o relacionamento e a utilização de dados secundários, e consequentemente, a possibilidade de subnotificações, não se podendo afirmar que o universo de pessoas com as duas doenças esteja representado na base de dados relacionada.

Entre os indivíduos com tuberculose e diabetes, assim como observaram outros estudos sobre tuberculose, houve predominância da raça/cor da pele preta/parda e do sexo masculino; diferentemente da diabetes no Brasil, para a qual a maior prevalência é feminina.^{5,6} Possivelmente, as mulheres com diabetes registradas no Hiperdia mantiveram um maior vínculo com os serviços, daí sua maior participação no banco relacionado quando comparados os dois grupos de gênero. Os resultados do presente estudo reiteram a importância de relacionar informações advindas de diferentes sistemas de informações em saúde, principalmente para a detecção e qualificação de casos registrados em outros sistemas – não captados pela vigilância da doença –, contribuindo para o aperfeiçoamento da gestão no SUS.

Em conclusão, a aplicação de relacionamentos de bancos de dados pode se transformar em uma ferramenta importante para o monitoramento da situação de saúde das pessoas com tuberculose e com diabetes em tratamento. Estudos futuros devem aprofundar o conhecimento dos diferentes perfis dessa comorbidade.

Contribuição dos autores

Abreu RG e Sousa AIA contribuíram no delineamento do estudo, análise e interpretação dos dados, elaboração de versões preliminares e redação final do manuscrito. Oliveira MRF e Sanchez MN contribuíram no delineamento do estudo, na interpretação dos dados e na revisão da redação. Todos os autores aprovaram a versão final para publicação e se responsabilizam integralmente por seu conteúdo.

Referências

1. Seiscento M. Tuberculose em situações especiais: HIV, diabetes mellitus e insuficiência renal. *Pulmão RJ*. 2012;21(1):23-6.
2. Organização Pan-Americana da Saúde. Manejo conjunto de tuberculose e diabetes mellitus: enfoque colaborativo nos serviços de saúde das cidades de Tlhuana (México) e a Bahia e São Paulo (Brasil) [Internet]. Brasília: Organização Pan-Americana da Saúde; 2012 [citado 2015 out 14]. Disponível em: http://www.paho.org/hq/index.php?option=com_docman&task=doc_view&gid=22820&Itemid=
3. Brasil. Ministério da Saúde. Portaria nº 204, de 17 de fevereiro de 2016. Define a Lista Nacional de Notificação Compulsória de doenças, agravos e eventos de saúde pública nos serviços de saúde públicos e privados em todo o território nacional, nos termos do anexo, e dá outras providências. *Diário Oficial da República Federativa do Brasil, Brasília (DF)*, 2016 fev 18; Seção 1:23.
4. Maciel ELN, Sales CMM. A vigilância epidemiológica da tuberculose no Brasil: como é possível avançar mais? *Epidemiol Serv Saude*. 2016 jan-mar;25(1):175-8.
5. World Health Organization. 2015 Global tuberculosis report [Internet]. Geneva: World Health Organization; [2016] [cited 2016 Jun 28]. Available from: http://apps.who.int/iris/bitstream/10665/191102/1/9789241565059_eng.pdf
6. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis. Panorama da tuberculose no Brasil: indicadores epidemiológicos e operacionais. Brasília: Ministério da Saúde; 2014.
7. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Detectar, tratar e curar: desafios e estratégias brasileiras frente à tuberculose. *Bol Epidemiol*. 2015;46(9):1-19.
8. Organização Mundial da Saúde. Cuidados inovadores para condições crônicas: componentes estruturais de ação. Brasília: Organização Mundial da Saúde; 2003.
9. Ministério da Saúde (BR). Secretaria de Atenção à Saúde. Departamento de Atenção Básica. Estratégias para o cuidado da pessoa com doença crônica: diabetes mellitus. Brasília: Ministério da Saúde; 2013. (Cadernos de Atenção Básica, 36).
10. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Departamento de Vigilância Epidemiológica. Manual de recomendações para o controle da tuberculose no Brasil. Brasília: Ministério da Saúde; 2011. (Série A. Normas e Manuais Técnicos).
11. Laguardia J, Domingues CMA, Carvalho C, Lauerman CR, Macário E, Glatt R. Sistema de Informação de Agravos de Notificação (Sinan): desafios no desenvolvimento de um sistema de informação em saúde. *Epidemiol Serv Saude*. 2004 jul-set;13(3):135-46.
12. Ferreira CLRA, Ferreira MG. Características epidemiológicas de pacientes diabéticos da rede pública de saúde - análise a partir do sistema HiperDia. *Arq Bras Endocrinol Metab*. 2009 fev;53(1):80-6.
13. Ministério da Saúde (BR). Secretaria Executiva. Departamento de Informática do SUS. HiperDia: Sistema de Cadastramento e Acompanhamento de Hipertensos e Diabéticos: manual de operação. Rio de Janeiro: Ministério da Saúde; 2002.
14. Ministério da Saúde (BR). e-SUS Atenção Básica [Internet]. Brasília: Ministério da Saúde; 2012 [citado 2015 nov 21]. Disponível em: <http://dab.saude.gov.br/portaldab/esus.php>.
15. Machado JP, Silveira DP, Santos IS, Piovesan MF, Albuquerque C. Aplicação da metodologia de relacionamento probabilístico de base de dados para a identificação de óbitos em estudos epidemiológicos. *Rev Bras Epidemiol*. 2008 mar;11(1):43-54.
16. Coeli CM, Pinheiro RS, Camargo Júnior KR. Conquistas e desafios para o emprego das técnicas de record linkage na pesquisa e avaliação em saúde no Brasil. *Epidemiol Serv Saude*. 2015 out-dez;24(4):795-802.
17. Pinheiro RS, Andrade VL, Oliveira GP. Subnotificação da tuberculose no Sistema de Informação de Agravos de Notificação (Sinan): abandono primário de bacilíferos e captação de casos em outras fontes de informação usando linkage probabilístico. *Cad Saude Publica*. 2012 ago;28(8):1559-68.
18. Bartholomay P, Oliveira GP, Pinheiro RS, Vasconcelos AMN. Melhoria da qualidade das informações sobre tuberculose a partir do relacionamento entre bases de dados. *Cad Saude Publica*. 2014 nov;30(11):2459-69.

19. Pinheiro RS, Andrade VL, Oliveira GP. Subnotificação da tuberculose no Sistema de Informação de Agravos de Notificação (SINAN): abandono primário de bacilíferos e captação de casos em outras fontes de informação usando linkage probabilístico. *Cad Saude Publica*. 2012 ago;28(8):1559-68.
20. Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Guia de Vigilância em Saúde. Brasília: Ministério da Saúde; 2014.
21. Ministério da Saúde (BR). Sistema de Gestão Clínica de Hipertensão Arterial e Diabetes Mellitus da Atenção Básica (Hiperdia) [Internet]. Brasília: Ministério da Saúde; ANO [citado 2016 out 24] Disponível em: <http://hiperdia.datasus.gov.br/>.
22. Brasil. Ministério da Saúde. Portaria Conjunta nº 02/SPS/SE, de 05 de março de 2002. Disponibiliza, no âmbito da atenção básica, instrumento de cadastro e acompanhamento dos portadores de Hipertensão Arterial e Diabetes Mellitus, Sishiperdia, para utilização pelos municípios. *Diário Oficial da República Federativa do Brasil, Brasília (DF)*, 2002 mar 20;Seção 1:38.
23. Brasil. Ministério da Saúde. Portaria Conjunta nº 112/SPS/SE, de 19 de junho de 2002. Aprova o fluxo de alimentação da base nacional do Hiperdia, obrigatório para todos os municípios que aderirem ao Programa de Assistência Farmacêutica à Hipertensão Arterial e ao Diabetes Mellitus conforme estabelecido no anexo da Portaria. *Diário Oficial da República Federativa do Brasil, Brasília (DF)*, 2002 jun 20;Seção 1:60.
24. Coeli CM, Camargo Júnior KR. Avaliação de diferentes estratégias de blocagem no relacionamento probabilístico de registros. *Rev Bras Epidemiol*. 2002;5(2):185-96.
25. Lucena FFA, Fonseca MGP, Sousa AIA, Coeli CM. O relacionamento de bancos de dados na implementação da vigilância da aids. *Cad Saude Publica*. 2006 abr-jun;14(2):305-12.
26. Queiroz OV, Guerra Júnior AA, Machado CJ, Andrade ELG, Meira Júnior W, Acúrcio FA, et al. A construção da Base Nacional de Dados em Terapia Renal Substitutiva (TRS) centrada no indivíduo: relacionamento dos registros de óbitos pelo subsistema de Autorização de Procedimentos de Alta Complexidade (Apac/SIA/SUS) e pelo Sistema de Informações sobre Mortalidade (SIM) – Brasil, 2000-2004. *Epidemiol Serv Saude*. 2009 abr-jun;18(2):107-20.
27. Peres SV, Latorre MRDO, Michels FAS, Tanaka LF, Coeli CM, Almeida ME. Determinação de um ponto de corte para a identificação de pares verdadeiros pelo método probabilístico de linkage de base de dados. *Cad Saude Colet*. 2014 out-dez;22(4):428-36.
28. Silveira DP, Artmann E. Acurácia em métodos de relacionamento probabilístico de bases de dados em saúde: revisão sistemática. *Rev Saude Publica*. 2009 out;43(5):875-82.
29. Oliveira GP, Pinheiro RS, Coeli CM, Barreira D, Codenotti SB. Uso do sistema de informação sobre mortalidade para identificar subnotificação de casos de tuberculose no Brasil. *Rev Bras Epidemiol*. 2012 set;15(3):468-77.
30. Queiroz OV, Guerra Júnior AA, Machado CJ, Andrade ELG, Meira Júnior W, Acúrcio FA, et al. Relacionamento de registros de grandes bases de dados: estimativa de parâmetros e validação dos resultados, aplicados ao relacionamento dos registros das autorizações de procedimentos ambulatoriais de alta complexidade com os registros de sistema de informações hospitalares. *Cad Saude Colet*. 2010;18(2):298-308.

Recebido em 30/06/2016
Aprovado em 22/09/2016

Errata

No artigo "Tuberculose e diabetes: relacionamento probabilístico de bases de dados para o estudo da associação entre ambas doenças", com número de doi: 10.5123/S1679-49742017000200013, publicado na revista *Epidemiologia e Serviços de Saúde*, 26(2):359-368, na página 364.

Onde se lia:

Não houve diferença estatisticamente significativa ($p > 0,05$) entre os casos de comorbidades identificados após o relacionamento e os casos identificados exclusivamente no Sinan, no que concerne a sexo, raça/cor da pele e anos de estudo (Tabela 2).

Leia-se:

Não houve diferença estatisticamente significativa ($p < 0,05$) entre os casos de comorbidades identificados após o relacionamento e os casos identificados exclusivamente no Sinan, no que concerne à sorologia para HIV (Tabela 2).

Para o artigo original acesse: <http://www.scielo.br/pdf/ress/v26n2/2237-9622-ress-26-02-00359.pdf>