

Codificação e seleção automáticas das causas de morte: adaptação para o uso no Brasil do software Iris

Automatic coding and selection of causes of death: an adaptation of Iris software for using in Brazil

Renata Cristófani Martins^I, Cassia Maria Buchalla^{II}

RESUMO: *Objetivo:* Elaborar um dicionário em Português a ser utilizado no *software* Iris e avaliar sua completude para a codificação das causas de morte. *Métodos:* Em primeiro lugar, criou-se um dicionário de doenças e lesões e seus códigos da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde - décima revisão (CID-10). Este foi baseado em duas fontes: o arquivo eletrônico do volume 1 da CID-10 e os dados do Tesouro da Classificação Internacional de Atenção Primária. A seguir, uma amostra de Declaração de Óbito do Programa de Aprimoramento das Informações de Mortalidade no Município de São Paulo (PRO-AIM) foi codificada manualmente e pelo Iris versão V4.0.34, e as causas foram comparadas. Quando o Iris não conseguiu codificar as causas de morte, ajustes foram feitos no dicionário. *Resultados:* O Iris foi capaz de codificar todas as causas mencionadas em 94,4% das Declarações de Óbito. No entanto, a codificação direta, sem ajustes, foi feita em 50,6% dos casos. Entre as declarações que o *software* não conseguiu codificar completamente, 89,2% incluíam diagnóstico de causas externas (capítulo XX da CID-10). Esse grupo de causas foi o que mostrou menor concordância ao se comparar a codificação feita pelo Iris com a manual. *Conclusão:* O *software* teve um bom desempenho, mas mostra necessidade de ajustes e de desenvolvimento de seu dicionário. A questão das causas externas constitui um problema que os programadores do Iris estão resolvendo nas novas versões. *Palavras-chave:* Causas de morte. Estatísticas vitais. Mortalidade. Sistemas de informação. Atestado de óbito. Automação. Registros de mortalidade.

^IPrograma de Pós-Graduação em Saúde Pública, Universidade de São Paulo – São Paulo (SP), Brasil.

^{II}Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São Paulo – São Paulo (SP), Brasil.

Autor correspondente: Renata Cristófani Martins. Rua Alcantarilla, 17, apto. 112, CEP 05717-170, São Paulo, SP, Brasil. E-mail: renatacristofani@gmail.com

Conflito de interesses: nada a declarar – **Fonte de financiamento:** nenhuma.

ABSTRACT: Objective: To prepare a dictionary in Portuguese for using in Iris and to evaluate its completeness for coding causes of death. **Methods:** Initially, a dictionary with all illness and injuries was created based on the International Classification of Diseases – tenth revision (ICD-10) codes. This dictionary was based on two sources: the electronic file of ICD-10 volume 1 and the data from Thesaurus of the International Classification of Primary Care (ICPC-2). Then, a death certificate sample from the Program of Improvement of Mortality Information in São Paulo (PRO-AIM) was coded manually and by Iris version V4.0.34, and the causes of death were compared. Whenever Iris was not able to code the causes of death, adjustments were made in the dictionary. **Results:** Iris was able to code all causes of death in 94.4% death certificates, but only 50.6% were directly coded, without adjustments. Among death certificates that the software was unable to fully code, 89.2% had a diagnosis of external causes (chapter XX of ICD-10). This group of causes of death showed less agreement when comparing the coding by Iris to the manual one. **Conclusion:** The software performed well, but it needs adjustments and improvement in its dictionary. In the upcoming versions of the software, its developers are trying to solve the external causes of death problem.

Keywords: Cause of death. Vital statistics. Mortality. Information systems. Death certificates. Automation. Mortality registries.

INTRODUÇÃO

As estatísticas de mortalidade são usadas para definir condições de saúde e parâmetros socioeconômicos¹. Por serem medidas de comparações internacionais, se faz necessária uma padronização de conceitos, de métodos de coleta e de análises dos dados. Como marco para as estatísticas de mortalidade, em 1948, a Organização Mundial da Saúde (OMS) definiu o modelo internacional do atestado de óbito, que passou a ser utilizado em muitos países, inclusive no Brasil. Além disso, a OMS definiu as regras de seleção para a Causa Básica de Morte (CB), causa essa que deve ser a origem da cadeia de eventos que termina na morte² e que vai ser a causa usada para as estatísticas de saúde.

No Brasil, a Declaração de Óbito (DO) é preenchida pelo médico e analisada por um codificador. Utilizando a Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde - décima revisão (CID-10), esse profissional codifica todas as causas de morte presentes na DO, aplica as regras de seleção da OMS e seleciona a CB. Para realizar essas tarefas, os codificadores passam por um treinamento e têm acesso aos manuais da OMS.

Nesse processo de codificação e seleção, podem ocorrer falhas que comprometem a qualidade das informações de mortalidade. Uma das falhas possíveis é o erro no registro das causas, feito pelo médico, que leva a erros de codificação pela falta de especificação das doenças³. Outra falha possível ocorre na seleção da CB, uma vez que as regras de seleção são complexas e abrangentes. A possibilidade de interpretar de formas diferentes as regras³ e a necessidade de considerar as numerosas exceções tornam difícil determinar a CB⁴, mesmo para codificadores treinados e experientes.

Para reduzir as falhas e melhorar a qualidade dos dados, existem diversas estratégias, entre elas, a sensibilização dos profissionais da área de saúde para os erros cometidos⁴. Outra solução são os *softwares* que simulam o papel do codificador, podendo apenas selecionar a CB ou codificar e selecionar automaticamente as causas mencionadas na DO.

O primeiro tipo de *software* seleciona a CB a partir dos códigos da CID-10 das causas de morte da DO, definidos pelo codificador. Esse tipo de *software* contém as tabelas de decisão que simulam as regras de seleção da OMS. O programa *Automated Classification of Medical Entities* (ACME) desenvolvido pelos Estados Unidos é o mais conhecido e utilizado mundialmente. O Sistema de Seleção da Causa Básica de Morte (SCB) é o programa usado no Brasil e também se enquadra nessa categoria. Ele foi desenvolvido a partir das tabelas de decisão do ACME que estão alteradas para a realidade brasileira⁵.

O segundo tipo de *software* faz a codificação das causas de morte mencionadas no atestado de óbito e depois seleciona a CB. Para isso, é necessário um dicionário de termos médicos com seus respectivos códigos da CID-10. São desse grupo o *Mortality Medical Data System* (MMDS)⁶, desenvolvido pelos Estados Unidos, e o Iris, desenvolvido em parceria entre instituições da França, Alemanha, Hungria, Itália, Espanha, Suécia e Estados Unidos⁷. Esse tipo de *software*, por fazer a codificação e a seleção automáticas, permite uma grande melhora da qualidade dos dados de mortalidade⁸.

Tanto o Iris quanto o MMDS utilizam as tabelas de decisão do ACME para codificar a CB. A diferença entre eles é que o MMDS tem o dicionário intrínseco à sua programação, enquanto o Iris não. Neste último, o dicionário é independente do programa, pode ser facilmente configurado em vários idiomas e, conseqüentemente, utilizado em vários países. Esse é um dos motivos que faz do Iris o *software* mais adaptável para o uso mundial. Além disso, por utilizar as regras de seleção da OMS e o modelo internacional de atestado médico, ele facilita a comparação dos dados de mortalidade entre diferentes países.

Considerando que o uso do Iris constitui um facilitador para se obterem estatísticas mais confiáveis e para melhorar a rotina diária dos codificadores, este estudo teve como objetivo descrever as etapas da adaptação do *software* Iris para o Brasil.

MÉTODOS

A preparação do *software* para uso no Brasil foi feita em etapas; primeiro, preparando o dicionário em Português; depois, testando a completude desse dicionário e adequando-o aos termos usados no país; e, finalmente, comparando sua aplicação à codificação manual das causas de morte.

DESENVOLVENDO O DICIONÁRIO

A versão do Iris utilizada foi a V4.0.34, de 30/03/2011⁹, com interface em Inglês. O dicionário do programa é um arquivo eletrônico que contém duas tabelas, uma com o dicionário propriamente dito e outra com um conjunto de regras de padronização (*Standardisation*). A tabela com o dicionário relaciona cada categoria (as possíveis causas de mortes, termos médicos e diagnósticos) com um código da CID-10. A tabela de padronização contém as

regras de padronização que produzem filtros para que as expressões médicas possíveis de serem escritas na DO possam ser localizadas no dicionário¹⁰.

Para a criação do dicionário em Português, foram utilizadas duas fontes. A primeira foi o arquivo eletrônico do PESQCID¹¹, que contém as categorias de diagnósticos e os seus respectivos códigos listados conforme o volume 1 da CID-10. A segunda fonte foi o Tesouro da Classificação Internacional de Atenção Primária (CIAP-2) desenvolvido pela Organização Mundial de Médicos de Família (antiga WONCA)¹², que contém a lista dos seus diagnósticos com os códigos equivalentes da CID-10.

As duas fontes resultaram em um dicionário com 58.546 categorias. Destas, 12.211 (20,9%) são originárias do PESQCID e 46.335 (79,1%) do Thesaurus. Os termos repetidos, em número de 7.500, foram excluídos, restando um dicionário contendo 51.046 termos diagnósticos e seus respectivos códigos da CID-10.

A etapa seguinte foi a adaptação e padronização do dicionário para que os termos coincidam com os utilizados pelos médicos para preencher a DO. Além de modificar o dicionário, as tabelas de padronização foram utilizadas. A padronização reduz o tamanho do dicionário por utilizar termos-chave e é especialmente importante para países com grande variedade linguística como o Brasil. Como exemplo dessa etapa, pode-se citar a retirada dos sinais gráficos de todos os termos. Assim, o programa reconhece da mesma forma uma grafia com ou sem sinal. Isso é feito por meio de uma regra que faz com que todas as expressões sejam buscadas no dicionário sem o sinal. Além disso, vários sinônimos ou formas diferentes de identificar a mesma condição médica foram padronizados para um único termo.

No final desse processo, o dicionário continha 46.801 categorias com seus respectivos códigos da CID-10 e a tabela de padronização continha 621 regras.

DEFININDO A AMOSTRA DE ATESTADOS DE ÓBITO

Utilizou-se uma amostra de óbitos de residentes do Município de São Paulo, ocorridos no período de 1 a 4 de dezembro de 2010. Os dados foram coletados a partir das primeiras vias das DOs arquivadas no Programa de Aprimoramento das Informações de Mortalidade no Município de São Paulo (PRO-AIM) da Prefeitura do Município de São Paulo, responsável pelo processamento, análise e divulgação das informações de mortalidade da cidade.

As informações coletadas das DOs foram: número da DO; data de nascimento do falecido; data do óbito; sexo e os dados do atestado médico (bloco V). Não houve outra identificação, exceto o número da DO, de modo a manter a privacidade do falecido.

Os codificadores do PRO-AIM marcam, nas DOs, os códigos da CID-10 para cada afecção e indicam a CB selecionada. Essa codificação manual foi usada como base para comparação com o método automático.

Não foram utilizadas informações adicionais além das encontradas nas DOs. Mais especificamente, não foram usados laudos, relatórios médicos ou resultados de investigação de óbito realizada por comitês, para modificação das causas de morte.

PREPARAÇÃO DOS LOTES

Uma das etapas de uso do Iris é a elaboração dos lotes de DO. O lote é um arquivo eletrônico gerado a partir de uma tabela pré-formatada. Nessa tabela, existem células para cada informação utilizada, necessária e produzida pelo *software* sobre cada DO processada. O arquivo eletrônico pode ter vários lotes e cada lote pode ter várias DOs; cabe a quem está utilizando o programa organizar e nomear cada lote. No caso desse teste, os lotes foram nomeados segundo a data de óbito, tendo em média 167 DOs cada.

Preparar um lote significa inserir na tabela as informações referentes a cada DO: sexo; data de nascimento; data do óbito e número de referência. Esse grupo de informações é considerado o mínimo necessário para o funcionamento do Iris.

Para informações não existentes nas DOs, criou-se a seguinte regra: quando a DO era de desconhecido, a data de nascimento foi assumida como 01/01/1950, para que os dados não fossem perdidos. Caso o campo idade estivesse preenchido, a data de nascimento era calculada como se o aniversário do falecido fosse a data do óbito. Quando o campo sexo não estava preenchido na DO, a informação era considerada como ignorada.

USO DO IRIS — O TESTE PROPRIAMENTE DITO

Uma vez finalizada a primeira versão do dicionário em Português e preparados os lotes, foram inseridas todas as DOs no Iris para iniciar o processo de codificação e seleção da CB.

Nas situações em que o Iris era incapaz de identificar e codificar a causa mencionada, a DO era submetida a análise, adotando-se uma das seguintes decisões:

- adicionar o termo da causa de morte no dicionário com, sempre que possível, um código de quatro caracteres da CID-10. Exemplo: choque circulatório com o código R57.9;
- mudar o texto de uma categoria na tabela do dicionário. Exemplo: apagar a expressão “sem especificação(ões)” de todas as categorias; assim, “Enxaqueca, sem especificação” passa a ser “enxaqueca”;
- adicionar uma regra de padronização. Exemplo: padronizar que “sepse grave é sepse”, ou seja, se estiver escrito como causa de morte “sepse grave”, o Iris vai procurar “sepse” no dicionário;
- mudar uma regra de padronização. Exemplo: acrescentar “sepse irreversível” à regra de padronização “sepse grave é sepse”, de modo que tanto “sepse grave” quanto “sepse irreversível” sejam padronizadas para “sepse”.

As informações coletadas formaram um banco de dados tendo com unidade básica a DO. Foram utilizadas para formar o banco de dados os códigos da CID-10 identificados pela codificação manual e pelo Iris, de cada causa mencionada na Parte I e Parte II do atestado de óbito, respeitando sua posição no mesmo; o código CID-10 da CB; as regras de seleção utilizadas e se houve necessidade de adicionar e/ou mudar uma categoria do dicionário ou as regras de padronização.

A tabulação e o banco de dados foram feitos nos programas Microsoft® Excel e Epi Info (TM) 3.5.3. Para a comparação das causas entre os dois tipos de codificação, manual e automática, foi criada uma razão de concordância (RC). Essa RC é obtida pela divisão da proporção de causas de morte codificadas em um determinado capítulo da CID-10, pelo Iris e manualmente. Não houve pareamento das DOs para essa análise. Considerou-se a RC como ótima, moderada ou baixa, conforme descrito na Tabela 1.

A pesquisa foi aprovada pelo Comitê de Ética em Pesquisa da Faculdade de Saúde Pública da Universidade de São Paulo, sob protocolo nº 2.256. Não houve conflito de interesses.

RESULTADOS

O teste utilizou 666 DOs de residentes no Município de São Paulo. A média de idade dos falecidos foi de 65,4 anos e 52,2% eram de indivíduos do sexo masculino. Em 2 casos (0,3%), o campo sexo não estava preenchido ou foi registrado com indefinido. Em 10 (1,5%) outros casos, a data de nascimento não estava preenchida, sendo que, em duas declarações, a informação era da idade aproximada.

O Iris foi capaz de codificar completamente todas as linhas preenchidas do atestado de óbito em 629 (94,4%) declarações e, portanto, estas puderam ter suas CBs selecionadas. Nas declarações em que o *software* não conseguiu codificar todas as causas de morte, verificou-se que, em 89,2% destas, havia algum diagnóstico de causa externa (capítulo XX da CID-10).

O número de códigos completos (com 4 caracteres) da CID-10 usados para codificar essa amostra pelo Iris foi de 362 (2,9% das 12.451 subcategorias existentes) e a codificação manual usou 388 (3,1%). A proporção de códigos por DO foi praticamente igual nos dois sistemas de codificação: 3,4% no Iris e 3,3% na codificação manual.

Comparando as causas de morte codificadas pelos dois métodos, sem pareamento de DO, e avaliando os códigos segundo os capítulos da CID-10, constatou-se que houve uma ótima concordância (Tabela 2).

Ao analisar as DOs pareadas, a codificação das causas de morte pelos dois métodos, Iris e manual, verificou-se uma concordância total em 420 DOs (63,1%). Nesses casos, todos os códigos da CID-10 completos de todas as causas de morte, das Partes I e II do atestado de óbito, concordaram nos dois sistemas de codificação. Em 20 DOs (3%), houve discordância total da codificação de todas as causas de morte; em 226 (33,9%) DOs, houve discordância

Tabela 1. Intervalo de valores da razão de concordância considerada ótima, moderada ou baixa.

	Intervalo dos valores da razão de concordância
Ótima	0,900 – 1,100
Moderada	0,750 – 0,900 ou 1,100 – 1,250
Baixa	< 0,750 ou > 1,250

Tabela 2. Distribuição das causas de morte codificadas pelo Iris e manualmente segundo os capítulos da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde - décima revisão, a razão de concordância dos valores e qualificação da razão de concordância nos óbitos dos residentes no Município de São Paulo de 1 a 4 de dezembro de 2010.

Capítulos da CID-10	Iris	Manual	RC Iris/ manual	QRC
I. Algumas doenças infecciosas e parasitárias	10,8% (245)	9,6% (216)	1,13	M
II. Neoplasias	8,2% (186)	8,3% (186)	0,99	O
III. Doenças do sangue e dos órgãos hematopoéticos e alguns transtornos imunitários	1,0% (22)	0,9% (21)	1,04	O
IV. Doenças endócrinas, nutricionais e metabólicas	4,8% (108)	4,9% (111)	0,97	O
V. Transtornos mentais e comportamentais	1,5% (33)	1,5% (34)	0,96	O
VI. Doenças do sistema nervoso	2,5% (57)	2,5% (57)	0,99	O
VII. Doenças do olho e anexos	0,0% (1)	0,0% (1)	0,99	O
VIII. Doenças do ouvido e da apófise mastóide	0,0% (0)	0,0% (0)	-	-
IX. Doenças do aparelho circulatório	27,4% (621)	26,2% (590)	1,04	O
X. Doenças do aparelho respiratório	18,6% (421)	18,5% (415)	1,01	O
XI. Doenças do aparelho digestivo	4,4% (100)	4,4% (100)	0,99	O
XII. Doenças da pele e do tecido subcutâneo	0,5% (11)	0,5% (11)	0,99	O
XIII. Doenças do sistema osteomuscular e do tecido conjuntivo	0,3% (8)	0,4% (9)	0,88	M
XIV. Doenças do aparelho geniturinário	4,6% (105)	4,8% (107)	0,97	O
XV. Gravidez, parto e puerpério	0,0% (0)	0,0% (0)	-	-
XVI. Algumas afecções originadas no período perinatal	1,1% (26)	1,3% (30)	0,86	M
XVII. Malformações congênitas, deformidades e anomalias cromossômicas	0,3% (8)	0,3% (7)	1,13	M
XVIII. Sintomas, sinais e achados anormais de exames clínicos e de laboratório, não classificados em outra parte	10,0% (226)	9,6% (217)	1,03	O
XIX. Lesões, envenenamento e algumas outras consequências de causas externas	2,2% (50)	2,5% (57)	0,87	M
XX. Causas externas de morbidade e mortalidade	1,5% (35)	3,5% (79)	0,44	B
XXI. Fatores que influenciam o estado de saúde e o contato com os serviços de saúde	0,0% (1)	0,0% (0)	-	-
XXII. Códigos para propósitos especiais	0,0% (0)	0,0% (0)	-	-
Total	100% (2.264)	100% (2.248)	-	-

CID-10: Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde - décima revisão; RC: razão de concordância; QRC: qualificação da razão de concordância; Notas - O: ótima; M: moderada; B: baixa.

em alguma causa de morte. A média de discordância da codificação com as DOs pareadas foi de 35,4% considerando-se o código da CID-10 completo.

Comparando a codificação dos termos de cada linha da DO, encontramos uma diferença de 14,1% considerando-se a codificação completa com os 4 caracteres que compõem o código da CID-10. Se considerarmos apenas concordância em nível de capítulo, essa diferença diminui para 9,0%.

O Iris foi capaz de codificar diretamente e por completo, sem necessidade de ajustes, 337 (50,6%) DOs. Para as demais, foi necessário algum tipo de alteração ou acréscimo de termo no dicionário ou na tabela de padronização. No total, houve 582 ajustes, o que indica que, em média, foi feito quase um ajuste por declaração (Tabela 3).

Foram acrescentados 433 termos na tabela do dicionário, sendo que os capítulos que mais tiveram categorias acrescidas foram: Doenças do aparelho circulatório, com 97 adições (22,4%); Doenças endócrinas nutricionais e metabólicas, com 56 novos termos (12,9%); Neoplasias, com 54 adições (12,5%); e Causas externas de morbidade e mortalidade, com 30 adições (6,9%).

As alterações realizadas após esse primeiro teste resultou em um dicionário com 47.020 categorias e 859 regras na tabela de padronização.

DISCUSSÃO

O Iris foi capaz de codificar e selecionar a causa básica de 94,4% das declarações amostradas. A concordância com a codificação manual foi de 63,1%, indicando a potencialidade do *software* para codificar e selecionar corretamente a CB. Quanto mais for usado, mais adequado pode se tornar.

A baixa razão de concordância na codificação das causas externas (capítulo XX) se deve ao fato de o Iris ainda apresentar dificuldade em codificar esse grupo de causas. Essa limitação é um problema mundial e está sendo corrigida pela equipe que desenvolve o programa. No Brasil, particularmente, existem variações de escrita e especificações de uma causa externa de morte que influenciam na sua codificação. Outra dificuldade sentida é que, na maioria

Tabela 3. Número de ajustes necessários para o Iris codificar, segundo tipo e porcentagem de ajuste por Declaração de Óbito nos óbitos dos residentes no Município de São Paulo de 1 a 4 de dezembro de 2010.

Tipo de ajuste realizado	Frequência	Porcentagem
Adicionar o termo no dicionário	235	35,29
Mudar o texto de um termo do dicionário	73	10,96
Adicionar uma regra de padronização	164	24,62
Mudar uma regra de padronização	110	16,52

desses casos, os dados informados no atestado de óbito são insuficientes para realizar a correta codificação da causa de morte. Os codificadores tiveram acesso a várias informações de outras fontes que o Iris não obteve.

Uma das principais dificuldades relacionadas a esse grupo de causas são as diferentes formas de escrever uma mesma causa de morte. As seguintes expressões, por exemplo, podem ser sinônimas: choque elétrico, eletrocussão, energia físico-química; agente físico eletricidade; queimadura ou outra lesão devida a corrente elétrica. Além disso, um mesmo termo pode ter vários códigos, dependendo da situação da morte. Agente contundente pode ser codificado em Y29 se a circunstância não é definida, X59 se acidente, X70 se suicídio, Y00 se homicídio.

O dicionário inicial não contemplava essa diversidade de termos e, conseqüentemente, os artifícios de padronização ou adição/modificação de um termo do dicionário foram utilizados. Como as variações dos sinônimos são amplas, torna-se difícil criar algumas regras de padronização para um único termo-chave. Além disso, devido a problemas de programação, o Iris ainda não aceita a inclusão de alguns termos no dicionário. Esse é o caso dos códigos iniciados em “W”, “X” ou “Y”, que, na CID-10, representam parte do capítulo de causas externas de morbidade e de mortalidade (capítulo XX). As inclusões feitas no dicionário referentes a termos do Capítulo XX foram, em sua grande maioria, relacionadas a procedimentos médicos, ambulatoriais e/ou cirúrgicos.

Outro fator que dificulta a codificação da causa externa é a ausência de informação na DO sobre a circunstância do óbito e a real causa de morte¹³. Essa é uma lacuna que dificulta a codificação tanto manual como automática. A falha de preenchimento desses dados leva o codificador a utilizar categorias menos específicas para a codificação das causas de morte¹³.

Para melhorar a qualidade dos dados de mortalidade por causa externa, o PRO-AIM, desde 1996, efetua investigações junto ao Instituto Médico Legal (IML) da Cidade de São Paulo, visando esclarecer as circunstâncias dessas mortes. Os documentos consultados para buscar informações adicionais são o boletim de ocorrência, o laudo de conclusão da necropsia e, em alguns casos, o relatório de encaminhamento do corpo, preenchido pelo Serviço de Verificação de Óbitos ou pelo hospital onde ocorreu a morte¹⁴. A partir dessas informações adicionais, a DO é recodificada.

Neste estudo, 63,6% das DOs que apresentaram alguma menção de causa externa haviam passado por necropsia. A codificação manual desses óbitos só foi feita após o acesso às informações adicionais coletadas no IML. No entanto, o Iris utilizou apenas a informação que constava na DO original. Esse fato teve grande impacto e pode explicar a baixa taxa de concordância desse grupo de causas, quando comparados os dois sistemas de codificação.

Quanto à discordância observada dentro de um mesmo capítulo da CID-10, um estudo brasileiro comparou a codificação manual do Sistema de Informação de Mortalidade (SIM) com um codificador manual e constatou uma discordância de 3,7%³, menor do que os 9% encontrados neste estudo. Essa diferença pode ser explicada pelo fato de o Iris ser um *software* recente e estar em processo de aprimoramento.

Um estudo norueguês⁸ comparou a codificação manual de todas as causas de morte mencionadas nas DOs, realizada por quatro codificadores, e obteve uma discordância de

24,2% quando considerada a codificação em nível de 4 caracteres e de 11,2% em nível de 1 caractere do código da CID-10. Uma explicação para a discrepância entre este estudo e o norueguês é a diferença do método utilizado, pois sabe-se que, à medida que aumenta o número de codificadores, diminui a concordância entre eles⁸.

Com frequência, algumas sequências de causas de morte são interpretadas de forma diferente pelos codificadores e até pelo mesmo codificador, em momentos diferentes. O uso do Iris permite que as sequências sejam sempre interpretadas da mesma maneira, aumentando a chance de comparações.

O estudo norueguês⁸ apontou também 15 a 20% de rejeição ou falha na seleção das CBs das DOs quando codificadas automaticamente. A explicação usada foi o fato de o *software* ter encontrado relações causais ambíguas e ser incapaz de solucioná-las. Nesses casos, a codificação manual precisará acontecer.

Soma-se às dificuldades listadas o fato de a versão do Iris utilizada neste estudo ser incapaz de lidar com causas externas ou complicações de quadros médicos ou cirúrgicos. Isso se deve ao fato de o Iris ser um *software* recente, que vem sendo aprimorado gradualmente à medida que é utilizado. Um grupo de especialistas de vários países é responsável pelo desenvolvimento e atualização do *software*. Esse grupo controla também as adaptações dos países, o que permite que a comparação das informações produzidas seja possível. Faz parte das funções do grupo incluir no programa as atualizações da CID-10, solicitar dos usuários do *software* informações para o aprimoramento do programa e melhorar as funções do mesmo, como adaptar seu uso para autópsia verbal¹⁵. Cabe à equipe de mortalidade dos países que usam o Iris apenas atualizar e adaptar melhor o dicionário e suas funções para sua realidade, sem interferir nas tabelas de seleção.

Várias modificações foram realizadas desde a primeira versão do Iris e as perspectivas são de continuidade desses avanços nos próximos anos. O grupo responsável pelo aprimoramento do *software* tem se empenhado em resolver as dificuldades de codificação das causas externas de óbito. As tabelas do dicionário estão sendo adequadas para que uma causa de morte listada na DO possa ser recodificada baseada em informações contidas em outros campos ou até mesmo por causa de outra causa de morte. Outro ponto importante que deve ser mencionado é a obrigatoriedade do uso de um *software* de apoio não livre (Microsoft® Access) para a formação de lotes. Sabe-se que novas alternativas, entre elas, uma plataforma na internet, são consideradas para substituição¹⁰.

O Sistema de Informações de Mortalidade do Brasil aceita uma DO incompleta. Informações como a data de nascimento ou a idade podem não estar preenchidas e, mesmo assim, o sistema processa essa DO. A versão do Iris utilizada neste estudo não permitiu a preparação de lote sem os dados de data de nascimento ou a identificação do sexo. No entanto, há uma nova versão, lançada recentemente, que não necessita da preparação do lote, facilitando seu uso. Caso, por qualquer motivo, haja incompatibilidade de idade ou sexo com alguma causa informada na DO, o programa abre uma tela com perguntas sobre a veracidade das informações digitadas.

Uma outra característica que torna o programa atraente é a possibilidade de tradução da tela adaptando-a às características da DO utilizada pelo país.

CONCLUSÃO

Considerando-se que este foi o primeiro teste do Iris no Brasil, o fato de ele ter sido capaz de codificar diretamente 50,6% das DOs é um ótimo indicador, principalmente pela possibilidade de melhora após ajustes e acréscimos no dicionário e na tabela de padronização. Além disso, o *software* apresentou 63,1% de concordância nas DOs pareadas considerando-se os códigos da CID-10 com 4 caracteres.

São esperados, como parte do desenvolvimento do dicionário, os acréscimos dos termos utilizados na rotina médica e os outros ajustes de adequação ao idioma. Portanto, nos próximos testes, os ajustes irão reduzir conforme aumenta a completude do dicionário.

Além disso, a utilização de uma nova versão do programa com atualizações, principalmente no manejo da codificação das causas externas, certamente mostrará melhor resultado.

REFERÊNCIAS

1. Laurenti R, Mello Jorge MHP, Lebrão ML, Gotlieb SLD. Estatísticas de saúde. 2ª edição. São Paulo: EPU; 2005. p. 54-7.
2. Organização Mundial da Saúde. Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde: CID-10. 3ª edição. São Paulo: Editora da Universidade de São Paulo; 1996. 2 v.
3. Fajardo S, Aerts DRG de C, Bassanesi SL. Acurácia da equipe do Sistema de Informações sobre Mortalidade na seleção da causa básica do óbito em capital no Sul do Brasil. *Cad Saúde Pública* 2009; 25(10): 2218-28.
4. Soares JAS, Horta FMB, Caldeira AP. Avaliação da qualidade das informações em declarações de óbito infantis. *Rev Bras Saúde Matern Infant* 2007; 7(3): 289-95.
5. Santo AH, Pinheiro CE. Uso de microcomputador na seleção da causa básica de morte. *Bol Oficina Sanit Panam* 1995; 119(4): 319-27.
6. U.S. Department of Health and Human Services. Centers for Disease Control and Prevention. National Center for Health Statistics. Proceedings of the International Collaborative Effort on Automating Mortality Statistics, Volume III. Hyattsville: U.S. Department of Health and Human Services; 2006.
7. Lefeuvre D, Pavillon G, Aouba A, Lamarche-Vadel A, Fouillet A, Jouglu E, et al. Quality comparison of electronic versus paper death certificates in France, 2010. *Popul Health Metr* 2014; 12: 3.
8. Harteloh P, Bruin K de, Kardaun J. The reliability of cause-of-death coding in the Netherlands. *Eur J Epidemiol* 2010; 25: 531-8.
9. International Coding System for Causes of Death - IRIS. [software on internet]. Version 4.0.23. França: Institut National de La Santé et de la Recherche Médicale; 2010. Disponível em: <http://www.cepdc.inserm.fr/site4/index2.php?p=download> (Atualizado em 19 de dezembro de 2010; Acessado em 30 de janeiro de 2014).
10. Johansson LA, Pavillon G, Pelikán L, Weber S, Witting B. Iris User Reference Manual V4.1.3. IRIS – Automated coding system for causes of death. Disponível em: <http://www.dimdi.de/dynamic/en/klassi/koop/irisinstitute/downloadcenter/manuals/user-guide/> (Acessado em 28 de janeiro de 2014).
11. Centro Colaborador da OMS para a Classificação de Doenças em Português - CBCD. Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde - CID-10. Versão 2008 eletrônica. Disponível em: <http://www.datasus.gov.br/cid10/v2008/cid10.htm> (Acessado em 25 de fevereiro de 2012).
12. Classificação Internacional de Atenção Primária - CIAP 2. [software em CD-ROM]. 2ª edição. Florianópolis: Comitê Internacional da Classificação da WONCA; 2009.
13. Matos SG, Proietti FA, Barata R de CB. Confiabilidade da informação sobre mortalidade por violência em Belo Horizonte, MG. *Rev Saúde Pública* 2007; 41(1): 76-84.
14. Drumond Jr M, Lira MMTA, Nitri TMV, Shibao K, Taniguchi M, Bourroul MLM. O novo modelo da declaração de óbito e a qualidade das informações sobre causas externas. In: VI Congresso Brasileiro de Saúde Coletiva; 2000 jul; Salvador (BR). Disponível em: http://www.prefeitura.sp.gov.br/cidade/secretarias/upload/saude/arquivos/publicacoes/causas_externas.pdf (Acessado em 30 de janeiro de 2014).
15. Pavillon G. IRIS. Newsletter on the WHO-FIC, Volume 10, number 1, 2012. Disponível em: <http://www.rivm.nl/who-fic/newsletter/newsletter%202012-1.pdf> (Acessado em 30 de janeiro de 2014).

Recebido em: 23/05/2014

Versão final apresentada em: 28/01/2015

Aceito em: 16/03/2015