

## O desafio da aplicação da metodologia de captura-recaptura na vigilância do diabetes mellitus em idosos: lições de uma experiência no Brasil

The challenge of applying capture-recapture methodology to diabetes mellitus surveillance in the elderly: lessons from a Brazilian experience

Cláudia Medina Coeli <sup>1,2</sup>  
Evandro Silva Freire Coutinho <sup>3</sup>  
Renato Peixoto Veras <sup>4</sup>

### Abstract

*The purpose of this study was to evaluate the feasibility of establishing a surveillance system for diabetes mellitus in the elderly using data sources on morbidity (hospital and outpatient data) and mortality and the capture-recapture methodology. Log-linear models were adjusted for the entire sample and for gender and age strata. Seven hundred and forty patients were identified in the three combined sources. The model selected for the whole sample when the first linkage strategy was deployed estimated a total of 22,925 cases (90%CI: 11,354-57,269), an implausibly high figure that suggested a strong final negative dependency between sources. Age and gender stratification led to more plausible point values, which nevertheless still lacked precision. The only subset that displayed precise values was the 60-to-69-year-old female stratum (1,290 cases; 90%CI: 779-2,476). The factors that possibly explain these results are: (a) the low number of cases originated in the specialized and more complex facilities in the outpatient data source and (b) overall underreporting of diabetes mellitus cases.*

*Diabetes Mellitus; Aged; Population Surveillance*

### Introdução

As aplicações da metodologia da captura-recaptura na área de Saúde Pública envolvem a vigilância de doenças transmissíveis e não-transmissíveis, a avaliação da cobertura de registros (registros de doenças de base populacional, por exemplo), o refinamento de estimativas de incidência e prevalência e a derivação de valores plausíveis para os limites superior e inferior de estimativas, nas situações onde estimativas pontuais precisas não possam ser obtidas <sup>1,2</sup>.

O uso rotineiro dessa metodologia no âmbito da Saúde Pública vem sendo defendido por diversos autores <sup>3,4,5</sup>, embora limitações da sua utilização neste contexto sejam apontadas <sup>6,7,8,9,10</sup>.

A utilização da metodologia de captura-recaptura para a implantação de sistemas rotineiros para a vigilância de doenças é certamente uma das aplicações mais interessantes da técnica, embora também seja a de mais difícil implementação. Para que um sistema possa ser implantado a um custo que viabilize a sua manutenção é necessário que estejam disponíveis fontes de dados em saúde diversificadas que permitam cobrir diferentes segmentos da população. Por outro lado, o sistema proposto deve ser avaliado buscando a identificação de possíveis violações das premissas básicas do método (população fechada, validade do diagnóstico, relacionamento de registros eficiente, independência entre fontes e capturabilidade

<sup>1</sup> Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

<sup>2</sup> Núcleo de Estudos de Saúde Coletiva, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

<sup>3</sup> Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

<sup>4</sup> Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil.

#### Correspondência

C. M. Coeli  
Núcleo de Estudos de Saúde Coletiva, Universidade Federal do Rio de Janeiro.  
Av. Brigadeiro Trompowski s/n, Hospital Universitário Clementino Fraga Filho, 5ª andar, Ala Sul, Rio de Janeiro, RJ 21931-059, Brasil.  
coeli@acd.ufrj.br

homogênea), devendo nestes casos ser empregadas soluções metodológicas (utilização de modelos loglineares, por exemplo) que viabilizem a obtenção de estimativas válidas <sup>2</sup>.

Em trabalho anterior, onde uma revisão sobre a metodologia é apresentada <sup>2</sup>, propusemos um modelo teórico para a implantação de um sistema de vigilância do diabetes mellitus na população idosa, baseado na utilização de fontes de morbidade e mortalidade usualmente disponíveis em nosso meio, e na aplicação da metodologia de captura-recaptura. O objetivo do presente estudo é avaliar a factibilidade da implantação deste sistema para a vigilância do diabetes mellitus na população idosa da Área Programática 2.2 (AP2.2), do Município do Rio de Janeiro, usuária do Sistema Único de Saúde (SUS).

## Metodologia

### Base populacional

A base populacional do estudo é formada pelos idosos (idade  $\geq 60$  anos), usuários do SUS, residentes na AP2.2 da cidade do Rio de Janeiro, durante os anos de 1994 e 1995. Tomando por base os resultados obtidos em estudos realizados em outros locais <sup>11,12,13,14</sup>, empregamos a proporção de 70,0% para a definição da população idosa da área usuária do SUS.

### Fontes, captação e processamento de dados

Foram empregadas as bases do Sistema de Informação sobre Mortalidade (SIM) relativas aos anos de 1994 e 1995. A base de 1994, divulgada pela Secretaria de Estado de Saúde do Rio de Janeiro (SES-RJ), não trazia o nome do paciente digitado, sendo necessário proceder à revisão das declarações. Para o ano de 1995, a informação sobre o nome do paciente foi fornecida em meio magnético pela Subgerência de Dados Vitais da Coordenação de Epidemiologia da Secretaria Municipal de Saúde do Rio de Janeiro (SMS-RJ), não sendo necessário proceder à revisão das declarações de óbito (DOs). No ano de 1994 foram revisadas todas as 2.974 DOs relativas a idosos residentes na AP2.2, sendo selecionados para o estudo os 150 óbitos onde o diabetes mellitus constou como causa básica e os 141 onde o diabetes foi registrado como causa associada do óbito. No ano de 1995 foram identificadas 2.857 DOs de idosos residentes na AP2.2, das quais 140 apresentavam o diabetes mellitus como causa básica do óbito. As DOs de idosos que faleceram em hospitais

privados não credenciados ao SUS e em hospitais militares foram excluídas da análise, em função da restrição da base populacional aos idosos usuários do SUS. Quanto aos óbitos ocorridos no domicílio ou em vias públicas, não foi possível identificar aqueles relativos aos pacientes usuários do SUS, sendo todos incluídos na amostra. A base final empregada para a estimativa do número de casos de diabetes entre idosos por meio da metodologia de captura-recaptura foi composta por 291 registros, sendo 190 relativos aos óbitos ocorridos em 1994 e 101 aos ocorridos em 1995.

Com relação à base hospitalar, foram empregados os bancos dos formulários de autorização de internação hospitalar (AIH). As estratégias utilizadas para a seleção de registros em 1994 e 1995 foram semelhantes, i.e., inicialmente foram selecionados todos os registros de formulários AIH do tipo 1 (AIH-1), onde a idade do paciente era  $\geq 60$  anos. Foram excluídos os registros apresentados para faturamento em 94 ou 95, mas com data da alta hospitalar anterior a 1<sup>o</sup> de janeiro de 1994. Adicionalmente, foram incluídos os registros apresentados em 1996 mas que referiam-se a altas no período compreendido entre 1<sup>o</sup> de janeiro de 1994 e 1<sup>o</sup> de dezembro de 1995.

O arquivo gerado nesta etapa foi relacionado baseando-se no campo número da AIH com uma base de dados de nomes e endereços fornecida pelo Departamento de Informação e Informática do SUS (DATASUS). A seguir, foram selecionados os registros que atendiam aos seguintes critérios: procedimento realizado de internação do grupo do diabetes mellitus em adultos (Tabela de Procedimentos – SIH-SUS, código 82.300-04-6) e/ou diagnóstico apresentando o código da Classificação Internacional de Doenças (CID-9) <sup>15</sup>, pertencente à categoria 250 (diabetes mellitus) ou às subcategorias 357.2 (polineuropatia do diabetes mellitus), 362.0 (retinopatia diabética), 251.0 (coma hipoglicêmico), residência na AP2.2. Entretanto, nas bases de 1994 só foi possível a seleção de registros a partir do diagnóstico principal, já que essas bases (arquivos reduzidos) não traziam o campo do diagnóstico secundário, o qual estava presente nas bases de 1995 (arquivos completos). Somando-se os registros de 1994 aos de 1995 e excluindo-se os registros duplicados (reinternações), chegamos a um total de 134 pacientes que constituíram a base de internações hospitalares utilizada para a análise de captura-recaptura.

Ao contrário do observado para as fontes de dados apresentadas anteriormente, não existe um Sistema de Informações Ambulatoriais que

inclua dados sobre morbidade padronizado para todas as unidades credenciadas pelo SUS. Sendo assim, a captação desses dados requereu inicialmente a avaliação da viabilidade da obtenção das informações necessárias junto às unidades de saúde. Esta característica, por outro lado, trouxe uma limitação operacional para o estudo tornando impossível a avaliação de todas as unidades de saúde credenciadas pelo SUS que prestam atendimento ambulatorial. Sendo assim, optamos por eleger para o estudo as unidades que apresentavam a maior probabilidade de captar um caso de diabetes mellitus em idosos residentes na AP2.2 (unidades sentinela). Dentro deste critério foram eleitas as seguintes unidades: unidades municipais de saúde da área que prestam atendimento ambulatorial e que são pólos de referência para o Programa de Diabetes Mellitus (quatro unidades); unidades da área que apresentem ambulatório na especialidade de endocrinologia (das três presentes apenas uma, que apresentava registros disponíveis para a identificação do local de residência dos pacientes, foi selecionada); uma unidade especializada localizada na área dedicada ao atendimento aos pacientes idosos; um hospital especializado da rede estadual, referência para o tratamento de diabetes mellitus, localizado fora da AP2.2 no bairro do Centro (AP1).

Em todas as demais unidades avaliadas procuramos identificar os pacientes que atendiam aos seguintes critérios: idade  $\geq 60$  anos, diagnóstico de diabetes, endereço de residência na AP2.2 e pelo menos um atendimento dentro da janela temporal de interesse para o presente estudo (1<sup>o</sup> de janeiro de 1994 a 31 de dezembro de 1995). Do total de 373 pacientes identificados, dez foram captados em mais de uma unidade. Sendo assim, excluindo-se as repetições, ficamos com 363 pacientes diferentes que constituíram a base de pacientes ambulatoriais que foi utilizada para o ajuste dos modelos de captura-recaptura. A maior proporção dos casos (76,3%, 267/363) que integrou a base ambulatorial foi captada junto às unidades de saúde da rede básica.

Os bancos foram relacionados empregando-se o programa RecLink<sup>16</sup>, que implementa o método de relacionamento probabilístico. Foram utilizados como identificadores as variáveis nome, sexo e data de nascimento.

### **Análise de dados**

Foram ajustados modelos loglineares para o conjunto da amostra e nos estratos formados pela faixa etária e sexo (60 a 69 anos, masculi-

no e feminino; 70 anos ou mais, masculino e feminino), sendo empregado o fator de ajuste proposto por Hook & Regal<sup>17</sup>. Para cada modelo ajustado procedeu-se ao cálculo da estatística da razão de verossimilhança e dos seguintes critérios de informação: “*Akaike Information Criterion*” (AIC) e duas versões do “*Bayesian Information Criterion*” (BIC), denominadas DIC e SIC<sup>17,18,19</sup>. Adicionalmente, foram calculadas as estimativas ponderadas utilizando a metodologia proposta por Draper<sup>20</sup> e empregando-se os diferentes critérios de informação (DIC ponderado, SIC ponderado e AIC ponderado)<sup>17</sup>. Os intervalos de confiança de 90,0% das estimativas foram calculados empregando-se o método da razão de verossimilhanças proposto por Regal & Hook<sup>21</sup>. As análises foram realizadas empregando-se o programa Stata, versão 5.

### **Resultados**

A distribuição dos casos segundo a situação de captação nas diferentes fontes de dados é apresentada na Figura 1. Ao todo foram selecionados 740 pacientes diferentes no conjunto das três fontes estudadas; destes, 695 foram captados por cada uma das fontes uma única vez, enquanto os demais foram captados simultaneamente em mais de uma fonte (Figura 1).

A amostra apresentou um predomínio do sexo feminino ( $n = 478$ ; 64,6%) e de idosos na faixa etária de 60 a 69 anos ( $n = 322$ ; 43,5%). Na Figura 2 é apresentada a proporção de casos captados por cada fonte de dados dentro dos diferentes estratos formados pela combinação das variáveis faixa etária e sexo. Nota-se a grande importância da fonte ambulatorial, especialmente no sexo feminino, na captação dos casos de idosos na faixa de 60 a 69 anos. Uma proporção expressiva de casos no estrato de idosas com 70 anos ou mais também foi captada a partir da fonte ambulatorial. Essa fonte apresentou importância reduzida no estrato dos idosos do sexo masculino com 70 anos ou mais. A fonte hospitalar foi menos importante para a captação dos casos no estrato das idosas com 60 a 69 anos, e respondeu por aproximadamente 20,0% dos casos captados nos demais estratos. Por fim, a fonte de mortalidade respondeu pela maioria dos casos captados nos estratos de idosos com 70 anos ou mais, principalmente entre os homens. Essa fonte apresentou uma importância reduzida no grupo das idosas jovens.

Nas Tabelas 1, 2 e 3 são apresentados os diferentes modelos loglineares ajustados para o conjunto da amostra e em cada estrato forma-

Figura 1

Número de idosos com diagnóstico de diabetes segundo a situação de captação nas diferentes fontes de dados.

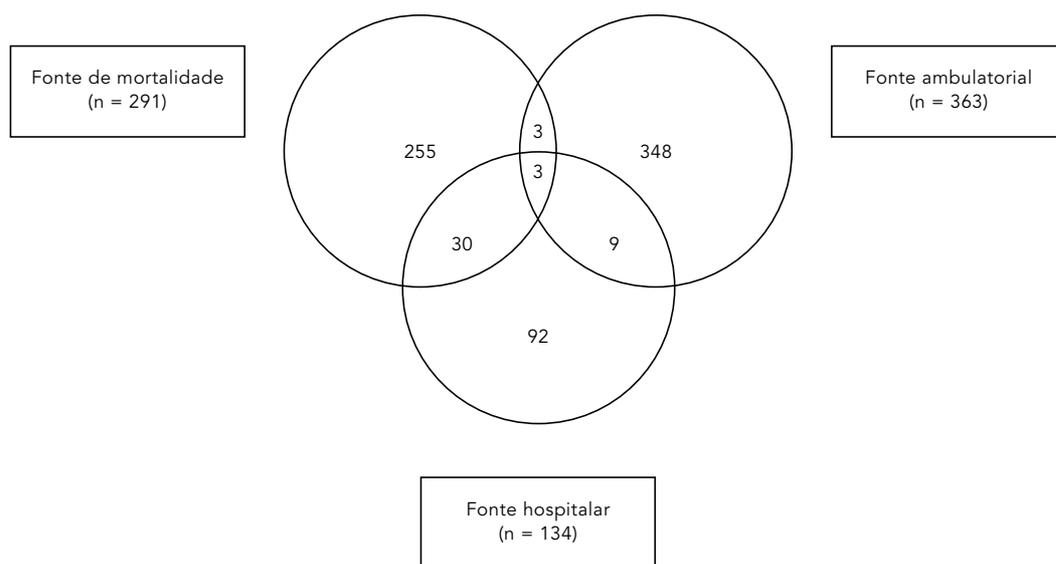
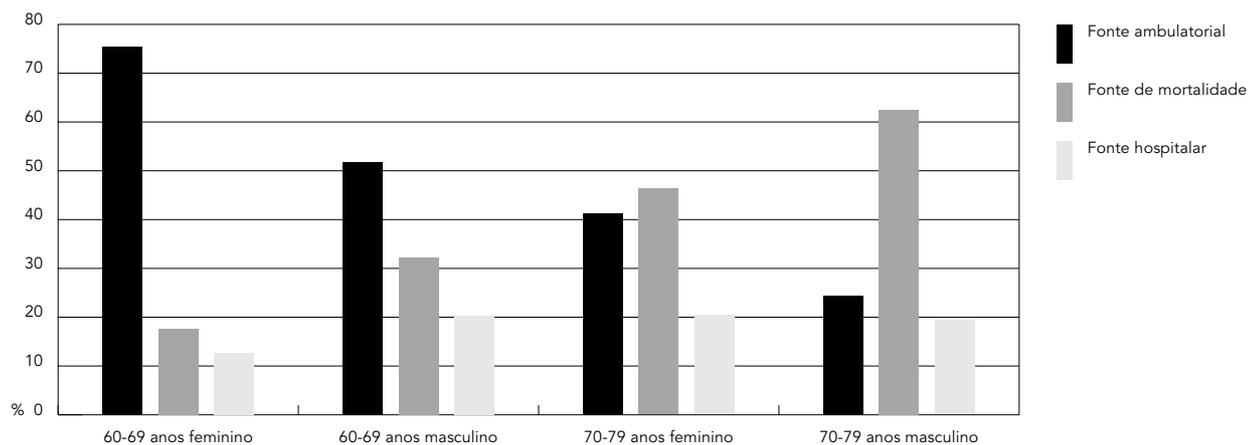


Figura 2

Proporção de casos captados nas diferentes fontes de dados em cada estrato formado pela combinação das variáveis faixa etária e sexo\*.



\* A soma dos percentuais em cada estrato ultrapassa 100%, pois um caso pode ter sido captado por mais de uma fonte.  
 Total de casos em cada estrato: 60-69 anos feminino (204); 60-69 anos masculino (118); 70-79 anos feminino (274); 70-79 anos masculino (144).

Tabela 1

Modelos loglineares calculados para o conjunto da amostra.

Modelo	gl	G <sup>2</sup>	AIC	SIC	DIC	$\hat{n}$	$\hat{N}$	IC90%	
<b>Global</b> (número de casos observados = 740)									
1. [A] [H] [M]	3	70,27	64,27	50,45	55,96	3.033	3.773	3.093	4.700
2. [A,H] [M]	2	70,27	66,27	57,06	60,73	3.020	3.760	3.000	4.802
3. [A,M] [H]	2	24,94	20,94	11,73	15,40	1.269	2.009	1.671	2.482
4. [H,M] [A]	2	13,18	9,18	-0,03	3,64	7.738	8.478	5.449	12.780
5. [A,H] [A,M]	1	11,40	9,40	4,79	6,63	757	1.497	1.270	1.836
6. [A,M] [H,M]	1	4,51	2,51	-2,10	-0,26	3.202	3.942	2.642	6.561
7. [A,H] [H,M]	1	0,03	-1,97	-6,58	-4,74	22.185	22.925	11.354	57.269
8. [A,M] [A,H] [H,M]	0	0,00	0,00	0,00	0,00	19.752	20.492	5.297	79.760
DIC Ponderado							20.840	10.026	53.856
SIC Ponderado							20.618	10.162	51.834
AIC Ponderado							20.869	9.171	59.046

Aplicação do fator de ajuste para "amostras pequenas" proposto por Hook & Regal<sup>17</sup>.

Modelo selecionado empregando-se os critérios AIC, DIC ou SIC: [ ] efeito principal; [ , ] interação.

Os modelos são representados a partir dos termos de maior ordem. A presença de um termo de interação implica obrigatoriamente a inclusão dos termos de menor ordem envolvidos na interação.

Legenda: A = fonte ambulatorial; H = fonte hospitalar, M = fonte de mortalidade; **gl** = graus de liberdade do modelo em ajuste; **G<sup>2</sup>** = estatística da razão de verossimilhanças, calculada em relação ao modelo em ajuste; **AIC** = Akaike Information Criterion (Critério de Informação Akaike); **DIC** = Bayesian Information Criterion (Critério de Informação Bayesiano) proposto por Draper<sup>20</sup>; **SIC** = Bayesian Information Criterion (Critério de Informação Bayesiano) proposto por Schwarz<sup>17</sup>;  $\hat{n}$  = valor estimado para casos não captados por nenhuma fonte;  $\hat{N}$  = valor estimado do número total de casos; **IC90%** = intervalo de confiança de 90% do número total de casos estimados, calculado pelo método baseado na estatística da razão de verossimilhanças proposto por Regal & Hook<sup>21</sup>.

do pela faixa etária e sexo. Os critérios de informação AIC, DIC e SIC coincidiram na seleção do melhor modelo, i.e., apresentaram os menores valores para o mesmo modelo, tanto na análise relativa ao conjunto da amostra (Tabela 1), como nas análises realizadas em cada estrato (Tabelas 2 e 3). Os modelos mais complexos, e especialmente os modelos saturados, tanto na análise global como nos diferentes estratos, apresentaram intervalos de confiança mais amplos. Especificamente, nos dois estratos envolvendo o sexo masculino, os modelos saturados estimaram um número de casos desconhecidos nulo e apresentaram os limites superiores dos intervalos de confiança excessivamente elevados. Isto se deu em função de não ter sido observado nenhum caso nas interseções entre as três fontes em conjunto nesses estratos. Levando em consideração este achado, optamos por excluir o modelo saturado do cálculo das estimativas ponderadas nos estratos.

Ainda com relação às estimativas ponderadas, nota-se que os valores obtidos com as diferentes formulações são semelhantes. Em quase todos os estratos as estimativas ponderadas obtidas com a utilização do critério AIC (AIC ponderado) apresentaram os maiores valores, enquanto as obtidas com a utilização do crité-

rio SIC (SIC ponderado) apresentaram os menores valores.

No conjunto da amostra, o modelo selecionado incluiu dois termos de interação envolvendo as fontes ambulatorial e hospitalar e as fontes hospitalar e de mortalidade. O valor da estimativa associada a este modelo (22.925 casos), assim como os valores das estimativas ponderadas, foram muito elevados (Tabela 1).

Para os idosos na faixa etária de 60 a 69 anos, em ambos os sexos, o modelo selecionado foi o que incluiu o termo de interação entre a fonte hospitalar e de mortalidade (Tabela 2). No estrato de idosos com 70 anos ou mais, o modelo selecionado incluiu o termo de interação entre as fontes ambulatorial e de mortalidade, enquanto para as idosas nesta faixa etária o modelo selecionado foi igual ao escolhido para o conjunto da amostra, i.e., selecionou-se o modelo que incluiu os termos de interação envolvendo as fontes ambulatorial e de mortalidade, e as fontes hospitalar e de mortalidade (Tabela 3).

Na Tabela 4 são apresentadas para cada estrato formado pelas variáveis faixa etária e sexo as estimativas populacionais, as estimativas do número e casos de diabetes mellitus, obtidas tanto a partir dos modelos selecionados em ca-

Tabela 2

Modelos loglineares calculados para os estratos de idosos com 60 a 69 anos.

Modelo	gl	G <sup>2</sup>	AIC	SIC	DIC	$\hat{n}$	$\hat{N}$	IC90%	
<b>60 a 69 anos feminino</b> (número de casos observados = 204)									
1. [A] [H] [M]	3	22,11	16,11	6,16	11,67	535	739	531	1.110
2. [A,H] [M]	2	20,78	16,78	10,14	13,82	415	619	433	985
3. [A,M] [H]	2	18,95	14,95	8,31	11,99	322	526	370	846
4. [H,M] [A]	2	1,23	-2,77	-9,41	-5,73	1.086	1.290	779	2.476
5. [A,H] [A,M]	1	9,20	7,20	3,88	5,72	81	285	238	403
6. [A,M] [H,M]	1	1,09	-0,91	-4,23	-2,39	1.341	1.545	660	6.614
7. [A,H] [H,M]	1	0,97	-1,03	-4,35	-2,51	1.341	1.545	691	4.522
8.[A,M] [A,H] [H,M]	0	0,00	0,00	0,00	0,00	4.023	4.227	695	35.511
DIC Ponderado							1.358	749	3.321
SIC Ponderado							1.323	765	2.881
AIC Ponderado							1.400	731	3.832
<b>60 a 69 anos masculino</b> (número de casos observados = 118)									
1. [A] [H] [M]	3	12,90	6,90	-1,41	4,10	516	634	395	1.158
2. [A,H] [M]	2	11,08	7,08	1,54	5,21	382	500	311	940
3. [A,M] [H]	2	7,30	3,30	-2,24	1,43	258	376	242	690
4. [H,M] [A]	2	0,57	-3,43	-8,97	-5,30	1.769	1.887	752	7.657
5. [A,H] [A,M]	1	0,03	-1,97	-4,74	-2,90	87	205	163	371
6. [A,M] [H,M]	1	0,33	-1,67	-4,44	-2,60	1.159	1.277	407	11.380
7. [A,H] [H,M]	1	0,05	-1,95	-4,72	-2,88	2.013	2.131	635	19.458
8.[A,M] [A,H] [H,M]	0	0,00	0,00	0,00	0,00	0	118	118	>1E+304
DIC Ponderado							1.539	580	8.305
SIC Ponderado							1.656	646	7.792
AIC Ponderado							1.468	539	8.577

Aplicação do fator de ajuste para "amostras pequenas" proposto por Hook &amp; Regal 17.

Modelo selecionado empregando-se os critérios AIC, DIC ou SIC: [ ] efeito principal; [ , ] interação.

Os modelos são representados a partir dos termos de maior ordem. A presença de um termo de interação implica obrigatoriamente a inclusão dos termos de menor ordem envolvidos na interação.

Legenda: A = fonte ambulatorial; H = fonte hospitalar, M = fonte de mortalidade; **gl** = graus de liberdade do modelo em ajuste; **G<sup>2</sup>** = estatística da razão de verossimilhanças, calculada em relação ao modelo em ajuste; **AIC** = Akaike Information Criterion (Critério de Informação Akaike); **DIC** = Bayesian Information Criterion (Critério de Informação Bayesiano) proposto por Draper 20; **SIC** = Bayesian Information Criterion (Critério de Informação Bayesiano) proposto por Schwarz 17;  $\hat{n}$  = valor estimado para casos não captados por nenhuma fonte;  $\hat{N}$  = valor estimado do número total de casos; **IC90%** = intervalo de confiança de 90% do número total de casos estimados, calculado pelo método baseado na estatística da razão de verossimilhanças proposto por Regal & Hook 21.

da estrato como por meio da formulação ponderada empregando-se o critério DIC (DIC ponderado) e os valores dos limites inferior e superior dos intervalos de confiança de 90% das estimativas.

As estimativas associadas ao melhor modelo e as estimativas ponderadas apresentaram valores aproximadamente semelhantes nos estratos de idosas e idosos com idade entre 60 e 69 anos e no de idosos com 70 anos ou mais. No grupo das idosas com 70 anos ou mais, a estimativa obtida com base no melhor modelo foi bem superior àquela calculada pelo critério

DIC ponderado, além de estar associada a um intervalo de confiança bastante amplo. Com relação a este último aspecto, nos demais estratos também foram observados intervalos de confiança bastante amplos, sendo o estrato de idosas de 60 a 69 anos aquele que apresentou a estimativa mais precisa.

A prevalência pontual estimada para o conjunto da amostra obtida a partir da soma das estimativas de cada estrato foi de 22,2%, quando se empregou para o cálculo da prevalência as estimativas do número de casos associadas ao melhor modelo, e de 14,4% quando se em-

Tabela 3

Modelos loglineares calculados para os estratos de idosos com 70 anos ou mais.

Modelo	gl	G <sup>2</sup>	AIC	SIC	DIC	$\hat{n}$	$\hat{N}$	IC90%	
<b>≥ 70 anos feminino</b> (número de casos observados = 274)									
1. [A] [H] [M]	3	28,79	22,79	11,95	17,46	852	1.126	860	1.541
2. [A,H] [M]	2	28,79	24,79	17,56	21,24	857	1.131	836	1.621
3. [A,M] [H]	2	6,65	2,65	-4,58	-0,90	358	632	498	849
4. [H,M] [A]	2	6,44	2,44	-4,79	-1,11	2.167	2.441	1.494	4.520
5. [A,H] [A,M]	1	2,44	0,44	-3,17	-1,34	248	522	420	700
6. [A,M] [H,M]	1	0,83	-1,17	-4,78	-2,95	770	1.044	640	2.151
7. [A,H] [H,M]	1	0,57	-1,43	-5,04	-3,21	5.885	6.159	2.425	25.186
8. [A,M] [A,H] [H,M]	0	0,00	0,00	0,00	0,00	2.648	2.922	476	22.955
DIC Ponderado							2.838	1.305	9.942
SIC Ponderado							2.519	1.233	8.117
AIC Ponderado							3.031	1.349	11.046
<b>≥ 70 anos masculino</b> (número de casos observados = 144)									
1. [A] [H] [M]	3	12,51	6,51	-2,40	3,11	445	589	404	940
2. [A,H] [M]	2	7,26	3,26	-2,68	1,00	668	812	494	1.551
3. [A,M] [H]	2	1,78	-2,22	-8,16	-4,48	202	346	249	538
4. [H,M] [A]	2	11,77	7,77	1,83	5,51	568	712	423	1.435
5. [A,H] [A,M]	1	0,29	-1,71	-4,68	-2,84	269	413	269	775
6. [A,M] [H,M]	1	0,14	-1,86	-4,83	-2,99	118	262	193	451
7. [A,H] [H,M]	1	2,47	0,47	-2,50	-0,66	2.635	2.779	825	24.982
8. [A,M] [A,H] [H,M]	0	0,00	0,00	0,00	0,00	0	144	144	>3E+244
DIC Ponderado							525	289	2.296
SIC Ponderado							461	278	1.568
AIC Ponderado							566	295	2.765

Aplicação do fator de ajuste para "amostras pequenas" proposto por Hook &amp; Regal 17.

Modelo selecionado empregando-se os critérios AIC, DIC ou SIC: [ ] efeito principal; [ , ] interação.

Os modelos são representados a partir dos termos de maior ordem. A presença de um termo de interação implica obrigatoriamente a inclusão dos termos de menor ordem envolvidos na interação.

Legenda: A = fonte ambulatorial; H = fonte hospitalar, M = fonte de mortalidade; **gl** = graus de liberdade do modelo em ajuste; **G<sup>2</sup>** = estatística da razão de verossimilhanças, calculada em relação ao modelo em ajuste;**AIC** = Akaike Information Criterion (Critério de Informação Akaike); **DIC** = Bayesian Information Criterion (Critério de Informação Bayesiano) proposto por Draper 20; **SIC** = Bayesian Information Criterion (Critério de Informação Bayesiano) proposto por Schwarz 17;  $\hat{n}$  = valor estimado para casos não captados por nenhuma fonte;  $\hat{N}$  = valor estimado do número total de casos; **IC90%** = intervalo de confiança de 90% do número total de casos estimados, calculado pelo método baseado na estatística da razão de verossimilhanças proposto por Regal & Hook 21.

pregou as estimativas ponderadas calculadas baseando-se no critério DIC (DIC ponderado). As prevalências pontuais nos demais estratos são apresentadas na Tabela 4.

## Discussão

Os valores de prevalência do diabetes encontrados por estudo realizado no Rio de Janeiro ao final da década de 80 22, fornecem uma importante base de comparação para os resultados por nós encontrados. A revisão de estudos

sobre a prevalência de diabetes mellitus entre idosos realizada nas últimas décadas em diferentes países mostra resultados bastantes variáveis, tornando fundamental a existência de um padrão local 23,24,25,26,27. Esta variabilidade reflete tanto diferenças reais na prevalência como também na metodologia utilizada e nas faixas etárias avaliadas. Os que foram desenvolvidos com base em entrevistas ou em fontes de dados secundários permitiram a avaliação do diabetes conhecido. Estes tenderam a apresentar prevalências menores do que aquelas observadas pelos estudos que se basearam na

Tabela 4

Estimativas da população idosa, do número de casos de diabetes e da prevalência empregando o modelo selecionado e as estimativas ponderadas.

Estratos	População	Estimativas obtidas com o modelo selecionado			Estimativas ponderadas (DIC ponderado)		
		Ñ	IC90%	Prevalência	Ñ	IC 90%	Prevalência
60-69 anos feminino	14.262	1.290	779-2.476	9,0%	1.358	749-3.321	9,5%
60-69 anos masculino	9.611	1.887	752-7.657	19,6%	1.518	580-8.305	15,8%
≥ 70 anos feminino	12.849	6.159	2.425-25.186	47,9%	2.838	1.305-9.942	22,9%
≥ 70 anos masculino	6.554	346	249-538	5,3%	525	289-2.296	8,0%
<b>Total geral</b>	<b>43.276</b>	<b>9.682</b>	<b>4.205-35.857</b>	<b>22,2%</b>	<b>6.239</b>	<b>2.922-23.864</b>	<b>14,4%</b>

realização do teste de tolerância oral à glicose e na aplicação dos critérios diagnósticos propostos pela Organização Mundial da Saúde (OMS) <sup>28</sup>. Os resultados destes últimos estudos apontaram uma proporção de desconhecimento do diabetes entre idosos variando de 25,0% a 76,0%, embora a grande maioria tenha identificado valores em torno de 40,0%. Em outras palavras, tomando por base os resultados dos estudos que empregaram os critérios da OMS para o diagnóstico, espera-se em média encontrar aproximadamente 1,7 vez mais idosos diabéticos (casos conhecidos e desconhecidos) do que o número de casos conhecidos.

Considerando o conjunto da amostra, o modelo selecionado por todos os três critérios de informação foi o que envolveu os termos de interação entre as fontes ambulatorial e hospitalar e entre as fontes hospitalar e de mortalidade. Este modelo estimou um número de casos implausivelmente elevado ( $n = 22.925$ ; prevalência de casos conhecidos = 53,0%), já que o inquérito realizado no Rio de Janeiro <sup>22</sup> evidenciou uma prevalência de diabetes total de 17,4% na faixa etária de 60 a 69 anos, e uma prevalência de diabetes auto-referido de 16,6% na faixa etária de 70 anos ou mais. A ocorrência de dependência negativa entre as fontes pode explicar o valor hiperestimado de casos, i.e., encontrou-se um número de casos nas interseções entre fontes menor do que o esperado ao acaso <sup>2</sup>.

As estimativas ponderadas também apresentaram valores hiperestimados e intervalos de confiança amplos. Este resultado era esperado, já que poucos modelos apresentaram os critérios de informação com valores negativos. Sendo assim, a estimativa associada ao modelo selecionado apresentou um fator de ponderação bem mais expressivo que as estimativas dos demais modelos, exercendo uma maior influência no cálculo das estimativas ponderadas.

A análise do perfil de casos nas diferentes fontes corroborou a presença de heterogeneidades de capturabilidade associadas às variáveis sexo e idade. Por outro lado, o ajuste de modelos loglineares dentro de cada estrato e a soma das estimativas obtidas em cada um deles levou à obtenção de um número total de casos estimado bem menor do que o observado durante a análise bruta (9.682 casos quando se empregou as estimativas do melhor modelo, e 6.239 quando se empregou as estimativas ponderadas calculadas com base no critério DIC). Este fato sugere que, pelo menos parcialmente, a forte dependência final negativa entre as fontes observadas na análise bruta foi determinada pela variabilidade de capturabilidade induzida pelas variáveis sexo e faixa etária, ou por outros fatores associados a estas duas variáveis.

Infelizmente, o problema do número de casos reduzidos nas interseções entre as fontes se intensificou com a estratificação, o que, associado ao menor número de casos por estrato, dificultou o ajuste dos modelos. Chama a atenção o fato de que, à exceção do estrato das idosas com 70 anos ou mais, que apresentou o maior número de casos, em todos os demais estratos os modelos selecionados foram os que envolveram uma única interação. Mais do que indicar um comportamento heterogêneo da dependência entre as fontes nos diferentes estratos, este achado reflete o baixo poder para a seleção de modelos mais complexos.

No estrato das mulheres com 70 anos ou mais, o modelo selecionado foi o que envolveu a interseção entre as fontes ambulatorial e hospitalar e entre as fontes hospitalar e de mortalidade. Como somente um caso foi observado na célula relativa à interseção entre a fonte ambulatorial e de mortalidade, este modelo produziu uma estimativa implausivelmente elevada, além de um intervalo de confiança bastan-

te amplo. Entretanto, ao contrário do verificado na análise para o conjunto da amostra, um maior número de modelos apresentou os valores de critérios de informação negativos, o que fez com que as estimativas ponderadas se tornassem menores e menos imprecisas.

O estrato das idosas com 60 a 69 anos foi o que apresentou os melhores resultados, o que em parte se deve ao fato dos efetivos das células neste grupo serem mais balanceados. A estimativa associada ao melhor modelo e a estimativa ponderada permitiram calcular uma prevalência de diabetes conhecido, respectivamente, de 9,0% e 9,5%. Estes resultados são compatíveis com os observados no estudo do Rio de Janeiro <sup>22</sup> acima citado, i.e., aplicando-se o fator de correção de 1,7, obteríamos uma prevalência total no estrato próxima a 16,0%. Adicionalmente, estas estimativas apresentaram um nível de precisão aceitável.

A interpretação dos resultados por nós obtidos deve ser feita levando em consideração o fato das premissas impostas para a aplicação da metodologia de captura-recaptura terem sido ou não respeitadas <sup>2</sup>. Em função do período de estudo ter sido curto (dois anos) e de se esperar uma baixa migração no grupo dos idosos, acreditamos que o “fechamento demográfico” tenha sido relativamente respeitado. Neste trabalho consideramos que a população avaliada, apesar de dinâmica, tenha se mantido estável ao longo do período de observação.

Já a suposição do “fechamento geográfico” não foi atendida integralmente. O primeiro problema que enfrentamos foi a inacessibilidade às informações sobre as estatísticas ambulatoriais e hospitalares nas classes de maior renda. Isto por um lado levou a uma limitação do sistema de vigilância proposto, que passou a ter o seu escopo de abrangência reduzido, e por outro, tornou necessário conhecer a proporção dos idosos usuários do SUS na AP2.2. Nós escolhemos o percentual da cobertura do SUS tomando por base os dados de literatura, e não os resultados de um estudo realizado na área. É interessante ressaltar, que o valor por nós empregado (70,0%) é coerente com os percentuais dos óbitos que ocorreram em hospitais privados e militares durante os anos de 1994 e 1995, os quais em conjunto representaram aproximadamente 30,0% do total de óbitos em idosos nestes anos. No entanto, pode ser esperado que este percentual esteja subestimado para as idosas, já que estas geralmente apresentam renda inferior e escolaridade menor do que os idosos <sup>29</sup>.

Com relação ao “fechamento geográfico”, o maior problema, entretanto, foi que a estratê-

gia amostral empregada não permitiu captar de forma adequada toda a experiência da base populacional. Na fonte ambulatorial, em função do número restrito de unidades avaliadas, ficaram sub-representados: os casos graves do diabetes; os casos leves, porém associados a outras doenças que demandavam um atendimento mais complexo; os pacientes do sexo masculino; e os pacientes com idade mais avançada. Já nas fontes hospitalar e de mortalidade, em função do esperado sub-registro de casos, ficaram sub-representados os casos mais leves do diabetes, e possivelmente, também os pacientes do sexo masculino com idade mais avançada <sup>30</sup>.

A violação da suposição do “fechamento geográfico” foi responsável pelo reduzido número de registros nas interseções entre as fontes ambulatorial e as demais fontes, e também pela atenuação da dependência positiva entre as fontes hospitalar e de mortalidade. Estes dois fatores em conjunto determinaram a ocorrência de uma dependência final negativa bastante acentuada, levando a resultados hiperestimados e imprecisos. A estratificação segundo sexo e faixa etária produziu estimativas pontuais mais plausíveis, porém imprecisas.

Cabe ressaltar que o estrato das idosas jovens, único a apresentar resultados plausíveis e precisos, possivelmente foi o grupo onde o “fechamento geográfico” foi melhor alcançado. Como as mulheres utilizam serviços ambulatoriais mais intensamente que os homens, poderia ser esperado o diagnóstico e tratamento mais precoce tanto do diabetes como de outras doenças, fazendo com que preponderasse neste estrato o perfil de casos que usualmente são atendidos na rede básica de saúde.

O presente estudo teve por objetivo avaliar a factibilidade da implantação de um sistema de vigilância do diabetes mellitus na população idosa em nosso meio, baseado na utilização da metodologia de captura-recaptura. Com base nesta análise, pretendíamos ainda determinar até que ponto esta nova metodologia poderia representar, tal qual proposto por La-Porte <sup>31</sup>, a “solução” para o problema da vigilância epidemiológica das doenças não-transmissíveis, permitindo o monitoramento eficiente de várias doenças simultaneamente em áreas geográficas extensas. Com estes objetivos em mente, buscamos propor um sistema que tivesse como principais características a agilidade para a captação e o processamento dos dados, o que seria fundamental para viabilizar a sua implantação no futuro.

O sub-registro nas fontes hospitalar e de mortalidade, a baixa representatividade na fonte

ambulatorial dos casos mais graves do diabetes e dos casos leves que cursam associados a outras doenças graves, foram os fatores que maior impacto tiveram sobre os resultados e que também são de mais difícil solução, pelo menos a curto prazo.

Para minimizar o problema do sub-registro seria necessário que todos os diagnósticos mencionados na declaração de óbito passassem a ser digitados, o que no Rio de Janeiro vem sendo feito a partir do ano de 1999. Já no caso do SIH-SUS, seria necessário ampliar o número de campos para o registro do diagnóstico e, principalmente, iniciar um programa de estímulo ao preenchimento dos mesmos. Quanto à maior cobertura das diferentes formas de manifestações do diabetes, seria necessário aumentar o número de unidades especializadas no tratamento da doença, incluindo as unidades localizadas fora da área de interesse, e dentro das unidades já escolhidas, avaliar os ambulatórios de outras especialidades (ex. serviços de cardiologia e fisioterapia). Isto só será factível, porém, se as unidades passarem a contar com sistemas informatizados para o gerenciamento do atendimento ambulatorial. Adicionalmente, casos em tratamento de terapia renal substitutiva poderiam ser identificados nas bases do Sistema de Informação de Alto Custo/Alta Complexidade (APAC), que se tornaram disponíveis a partir de 1996.

A disponibilidade de fontes de dados representa, portanto, o principal fator limitante pa-

ra a implantação de sistemas rotineiros de vigilância de doenças crônico-degenerativas baseados na utilização da metodologia de captura-recaptura. Apesar do considerável avanço que assistimos neste campo em nosso meio nos últimos anos, muito ainda precisa ser feito para que os sistemas de vigilância possam ser implantados a um custo aceitável. O diabetes mellitus em idosos certamente representa uma das doenças mais difíceis de serem avaliadas, em função das múltiplas formas com que se apresenta e de sua elevada prevalência. Considerando a analogia com as aplicações da área de ecologia<sup>2</sup>, a população de diabéticos se comportaria como a população de peixes de um reservatório grande, onde seria necessário utilizar um número grande de armadilhas de forma a garantir que toda a extensão do reservatório fosse coberta. Aspectos especiais do diabetes como a insuficiência renal em terapia substitutiva e a amputação de membros inferiores, esta última já estudada em nosso meio<sup>32</sup>, representam eventos menos prevalentes e com fontes de dados adicionais disponíveis, sendo mais indicados para a aplicação da metodologia de captura-recaptura. Adicionalmente, outras doenças (alguns tipos de câncer, por exemplo) cujo atendimento é realizado dentro de unidades especializadas e para as quais os erros de diagnósticos são de menor importância, certamente representariam situações mais fáceis para a implementação da vigilância.

## Resumo

*O presente estudo buscou avaliar a factibilidade da implantação de um sistema de vigilância do diabetes mellitus na população idosa, empregando para tal fontes de dados sobre morbidade (hospitalar e ambulatorial) e mortalidade e a metodologia de captura-recaptura. Modelos loglineares foram ajustados para o conjunto da amostra e nos estratos formados pelo sexo e a faixa etária. Setecentos e quarenta pacientes foram identificados no conjunto das fontes. O modelo selecionado para o conjunto da amostra estimou um total de 22.925 casos (IC90%: 11.354-57.269), que é implausivelmente elevado sugerindo uma forte dependência*

*negativa final entre as fontes. A estratificação segundo sexo e faixa etária levou a valores pontuais mais plausíveis, embora ainda imprecisos. O único estrato que apresentou valores precisos foi o de mulheres com 60 a 69 anos (1.290 casos; IC90%: 779-2.476). Os fatores que possivelmente explicam estes resultados são: (a) o número pequeno dos casos captados na fonte ambulatorial atendidos em unidades especializadas e de maior complexidade; (b) o sub-registro do diabetes.*

*Diabetes Mellitus; Idosos; Vigilância da População*

## Colaboradores

C. M. Coeli participou das fases de planejamento, coleta de dados, análise de resultados e redação do texto final. R. P. Veras e E. S. F. Coutinho participaram das fases de planejamento, análise de resultados e redação do texto final.

## Agradecimentos

A Angela Maria Cascão e Cláudio P. Noronha, por permitirem o acesso às bases de dados de mortalidade. A Jacques Levin, por permitir acesso às bases de dados hospitalares. A Anelise Fiad, Berenice Ribeiro, Eurydice Germano Marque, Flávia M. Leal, Leda Suely Barbosa César, Maria Inez Padula Anderson e Ricardo Meireles, por permitirem o acesso aos dados ambulatoriais. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (processo nº 474845/2001-0) pelo financiamento da pesquisa.

## Referências

1. Hook EB, Reagal RR. Capture-recapture methods in epidemiology: methods and limitations. *Epidemiol Rev* 1995; 17:243-64.
2. Coeli CM, Veras RP, Coutinho E. Metodologia de captura-recaptura: uma opção para a vigilância das doenças não-transmissíveis na população idosa. *Cad Saúde Pública* 2000; 16:1071-82.
3. Bruno G, LaPorte RE, Mereletti F, Biggeri A, McCarty D, Pagano G. National diabetes programs. Application of capture-recapture to count diabetes? *Diabetes Care* 1994; 17:548-56.
4. LaPorte RE, Dearwater SR, Chang Y, Aaron DJ, Anderson RL. Efficiency and accuracy of disease monitoring systems: application of capture-recapture methods to injury monitoring. *Am J Epidemiol* 1995; 142:1069-77.
5. McCarty DJ, Tull E, Moy CS, Kwok CK, LaPorte RE. Ascertainment corrected rates: applications of capture-recapture methods. *Int J Epidemiol* 1993; 22:559-65.
6. Black JFP, McLarty DG, Mtasiwa D. Difficult to use in developing countries. *BMJ* 1994; 308:531.
7. Hook EB, Regal RR. Capture-recapture methods. *Lancet* 1992; 339:742.
8. Papoz L, Balkau B, Lellouch J. Case counting in epidemiology: limitations of methods based on multiple data sources. *Int J Epidemiol* 1996; 25: 474-8.
9. Schouten LJ, Straatman H, Kiemeny LLM, Gimbrere CHF, Verbeek AALM. The capture-recapture method for estimation of cancer registry completeness: a useful tool? *Int J Epidemiol* 1994; 23: 1111-6.
10. Waters WE. More unreliable in humans than birds. *BMJ* 1994; 308:531.
11. Anderson MIP. Saúde e condições de vida do idoso no Brasil [Dissertação de Mestrado]. Rio de Janeiro: Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro; 1997.
12. Cesar CLG, Tanaka OY. Inquérito domiciliar como instrumento de avaliação de serviços de saúde: um estudo de caso na região sudoeste da área metropolitana de São Paulo, 1989-1990. *Cad Saúde Pública* 1996; 12 Suppl 2:59-70.
13. Soares MS, Andreazzi MFS. Utilização dos sistemas de informações de saúde em avaliações de impactos ambientais: limitações, perspectivas e soluções metodológicas – Sistema de Informação Ambulatorial. In: Heller L, Monteiro TCN, Salles MJ, Almeida LM, Cancio J, organizadores. Saneamento e saúde nos países em desenvolvimento. Rio de Janeiro: CC&P Editores; 1997. p. 49-65.
14. Coeli CM. O Sistema de Informações Hospitalares do SUS (SIH-SUS): um instrumento para a avaliação do impacto do Programa de Despoluição da Baía de Guanabara sobre a saúde. In: Heller L, Monteiro TCN, Salles MJ, Almeida LM, Cancio J, organizadores. Saneamento e saúde nos países em desenvolvimento. Rio de Janeiro: CC&P Editores; 1997. p. 76-87.
15. Organização Mundial da Saúde. Classificação estatística internacional de doenças, lesões e causas de óbitos – nona conferência de revisão. São Paulo: Centro Brasileiro de Classificação de Doenças em Português; 1985.
16. Camargo Jr. KR, Coeli CM. RecLink: aplicativo

- para o relacionamento de bases de dados, implementando o método probabilistic record linkage. *Cad Saúde Pública* 2000; 16:439-47.
17. Hook EB, Regal RR. Validity of methods for model selection, weighting for model uncertainty, and small sample adjustment in capture-recapture estimation. *Am J Epidemiol* 1997; 145:1138-44.
  18. Agresti A. *Categorical data analysis*. New York: John Wiley & Sons; 1990.
  19. Coeli CM. *Vigilância do diabetes mellitus em uma população idosa: aplicação da metodologia de captura-recaptura [Tese de Doutorado]*. Rio de Janeiro: Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro; 1998.
  20. Draper D. Assessment and propagation of model uncertainty. *J R Stat Soc* 1995; 57:45-97.
  21. Regal RR, Hook EB. The effects of model selection on confidence intervals for the size of a closed population. *Stat Med* 1991; 10:717-21.
  22. Oliveira JEP. *Prevalência do diabetes mellitus no Rio de Janeiro [Tese de Doutorado]*. Rio de Janeiro: Faculdade de Medicina, Universidade Federal do Rio de Janeiro; 1992.
  23. Galting W, Tufail S, Mullee MA, Wetacott TA, Hill RD. Mortality rates in diabetic patients from a community-based population compared to local age/sex matched controls. *Diabet Med* 1997; 14: 316-20.
  24. Blanchard JF, Ludwig S, Waja A, Dean H, Anderson K, Kendall O, et al. Incidence and prevalence of diabetes in Manitoba, 1986-1991. *Diabetes Care* 1996; 19:807-11.
  25. Bourdel-Marchasson I, Dubroca B, Manciet G, Decamps A, Emerian J, Dartigues F. Prevalence of diabetes and effect on quality of life in older french living in the community: the PAQUID Epidemiological Survey. *J Am Geriatr Soc* 1997; 45: 295-301.
  26. Stolk RP, Pols HAP, Lamberts SWJ, Jong PTVM, Hofman A, Grbbee DE. Diabetes mellitus, impaired glucose tolerance, and hyperinsulinemia in an elderly population. The Rotterdam Study. *Am J Epidemiol* 1997; 145:24-32.
  27. Tuomilehto J, Nissinen A, Kivelä S-L, Pekkanen J, Kaarsalo E, Wolf E, et al. Prevalence of diabetes mellitus in elderly men aged 65 to 84 years in eastern and western Finland. *Diabetologia* 1986; 29:611-5.
  28. World Health Organization. *Diabetes mellitus*. Geneva: World Health Organization; 1985. (Technical Report Series 727).
  29. Veras RP. *País jovem com cabelos brancos. A saúde do idoso no Brasil*. Rio de Janeiro: Relume Dumará/Universidade do Estado do Rio de Janeiro; 1994.
  30. Coeli CM, Ferreira LGFD, Drbal MM, Veras RP, Caramargo Jr. KR, Cascão AM. Mortalidade em idosos por diabetes mellitus como causa básica e associada. *Rev Saúde Pública* 2002; 36:135-40.
  31. LaPorte RE. Assessing the human condition: capture-recapture techniques. *BMJ* 1994; 308:5-6.
  32. Spichler ER, Spichler D, Lessa I, Costa E, Forti A, Franco LJ, et al. Capture-recapture method to estimate lower extremity amputation rates in Rio de Janeiro, Brazil. *Rev Panam Salud Publica* 2001; 10:334-40.

---

Recebido em 25/Set/2003

Versão final reapresentada em 26/Mai/2004

Aprovado em 19/Jul/2004