

Utilização de dados secundários do SIM, Sinasc e SIH na produção científica brasileira de 1990 a 2006

Eliane de Freitas Drumond*
Carla Jorge Machado**
Maria do Rosário Vasconcelos***
Elisabeth França****

Trata-se de estudo descritivo da distribuição e de algumas características dos artigos científicos publicados de 1990 a 2006 que utilizaram dados secundários dos Sistemas de Informação sobre Mortalidade (SIM), Nascidos Vivos (Sinasc) e/ou Informações Hospitalares (SIH). Após pesquisa bibliográfica no banco de dados MEDLINE/PubMed, com base no MeSH (Medical Subject Headings), selecionaram-se 294 artigos. Foram analisadas variáveis referentes aos artigos, autores e periódicos. Observou-se aumento de quatro vezes no volume de publicações e nas edições em inglês. Houve crescimento no número de periódicos utilizados, nas instituições de origem dos pesquisadores, na utilização conjunta de duas bases de dados e nas pesquisas dirigidas às doenças e agravos não-transmissíveis (especialmente as cardiovasculares e causas externas). Os resultados apontam expressiva ampliação da produção científica brasileira que utiliza dados do SIM, Sinasc e/ou SIH.

Palavras-chave: Bibliometria. Estatísticas vitais. Pesquisa/tendências. Indicadores de produção científica. Mortalidade.

Introdução

Os Sistemas de Informação em Saúde (SIS), desenvolvidos e implantados pelo Ministério da Saúde em todo o Brasil, têm se tornado, cada vez mais, ferramentas necessárias para gestão loco-regional do sistema de saúde e vêm apresentando consolidação e qualificação progressivas (LEVY, 2007; PAES, 2007; THEME FILHA et al., 2007; ROMERO; CUNHA, 2006). Cada um dos SIS atende a diferentes finalidades, sendo desenvolvido, implantado e operacio-

nalizado de forma distinta dos demais. No Sistema de Informação sobre Mortalidade (SIM) e no Sistema de Informação sobre Nascidos Vivos (Sinasc), por exemplo, são armazenados e processados dados sobre as estatísticas vitais. Além dos números de nascidos vivos e óbitos que possibilitam o cálculo direto de importantes indicadores de saúde, como a mortalidade infantil, o Sinasc e o SIM disponibilizam outras informações fundamentais, como o peso ao nascer e a causa básica da morte, respectivamente. Já no Sistema de Informação Hospitalar

* Técnica da Gerência de Epidemiologia e Informação, da Secretaria Municipal de Saúde de Belo Horizonte.

** Professora do Departamento de Demografia, Cedeplar, Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais.

*** Bibliotecária da Biblioteca J. Baeta Viana da Faculdade de Medicina da Universidade Federal de Minas Gerais.

**** Professora do Departamento de Medicina Preventiva e Social da Faculdade de Medicina da Universidade Federal de Minas Gerais.

(SIH), único de abrangência nacional sobre morbidade hospitalar, estão os registros sobre internações nos hospitais públicos municipais, estaduais ou federais e naqueles contratados/conveniados integrantes da rede SUS.

A disponibilização, no sítio da Internet do Departamento de Informática do Sistema Único de Saúde (Datasus), dos dados do SIM, Sinasc e SIH desagregados até o nível dos municípios, mas sem as variáveis identificadoras, impulsionou sua utilização para análises epidemiológicas.

O acesso gratuito e em meio eletrônico a essas bases de dados também tem facilitado as avaliações de cobertura, de completitude do preenchimento das variáveis e de confiabilidade desses SIS. Por sua vez, quando necessários e atendidos os preceitos éticos inerentes às pesquisas científicas e após aprovação pelos Comitês de Ética em Pesquisa, os dados contendo variáveis identificadoras podem ser obtidos junto aos serviços de saúde com relativa facilidade.

O SIM foi desenvolvido em 1976 e sua implantação iniciou-se em 1977. Encontra-se disponível na Internet a série histórica, que se inicia em 1979, com informações importantes e necessárias para as estatísticas de mortalidade, inclusive as causas básicas de morte.

As causas de morte declaradas pelo médico foram codificadas segundo regras estabelecidas pela Organização Mundial da Saúde, por meio da Classificação Estatística Internacional de Doenças e Problemas Relacionados (CID). Embora a qualidade das informações do SIM ainda seja controversa (ROMERO; CUNHA, 2006), observa-se crescente ampliação da sua cobertura (CARDOSO et al., 2005; PAES, 2005) e confiabilidade (LAURENTI et al., 2005).

No Sinasc, implantado pelo Ministério da Saúde a partir de 1990, podem ser obtidas informações sobre condições de saúde materno-infantil, por meio de variáveis com as características da mãe – como parturição, duração da gravidez e tipo de parto – e do recém-nascido – como peso de nascimento e Apgar. Esse sistema é reconhecido pela sua boa cobertura (HARAKI et al., 2005),

maior completitude (ALMEIDA et al., 2006) e confiabilidade (LEVY, 2007; THEME FILHA et al., 2004; HARAKI et al., 2005) em comparação ao SIM. A maior qualidade do Sinasc pode estar relacionada às diferenças nas rotinas de registro dos seus dados pelos serviços de saúde e na forma de coleta e processamento pelas secretarias municipais de saúde (HARAKI et al., 2005).

O SIH foi desenvolvido, principalmente, para o processamento das informações administrativas sobre as internações hospitalares na rede SUS. Nele também podem ser obtidos dados epidemiológicos que possibilitam a caracterização da população atendida e da morbimortalidade nos hospitais da rede.

Contudo, as possibilidades de exploração do SIH para ampliação das análises epidemiológicas focalizadas no atendimento hospitalar da rede SUS têm sido pouco utilizadas, devido, principalmente, às controvérsias a respeito da qualidade dos dados sobre o diagnóstico da internação e ao pequeno número de estudos sobre a confiabilidade das suas informações (BITTENCOURT et al., 2006). Estima-se que, em razão das suas características contábeis, o sub-registro de internações neste sistema seja pequeno (BITTENCOURT et al., 2006) e que a média nacional de cobertura do SIH varie entre 70% e 80% (BITTENCOURT et al., 2006; MINISTÉRIO DA SAÚDE, 2007), sendo inversamente proporcional ao desenvolvimento socioeconômico de cada localidade.

A descentralização das ações, dos serviços e da gestão, impulsionada na década de 1990 no Brasil, teve como um de seus resultados a evolução nos conceitos de informação em saúde e de epidemiologia nos serviços de saúde. A epidemiologia passou a ser reconhecida, cada vez mais, como recurso estratégico para a construção e consolidação do SUS e as informações contidas nas bases de dados nacionais valorizaram-se, ao serem cada vez mais necessárias para as formulações e avaliações de políticas públicas voltadas para a melhoria das condições de saúde da população. Consequentemente, aumentaram as necessidades de qualifica-

ção das informações e de conhecimento e uso de ferramentas epidemiológicas pelos serviços de saúde (DRUMONG JÚNIOR, 2006). Deve-se observar, no entanto, que, a despeito da crescente importância atribuída à informação produzida pelos serviços de saúde, os problemas relativos à produção e à gestão de registros de qualidade ainda são negligenciados.

Embora nos estudos científicos realizados a partir dos SIS possam ser observadas as vantagens ligadas à utilização de dados secundários em pesquisas científicas (como o viés de informação, por exemplo), uma das maiores vantagens dessa utilização é permitir aos pesquisadores a realização, a baixo custo, de estudos de base populacional e com abrangência nacional. Com base em perguntas adequadas e reconhecendo as potencialidades e limitações dos SIS, os trabalhos científicos possibilitam e disponibilizam novas formas de análises da situação de saúde e de avaliação dos serviços (PAIM, 2006).

Artigos científicos publicados em periódicos indexados representam apenas uma parcela da produção científica. Reconhece-se, no entanto, a importância dessas publicações para a divulgação do conhecimento científico, pois são caracterizadas pelo controle de qualidade feito por pares. A possibilidade de acesso a tais publicações foi ampliada graças à democratização do uso de microcomputadores, à sua crescente divulgação em meio eletrônico e à maior facilidade de utilização da Internet.

Por meio da Internet, pode-se ter acesso imediato, de forma parcial ou integral, a artigos científicos publicados em periódicos indexados e disponibilizados por bases de dados bibliográficas. Estão disponíveis, pela Biblioteca Virtual em Saúde (BVS) – *website* ligado à Organização Mundial de Saúde –, pesquisas em bases de dados como MEDLINE, LILACS, WHOLIS, PAHO, SciELO, entre outras.

A MEDLINE/PubMed, coordenada pela *National Library of Medicine* (NLM), é a maior base mundial de dados na área da saúde e contém *abstracts* e citações bibliográficas de mais de 5.000 periódicos biomédicos sobre medicina, enfermagem, odontologia

e veterinária, entre outros, sendo atualizada semanalmente. Além de possibilitar a recuperação de artigos científicos, a MEDLINE/PubMed oferece a possibilidade de estimar e avaliar a evolução da produção científica em determinada área, identificar o ano da publicação, o periódico e o assunto, entre outros. Por outro lado, estudos de citações e de afiliação dos autores não podem ser realizados por meio dessa base de dados. Importantes periódicos brasileiros para as áreas de saúde coletiva e epidemiologia ainda não estão incluídos nesta base. Mas, apesar de não representar o universo dos mais importantes periódicos nacionais, tem sido observado aumento do volume das publicações científicas brasileiras, em geral, e das áreas das ciências da saúde, em especial, em bases de dados bibliográficas como ISI (*Institute for Scientific Information*) e MEDLINE/PubMed (RODRIGUES et al., 2000; BARRETO, 2006; BLANK et al., 2006), a despeito das limitações orçamentárias e da ainda pequena comunidade científica nacional, se comparadas aos países desenvolvidos.

O objetivo do presente estudo foi realizar análise descritiva sobre a distribuição e algumas características da produção científica nacional com utilização de dados secundários provenientes do SIM, Sinasc e/ou SIH referenciadas na base de dados MEDLINE/PubMed, no período de 1990-2006.

Metodologia

Realizou-se estudo descritivo a partir do levantamento do universo de registros da produção científica internacional e nacional e dos registros nos quais as fontes de dados utilizadas foram o SIM, o Sinasc e/ou o SIH, recuperados a partir da base de dados MEDLINE/PubMed em janeiro de 2007, correspondente ao período de 17 anos (1990 a 2006).

Fontes de informação

Os artigos foram recuperados na base MEDLINE, interface PubMed, subconjunto da base PubMed, que indexava 5.164 periódicos em janeiro de 2007. Devido às

atualizações semanais, o total de registros da MEDLINE/PubMed para determinado período pode variar (PACKER et al., 2007). Os descritores, que podem ser utilizados em pesquisas bibliográficas e estudos bibliométricos, representam os assuntos discutidos nos artigos e são extraídos do vocabulário MeSH – *Medical Subject Headings*.

Filtros temáticos

O ano de início do período do estudo (1990) coincidiu com o da implantação do Sinasc. Os dados foram extraídos para cada ano, de primeiro de janeiro a 31 de dezembro de 1990 a 2006. Cada ano foi definido por meio da opção *Limits* da interface PubMed.

A pesquisa foi feita em duas etapas. Na primeira, destinada à obtenção da participação anual de artigos com estudos brasileiros em relação ao total da base de dados bibliográfica, foram realizados o levantamento do universo de citações e o das citações de artigos referentes ao país Brasil, no campo de palavras *Brazil*.

A segunda etapa teve dois objetivos: caracterizar se um artigo utilizou dados secundários provenientes do SIM, Sinasc e/ou SIH; e calcular a participação dos artigos selecionados no universo da produção científica brasileira, resultado obtido na etapa anterior.

Para atingir o primeiro objetivo, foram utilizados, concomitantemente, os descritores do MESH *Mortality OR Pregnancy OR Pregnancy Rate OR Pregnancy Outcome OR Live Birth OR Infant, newborn OR Hospital Information Systems OR Information Systems OR Databases, factual* e no campo de palavras *SIM OR SINASC OR SIH*, combinados (AND) com *Brazil* no campo de palavras. Foram incluídos apenas artigos originais e selecionados tão somente aqueles cujas pesquisas utilizaram fontes de dados oficiais do Ministério da Saúde.

Análise dos dados

O software SPSS versão 10.0 para Windows (SPSS Inc., Chicago, Estados Unidos) foi utilizado para armazenamento e

análise dos dados. Para fins de análise, os 17 anos do estudo foram divididos em dois períodos, segundo a década de publicação do artigo: de 1990 a 1999 (dez anos) e de 2000 a 2006 (sete anos).

As variáveis selecionadas em cada um dos artigos foram: ano de publicação; idioma; bases de dados secundárias utilizadas; faixa etária da população estudada; e classificação do conteúdo. Dos periódicos, selecionaram-se os nomes e o país de origem. Em relação ao primeiro autor, analisaram-se as instituições de origem e as regiões do país onde se localizavam. As faixas etárias da população do estudo foram classificadas como perinatal (de 22 semanas a sete dias de vida), neonatal (de zero a 27 dias de vida), infantil (de zero a um ano de vida), criança (de zero a dez anos), adolescente (de 11 a 18 anos), adultos (de 19 a 60 anos) e idosos (mais de 60 anos). Quando na metodologia do estudo não havia especificação sobre a idade da população estudada, classificou-se a faixa etária como não especificada.

Para classificação do seu conteúdo, os artigos foram agrupados em cinco categorias temáticas, obtidas a partir dos descritores de cada artigo e consideradas mutuamente excludentes, as quais foram desagregadas em subcategorias: causas de morbimortalidade (doenças transmissíveis, doenças crônicas, causas externas); saúde materno-infantil (mortalidade perinatal, neonatal e infantil e mortalidade materna); sistemas de informação (SIM, Sinasc, SIH); assistência à saúde (cuidado pré-natal/tocologia, outras); e a categoria outros (composta das subcategorias desigualdade em saúde, evitabilidade, poluição, doenças profissionais, etnia e saúde e demais), criada para enquadrar artigos que não pertenciam às categorias anteriores.

Os resultados foram tabulados pela evolução ano a ano do total de artigos recuperados por sistema de informação utilizado e sua representação no total dos artigos indexados; faixa etária estudada e temas em destaque, a partir dos demais descritores utilizados; periódicos em que foram publicados os artigos; instituição de origem do primeiro autor e sua localização nas grandes regiões geográficas do país; e idioma

das publicações. Finalmente, realizou-se a comparação dos resultados obtidos entre os dois períodos selecionados.

Resultados

Em janeiro de 2007, na primeira etapa da busca, foram recuperados 8.522.732 artigos publicados na base MEDLINE/PubMed, no período de 1990 a 2006, dos quais 0,29% (24.351) eram referentes ao Brasil. A partir de janeiro de 2007, novos artigos publicados no mesmo período do estudo ingressaram na base de dados, mas não foram incluídos neste estudo.

A Tabela 1 apresenta a distribuição anual do total de artigos e dos artigos brasileiros recuperados nesta base de dados bibliográfica. A participação anual de artigos

brasileiros em relação ao total aumentou de 0,13%, em 1990, para 0,40%, em 2006. Em números absolutos, os artigos brasileiros passaram de 516 para 2.856, nesse mesmo período, correspondendo a um aumento de 453,5%, enquanto o crescimento do total de artigos publicados na mesma base foi de 78,9%.

Por meio da combinação *booleana* de descritores do *MESH*, foram recuperados 2.373 artigos. Os resumos estavam disponibilizados em 2.156 artigos (90,8%), dos quais 1.712 (75%) também apresentavam o texto completo. Quando o resumo não permitia a inclusão/exclusão do artigo, procedeu-se a leitura dos textos na íntegra. Assim, foram selecionados 294 artigos que utilizaram as fontes de dados oficiais do Ministério da Saúde (Tabela 1). A média

TABELA 1
Total de artigos, artigos referentes ao Brasil e artigos selecionados na base MEDLINE/PubMed 1990-2006

Ano da publicação	Total de artigos	Artigos brasileiros		Artigos selecionados	
		N. Abs.	Participação s/ o total (%)	N. Abs.	Participação s/ o total de artigos brasileiros (%)
1990	399.926	516	0,13	8	1,55
1991	401.431	642	0,16	4	0,62
1992	405.750	768	0,19	2	0,26
1993	413.360	715	0,17	4	0,56
1994	423.609	733	0,17	5	0,68
1995	434.111	760	0,18	3	0,39
1996	444.408	834	0,19	5	0,60
1997	442.652	898	0,20	10	1,11
1998	461.410	1186	0,26	10	0,84
1999	479.798	1253	0,26	18	2,08
2000	521.197	1536	0,29	32	1,98
2001	533.456	1666	0,31	33	1,46
2002	553.508	2050	0,37	30	1,32
2003	584.071	2341	0,40	31	1,16
2004	624.265	2664	0,43	31	1,26
2005	684.156	2933	0,43	37	1,09
2006	715.624	2856	0,40	31	1,40
Total	8.522.732	24.351	0,29	294	1,20
Crescimento de 1990 a 2006 (%)	78,94	453,49		287,50	

Fonte: Base de dados MEDLINE/PubMed.

Nota: Os artigos foram contabilizados mais de uma vez, quando utilizada mais de uma fonte entre SIM, Sinasc e SIH.

anual de artigos publicados no período correspondeu a 17,3. Observa-se, em 2000, um elevado número de artigos em comparação aos anos anteriores. Apesar disso, verifica-se aumento de 287,5% no número anual de publicações, que pode ser constatado pela diferença entre o número de artigos publicados em 1990 (8) e em 2006 (31).

Nesta tendência ascendente a partir de 1996, houve oscilações no período 1996-2001 e, em seguida, uma estabilização após o forte aumento de publicações detectadas no MEDLINE (com exceção de 2005, em que foram recuperados 37 artigos).

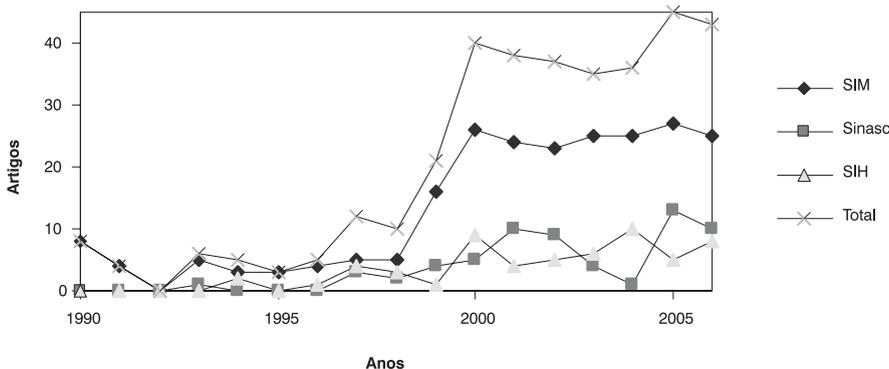
No Gráfico 1, observa-se a distribuição anual de artigos científicos recuperados na base MEDLINE/PubMed, no período de 1990 a 2006, segundo o SIS utilizado. A base de dados mais usada foi o SIM (232 artigos), seguida pelo Sinasc (63) e pelo SIH (59). Em 80,3% dos artigos (236), foi empregado somente um SIS e, em 19,7% (58), houve utilização simultânea do SIM/Sinasc (37), SIM/SIH (19), Sinasc/SIH (1) e SIM/Sinasc/SIH (1).

Aproximadamente 70% dos textos publicados foram escritos exclusivamente em português. Os periódicos que apresentaram um ou mais artigos sobre o tema totalizaram 54. As publicações nacionais veicularam 81% (238) do total de textos selecionados, destacando-se, com números bastante similares, a *Revista de Saúde Pública* (98) e os *Cadernos de Saúde Pública* (94). Os

outros periódicos com maior volume de publicações foram os *Arquivos Brasileiros de Cardiologia* (16), *Arquivos de Neuropsiquiatria* (5), *São Paulo Medical Journal* (4), *Revista da Associação Médica Brasileira* (4), *Jornal de Pediatria* (3) e as *Revistas da Sociedade Brasileira de Medicina Tropical e Brasileira de Enfermagem* (dois artigos em cada uma). A *Revista Panamericana de Salud Pública* (8), o *Journal of Epidemiology and Community Health* (4) e o *International Journal of Epidemiology* (3) foram os periódicos internacionais que publicaram o maior número de artigos.

Sessenta e quatro instituições de origem dos primeiros autores foram citadas pelo menos uma vez, entre as quais se encontravam instituições de ensino, de pesquisa ou serviços de saúde públicos ou privados. Em 72,1% dos casos (212) eram instituições localizadas na Região Sudeste, seguidas pelo Nordeste (42) e Sul (31). Na Região Sudeste foram expressivos os números de trabalhos publicados pela Universidade de São Paulo (59), Escola Nacional de Saúde Pública/Centro de Informação Científica e Tecnológica/Fundação Oswaldo Cruz (50), Universidade Federal do Rio de Janeiro (13) e Universidade Federal de Minas Gerais (12). Também verificaram-se publicações que tiveram como primeiros autores pesquisadores originários dos serviços de saúde, como as Secretarias Estaduais de Saúde do Rio de Janeiro (13)

GRÁFICO 1
Artigos científicos recuperados na MEDLINE/PubMed, segundo SIS utilizado 1990-2006



Fonte: Base de dados MEDLINE/PubMed.

e de São Paulo (12) e as Secretarias Municipais de Saúde do Rio de Janeiro (7) e de Belo Horizonte (5). Nas Regiões Nordeste e Sul destacaram-se a Universidade Federal da Bahia (17) e a Universidade Federal do Rio Grande do Sul (7), respectivamente. O maior número de artigos da Região Norte

pertence a pesquisadores da Universidade Federal do Maranhão (3).

A Tabela 2 apresenta os resultados encontrados na classificação dos temas dos artigos. Considerando-se o enfoque dado pelos textos a partir da análise dos seus conteúdos, observou-se maior número de

TABELA 2
Artigos selecionados na MEDLINE/PubMed, segundo classificação do conteúdo
1990-2006

Conteúdo	1990-1999	2000-2006	Total
Causas			
Doenças transmissíveis			
Aids	1	9	10
Tuberculose	0	6	6
Diarréias	0	5	5
Sífilis	0	2	2
Outras	2	5	7
Total	3	27	30
Doenças crônicas			
Cardiovasculares	4	27	31
Câncer	3	11	14
Respiratórias/Asma	1	5	6
Diabetes	2	3	5
Outras	0	5	5
Total	10	51	61
Causas externas			
Homicídios	1	12	13
Suicídios	0	3	3
Ac. Trânsito	1	2	3
Todas as causas ext.	10	14	23
Total	12	31	42
Materno-infantil			
Perinatal/neonatal/infantil	5	23	28
Maternas/aborto/ fertilidade	4	8	12
Total	9	31	40
Sistemas de informação			
SIM	12	12	24
Sinasc	4	4	8
SIH	3	1	4
Total	19	17	36
Assistência à saúde			
Assist. pré-natal/ parto	0	7	7
Outras	2	17	19
Outros			
Desigualdade em saúde	4	22	26
Evitabilidade	1	5	6
Poluição	2	7	9
Doenças profissionais	1	2	3
Etnia e saúde	2	1	3
Demais	4	7	11
Total	14	47	61
Total	69	225	294

Fonte: Base de dados MEDLINE/PubMed.

estudos sobre causas de morbimortalidade (133). As doenças cardiovasculares (31), as causas externas de morte (23), os cânceres (14), os homicídios (13) e a Aids (10) foram as causas mais encontradas nesta categoria. A categoria outros respondeu pela segunda maior frequência de artigos (61), principalmente devido ao significativo número de textos que abordaram o tema das desigualdades em saúde (26). Em terceiro lugar apareceu a categoria referente à saúde materno-infantil, com 40 artigos, dos quais 70% (28) estudaram os períodos perinatal, neonatal e/ou infantil. Os artigos que realizaram avaliações dos SIS ocuparam a quarta colocação em número de publicações (36). Entre estes, 24 (66,6%) avaliaram a cobertura, completude e/ou confiabilidade dos dados. Finalmente, na categoria assistência à saúde, foram selecionados 26 textos que, em sete casos, abordaram questões relacionadas à assistência pré-natal e/ou ao parto. As faixas etárias da população estudada não foram especificadas em 129 artigos. Nos casos com grupos etários identificados, os textos eram dedicados ao período perinatal (7), ao período neonatal (20), aos menores de um ano (30), às crianças (14), aos adolescentes (13), aos adultos (37) e aos idosos (14).

A divisão da série histórica estudada em dois períodos, segundo a década de sua publicação, ou seja, de 1990 a 1999 e de 2000 a 2006, possibilitou a observação de algumas diferenças entre eles. O número médio do total de artigos brasileiros publicados passou de 830, em 1990-1999, para 2.292, no período seguinte, o que representou um aumento de 1,8 vez. Em relação aos textos selecionados para este estudo, o crescimento foi de 3,6 vezes, tendo em vista que a média do primeiro período correspondeu a 6,9 artigos e a do segundo foi de 32,1. Entre 2000 e 2006, observou-se concentração dos artigos (91,3%) que utilizaram simultaneamente mais de um dos SIS.

Houve aumento de mais de quatro vezes no número de artigos em inglês no período 2000-2006, se comparado ao anterior. A partir de 2000, foram recuperadas 13 publicações bilíngues português/inglês, uma em francês e duas em espanhol. Ressalte-se

a diversificação dos periódicos utilizados a partir de 2000, quando foram observados os primeiros artigos científicos com dados do SIM, Sinasc e/ou SIH em 36 revistas, sendo que algumas destas são dirigidas a clínicos, como os *Arquivos Brasileiros de Cardiologia* (16 textos) e o *Jornal de Pediatria* (3). A partir de 2000, também pela primeira vez, foram citadas 39 novas instituições, destacando-se a Secretaria de Estado da Saúde de São Paulo (12) e a Secretaria Municipal de Saúde de Belo Horizonte (5). A produção científica das Regiões Centro-Oeste e Norte também ganharam impulso a partir de 2000, com a publicação de sete artigos.

Em relação às faixas etárias estudadas, ressaltam-se os significativos aumentos observados, a partir de 2000, nas publicações dedicadas ao período perinatal (7), aos adolescentes (13) e à população idosa (14). O interesse pelas populações infantil e neonatal foi verificado por meio dos 12 artigos publicados entre 1990 e 1999 e dos 50 editados no período seguinte.

Alguns aspectos foram enfocados pela primeira vez em artigos publicados a partir de 2000, tais como as avaliações da assistência ao pré-natal e/ou ao parto (7 textos) e o suicídio (3). Houve aumento no número de estudos dedicados às doenças crônicas, especialmente as cardiovasculares (27) e às causas externas (31), embora as doenças transmissíveis (especialmente a Aids) também tenham permanecido em foco. Os SIS também foram utilizados em maior número de vezes em abordagens sobre desigualdade em saúde (22), evitabilidade (6) e poluição ambiental (7), no período 2000-2006 (Tabela 2).

Discussão

Os resultados indicam tendência de aumento na utilização das bases de dados secundárias do SIM, Sinasc e/ou SIH em pesquisas no período 1990-2006. Esse crescente interesse pode se relacionar à maior facilidade e agilidade de acesso às informações processadas em meio eletrônico e à qualidade dos dados, observada com as ações de descentralização da sua gestão. Estes resultados corroboram evidências

apresentadas em outros estudos sobre o crescimento da produção científica em epidemiologia (BARRETO, 2006), da utilização do SIM e do Sinasc em pesquisas brasileiras sobre mortalidade perinatal (FONSECA; COUTINHO, 2004) e do SIH na Saúde Coletiva (BITTENCOURT et al., 2006).

Coura e Wilcox (2003) observaram que a publicação de artigos de autores brasileiros em periódicos internacionais indexados pelo ISI quintuplicou no período 1981-2001 e que o número de pessoas qualificadas (doutores e mestres) triplicou na última década. A limitação da busca à base de dados MEDLINE/PubMed não possibilitou o acesso a importantes periódicos da área que também são muito consultados e que não estavam indexados nesta base bibliográfica, como a *Revista Brasileira de Epidemiologia* e os periódicos *Ciência & Saúde Coletiva* e *Epidemiologia e Serviços de Saúde*, bem como a *Revista Brasileira de Estudos de População*, no campo de demografia, que tradicionalmente utiliza dados secundários – cabe comentar que todas foram, em anos recentes, indexadas à base de dados SciELO, um caminho para sua indexação à MEDLINE/PubMed.

De fato, a MEDLINE/PubMed é a fonte bibliográfica mais consultada no campo da saúde (BARRETO, 2006) e nela estão indexadas os *Cadernos de Saúde Pública* e a *Revista de Saúde Pública*, nas quais se concentrou grande parte das publicações na área. Assim, os resultados aqui apresentados cumpriram o objetivo de evidenciar o crescimento dos estudos na área, embora a ampliação da busca em outras bases bibliográficas pudesse resultar na obtenção de maior magnitude desse aumento.

Mais de 70% da produção científica capturada na base de dados MEDLINE/PubMed por este estudo foi escrita em português, o que pode representar uma barreira para a leitura destes artigos por pesquisadores de outras nacionalidades. Textos publicados em idiomas diferentes do inglês são menos frequentemente lidos e citados, mesmo que seja possível o acesso ao artigo completo, via SciELO (CUNHA-MELO et al., 2006).

Após sua publicação, deseja-se que o artigo atinja o maior número de leitores e

que seu conteúdo promova e/ou participe do diálogo da comunidade científica nacional e internacional. Em busca dessas metas, os conselhos editoriais das revistas científicas trabalham intensamente para que seus periódicos tenham maior visibilidade internacional, por meio da indexação em bases bibliográficas de acesso internacional. Neste sentido, Blank e colaboradores (2006) avaliaram que a indexação do *Jornal de Pediatria* ao MEDLINE/PubMed propiciou um crescimento significativo das submissões de artigos brasileiros e estrangeiros a este periódico.

A tendência de aumento da pluralidade da língua (ou internacionalização) das publicações foi observada neste estudo por meio de duas importantes iniciativas. Uma delas foi verificada nos *Arquivos Brasileiros de Cardiologia* e no *Jornal de Pediatria*, que, ao disponibilizarem na Internet a opção da publicação bilingue (português-inglês), permitiram acesso aos artigos completos à audiência internacional e a todo público brasileiro. A outra iniciativa que também deve ser ressaltada refere-se ao aumento dos artigos em inglês publicados pela *Revista de Saúde Pública* e pelos *Cadernos de Saúde Pública*.

O maior emprego do SIM como fonte de pesquisa é coerente com a reconhecida importância das estatísticas de mortalidade, embora, a partir de 2000, não tenha sido observada diferença significativa na tendência de utilização desse sistema em relação aos demais SIS. Segundo avaliações recentes, o SIM apresentou ampliação da cobertura de óbitos (PAES, 2005) e redução do número de óbitos por causas maldefinidas (LAURENTI et al., 2005), embora ainda persistam muitas deficiências na qualidade da informação sobre mortalidade em nosso país (ALMEIDA et al., 2006).

Deve ser ressaltado o grande potencial demonstrado pela utilização conjunta de duas ou mais bases de dados, evidenciado neste estudo pelo aumento do número de artigos que cruzaram informações de diferentes origens, principalmente por meio de *linkages* ou de relacionamentos probabilísticos de dados.

Outro aspecto a se destacar é a relativa estabilidade no número absoluto de

artigos com a utilização do SIS, entre 2000 e 2006, com média de 32,1 textos por ano, com moda igual a mediana (31 artigos por ano). Observou-se, inclusive, queda na proporção de artigos (de 2,1%, em 1999, para 1,1% e 1,4%, em 2005 e 2006, respectivamente). Vários fatores podem ter atuado nesta tendência de “salto” seguido por estabilização, tais como: a revisão da CID 10 implementada em 1996, que pode ter impulsionado um número de trabalhos extras visando contrastar mudanças de categorizações, por exemplo (de fato, segundo Grassi e Laurenti (1998), a mera revisão pode alterar a tendência de uma doença, muitas vezes de forma brusca); aceleração das atualizações dos bancos de dados, bem como aumento da cobertura e qualidade das informações (MATHIAS et al., 2005); crescimento dos programas de ensino e pesquisa nas áreas de Demografia e Saúde Pública; e o fluxo de incorporação dos periódicos, no MEDLINE, com aumento da visibilidade das revistas eletrônicas (CASTRO, 2006).

Deve ser destacada ainda a crescente facilitação para utilização dos dados do SIS, adicionada às facilidades computacionais, disponíveis aos gestores, pesquisadores da área de saúde em geral, docentes e alunos de pós-graduação. Esta facilitação pode elevar o número e a proporção de artigos com uso do SIS em relação ao total de artigos brasileiros. A utilização dos correios eletrônicos, cabe ainda destacar, pode trazer grandes melhorias e maior rapidez na condução de pesquisas em colaboração, o que facilmente incrementaria a produção dos profissionais da área e multiplicaria sua produção acadêmica (CUENCA, 2004).

A concentração da produção científica em instituições dos estados da Região Sudeste, notadamente São Paulo e Rio de Janeiro, observada neste estudo, corroborou os resultados obtidos por pesquisas em outras áreas. Essa concentração foi ainda mais acentuada nas publicações com dados do SIH, conforme já relatado por Bittencourt e colaboradores (2006).

Em levantamento sobre a produção científica em saúde em seis países da América Latina, no período 1973-1992, Pellegrini

Filho e Goldbaum (1997) demonstraram predomínio de trabalhos produzidos por autores de uma única instituição (73,4%), mas também observaram tendência de aumento dos estudos em colaboração. Ao se limitar a considerar apenas a instituição de origem do primeiro autor, os resultados obtidos neste estudo podem ter subestimado o papel de colaboradores de instituições de pesquisa e/ou centros formadores de docentes e pesquisadores.

Verificou-se haver ampliação na variedade temática nos artigos analisados. Ao longo dos anos foram sendo abordados novos temas, como as doenças crônicas (as cardiovasculares, principalmente), a violência (em especial os homicídios), as desigualdades em saúde (inclusive as raciais) e as abordagens sob as óticas da evitabilidade, dos riscos ambientais e da saúde ocupacional. As faixas etárias extremas da vida (perinatal e idosos) e a adolescência também surgiram como novos temas.

Embora fuja ao escopo desse estudo, deve ser citada a riqueza de alternativas metodológicas adotadas para ampliação da compreensão sobre os problemas de saúde da população. Em especial, devem ser referidos os crescimentos das análises de tendências temporais de causas de morte, das abordagens ecológicas com análises espaciais, das análises baseadas em diferentes técnicas de relacionamento de dado e dos estudos caso-controle e de coorte.

Finalmente, deve-se ressaltar que o presente estudo, ao avaliar a tendência de utilização dos sistemas de informação em saúde em pesquisas científicas, visou prioritariamente valorizar e estimular o uso dessas bases de dados. Como resultado da progressiva descentralização da gestão, as análises de situação de saúde são realizadas pelos serviços de saúde em níveis cada vez mais desagregados, tendo como metas as intervenções específicas a serem adotadas pelos níveis locais.

A demanda pela aproximação entre os diagnósticos de situação e os níveis mais descentralizados da atenção criou a necessidade de construção de indicadores de saúde adequados para intervenções específi-

cas. Consequentemente, aumentou também a demanda pela qualificação das bases de dados, com ampliação da cobertura e maior qualidade no preenchimento das variáveis. Para atender a estas demandas, observa-se, nos três níveis de governo, que os setores responsáveis pela coleta, processamento e divulgação das informações das bases de dados nacionais têm realizado grandes

investimentos em qualificações das equipes técnicas e das suas infraestruturas.

Assim, aos ganhos mais facilmente mensuráveis que podem advir da utilização das bases de dados como fonte de pesquisa científica, soma-se mais um: o maior estímulo dado aos profissionais destas equipes técnicas resultante da ampliação do significado do produto de seu trabalho.

Referências

ALMEIDA, M.F.; ALENCAR, G.P.; NOVAES, H.M.D.; ORTIZ, L.P. Sistemas de informação e mortalidade perinatal: conceitos e condições de uso em estudos epidemiológicos. **Revista Brasileira de Epidemiologia**, São Paulo, 9(1): 56-68, 2006.

BARRETO, M.L. Crescimento e tendência da produção científica em epidemiologia no Brasil. **Revista de Saúde Pública**, São Paulo, 40, número especial, p. 79-85, 2006.

BITTENCOURT, S.A.; CAMACHO, L.A.B.; LEAL, M.C. O Sistema de Informação Hospitalar e sua aplicação na saúde coletiva. **Cadernos de Saúde Pública**, Rio de Janeiro, 22(1): 19-30, 2006.

BLANK, D.; ROSA, L.O.; GURGEL, R.Q.; GOLDANI, M.Z. Brazilian knowledge production in the field of child and adolescent health. **Jornal de Pediatria**, Rio de Janeiro, 82(2): 97-102, 2006.

CARDOSO, A.M.; SANTOS, R.V.; COIMBRA Jr., C.E.A. Mortalidade infantil segundo raça/cor no Brasil: o que dizem os sistemas nacionais de informação? **Cadernos de Saúde Pública**, Rio de Janeiro, 21(5): 1.602-1.608, 2005.

CASTRO, R.C.F. Impacto da Internet no fluxo da comunicação científica em saúde. **Revista de Saúde Pública**, São Paulo, 40, número especial, p. 57-63, 2006.

COURA, J.R.; WILCOX, L.C.B. Impact factor, scientific production and quality of brazilian medical journals. **Memórias do Instituto Oswaldo Cruz**, Rio de Janeiro, 98(3): 293-297, 2003.

CUENCA, A.M.B. **O uso da internet por docentes da área de saúde pública no**

Brasil. Tese (Doutorado em Saúde Pública). São Paulo, Faculdade de Saúde Pública da USP, 2004.

CUNHA-MELO, J.R.; SANTOS, G.C.; ANDRADE, M.V. Brazilian medical publications: citation patterns for Brazilian-edited and non-Brazilian literature. **Brazilian Journal of Medical and Biological Research**, Ribeirão Preto, 39(8): 997-1.002, 2006.

DRUMOND Jr., M. Epidemiologia em serviços de saúde. In: CAMPOS, G.W.S.; MINAYO, M.C.S.; AKERMAN, M.; DRUMOND Jr., M.; CARVALHO, Y.M. (Orgs.). **Tratado de saúde coletiva**. São Paulo: Hucitec; Rio de Janeiro: Fiocruz, 2006.

FONSECA, S.C.; COUTINHO, E.S.F. Pesquisa sobre mortalidade perinatal no Brasil: revisão da metodologia e resultados. **Cadernos de Saúde Pública**, Rio de Janeiro, 20: Sup 1: S7-S19, 2004.

GRASSI, P.R.; LAURENTI, R. Implicações da introdução da décima revisão da classificação internacional de doenças em análise de tendência da mortalidade por causas. **Informe Epidemiológico do SUS**, Brasília, 7 (3): 43-47, 1998.

GUIMARÃES, J.A. A pesquisa médica e biomédica no Brasil. Comparações com o desempenho científico brasileiro e mundial. **Ciência & Saúde Coletiva**, Rio de Janeiro, 9(2): 303-27, 2004.

HARAKI, C.A.C.; GOTLIEB, S.L.D.; LAURENTI, R. Confiabilidade do Sistema de Informações sobre Mortalidade em município do sul do Estado de São Paulo. **Revista Brasileira de Epidemiologia**, São Paulo, 8(1): 19-24, 2005.

LAURENTI, R.; MELLO-JORGE, M.H.P.; LEBRÃO, M.L.; GOTLIEB, S.L.D.; ALMEIDA, M.F. Estatísticas vitais: contando os nascimentos e as mortes. **Revista Brasileira de Epidemiologia**, São Paulo, 8(2): 108-110, 2005.

LEVY, S. Pense grande faça pequeno. **Revista Brasileira de Epidemiologia**, São Paulo, 10(1): 127-128, 2007.

MANSUR, J.A.; ABUD, A.S.; ALBUQUERQUE, A.P. Tendências das publicações na fase trimestral, bimensal e mensal de cinco décadas dos Arquivos Brasileiros de Cardiologia. **Arquivos Brasileiros de Cardiologia**, Rio de Janeiro, 75(1): 1-7, 2000.

MATHIAS, T.A.F.; MELLO-JORGE, M.H.P.; LAURENTI, R.; AIDAR, T. Considerações sobre a qualidade de informações de mortalidade na população idosa residente no Município de Maringá, Estado do Paraná, Brasil, no período de 1979 a 1998. **Epidemiologia e Serviços de Saúde**, Brasília, 14(3): 159-169, 2005.

MINISTÉRIO DA SAÚDE. Agência Nacional de Saúde Suplementar. **Caderno de Informação da Saúde Suplementar: beneficiários, operadoras e planos**. Disponível em: <http://www.ans.gov.br/portal/upload/informacoess/caderno_informaca_03_2007.pdf>. Acesso em: 02 jun. 2007.

PACKER, A.L.; TARDELLI, A.O.; CASTRO, R.C.F. A distribuição de conhecimento científico público em informação, comunicação e informática em saúde indexado nas bases de dados MEDLINE/PubMed e LILACS. **Ciência & Saúde Coletiva**, Rio de Janeiro, 12(3): 587-599, 2007.

PAES, N.A. Avaliação da cobertura dos registros de óbitos dos estados brasileiros

em 2000. **Revista de Saúde Pública**, São Paulo, 39(6): 882-90, 2005.

_____. Qualidade das estatísticas de óbitos por causas desconhecidas dos Estados brasileiros. **Revista de Saúde Pública**, São Paulo, 41(3): 436-445, 2007.

PAIM, J.S. Atenção à saúde no Brasil. In: PAIM, J.S. **Desafios da saúde coletiva no século XXI**. Salvador, EDUFBA, 2006.

PELLEGRINI FILHO, A.; GOLDBAUM, M.; SILVI, J. Production of scientific articles on health in six Latin American countries, 1973-1992. **Revista Panamericana de Salud Publica**, United States, 2:121-9, 1997.

RODRIGUES, O.S.; FONSECA, L.; CHAIMOWICZ, H. Mapping cancer, cardiovascular and malaria research in Brazil. **Brazilian Journal of Medical and Biological Research**, Ribeirão Preto, 33:853-867, 2000.

ROMERO, D.E.; CUNHA, C.B. Avaliação da qualidade das variáveis sócio-econômicas e demográficas dos óbitos de menores de um ano registrados no Sistema de Informações sobre Mortalidade (1996/2001). **Cadernos de Saúde Pública**, Rio de Janeiro, 22(3): 673-84, 2006.

SILVA, A.A.M.; RIBEIRO, V.S.; BORBA Jr., A.F.; COIMBRA, L.C.; SILVA, R.A. Avaliação da qualidade dos dados do Sistema de Informações sobre Nascidos Vivos em 1997-1998. **Revista de Saúde Pública**, São Paulo, 35(6): 508-14, 2001.

THEME FILHA, M.M.; GAMA, S.G.N.; CUNHA, C.B.; LEAL, M.C. Confiabilidade do Sistema de Informações sobre Nascidos Vivos Hospitalares no Município do Rio de Janeiro, 1999-2001. **Cadernos de Saúde Pública**, Rio de Janeiro, 20 (supl.): S83-S91, 2004.

Resumen

Utilización de datos secundarios del SIM, Sinasc y SIH en la producción científica brasileña de 1990 a 2006

Se trata de un estudio descriptivo de la distribución y de algunas características de los artículos científicos publicados de 1990 a 2006, que utilizaron datos secundarios de los Sistemas de

Información sobre Mortalidad (SIM), Nacidos Vivos (Sinasc) y/o Informaciones Hospitalarias (SIH). Después de una investigación bibliográfica en el banco de datos MEDLINE/PubMed, con base en el MeSH (Medical Subject Headings), se seleccionaron 294 artículos. Fueron analizadas variables referentes a los artículos, autores y periódicos. Se observó un aumento cuádruple en el volumen de publicaciones y en las ediciones en inglés. Hubo un crecimiento en el número de periódicos utilizados, en las instituciones de origen de los investigadores, en la utilización conjunta de dos bases de datos y en las investigaciones dirigidas a las enfermedades y agravamientos no-transmisibles (especialmente las cardiovasculares y causas externas). Los resultados señalan una expresiva ampliación de la producción científica brasileña que utiliza datos del SIM, Sinasc y/o SIH.

Palabras-clave: Bibliometría. Estadísticas vitales. Investigación / tendencias. Indicadores de producción científica. Mortalidad.

Abstract

The use of secondary data from SIM, Sinasc, and SIH in Brazilian scientific production between 1990 and 2006

The present study describes the distribution and several of the characteristics of scientific articles published between 1990 and 2006 that describe secondary data from the Brazilian Information Systems on Mortality (SIM), from the Information System on Live Births (SINASC), and the Hospital Information System (SIH). A bibliographic search on the MEDLINE/PubMed database, using MeSH (Medical Subject Headings), called up 294 articles. We analyzed variables referring to the articles, the authors, and the periodicals in which the articles were published. The number of publications and editions in English increased fourfold during this period. There was also an increase in the number of journals studied at the researchers' host institutions, in the simultaneous use of databases and in research on non-transmittable diseases (especially from cardiovascular and external causes). The results suggest a considerable increase in Brazilian scientific production using data from SIM, SINASC, and/or SIH.

Keywords: Journal articles. Publications. Research. Trends. Information systems.

Recebido para publicação em 20/04/2008.

Aceito para publicação em 05/03/2009.