

Georreferenciamento de dados de saúde na escala submunicipal: algumas experiências no Brasil

Geocoding Health Data in Sub-Municipal Scale: Some Brazilian Experiences

Christovam Barcellos

Departamento de Informações em Saúde, Centro de Informação Científica e Tecnológica, Instituto Oswaldo Cruz, Ministério da Saúde, Rio de Janeiro-RJ, Brasil

Walter Massa Ramalho

Departamento de Análise de Situação de Saúde, Secretaria de Vigilância em Saúde, Ministério da Saúde, Brasília-DF, Brasil

Renata Gracie

Departamento de Informações em Saúde, Centro de Informação Científica e Tecnológica, Instituto Oswaldo Cruz, Ministério da Saúde, Rio de Janeiro-RJ, Brasil

Mônica de Avelar F. M. Magalhães

Departamento de Informações em Saúde, Centro de Informação Científica e Tecnológica, Instituto Oswaldo Cruz, Ministério da Saúde, Rio de Janeiro-RJ, Brasil

Márcia Pereira Fontes

Gerência de Vigilância Epidemiológica, Secretaria Municipal de Saúde do Rio de Janeiro-RJ, Brasil

Daniel Skaba

Coordenação de Estruturas Territoriais, Diretoria de Geociências, Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro-RJ, Brasil

Resumo

Neste trabalho, são relatadas experiências de georreferenciamento de dados de saúde em alguns Municípios brasileiros, analisadas segundo a disponibilidade e atualização de bases cartográficas e o tratamento de dados de endereço nos sistemas de informações em saúde (SIS). A diversidade de estratégias de georreferenciamento de dados no Brasil é resultado das condições particulares de desenvolvimento desses projetos nos Municípios, o que influi na eficiência e precisão da localização dos eventos de saúde. São sugeridas estratégias para captação e armazenamento de dados de endereço nos SIS e sua compatibilização com os cadastros de logradouros existentes, bem como o desenvolvimento de programas e aplicativos que permitam buscas e relacionamentos entre essas duas bases de dados.

Palavras-chave: georreferenciamento; sistemas de informações em saúde; cartografia; geoprocessamento.

Summary

This work describes experiences in geocoding health data in Brazilian municipalities, analyzed according to the availability and update of cartographic databases and the treatment given to addresses data in health information systems (SIS). The diversity of current geocoding strategies in Brazil results of local conditions on which these projects are developed, presenting variable efficiency and accuracy for health event location. The authors suggest strategies to capture and store addresses data in SIS, compatible with the existent street registries, as well as the development of programs and scripts to search and link these two databases.

Key words: geocoding; health information systems; cartography; geoprocessing.

Endereço para correspondência:

Ministério da Saúde, Instituto Oswaldo Cruz, Centro de Informação Científica e Tecnológica, Departamento de Informações em Saúde, Av. Brasil, 4365, Manguinhos, Rio de Janeiro-RJ, Brasil. CEP: 21045-900
E-mail: xris@cict.fiocruz.br

Introdução

O geoprocessamento é definido como um conjunto de tecnologias voltadas para a coleta e tratamento de informações espaciais com determinado objetivo, executadas por sistemas específicos para cada aplicação. Nas últimas décadas, esses sistemas têm sido empregados para avaliação ambiental, planejamento urbano, meteorologia, entre outros campos de aplicação. Ao contrário do que acontece nessas áreas, na Saúde, os dados não são obtidos por meios remotos. Dados sobre as condições de saúde das pessoas devem ser adquiridos mediante inquéritos e censos demográficos, ou pelos sistemas de vigilância.¹ São dados de atributos da pessoa e um grande esforço tecnológico e metodológico tem se desenvolvido para captá-los e tratá-los como característica de território. Essa transformação resulta em uma abstração e simplificação de processos sociais e ambientais presentes na determinação de doenças.² Assim, as bases cartográficas digitais, que são, muitas vezes, o produto final de projetos de geoprocessamento de outros setores, constituem apenas o ponto de partida das análises espaciais de saúde. Para que sejam utilizadas como meio de análise, as bases de dados de saúde devem ser georreferenciadas,^{3,4} integradas a dados ambientais e socioeconômicos,^{4,6} e submetidas a procedimentos de avaliação de sua distribuição espacial.^{7,8} O georreferenciamento de um dado com endereço é definido como o processo de associação desse dado a um mapa e pode ser efetuado de três formas básicas: associação a um ponto, a uma linha ou a uma área. O resultado desse processo é a criação de elementos gráficos que podem ser usados para a análise espacial.

Na área da Saúde, os sistemas de informações geográficas (SIG) têm se tornado ferramentas de grande utilidade. Sua capacidade de integrar diversas operações, como captura, armazenamento, manipulação, seleção e busca de informação, análise e apresentação de dados, auxilia o processo de entendimento da ocorrência de eventos, predição, tendência, simulação de situações, planejamento e definição de estratégias no campo da Vigilância em Saúde. A incorporação dos SIG pela Saúde tem história relativamente recente e ainda depende de um conjunto de bases tecnológicas e metodológicas em fase de implementação.

Os últimos anos vêm sendo marcados pela crescente disponibilidade e facilidade de acesso e análise de

dados mediante sistemas computacionais simples. No Brasil, o setor Saúde é detentor de um extenso banco de dados, que abrange dados vitais, de morbidade, gerenciais e contábeis. Esses dados vêm sendo armazenados em diversos sistemas de informações: Sistema de Informações sobre Mortalidade (SIM), Sistema de Informações sobre Nascidos Vivos (Sinasc), Sistema de Informação de Agravos de Notificação (Sinan), Sistema de Informações Hospitalares do Sistema Único de Saúde (SIH/SUS), Sistema de Informações Ambulatoriais do Sistema Único de Saúde (SIA/SUS), entre outros. Seguindo a lógica do Sistema Único de Saúde (SUS), de um sistema de cobertura nacional e arquitetura hierárquica, os dados desses sistemas de informações são gerados no nível local e repassados, no sentido ascendente, às demais esferas de governo. É de responsabilidade da gestão municipal, portanto, a captação dos dados e o correto preenchimento dos campos dos formulários desses sistemas de informações, inclusive daqueles reservados à localização geográfica, como nome e código de logradouro e bairro de residência.

Para que sejam utilizadas como meio de análise, as bases de dados de saúde devem ser georreferenciadas, integradas a dados ambientais e socioeconômicos, e submetidas a procedimentos de avaliação de sua distribuição espacial.

No Brasil, o cenário atual da aplicação dos SIG em saúde é extremamente favorável e pode ser resumido segundo quatro eixos de desenvolvimento: disponibilidade de bases de dados; aperfeiçoamento de programas computacionais; desenvolvimento tecnológico; e capacitação de pessoal. Esses eixos são inter-relacionados e cada solução tecnológica pode ter reflexos sobre os programas e exigir um redirecionamento das iniciativas de capacitação. As soluções para a democratização desse conjunto de ferramentas são, portanto, integradas e exigem a coordenação de esforços entre a Saúde e outros setores.⁹

Para que os dados gerados pelos sistemas de informações em saúde sejam mapeados, os eventos

de saúde devem ser relacionados a um conjunto de objetos geográficos ou unidades espaciais previamente construídas, como bairros, setores censitários, lotes ou trechos de logradouros. Assim, um dos primeiros passos para o georreferenciamento desses dados é o reconhecimento do estágio atual da cartografia urbana existente nas cidades. E os sistemas de informações em saúde, por sua vez, devem coletar e armazenar dados de endereço compatíveis com essa estrutura de dados cartográficos. Ao longo do processo de georreferenciamento, diversas decisões são tomadas, tais como a escolha de uma unidade espacial de referência, a solução de alguma incoerência ou complementação de endereço incompleto ou, ainda, a aproximação de sua numeração.¹⁰ Essas decisões afetam a disposição final dos eventos sobre a base cartográfica e, por conseguinte, os possíveis resultados da análise espacial desses eventos.

No presente trabalho, são discutidos os problemas enfrentados no georreferenciamento de dados de saúde a partir da análise das principais experiências desenvolvidas em Municípios do Brasil. São sugeridas estratégias para a captação e tratamento de dados de endereço no país – com suas principais alternativas de padronização – e para o georreferenciamento de dados.

Metodologia

Este trabalho é um relatório conclusivo da “Oficina sobre georreferenciamento da base de dados da nova versão do Sistema Nacional de Informação de Agravos de Notificação (Sinan XP)”, realizada em Brasília, no mês de junho de 2003, quando foram discutidas as experiências de algumas cidades, as quais, reunidas, representavam a grande diversidade de estruturas urbanas possíveis entre os Municípios brasileiros. Também foram relatados os avanços do geoprocessamento em saúde no Brasil, recursos de publicação de dados gráficos pela Internet e aplicações do geoprocessamento de dados de saúde, principalmente para vigilância em saúde. As discussões e recomendações deste documento não se limitam, portanto, ao Sinan; elas podem refletir situações semelhantes encontradas nos diversos sistemas de informações em saúde.

Após levantamento, por meio de contatos institucionais e da busca via ferramentas da Internet, foram convidados e participaram da Oficina representantes

das Secretarias Municipais de Saúde (SMS) de alguns dos Municípios brasileiros de maior tradição no uso do geoprocessamento em saúde: Porto Alegre; Rio de Janeiro; Belo Horizonte; Goiânia; e Salvador.

Foram identificados outros Municípios com experiência nessa tecnologia, como Manaus-AM, Vitória da Conquista-BA, Brasília-DF, Campo Grande-MS, Montes Claros-MG, Betim-MG, Ipatinga-MG, Belém-PA, Campina Grande-PB, Curitiba-PR, Londrina-PR, Paranaguá-PR, Maringá-PR, Natal-RN, Dom Pedrito-RS, Santana de Parnaíba-SP, Recife-PE, Olinda-PE, Campinas-SP e Suzano-SP, entre outros, os quais não participaram da Oficina de Brasília-DF. Essa lista não é conclusiva; certamente, diversos outros Municípios utilizam ferramentas de geoprocessamento em saúde, embora não tenham seus projetos divulgados.

Os representantes das SMS e órgãos do Ministério da Saúde apresentaram levantamentos sobre as condições em que são realizados os trabalhos de georreferenciamento de dados de saúde, que incluem a padronização de dados de endereço e sua busca em cadastros municipais. Durante a Oficina, o encaminhamento das discussões de aperfeiçoamento desses projetos desenvolveu-se sobre três eixos: dados de endereço dos SIS; cadastros de endereços (base cartográfica digital e cadastro com dados alfanuméricos); e programas de georreferenciamento.

Resultados

A diversidade de estratégias de georreferenciamento apresentadas demonstra as condições particulares de desenvolvimento desses projetos nos Municípios do Brasil. Várias são as estratégias de microlocalização de dados de saúde adotadas por esses Municípios, nos últimos anos (Tabela 1). A Prefeitura Municipal de Porto Alegre-RS, por exemplo, desenvolveu, por meio de um SIG, um sistema de localização de eventos de saúde por trechos de logradouros (a partir do endereço), para registros do SIM e do Sinasc. Estratégias semelhantes estão sendo adotadas no Município do Rio de Janeiro-RJ (SMS), em Curitiba-PR e cidades de médio porte dos Estados de São Paulo e do Paraná. Em Goiânia-GO e Belo Horizonte-MG, os diversos sistemas (SIM, Sinasc, Sinan, SIH/SUS, etc.) também são estruturados para o referenciamento do lote de endereço (nome do logradouro e número de porta). A representação geográfica desses eventos a partir

Tabela 1 - Principais recursos e características dos sistemas de geoprocessamento de dados de saúde em algumas capitais do Brasil^a

Cidade	Instituição	Ano de implantação	Programa de georreferenciamento	Cadastro de endereços	Eficiência
Porto Alegre-RS	SMS e Procempa	1994	SIM e Sinasc modificados	CTM	60 a 90%
Goiânia-GO	SMS e Comdata	1997	–	Cartão SUS	90% ^c
Rio de Janeiro-RJ	SMS e SMU/IPP	2001	EpiGIS ^b	CTM (SMF/IPTU)	60 a 90%
Belo Horizonte-MG	SMS e Prodabel	1992	Os próprios sistemas de entrada de dados: SIM, Sinasc, Sinan, SIH/SUS, CadSUS ou Sisgeo ^b	CTM	90 a 94% (considerando aproximação)
Região Metropolitana do Rio de Janeiro-RJ ^a	Fiocruz/MS	1995	Georreferenciador ^b	Censos de 2000 e de 1991 (IBGE)	50 a 90%
Salvador-BA	UFBa e Conder	2000	Localiza ^b	Censos de 2000 e de 1991 (IBGE)	40 a 80%

a) Elaborado durante a "Oficina de Georreferenciamento da Base de Dados da Nova Versão do Sistema de Informação de Agravos de Notificação (Sinan XP)", realizada em Brasília-DF, em junho de 2003.

b) Possui um programa específico para busca de endereços e georreferenciamento de dados de saúde

c) Depende da cobertura do cadastro do Cartão SUS

Nota de tradução de siglas presentes na tabela:

CadSUS: aplicativo de cadastro e manutenção de usuários do Sistema Único de Saúde e seus domicílios de residência

Comdata: Companhia de Processamento de Dados do Município de Goiânia

Conder: Companhia de Desenvolvimento Urbano do Estado da Bahia

CTM: cadastro técnico municipal

EpiGIS: sistema de integração de bases de dados para controle do dengue por regiões do Município do Rio de Janeiro

Fiocruz/MS: Instituto Oswaldo Cruz, do Ministério da Saúde

IBGE: Instituto Brasileiro de Geografia e Estatística

Procempa: Companhia de Processamento de Dados do Município de Porto Alegre

Prodabel: Empresa de Informática e Informação do Município de Belo Horizonte

SIH/SUS: Sistema de Informações Hospitalares do Sistema Único de Saúde

SIM: Sistema de Informações sobre Mortalidade

SMF/IPTU: Secretaria Municipal de Fazenda, Imposto Predial e Territorial Urbano, da Prefeitura da Cidade do Rio de Janeiro

SMS: Secretaria Municipal de Saúde

SMU/IPP: Secretaria Municipal de Urbanismo e Instituto Municipal de Urbanismo Pereira Passos, da Prefeitura da Cidade do Rio de Janeiro

Sinan: Sistema de Informação de Agravos de Notificação

Sinasc: Sistema de Informações sobre Nascidos Vivos

Sisgeo: sistema de indicadores sociais georeferenciados

UFBa: Universidade Federal da Bahia

do endereço permite uma variedade de agregações: setor censitário; bairro; região administrativa; áreas programáticas; entre outras possibilidades, facilmente construídas a partir dos recursos de análise espacial oferecidos pelos SIG.

Em diversos Municípios não presentes na Oficina, como Recife-PE, Olinda-PE e Londrina-PR, o bairro vem sendo utilizado como unidade espacial mínima

de agregação de dados. Essa é uma solução de georreferenciamento mais simples, apesar de exigir alguns procedimentos de padronização de entrada de dados. Nesses Municípios, o campo correspondente foi implantado nos principais sistemas de informações e seu preenchimento vem sendo padronizado para impedir incongruências entre a base cartográfica (mapa) e a base de dados não gráficos (tabelas). Diversos Mu-

nicípios do Brasil, como São Paulo-SP e Goiânia-GO, não possuem bairros oficializados, o que dificulta o traçado de seus limites oficiais em mapas, bem como a captação desse dado pelos sistemas de informações em saúde. Em Goiânia-GO, a associação do número do Cartão SUS ao lote georreferenciado tem permitido o georreferenciamento dos dados de saúde.

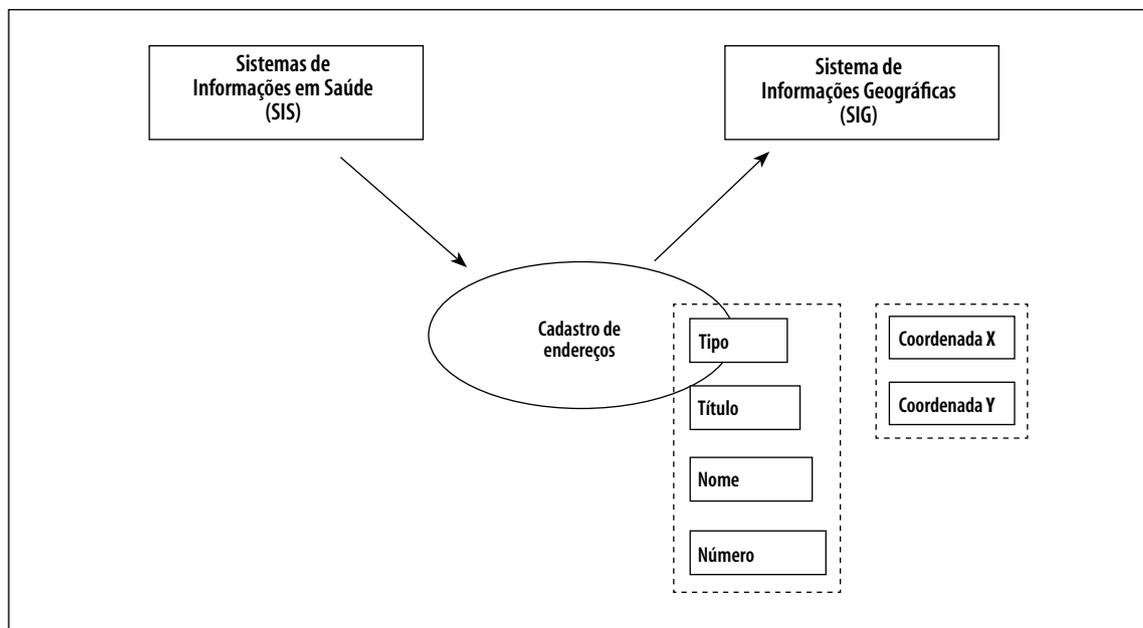
No Rio de Janeiro-RJ [Instituto Oswaldo Cruz (Fiocruz/MS), do Ministério da Saúde] e em Salvador-BA [Instituto de Saúde Coletiva da Universidade Federal da Bahia (ISC/UFBa)], projetos de pesquisa vêm desenvolvendo meios de georreferenciamento de dados de saúde, por meio da identificação dos setores censitários que corresponde ao endereço de residência registrado nos SIS.

Os programas de georreferenciamento de dados utilizados pelos Municípios representados na Oficina comparam bases de dados originadas dos SIS com um cadastro de endereços padronizado, existente no Município. A Figura 1 mostra o relacionamento entre os sistemas de informações em saúde e o SIG, com base em um cadastro de endereços. Entre os cadastros utilizados, pode-se citar o cadastro técnico municipal, usado principalmente para a arrecadação de impostos territoriais, o cadastro de usuários do SUS (Cartão

SUS) e o cadastro de segmentos de logradouros, este último construído pelo Instituto Brasileiro de Geografia e Estatística (IBGE) para a organização do Censo Demográfico de 2000.

Esse cadastro de endereços, por sua vez, pode estar relacionado a uma tabela que contém os pares de coordenadas do objeto geográfico, de modo a se obter uma localização (exata ou aproximada) do evento de saúde. Por exemplo, se um registro de saúde for relacionado a um cadastro de lotes e se a prefeitura dispuser de uma base cartográfica de lotes da cidade, pode-se referenciar esse evento de saúde ao lote correspondente. Igualmente, os setores censitários ou trechos de rua podem ser utilizados para o georreferenciamento desses dados. Dessa maneira, um programa de georreferenciamento deve incluir, como campo-chave do registro de saúde, a identificação de um objeto geográfico: trecho de rua, setor censitário, lote, etc. O mesmo programa pode, ainda, inserir campos de coordenadas no registro de saúde para que possam ser apresentados em um mapa. Atualmente, é bastante simples calcular as coordenadas dos diversos objetos geográficos usadas em um SIG.

Cabe ressaltar que o setor censitário tem sido utilizado como unidade de análise em vários estudos



a) Elaborado durante a "Oficina de Georreferenciamento da Base de Dados da Nova Versão do Sistema de Informação de Agravos de Notificação (Sinan XP)", realizada em Brasília-DF, em junho de 2003.

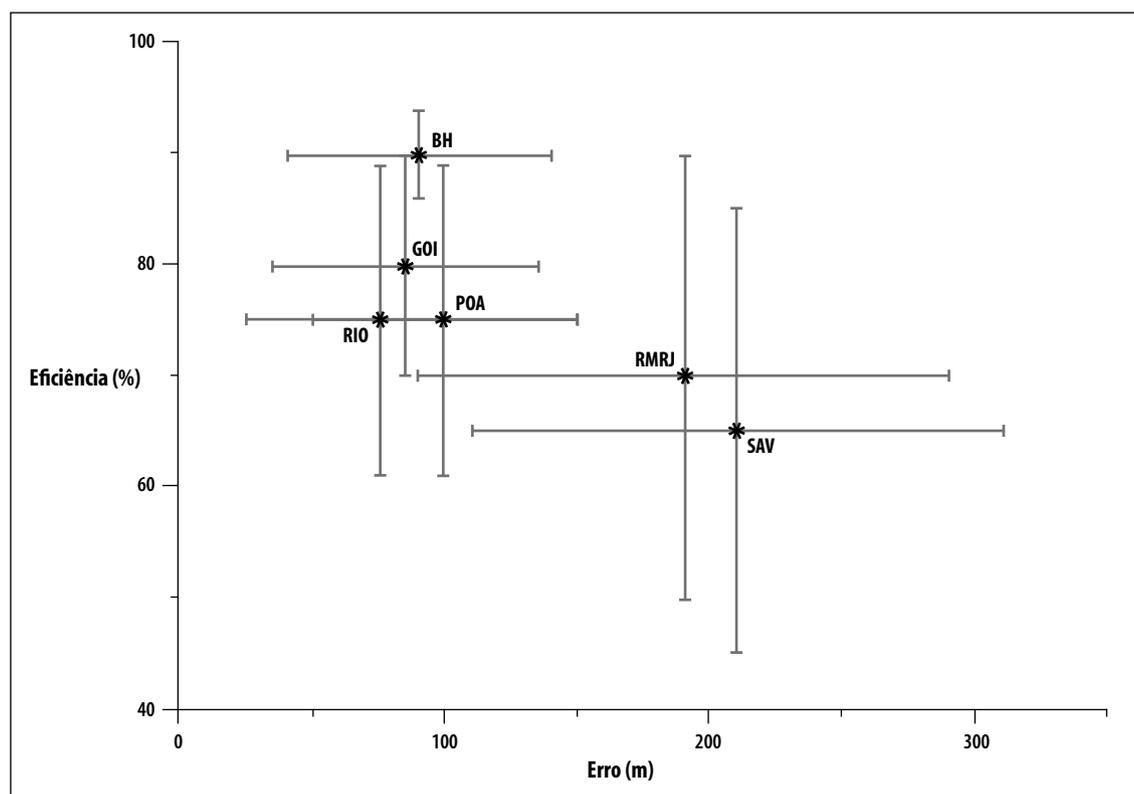
Figura 1 - Esquema de georreferenciamento de dados de saúde ^a

de distribuição de eventos de saúde.¹¹ Essa escolha decorre da possibilidade de utilizar as informações socioeconômicas resultantes dos questionários do censo. É mister ressaltar, contudo, que a delimitação dos setores tem objetivo operacional e pode variar grandemente, a cada censo demográfico realizado.¹²

Como nem todo evento de saúde contido em uma base de dados pode ser efetivamente localizado, as diferentes estratégias de georreferenciamento adotadas podem resultar em graus de eficiência variáveis, qual seja, maiores ou menores perdas de registros em decorrência da capacidade do sistema para localizar determinado endereço em uma base cartográfica digital do Município. Obviamente, essa eficiência depende da qualidade dos dados de endereço captados nos sistemas de informações em saúde, da cobertura e atualização do cadastro de endereços utilizado e

da capacidade do próprio programa em relacionar essas bases de dados, levando em consideração possíveis erros e formas de grafia de endereços. As eficiências (entendidas como proporção de eventos localizados) dos programas nas cidades são mostradas na Tabela 1.

A Figura 2 mostra o Município de Belo Horizonte-MG como aquele que apresenta maior eficiência no processo de georreferenciamento de dados de saúde. Há vários anos, a capital mineira vem investindo na construção de sistemas de informações geográficas, estruturando-os em mapas digitais do Município, em escala cadastral (entre 1:2000 e 1:10000), além de dispor de boa qualidade de dados nos sistemas de informações de saúde. O Município de Goiânia-GO, que apresenta boa eficiência em georreferenciamento, conta com uma cidade planejada, de endereços pa-



a) Elaborado durante a "Oficina de Georreferenciamento da Base de Dados da Nova Versão do Sistema de Informação de Agravos de Notificação (Sinan XP)", realizada em Brasília-DF, em junho de 2003, de acordo com a experiência dos Municípios de: Belo Horizonte – BH –, segundo a Secretaria Municipal de Saúde (SMS); Rio de Janeiro – RIO –, segundo a SMS; Goiânia – GOI –, segundo a SMS; Porto Alegre – POA –, segundo a SMS; Região Metropolitana do Rio de Janeiro – RMRJ –, segundo o Laboratório de Geoprocessamento do Instituto Oswaldo Cruz; e Salvador – SAV –, segundo o Instituto de Saúde Coletiva da Universidade Federal da Bahia.

Figura 2 - Relação entre eficiência e erro dos processos de georreferenciamento de dados segundo algumas experiências de Municípios do Brasil^a

dronizados. O georreferenciamento realizado no Rio de Janeiro-RJ (SMS), com base no mapa digital de logradouros da base cartográfica do Município – criada e mantida pela Secretaria Municipal de Urbanismo e o Instituto Municipal de Urbanismo Pereira Passos –, ocupa o terceiro lugar em eficiência de projetos. Porto Alegre-RS apresenta resultados semelhantes.

Os dois últimos programas, da Região Metropolitana do Rio de Janeiro-RJ e do Município de Salvador-BA, utilizam a mesma base de comparação de endereços, o Cadastro de Logradouros do IBGE (referente a 1991, para o Município do Rio de Janeiro; e referente ao ano 2000, para a Região Metropolitana do Rio de Janeiro-RJ e para o Município de Salvador-BA). Essa base apresenta alguns problemas, como a falta de preenchimento de numeração em alguns logradouros, resultando em alto número de registros não encontrados.

A precisão do processo de georreferenciamento, por sua vez, é medida pela capacidade de se localizar corretamente um evento de saúde, o mais próximo possível do local de sua ocorrência. Essa precisão depende da escala utilizada para o georreferenciamento: ela é tanto maior quanto maior é o detalhamento do objeto geográfico utilizado como referência. Um lote (escala maior) tem dimensões menores que um bairro (escala menor). Por isso, o georreferenciamento de eventos tendo como referência lotes tem maior precisão, comparativamente àquele que trabalha com bairros.¹³

Vários Municípios do Brasil contam com cadastros de endereço para a cobrança de impostos territoriais urbanos. Obviamente, essa base de dados é parcial e reflete, tão-somente, os lotes ou logradouros legalizados. Nesse caso, há uma sensível perda de dados relativos às áreas ocupadas por populações de baixa renda (classificadas como favelas) e rurais. Em algumas localidades, esse cadastro vem sendo atualizado e complementado por diversos setores, inclusive o da Saúde. O Rio de Janeiro-RJ está a viabilizar recursos de *software* e tecnológicos para que sua base cartográfica digital passe a dispor dos endereços das áreas informais do Município. Atualmente, o endereçamento é incipiente nessas regiões, seja por falta de dados de endereçamento nos sistemas de informações em saúde, seja por deficiência na atualização da base cartográfica digital. A eficiência para eventos ocorridos na área formal é de cerca de 95%.

Algumas cidades produziram, durante a fase de cadastramento do Cartão SUS, uma base de dados

de endereços relacionada aos usuários do Sistema Único de Saúde (mediante o número do cartão). Infelizmente, esse cadastro não tem sido realizado de forma padronizada pela maioria dos Municípios. Em diversos deles, o cadastramento é espontâneo e não tem base territorial.

Uma terceira possibilidade é o uso da base territorial do censo (Cadastro de Logradouros), produzida pelo IBGE com a regularidade de dez anos, atualizada a cada cinco anos. Essa base contém a lista de logradouros das cidades com mais de 25.000 habitantes, separados – para cada setor censitário – nos seguintes campos: nome; tipo; título; início da numeração ímpar; início da numeração par; fim da numeração ímpar; fim da numeração par; e código de endereçamento postal (CEP), definido pela Empresa Brasileira de Correios e Telégrafos (EBCT). Em oficina organizada pela Rede Interagencial de Informações para a Saúde (Ripsa), avaliou-se a qualidade dessa base e sua adequação para o georreferenciamento de dados de saúde em Goiânia-GO. Assim, espera-se o aprimoramento da base territorial com vistas a uma maior eficiência no próximo censo. Para tanto, é necessária a participação do IBGE nos projetos de localização de eventos de saúde dentro de Municípios.

Outros cadastros possíveis de serem usados são gerados por concessionárias de serviços de saneamento, energia elétrica e telefone. Esses cadastros cobrem somente as áreas servidas da cidade e apresentam grande atualização. Por exemplo, na maioria dos Municípios, cerca de 90% da população urbana faz parte do cadastro das companhias de energia elétrica.

Discussão

Em função dos problemas apontados neste levantamento, são sugeridas modificações e a introdução de procedimentos para o aperfeiçoamento de projetos de georreferenciamento segundo três linhas de atuação: captação e tratamento de dados de endereço pelos sistemas de informações em saúde; padronização de dados de endereço e uso de cadastros de endereços; e desenvolvimento de estratégias e programas de georreferenciamento.

Recomenda-se que sejam desenvolvidas formas de entrada padronizada de dados de endereço. Essa padronização pode se realizar durante a digitação ou após, em uma fase de análise crítica e consolidação

dos dados. Um endereço tipicamente urbano contém subcampos passíveis de serem separados no momento da digitação, como tipo (rua, avenida, beco, praça, etc.), título (senador, padre, dona, santo, etc.), preposição (de, dos, das, etc.), nome e número; ou agregados em um único campo, como “RUA PADRE JOSÉ”. Segundo o IBGE, existem 107 tipos de logradouros e 256 títulos. Essas listas podem ser expandidas e fornecidas aos responsáveis pelo desenvolvimento dos programas. Os Municípios do Rio de Janeiro-RJ (Fiocruz/MS e SMS), Belo Horizonte-MG, Porto Alegre-RS e Salvador-BA (ISC/UFBa) têm criado listas próprias de tipos e títulos, colocadas à disposição dos responsáveis pelo desenvolvimento do programa. A principal vantagem da separação do campo de endereço em subcampos está na possibilidade de combinação entre estes, na busca das diferentes formas de grafia de um endereço. Outra vantagem da separação do campo do endereço aparece no momento de preencher o formulário. O treinamento para esse preenchimento é mais efetivo quando se apresentam formas padronizadas: o conhecimento é rapidamente absorvido e posto em prática, gerando um maior número de endereços localizados.

A implantação de um sistema efetivo de geoprocessamento depende da incorporação de listas de endereços, da padronização e melhoria do preenchimento dos respectivos campos, da atualização da base cartográfica digital – logradouros – e da difusão de uma cultura de geoprocessamento de informações em saúde por todos os níveis do SUS.

Os SIS podem utilizar mais de um cadastro de endereços. Cada cidade que disponha de um cadastro mais atualizado e completo de endereços poderá usá-lo na padronização e busca de dados dos SIS. Em Goiânia-GO, por exemplo, o cadastro do Cartão SUS é a base de endereços mais adequada para localização de dados de saúde. Outros Municípios usam a base de endereços do cadastro técnico municipal ou do Censo Demográfico de 2000.

Os SIS devem captar outros campos complementares de endereço que possam ser aplicados no aperfeiçoamento das buscas, como ‘bairro’, ‘localidade’ e ‘distrito’ ou ‘área de abrangência’, além do código do Município, que já vem sendo preenchido adequadamente nos sistemas. Os campos ‘bairro’ e ‘distrito’ já existem em alguns sistemas, como no Sinan. Se o mesmo nome de logradouro for encontrado em diferentes áreas da cidade, os dados desses campos podem servir como critério de diferenciação de endereços.

A melhoria dos dados de endereço depende do uso dessa informação pelo próprio Município. Com a intensificação de ações de saúde baseadas na localização de eventos, pode-se incentivar o preenchimento correto dos campos de endereço. Por isso, é importante promover um retorno das informações geradas no nível local. Essa experiência tem alcançado êxito em Municípios como Goiânia-GO, Belo Horizonte-MG e Porto Alegre-RS.

A implantação desse tipo de sistema depende da incorporação de listas de endereços, da padronização e melhoria do preenchimento dos respectivos campos, da atualização da base cartográfica digital – no que se refere aos logradouros – e da difusão de uma cultura de geoprocessamento de informações em saúde por todos os níveis do SUS. A consecução deste último item, talvez o mais desafiador, deve ser lograda com o retorno de análises e diagnósticos espaciais para os geradores de dados, de modo a auxiliar o nível local no planejamento de ações de saúde. A utilização permanente dos bancos de dados para espacialização geográfica de eventos deve levar ao aperfeiçoamento de sua qualidade. Em vários países da Europa e nos Estados Unidos da América, os investimentos, tanto para localização de eventos quanto para estruturação de bases cartográficas, provêm de diversos setores de governo, que compartilham os custos de implantação do sistema e ajudam em seu aperfeiçoamento.¹²

A dificuldade em localizar dados referentes às áreas ocupadas por populações de baixa renda (favelas) abre uma oportunidade à criação de listas dessas áreas, comunidades, vilas e sub-bairros, com o objetivo de complementar os cadastros de logradouros. Também devem ser incorporados os chamados “endereços de entrada de favelas” ou de associações de moradores, geralmente usados pela comunidade para a entrega de correspondência, não necessariamente localizados nas favelas mas em suas proximidades. Essa experiência

vem sendo realizada em Salvador-BA (ISC/UFBA) e no Rio de Janeiro-RJ (Fiocruz/MS). Esses endereços devem compor uma lista de prováveis endereços de áreas carentes – na possibilidade de haver um mascaramento do verdadeiro local da ocorrência e respectiva população sob risco, indicando a área formal da cidade e não a referida comunidade.¹⁴

Vários programas de georreferenciamento foram apresentados durante a Oficina (Tabela 1). Geralmente, esses programas permitem o relacionamento entre uma base de dados de saúde e um cadastro georreferenciado de logradouros, incorporando novos campos aos registros de saúde e, dessa forma, possibilitando o georreferenciamento desses dados. Entre esses novos campos, podem-se inserir os pares de coordenadas correspondentes ao centro do objeto geográfico ou à unidade territorial de referência (setor censitário, lote ou trecho de rua). Algumas características possíveis de serem incorporadas a um programa a ser distribuído aos Municípios são apresentadas de forma sucinta, nos parágrafos a seguir.

Os programas podem fragmentar o campo de endereço em subcampos, conforme descrito anteriormente. A ação será facilitada se estiverem disponíveis tabelas de tipos, títulos e preposições como as utilizadas pelo IBGE e pela EBCT.

As principais falhas verificadas no processo de georreferenciamento são: nomes de logradouros não encontrados ou encontrados com grafias diferentes; numeração do endereço fora da faixa existente no cadastro de logradouros; endereço insuficiente; endereço com mais de um correspondente no cadastro de logradouros; e pequenos erros de grafia, como troca ou inversão de letras. O programa deve contar com um campo para caracterizar o tipo de problema encontrado no processo de georreferenciamento de dados, o que permite a geração das listas de erros a serem corrigidos *a posteriori*, pelo pessoal de campo ou de postos de saúde.

Também é interessante introduzir alguns instrumentos que facilitem a localização de endereços. No caso de Belo Horizonte-MG, o programa admite uma tolerância de numeração de endereço que permite localizar o vizinho mais próximo, ainda que o número exato do lote não seja encontrado no cadastro de endereços. E no Rio de Janeiro-RJ (SMS), quando o número exato do endereço não é encontrado, o SIG realiza interpolação na numeração existente no logra-

douro, permitindo o georreferenciamento aproximado da ocorrência.

Outrossim, deve-se promover um esforço especial no sentido de desenvolver ferramentas de inteligência artificial e busca fonética, para facilitar a identificação de grafias semelhantes de um mesmo logradouro ou erros de escrita, fato comentado anteriormente. Por exemplo, no mesmo Rio de Janeiro-RJ (SMS), realiza-se pesquisa fonética (*soundex*) para identificação das diferentes grafias usadas para nomear um logradouro. O programa local, ademais, permite criar o artifício de armazenamento, em um campo separado, de outros nomes considerados “sinônimos”, facilitando a identificação automática do logradouro e suas várias denominações. A mesma experiência foi desenvolvida pelo Departamento de Informática do SUS (Datusus/MS), do Ministério da Saúde.

Outra possibilidade de pareamento de dados de endereço com cadastros é a aproximação probabilística entre registros contidos nas duas tabelas consideradas. As posições relativas dos caracteres que compõem o campo de endereço são comparadas, gerando um escore que mede a probabilidade de acerto de cada registro, cuja decisão caberá ao operador do sistema. A experiência, realizada com o programa Reclink [desenvolvido pelo Instituto de Medicina Social da Universidade do Estado do Rio de Janeiro (IMS/UERJ)], tem alcançado êxito no pareamento de dados de diferentes sistemas de informações em saúde.¹⁵

O programa também pode produzir uma saída gráfica, na qual, rapidamente, são verificadas as localizações dos eventos de saúde, tendo como fundo o desenho das ruas ou quadras. Essa tecnologia também se encontra disponível em programas de acesso livre pela Internet.

Municípios brasileiros vêm adotando diversas estratégias para o georreferenciamento de dados de saúde, a depender dos recursos disponíveis e de sua base cartográfica digital. A maior parte deles, entretanto, não possui qualquer dos requisitos necessários à implantação de sistemas de geoprocessamento voltados para a saúde. Nesses casos, uma alternativa simples – e de baixo custo – está na busca ao Censo de 2000, precisamente em sua base cartográfica e base de endereços, adotando-se o Cadastro de Logradouros do IBGE, proposta defendida em diversos fóruns de discussão sobre geoprocessamento na área da Saúde. Além de facilitar o georreferenciamento de dados, o uso do Cadastro

de Logradouros do IBGE permite calcular populações, as quais serão usadas como número denominador de indicadores epidemiológicos, em diferentes níveis de agregação submunicipal [setor censitário; bairro; área de abrangência de unidade de saúde; área de cobertura do Programa Saúde da Família (PSF); área de controle da dengue; e distrito sanitário].

A demanda da Saúde pelo geoprocessamento é crescente e depende da disponibilidade e qualidade de dados georreferenciados.

A análise desses dados pelo nível local tem incentivado a melhoria dos sistemas de informações e das bases cartográficas.

Outra questão importante para discussão resulta do grau de urbanização e consolidação de vias existente. Centros urbanos com maior população, cujo crescimento não foi – ou foi parcialmente – planejado –, a grande maioria no país – tendem a apresentar um sistema viário mais complexo e heterogêneo. Suas áreas periféricas e favelas contam com uma lógica viária distinta à do restante da cidade. Essa estrutura, em sua informalidade, dificulta a coleta e análise de dados de endereço, o que pode explicar a baixa eficiência de georreferenciamento de dados em cidades como Salvador-BA e Rio de Janeiro-RJ, por exemplo, quando considerados somente os logradouros constantes no cadastro do censo demográfico. Tal eficiência pode aumentar consideravelmente com o desenvolvimento de um trabalho de compatibilização desses endereços informais com listas de endereços de favelas e loteamentos recentes, inclusive ocupações. Essa fase, chamada de georreferenciamento semi-automático, exige grande esforço técnico na atualização e busca de endereços alternativos aos constantes dos cadastros urbanos convencionais.

Os sistemas de informações em saúde devem permitir a entrada padronizada de dados de endereço. Essa padronização pode ser feita mediante a incorporação de listas de tipos, títulos e logradouros, fornecidas pela Prefeitura Municipal; ou pelo IBGE, no caso do Cadastro de Logradouros do Censo de 2000. Esse

Cadastro ainda pode ter várias outras aplicações, no âmbito das ações do Ministério da Saúde.

Ressalta-se, ademais, a oportunidade de desenvolvimento de um programa específico de georreferenciamento de dados de saúde aproveitando as experiências acumuladas em alguns Municípios (Tabela 1). Esse programa, a ser distribuído a todas as SMS, incentivaria o uso de técnicas de geoprocessamento e a melhoria da qualidade de dados de endereço. Para os Municípios sem cadastro de logradouros, o programa incorporaria a base do Censo de 2000.

Os resultados obtidos no processo de georreferenciamento de dados de saúde são, em grande parte, dependentes da qualidade desses dados e da base cartográfica sobre a qual serão dispostos. No nível local, têm-se verificado a atuação de projetos de produção de bases cartográficas digitais, que, salvo algumas exceções, é diretamente proporcional ao tamanho e nível de organização do Município. Nessa escala, o geoprocessamento em saúde tem se beneficiado da digitalização dos componentes urbanos (arruamento e divisões internas do Município) e da estruturação dos códigos de logradouros, para o planejamento urbano e a arrecadação de impostos, respectivamente.

No Brasil, em Municípios onde os investimentos para o desenvolvimento de sistemas de informações geográficas – SIG – foram maciços, como é caso de Belo Horizonte-MG, o georreferenciamento parte dessa base cartográfica detalhada. Seu custo, entretanto, é muito alto e transcende a capacidade do setor Saúde. Uma estratégia alternativa e efetiva no alcance de bons resultados encontra-se na microlocalização de eventos de saúde, baseada na criação de bons cadastros alfanuméricos e na padronização da entrada de endereços. Na maioria das cidades brasileiras, as regras básicas de padronização de endereços não são obedecidas, principalmente em suas periferias e áreas de ocupação irregular, como favelas. Algumas cidades apresentam alto índice de numeração irregular; outras, como Brasília-DF e Palmas-TO, utilizam endereços por quadra e não por logradouro. E os mapas urbanos digitais, quando disponíveis, não seguem um padrão único.

A demanda do setor Saúde por ferramentas de geoprocessamento é crescente. A plena utilização dessas técnicas depende, todavia, da disponibilidade e qualidade de dados georreferenciados. A análise desses dados pelo nível local, de sua parte, é o que tem incentivado a melhoria da qualidade dos sistemas

de informações e das bases cartográficas, gerando um mecanismo de retroalimentação que garante o aperfeiçoamento de projetos de geoprocessamento de dados de saúde.

O Ministério da Saúde, em ação conjunta com a Organização Pan-Americana da Saúde/Organização Mundial da Saúde (OPAS/OMS), vem desenvolvendo uma estratégia de articulação interinstitucional focada na criação da Rede Interagencial de Informações para a Saúde – Ripsa –, já mencionada neste relato. Integrada por entidades e instituições envolvidas com a produção e análise de dados em saúde, a Ripsa viabiliza parcerias e intercâmbio visando ao aperfeiçoamento da compreensão do quadro sanitário do Brasil. A Rede é composta de subgrupos de trabalho, os denominados Comitês Técnicos Interdisciplinares (CTI), definidos segundo necessidades identificadas. O Comitê Temático Interdisciplinar sobre Geoprocessamento e Dados Espaciais em Saúde¹⁶ tem, como integrantes, a OPAS/OMS, a Secretaria de Vigilância em Saúde do Ministério da Saúde (SVS/MS), o Datasus/MS, a Fiocruz/MS, o IBGE e representantes das Secretarias de Estado e Municipais de Saúde. Sua missão primordial

reside na coordenação do movimento nacional para o aperfeiçoamento da análise dos sistemas de informações em saúde, bem como na articulação entre instituições produtoras de dados de interesse para o geoprocessamento em saúde, consubstanciada na reunião de dados textuais e cartográficos e na divulgação de técnicas e programas de análise espacial e dados.

Agradecimentos

Os autores agradecem a participação de técnicos do Ministério da Saúde e das Secretarias Municipais de Saúde, por sua contribuição para o levantamento do estágio atual dos projetos de georreferenciamento em cidades brasileiras e proposição de alternativas de aperfeiçoamento desses sistemas. São eles: Carla Domingues; Carlos Rodrigo; Cláudia Risso; Claunara Schilling Mendonça; Davi Felix Martins Junior; Denise Santos Correia de Oliveira; Eduardo Macário; Hudson Carrano; Maria Cristina de Matos Almeida; Rui Flores, Ruth Glatt; Vera Regina Barêa; Walter Luis Batista Ferraz; Wendell Kill; e Werner Leyh.

Referências bibliográficas

1. Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Race/ethnicity, gender, and monitoring socioeconomic gradients in health: a comparison of area-based socioeconomic measures. The public health disparities geocoding project. *American Journal of Public Health* 2003; 93(10):1655-1671.
2. Bennett D. Explanation in medical geography. Evidence and epistemology. *Social Science and Medicine* 1991;33:339-346.
3. Barcellos C, Santos SM. Colocando dados no mapa: a escolha da unidade espacial de agregação e integração de bases de dados em saúde e ambiente através do geoprocessamento. *Informe Epidemiológico do SUS* 1997;VI(1):21-29.
4. Malta DC, Almeida MCM, Dias MAS, Merhy EE. A mortalidade infantil em Belo Horizonte, Minas Gerais, Brasil, por área de abrangência dos Centros de Saúde (1994-1996). *Cadernos de Saúde Pública* 2001;17(5):1189-1198.
5. Jones HM. Applications of remote sensing to the identification of the habitats of parasites and disease vectors. *Parasitology Today* 1989;5(8):244-250.
6. Vine MF, Degnan D, Hanchette C. Geographic Information Systems: their use in environmental epidemiologic research. *Environmental Health Perspectives* 1997;105(6):598-605.
7. Bailey TC. A review of statistical spatial analysis in geographical information systems. In: Fotheringham S, Rogerson P. *Spatial Analysis and GIS*. Londres: Taylor & Francis; 1994. p. 13-44.
8. Câmara G, Monteiro AMV. Geocomputation techniques for spatial analysis: are they relevant to health data? *Cadernos de Saúde Pública* 2001;17(5):1059-1081.
9. Barcellos C, Ramalho WM. Situação atual do geoprocessamento e da análise de dados espaciais em saúde no Brasil. *Revista de Informática Pública* 2002;4(2):221-230.
10. Dramowicz E. Three standard geocoding methods. *Directions Magazine* [monography on the Internet]

- [cited 2004 Out]. Available from: http://www.directionsmag.com/article.php?article_id=670.
11. Ximenes RAA, Martelli CMT, Souza WV, Lapa TM, Albuquerque MFM, Andrade ALSS, Morais Neto OL, Silva AS, Lima MLC, Portugal JL. Vigilância de doenças endêmicas em áreas urbanas: a interface entre mapas de setores censitários e indicadores de morbidade. *Cadernos de Saúde Pública* 1999;15(1):53-62.
 12. Skaba DA, Carvalho MS, Barcellos C, Martins PC, Terron SL. Geoprocessamento dos dados da saúde: o tratamento dos endereços. *Cadernos de Saúde Pública* 2004;20(6):1753-1756.
 13. Davis CA, Fonseca FT, Borges KAV. A Flexible addressing system for approximate geocoding. In: *Proceedings of the 5th Brazilian Symposium on Geoinformatics*; 2003; Campos do Jordão-SP, Brazil. Campos do Jordão-SP; 2003.
 14. Gracie R, Barcellos C, Pina F, Magalhães M. Problemas de localização de eventos de saúde nas favelas do Município do Rio de Janeiro. In: *Anais do 2º Simpósio Nacional de Geografia da Saúde*; 2005; Rio de Janeiro-RJ, Brasil. Rio de Janeiro-RJ; 2005.
 15. Camargo-Jr KR, Coeli CM. Reclink: aplicativo para o relacionamento de bases de dados, implementando o método probabilistic record linkage. *Cadernos de Saúde Pública* 2000;16(2):439-447.
 16. Geosaúde. Comitê Temático Interdisciplinar sobre Geoprocessamento e Dados Espaciais em Saúde [monografia na Internet]. Rio de Janeiro-RJ: Fiocruz; 2003 [acessado 2005 out.]. Disponível em: <http://www.geosaude.cict.fiocruz.br>

Recebido em 22/09/2006

Aprovado em 15/06/2007