

GEOREFERENCIAÇÃO DE DADOS EM MICRO-ÁREAS. DESAFIO À UTILIZAÇÃO DOS SIG NA ÁREA DA SAÚDE

Maria de Fátima de Pina¹
Andreia Olhero^{1,2}

¹ Instituto de Engenharia Biomédica – INEB

² Serviço de Higiene e Epidemiologia, Faculdade de Medicina do Porto

INTRODUÇÃO

A utilização de técnicas de análise espacial, seja através dos Sistemas de Informação Geográfica seja através de estatística espacial ou de uma combinação das duas, tem despertado muito interesse entre os epidemiologistas e outros profissionais da saúde uma vez que permite ter uma visão abrangente da saúde, dos indivíduos no contexto sócio-cultural-ambiental em que estão inseridos.

A georeferenciação é quase invariavelmente, a etapa inicial do processo de análise espacial de dados em que se pretende correlacionar prevalência e incidência de algumas doenças, com as condições socioeconómicas e ambientais dos lugares de residência dos participantes. Não raras vezes, é nesta etapa que se verificam os primeiros entraves à utilização dos SIG na área da saúde.

Quando o interesse é em estudos ecológicos, em que o objecto de análise é uma população associada a uma área espacialmente delimitada, a disponibilidade tanto de mapas, quanto de dados estatísticos é muito grande e é com relativa facilidade que se realiza a georeferenciação das bases de dados (normalmente utilizando os códigos de freguesias, concelhos, distritos, NUTS, etc.). No entanto, em alguns estudos, há interesse em localizar eventos de saúde em micro-áreas (como subsecção geográfica ou segmentos de ruas) e nestes casos, a georeferenciação normalmente é realizada através de uma morada e as dificuldades que daí advêm ainda hoje limitam a plena utilização dos SIG na área da saúde [1; 2].

Este trabalho tem por objectivo relatar a experiência de georeferenciar, uma base de dados com informações sobre os participantes de um estudo epidemiológico denominado EpiPorto, desenvolvido pelo serviço de *Higiene e Epidemiologia da Faculdade de Medicina do Porto*. A georeferenciação foi realizada no *Instituto de Engenharia Biomédica* da Universidade do Porto, no âmbito de uma nova área de pesquisa intitulada “Estudo espacial das doenças osteoarticulares e outros problemas de saúde”.

MATERIAL E MÉTODOS

DADOS

Utilizamos a base de dados dos participantes do estudo EpiPorto, que contém informações clínicas e socioeconómicas dos participantes, além das suas moradas.

Para facilitar a georeferenciação, a morada foi segmentada em diferentes campos, entre os quais o “*tipo*” (Alameda, Rua, Travessa, Bairro, Viela, Caminho, Beco, Calçada, Praça, Largo, Passeio, etc.) o “*título*” (Professor, Engenheiro, Doutor, Arquitecto, Actor, Capitão, General, Maestro, Padre, Senhora, Santo, etc.), o “*nome*”, o “*numero de polícia*”, o “*código postal*” o “*concelho*” e a “*freguesia*”. A estruturação da morada nestes campos em vez de um único, permite que se utilizem alguns critérios de busca quando a morada do participante não é encontrada automaticamente no mapa.

Utilizamos também um mapa da cidade do Porto (em suporte digital) com os eixos de via (centerlines) e as informações referentes ao nome da via e numeração inicial e final de cada segmento de via. É de referenciar ainda a utilização de alguns roteiros do Porto, em formato analógico que contêm os nomes e as ruas da cidade.

METODOLOGIA

O processo de georeferenciação iniciou-se com uma preparação prévia da base de dados, que consistiu na verificação e correcção de vários erros encontrados, entre os quais, nomes de ruas incompletos, erros ortográficos e erros de digitação do nome das vias tanto nas moradas dos participantes quanto no mapa digital.

A georeferenciação foi realizada em três fases: a primeira, automática, consistiu na utilização da funcionalidade do SIG para localizar cada morada no mapa. A segunda fase consistiu numa georeferenciação semi-automática dos registos não georeferenciados na primeira fase, por problemas de grafia ou porque o número de polícia não coincidia com a numeração constante no mapa e foi realizada de maneira interactiva através da selecção de algumas opções que o programa de georeferenciação apresentou. A terceira fase, manual, foi efectuada nos registos que não puderam ser georeferenciados nem na primeira nem na segunda fase e consistiu na busca dos nomes das vias em guias de rua da cidade do Porto. Por fim, quando a georeferenciação não foi possível, em nenhuma das situações anteriores, efectuaram-se telefonemas aos participantes solicitando as informações correctas para localizar a morada.

A situação mais trabalhosa e que levantou mais problemas foi sem dúvida a georeferenciação manual. Os principais problemas impeditivos à georeferenciação automática ou semi-automática são variados, embora se destaquem as seguintes situações:

- Nomes de ruas completamente desconhecidas, por não aparecerem nem no mapa digital nem nos mapas analógicos. Nestes casos o problema ficou a dever-se à errada digitação da morada, ou à inserção indevida na base de dados de uma morada não pertencente ao concelho do Porto, ou ainda, moradas fornecidas incorrectamente por parte dos participantes. Sempre que necessário recorreu-se ao contacto telefónico para solucionar estas situações. Em alguns casos, em que o contacto telefónico não foi possível (por exemplo, porque o participante não conhecia a morada correcta, ou o número de telefone tinha mudado ou simplesmente não atendeu a chamada) foram considerados dados perdidos;
- Ruas não identificadas no mapa digital mas presentes nos mapas em papel e por outro lado, ruas delineadas no mapa digital mas sem qualquer nome identificativo. Nestas duas situações o problema residiu na falta de actualização do mapa digital fornecido, tornando imprescindível o uso dos mapas em formato analógico. No primeiro caso referido, os mapas analógicos serviram para fazer uma georeferenciação manual aproximada, através do estabelecimento de comparações entre os mapas digital e analógico. No segundo caso, o mapa analógico, ajudou-nos a actualizar a base cartográfica digital. As situações que suscitaram mais dúvidas não dispensaram também o contacto telefónico;
- Moradas incorrectas ou incompletas e portanto impossíveis de georeferenciar. São exemplo disso, casos de moradas sem o número da porta, outros apenas com o bloco ou só o andar. Nestas situações o problema foi solucionado mais uma vez através do contacto telefónico ou através de informações contidas em moradas de outros participantes vizinhos.

O problema das moradas fornecidas incorrectamente, deve-se essencialmente ao facto da cidade do Porto ter sofrido algumas alterações ao nível da toponímia das suas ruas, embora alguns participantes não tenham adoptado o novo nome, continuando, por isso, a usar o nome antigo. Esta situação foi bastante mais visível no caso dos “bairros”, cujas ruas foram, recentemente, identificadas com nomes individuais.

Foram também frequentes os casos de numeração repetida na mesma rua, o que fica a dever-se (segundo as pessoas contactadas) ao facto de não haver um controle rígido por parte

das entidades responsáveis, o que permite às pessoas colocarem os números que desejam, nas suas habitações.

Nos casos em que, os indivíduos contactados desconheciam o número da sua casa ou nos casos em que as moradas dos participantes se revelaram completamente desconhecidas, pedimos às pessoas referências que nos pudessem ajudar a identificar, o mais aproximadamente possível, o ponto de localização da sua habitação. Para isso, perguntamos o nome das ruas circundantes e questionamos a posição da habitação (no princípio, no meio ou no fim da rua) em função das ruas vizinhas. A georeferenciação foi realizada maneira aproximada.

RESULTADOS

O total de regtos na base de dados do EpiPorto é de 2427, dos quais 6 foram excluídos por não pertencem ao concelho do Porto. Dos restantes regtos, 4 (0.2%) não puderam ser georeferenciados por se tratarem de moradas incompletas ou incorrectas e cujas tentativas de localização não surtiram efeito.

Os 2417 regtos restantes foram georeferenciados em diversas etapas conforme descrito na metodologia e os resultados de cada etapa estão apresentados na tabela 1, onde se pode verificar que a grande maioria dos regtos (71%) foram georeferenciados sem problemas, de maneira automática (FASE 1) sendo que uma pequena parcela (13%) foi realizada de modo semi-automático (FASE 2).

Tabela 1: Georeferenciação no projecto EpiPorto

	TOTAL	%
FASE 1	1717	71
FASE 2	315	13
FASE 3	385	16
PERDAS	4	0.2
TOTAL	2427	100

Dos 385 casos georreferenciados manualmente (FASE 3), a maior percentagem (75%) diz respeito a ruas não existentes inicialmente no mapa digital mas identificadas nos mapas analógicos. Nestes casos a georeferenciação foi feita de maneira aproximada identificando no

mapa digital a posição da morada. Outra situação (17%) diz respeito a ruas que estão desenhadas no mapa digital mas sem qualquer nome identificativo, neste caso o mapa digital foi actualizado com o nome da rua e procedeu-se a nova georeferenciação. Os restantes 8% dizem respeito a moradas desconhecidas, uma vez que, não existiam nem no mapa digital nem nos mapas em papel. Nestas situações recorreu-se aos telefonemas e à georeferenciação aproximada. O total de telefonemas realizados nesta fase foi de 216.

DISCUSSÃO

A utilização dos SIG na saúde em Portugal é ainda muito pouco usual. Este trabalho é inovador e dará aos gestores da área da saúde uma importante contribuição no planeamento de acções de prevenção da saúde.

No caso do projecto EpiPorto, o processo de georeferenciação foi bastante demorado, essencialmente devido à falta de qualidade no preenchimento do campo relativo à morada dos participantes e às inconsistências entre o nome constante no mapa digital e a base de dados do EpiPorto.

Para prever esta situação seria importante que as câmaras municipais, tanto do Porto como de outros concelhos, disponibilizassem bases de dados com todos os nomes de ruas oficialmente reconhecidos de maneira a padronizar a sua utilização por todos os utilizadores.

Neste trabalho, foram encontradas várias incompatibilidades não só no nome das ruas, fornecidas pelos participantes, mas também nos nomes constantes nos mapas, tanto digital como analógico.

Estas incompatibilidades dificultam o processo de georeferenciação automático, tornando-o muito mais lento e menos preciso.

A georeferenciação em micro-áreas depende da utilização de bases cartográficas em grande escala, com arruamentos (centerlines) e informações sobre o nome das ruas, assim como o número de polícia inicial e final dos segmentos das ruas. Porém, o acesso a estas bases é difícil e limitado uma vez que a sua aquisição pode ser muito dispendiosa e em alguns concelhos não existe ou não está disponibilizada. O acesso às bases cartográficas é mais fácil quando se trata de grandes centros urbanos, como Porto e Lisboa, o que, como facilmente se percebe, põe em causa a execução de projectos que envolvam outras áreas do país.

REFÊRÊNCIAS BIBLIOGRÁFICAS

1. BAILEY, TREVOR; *Spatial statistical methods in helth*; 2001; Cadernos de Saúde Pública, Análise de Dados Espaciais em Saúde; Vol 17.
2. ROCHA, CARLOS; PINA, FÁTIMA; OLIVEIRA, GILBERTO; MAGALHÃES MÔNICA; CARRIJO, RENATA; 2004. *Automatização do Georeferenciamento de eventos em Micro-áreas*, (artigo não publicado)