

Concepção de um modelo geoespacial aplicado na gestão à saúde

Design of a geospatial model applied to Health management

La producción de un modelo geoespacial aplicado a la gestión de la salud

Marcelo Rosano Dallagassa^{I,II}

ORCID: 0000-0003-0570-4405

Franciele Iachecen^{II}

ORCID: 0000-0002-2019-1479

Deborah Ribeiro Carvalho^I

ORCID: 0000-0002-9735-650X

Sergio Ossamu Ioshii^I

ORCID: 0000-0003-1208-2448

^IPontifícia Universidade Católica do Paraná. Curitiba-PR, Brasil.

^{II}Unimed Federação do Estado do Paraná,
Núcleo de Inteligência em Saúde. Curitiba-PR, Brasil.

Como citar este artigo:

Dallagassa MR, Iachecen F, Carvalho DR, Ioshii SO.
Design of a geospatial model applied to Health
management. Rev Bras Enferm. 2019;72(2):420-6.
doi: <http://dx.doi.org/10.1590/0034-7167-2018-0589>

Autor Correspondente:

Marcelo Rosano Dallagassa
E-mail: dallagassa@unimedpr.coop.br



RESUMO

Objetivo: Identificar geograficamente os beneficiários categorizados como propensos à doença Diabetes Mellitus Tipo 2, utilizando o reconhecimento de padrões em uma base de dados de uma operadora de plano de saúde, por meio da mineração de dados. **Método:** Desenvolveram-se as seguintes etapas: fase inicial, levantamento de informações. Desenvolvimento, construção do processo de extração, transformação e carga do banco de dados. Implantação, apresentação das informações geográficas por meio da ferramenta de georreferenciamento. **Resultados:** Como resultados, apresenta-se o mapeamento do Paraná em relação a sua rede assistencial e a concentração de Diabetes Mellitus Tipo 2, oportunizando a identificação de relações de causa-efeito. **Conclusão:** Conclui-se que a análise de informações georreferenciadas, vinculadas às informações de saúde obtidas por meio da técnica de mineração de dados, pode ser um excelente instrumento para a gestão da saúde de uma operadora de plano de saúde, contribuindo para o apoio à tomada de decisões em saúde.

Descritores: Atenção à Saúde; Mineração de Dados; Georreferenciamento; Saúde Suplementar; Doenças Crônicas.

ABSTRACT

Objective: To identify geographically the beneficiaries categorized as prone to Type 2 Diabetes Mellitus, using the recognition of patterns in a database of a health plan operator, through data mining. **Method:** The following steps were developed: the initial step, the information survey. Development, construction of the process of extraction, transformation, and loading of the database. Deployment, presentation of the geographical information through a georeferencing tool. **Results:** As a result, the mapping of Paraná according to its health care network and the concentration of Type 2 Diabetes Mellitus is presented, enabling the identification of cause-and-effect relationships. **Conclusion:** It is concluded that the analysis of georeferenced information, linked to health information obtained through the data mining technique, can be an excellent tool for the health management of a health plan operator, contributing to the decision-making process in Health.

Descriptors: Health Care; Data Mining; Geographic Mapping; Supplementary Health; Chronic Disease.

RESUMEN

Objetivo: Identificar geográficamente a los beneficiarios categorizados como propensos a la enfermedad Diabetes mellitus tipo 2, utilizándose el reconocimiento de patrones en una base de datos de cierta compañía de seguro médico por medio de la minería de datos. **Método:** Se desarrollaron las siguientes etapas: fase inicial, levantamiento de información. Desarrollo, construcción del proceso de extracción, transformación y carga en la base de datos. Implantación, presentación de la información geográfica mediante la herramienta de georreferenciación. **Resultados:** Se presenta el mapeo de Paraná (Brasil) con relación a su red asistencial y la concentración de Diabetes mellitus tipo 2, proporcionando la identificación de las relaciones de causa-efecto. **Conclusión:** Se concluyó que el análisis de las informaciones georreferenciadas, vinculadas a las informaciones de salud obtenidas por la técnica de minería de datos, puede ser una excelente herramienta en la gestión de salud de cierta compañía de seguro médico, lo que contribuye al apoyo a la toma de decisiones en salud.

Descriptorios: Atención a la Salud; Minería de Datos; Mapeo Geográfico; Salud Suplementaria; Enfermedades crónicas.

Submissão: 17-07-2018

Aprovação: 04-09-2018

INTRODUÇÃO

A epidemiologia consiste no estudo das inter-relações dos vários determinantes da frequência e distribuição de doenças em um conjunto populacional. Esse conhecimento é fundamental na prevenção e no tratamento de doenças, oferecendo elementos consistentes para encaminhamentos na área da saúde. A identificação de padrões sobre a ocorrência de doenças nas populações humanas e dos fatores que as influenciam e condicionam definem o objeto de estudo da epidemiologia⁽¹⁾.

Em 1854, na cidade de Londres, o número de casos de cólera era considerado estável e com baixa incidência, porém, em um determinado momento, se tornou uma grande epidemia, chegando a registrar mais de 500 casos fatais em aproximadamente 12,5 hectares em um período de dez dias. John Snow, considerado por muitos como o pai da epidemiologia, desenvolveu na época, um estudo observacional que vinculava os casos geograficamente, identificando o ponto médio – uma bica de água do Bairro de *Soho* – como sendo o responsável pela disseminação da doença. Pela relação de causa-efeito, após cuidadosas e intensas investigações, concluiu que deveria rejeitar as demais hipóteses sobre a origem da epidemia, afirmando que a via hídrica era a responsável pela transmissão do *Vibrio cholerae*⁽¹⁾.

Outro exemplo da associação de posicionamento geográfico com uma relação de causa-efeito é a síndrome Li-Fraumeni, patologia que torna os indivíduos mais vulneráveis a determinados tipos de neoplasias, identificada na mutação da expressão gênica. Observou-se que esses casos localizavam-se geralmente na região sul e sudeste do Brasil e, particularmente, os pontos geográficos dos indivíduos portadores da síndrome, vinculavam-se ao trajeto de um imigrante português no século XVIII, teoria hoje comprovada da descendência genética da doença⁽²⁾.

Atualmente, o Brasil passa por uma transição demográfica, decorrente da diminuição da mortalidade infantil e o aumento da expectativa de vida da população, que passou de 50 para 73 anos, em média. Outros fatores também influenciam esse cenário, como por exemplo, a diminuição da taxa de fecundidade entre as mulheres, que nos anos 60 era de, em média, seis filhos e atualmente passou para menos que dois. Esses fatores impactam diretamente no cenário nacional futuro, estimando-se que em 2050, haverá 64 milhões de idosos, representando 29,7% da população brasileira⁽³⁾.

Com o aumento da expectativa de vida, observa-se a mudança no perfil epidemiológico, elevando a prevalência de doenças crônicas não transmissíveis na população em geral⁽⁴⁾. O envelhecimento populacional impacta diretamente nos gastos com a saúde, motivado pela crescente proporção de idosos na população e pela crescente utilização de recursos assistenciais. A magnitude do aumento dos custos com a saúde dependerá da qualidade de vida das pessoas e da existência ou não de doenças e comorbidades⁽⁵⁾. O Banco Mundial, frente a esses cenários, destaca a importância da organização do sistema de saúde de se adaptar ao novo perfil epidemiológico, afirmando que a promoção da saúde e a prevenção de doenças continuarão sendo um dos maiores desafios para o setor.

No âmbito da saúde suplementar há um desafio para as operadoras de planos de saúde em relação a sua sustentabilidade. Essa tarefa tem se dificultado devido às imposições legais da Agência Nacional de Saúde Suplementar (ANS), as quais impactam diretamente em seus custos, devido a limitação dos reajustes dos contratos superarem a

inflação médico-hospitalares Variação de Custo Médico-Hospitalar (VCHM), que em dezembro de 2016, no período de 12 meses, apresentou um acúmulo de 20,4%⁽⁵⁾, índice muito superior à variação da inflação geral Índice Nacional de Preços ao Consumidor Amplo (IPCA), que para o mesmo período, foi de 6,29%⁽⁶⁾. Além disso, bianualmente, há a introdução de novas e onerosas tecnologias e procedimentos que ampliam o rol de cobertura dos serviços obrigatórios para as operadoras de planos de saúde.

Diante desse contexto, há a necessidade da utilização de aplicações que viabilizem a atuação de medidas rápidas e inteligentes a fim de minimizar os custos diretos da operação, buscando o equilíbrio com a satisfação do cliente e os prestadores de serviços, pois “garantir as condições mínimas de solvência das seguradoras é condição essencial para viabilizar a existência de um mercado segurador que atenda o objetivo de proteção aos interesses dos segurados”⁽⁷⁾.

No entanto, as operadoras de planos de saúde encontram dificuldades para atuar na prevenção de doenças crônicas. Uma das razões está diretamente associada à ausência de dados clínicos de seus beneficiários, sistematizados em suas bases de dados, impossibilitando a extração de informações e o conhecimento de maneira mais automatizada. Isso também se deve ao fato da construção dos sistemas de informação dessas empresas serem direcionados apenas para o controle administrativo, visando somente o pagamento dos prestadores de serviços e a gestão de contratos⁽⁸⁾.

No agravamento desse quadro, em 2007 foi publicada a resolução de número 1.819/2007 do Conselho Federal de Medicina⁽⁹⁾, a qual proíbe a colocação da Classificação Internacional de Doenças (CID) no preenchimento das guias de Troca de Informação da Saúde Suplementar (TISS) em atendimentos ambulatoriais.

O grande volume de informações gerados pelos sistemas de informações de saúde, agregadas às informações dos ambientes externos que, em muitas vezes, são alimentados em tempo real e disponibilizados em diversos formatos (textos, vídeos, imagens, mensagens, expressões gênicas etc.), definem o termo “*Big Data*”⁽¹⁰⁾.

A descoberta de padrões no ambiente “*Big Data*”, quando se utiliza os métodos tradicionais de análise, além de demandar muito tempo e recursos, não garante a plena exploração de seus potenciais⁽¹⁰⁾. Com o uso da mineração de dados, um processo não trivial de descoberta de informações escondidas e possivelmente úteis, esse trabalho torna-se mais eficiente, possibilitando o apoio à tomada de decisão⁽¹¹⁾.

Integrar as informações de saúde, vinculadas a dados geográficos, ambientais, socioeconômicos e demográficos permite criar hipóteses para investigação científica sobre as causas e as origens de determinadas doenças, oportunizando o conhecimento de prevalências, incidências, controle de transmissão e tratamento de enfermidades. A análise das informações georreferenciadas também possui uma grande importância acerca da descoberta das relações de causa-efeito, proporcionando uma análise dinâmica e tornando possível a identificação de grupos mais vulneráveis e o conhecimento da atual situação de saúde da população⁽¹²⁻¹³⁾.

Existem inúmeras bases de dados que quando agregadas e enriquecidas com outras informações vinculadas a um registro de localização geográfica, permitem a descoberta de características importantes para a identificação de padrões de determinada região.

Associar as metodologias de reconhecimento de padrões com ferramentas de georreferenciamento a uma questão de

pesquisa bem formulada pode gerar instrumentos úteis para a descoberta de conhecimentos na área da saúde, possibilitando a implantação de práticas importantes para a redução de custos e melhoria da qualidade de vida da população, mesmo considerando as dificuldades em relação à ausência de informações clínicas.

Assim, a partir das dificuldades e necessidades apontadas, surge a questão de pesquisa que se pretende abordar. Quais são as técnicas e métodos de descoberta do conhecimento, que possibilitem a criação de indicadores referenciados geograficamente para o monitoramento da atenção à saúde da carteira de beneficiários de uma operadora de plano de saúde, direcionados a promoção da saúde e prevenção de doenças.

Foram identificados alguns trabalhos relacionados para o reconhecimento de doenças por meio da base de dados, que consideram a utilização de procedimentos associados, algoritmos e métodos que possibilitam identificar nos registros de pagamentos de contas, a descoberta de padrões de utilização vinculada a determinadas doenças, dentre os quais, podemos citar um estudo na região de Pávia, Itália, que consistiu em uma pesquisa com dados administrativos e clínicos sobre diabetes, utilizando regras de associação temporal. A pesquisa teve como objetivo analisar dados do sistema de saúde da região. O método também destacou as associações temporais frequentes de interesse sobre o diagnóstico relacionado com a condição clínica do paciente⁽¹⁴⁾.

Identificou-se outro estudo, também desenvolvido na Universidade de Pávia, Itália, que consiste em um método para a identificação de padrões baseado em dados temporais, que foram extraídos em forma de regras de associação temporal. Essa pesquisa concluiu que há muito potencial para mineração de dados, em busca de regras de associação temporais e sugere que estes métodos podem ser mais explorados, uma vez que é crescente a demanda por ferramentas que descubram regras de potencial interesse para gestores⁽¹⁵⁾.

Cita-se ainda, outra metodologia para identificação de beneficiários em uma operadora de planos de saúde com indicativos de Diabetes Mellitus Tipo 2 no estado do Paraná, Brasil. Por meio de um histórico de utilização realizou-se a seleção das variáveis relevantes para a geração dos dados. A seleção foi submetida ao algoritmo J48 para o encontro de regras, avaliadas por um conjunto de especialistas. A partir dessa técnica foi possível realizar a extração de padrões para outras doenças crônicas, que hoje fazem parte de aplicações para identificação e categorização de beneficiários⁽⁸⁾.

A partir dessa problematização e da definição desses conceitos, surge a necessidade da construção de um ambiente que agregue informações de diversas fontes para o enriquecimento das bases de dados internas, integradas com técnicas de reconhecimento de padrões e associadas a um referencial geográfico, possibilitando a identificação sobre doenças que afetam a população, a fim de promover um monitoramento mais efetivo em relação às doenças crônicas não transmissíveis e viabilizar o desenvolvimento de ações regionais dos programas de promoção da saúde.

OBJETIVO

Identificar geograficamente os beneficiários categorizados como propensos à doença Diabetes Mellitus (DM) Tipo 2, utilizando o reconhecimento de padrões em uma base de dados de uma operadora de plano de saúde, por meio da mineração de dados.

MÉTODO

Aspectos éticos

Esta pesquisa foi realizada com assentimento da instituição e submissão do projeto ao Comitê de Ética em Pesquisa da Pontifícia Universidade Católica do Paraná (PUCPR).

Desenho, local do estudo e período

Tratou-se de um estudo descritivo, de abordagem quantitativa, de natureza retrospectiva que utilizou a base de dados de uma operadora de planos de saúde de grande porte do estado do Paraná, no ano de 2017.

População ou amostra: critérios de inclusão e exclusão

A amostra teve como base beneficiários do plano de saúde da operadora, que no período observado encontravam-se ativos e que devido a sua utilização e frequência dos serviços utilizados, foram categorizados como propensos à DM Tipo 2. Como critérios de inclusão, estabeleceu-se a seleção de beneficiários ativos no plano de saúde no ano de 2017, que se enquadravam na categoria de propensos à DM Tipo 2; como critérios de exclusão: elegeu-se os beneficiários inativos no ano de 2017 não categorizados para DM Tipo 2. A escolha da patologia deve-se ao fato de ser uma doença em constante crescimento, motivado por diversos fatores, como a obesidade, sedentarismo, envelhecimento populacional e aumento da sobrevivência do paciente portador da doença⁽⁸⁾. Além disso, estimou-se que as doenças crônicas não transmissíveis motivaram 38 milhões de óbitos em 2012 e somente a DM Tipo 2 foi a causa básica de 5,3% deles⁽¹⁶⁾.

Protocolo do estudo

Após a estruturação da base de dados, dividiu-se o estudo em três etapas: etapa inicial, desenvolvimento e implantação. Na etapa inicial, realizou-se um levantamento das potenciais bases de pesquisa públicas de registros de saúde, com o auxílio dos setores da instituição, com a intenção da construção do referencial teórico, objetivando identificar e justificar possíveis aplicações e oportunidades de utilização para o desenvolvimento do projeto com a devida aprovação junto aos gestores da organização; em planejamento e em conjunto com os gestores, elencaram-se as possíveis bases de dados com as áreas da organização. As bases elencadas foram obtidas por meio de um questionário. A pesquisa foi formulada pelos autores da pesquisa, com o objetivo de levantar as bases de informações necessárias para compor o projeto. A aplicação do questionário ocorreu em 3 de outubro de 2016 na sede da operadora de planos de saúde. Antecedendo a entrega dos questionários, realizou-se uma apresentação, onde se discorreu sobre o principal objetivo do projeto, bem como as bases públicas já contempladas. Na sequência, foram entregues 27 questionários, destes, dez (41,6%) foram devolvidos preenchidos. A pesquisa teve como público alvo especialistas que atuam nos setores de informação da OPS em diferentes municípios do estado do Paraná. O questionário foi estruturado com três questões básicas: uma referente à área da saúde, outra à área de mercado e outra à área estratégica. Cada pergunta continha uma avaliação

sobre importância de cada área, que variava de 0 a 5, sendo 0, pouca importância, 3, média importância e 5, muita importância e uma questão aberta, solicitando ao entrevistado que descrevesse outras possibilidades de fontes de informações passíveis de aplicação. Ainda nessa etapa, definiu-se a ferramenta utilizada para elaborar o processo de desenvolvimento e implantação com a premissa de utilização de um *software* livre. Na etapa de desenvolvimento, realizou-se a instalação do *software* previamente elencado e aplicou-se treinamento dos *softwares* Quantum GIS⁽¹⁷⁾ e Pentaho PDJ⁽¹⁸⁾ para a equipe interna responsável pelo projeto, posteriormente, deu-se início à etapa e preparação do ambiente, com a disposição dos dados de beneficiários categorizados como propensos à DM Tipo 2, segundo seu município de residência. Por fim, iniciou-se a integração das informações de interesse obtidas com o resultado do questionário em camadas georreferenciadas, com a construção de um painel para visualização dessas informações geograficamente. Na fase de implantação, tem-se como atividade a homologação da ferramenta, com a apresentação dos resultados obtidos, com o mapeamento das camadas de informações, onde conforme a intensidade da cor concentra o maior número de beneficiários propensos à DM Tipo 2.

Análise dos resultados e estatística

Os dados foram organizados em tabelas, para desenvolvimento dos estudos e discutidos segundo a literatura disponível acerca do tema. Para o desenvolvimento das camadas georreferenciadas, utilizou-se o *software* de georreferenciamento Q-GIS. Os dados contidos na tabela foram transferidos a esse *software* por meio de uma junção com uma camada estadual com os 399 municípios do Paraná, obtida no site do IBGE. Os resultados dos beneficiários categorizados, por meio do reconhecimento de padrões utilizando mineração de dados, como DM Tipo 2 foram dispostos por município do Paraná. A prevalência de cada município foi calculada com base na população que possuía o plano de saúde e teve como numerador os casos de DM Tipo 2 e como denominador, o número de beneficiários nesse período, multiplicado por 1.000.

RESULTADOS

Fase Inicial – Referencial teórico e planejamento

Na atividade de planejamento, por meio de entrevistas e reuniões com os setores da organização, identificaram-se algumas oportunidades que poderiam ser úteis para compor a base de dados geográficos, entre elas, destacam-se as seguintes fontes: ANS⁽¹⁹⁾, MS (Ministério da Saúde)⁽²⁰⁾, SESAPR (Secretaria de Estado da Saúde do Paraná)⁽²¹⁾, PMC (Prefeitura Municipal de Curitiba)⁽²²⁾, Datasus (Departamento de Informática do SUS) – Tabnet (Informações de Saúde)⁽²³⁾, DW, Simepar (Sistema Meteorológico do Paraná)⁽²⁴⁾, Pesquisas de Satisfação e Utilização, entre outras. A intenção da escolha das bases de dados consiste em compor as informações referentes ao perfil epidemiológico, rede assistencial, avaliação de tecnologia em saúde, dados sobre clima, diretrizes de saúde e protocolos e sobre a qualidade de atendimento.

Na aplicação do questionário, obteve-se o seguinte resultado: na área da saúde, 100% dos entrevistados pontuaram como 5 (muita importância). Na área de mercado, 80% dos entrevistados pontuaram

como 5 (muita importância), 10% como média importância e 10% não responderam. Na área de estratégia, 80% dos entrevistados pontuaram como 5 (muita importância), 20% não responderam.

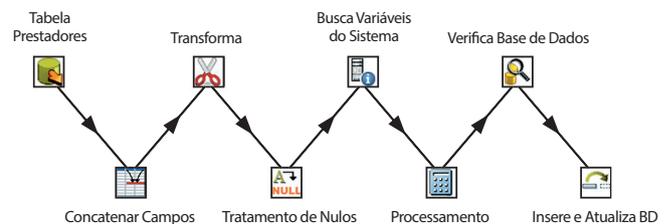
Nas questões abertas, obteve-se o seguinte resultado: na área da saúde, foram citados como fontes relevantes de aplicação ao estudo: conhecimento do local de moradia, saneamento básico, acesso a bens e serviços, rede de telefonia celular, indicadores de cobertura vacinal, pesquisas de cobertura parcial temporária com beneficiários, aplicativos de saúde, taxas de mortalidade, morbidade e natalidade. Na área de mercado foram citados: densidade demográfica dos municípios, mapeamento de indústrias e empresas, população dos municípios, população de beneficiários e associações comerciais. Na área estratégica foram citados: conhecimento de maior circulação de pessoas, monitoramento de um grupo controle no estado.

Como resultante da aplicação do questionário, obteve-se o direcionamento da atenção do projeto na área da saúde e a adição de novas fontes de informação a serem incorporadas na base de dados georreferencial.

Fase de desenvolvimento

Por meio do método de identificação de padrões das doenças crônicas utilizando a técnica de mineração de dados – tarefa de classificação⁽⁸⁾, foi possível categorizar as doenças crônicas: DM tipo 2, Neoplasias, Doenças Pulmonares, Cerebrovasculares, Hipertensão, Obesidade, Psiquiátricas e Isquêmicas do Coração.

Dessa maneira, selecionou-se um conjunto de registros de interesse com doenças crônicas, sem a identificação do indivíduo, analisando apenas as informações de localizações geográficas. Também foram implementadas as informações sobre a rede assistencial (figura 1); para o processo de integração das informações da rede utilizou-se uma ferramenta de ETL (extração, transformação e carga de dados), o Pentaho, que permite de maneira intuitiva defini-lo de forma gráfica, possibilitando, dessa maneira, documentar todo o ambiente⁽¹⁸⁾.



Fonte: Elaborado pelos autores.

Figura 1 – Exemplo de um processo de extração, transformação e carga de dados para o banco de dados geográfico.

Com as informações das camadas da rede assistencial, sobre identificação de doenças crônicas, associadas às informações externas, integram-se as camadas em uma base de dados geográfica para a construção de uma ferramenta de visualização.

A partir da base formada, criou-se, por meio de uma ferramenta de sistema de informações geográfica – Quantum GIS⁽¹⁷⁾, *software* livre com licença geral pública (GNU – *General Public Licence*) – diversas camadas para análise geográfica.

Fase de implantação

Por meio da aplicação das regras de categorização para a identificação dos beneficiários com DM Tipo 2, obteve-se uma amostra de 18.013 indivíduos. O perfil epidemiológico desses indivíduos está exposto na Tabela 1. Entre os indivíduos da amostra, 10.495 (58,3%) são do sexo feminino e 7.518 (41,7%) do sexo masculino. A DM Tipo 2 é predominante nas faixas etárias de 60 a 69 anos, com 3.910 casos (21,7%), seguido de 50 a 59 anos, com 3.484 casos (19,3%), 40 a 49 anos, com 2.742 casos (15,2%), 30 a 39 anos, com 2.553 casos (14,2%) e 70 a 79 anos, com 2.494 casos (13,8%). Cabe ressaltar que a essas cinco faixas etárias acumulam 84,3% dos casos. As demais faixas etárias analisadas (80 anos ou mais e abaixo de 30 anos) representam 15,7% dos casos, onde 8,5% representa idade inferior a 30 anos, situações de valores na base de dados que reconheceram possíveis casos de outros tipos de DM. Um fator de risco importante para a diabetes é a obesidade; na população estudada, 9,4% dos casos (1.702) apresentaram atendimentos com o CID de obesidade em seu registro de saúde, podendo ser os CIDs E66 - Obesidade, E66.0 - Obesidade devida a excesso de calorias, E66.1 - Obesidade induzida por drogas, E66.2 - Obesidade extrema com hipoventilação alveolar, E66.8 - Outra obesidade e E66.9 - Obesidade não especificada.

Tabela 1 – Perfil epidemiológico e associação com registro de obesidade dos beneficiários com Diabetes Mellitus Tipo 2 da operadora de planos de saúde, Paraná, Brasil, 2017.

| Variáveis | n (N= 18.013) | % |
|---------------------|------------------|------|
| Sexo | | |
| Sexo Feminino | 10.495 | 58,3 |
| Sexo Masculino | 7.518 | 41,7 |
| Faixa Etária | | |
| 60 a 69 anos | 3.910 | 21,7 |
| 50 a 59 anos | 3.484 | 19,3 |
| 40 a 49 anos | 2.742 | 15,2 |
| 30 a 39 anos | 2.553 | 14,2 |
| 70 a 79 anos | 2.494 | 13,8 |
| 80 anos ou mais | 1.305 | 7,2 |
| Abaixo de 30 anos | 1.525 | 8,5 |
| Obesidade associada | | |
| Sim | 1.702 | 9,4 |
| Não | 16.311 | 90,6 |

Fonte: Elaborado pelos autores.

O mapeamento geográfico dos beneficiários analisados teve como base o estado do Paraná, principal área de atuação da operadora de planos de saúde. Os indivíduos da amostra foram dispostos conforme seu município de residência e comparado ao número total para buscar conhecer em quais municípios os indivíduos elegíveis se concentram. Os resultados podem ser analisados por meio da tabela 2. Vinte e cinco municípios do estado do Paraná concentram 80,65% dos beneficiários categorizados como DM Tipo 2 no ano de 2017. O restante, 374 municípios concentram 19,35% dos indivíduos com DM Tipo 2. Nesta tabela, o N de beneficiários diminui, pois nem todos os beneficiários da OPS residem no estado do Paraná.

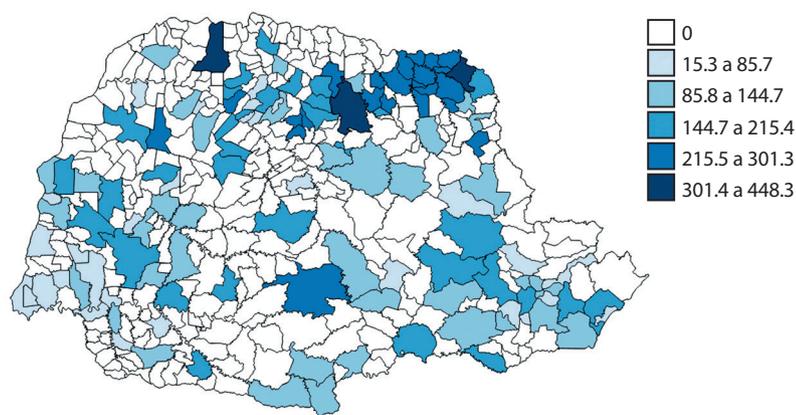
Tabela 2 – Beneficiários da operadora de planos de saúde categorizados como Diabetes Mellitus Tipo 2, Paraná, Brasil, 2017, segundo município de residência.

| Município | n (N= 15.572) | % | Acumulado % |
|---------------------------------------|------------------|-------|----------------|
| Curitiba | 5850 | 37,57 | 37,57 |
| Londrina | 1663 | 10,68 | 48,25 |
| Maringá | 744 | 4,78 | 53,02 |
| Cascavel | 626 | 4,02 | 57,04 |
| Ponta Grossa | 346 | 2,22 | 59,27 |
| Guarapuava | 286 | 1,84 | 61,10 |
| Foz do Iguaçu | 265 | 1,70 | 62,81 |
| São José dos Pinhais | 230 | 1,48 | 64,28 |
| União da Vitória | 213 | 1,37 | 65,65 |
| Pinhais | 200 | 1,28 | 66,93 |
| Bandeirantes | 183 | 1,18 | 68,11 |
| Paranavaí | 177 | 1,14 | 69,25 |
| Apucarana | 173 | 1,11 | 70,36 |
| Jacarezinho | 165 | 1,06 | 71,42 |
| Paranaguá | 153 | 0,98 | 72,40 |
| Campo Mourão | 148 | 0,95 | 73,35 |
| Assaí | 147 | 0,94 | 74,29 |
| Toledo | 140 | 0,90 | 75,19 |
| Umuarama | 139 | 0,89 | 76,09 |
| Cornélio Procopio | 137 | 0,88 | 76,97 |
| Ubiratã | 123 | 0,79 | 77,75 |
| Cambará | 121 | 0,78 | 78,53 |
| Uraí | 114 | 0,73 | 79,26 |
| Andirá | 108 | 0,69 | 79,96 |
| Arapongas | 108 | 0,69 | 80,65 |
| Demais municípios do estado do Paraná | 270 | 19,35 | 100 |

Fonte: Elaborado pelos autores.

Buscando atender o objetivo do estudo e proporcionar diferentes percepções de visualização, os indivíduos da amostra foram dispostos de forma georreferenciada com a utilização do *software* elencado. Na Figura 1 é possível observar que as regiões norte pioneiro, norte central e centro sul concentram a maior porcentagem dos indivíduos da amostra. A análise georreferenciada permite a visualização geoespacial dos casos, tornando possível saber em quais regiões do estado deve-se investir em ações de prevenção e promoção da saúde.

Na figura 2, demonstra-se um exemplo da aplicação do modelo de visualização, com o mapeamento dos beneficiários da OPS categorizados com DM Tipo 2, que conforme a intensidade das cores representa a maior incidência da doença por município. A taxa variou de 15,38/1.000 a 448,37/1.000 beneficiários. Os municípios que apresentaram a maior incidência foram: Jacarezinho (448,37/1.000 beneficiários), Londrina (355,42/1.000 beneficiários), Paranavaí (340,38/1.000 beneficiários) e Guarapuava (301,57/1.000 beneficiários). A pesquisa foi realizada no banco de dados próprio, com os casos categorizados com DM Tipo 2. A incidência foi calculada com base no número de beneficiários da OPS, residentes em cada município, e teve como numerador os casos de DM Tipo 2 e como denominador, o número de beneficiários no mesmo período, multiplicado por 1.000. Foram excluídos os dados dos municípios do estado com uma amostra menor que 50.



Fonte: Elaborado pelos autores.

Figura 2 – Distribuição geográfica de casos (incidência) de beneficiários da operadora de planos de saúde categorizados com Diabetes Mellitus Tipo 2, Paraná, Brasil, 2017, segundo município de residência.

DISCUSSÃO

Uma informação obtida pela identificação de padrões, com o uso da técnica de mineração de dados e apresentados geograficamente potencializa a descoberta de novos conhecimentos em base de dados e dessa maneira oportuniza ações dinâmicas e ágeis da área de gestão da saúde.

Cita-se como essa prática, o exemplo do mapeamento geográfico demonstrado na figura 2, em que identifica-se a necessidade do desenvolvimento de ações nas regiões norte e norte pioneiro do Paraná em relação à DM Tipo 2, devido a sua alta prevalência. Em outro estudo, percebeu-se a preocupação em relação à alta prevalência de casos de hepatite B na região sudoeste do Paraná. Como ação a partir dessa descoberta, em conjunto com a área de saúde da OPS, buscou-se identificar se haviam beneficiários residentes naquela região que não haviam se imunizado contra a Hepatite B. Por meio de um levantamento realizado no banco de dados, com posterior contato telefônico aos beneficiários identificados, verificou-se que grande parte desses beneficiários, cerca de 200, que residem naquela região não era imunizada contra Hepatite B. Com base nessa informação, o profissional da área de gestão de saúde via tele monitoramento, orientou o beneficiário quanto à importância da imunização, agindo de maneira proativa na educação em saúde. Muitos beneficiários desconheciam sobre o cenário da doença naquela região e também de que a vacina é disponibilizada para todas as faixas etárias, sem custo nas unidades de saúde de sua cidade.

Os resultados da aplicação desse modelo em distribuição geográfica, utilizando o método de mineração de dados, se apresentou eficiente para a elegibilidade de beneficiários que potencialmente poderiam evoluir para doenças crônicas. Conforme a referência⁽⁶⁾, foram indicados 5.953 beneficiários em 2011 elegíveis para o programa de diabéticos, representando 5,7% do total de beneficiários da carteira. Este resultado é compatível com a Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (Vigitel) de 2012, que apresentava 5,6% de adultos (≥ 18 anos) com diagnóstico médico referido para o Diabetes⁽²⁵⁾.

Tratar preventivamente situações de alertas epidemiológicos e identificar, geograficamente, potenciais riscos que possam ser mitigados com a adoção de ações proativas é um papel fundamental

para a sustentabilidade das OPS, colaborando e agindo em parceria com as instituições públicas com o foco no cuidado da população.

Limitações do estudo

O estudo apresentou algumas limitações. Na integração de algumas bases públicas elencadas com o auxílio dos gestores da instituição, observou-se que algumas delas não haviam sido atualizadas. Algumas bases apresentavam dados até o ano de 2012, que refletem à questão da qualidade do banco de dados.

Contribuições para a área da enfermagem e saúde

O estudo trouxe contribuições para a área da enfermagem e da saúde pública, pois a utilização de ferramentas de georreferenciamento é uma necessidade atual dos serviços de saúde. Destaca-se que a pesquisa presou pela utilização de uma ferramenta de *software* livre, sem custo de aquisição e manutenção, oportunidade que torna viável o acesso de qualquer serviço de saúde.

CONCLUSÃO

O modelo proposto e descrito nesse trabalho, baseado em ferramenta de georreferenciamento aplicada à saúde, demonstrou-se eficiente e poderá ser útil para a descoberta de padrões a níveis regionais e que, por relação de causa e efeito, contribuirá para a formulação e identificação de situações passíveis de atuação na prevenção e promoção à saúde.

A identificação geográfica dos beneficiários categorizados como propensos às doenças crônicas, utilizando metodologias de reconhecimento de padrões em base de dados agregando as informações de ocorrências dos serviços, por meio dos registros de contas médicas e os pedidos de liberação da OPS, de alguma maneira, será possível identificar situações de alerta, promovendo antecipadamente ações dirigidas de prevenção e promoção à saúde, que estão vinculadas ao principal objetivo da metodologia proposta.

Produtos derivados dessa metodologia serão úteis para o gerenciamento dos serviços da OPS, como a gestão de rede, a avaliação de tecnologias em saúde, a análise de especialidades médicas, entre outros, possibilitando que os gestores utilizem informações ágeis e oportunas para o processo de apoio à tomada de decisões.

Em trabalho futuro pretende-se, ainda, registrar os alertas e situações identificadas pelo ambiente para avaliar os resultados que a ferramenta proporcionou em termos de otimização dos custos para a OPS e como benefício para os seus clientes.

FOMENTO

Os autores agradecem o apoio e financiamento, dessa pesquisa, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pela concessão da bolsa de estudo.

REFERÊNCIAS

1. Almeida Filho N, Rouquayrol MZ. Introdução à Epidemiologia. 4th ed. Rio de Janeiro: Guanabara Koogan; 2006.
2. Ashton-Prolla P, Vargas FR. Prevalence and impact of founder mutations in hereditary breast cancer in Latin America. *Genetics Molecular Biol*[Internet]. 2014[cited 2017 Feb 03];37(1):234-40. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3983579/pdf/gmb-37-234.pdf>
3. Banco Mundial. Banco Internacional para a Reconstrução e o Desenvolvimento. Envelhecendo em um Brasil mais velho: implicações do envelhecimento populacional para o crescimento econômico, a redução da pobreza, as finanças públicas e a prestação de serviços. Washington: World Bank; 2011.
4. Instituto de Estudos de Saúde Suplementar (BR). Envelhecimento populacional e os desafios para o sistema de saúde brasileiro. São Paulo: IESS; 2013.
5. Instituto de Estudos de Saúde Suplementar (BR). Variação de Custos Médico-Hospitalares[Internet]. 2016[cited 2017 Dec 12]. Available from: https://www.iess.org.br/cms/rep/VCMH_set17.pdf Portuguese.
6. Instituto Paranaense de Desenvolvimento Econômico e Social. Índice nacional de preços ao consumidor (INPC) e índice nacional de preços ao consumidor amplo (IPCA) – Brasil – janeiro 1994 – novembro 2017[Internet]. 2017[cited 2017 Dec 12]. Available from: http://www.ipardes.gov.br/pdf/indices/inpc_ipca.pdf Portuguese.
7. Carneiro LAF. Planos de saúde: aspectos jurídicos econômicos. Rio de Janeiro: Forense; 2012.
8. Carvalho DR, Dallagassa MR, Silva SH. Uso de técnicas de mineração de dados para a identificação automática de beneficiários propensos ao diabetes mellitus tipo 2.[Internet]. 2015[cited 2017 Feb 02];20(3):274-96. Available from: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/16018/17648> Portuguese.
9. Conselho Federal de Medicina (CFM). Resolução nº 1.819, de 22 de maio de 2007. Proíbe a colocação do diagnóstico codificado (CID) ou tempo de doença no preenchimento das guias da TISS de consulta e solicitação de exames de seguradoras e operadoras de planos de saúde concomitantemente com a identificação do paciente e dá outras providências. Diário Oficial da União 22 maio 2007; Seção 1.
10. Bellazzi R. Big Data and Biomedical Informatics: a challenging opportunity. *Yearb Med Inform*[Internet]. 2014[cited 2017 Feb 05];9(1):8-13. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4287065/pdf/yimi-09-0008.pdf>
11. Fayyad U, Piatetsky-Shapiro G, Smyth P. From data mining to knowledge discovery in databases. *AI Mag*[Internet]. 1996[cited 2017 Feb 05];17(3):37-54. Available from: <https://www.aaai.org/ojs/index.php/aimagazine/article/view/1230/1131>
12. Casters M, Bouman R, Van Dongen J. Pentaho Kettle solutions: building open source ETL solutions with Pentaho Data Integration. Hoboken: John Wiley & Sons; 2010.
13. Müller EPL, Cubas MR, Bastos LC. Geoprocessing of data as a management tool in a family health unit. *Rev Bras Enferm* [Internet]. 2010[cited 2017 Feb 05];63(6):978-82. Available from: <http://www.scielo.br/pdf/reben/v63n6/17.pdf>
14. Concaro S, Sacchi L, Cerra C, Bellazzi R. Mining administrative and clinical diabetes data with temporal association rules. *Stud Health Technol Informatics*[Internet]. 2009[cited 2017 Feb 05];150:574-8. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/19745376>
15. Bellazzi R, Sacchi L, Concaro S. Methods and Tools for Mining Multivariate Temporal Data in Clinical and Biomedical Applications, 31st Annual International Conference of the IEEE EMBS. Minnesota 2009:5629-32.
16. Ministério da Saúde (BR). Vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico: estimativas sobre frequência e distribuição sócio demográfica de fatores de risco e proteção para doenças crônicas nas capitais dos 26 estados brasileiros e no distrito federal em 2016. Brasília: Ministério da Saúde; 2017.
17. QGIS. A Free and Open Source Geographic Information System [Internet]. 2016[cited 2016 Jun 05] Available from: <https://www.qgis.org/en/site/index.html>
18. Pentaho. Data Integration, Business Analytics and Big Data[Internet]. c2016[cited 2016 Jun 05]. Available from: <https://www.hitachivantara.com/en-us/products/big-data-integration-analytics.html>
19. Agência Nacional de Saúde Suplementar (ANS) [Internet]. Rio de Janeiro: Agência Nacional de Saúde Suplementar (ANS), c2018 [cited 2019 Mar 17]. Available from: <http://www.ans.gov.br>
20. Ministério da Saúde (BR) [Internet]. c2013-2018[cited 2016 Sep 09]. Available from: <http://www.portalms.saude.gov.br/>
21. Secretaria da Saúde (PR) [Internet]. Curitiba (PR): Secretaria da Saúde; c2019 [cited 2019 Mar 17]. Available from: <http://www.saude.pr.gov.br>
22. Secretaria Municipal da Saúde (PR) [Internet]. Curitiba (PR): Secretaria Municipal da Saúde; c2019 [cited 2019 Mar 17]. Available from: <http://www.saude.curitiba.pr.gov.br>
23. Ministério da Saúde (BR). Departamento de Informática do SUS [Internet]. 2016[cited 2016 Sep 09]. Available from: <http://datasus.saude.gov.br/>
24. SIMPAR: Sistema Meteorológico do Paraná [Internet]. Curitiba (PR): SIMPAR; c2019 [cited 2019 Mar 17]. Available from: <http://www.simepar.br/prognozweb/simepar/home>
25. Ministério da Saúde (BR). Vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico: Estimativas sobre frequência e distribuição sócio demográfica de fatores de risco e proteção para doenças crônicas nas capitais dos 26 estados brasileiros e no distrito federal em 2012. Brasília: Ministério da Saúde; 2013.