

Projeto MODAU (modern autopsy techniques to investigate human diseases): Data Science na prática - Motivação do projeto, fontes de informação e constituição da equipe

Carmen Diva Saldiva de André ¹

Os objetivos gerais do projeto MODAU são:

- Explorar o conceito de autópsia minimamente invasiva para entender processos de doenças, avaliar o diagnóstico clínico e prescrever tratamento;
- Aprimorar das medidas de exposição à poluição atmosférica na cidade de São Paulo;
- Avaliar do papel da autópsia como ferramenta de investigação dos efeitos do ambiente urbano na determinação de doenças humanas.

O projeto é dividido em dois grandes subprojetos: BIAS e METROHEALTH

O diagnóstico post-mortem é essencial para o estabelecimento de políticas públicas de saúde que possam evitar a morte evitável. O projeto BIAS (Brazilian Image-Auopsy Study) tem como objetivo avaliar o papel de diferentes ferramentas de imagem no diagnóstico post-mortem, por meio de autópsias pouco invasivas. Especificamente, compara-se autópsia convencional (AC) com autópsia minimamente invasiva (AMI), incluindo tomografia computadorizada, angiotomografia e biópsia guiada por imagem (AMI-CT) em sujeitos mortos por causas naturais. Apesar de serem ferramentas poderosas para o apoio à autópsia, a implementação dessas técnicas pode ser custosa, especialmente para centros mais remotos onde falta expertise ou com número muito restrito de casos, não havendo compensação financeira que garanta sustentabilidade. Pensando nestas limitações, foram também consideradas duas alternativas de menor custo para autópsia minimamente invasiva: AMI-US com protocolo semelhante ao descrito para o AMI-CT, com ultrassonografia substituindo angiotomografia e autópsia verbal (AV). O questionário AV tem sido cada vez mais usado para medir causas de morte por óbitos não hospitalares ou não certificadas ou para substituir certificados de óbito de baixa qualidade, principalmente em países de baixa e média renda. Os desempenhos de dois procedimentos de atribuição de causa de óbito com base na AV estão sendo avaliados: a atribuição das causas de óbito por um clínico e o método automático Tariff, no qual a causa é determinada por meio de um aplicativo computacional. Ainda dentro do projeto BIAS, foi também criado um sistema virtual de lâminas (DigiDeath) que permite a integração de patologistas de diferentes locais e possibilita a sua participação no projeto. Esta plataforma será também utilizada com a finalidade de ensino na área de Patologia.

¹ Instituto de Matemática e Estatística da USP.

Todos os casos AMI-US têm questionário AV, possibilitando a comparação de seus desempenhos na atribuição da causa básica de morte. Todos os casos com questionário AV preenchido foram geocodificados. A alguns casos foi também aplicada a escala de avaliação clínica da demência CDR (Clinical Dementia Rating).

O projeto METROHEALTH está relacionado à avaliação de medidas de exposição à poluição atmosférica na cidade de São Paulo e consta dos seguintes subprojetos:

- Obtenção das concentrações de elementos químicos em cascas de árvores para identificar fontes e avaliar a variabilidade espacial da poluição atmosférica em um cenário urbano e reconstrução temporal do clima e da poluição atmosférica baseada nos registros de anéis de crescimento de árvores urbanas.
- Utilização de dados de antracose como indicador de exposição (Biomonitoramento baseado em pulmão).
- Doença de Alzheimer: correlação com o meio ambiente por meio da análise do epitélio olfatório;
- Modelagem da qualidade do ar: aprimoramento da descrição das variações espaciais e temporais de emissões veiculares com a utilização de dados de fatores de emissão, classificação de vias urbanas, congestionamentos e melhoria do modelo de qualidade do ar com previsões de concentração de poluentes.
- Construção de um modelo espaço-temporal para a predição da concentração de material particulado fino na cidade de São Paulo.
- Desenvolvimento de um monitor passivo químico para monitoramento ambiental de gases, ozônio e dióxido de nitrogênio;
- A relação entre a proporção etanol/gasolina no combustível da frota leve na cidade de São Paulo e a poluição atmosférica;
- Desenvolvimento de um índice sócio-econômico que, além da renda, inclua detalhes (como presença ou ausência da poluição atmosférica local) sobre o entorno em que os habitantes vivem e trabalham.
- Identificação de agrupamentos espaciais de recém-nascidos pré-termo não gemelares e não indígenas com o objetivo de definir agrupamentos adequados para a realização de estudos caso-controle;
- Análise de riscos relativos de mortalidade por doenças do aparelho circulatório considerando apenas indicadores socioeconômicos do contexto do lugar de residência das pessoas.

Os subprojetos do MODAU geram uma grande quantidade de dados que estão sendo integrados ao sistema SISAUT, desenvolvido com o objetivo de armazenar e integrar dados de óbitos, de coleta e processamento de materiais biológicos e de questionários. Estão sendo

também integrados ao sistema concentrações de poluentes obtidas a partir da rede de monitoramento da CETESB, variáveis meteorológicas (CETESB e IAG) e dados de tráfego (CET).

O projeto MODAU é um exemplo típico de Ciência de Dados pois requer o planejamento dos diferentes subprojetos, a coleta de dados, a construção de um banco de dados, a análise desses dados e a divulgação dos resultados. Com essa finalidade precisa a colaboração interativa de profissionais de diferentes áreas (Medicina, Estatística, Computação, Geografia, Engenharia, Biologia, Meteorologia, Enfermagem, Psicologia e Química).

Projeto MODAU (Modern Autopsy Techniques to Investigate Human Diseases):**Data Science na prática****A construção do banco de dados e Sistema de aquisição de informações e interface com o Banco de Dados****Paulo Afonso de André¹**

O SVOC – Serviço de Verificação de Óbitos da Capital (SP), instalado na Faculdade de Medicina da USP, é o maior serviço de autópsia do Brasil e sedia o Projeto MODAU, no qual são avaliados métodos minimamente invasivos de autópsia e a associação de doenças com o meio ambiente. Dispõe de uma completa plataforma de imagens da sala de autópsia – PISA, o que permite complementar as autópsias convencionais com outros exames de imagem e coletas de materiais.

O SISAUT (Sistema transacional de dados de autópsia e material biológico) está sendo desenvolvido para gerenciar e controlar os fluxos de trabalho dentro do SVO e para seguir e monitorar os diversos subprojetos do MODAU. Dispõe de um módulo para capturar do sistema do SVO todas as informações de interesse aos subprojetos, em tempo real, o que evita erros e otimiza o desempenho do SISAUT.

Uma extensão foi desenvolvida para incluir a aplicação eletrônica dos questionários AV – Autópsia Verbal e CDR – Clinical Dementia Rating, no ambiente LimeSurvey, próprio para essa aplicação.

Conta ainda com a plataforma DigiDeath, desenvolvida para armazenar informações dos questionários e aquelas oriundas do SVO, já disponíveis no SISAUT, como também imagens de lâminas dos tecidos coletados nas autópsias. Essas imagens poderão ser acessadas remotamente via internet padrão (menor velocidade). As equipes médicas poderão acessar essas imagens como se estivessem fazendo uso de microscópio comum, com grande aplicação de zoom e capacidade de varredura de toda a lâmina com máxima resolução. Nessas condições, poderão imputar no DigiDeath os laudos de causa de morte e sua codificação em padrões internacionais (CID10), para cada um dos processos de autópsia: convencional, alternativos e automatizados, que serão transmitidos em tempo real para o SISAUT.

Outras imagens, que não carecem do uso das funções da microscopia, também serão armazenadas e estarão acessíveis no SISAUT, sem o acesso ao DigiDeath, como as do projeto da Antracose, que requer o registro simples da imagem dos tecidos.

Para a alimentação de outras bases com grande quantidade de dados foi criado o sistema auxiliar ModauData, que facilita a integração de dados meteorológicos, poluentes ambientais, análises elementares em cascas de árvores, e outras fontes de dados armazenados em planilhas eletrônicas.

¹ *Pesquisador Colaborador, Faculdade de Medicina da USP (Paulo.afonso@fm.usp.br/11-98326-4843)*

A abordagem espacial do Projeto MODAU

Ligia Vizeu Barrozo¹

O componente espacial do Projeto MODAU é central para a integração dos dados das diversas fontes. Tal integração só é possível mediante a geocodificação dos endereços dos casos autopsiados ou do georreferenciamento dos atributos estudados, tais como características físicas (uso do solo, malha viária, bioindicadores, estações de monitoramento da qualidade do ar etc.) e socioeconômicas, a partir da malha espacial dos setores censitários do Censo Demográfico. Como todas as bases estão integradas espacialmente, associações entre os desfechos de interesse e as condições do ambiente urbano podem ser avaliadas. Exemplo desta abordagem permitiu constatar que a exposição ao tráfego veicular é uma fonte significativa de ingestão de poluição do ar (TAKANO et al., 2019). Para esta análise foram considerados o índice de antracose pulmonar obtido dos casos autopsiados, a exposição ao tabaco e o tempo de exposição ao tráfego veicular, obtidos a partir dos questionários aplicados aos familiares do falecido e condições socioeconômicas e proximidade das vias de maior intensidade de tráfego, obtidas a partir de análise espacial do endereço de residência do indivíduo. As associações foram estimadas a partir de modelos de regressão múltipla.

Outro exemplo de aplicação deste banco de dados foi o estudo caso-controle espacial que permitiu verificar a associação entre o risco relativo de nascimento de bebês prematuros e poluição do ar como fator de risco independente para a prematuridade (SALDIVA et al., 2018).

Referências

SALDIVA, S.; BARROZO, L.; LEONE, C.; FAILLA, M.; BONILHA, E.; BERNAL, R.; OLIVEIRA, R.; SALDIVA, P. Small-Scale Variations in Urban Air Pollution Levels Are Significantly Associated with Premature Births: A Case Study in São Paulo, Brazil. **International Journal of Environmental Research and Public Health**, v. 15, n. 10, p. 2236, 12 out. 2018. Disponível em: <<http://www.mdpi.com/1660-4601/15/10/2236>>.

TAKANO, A. P. C.; JUSTO, L. T.; DOS SANTOS, N. V.; MARQUEZINI, M. V.; DE ANDRÉ, P. A.; DA ROCHA, F. M. M.; PASQUALUCCI, C. A.; BARROZO, L. V.; SINGER, J. M.; DE ANDRÉ, C. D. S.; SALDIVA, P. H. N.; VERAS, M. M. Pleural anthracosis as an indicator of lifetime exposure to urban air pollution: An autopsy-based study in Sao Paulo. **Environmental Research**, v. 173, p. 23–32, jun. 2019. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0013935119301343>>.

¹ Departamento de Geografia, FFLCH e Instituto de Estudos Avançados da Universidade de São Paulo; e-mail: lija@usp.br

A construção do sistema de autópsia verbal

Lúcia Pereira Barroso¹

A autópsia verbal (AV) consiste em atribuir causa de morte com base em questionário respondido por familiares sobre sintomas e doenças sofridas pelo falecido. O questionário da AV tem sido cada vez mais usado em caso de óbitos não hospitalares ou não certificados, principalmente em países de baixa e média renda.

O objetivo deste subprojeto é avaliar a eficácia do instrumento de AV, versão reduzida do questionário desenvolvido pelo PHMRC - Population Health Metrics Research Consortium e adaptada para o Brasil, para a determinação da causa de morte tendo a autópsia estendida como padrão ouro. O instrumento contém questões fechadas e relato aberto feito pelos familiares.

Este subprojeto do MODAU vem sendo desenvolvido em parceria com o Ministério da Saúde, Universidade Federal de Minas Gerais, University of Melbourne e Institute for Health Metrics and Evaluation - University of Washington. A causa básica de morte é atribuída de 3 formas: autópsia com microscopia e macroscopia mais questões fechadas mais relato aberto (padrão ouro), por médico clínico - Physician Certified Verbal Autopsy (PCVA) e pelo método automático Tariff por meio do aplicativo computacional SmartVA, desenvolvido pelo PHMRC.

A eficácia do instrumento da AV será verificada com relação à determinação de 22 causas de óbito e será feita por análise de concordância das causas individuais atribuídas e por comparação da distribuição das causas na população.

¹ Universidade de São Paulo, Instituto de Matemática e Estatística, lbarroso@ime.usp.br.

A relação entre a proporção de etanol/gasolina e poluição atmosférica

William Nilson de Amorim¹

Nos últimos anos, a área conhecida como Ciência de Dados vem exigindo de estatísticos e não estatísticos habilidades que vão muito além de modelagem, começando na obtenção e estruturação das bases de dados, passando pela manipulação e visualização dos dados e terminando na divulgação dos resultados em forma de relatórios, *painéis* ou publicações na internet. Dentro dela, uma abordagem chamada de aprendizado automático (*machine learning*) reuniu diversas técnicas e estratégias para modelagem preditiva, que, com alguns cuidados, podem ser aplicadas também para inferência. Essas novas visões da Estatística foram pouco absorvidas pela comunidade científica até então, principalmente pela ausência de estatísticos em grande parte dos estudos. Como exemplo de aplicação dessas novas estratégias de análise, vamos investigar a complexa associação entre a concentração troposférica de ozônio e a proporção de carros a gasolina na cidade de São Paulo.

Referencias

Alberto Salvo e Franz M. Geiger. Reduction in local ozone levels in urban São Paulo due to a shift from ethanol to gasoline use. *Nature Geoscience*, 7: 450-458

Alberto Salvo, Joel Brito, Paulo Artaxo e Franz M. Geiger. Reduced ultrafine particle levels in São Paulo's atmosphere during shifts from gasoline to ethanol use. *Nature Communications*, 8: 1-14

Gareth James, Daniela Witten, Trevor Hastie e Robert Tibshirani. An Introduction to Statistical Learning. Springer Series in Statistics. Springer, New York.

Trevor Hastie, Robert Tibshirani e Jerome Friedman. The Elements of Statistical Learning. Springer Series in Statistics. Springer.

¹ Instituto de Matemática e Estatística da USP, william.amorim@usp.br

A utilização de cascas e anéis de árvores para avaliação da poluição atmosférica

Giuliano Maselli Locosselli¹

A poluição do ar é considerada um dos principais fatores de risco de morte em todos os gêneros e faixas de idade. Nos grandes centros urbanos, as concentrações dos poluentes atmosféricos chegam a ser até sete vezes maiores do que os níveis recomendados pela Organização Mundial da Saúde, afetando diretamente a saúde população. Porém, estes poluentes não possuem uma distribuição homogênea tanto sobre o espaço urbano, quanto no tempo, pois as concentrações dos poluentes variam de acordo com os tipos de fonte presentes e os históricos delas. Estudos que visam a compreensão dos efeitos da poluição sobre a saúde da população urbana necessitam de dados que representem esta heterogeneidade da poluição urbana. Redes de estações de monitoramento de qualidade do ar suprem parcialmente este requisito, mas em geral são insuficientes em número para representar a exposição real à poluição. Esta limitação é uma consequência dos custos altos relativos a aquisição de plataformas de coletas de dados de poluição e sua manutenção. Os estudos de biomonitoramento são uma alternativa viável às redes de monitoramento convencionais por permitir uma caracterização espacial mais fina da poluição nas cidades, além de permitir obter informações sobre tendências temporais de poluentes para períodos pré-instrumentais. As árvores são um exemplo de biomonitores presentes em grande parte das cidades, e cada uma delas tem o potencial de registrar a variabilidade da poluição local. Como um exemplo, a cidade de São Paulo possui cerca de 650 mil árvores viárias que são potenciais monitores da poluição na cidade. A variabilidade espacial da poluição pode ser obtida por meio da análise da composição química da porção externa da casca das árvores. A rugosidade e a porosidade das cascas das árvores retêm parte dos elementos químicos presentes na atmosfera do entorno, em geral metais, que representam a exposição em um determinado local. Estes elementos também podem ser absorvidos pelas plantas e depositados nos anéis de crescimento presentes no interior do tronco de cada árvore. Cada anel de crescimento das árvores representa um ano de vida, e as concentrações dos metais no anel de crescimento são um registro da exposição da árvore às concentrações destes metais no ambiente. A combinação destas duas abordagens permite a obtenção de dados inéditos para estimar a exposição da população à poluição atmosférica e os seus efeitos sobre a saúde.

¹ Instituto de Biociências USP, locosselli@yahoo.com.br.