

O futuro está chegando: perspectivas promissoras sobre o uso de machine learning no transplante renal

The future is coming: promising perspectives regarding the use of machine learning in renal transplantation

Autores

Pedro Guilherme Coelho Hannun¹ 
Luis Gustavo Modelli de Andrade¹

¹ Universidade Estadual Paulista, Departamento de Medicina Interna, São Paulo, SP, Brasil.

RESUMO

Introdução: A predição de resultados pós-transplante é clinicamente importante e envolve vários problemas. Os atuais modelos de previsão baseados em padrões estatísticos são muito complexos, difíceis de validar e não fornecem previsões precisas. Machine Learning, é uma técnica estatística que permite que o computador faça previsões futuras usando experiências anteriores, está começando a ser usada para resolver essas questões. No campo do transplante renal, o uso da previsão computacional foi relatado na predição de rejeição crônica de aloenxerto, função tardia do enxerto e sobrevida do enxerto. Este artigo descreve os princípios e etapas de machine learning para fazer uma previsão e realiza uma breve análise das aplicações mais recentes de seu uso na literatura. **Discussão:** Existem evidências convincentes de que as abordagens de machine learning baseadas nos dados do doador e do receptor são melhores para proporcionar melhor prognóstico dos resultados do enxerto do que a análise tradicional. As expectativas imediatas que emergem dessa nova técnica de modelagem de previsão são que ela gerará melhores decisões clínicas baseadas em dados de práticas dinâmicas e locais e aperfeiçoará a alocação de órgãos, bem como o gerenciamento de cuidados pós-transplante. Apesar dos resultados promissores, ainda não há um número substancial de estudos para determinar a viabilidade de sua aplicação em um cenário clínico. **Conclusão:** A forma como lidamos com dados de armazenamento em prontuários eletrônicos de saúde mudará radicalmente nos próximos anos e a machine learning fará parte da rotina clínica diária, seja para prever resultados clínicos ou sugerir um diagnóstico baseado na experiência institucional.

Palavras-chave: Aprendizado de Máquina; Transplante de Rim; Modelos Estatísticos.

ABSTRACT

Introduction: The prediction of post transplantation outcomes is clinically important and involves several problems. The current prediction models based on standard statistics are very complex, difficult to validate and do not provide accurate prediction. Machine learning, a statistical technique that allows the computer to make future predictions using previous experiences, is beginning to be used in order to solve these issues. In the field of kidney transplantation, computational forecasting use has been reported in prediction of chronic allograft rejection, delayed graft function, and graft survival. This paper describes machine learning principles and steps to make a prediction and performs a brief analysis of the most recent applications of its application in literature. **Discussion:** There is compelling evidence that machine learning approaches based on donor and recipient data are better in providing improved prognosis of graft outcomes than traditional analysis. The immediate expectations that emerge from this new prediction modelling technique are that it will generate better clinical decisions based on dynamic and local practice data and optimize organ allocation as well as post transplantation care management. Despite the promising results, there is no substantial number of studies yet to determine feasibility of its application in a clinical setting. **Conclusion:** The way we deal with storage data in electronic health records will radically change in the coming years and machine learning will be part of clinical daily routine, whether to predict clinical outcomes or suggest diagnosis based on institutional experience.

Keywords: Machine Learning; Kidney Transplantation; Models, Statistical.

Data de submissão: 26/02/2018.
Data de aprovação: 05/08/2018.

Correspondência para:

Luis Gustavo Modelli de Andrade
E-mail: landrade@fmb.unesp.br

DOI: 10.1590/2175-8239-JBN-2018-0047



INTRODUÇÃO

Machine learning é uma técnica estatística que permite que um computador faça previsões futuras baseadas em experiências passadas. As estatísticas convencionais são combinações estruturadas com amostras baseadas em uma distribuição conhecida para julgar se as diferenças nos dados corroboram a existência de um efeito na população que elas representam. Ronald Fisher estabeleceu que uma probabilidade de menos de 5% é estatisticamente significativa ou, em outras palavras, que a amostra estudada difere da população ($p = 0,05$)¹. O trabalho inicial de Fisher foi baseado no cálculo do desvio padrão, que pressupõe que os dados têm distribuição normal. A distribuição normal é representada como uma curva em forma de sino, com a média no topo do sino e as “caudas” caindo para os lados. O desvio padrão é simplesmente a “média” do desvio absoluto de um valor a partir da média, e é uma medida de incerteza. Portanto, essa abordagem estatística formal considera que as amostras têm uma distribuição normal, também conhecida como modelo paramétrico. O próximo passo na estatística formal é fazer comparações usando uma distribuição conhecida (teste-*t* de *Student*, distribuição de Fisher ou qui-quadrado).

Machine learning, por outro lado, usa principalmente modelos não-lineares, que são menos dependentes do tipo de distribuição de dados. Alguns exemplos são árvores de decisão, rede neural artificial, algoritmo do vizinho-mais-próximo e máquinas de vetores de suporte.² As principais diferenças entre essas duas estratégias é que a estatística convencional busca encontrar diferenças entre as amostras, enquanto a machine learning visa criar modelos que possam prever eventos futuros. As técnicas de machine learning concentram-se na importância do preditor e da resposta direcionada, em vez de apenas a análise estatística para valores de *p*.

As etapas envolvidas na machine learning podem ser resumidas em aquisição de dados, limpeza, preparação, modelagem e implementação². Os primeiros passos envolvidos na aquisição de dados podem ser feitos automaticamente usando bancos de dados existentes. Para a segunda etapa, procedimentos de gerenciamento de dados como remoção de variáveis não relevantes, imputação de dados, falta de dados, remoção de outliers e seleção das variáveis mais importantes (preditores), são geralmente feitos na preparação para a etapa de modelagem. Na última etapa, vários modelos são projetados para prever a

resposta; aqueles que incorporam os melhores algoritmos de previsão são selecionados e seus resultados são combinados (modelo de ensemble). Geralmente, um modelo é construído com 70% de todos os dados disponíveis, o que é chamado de “treinamento”, e os 30% restantes são dados de teste (não usados na elaboração do modelo). Os dados de teste são usados para validar o modelo e evitar exageros de adaptação, definido como a produção de uma análise que corresponde muito de perto a um conjunto específico de dados.

APLICAÇÃO

Como muitos fatores influenciam o curso de uma doença, é difícil prever com precisão se o tratamento resultará em um ou mais desfechos usando padrões estatísticos formais. Novas abordagens, como o uso de técnicas de data mining, podem melhorar a precisão e acurácia na previsão de resultados de diferentes tipos de tratamento, com a consideração simultânea de vários fatores, bem como interações complexas que ocorrem entre eles. Machine learning, que é uma grande base técnica para data mining, fornece um método para extrair informações de dados brutos dentro de registros médicos.³

Bihorac *et al.* recentemente, em uma coorte de 51.457 pacientes, pesquisadores desenvolveram e validaram um algoritmo (MySurgeryRisk) para prever escores de risco probabilístico para 8 complicações pós-operatórias (lesão renal aguda, tromboembolismo venoso, internação em terapia intensiva superior a 48 horas, ventilação mecânica por 48 horas, ferida, complicações neurológicas e cardiovasculares, e óbito até 24 meses após a cirurgia) usando apenas dados clínicos armazenados em registros eletrônicos de saúde em instituições.⁴ O desenvolvimento do algoritmo consistiu em dois estágios classificados como transformador de dados e análise de dados. A etapa transformadora de dados integrou os dados disponíveis (variáveis demográficas, pré-operatórias, socioeconômicas, administrativas, médicas, farmacêuticas e laboratoriais) para que o procedimento de pré-processamento dos dados e a seleção do modelo pudessem ser realizados posteriormente para aperfeiçoar a análise dos dados. A etapa de análise de dados utilizou vários algoritmos computacionais para calcular probabilidades de risco para pós-operatório e mortalidade para um paciente individual. Finalmente, o MySurgeryRisk foi submetido à validação cruzada, ou seja, um modelo foi testado em 250 coortes diferentes e aleatoriamente

selecionou pacientes dentro da amostra (um total de 10.291 pacientes em cada coorte de validação). Os resultados mostraram um poder preditivo para as oito complicações de 0,82 a 0,94 (intervalo de confiança de 99% [IC], 0,81 a 0,94); no caso de óbito, o risco em diferentes tempos de pós-operatório variou de 0,77 e 0,83 (IC de 99%, 0,76 a 0,85).⁴

No transplante renal, estabelecer uma associação com desfechos clínicos é ainda mais difícil do que na prática clínica cotidiana, devido à alta complexidade do tratamento. Uma questão exaustivamente estudada é a previsão da sobrevida do enxerto a longo prazo.⁵⁻⁷

Nas últimas décadas, embora as melhorias na terapia imunossupressora tenham reduzido significativamente as taxas de rejeição aguda e aumentado a sobrevida a curto prazo do enxerto, benefícios substanciais na sobrevida a longo prazo não ocorreram. Causas especuladas incluem influência de novos perfis de doadores e receptores; entretanto, até o momento, fatores que contribuem para essa observação não são bem definidos, assim como a capacidade de fazer um prognóstico preciso e prever a duração do transplante.⁶⁻⁸ Os modelos atuais de sobrevida a longo prazo do enxerto são limitados por múltiplos fatores, incluindo: dependência de fatores pré-transplante, sem consideração de fatores imunológicos;⁶⁻¹⁰ amostras relativamente pequenas usadas para construir os modelos;⁸⁻¹⁰ e falha na utilização precisa de dados de pacientes censurados.⁶⁻¹⁰ Além disso, o tempo de observação dos modelos existentes é relativamente curto, enquanto que um longo período de observação seria essencial para a sobrevida do enxerto a longo prazo.^{6,7}

Nesse contexto, Yoo *et al.* avaliaram o uso de técnicas de machine learning para prever a sobrevida.¹¹ Aplicando métodos de data mining combinados com estatísticas de sobrevida, os pesquisadores construíram modelos preditivos de sobrevida do enxerto que incluíam fatores imunológicos, bem como variáveis conhecidas de receptores e doadores, seguindo um método semelhante à construção do algoritmo descrito anteriormente. Para esse propósito, eles usaram uma análise retrospectiva dos dados de um estudo de coorte multicêntrico envolvendo 3.117 pacientes, e analisaram o poder preditivo de um algoritmo de aprendizagem conjunta. Posteriormente, os resultados foram comparados com os dos modelos convencionais. A análise por modelo convencional revelou que a associação dos níveis séricos de creatinina 3 meses após o transplante foi de 77,8% hada 0,7 concordância (um índice que mede quão bem o modelo discrimina entre

diferentes respostas, ou seja, entre a resposta esperada e a resposta observada). Curiosamente, o uso de uma árvore de decisão de sobrevida, outro modelo padrão de um método de data mining, aumentou a concordância da previsão em relação ao primeiro algoritmo (concordância de 0,8), incluindo a incidência de rejeição aguda no primeiro ano após o transplante.¹¹

Outras tentativas foram feitas para usar ferramentas de machine learning para data mining, a fim de prever a sobrevida do enxerto a longo prazo. Entre essas tentativas, uma boa correlação entre a sobrevida prevista do enxerto e a taxa de sobrevida observada em 10 anos, calculada a partir de dados de sobrevida no Sistema de Dados Renais dos Estados Unidos, foi identificada usando a modelagem de árvore de decisão.⁶ Em outro estudo de modelagem que utilizou dados de 1.542 receptores de transplante renal do Registro de Diálise e Transplante da Austrália e Nova Zelândia, o sucesso ou fracasso de um transplante foi previsto com uma precisão de 85% usando redes neurais artificiais.⁹ Os classificadores bayesianos foram capazes de prever o sucesso ou insucesso do transplante com precisão de 97%. No entanto, a precisão da predição da duração da sobrevida do enxerto a longo prazo foi menor em 68% dos estudos.¹⁰

DISCUSSÃO

O consenso da maioria dos autores citados nesses estudos anteriores é semelhante: apesar dos vários achados em grupos de pacientes altamente representativos, mais pesquisas são necessárias para validar externamente essa abordagem, determinar a viabilidade de sua aplicação em um cenário clínico e avaliar se seu uso poderia levar a melhores resultados quando comparados às práticas atuais. Ainda assim, tais estudos consideram que os métodos de machine learning podem fornecer ferramentas flexíveis e viáveis para prever resultados envolvendo múltiplas variáveis.

Indubitavelmente, existem evidências sobre o benefício derivado dessas técnicas computacionais no campo da saúde, mas como essa evidência pode ser implementada na rotina médica e, mais especificamente, no manejo de receptores de transplante renal, ainda não está claro. Além disso, quais áreas devem esses novos estudos abordar? As respostas a essas perguntas incluem a geração de decisões clínicas baseadas em dados dinâmicos e de práticas locais, bem como a otimização da alocação de órgãos e cuidados pós-transplante.

Atualmente, muitas decisões clínicas são frequentemente baseadas na experiência e na intuição do médico.⁵ Essas práticas comuns levaram a erros e custos médicos excessivos que afetam a qualidade do serviço prestado aos pacientes. Os métodos de machine learning poderiam ajudar a gerar padrões de evolução baseados em grandes conjuntos de dados locais e multicêntricos, o que seria muito útil para melhorar a qualidade das decisões clínicas. Hoje, uma riqueza de dados está prontamente disponível em prontuários eletrônicos hospitalares e grandes bases de dados nacionais, como a mantida pela Rede Unida para Compartilhamento de Órgãos (UNOS), que é notavelmente completa e atualmente não está sendo usada para beneficiar pacientes.⁵ Esse tipo de análise deve fornecer cuidados direcionados a determinados tipos ou subpopulações de pacientes com risco aumentado de perda ou morte do enxerto. Esta aplicação foi mostrada em certa medida por Taber *et al.*, que, em uma análise de risco, os dados dinâmicos no nível do paciente melhoraram a precisão da previsão de reinternação dentro de 30 dias após o transplante renal.¹²

Uma possibilidade ainda melhor seria desenvolver um esquema de alocação de órgãos direcionado, que focaria o resultado pós-transplante como medida de desempenho.¹³ Todo o processo seria baseado na otimização do transplante para garantir que o transplante fosse realizado apenas em pacientes que teriam benefícios a longo prazo. Grande parte dessa necessidade decorre do fato de que, devido à escassez na doação de órgãos⁵⁻⁷ há uma demanda crescente para o desenvolvimento de procedimentos eficazes e eficientes para selecionar o receptor de órgão ideal e garantir a sobrevivência máxima possível. No futuro, uma nova ferramenta baseada nessas técnicas poderia ser projetada para auxiliar no complexo processo decisório usado para identificar bons candidatos a transplantes baseado nas características específicas dos rins disponíveis.

Diversas técnicas de modelagem preditiva poderiam ser empregadas para os componentes estatísticos da avaliação, como vetores de suporte de máquina, redes neurais artificiais, classificadores de Bayes e árvores de regressão,⁶⁻¹⁰ a fim de desenvolver modelos preditivos e extrair as variáveis mais úteis em uma análise de sensibilidade usando o melhor modelo de desempenho (como mostrado em estudos anteriores). A análise de sobrevida pode ser estimada usando modelos de regressão baseados em dados coletados de candidatos e doadores (como o modelo de regressão de riscos proporcionais de Cox); isso forneceria informações sobre o benefício em termos de sobrevida que um determinado transplante poderia

proporcionar a um paciente. Acreditamos que o índice prognóstico poderia ser concebido, o que classificaria os pacientes que se submetem a transplantes a diversas categorias de risco (baixo, médio e alto), entre muitas outras possibilidades.¹³

CONCLUSÃO

Existe um grande volume de informações em formato digitalizado em bancos de dados de saúde eletrônicos. O próximo desafio é como derivar uma análise útil a partir dessas informações. A machine learning parece ser a opção mais viável para essa análise. A maneira como os profissionais de saúde lidam com os dados hospitalares mudará radicalmente nos próximos anos, e essa tendência é altamente relevante para decisões clínicas e cirúrgicas baseadas em padrões computacionais obtidos em cada prática local.

REFERÊNCIAS

1. Fisher RA. Statistical Methods for Research Workers. Edinburgh: Oliver and Boyd; 1934.
2. Nisbet R, Elder J, Miner G. Handbook of Statistical Analysis and Data Mining Applications. 1st ed. San Diego: Academic Press; 2009.
3. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med* 2016;375:1216-9.
4. Bihorac A, Ozrazgat-Baslanti T, Ebadi A, Motaei A, Madkour M, Pardalos PM, et al. MySurgeryRisk: Development and Validation of a Machine-learning Risk Algorithm for Major Complications and Death After Surgery. *Ann Surg* 2018 Feb 27. DOI: 10.1097/SLA.0000000000002706 [Epub ahead of print]
5. Srinivas TR, Taber DJ, Su Z, Zhang J, Mour G, Northrup D. et al. Big Data, Predictive Analytics, and Quality Improvement in Kidney Transplantation: A Proof of Concept. *Am J Transplant* 2017;17:671-81.
6. Krikov S, Khan A, Baird BC, Barenbaum LL, Leviatov A, Koford JK, et al. Predicting kidney transplant survival using tree-based modeling. *ASAIO J* 2007;53:592-600.
7. Goldfarb-Rumyantzev AS, Scandling JD, Pappas L, Smout RJ, Horn S. Prediction of 3-yr cadaveric graft survival based on pre-transplant variables in a large national dataset. *Clin Transplant* 2003;17:485-97.
8. Fouad M, Ellatif MMA, Hagag M, Akl A. Prediction of long term living donor kidney graft outcome: Comparison between rule based decision tree and linear regression. *Int J Adv Comp Res* 2015;3:185-92.
9. Petrovsky N, Tam SK, Brusica V, Bajic V, Socha-Hernandez LA, Russ G. Use of artificial neural networks in improving renal transplantation outcomes. *Graft* 2002;25:6-13.
10. Jiakai L, Serpen G, Selman S, Franchetti M, Riesen M, Schneider C. Bayes net classifiers for prediction of renal graft status and survival period. *Int J Med Med Sci* 2010;1:215-21.
11. Yoo KD, Noh J, Lee H, Kim DK, Lim CS, Kim YH, et al. A Machine Learning Approach Using Survival Statistics to Predict Graft Survival in Kidney Transplant Recipients: A Multicenter Cohort Study. *Sci Rep* 2017;7:8904.
12. Taber DJ, Palanisamy AP, Srinivas TR, Gebregziabher M, Odeghe J, Chavin KD, et al. Inclusion of dynamic clinical data improves the predictive performance of a 30-day readmission risk model in kidney transplantation. *Transplantation* 2015;99:324-30.
13. Ravikumar A, Saritha R, Chandra V. Recent Trends in Computational Prediction of Renal Transplantation Outcomes. *Int J Comp Appl* 2013;63:33-7.