MEDINFO 2019: Health and Wellbeing e-Networks for All
L. Ohno-Machado and B. Séroussi (Eds.)

doi:10.3233/SHTI190173

Graft Rejection Prediction Following Kidney Transplantation Using Machine Learning Techniques: A Systematic Review and Meta-Analysis

Aldilas Achmad Nursetyo[a,b], Shabbir Syed-Abdul[a,b], Mohy Uddin[c], Yu-Chuan (Jack) Li[a,b,d]

[a] *Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taiwan*
[b] *International Center for Health Information Technology, Taipei Medical University, Taiwan*
[c] *King Abdullah International Medical Research Center, King Saud bin Abdulaziz University for Health Sciences, Executive Office, King Abdulaziz Medical City, Ministry of National Guard Health Affairs, Riyadh, Kingdom of Saudi Arabia*
[d] *Research Center of Cancer Translational Medicine, Taipei Medical University, Taiwan*

Abstract

Kidney transplantation is recommended for patients with End-Stage Renal Disease (ESRD). However, complications, such as graft rejection are hard to predict due to donor and recipient variability. This study discusses the role of machine learning (ML) in predicting graft rejection following kidney transplantation, by reviewing the available related literature. PubMed, DBLP, and Scopus databases were searched to identify studies that utilized ML methods, in predicting outcome following kidney transplants. Fourteen studies were included. This study reviewed the deployment of ML in 109,317 kidney transplant patients from 14 studies. We extracted five different ML algorithms from reviewed studies. Decision Tree (DT) algorithms revealed slightly higher performance with overall mean Area Under the Curve (AUC) for DT (79.5% ± 0.06) was higher than Artificial Neural Network (ANN) (78.2% ± 0.08). For predicting graft rejection, ANN and DT were at the top among ML models that had higher accuracy and AUC.

Keywords:

Kidney Transplantation; Graft Rejection; Machine Learning

Introduction

Kidney transplantation provides high-quality life years to patients with ESRD. Outcomes following kidney transplantation are evaluated by renal function and graft rejection. Recipients' clinical status and outcomes after the transplant are influenced by recipients' ages, Human Leukocyte Antigen (HLA) matching, HLA immunization, ethnic background, time on dialysis, and cardiovascular comorbidities [10; 19].

Graft rejection is the most common problem for kidney transplant recipients. Antibody-mediated rejection requires a distinct therapy as compared to the therapy for usual T-cell-mediated acute rejection. Renal function, based on estimated Glomerular Filtration Rate (GFR) and/or proteinuria values, is a result of these factors. Renal function impairment, whether in a stable condition or as a progressing dysfunction, also has an impact. Confirmation of graft rejection needs renal biopsy, which in turn requires expenditure and time. Antibiotics and immunosuppressive drugs could decrease acute graft rejection incidence, but chronic graft rejection is still a major problem [7; 28]. Specific and nonspecific (diabetes, nephrotoxicity, infection and cancer) conditions could have significant negative long-term consequences [23].

Due to the increased availability clinical data and rapid development of computing technology, artificial inteligence (AI) has been successfully applied in the healthcare domain. AI uses advanced learning algorithms to analyze large volumes of healthcare data that facilitates decision making in clinical practice. Various forms of clinical data (such as, diagnosis, screening, and treatment assignment) can be used as training data before AI systems can be applied in daily healthcare settings. By doing so, the system can learn and apply AI to similar groups of subjects, associations between subject features, and outcomes of interest. Training data is not only limited to clinical data, but it also includes demographics, medical notes, electronic recordings from medical devices, physical examinations, and clinical laboratory and images [15].

Machine learning (ML) techniques are used for specific purposes. Each unique feature of ML can be used for a different function, therefore different results may be obtained with different ML models. In the literature review, we did not come across any meta-analysis study specifically evaluating the use of ML algorithms for predicting kidney transplant outcomes [17]. The objectives of this study are: (1) to review the role of ML in predicting graft rejection following kidney transplantation, and (2) to specifically identify ML algorithms that have a higher accuracy and performance for predicting graft rejection following kidney transplantation, reported in the literature.

Methods

This review paper followed the flowchart and checklist provided by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [24].

Search Strategy and Data Sources

In order to get accurate results from the vast biomedical and health research databases—PubMed and Scopus, we used combined search keywords, such as "kidney transplants AND machine learning", "kidney transplants AND data mining", and "kidney transplants AND artificial intelligence". For covering the domain of computer science, we also used DBLP, which is a database for scientific journals in the field of computer science that are not yet indexed by PubMed, although the proposed published methods are applied on biomedical datasets [16]. In DBLP, we used the keyword "kidney transplants" to get all the related studies with this specific search criterion. We included all studies found in the literature until July 9, 2018, and checked for duplicate findings.

Eligibility Criteria

The articles were included based on the following criteria:

1. Written in English.
2. Using ML techniques to predict graft rejection in patients who had undergone kidney transplantations.
3. Building model extracted on medical record (eg. renal registry).

Articles were excluded based on the following criteria:

1. Used features other than clinical features extracted from medical record (eg. –omics features, radiological imaging).
2. Using ML techniques other than graft rejection prediction purpose.
3. Did not mention the result of ML performance.

Data Extraction and Synthesis

The selected studies were summarized based on the author name, year of publication, number of patients, dataset details, types of input variables, used ML techniques, and also validation method. The aim of the summarizing process was to get a detailed description of each ML model generated by the studies. Performance of the ML models was also recorded by mentioning each performance metric (accuracy and AUC). Comparison of each ML techniques was analyzed based on a previous study done by Malhotra [22]. Performance was visualized for each ML techniques and performance metrics.

Results

By applying inclusion and exclusion criteria mentioned above, 14 articles were identified for detailed study analysis (Figure 1).

Datasets and Patients

From 14 selected studies, most studies used publicly available datasets, such as United States Renal Data System (USRDS), United Network for Organ Sharing (UNOS) and the Eurotransplant database. While other studies used self collected data from cohorts hosted by their organizations. Table 2 describes the datasets and number of patients used in each study. Number of patients used in studies ranged from 80 to 57,389. A total of 109,317 patients were described in this study.

USRDS is a national data system that collects, analyzes, and distributes information about ESRD and CKD in United States population [4]. It collects data on patient demographic characteristics, contact information, treatment, laboratory values, quality-of-life survey interviews and nutrition survey interviews for dialysis patients, and also facilitates data sharing by filling request form provided in their web page [9]. UNOS is a private, non-profit organization based in US focused on organ transplant procurement by maintaining contact with volunteers. UNOS also maintains Organ Procurement and Transplantation Network which contains pre-transplant data pertains to transplant candidates [3].

ANZDATA is a registry that holds records of the incidence, prevalence and outcome of dialysis and transplant treatment for patients with end stage renal failure in Australia and New Zealand population [1]. Like UNOS, Eurotransplant is also a non-profit organization that facilitates procurement organ transplant across Europe, especially post-mortem donor organs. It collects various data both from the donors' side and the recipients' side along with the outcomes of procedures [2]. The other included studies collected longitudinal data of kidney transplantation from local organization (e.g. teaching hospital) in order to conduct their research.

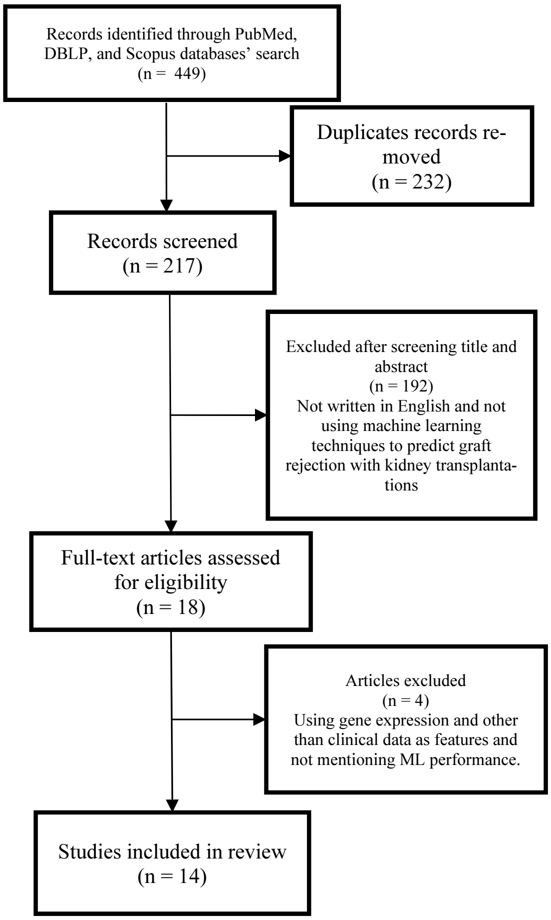

Figure 1. PRISMA Flow Diagram for Study Selection Process

ML Techniques and Overall Performance

Overall, 5 ML techniques that were used in the 14 included studies are: Artificial Neural Network (ANN), Support Vector Machine (SVM), Bayesian Belief Network (BBN), Decision Tree (DT) and an ensemble learning method called Random Forests. ANN was the topmost technique used in 7 studies. The second most commonly used technique was DT algorithm that appeared with different type such as C.50, Classification & Regression Tree (CART/C&RTree) and Random Forests.

Figure 2 describes the ML performances in the form of box plots. From the box plots, it is clear that DT and ANN mostly outperform all other techniques that had been used in studies. Tang et al [31] showed that ANN could perform better than statistical learning methods, such as Logistic Regression (LR). While Shaikhina et al. [27] showed that DT still can be the technique of choice even after being applied in ensemble methods, such as Random Forest. Study done by Esteban et al. [8] showed high performance by utilizing Recurrent Neural

Table 2. Dataset and number of patients included in studies.

Study No.	Author	Year	Dataset	Number of patients
1	Lin et al. [20]	2008	USRDS (2003) + UNOS	57389
2	Topuz et al. [33]	2017	UNOS (2004-2015)	31207
3	Brown et al. [6]	2012	USRDS (2004)	7348
4	Tang et al. [31]	2011	USRDS (2002)	4754
5	Yoo et al. [34]	2017	Misc.	3117
6	Esteban et al. [8]	2016	Misc.	2061
7	Shadabi et al. [25]	2004	ANZDATA Registry Database (2000)	1344
8	Lasserre et al. [18]	2012	Eurotransplants database (1998-2008)	707
9	Shahmoradi et al. [26]	2016	Misc.	513
10	Tapak et al. [32]	2017	Misc.	378
11	Greco et al. [13]	2010	Misc.	194
12	Hummel et al. [14]	2010	Misc.	145
13	Lofaro et al. [21]	2010	Misc.	80
14	Shaikhina et al. [27]	2017	Misc.	80

USRDS = United States Renal Data System; UNOS = United Network for Organ Sharing; ANZDATA = Australia & New Zealand Dialysis and Transplant Registry

Figure 2 (a,b) Box plots showing ML performances of studies based on (a) Accuracy and (b) AUC. Outliers are shown by study number depicted in Table 2. ANN = Artificial Neural Network, BBN = Bayesian Belief Network, SVM = Support Vector Machine

Network (RNN). The model combined non-linear and linear features (medication perscriptions and laboratory results), along with with static features (gender, age, weight) showing 82% of AUC performance.

Discussion

This paper reviewed the role of applying AI techniques (ML methods) in predicting graft rejection following kidney transplantation, and described the algorithms used by critically reviewing their performances. It is important to be clear about the specific outcomes to be studied, before deploying ML methods.

Based on our results, ANN and DT were the most commonly used models. These techniques showed better performance than SVM, Random Forest and BBN. As DT has the robustness to noise, low computational cost, and ability to deal with redundant features; it has advantages over other learning algorithms. DT could be inducted in various ways, such as C5.0 and CART, but none have been shown to be superior to other methods [5].

ANN is an mathematical algorithm that represents the human neural architecture and resembling the function like learning and generalizing ability. Nowadays, these techniques are widely applied in various research fields because they can show good performance in finding relationship among unknown or complex variables, such as non-linear variables. ANN can be applied in various ways, the most used techniques are Multi-layer Perceptron (MLP), with 3 important layers: input layer, hidden layer, and output layer. This technique is described as being fully connected to every node in the next and previous layer. MLP are trained by selecting suitable connecting weights and transfer functions between the input and output vectors [11]. In this group, prediction model using RNN algorithm developed by Esteban et al. [8] are the most powerful in classification power. RNN are kind of neural networks that usually applied in sequential data such as voice recognition and natural language processing (NLP). The algorithm elaborate both dynamic and static data from medical record that are relevant to predict future outcomes [12].

While doing the literature search, we also found some other systematic reviews related to AI techniques and transplantation. Sousa et al. [30] has reviewed AI techniques used for analysing organ transplant databases from 2009 to 2010 from PubMed and Web of Knowledge. They inferred that the main techniques used were: ANN, LR, DT, Markov Models (MM), and Bayesian Networks (BN). ANN was most preferred for knowledge extraction. Singh et al. [29] provided a systematic review of clinical prediction models of patient and graft survival in kidney transplant recipient using Medline and EMBASE databases covering the time period from 1966 to 2013. They showed the model discrimination with 'C' statistics for patient survival models and graft survival models and reported calibration and external validation of the methods. They also deduced modest discriminatory ability in most clinical prediction models, variability in other measures of model performance, and inconsistency for external validation of models. While Sousa [30] focused on AI techniques that were applied to extract knowledge from transplantation databases, Singh [29] reviewed articles that developed clinical prediction models of patient and graft survival in kidney transplant recipients. In comparison to these articles, our review article specifically studied the role of AI techniques (ML methods) utilized in predicting outcomes following kidney transplantation, and also evaluated the performances of the algorithms used.

More studies are desirable to compare different models. Hybrid models could be used for prediction enhancements. Ensemble and deep learning methods could also be considered in the future.

Conclusions

Based on the PRISMA guidelines, this study evaluated the role of AI techniques (ML algorithms) in predicting treatment outcome following kidney transplantation by examining the available literature. From the literature and our results it was clear that there is no 'One size fits all' approach for applying ML methods. Selection of the right algorithm, provided input variables and volume, and accuracy of the training datasets are critical. Based on performance measured by sensitivity, specificity, accuracy, and AUC, we concluded that ANN and DT were the most suitable and prevalent methods to predict graft rejection following transplantation procedure. A new model built with features taken from both donors' and recipients' side is desirable. Comparison of various models, especially Ensemble method and Deep Learning is required for the future work.

Acknowledgements

This research is sponsored in part by Ministry of Science and Technology (MOST) under grant MOST 106-2221-E-038-005, 106-2923-E-038-001-MY2, 107-2923-E-038 -001 -MY2, Taipei Medical University under grant number 106-3805-004-111 and Wanfang hospital under grant number 106TMU-WFH-01-4.

Address of Correspondence

Shabbir Syed-Abdul

E-mail : drshabbir@tmu.edu.tw

Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, 15F., No. 172-1, Sec. 2, Keelung Rd., Da'an Dist., Taipei, Taiwan 10675

References

[1] ANZDATA Registry, in.

[2] Eurotransplant Database, in.

[3] United Network for Organ Sharing, in.

[4] United States Renal Data System (USRDS), in.

[5] R.C. Barros, M.P. Basgalupp, A.C.P.L.F. de Carvalho, and A.A. Freitas, A Survey of Evolutionary Algorithms for Decision-Tree Induction, *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **42** (2012), 291-312.

[6] T.S. Brown, E.A. Elster, K. Stevens, J.C. Graybill, S. Gillern, S. Phinney, M.O. Salifu, and R.M. Jindal, Bayesian modeling of pretransplant variables accurately predicts kidney graft survival., *American Journal of Nephrology* **36** (2012), 561-569.

[7] R.B. Colvin, Antibody-mediated renal allograft rejection: diagnosis and pathogenesis, *J Am Soc Nephrol* **18** (2007), 1046-1056.

[8] C. Esteban, O. Staeck, S. Baier, Y. Yang, and V. Tresp, Predicting Clinical Events by Combining Static and Dynamic Information Using Recurrent Neural Networks, (2016), 93-101.

[9] R.N. Foley and A.J. Collins, The USRDS: what you need to know about what it can and can't tell us about ESRD, *Clin J Am Soc Nephrol* **8** (2013), 845-851.

[10] G. Garcia-Garcia, P. Harden, and J. Chapman, The global role of kidney transplantation, *Indian Journal of Nephrology* **22** (2012), 77-82.

[11] M.W. Gardner and S.R. Dorling, Artificial Neural Networks (The Multilayer Perceptron)—A Review Of Applications In The Atmospheric Sciences, *Atmospheric Environment* **32** (1998), 2627-2636.

[12] A. Graves, A.-r. Mohamed, and G. Hinton, Speech recognition with deep recurrent neural networks, in: *2013 IEEE international conference on acoustics, speech and signal processing*, IEEE, 2013, pp. 6645-6649.

[13] R. Greco, T. Papalia, D. Lofaro, S. Maestripieri, D. Mancuso, and R. Bonofiglio, Decisional trees in renal transplant follow-up, *Transplant Proc* **42** (2010), 1134-1136.

[14] A.D. Hummel, R.F. Maciel, R.G. Rodrigues, and I.T. Pisa, Application of artificial neural networks in renal transplantation: classification of nephrotoxicity and acute cellular rejection episodes, *Transplant Proc* **42** (2010), 471-472.

[15] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen, and Y. Wang, Artificial intelligence in healthcare: past, present and future, *Stroke Vasc Neurol* **2** (2017), 230-243.

[16] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, Machine Learning and Data Mining Methods in Diabetes Research, *Comput Struct Biotechnol J* **15** (2017), 104-116.

[17] K. Kourou, T.P. Exarchos, K.P. Exarchos, M.V. Karamouzis, and D.I. Fotiadis, Machine learning applications in cancer prognosis and prediction, *Comput Struct Biotechnol J* **13** (2015), 8-17.

[18] J. Lasserre, S. Arnold, M. Vingron, P. Reinke, and C. Hinrichs, Predicting the outcome of renal transplantation., *Journal of the American Medical Informatics Association* **19** (2012), 255-262.

[19] C. Legendre, G. Canaud, and F. Martinez, Factors influencing long-term outcome after kidney transplantation, *Transplant International* **27** (2014), 19-27.

[20] R.S. Lin, S.D. Horn, J.F. Hurdle, and A.S. Goldfarb-Rumyantzev, Single and multiple time-point prediction models in kidney transplant outcomes, *Journal of Biomedical Informatics* **41** (2008), 944-952.

[21] D. Lofaro, S. Maestripieri, R. Greco, T. Papalia, D. Mancuso, D. Conforti, and R. Bonofiglio, Prediction of Chronic Allograft Nephropathy Using Classification Trees, *Transplantation Proceedings* **42** (2010), 1130-1133.
[22] R. Malhotra, A systematic review of machine learning techniques for software fault prediction, *Applied Soft Computing* **27** (2015), 504-518.
[23] N. Mohammadpour, S. Elyasi, N. Vahdati, A.H. Mohammadpour, and J. Shamsara, A Review on Therapeutic Drug Monitoring of Immunosuppressant Drugs, *Iranian Journal of Basic Medical Sciences* **14** (2011), 485-498.
[24] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, and P.G. The, Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement, *PLOS Medicine* **6** (2009), e1000097.
[25] F. Shadabi, R. Cox, D. Sharma, and N. Petrovsky, Use of Artificial Neural Networks in the Prediction of Kidney Transplant Outcomes, in: *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, 2004.
[26] L. Shahmoradi, M. Langarizadeh, G. Pourmand, Z.A. Fard, and A. Borhani, Comparing Three Data Mining Methods to Predict Kidney Transplant Survival, *Acta Inform Med* **24** (2016), 322-327.
[27] T. Shaikhina, D. Lowe, S. Daga, D. Briggs, R. Higgins, and N. Khovanova, Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation, *Biomedical Signal Processing and Control* (2017).
[28] B. Shrestha, J. Haylor, and A. Raftery, Historical perspectives in kidney transplantation: an updated review., *Progress in Transplantation* **25** (2015), 64-69, 76.
[29] S. Singh, D. Naimark, J. Victor, and S. Kim, Clinical Prediction Models of Patient and Graft Survival in Kidney Transplant Recipients: A Systematic Review.: Abstract# D2483, *Transplantation* **98** (2014), 631-632.
[30] F.S. Sousa, A.D. Hummel, R.F. Maciel, F.M. Cohrs, A.E. Falcao, F. Teixeira, R. Baptista, F. Mancini, T.M. da Costa, D. Alves, and I.T. Pisa, Application of the intelligent techniques in transplantation databases: a review of articles published in 2009 and 2010, *Transplant Proc* **43** (2011), 1340-1342.
[31] H. Tang, M.R. Poynton, J.F. Hurdle, B.C. Baird, J.K. Koford, and A.S. Goldfarb-Rumyantzev, Predicting Three-Year Kidney Graft Survival in Recipients with Systemic Lupus Erythematosus, *American Society of Artificial Internal Organs* (2011).
[32] L. Tapak, O. Hamidi, P. Amini, and J. Poorolajal, Prediction of Kidney Graft Rejection Using Artificial Neural Network., *Healthcare informatics research* **23** (2017), 277-284.
[33] K. Topuz, F.D. Zengul, A. Dag, A. Almehmi, and M.B. Yildirim, Predicting graft survival among kidney transplant recipients: A Bayesian decision support model, *Decision Support Systems* **106** (2017), 97-109.
[34] K.D. Yoo, J. Noh, H. Lee, D.K. Kim, C.S. Lim, Y.H. Kim, J.P. Lee, G. Kim, and Y.S. Kim, A Machine Learning Approach Using Survival Statistics to Predict Graft Survival in Kidney Transplant Recipients: A Multicenter Cohort Study, *Scientific Reports* **7** (2017), 8904.