Use of Artificial Neural Networks in Improving Renal Transplantation Outcomes

Nikolai Petrovsky, Soh Khum Tam, Vladimir Brusic, Graeme Russ, Luis Socha, and Vladimir B. Bajic

Recent advances in renal transplantation, including the matching of major histocompatibility complex or new immunosuppressants, have improved 1-year survival of cadaver kidney grafts to more than 85%. Further optimization of kidney transplant outcomes is necessary to enhance both the graft survival time and the quality of life. Techniques derived from the artificial intelligence enable better prediction of graft outcomes by using donor and recipient data. The authors used an artificial neural network (ANN) to model kidney graft rejection and trained it with data on 1542 kidney transplants. The ANN correctly predicted 85% of successful and 72% of failed transplants. Also, ANN correctly predicted the type of rejection (hyperacute, acute, subacute, and chronic) for approximately 60% of the failed transplants. These results indicate that the ANN-based approach is useful for prediction of both the general outcomes of kidney transplants and the prediction of the type of rejection.

Introduction

Renal failure is an increasing problem around the world. Whereas the incidence of renal failure due to infectious causes is stable or falling in the developed countries, failure secondary to diabetes is on the increase. This increase is in parallel with the increase in prevalence of type 2 diabetes. Although dialysis is a short-term solution, renal transplantation remains the optimum solution both for restoring quality of life and for increasing life expectancy of patients. A major limitation to renal transplantation is the supply of donor kidneys. Although success rates from renal transplantation continue to improve, many donor kidneys continue to be lost due to rejection or recurrent disease. Consequently, many renal transplant recipients require a 2nd or subsequent graft. The ability to improve the graft success rate and thereby reduce the number of patients requiring multiple grafts would both improve patient outcomes and increase availability of donor kidneys for primary recipients.

Although advances in renal transplantation such as HLA matching and improved immunosuppressive medication have reduced transplant failure rates, further optimization of renal transplant outcomes is necessary to improve both the survival time of the graft and the quality of life of recipients. Techniques derived from the artificial intelligence (AI) field, such as artificial neural networks (ANNs), may better predict graft outcomes after training on a combination of donor and recipient data and thereby may be used to optimize donor recipient selection. They could also be used to identify recipients at increased risk for acute rejection and target them for more aggressive immunosuppression regimens. Combining the strength of AI methods with patient and transplant data may provide a breakthrough tool for kidney allocation and the management of posttransplantation immunosuppression.

Factors That Determine Graft Outcome

Two of the most important determinants of the outcome of renal transplantation are the degree of HLA matching and the cold ischemia time (CIT), which is the total time between removal of the kid-
ney from the donor and its transplantation into the recipient. Other factors that influence short-term graft outcome include presence of donor-specific cytotoxic antibodies, blood group matching, number of prior grafts, age of donor and recipient, time on dialysis prior to transplantation, diabetes in the recipient, race, living or cadaver donor, and transplant center. Although the importance of HLA matching has diminished with the advent of improved immunosuppression regimens and, in particular, the advent of cyclosporin, grafts with no HLA mismatches continue to demonstrate superior graft survival.4

Part of the difficulty in analyzing transplant data is that this is a rapidly changing field. Transplantation practices keep changing over time as improvements are identified and new treatments become available. Historic data from the precyclosporin era that show the importance of HLA matching may not be relevant today, given the availability of much more effective antirejection regimens. Similarly, whereas transfusion prior to transplantation used to have a positive effect on outcome, recent data suggest this is no longer the case.5

Current Methods of Donor Kidney Allocation

Given the marked imbalance between supply and demand, the allocation of cadaver donor organs becomes a conflict between utility and equity.6 Factors influencing donor organ distribution can be classified as medical, ethical, utilitarian, financial, or political. Medical factors include likelihood of patient survival and long-term graft survival, suitability of recipient, and medical urgency. Ethical issues include equality of access regardless of age, religion, race, or socioeconomic status. Utilitarian issues include proximity of nearest transplant center and dependency of organ donation on perceived local access to those organs. Financial issues include who pays for the cost of transplantation and posttransplant treatment. Political issues may involve attempts to ensure good electoral region access relative to other electorates.

Methods of kidney allocation vary both within and between countries. Most methods of donor organ allocation involve the use of simple algorithms designed to take into account major factors thought to influence graft outcome such as donor/recipient matching,7 plus various equity algorithms designed to ensure that minority groups are not disadvantaged,8 or that regions with high kidney donor rates retain some of the benefits by way of greater accessibility. To address perceived problems of inadequate supply and inequitable distribution, the U.S. Congress in 1984 enacted the National Organ Transplant Act under which the United Network for Organ Sharing (UNOS), a nonprofit organization, regulates, coordinates, and supervises regional organ procurement agencies. UNOS administers a kidney allocation policy using algorithms based on HLA-matching, time on the waiting list, medical urgency, and geographic area. With the permission of UNOS, regions such as New England have adopted slightly different algorithms to achieve particular targets such as the number of local versus remote transplants by downplaying the emphasis on HLA matching and placing greater emphasis on distance from donor source.9 Because algorithms based on outcome and equity are calculated independently and then combined, these methods cannot provide optimal outcomes. Regional or national sharing of organs to achieve better HLA matching, for example, may lead to greater CIT and consequently a lower rate of graft survival than if kidneys were transplanted locally. Similarly, the provision of suboptimally matched kidneys to minority groups will inevitably lead to poorer graft outcomes. This problem may be compounded by greater CIT if these groups are geographically distant from the major sites of organ donation. The biggest problem of such algorithms, however, is the fact that they are based on broad population rather than individual data. Decisions made on the basis of such algorithms may never, therefore, provide optimal graft outcomes at the individual level.

Renal Transplantation Outcomes

Over the 35 years of the Michigan Transplant Program, 1-year survival for cadaver kidney grafts has improved from 25% to more than 85%.10 An analysis of UNOS data between 1987 and 1999 reveals an estimated 10-year rate of cadaveric graft survival of 52% for HLA-matched transplants in comparison to 37% for mismatched transplants. The estimated half-lives of the transplants were
12.5 and 8.6 years, respectively. Among more than 3500 pairs of kidneys, HLA-matched transplants had a lower incidence of episodes of rejection as well as a higher rate of survival. For primary cadaver grafts performed in 1998 in Australia, 12-month patient and graft survival was 95% and 91%, respectively. Likewise, 5-year graft survival for primary cadaver grafts performed in 1994 was 72% (according to the 23rd report of the Australia and New Zealand Dialysis and Transplant Registry).

Use of AI in Clinical Decision Making

The reasons simple algorithms continue to be relied on in donor kidney allocation include the complexity of the problem of predicting outcomes at the individual level while attempting to resolve equity issues. Furthermore, this analysis needs to be done in an extremely short time frame, and often with limited data. Traditional statistical methods are inadequate for handling problems of such complexity. AI modeling methods have been used for solving problems of high complexity. These methods provide advanced tools in support of decision making for complex problems, such as transplant allocation. The advantages of AI methods and tools, such as ANNs, are that they are capable of dealing with large quantities of complex, incomplete, and often imprecise data and are able to identify optimized solutions taking into account all relevant information. ANNs have previously been used in a wide variety of clinical settings including disease diagnosis, classification and staging, image analysis, and disease outcome prediction. A brief description of ANNs is provided in the appendix.

Uses of ANNs in Medical Outcome Prediction

ANNs have been used in various settings to predict medical outcomes. In this section, we list just some of the applications of ANNs in the prediction of medical outcomes. ANNs are significantly more accurate than the standard TNM staging system in predicting survival of patients with breast cancer or advanced colon cancer. ANNs can predict self-sufficiency in a geriatric population, discharge destination from a postacute geriatric rehabilitation, and 2-year survival in elderly people. They predict survival in patients with squamous cell carcinoma of the head and neck more accurately than logistic regression. They improve the prediction of mortality in patients with intracerebral hemorrhage and perform better than emergency department physicians in predicting which patients with closed-head injuries have abnormalities on computed tomography scans. ANNs were 91% accurate in predicting subsequent mortality in trauma patients admitted to an emergency room, and ANNs outperformed other methods in predicting survival of trauma patients using prehospital, emergency room data and the Injury Severity Score. ANNs trained on data at the time of hospital admission predict patient survival following in-hospital cardiopulmonary resuscitation. ANNs have been used in the prediction of postoperative seizures in patients undergoing surgery for epilepsy, which patients with acute pancreatitis have a length of stay of greater than 7 days, the duration of the 1st stage of labor, and cancer progression after radical prostatectomy.

Use of ANNs in Transplantation Settings

Examples of ANN applications in transplantation include prediction of liver transplant rejection, prediction of tacrolimus blood levels in liver transplantation, diagnosis of early acute renal allograft rejection and evaluation of complications of renal transplants and evaluation of complications of renal transplants, prediction of cytomegalovirus disease after renal transplantation, prediction of pancreas transplant outcome, and MHC haplotype matching. ANNs of Kohonen type were able to classify patients as having glomerular or tubular kidney disease better than a team of 3 nephrologists. ANNs can be used to predict delayed renal allograft function as a guide for induction of immunosuppression therapy.

Data from 100 transplant recipients were used to train an ANN to predict delayed graft function (DGF). In 20 prospective patients, the ANN predicted DGF in 80% of cases, and serum creatinine was less than 2.5 mg/dL in all of the patients. The ANN model of DGF was then used to test the effect of cold ischemia time on DGF. It showed that a short CIT (<12 hours) could compensate for most other donor variables influencing DGF apart from anuria and increasing serum creatinine. If all other donor variables were favorable, then up to 36...
hours of CIT there was no effect on DGF. ANNs were also used to select the most important variables in the prediction of chronic renal allograft rejection progression based on retrospective data from 27 graft patients. This analysis showed that chronic glomerular changes, mean value of serum creatinine, cholesterol, triglycerides, blood pressure on day of biopsy, duration of dialysis before transplantation, mean dose of cyclosporin, and number of days of serum creatinine >150 mmol/L to biopsy all had a strong influence on chronic allograft rejection progression.

The aforementioned studies, however, were performed using data from extremely small patient numbers. There are currently no published reports of the use of large-scale renal transplant databases to train ANNs for optimal matching of donor kidneys. In a pilot study using retrospective data on 1542 renal transplant recipients, we trained an ANN and showed that such a project is feasible, obtaining a predictive accuracy of more than 80%. We performed an analysis on renal transplant data from the Australian and New Zealand Dialysis and Transplant Registry, which contains data on all transplants performed in Australia over the past 30 years, to see whether an ANN could be trained to predict graft survival. For this pilot study, a relatively small subset of records of successful and unsuccessful transplants (771 in each group) were split in 2 sets. From each of the 2 groups, 70% of records were used for training and the remaining 30% for testing. We used a 3-layer feedforward network architecture. (see the appendix). Data preprocessing included the normalization and principal components analysis. The ANN (Fig. 1) was trained to predict kidney transplant outcomes using 22 variables. These variables included data on donor (age, sex, graft source), recipient (age, CMV and EBV antibody status, other organ transplants, earlier transfusions), institutions (referring, donor, transplant hospital, state), graft (total ischemia, kidney preservation), and HLA matching (HLA-A, -B, -DR, and DQ). The best performing network had 8 neurons in the hidden layer and a single output neuron.

![Diagram of an artificial neural network (ANN) with 3 layers: output layer (in this case having only 1 neuron), 1 hidden layer with many neurons, and an input layer which contains signal input points. For clarity, only representative connections are shown. The input vector is represented by signals $I_1, I_2, \ldots, I_n$. The output signal is represented by $O_1$, whereas representative connection weights are represented by $v_{11}, v_{12}, w_{11},$ and $w_{12}$. The type of ANN shown here is the multilayer perceptron. In the example presented in this article, we used the “error back-propagation” training algorithm.](image-url)
After training, the ANN was able to correctly predict 84.95% of successful transplants and 71.7% of unsuccessful transplants. These results show that a properly trained ANN is capable of predicting the success and failure of renal transplants with high accuracy. The ANN is better in predicting successful than unsuccessful transplants, suggesting that the factors that determine graft success may be inherently more predictable than the factors that determine graft failure.

We then used an ANN system to predict the type of graft rejection. We considered 4 classes of rejection: hyperacute (HA), acute (A), subacute (SA), and chronic (C). All rejection data were for patients who were alive at 6 months after transplant and rejection was the outcome at the end of the survey. The system for prediction is made up of a number of smaller ANNs, as depicted in Figure 2. Six networks were trained to distinguish between 6 pairs of rejection outcomes (HA-A, HA-SA, HA-C, A-SA, A-C, SA-C), and 4 networks were trained to distinguish successful cases from the 4 types of rejection outcomes. These ANNs, each having identical architecture (described in the appendix), were combined in a prediction system for the prediction of graft rejection outcomes. The results for recognition of the type of the graft rejection are shown in Table 1. In the case of random predictions of transplant outcome, the probability of correctly predicting any 1 of the 4 possible outcomes will be 25%. However, the ANN-based prediction correctly predicted 59% of rejection outcomes with respect to the type of rejection. These results indicate that the ANN-based approach is useful for both the prediction of general outcomes of kidney transplants and the prediction of the type of rejection. With the in-
creased number of training data, the prediction accuracy typically improves.

**Future Use of AI in the Allocation of Donor Kidneys**

The biggest problem in improving renal transplant allocation may ironically not be the development of more reliable graft survival predictions but, rather, the gaining of acceptance of such predictions by the organ allocation bodies. These bodies are by their nature conservative and may be resistant to the adoption of new technology, particularly if it is seen to supplant their decision-making process. Such bodies are likely to require prospective validation of the technology, although testing with blinded retrospective data should be sufficient to determine the predictive power of the technology and would obviate several years of delay in its implementation. The impartiality of an AI-based allocation system should ultimately be its greatest strength, as it would prevent bias creeping into organ allocation. The existence of such bias is suggested by data such as, for example, evidence that black patients in the United States are half as likely as white patients to be rated as suitable candidates for transplantation when assessed by expert panels. Even when assessed suitable for transplantation, black patients are less likely to be referred for evaluation, placed on a waiting list, or transplanted. Clearly, no system that is required to balance the conflicting needs for utility and equity is ever going to be easy to develop and implement. Such difficulties should not stop us, however, from improving the current organ allocation system by incorporating more accurate assessment of the likely graft outcome. The technological advances supported by the AI methods for graft allocation will ultimately benefit the many patients currently awaiting organ transplants.

**Appendix: Artificial Neural Networks**

An artificial neural network (ANN) is an information-processing system composed of a large number of densely connected units. An ANN uses a set of nonlinear equations and inequalities that mimic the neuronal connections of biological systems. The process of ANN training emulates basic properties of biological neuronal systems, and thus ANNs exhibit adaptive learning properties. An ANN structure (Fig. 1) resembles the structure of a collection of neurons organized in layers. Artificial neurons are represented as circles in Figure 1, with the links between neurons resembling axons and synaptic junctions. Each link is characterized by its
strength, similar to the synapse strength in a living neuron. The interconnected artificial neurons can be trained to learn by example and to produce an appropriate response (i.e., produce an output signal) to a specific input stimulus.

An ANN comprises layers of units, which transmit signals through connecting links. The activation pattern of units depends on the strength of input signals, connection weights, and unit activation thresholds. Input to this ANN is a representation of data relevant for kidney transplants: output signal represents a decision—rejection versus successful transplant, or the type of rejection versus other type of rejection. The training of an ANN involves adjusting the connection weights and neuronal activation thresholds (parameters) until the ANN learns patterns that can predict transplant success or rejection. If in the training process the ANN does not respond properly (i.e., if it produces a response that is different from the expected one), then small adjustments are made to parameters in an attempt to reduce the error (Fig. 3). This process is repeated until the ANN achieves the predefined error level on the set of known examples (training set), at which point we say that the ANN is trained.

When new cases are presented to the trained ANN, it is able to generalize from the training examples and make accurate predictions with respect to these previously unseen cases. More detailed descriptions of ANNs are available in the literature.36,37

Acknowledgment

We wish to acknowledge the staff and management of the Australian and New Zealand Dialysis and Transplant Registry for their kind assistance with provision of transplant data for analysis.

Table 1 | Analyzing the Accuracy of the System for Predicting Outcomes of Kidney Transplants (see Figure 1)

<table>
<thead>
<tr>
<th>Training Set</th>
<th>HA</th>
<th>A</th>
<th>SA</th>
<th>C</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>82.0%</td>
<td>94.7%</td>
<td>78.4%</td>
<td>96.6%</td>
<td>91.2%</td>
</tr>
<tr>
<td>Test Set</td>
<td>62.5%</td>
<td>59.7%</td>
<td>56.3%</td>
<td>57.1%</td>
<td>59.3%</td>
</tr>
</tbody>
</table>

HA, A, SA, and C stand for hyperacute, acute, subacute, and chronic rejection. The training set results show the degree of data fit to the model. The test set results provide the assessment of the accuracy of the system on previously unseen data.

References