Recent Trends in Computational Prediction of Renal Transplantation Outcomes

Aswathy Ravikumar
Dept of Computer Science,
College of Engineering,
Thiruvananthapuram, India

Saritha R
Dept of Computer Science,
College of Engineering,
Thiruvananthapuram, India

Vinod Chandra S S
Computer Center,
University of Kerala,
Thiruvananthapuram, India

ABSTRACT
Renal transplantation has become the treatment of choice for most patients with end-stage renal disease. Recent advances in renal transplantation notably, the matching of Major Histocompatibility Complex (MHC) and improved immunosuppressants have improved short-term and long-term graft survival rates. In light of recent developments optimization of kidney transplant outcomes is paramount to further augment the graft survival time and the quality of life of the patient. An intuitive understanding of the post transplantation interaction mechanisms involving graft and host is intricate and on account of this prognosis of planned organ transplantation outcomes is an involved problem. Consequently, machine learning approaches based on donor and recipient data are indespensible for improved prognosis of graft outcomes. This study proposes improved data mining-based models for variable filtering and for prediction of graft status and survival period in renal transplantation using the patient profile information prior to the transplantation.

Keywords
Prediction model, Survival analysis, machine learning, Data mining, Renal Transplantation.

1. INTRODUCTION
Renal transplantation and maintenance dialysis are the only treatments for end-stage renal disease (ESRD). Despite being a major surgical procedure, transplantation is the only treatment of choice for ESRD patients, as a successful transplantation improves their quality of life. In particular, dialysis treatment requires patient visits of 12hrs a week to the dialysis facility, whereas transplantation typically allows the patient to resume regular life activities. Furthermore, research and clinical studies have statistically shown that transplantation also reduces the mortality risk for patients. Accordingly, a kidney transplant is considered by many as a potential life-saver.

Medical sector is one of the most vital sectors in service sector since it is life-crucial and errors can lead to fatal risks. Improper resource allocation has been identified as one of the perpetual problems in medical services. Particularly, the allocation of limited organs for transplantation requires foremost attention. Notwithstanding the advances in medical research organ transplantation is sole viable therapy for numerous end-stage diseases. Unfortunately the number of donor organs seldom meets the requirements. Consequently significant loss of human life results due want of an optimal match and this paucity often leads to suboptimal matching between the donor and recipients thereby further accentuating the risks involved in transplantation. The recent shortage in availability of organs for transplant can be attributed to two major causes. Primarily, due to advances in clinical research success rate of organ transplantations have increased and subsequently number of patients requesting a transplant also increased. Secondly, the fundamental problem of available donated organs always falling short of the increased requirements. This ever widening gap between the number of patients waiting for an organ transplant and donor supplies has increased waiting times, leading to the death of patients in waiting for a transplant [1].

Survival analysis is a branch of statistics that studies the surviving period of patient post transplantation. It is the primary evaluation method measuring the effectiveness of such an operation. Owing to the paucity of donor organs, survivability prediction is gaining importance in medical research [2]. In particular, the prediction of expected survival period and prognostic analysis of medical treatments are clinically important and involved problems [3]. Organ paucity demands development of effective and efficient procedures to select the most optimal organ receiver since the demand for organs might not be satisfied. In light of these the idea of benefit driven organ allocation scheme gains importance since it dwells on post-transplant outcome as the performance measure. Here the whole process is based on optimization of entire transplantation ensuring that transplantations will be performed only on patients who benefit from them in long term. Accordingly transplantation success prognosis and long term survivability prediction are pivotal in decision making.

The prediction of survival and the quality of life post transplantation are clinically important and involved problems. The design of such models is very complex and even more difficult to validate, verify and control so as to effectively predict the outcome of the transplantation [5]. Hence, modeling such a system necessitates effective procedures for the selection of best organ recipients since currently it is not possible to satisfy all organ demands.

The central objective of our study is to develop an integrated data mining method to predict accurate survival rate and to analyze the prognostic factors for different risk groups of transplant patients. We aim to analyze the relevant collated historical transplantation data set from a new cause-and-effect perspective. In doing so, we propose to use copious data sets with hundreds of determinative variables pertaining the donors, potential recipients, and transplantation procedures. Our study deals with the benefit-driven organ allocation schemes in terms of "causality" perspective because such a methodology would give clearer interpretation ability as well as a better prediction accuracy of the transplant success. While the former is extremely important to the medical professionals, the latter is essential for designing a satisfying optimal allocation scheme [8].

2. STRUCTURE AND FUNCTION OF KIDNEY

Kidney is one of the most complex organs in the human body; the kidneys perform several tasks including the excretion of waste products, the homeostatic balance in the body is maintained and hormone releases. Human kidneys are highly developed with refined anatomy and physiology. In order to accommodate the liver the right kidney is lower than the left kidney. To provide protection to the kidney it is enclosed in a membranous renal capsule made of fibrous connective tissue. In addition to that there is a protective layer of fat called the adipose capsule around the kidneys. An outermost layer of connective tissue (fascia) helps to anchor the kidney to the peritoneum and abdominal wall. The kidneys and the ureters, lie posterior to the peritoneum. It is due to this that the kidneys are in an area called theretroperitoneal space and not in the peritoneal cavity. The kidney is a somewhat flattened organ about four inches long, two inches wide, and one inch thick .Hilum is a notch on the medial border, where the renal artery, the renal vein, and the ureter connect with the kidney. The entire organ a bean-shaped appearance since lateral border is convex.

The kidney has mainly two regions: the renal cortex and the renal medulla. The kidney's outer portion is the renal cortex. The urine is formed and collected in the tubes in renal medulla. These tubes form a number of cone-shaped structures called renal pyramids. The tips of the pyramids point toward the renal pelvis, a funnel-shaped basin that forms the upper end of the ureter. Urine is collected in the cuplike extensions of the renal pelvis surround the tips of the pyramids. The urine that collects in the pelvis then passes down the ureters to the bladder.

Nephron is the basic unit, which actually does the kidney's work. The nephron is essentially a tiny coiled tube with a bulb at the end. This bulb, known as the glomerular capsule, surrounded by a cluster of capillaries called glomerulus. Each kidney contains millions of nephrons.

The afferent arteriole which supplies blood to the glomerulus is a small blood vessel; another small vessel, called the efferent arteriole is used to carry the blood from the glomerulus. From the glomerulus the blood flows into a capillary network that surrounds the nephrons tubular portion and not to the heart. The tubular portions of the nephron have several parts. The coiled part leading from the glomerular capsule is called the proximal convoluted tubule (PCT, or proximal tubule). The loop of Henle is a hairpin shaped segment. The first part of the loop is used to carry the fluid toward the medulla and it is the descending limb. The other part that continues from the loop's turn and carries fluid away from the medulla, is the ascending limb. Continuing from the ascending limb, the tubule coils again into the distal convoluted tubule (DCT, or distal tubule). The DCT empties into a collecting duct, which then continues through the medulla into the renal pelvis. The renal cortex contains the glomerular capsule, and the proximal and distal convoluted tubules of the nephron. The medulla contains loop of Henle and collecting. The detailed structure of a nephron is shown in fig1.

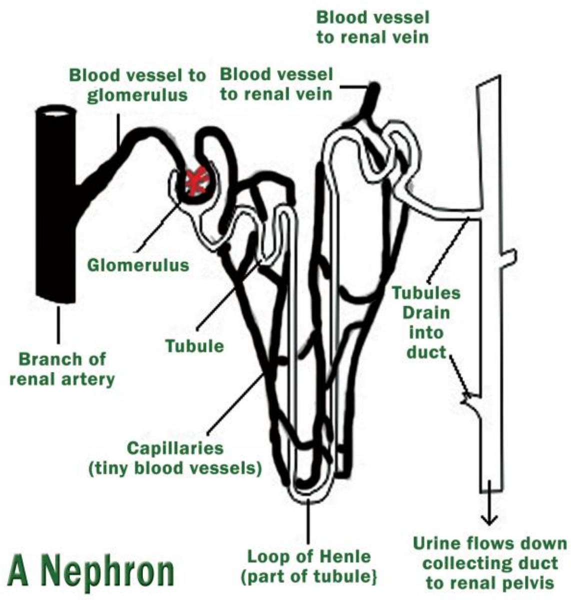

Fig 1: The Structure of nephron and glomerulus

The well–being of the body cells mainly depend on the chemical composition of body fluids, which is maintained with the help of renal system. The circulatory system of the body helps in the transportation of fluids but not for the composition of those fluids. The main functions of the kidneys are: to remove the waste and to maintain the body's water balance. Kidneys help to regulate the blood pressure via the renin-angiotensin-aldosterone system. Kidneys also help to regulate the blood electrolyte balance – Sodium ions, Potassium ions and the Calcium ions etc. Excretion of metabolic wastes such as urea, creatinine and foreign substances and the chemicals we ingest with our food are done by the kidney. Kidneys help in the regulation of the body's acid base balance. Kidneys also help to regulate the amount of RBC production via the hormone erythropoietin. It also helps in the production of vitamin D.

3. RELATED WORK

Organ Transplantation mainly focused to transplantation and its survival rate.

3.1 Computational Prediction of Organ Transplantation Outcomes

Earlier advanced statistical techniques were used for prediction like nonparametric or nonlinear statistical modelling, they are computationally very expensive especially in the cases of very crowded set of variables, and they require a good base knowledge about the data, to set the initial parameters and to make initial assumptions for the model. While, compared to that the data mining techniques provide much faster solution and do not need prior knowledge about the data. In the work reported in Asil et al. [16] an integrated data mining approach was used for the prediction of the graft outcome for combined heart–lung transplantation patients. The decision makers were assisted with the prediction of graft survival rate for a given set of possible recipients for an organ transplant such as a kidney. Techniques derived from the artificial intelligence increases the prediction accuracy of graft outcomes by using donor and recipient data. In [17] an artificial neural network (ANN) was used to design a model for the prediction of kidney graft and trained it with data on 1542 kidney transplants data.

3.2 Survival Analysis for Organ Transplantation

Recently, a large number of studies were conducted using medical data and the analysis of various organ transplantation datasets was done. Hariharan et al. [9] mainly focused towards graft survival rate analysis using the immunosuppressant drug (cyclosporine) after renal transplantation in both long duration (more than 1 year) and short duration (less than 1 year) and he used a regression analysis for that. R.A .Wolfe et.al [19] discussed a proposal for a new deceased donor kidney allocation system where the candidate was ranked based on his life span. Median lifespan was calculated based on objective medical criteria for each candidate. LYFT scores were higher for younger candidates and small for diabetic patient who received a kidney transplant. Jiakai Li [20] designed a bayes classifier for the prediction of the renal graft and they used mainly 70 variables for the prediction. But the model showed only 68% accuracy in the prediction of graft survival period. N. Hoot [21] conducted a study for the prediction of the graft survival rate of liver transplant candidates and they used the Bayesian belief network (BBN) .The main drawback of this study was that it used only small number of variables and it had only 67% accuracy. Similarly in [22] also a Bayesian belief network to develop a predictor for renal graft survival period. The study clearly showed that for a patient who is younger in age, had few HLA mismatch , low PRA level and had no previous history of transplantation had more life span compared to others.

Herrero et al. [10] conducted the study on liver transplant candidates, they compared between two groups (younger and older). The study showed that the patients belonging to the younger group have a greater lifespan after the transplantation. The main drawback of this study was that they used a very small dataset (166 patients). Survival analysis of liver transplant patients in Canada was done by Hong et al. [11] .In this study they considered many clinical and physical parameters which included age, donor type, blood group, HLA etc . However, the main drawback was the limited number of variables. Kusiak et al. [12] compared two data mining techniques (rough sets and decision trees) to predict the survival time of patients undergoing kidney dialysis. The main drawback of this study was the limited data set and the lack of many important variables. Moreover, the main drawback of the above mentioned studies was the small datasets with limited number of prediction variables for survivability prediction of patients after transplantation. This can lead to a faulty model with the main deciding parameters lost.

The survival analysis is done either by an unadjusted or risk-adjusted estimate. In the Unadjusted rates we assume that all patients in the dataset have the same chance of surviving given a define time span after transplant. But in real life each patient differ and there life span varies, so a risk-adjusted estimate which allows for these differences must be used. Kaplan-Meier method is used for the unadjusted survival rate estimation. The main advantage of this method is that it allows including the patients with incomplete information in the computation. Cox Proportional Hazard models are used for calculating risk adjusted estimate.

4. FEATURE SELECTION

In the early years of kidney transplantation, the important factor was HLA matching, the most allocation schemes was mostly weighted towards close tissue matching. With the new researches and the introduction of better immunosuppression, the other factors are being considered and given greater importance [9]. Now the most important factors that determine 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 kidney from the donor and its transplantation into the recipient. Other factors that influence the graft outcome include presence of antibodies, number of prior grafts, donor-specific cytotoxic age of donor and recipient, time on dialysis prior to transplantation, blood group matching, diabetes in the recipient, race, living or cadaver donor, and transplant centre. Although the importance of HLA matching has decreased with the advent of better immunosuppression regimens and, in particular, the advent of cyclosporin, grafts with no HLA mismatches continue to demonstrate superior graft survival. Another difficulty is in analyzing the transplant data that is rapidly changing. Transplantation practices keep changing with time as new improvements are discovered and new treatments become available. The table given below contains the list and description of the main features that should be considered during the prediction of renal outcomes and the survival rate.

TABLE I
MAIN FEATURES USED

Feature	Description
ABO	Recipient and donor blood group
Age	Recipient and donor age
HLA	HLA match level
Gender	Male/Female
CIT	Cold Ischemic time
RUS	Renal Ultra Sound
Race	Donor/ recipient race
Dialysis Status	Boolean
Type 1 Diabetics	Boolean
PRA	Panel Reactive Antibody
Steroid dose	In mg
Blood sugar level	>120mg/dl(Boolean)
Donor Type	Live/Cadaveric
Medical history	primary renal disease

5. CHALLENGES AND ISSUES

Medical diagnosis is a significant yet intricate task that needs to be carried out precisely and efficiently. The automation of the medical diagnosis would be highly beneficial. Clinical decisions are often made based on doctor's experience and his intuition in that area rather than on the knowledge rich date that lie hidden in the database. These common practices have lead to unwanted errors and excessive medical costs which affected the quality of service given to patients. Machine learning methods help to generate a knowledge enriched environment which is very helpful in improving the quality of clinical decisions. The biggest problem in improving renal transplant allocation is not the development of more reliable graft survival predictions but, the gaining of acceptance of such predictions by the organ allocation bodies. These organisations are usually conservative and may be resistant towards the adoption of new technology, particularly if it is seen to be replacing their decision-making processes.

6. RECOMMENDATIONS

In India, which are a highly populated country, studies and surveys showed that more than 10% people have some form of chronic kidney disease? Hypertension and Diabetes are the major cause of chronic kidney disease. In India about 20-30% of the adults have hypertension and many of them will develop chronic kidney disease; large proposition of adults in urban areas are diabetics and of them a majority are likely to develop kidney disease. There being no effective treatments, chronic kidney disease leads to Kidney failure. Kidney transplant seems to be the only last hope for Kidney failure patients. Every year a large number of patients with ESRD requiring kidney transplant are added to the waiting list. Only a very few of them are able to get some form of effective treatment, while the rest of them die without getting any definite treatment. The major limitations for successful treatment of chronic kidney disease are the high cost of treatment and non-availability of donor organs. The Union Health Ministry of India have launched a National Organ Transplant Programme to promote major transplantations like kidney, liver and heart in the country. It would be very effective if an online site was maintained by the government for the registration of patients and the interested donors and computational prediction was used for finding the perfect match. This could also lead to more research in this area.

In the future a new tool can be designed to help with the complex decision making process associated with the identification of a good transplant candidate for an available kidney. Rather than including all the clinical variables collected, we must determine the variables that would mostly account for the variation in outcomes of the prediction model, thereby limiting the number of input variables to the models. An effective ensemble classifier can be used for the prediction and from that the best match can be obtained by doing the survival analysis for each of the matching donor. In this method we can improve the efficiency and the survival period of the recipient (patients). We can also employs various predictive modelling techniques such as support vector machines, artificial neural networks, Bayes classifier, and regression trees to develop prediction models and to extract the most important variables by means of sensitivity analysis through the best performing model. Survival analysis can be estimated using regression models based on candidate and donor data collected, it provides information about the survival benefit that a given transplant could provide a given patient. This measure can be estimated by Cox proportional hazards regression models. A critical prognostic index can be devised, which categorizes the transplant patients in terms of various risk groups, namely low, medium, and high. In the future we can develop more sophisticated simulation tool which can simultaneously handle various organ transplantation scenarios, namely heart–lung, liver, and kidney.

7. CONCLUSION

In this paper the problem of constraining and summarizing different algorithms of data mining used in the field of medical prediction are discussed. The focus is on using different algorithms and combinations of several target attributes for intelligent and effective renal transplantation prediction using data mining. The survey shows that the most of the existing studies for organ transplantation procedures utilize conventional statistical approaches with expert-selected variables to predict the survivability. However, organ transplantation procedures consist of a large number of variables (several hundred) that may have nontrivial impact on modelling the prognosis of the grafts/patients. We can use a somewhat comprehensive variable list may help discriminate patients from each other by placing them into proper risk groups. Omission of the important variables may lead to inaccurate classification of patient risk groups, which may, in turn, lead to suboptimal organ allocation policies and in effective treatments.

8. REFERENCES

[1] Abouna, G. M. (2003). Ethical issues in organ transplantation. *Medical Principle and Practice, 12,* 54-69.

[2] Sheppard, D., McPhee, D., Darke, C., Shrethra, B., Moore, R., Jurewitz, A., & Gray, A. (1999). Predicting cytomegalovirus disease after renal transplantation: An artificial neural network approach". International Journal of Medical Informatics, 54(1), 55-76.

[3] Lin, H. M., Kauffman, H. M., McBride, M. A., Davies, D. B., Rosendale, J. D., Smith, C. M., Edwards, E. B., Daily, O.P., Kirklin, J., Shield, C. F., & Hunsicker, L. G. (1998). Center-specific graft and patient survival rates: 1997 UNOS report. The Journal of the American Medical Association (JAMA), 280(13), 1153-1160.

[4] Locatelli, F. and al., Nephrology: Main Advances in the Last 40 Years, Journal of Nephrology, Vol. 19, p. 6-11 ., 2006.

[5] J.M.A. Smith, J. Vanhaecke, A. Haverich, E. De Veries, L. Roels, G. Persijn, G. Laufer, Waiting for a thoracic transplant in eurotransplant, Transplant International 19 (2006).

[6] IBM SPSS Modeler, A comprehensive data/text mining software environment,version 14.08.

[7] F.L. Grover, M.L. Barr, L.B. Edwards, F.J. Martinez, R.N. Pierson, B.R. Rosengard, S.Murray, Thoracic transplantation, American Journal of Transplantation 3 (2003)91–102.

[8] M. Schmitt, H.N. Teodorescu, A. Jain, S. Jain, L.C. Jain, Computational intelligence processing in medical processing, Studies in Fuzziness and Soft Computing,Springer-Verlag, 2002.

[9] Hariharan S, Johnson CP, Bresnahan BA, Taranto SE, McIntosh MJ, Stablein D. "Improved graft survival after renal transplantation in the United States, 1988 to 1996." ,The New England Journal of Medicine 2000;342:605–12.

[10] Herrero JI, Lucena JF, Quiroga J, Sangro B, Pardo F, Rotellar F, et al. "Liver transplant recipients older than 60 years have lower survival and higher incidence of malignancy. ", American Journal of Transplantation 2003;3:1407–12.

[11] Hong Z, WuJ, Smart G, Kaita K, Wen SW, Paton S, et al. Survival analysis of liver transplant patients in Canada. Transplantation Proceedings 2006;38:2951–6.

[12] Kusiak A, Dixon B, Shah S. "Predicting survival time for kidney dialysis patients: a data mining approach". Computers in Biology and Medicine 2005;35:311–27.

[13] Jenkins PC, Flanagan MF, Jenkins KJ, Sargent JD, Canter CE, Chinnock RE, et al. "Survival analysis and risk factors for mortality in transplantation and staged surgery for hypoplastic left heart syndrome.", Journal of the American College of Cardiology 2000;36:1178–85.

[14] Asil Oztekin, "An Analytical Approach to Predict the Performance of Thoracic Transplantations" ,JCC: The Business and Economics Research Journal ,Volume 5, Issue 2, 2012 , 185-206.

[15] Sarah E. Taranto,Ann M. Harper,Erick B. dwards,John D. Rosendale, Maureen A. McBride,0. Patrick Daily, "Developing a National Allocation Model For Cadaveric Kidneys", Proceedings of the 2000 Winter Simulation Conference.

[16] Asil Oztekin, Dursun Delen,Zhenyu, " Predicting the graft survival for heart–lung transplantation patients: An integrated data mining methodology", international journal of medical informatics 7 8 (2 0 0 9) e84–e96.

[17] N. Petrovsky, S. K. Tam, V. Brusic, G. Russ, L. Socha, and V. B. Bajic,"Use of Artificial Neural Networks in Improving Renal Transplantation Outcomes Outcomes," *Graft*, Vol. 5, Issue 1, pp. 6-13, 2002.

[18] S.K. Agarwal, R.K. Srivastava,, "Chronic Kidney Disease in India: Challenges and Solutions", *Nephron Clin Pract* 2009;111:c197-c203.

[19] R.AWolfe, K.P.Mcmullough et.al. "Calculating life years from Transplant Methods for kidney and kidney pancreas candidates", American Journal of Transplantation 2008; 8 (Part 2): 997–1011.

[20] Jiakai Li, Gursel Serpen, Steven Selman et.al. "Bayes Net Classifiers for Prediction of Renal Graft Status and Survival Period", World Academy of Science, Engineering and Technology 39 2010.

[21] N. Hoot, "Models to Predict Survival After Liver Transplantation," M.S. thesis, Vanderbilt University , Nashville, Tennessee, USA, 2005.

[22] J.-H. Ahn, J.-W. Kwon and Y.-S. Lee, "Prediction of 1-year Graft Survival Rates in Kidney Transplantation: A Bayesian Network Model," in *Proc. INFORMS & KORMS*, Seoul, Korea, 2000, pp. 505-513.