Modelling access to renal transplantation waiting list in a French healthcare network using a Bayesian method

Sahar BAYAT\textsuperscript{a}, Marc CUGGIA\textsuperscript{a}, Michel KESSLER\textsuperscript{b}, Serge BRIANÇON\textsuperscript{c}, Pierre LE BEUX\textsuperscript{\textcircled{c}}, Luc FRIMAT\textsuperscript{\textcircled{c}}

\textsuperscript{a}EA 3888, Université Rennes 1, IFR 140, Rennes, France
\textsuperscript{b}Service de néphrologie, Hôpitaux du Brabois, Nancy, France
\textsuperscript{c}EA 4003, Nancy Université, France

\textbf{Abstract}: Evaluation of adult candidates for kidney transplantation diverges from one centre to another. Our purpose was to assess the suitability of Bayesian method for describing the factors associated to registration on the waiting list in a French healthcare network. We have found no published paper using Bayesian method in this domain. Eight hundred and nine patients starting renal replacement therapy were included in the analysis. The data were extracted from the information system of the healthcare network. We performed conventional statistical analysis and data mining analysis using mainly Bayesian networks. The Bayesian model showed that the probability of registration on the waiting list is associated to age, cardiovascular disease, diabetes, serum albumin level, respiratory disease, physical impairment, follow-up in the department performing transplantation and past history of malignancy. These results are similar to conventional statistical method. The comparison between conventional analysis and data mining analysis showed us the contribution of the data mining method for sorting variables and having a global view of the variables' associations. Moreover these approaches constitute an essential step toward a decisional information system for healthcare networks.

\textbf{Keywords}: Data analysis-extraction tools; Decision support; Organization; Bayesian network; Healthcare network; renal transplant waiting list

\textbf{Introduction}

Incidence and prevalence of end-stage renal disease (ESRD) requiring renal replacement therapy (RRT), i.e. hemodialysis, peritoneal dialysis or kidney transplantation, are still increasing [1]. Kidney transplantation is associated with longer survival and lower long-term cost [2]. But, given the graft shortage, transplantation with a cadaveric donor kidney before start of RRT is not commonly achievable. On the other hand, all patients are not equally suited for transplantation and there is a body of evidence showing that selection criteria of the potential transplant recipient diverges from one centre to another [3]. Ideally, placement on the waiting list should be based solely on medical factors, in accordance with medical guidelines. However, previous
studies showed that blacks, Hispanics, women and elderly were less likely to receive a renal transplant in some countries. Moreover, distance from transplantation department, and private ownership of dialysis facilities have been associated with poor access to kidney transplant waiting lists [4-6]. The network of care NEPHROLOR was set up in the French administrative region Lorraine to improve quality of care of all patients with chronic kidney disease (CKD). NEPHROLOR has set up a Regional Health Information System in order to collect patients’ data at first RRT and during the follow up. These data are potentially useful for optimizing organization in NEPHROLOR. Indeed, Access to renal transplant waiting list is one key point to achieve optimal care. The purpose of the present study was to evaluate the feasibility and suitability of Bayesian networks for detecting factors associated with registration on the waiting list. These methods have been already used to predict survival of liver[7] or lung [8] transplant patients but we have found no published study using Bayesian approach to evaluate access to renal transplantation.

1. Material and method

1.1. Organisation of care for CKD in Lorraine region

The NEPHROLOR network of care was set up, in Lorraine, in 2002. It combines the public and private for-profit facilities, operating dialysis units in Lorraine. In Lorraine as in most French regions, renal transplantation is performed at the University hospital nephrology department which is the only transplant centre of the region.

1.2. Study population

Since June 1997, all adult patients living in Lorraine and starting RRT (dialysis or pre-emptive transplantation) in a NEPHROLOR facility were progressively registered in a regional database [1, 4]. For the present study, we included the incident patients between July 1, 1997 and June 30, 2003. The patients having missing data were not included. In order to identify all ESRD patients living in Lorraine and placed on the waiting list, the list of patients registered on the French national waiting list between January 1, 1996 and December 31, 2004 was extracted from the CRISTAL[9] database and the list of transplanted patients from the database of the transplantation department of Lorraine.

1.3. Data collection

For each inclusion of a new patient in the NEPHROLOR database, a standardized form is prospectively filled out at the initiation of RRT [1, 4]. Three categories of variables possibly related to registration on the transplant waiting list were studied. The first included social and demographic data: age, sex and residence at first RRT. The distance between the patient's residence and the department performing transplantation was calculated in kilometers. In France, legal regulations prohibit considering ethnic differences in the French ESRD registry. However, we know that, in Lorraine, almost all patients were Caucasians, with a significant proportion of natives from Italy, and North Africa. In the French ESRD registry, descriptions about income
or education were not available [1]. The second category included clinical, anthropometric and biological data at first RRT: existence of diabetes, cardiovascular disease (coronary artery disease, peripheral vascular disease, congestive heart failure and cerebrovascular disease), respiratory disease, hepatic disease, psychiatric disorder (severe depression or other psychiatric disorder) and past history of malignancy. Patients who were confined to a wheelchair or were bedridden were considered to have physical impairment of ambulation. Body Mass Index was categorized in <20 kg/m², 20-24.99, ≥25, hemoglobin in <11 g/dl, ≥11, and serum albumin in <3 g/dl, 3-3.49, ≥3.5. The third category included data related to medical follow up in the NEPHROLOR network: ownership of nephrology facility where the first RRT was performed: public, or private. We also took into account the effect of a medical follow-up in the department performing transplantation versus 12 other facilities without transplantation.

1.4. Statistical analysis

As the oldest patient registered on the waiting list was 78 years old, we decided to exclude patients older than 80 years on the first day of RRT from statistical analysis. A conventional statistical analysis was done using descriptive method and statistical tests. Impact of baseline characteristics on registration on the waiting list was univariately analyzed using the χ² test. Subsequently, all variables univariately significantly associated with registration with p of 0.10 or less were presented stepwise to a multiple logistic regression model to assess their independent value for registration. Within each step, significant risk factors were selected with a forward strategy using the likelihood ratio statistic, with p of 0.05 on the criterion level of selection. These analyses were performed with SAS software (version 9.1; SAS Institute Inc, cary, NC). We performed also a Bayesian network [10] analysis with BAYESIALAB Software (version 3.3; Bayesia SA). A Bayesian network describes a system of interest by specifying relationships of conditional dependence between its variables. These relationships are represented by a graph, in which, the nodes represent the variables and the arcs represent relationship between variables. The thickness of the arc between 2 variables is proportional to the contribution of their relationship in characterization of the network. The brightness of square represented in a node is proportional to the contribution of this node in the characterization of the target. The Bayes’ rules can be described with the following formula:

\[
P(R = r | e) = \frac{P(e|R = r)P(R = r)}{P(e)}
\]

Where P(R = r | e) denotes the probability that random variable R has value r given evidence e. In these models, a simple representation of a complex problem is possible. The analysis were performed in three steps: learning step (SopLEQ method)[11-13], analysis of associations, inference in order to characterize registration on the waiting list (target node). We used the parent-sons model to characterize the target node. These models allow us to deduce which variables are directly related with the target node.
2. Results

2.1. Patients’ characteristics

Eight hundred and nine patients were included in the study. Their mean age was 62.1 ± 14.2 years; 482 (59.6%) were male, 279 (34.5%) had diabetes, 358 (44.2%) cardiovascular disease and 90 (11.1%) respiratory disease. One hundred and fourteen (14.1%) had a past history of malignancy, 158 (19.5%) physical impairment and 48 (5.9%) psychiatric disease. Two hundred and twelve (26.2%) of them were registered on the transplant waiting list.

2.2. Logistic regression results

Table 1 presents factors associated with registration on the waiting list using a logistic regression model. The younger the patient, the more frequent was the registration on the waiting list. For example, compared with a patient older than 65 years, a patient younger than 44 years was 39 times more likely to be placed on the waiting list. Among co-morbidities, psychiatric disorders, cardiovascular disease, diabetes, past history of malignancy, and respiratory disease were independent factors associated with non registration. Patients with serum albumin greater than 3.5 g/dl were 2.7 times more likely to be placed on the waiting list, compared to those having serum albumin less than 3 g/dl. Patients followed in the nephrology department performing transplantation were 85% more likely to be registered.

Table 1. Factors associated with registration on the waiting list (multivariate logistic regression analysis, n=809 patients). For the co morbidities, presence of pathology represents the reference modality.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age categories</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 65 years</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>55 – 64</td>
<td>11.50</td>
<td>6.10 – 21.68</td>
</tr>
<tr>
<td>45 – 54</td>
<td>33.00</td>
<td>16.62 – 65.53</td>
</tr>
<tr>
<td>18 – 44</td>
<td>38.99</td>
<td>18.57 – 81.89</td>
</tr>
<tr>
<td>Psychiatric disorder</td>
<td>7.37</td>
<td>2.23 – 24.37</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>3.26</td>
<td>1.89 – 5.61</td>
</tr>
<tr>
<td>Serum Albumin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 3 g/dl</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3 – 3.49</td>
<td>1.54</td>
<td>0.76 – 3.11</td>
</tr>
<tr>
<td>≥ 3.5</td>
<td>2.74</td>
<td>1.46 – 5.16</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2.97</td>
<td>1.67 – 5.28</td>
</tr>
<tr>
<td>Past history of malignancy</td>
<td>2.73</td>
<td>1.29 – 5.79</td>
</tr>
<tr>
<td>Respiratory disease</td>
<td>3.50</td>
<td>1.16 – 10.56</td>
</tr>
<tr>
<td>Medical follow-up in the department</td>
<td></td>
<td></td>
</tr>
<tr>
<td>performing transplantation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.85</td>
<td>1.15 – 2.98</td>
</tr>
</tbody>
</table>
2.3. Bayesian model results

For the learning step, we used the SopLEQ method to build the network (Fig. 2). The graph of the variables showed that the probability of being registered on the waiting list (target node) is directly associated with 6 variables: age, diabetes, cardiovascular disease, past history of malignancy, serum albumin level and follow up in the department performing transplantation. The Parent-Son relationship (Fig. 3) demonstrated that one node appears to be highly related to the target node. This node is the ‘Age’ node. Diabetes, cardiovascular and respiratory diseases, physical impairment, serum albumin level, past history of malignancy and medical follow-up in the department performing transplantation were also related to registration on the waiting list. Relative contribution of the variables to characterization of the target node were as following: Age: 1.0000; Cardiovascular disease: 0.3387; Diabetes: 0.2016; Albumin: 0.0837; Respiratory disease: 0.0762; Physical impairment: 0.0504; Follow up in transplantation center: 0.0423; Past history of malignancy: 0.0395.

Figure 2 - Bayesian Network (SopLEQ learning method)

Figure 3 – Parent-Son relationship
3. Discussion

The results of the conventional analyses and the Bayesian methods were similar. It appears that in Lorraine, access to transplant waiting list is primarily associated with medical determinants. Differences due to sex, ownership of dialysis facility, and distance to transplant centre that have been identified in other studies [5, 6] were not found in Lorraine region. Conventional statistical analyses enable the clinician’s hypothesis to be tested and this can produce results with statistical and clinical significance notions. Conversely, data mining methods allow a complex universe to be untangled. While, data mining methods represent a global view of relationships between variables instantaneously, it is necessary to spend much more time to obtain the same conclusion with the statistical analysis. Moreover, the relationships between variables are provided without prior hypothesis. The sample of data ‘speaks for itself’ and illustrates alone which variables are associated with the others. The inference power offered by Bayesian network is well accepted by physicians and comprehensive. Furthermore, it allows the experts to imagine other relationships or to test impact of events for a diagnosis or a disease.

4. Conclusion

Bayesian networks allowed us to describe the determinants of access to renal transplant waiting list. Furthermore, compared to conventional statistical method, it provides a global view of variables associations. Data mining constitutes an essential step toward a decisional information system for healthcare networks. Indeed, the result of our study can be used for optimizing renal transplant registration process in NEPHROLOR.

References