

**RAFAEL FÁBIO MACIEL**

**APLICAÇÃO DE REDES NEURAIAS ARTIFICIAIS EM TRANSPLANTES RENAIIS:  
CLASSIFICAÇÃO DE NEFROTOXICIDADE E REJEIÇÃO CELULAR AGUDA.**

**SÃO PAULO**

**2010**

**RAFAEL FÁBIO MACIEL**

**APLICAÇÃO DE REDES NEURAS ARTIFICIAIS EM TRANSPLANTES RENAI:  
CLASSIFICAÇÃO DE NEFROTOXICIDADE E REJEIÇÃO CELULAR AGUDA.**

Tese de mestrado apresentada ao Programa de Pós-Graduação em Saúde Coletiva da Universidade Federal de São Paulo, como parte dos requisitos para a obtenção do título de Mestre em Saúde Coletiva.

ORIENTADOR: Professor Doutor Ivan Torres Pisa

CO-ORIENTADOR: Professor Doutor Renato Glauco de Souza Rodrigues

SÃO PAULO

2010

**Maciel, Rafael Fábio**

Aplicação de redes neurais artificiais em transplantes renais: classificação de nefrotoxicidade e rejeição celular aguda.

Rafael Fábio Maciel – São Paulo, 2010.

Tese (mestrado) – Universidade Federal de São Paulo.  
Programa de Pós-Graduação em Saúde Coletiva.

Título em inglês: Application of Artificial Neural Networks in Kidney Transplantation: Classification of Nephrotoxicity and Acute Cellular Rejection.

1. Transplante Renal.
2. Reconhecimento de Padrão.
3. Redes Neurais.
4. Monitoramento de Medicamentos/Efeitos Adversos.

**RAFAEL FÁBIO MACIEL**

**APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS EM TRANSPLANTES RENAIIS:  
CLASSIFICAÇÃO DE NEFROTOXICIDADE E REJEIÇÃO CELULAR AGUDA.**

Esta Tese foi julgada adequada para a obtenção do título de Mestre em Saúde Coletiva e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Banca Examinadora:

Prof. Dr. Frederico Cavalcanti Castelo Branco, UFPE  
Doutor pela USP – São Paulo, Brasil

Prof. Dr. Domingos Alves, FMRP  
Doutor pela USP – Ribeirão Preto, Brasil

Prof. Dr. Daniel Sigulem, UNIFESP  
Doutor pela UNIFESP – São Paulo, Brasil

Coordenador do PPGSC: Prof. Dr. Luiz Roberto Ramos

São Paulo, outubro de 2010.

## **DEDICATÓRIA**

Dedico este trabalho aos meus professores e a meu orientador, aos alunos do Programa de Pós-Graduação em Informática em Saúde - UNIFESP e, em especial, a minha esposa Maiara Maria Camerini Maciel e aos meus filhos Raphael Camerini Maciel e Gabriella Camerini Maciel pela dedicação e apoio em todos os momentos difíceis.

## **AGRADECIMENTOS**

Agradeço ao Programa de Pós-Graduação em Saúde Coletiva, UNIFESP – EPM pela oportunidade de realização de trabalhos em minha área de pesquisa.

Aos colegas do Programa de Pós-graduação em Informática em Saúde, UNIFESP pelo auxílio nas tarefas desenvolvidas durante o curso e apoio na revisão deste trabalho.

Em especial, ao meu orientador Ivan Torres Pisa pela dedicação e paciência dedicadas a este pesquisador.

## **LISTA DE ILUSTRAÇÕES**

Figura 1 – Metodologia de classificação: fases da classificação para definir se o paciente apresenta nefrotoxicidade ou RCA.

Figura 2 – Problema de classificação não separável linearmente com uma solução global.

Figura 3 – Problema de classificação não separável linearmente com 8 subsoluções triviais ou separáveis linearmente.

Figura 4 – Curva ROC dos Classificadores na Fase 2

## **LISTA DE TABELAS**

Tabela 1 – Algoritmos utilizados e suas respectivas implementações.

Tabela 2 – Odds ratio bruto e ajustado para todas as variáveis independentes, considerando o desfecho de rejeição celular aguda.

Tabela 3 – Odds ratio bruto e ajustado para todas as variáveis independentes, considerando o desfecho de nefrotoxicidade.

Tabela 4 – Modelos selecionados na fase 1.

Tabela 5 – Modelos selecionados na fase 2.



## **LISTA DE ABREVIATURAS**

ABTO: Associação Brasileira de Transplante de Órgãos

ABNT: Associação Brasileira de Normas Técnicas

IC: Inibidor de Calcineurina

HLA: Human Leukocyte Antigens

IRC: Insuficiência Renal Crônica

PPGSC: Programa de Pós-Graduação em Saúde Coletiva

RNA: Rede Neural Artificial

RCA: Rejeição Celular Aguda

TA: Tensão Arterial

TCLE: Termo de Consentimento Livre e Esclarecido

WEKA: Waikato Environment for Knowledge Analysis

SVM: Support Vector Machines

## LISTA DE PUBLICAÇÕES

A.D. Hummel, R.F. Maciel, R.G.S. Rodrigues, and I.T. Pisa. Application of Artificial Neural Networks in Renal Transplantation: Classification of Nephrotoxicity and Acute Cellular Rejection Episodes. **Transplantation Proceedings (ISSN 0041-1345)**, 2010; 42:471-472.

A.D. Hummel, R.F. Maciel, R.G.S. Rodrigues, and I.T. Pisa. Aplicação de técnicas computacionais em bases de dados de transplante: revisão de artigos publicados no biênio 2007-2008. **JBT J Bras Transpl (ISSN 1678-3387)**. 2009; 12:1045-1048.

Rafael Fabio Maciel, Anderson Diniz Hummel, Frederico Molina Cohrs, Msc. Felipe Mancini, Alex Jacoud Falcão, Fabio de Oliveira Teixeira, Thiago Martini da Costa, Fernando Sequeira de Sousa, Domingos Alves, Rafael Miranda, Ivan Torres Pisa. Aplicação de técnicas de inteligência artificial em transplantes renais: classificadores automáticos para nefrotoxicidade e rejeição celular aguda. **Journal of Health Informatics (ISSN 1446-4381)**. 2010; Vol. 2, No 3.

## **PARTICIPAÇÃO EM CONGRESSOS**

**XI Congresso Brasileiro de Transplantes, VIII Congresso Luso-Brasileiro de Transplantes, X Encontro de Enfermagem em Transplantes, II Encontro Multidisciplinar em Transplantes, Forum de Histocompatibilidade da ABH, Recife, 13 a 16 de outubro de 2009.**

Aplicação de Redes neurais Artificiais em Transplantes Renais: Classificação de Nefrotoxicidade e Rejeição Celular Aguda. Anderson Diniz Hummel, Rafael Fabio Maciel, Renato Glauco de Souza Rodrigues, Ivan Torres Pisa.

**XXIII International Congress of The Transplantation Society / August 15-19, 2010 / Vancouver, CA.**

Application of Artificial Neural Networks in Kidney Transplantation: Classification of Nephrotoxicity and Acute Cellular Rejection. Rafael Fabio Maciel, Anderson Diniz Hummel, Frederico Molina Cohrs, Felipe Mancini, Domingos Alves, Renato Glauco de Souza Rodrigues, Rafael Miranda, Ivan Torres Pisa.

**VI Congresso Brasileiro de Transplantes de Fígado e Pâncreas, Porto de Galinhas, 07 a 09 de outubro de 2010.**

Aplicação de Técnicas de Inteligência Artificial e de Seleção de Atributos no Pós-Transplante Renal: Classificação de Nefrotoxicidade e Rejeição Celular Aguda. Rafael Fabio Maciel, Anderson Diniz Hummel, Fernando Sequeira Sousa, Thiago Martini da Costa, Fabio Teixeira, Alex Esteves Jaccoud Falcão, Renato Glauco de Souza Rodrigues, Rafael Miranda, Domingos Alves, Ivan Torres Pisa.

Aplicação de Técnicas de Inteligência Artificial em Transplantes Renais: Triagem de biopsia em pacientes com suspeita de Nefrotoxicidade ou Rejeição Celular Aguda. Anderson Diniz Hummel, Rafael Fabio Maciel, Fernando Sequeira Sousa, Felipe Mancini, Roberto Baptista, Frederico Molina Cohrs, Renato Glauco de Souza Rodrigues, Rafael Miranda, Domingos Alves, Ivan Torres Pisa.

Aplicação de Técnicas Inteligentes em Bases de dados de Transplantes: uma revisão de artigos publicados em 2009 e 2010. Fernando Sequeira Sousa, Anderson Diniz Hummel, Rafael Fabio Maciel, Frederico Molina Cohrs, Alex Esteves Jaccoud Falcão, Fabio Teixeira, Roberto Silva Baptista, Thiago Martini Costa, Domingos Alves, Ivan Torres Pisa.

## ANEXOS

1. CEP 1677/08/UNIFESP
2. TCLE
3. Base de Dados de Treinamento dos Algoritmos de Inteligência Artificial para Complicações/Não Complicações
4. Base de Dados de Treinamento dos Algoritmos de Inteligência Artificial para Nefrotoxicidade/RCA

## 5. RESUMO

**INTRODUÇÃO:** Complicações associadas ao transplante de rim e imunossupressão podem ser prevenidas ou tratadas efetivamente quando são diagnosticadas em fases iniciais com o monitoramento pós-transplante. **OBJETIVO:** Apresentar resultados da comparação de classificadores automáticos de padrões utilizando diferentes técnicas de inteligência artificial para prever eventos de nefrotoxicidade e rejeição celular aguda (RCA) com até um ano após o transplante renal. **MÉTODOS:** Foram analisados dados de 135 prontuários de pacientes submetidos a transplante renal no período de fevereiro de 2006 a junho de 2010 e realizados testes estatísticos de incidência e regressão logística nas variáveis em relação à nefrotoxicidade e RCA. Foram utilizados diferentes classificadores: redes neurais, support vector machines (SVM), árvores de decisão, inferência bayesiana e vizinhos mais próximos, com o objetivo de prever RCA e nefrotoxicidade. Os classificadores foram avaliados segundo o valor de sensibilidade, especificidade e área sob a curva ROC (AUC). **RESULTADOS:** A incidência de RCA foi de 31,0% e de nefrotoxicidade de 26,9%. A técnica que apresentou o melhor valor de sensibilidade para prever se há nefrotoxicidade ou RCA foi a SVM (algoritmo LIBSVM) com sensibilidade 0,87 (taxa de acerto 79,86%; especificidade 0,70; AUC 0,79). A técnica que apresentou o melhor valor de AUC para prever nefrotoxicidade ou RCA foi a de inferência bayesiana (algoritmo *NaiveBayes*) com AUC 0,80 (taxa de acerto 75,92%). **CONCLUSÃO:** Os resultados são animadores, com taxas de erro e acerto condizentes com a determinação de rejeição celular aguda e nefrotoxicidade.

**Palavras chaves:** Transplante renal, reconhecimento de padrão, redes neurais, monitoramento de medicamentos/efeitos adversos.

## ABSTRACT

**BACKGROUND:** Complications associated with kidney transplantation and immunosuppression can be prevented or treated effectively if diagnosed in early stages with monitoring post-transplant. **OBJECTIVE:** To present the results of comparison of automatic pattern classifiers using different techniques of artificial intelligence to predict events of nephrotoxicity and acute cellular rejection (RCA), with up to one year of renal transplantation. **METHODS:** Statistical tests were performed on the prevalence and linear regression in variables regarding nephrotoxicity and RCA. We used different classifiers (neural networks, support vector machines (SVM), decision trees, Bayesian inference, and closest neighbors) in order to provide RCA and nephrotoxicity. The classifiers were evaluated according to the value of sensitivity, specificity and area under ROC curve (AUC). **RESULTS:** The prevalence of acute cellular rejection was 31.0% and 26.9% of nephrotoxicity. The technique had the highest sensitivity value prediction for the submission to the transplanted kidney biopsy was SVM (LIBSVM algorithm) with sensitivity rates of 0.87 (accuracy rate 79.86; specificity 0.70; AUC 0.79). The technique had the highest AUC for predicting nephrotoxicity and RCA was bayesian inference (NaiveBayes), with AUC rates of 0.8 (accuracy rate 75.92). **CONCLUSION:** The results are encouraging, with rates of trial and error consistent with the determination of acute cellular rejection and nephrotoxicity.

**Key words: Kidney transplantation, pattern recognition, neural networks, monitoring of drug/adverse effects.**

# SUMÁRIO

<b>INTRODUÇÃO</b>	<b>13</b>
<b>OBJETIVOS</b>	<b>18</b>
<b>MATERIAL E MÉTODOS</b>	<b>19</b>
<b>3.1 COLETA DE DADOS</b>	<b>19</b>
<b>3.3 CLASSIFICADORES DE PADRÕES</b>	<b>21</b>
<b>3.4 FERRAMENTAS DE MINERAÇÃO DE DADOS</b>	<b>26</b>
<b>3.5 DESENVOLVIMENTO DOS CLASSIFICADORES</b>	<b>27</b>
<b>3.6 AVALIAÇÃO DOS CLASSIFICADORES</b>	<b>27</b>
<b>RESULTADOS</b>	<b>28</b>
<b>DISCUSSÃO</b>	<b>32</b>
<b>CONCLUSÃO</b>	<b>34</b>



## INTRODUÇÃO

A Insuficiência Renal Crônica (IRC) representa um importante problema médico e de saúde pública. No Brasil, a prevalência de pacientes mantidos em programa crônico de diálise mais que dobrou nos últimos oito anos. A incidência de novos pacientes cresce cerca de 8% ao ano e representa um custo aproximado de um bilhão e meio de reais ao ano com terapia renal substitutiva (Romão Junior, 2004).

O impacto da IRC no indivíduo e na família, independente do aspecto da saúde física e mental, leva comumente à desagregação das relações familiares. O paciente com IRC, filho ou conjugue, passa a dominar as ações da família, pois necessita de cuidados permanentes, de remoções três vezes por semana para os serviços de hemodiálise, que, geralmente, são distantes de seu domicílio. De acordo com sua vulnerabilidade, necessita de dedicação quase que exclusiva de um cuidador, que, na maioria das vezes, são os pais ou o conjugue. E, assim, perde sua cidadania, por não poder trabalhar, estudar, viajar ou mesmo formar novo núcleo familiar, muitas vezes.

O transplante renal não é a cura da IRC. É um tratamento que resgata a cidadania, desatrela o indivíduo de uma máquina de hemodiálise e traz resultados positivos na família.

O transplante renal representa uma importante opção terapêutica do ponto de vista médico e socioeconômico e está indicado para os pacientes com IRC terminal em diálise ou na fase pré-dialítica. Consiste em implantar um rim obtido de um doador vivo ou de um doador falecido. Porém, o transplante renal não representa a cura da IRC e sim uma das modalidades terapêutica substitutiva renal, não isenta de complicações.

Existe uma probabilidade específica para a ocorrência de complicações de acordo com o período pós-transplante e o primeiro ano é o mais crítico. Nos três primeiros meses, a preocupação é com rejeição e perda do enxerto. A partir do quarto mês até o primeiro

aniversário, são as infecções e a nefrotoxicidade provocadas pelo uso dos imunossupressores, especialmente os inibidores de calcineurina (IC), além de episódios de rejeição celular aguda (RCA) (Kasiske et al., 2000). A rejeição ao enxerto continua sendo um problema significativo na perda precoce do órgão enxertado e fator relevante que compromete a sua longevidade. Entre as drogas imunossupressoras responsáveis pelo controle da rejeição do enxerto, as mais usadas são os inibidores de calcineurina, porém o custo é a nefrotoxicidade que contribui como uma das causas de perda do rim transplantado no período tardio.

O monitoramento clínico e laboratorial no acompanhamento ambulatorial do paciente transplantado de rim é uma ferramenta importante na prevenção e na administração das complicações associadas ao transplante e à terapia imunossupressora. Complicações associadas ao transplante de rim e imunossupressão podem ser prevenidas ou tratadas efetivamente quando são diagnosticadas em fases iniciais (Hariharan, 2006).

Na última década, houve significativa redução na incidência de RCA em transplantes de rim, de 14-20% para valores em torno de 4-10%, com novos protocolos de imunossupressão (Nashan, 2005). Porém, os efeitos adversos dos medicamentos trazem grandes impactos na sobrevida dos transplantes, como um dos fatores não imunológicos relacionados com as lesões que contribuem com a nefropatia crônica do enxerto, obrigando o monitoramento constante destas drogas e acompanhamento dos efeitos adversos (Johnson et al., 2000; Ciancio et al., 2004).

Tradicionalmente, têm-se utilizado métodos estatísticos para prever resultados em Medicina. No entanto, técnicas estatísticas têm baixo poder para resolver problemas de alta complexidade (Hummel, 2010). Recentemente, a atenção se voltou para uma variedade de técnicas de modelagem computacional, como rede neural artificial (RNA) (Lisboa PJ et al., 2002; Kordylewski et al., 2001). Uma importante aplicação é o sistema para alerta automático

de riscos de saúde e procedimentos necessários para cada caso a partir dos dados armazenados sobre o paciente (Dimitrov, 2003).

Vantagens oferecidas pelas RNAs em relação a outras técnicas de classificação incluem robustez, adaptabilidade, capacidade de auto-organização, alto grau de paralelismo distribuído e facilidade de implementação em *hardware*.

Portanto, RNAs podem ser utilizadas em situações nas quais a informação sobre um dado assunto não é totalmente acessível ou completa, ou, ainda, em casos em que o problema sendo tratado não é totalmente entendido do ponto de vista numérico.

RNAs são amplamente utilizadas na área médica como ferramenta padrão de análise. Executando-se uma pesquisa no site da PubMed <<http://www.ncbi.nlm.nih.gov>> com as palavras “*neural networks*” e “*medical*” foram encontrados 1.475 artigos referenciando a utilização de RNAs artificiais na Medicina. Na área de transplante renal podemos destacar trabalhos relacionados à predição do funcionamento do enxerto renal (Shoskes et al., 1998), no diagnóstico de episódios de RCA (Furness et al., 1999), predição de rejeição crônica (Simic-Ogrizovic et al., 1999) e na dosagem de imunossupressores (Camps-Valls et al., 2003). Este estudo busca um sistema com base em RNAs que seja capaz de generalizar conhecimentos sobre o diagnóstico diferencial das complicações no primeiro ano de transplante renal.

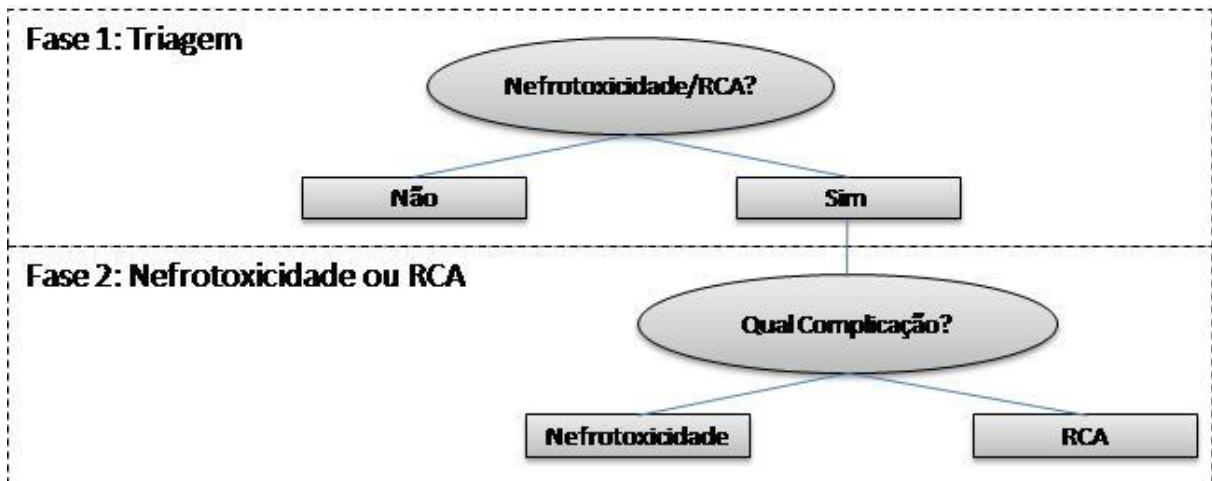
## **OBJETIVOS**

Apresentar resultados da comparação de classificadores automáticos de padrões utilizando diferentes técnicas de inteligência artificial para prever eventos de nefrotoxicidade e RCA a partir de dados de prontuários de pacientes no primeiro ano do transplante renal.

## MATERIAIS E MÉTODOS

A pesquisa foi realizada em três etapas: coleta de dados, análise estatística dos dados e desenvolvimento dos classificadores. A etapa do desenvolvimento dos classificadores foi dividida em duas fases: triagem (Fase 1) e nefrotoxicidade e RCA (Fase 2).

A Fase 1 tem como objetivo a triagem dos casos suspeitos de nefrotoxicidade e RCA. A Fase 2 é específica para prever qual das complicações o paciente apresenta, e, obrigatoriamente, prediz uma delas. A Figura 1 apresenta esta divisão em fases do processo de classificação.



**Figura 1** - Metodologia de classificação: fases da classificação para definir se o paciente apresenta nefrotoxicidade ou RCA.

O estudo foi conduzido na Universidade Federal de São Paulo (UNIFESP) e aprovado na Comissão de Ética em Pesquisa, processo no. CEP 1677/08 (Anexo 1) a partir de dados de pacientes transplantados do Centro de Transplante da Casa de Saúde Santa Efigênia, Caruaru PE. Os resultados desta investigação não influenciaram a conduta clínica realizada.

### 3.1 COLETA DE DADOS

A coleta de dados nos prontuários dos pacientes transplantados foi do período de março de 2006 a dezembro de 2009. No protocolo definido, foram avaliados os prontuários de 135

pacientes em pós-transplante renal, em acompanhamento durante o primeiro ano de transplantados. Como critérios de inclusão, foram considerados pacientes com suspeita clínica de diagnósticos de RCA e/ou nefrotoxicidade, e que depois se submeteram à biópsia renal. Foram excluídos pacientes que não fazem uso de inibidor de calcineurina. Por fim, o protocolo possibilitou incluir dados de 93 pacientes para estudo. Todos os pacientes assinaram o Termo de Consentimento Livre e Esclarecido – TCLE (Anexo 2) – autorizando o uso dos dados, sem identificação, para este estudo.

Os dados coletados incluem as seguintes variáveis: tempo da consulta pós-transplante, dosagem de inibidor de calcineurina, indução à imunossupressão com anticorpos, função inicial (diurese ou não), tipo de doador (vivo ou cadáver), sorologia para citomegalovírus no receptor, aumento ou queda da diurese no momento do diagnóstico, aumento de temperatura, edema, tremores, dosagem da ureia, dosagem de creatinina, glicemia, contagem de leucócitos, contagem de linfócitos, contagem de plaquetas, TA média e histocompatibilidade HLA (*Human Leukocyte Antigens*).

### **3.2 ANÁLISE ESTATÍSTICA DOS DADOS**

Os dados foram analisados no programa de estatística SPSS, 15.1. Inicialmente, utilizou-se, para as variáveis contínuas, o teste de Kolmogorov-Smirnov (Neal DK, 1996) para verificação da normalidade da distribuição. Como as variáveis apresentaram distribuição normal, foram adotados procedimentos paramétricos. Na descrição das proporções, a distribuição binomial foi aproximada à distribuição normal pelo intervalo de confiança de 95%. Com o objetivo de observar o comportamento das variáveis e confrontar com a construção dos classificadores, foram realizados testes estatísticos de prevalência e regressão linear nas variáveis em relação à nefrotoxicidade e RCA. Nos testes inferenciais, as

proporções foram comparadas pelo teste de chi-quadrado ou teste exato de Fisher, quando as respostas apresentadas foram inferiores a cinco casos. Foram realizadas duas regressões logísticas binárias, sendo os desfechos: RCA e nefrotoxicidade. Inicialmente, para cada um dos desfechos, foram calculadas as regressões brutas para cada uma das variáveis independentes e os seus respectivos intervalos de confiança. Posteriormente, realizou-se a regressão ajustada para todas as variáveis independentes, com os respectivos intervalos de confiança. Foi adotado um nível de significância de 5%.

### **3.3 CLASSIFICADORES DE PADRÕES**

Existem cinco técnicas de mineração de dados: a classificação, a estimativa, a previsão, a análise de afinidades e a análise de agrupamentos (Carvalho, 2005).

A classificação pode ser entendida como processo de discriminação entre classes ou categorias, em geral, essa tarefa é realizada pelo especialista e serve para tentar entender dados existentes a partir da subdivisão em classes e inferir conhecimento (Fryer, 2003).

Em muitos casos, a separação em classes é simples, como nos casos das classes linearmente separáveis, entretanto, a maioria dos problemas reais de interesse não são linearmente separáveis.

A Figura 2 mostra um caso de separação entre duas classes distintas, por exemplo, nefrotoxicidade (classe O) e rejeição (classe X), em que é possível obter uma solução global para um conjunto de dados não separável linearmente.

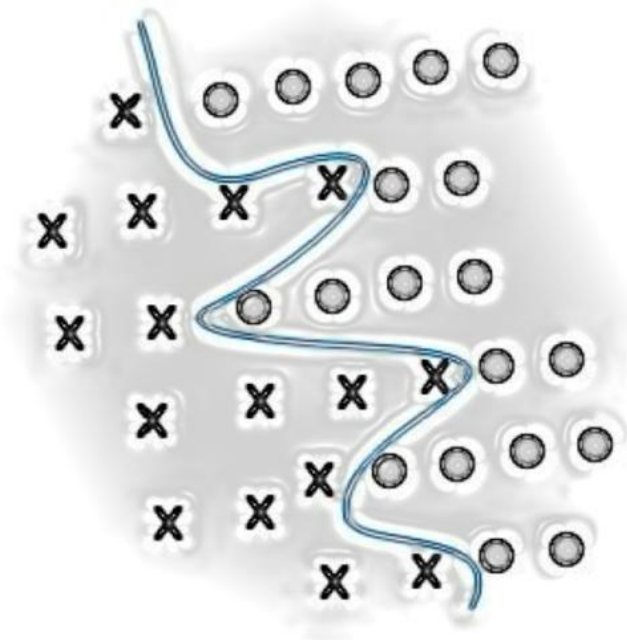


Figura 2 – Problema de classificação não separável linearmente com uma solução global.

A Figura 3 mostra o mesmo problema sendo abordado através da subdivisão em problemas triviais mais simples.

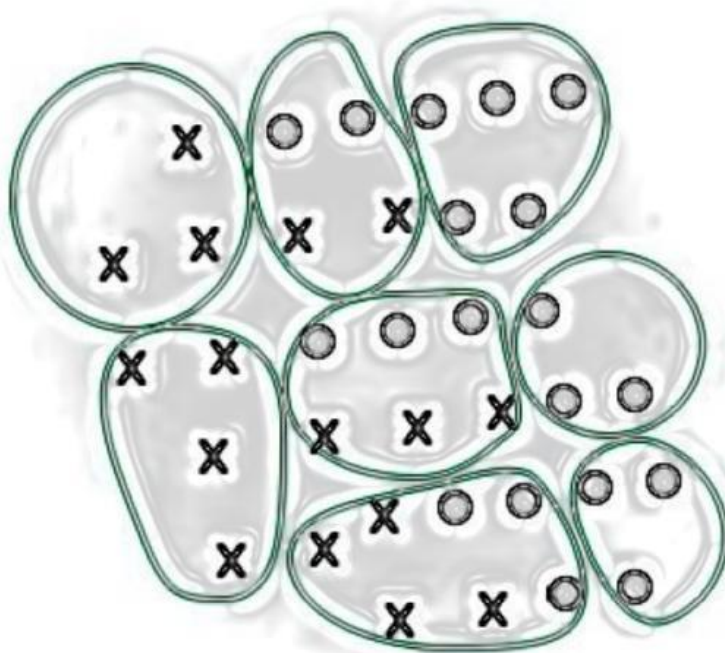


Figura 3 - Problema de classificação não separável linearmente com 8 subsoluções triviais ou separáveis linearmente.



Em geral, o problema de classificação é abordado por diversas técnicas de inteligência artificial para geração de classificadores. Neste trabalho, as técnicas estudadas foram: Redes Neurais Artificiais (RNAs) (Fryer, 2003), *Support Vector Machines* (SVM) (Mcgarry, 2001), Árvores de Decisão (Banerjee, 2003), Inferência Bayesiana (Berner, 1999) e vizinhos mais próximos (Bouchon-Meunier, 2000).

RNAs são modelos matemáticos, tratados computacionalmente devido à sua complexidade, inspirados em redes neurais biológicas (Haykin, 1999) cujo resultado é determinado pelas conexões entre os elementos de transformação (nós) e os parâmetros de entradas (Mohamad, 2005). RNAs são sistemas de equações que imitam a topologia dos neurônios no cérebro, podendo ser classificadas quanto à forma de treinamento, em RNAs supervisionadas e não supervisionadas. O treinamento se dá a partir de um conjunto de dados representativo do problema para que ela possa prever um determinado comportamento (Banerjee, 2003; Haykin, 1999; Mohamad, 2005).

*Support Vector Machine* (SVM) é um método de aprendizagem baseado em teorias de aprendizagem estatísticas e tem sido aplicado recentemente em problemas como de classificação em diversas áreas como Engenharia, Medicina e outras, tornando-se nos, últimos anos, uma das ferramentas padrão para aprendizagem de máquinas.

O método SVM consiste em uma abordagem para classificação binária (duas classes) ótima, tendo sido estendida para suportar regressão e problemas de classificação de múltiplas classes. SVM é um caso particular de métodos baseados em *kernel*, em que mapeia vetores de características em um espaço de alta dimensão usando alguma função *kernel*, e, então, para este novo espaço linear, cria-se um hiperplano ótimo separando as classes.

A solução é ótima no sentido de que a margem entre o hiperplano e os mais próximos vetores de características das duas classes é máxima. Os vetores de características que são mais

próximos do hiperplano são chamados de "*support vectors*", significando que as posições dos outros vetores não afetam o hiperplano.

Os fundamentos das Máquinas de Vetor Suporte (SVM) foram desenvolvidos por Vapnik e colaboradores (Vapnik, 2000). A formulação por ele apresentada se baseia no princípio de Minimização do Risco Estrutural (SRM), que tem um desempenho de generalização superior ao tradicional princípio de Minimização do Risco Empírico (ERM), empregado em redes neurais convencionais.

Basicamente, o funcionamento de uma SVM pode ser descrito da seguinte forma: dadas duas classes e um conjunto de pontos que pertencem a essas classes, uma SVM determina o hiperplano que separa os pontos de forma a colocar o maior número de pontos da mesma classe do mesmo lado, enquanto maximiza a distância de cada classe a esse hiperplano. A distância de uma classe a um hiperplano é a menor distância entre ele e os pontos dessa classe, e é chamada de margem de separação. O hiperplano gerado pela SVM é determinado por um subconjunto dos pontos das duas classes, chamado vetores de suporte.

Uma árvore de decisão é um modelo preditivo que utiliza um modelo gráfico para descrever as conclusões (Yuan, 1995). Amplamente utilizadas em algoritmos de classificação, as árvores de decisão são representações simples do conhecimento e um meio eficiente de construir classificadores que predizem classes baseadas nos valores de atributos de um conjunto de dados.

As árvores de decisão consistem de nodos que representam os atributos; de arcos, provenientes destes nodos e que recebem os valores possíveis para estes atributos; e de nodos folha, que representam as diferentes classes de um conjunto de treinamento. Classificação, neste caso, é a construção de uma estrutura de árvore, que pode ser usada para classificar corretamente todos os objetos do conjunto de dados da entrada (Stuart Russel, 2003).

Muitos são os algoritmos de classificação que constroem árvores de decisão. Não há uma forma de determinar qual é o melhor algoritmo, um pode ter melhor desempenho em determinada situação e outro algoritmo pode ser mais eficiente em outros tipos de situações.

Após a construção de uma árvore de decisão, é importante avaliá-la. Esta avaliação é realizada através da utilização de dados que não tenham sido usados no treinamento. Esta estratégia permite estimar como a árvore generaliza os dados e se adapta a novas situações, podendo, também, se estimar a proporção de erros e acertos ocorridos na construção da árvore.

A partir de uma árvore de decisão, é possível derivar regras. As regras são escritas considerando o trajeto do nodo raiz até uma folha da árvore. Estes dois métodos são, geralmente, utilizados em conjunto. Devido ao fato das árvores de decisão tenderem a crescer muito, de acordo com algumas aplicações, elas são, muitas vezes, substituídas pelas regras. Isto acontece em virtude das regras poderem ser facilmente modularizadas. Uma regra pode ser compreendida sem que haja a necessidade de se referenciar outras regras.

Uma árvore de decisão tem a função de particionar recursivamente um conjunto de treinamento, até que cada subconjunto obtido deste particionamento contenha casos de uma única classe. Para atingir esta meta, a técnica de árvores de decisão examina e compara a distribuição de classes durante a construção da árvore. O resultado obtido, após a construção de uma árvore de decisão, são dados organizados de maneira compacta, que são utilizados para classificar novos casos.

A árvore de decisão não assume nenhum modelo estatístico a priori, sendo a divisão do espaço de atributos feita de acordo com as amostras provenientes do treinamento. A Inferência Bayesiana consiste na aplicação do Teorema de Bayes, os quais são a base da estatística convencional.

A informação que se tem sobre uma quantidade de interesse é extraída a partir de um conjunto de medidas que se tem sobre ela. O verdadeiro valor da quantidade de interesse desconhecida é aproximando tentando reduzir este desconhecimento. A intensidade da incerteza a respeito da quantidade de interesse pode assumir diferentes graus. Do ponto de vista Bayesiano, estes diferentes graus de incerteza são representados através de modelos probabilísticos para a quantidade estudada. Neste contexto, é natural que diferentes estudos possam ter diferentes graus de incerteza sobre a grandeza de interesse (especificando modelos distintos). Sendo assim, não existe nenhuma distinção entre quantidades observáveis e os parâmetros de um modelo estatístico, todos são considerados quantidades aleatórias.

O classificador com base em vizinhos mais próximos, ou K-NN ("*K-Nearest Neighbor*" ou K-ésimo vizinho mais próximo), é um método de classificação supervisionado não-paramétrico no qual um dado é associado a uma classe, dependendo do número K de pontos vizinhos que pertençam a essa classe (Duda, R O, 1973), segundo um critério de distância. Em geral, a medida de distância utilizada é a distância Euclidiana.

Uma árvore de decisão é um modelo preditivo que utiliza um modelo gráfico para descrever as conclusões (Yuan, 1995).

### **3.4 FERRAMENTAS DE MINERAÇÃO DE DADOS**

Estão disponíveis, diversas ferramentas para mineração de dados. Nesse trabalho, foram utilizadas as seguintes ferramentas: RapidMiner (Mierswa,2006) e Waikato *Environment for Knowledge Analysis* – WEKA (Witten, 2005). O pacote de software WEKA foi desenvolvido para agregar algoritmos com diferentes abordagens e paradigmas de aprendizado de máquina com o intuito de obter novos conhecimentos de forma indutiva ou dedutiva a partir de uma base de dados por meio de uma interface gráfica amigável. É interessante destacar que, através do WEKA, é possível aplicar um lote de algoritmos e suas configurações a um

conjunto de dados. O WEKA foi desenvolvido utilizando linguagem Java e encontra-se sobre licença GPL (Witten, 2005).

### 3.5 DESENVOLVIMENTO DOS CLASSIFICADORES

Foram gerados classificadores para as fases 1 e 2. O classificador da Fase 1 teve como objetivo a triagem dos casos suspeitos de nefrotoxicidade e RCA, sem especificar o tipo de complicação. O classificador da Fase 2 foi específico para predizer se o paciente apresenta nefrotoxicidade ou RCA.

Foi utilizado o software livre WEKA (Witten, 2005) para treinar os classificadores automáticos e, como método de treinamento e teste, utilizou-se uma validação cruzada com 10 subgrupos (10 fold cross validation) (Haykin, 1999). Foram utilizadas as seguintes implementações para cada um dos algoritmos:

**Tabela 1 - Algoritmos utilizados e suas respectivas implementações.**

Algoritmo	Implementação
Árvore de Decisão	J48
Inferência Bayesiana	Bayesnet, NaivesBayes
Vizinhos mais próximos	IBK
RNAs	MLP
SVM	LIBSVM

### 3.6 AVALIAÇÃO DOS CLASSIFICADORES

Para avaliar e comparar os algoritmos quanto à confiança na classificação, utilizou-se como métrica para seleção do algoritmo a taxa de acerto, a sensibilidade, a especificidade e a área

sob a curva *Receiver Operating Characteristic* (ROC) (AUC) (McGarry K, 2001; Massad E, 2004). A curva ROC correlaciona a sensibilidade com o complemento da especificidade atingido por um classificador por meio da variação do corte de um ou mais parâmetros do teste. A curva ROC é utilizada para escolher o melhor classificador de padrões para uma determinada tarefa (Massad et al., 2004).

## RESULTADOS

A incidência de RCA foi de 31,0% e de nefrotoxicidade de 26,9%. Destacamos as variáveis estatisticamente significativas na Tabela 1 relacionadas à RCA.

**Tabela 2 - Odds Ratio** bruto e ajustado para todas as variáveis independentes, considerando o desfecho de Rejeição Celular Aguda.

Variável	OR bruto	IC95%	Valor p*	OR Ajustado	IC95%	Valor p**
Dosagem de TAC normal	0,87	0,27 – 2,98	0,87	0,39	0,50 – 3,91	0,37
Dosagem de TAC alto	0,15	0,03 – 0,75	0,02	0,02	0,01 – 0,49	0,01
Dosagem de TAC baixo	3,69	1,26 – 10,74	0,01	3,98	1,88 – 9,21	0,00
Indução	1,04	0,36 – 2,79	0,97	1,14	0,22 – 2,64	0,19
Tipo de doador	1,06	0,34– 3,28	0,91	0,78	0,08– 2,48	0,83
Queda da diurese	10,61	2,5 – 45,00	0,00	3,40	2,51 – 4,53	0,02
CMV receptor positivo	1,41	0,46 – 4,28	0,53	3,91	1,22 – 4,99	0,04
Temperatura aumentada	3,71	0,57 – 23,9	0,16	0,18	0,03 – 1,12	0,42
Edema	6,92	1,22 – 39,03	0,02	0,25	0,03 – 2,12	0,54
Tremores	0,98	0,27 – 3,63	0,98	1,81	0,27 – 1,53	0,58
Ureia aumentada	4,17	1,03 – 13,10	0,01	3,21	0,14 – 6,93	0,45
Creatinina aumentada	5,08	1,50 – 17,23	0,0	4,20	0,30 – 5,86	0,28
Leucócitos aumentados	2,39	0,81 – 7,06	0,11	2,40	0,44 – 8,17	0,33
Linfócitos aumentados	0,43	0,15 – 1,22	0,11	0,33	0,08 – 2,12	0,24
Plaquetas diminuídas	0,73	0,07 – 7,43	0,80	0,88	0,10 – 9,13	0,88
Temperatura média	5,30	1,76 – 16,14	0,00	8,76	1,08 – 7,23	0,04

TAC = Tacrolimo, CMV= Citomegalovírus

Níveis séricos elevados do IC Tacrolimo (TAC) conferiram baixo risco para RCA (OR 0,02 p=0,01), enquanto que os valores da dosagem no sangue de TAC baixos obtiveram um risco quatro vezes maior para RCA (OR 3,98 p=0,00). Em relação à queda de diurese, sorologia para citomegalovirus do receptor e tensão arterial média, os valores foram estatisticamente

significativos em relação ao risco de RCA (OR 3,40 p=0,02; OR 3,91 p=0,04; OR 8,76 p=0,04, respectivamente). A Tabela 2 apresenta resultados sobre a regressão de nefrotoxicidade;

**Tabela 3 - Odds Ratio bruto e ajustado para todas as variáveis independentes, considerando o desfecho de Nefrotoxicidade.**

Variável	OR bruto	IC95%	Valor p*	OR Ajustado	IC95%	Valor p**
DosagemTAC normal	2,45	0,77 – 7,80	0,13	2,06	0,75 – 4,15	0,80
DosagemTAC alto	4,14	1,13 – 12,68	0,01	7,31	2,67 – 13,31	0,01
DosagemTAC baixo	0,08	0,01 – 0,38	0,00	0,15	0,02 – 0,44	0,00
Indução	0,79	0,26 – 2,34	0,67	0,15	0,22 – 2,64	0,19
Tipo de doador	1,66	0,47 – 5,81	0,42	1,49	1,12 – 3,78	0,04
CMV do receptor positivo	1,05	0,34 – 3,28	0,92	0,88	0,82 – 2,75	0,20
Queda da diurese	0,20	0,02 – 1,72	0,14	0,05	0,02 – 3,63	0,18
Temperatura aumentada	1,92	0,29 – 12,41	0,49	1,90	0,33 – 11,12	0,33
Edema	6,91	1,22 – 3,90	0,02	5,27	1,10 – 2,98	0,04
Tremores	10,79	2,77 – 4,21	0,01	12,11	2,88 – 5,67	0,02
Ureia aumentada	2,16	0,71 – 6,57	0,17	1,12	0,43 – 5,54	0,45
Creatinina aumentada	3,75	1,09 – 12,85	0,03	4,33	1,66 – 11,21	0,01
Leucócitos aumentados	1,25	0,40 – 3,93	0,70	0,76	0,32 – 2,12	0,43
Linfócitos aumentados	4,37	1,27 – 14,97	0,01	6,91	1,88 – 16,12	0,00
Plaquetas diminuídas	0,01	0,00 – 3,66	0,80	0,01	0,00 – 2,12	0,61
Temperatura média	1,13	0,36 – 3,55	0,82	0,04	0,01 – 1,88	0,47

TAC = Tacrolimo, CMV = Citomegalovirus

Ressaltamos que os valores da dosagem no sangue de TAC altos apresentaram elevado risco para nefrotoxicidade (OR 7,31 p=0,01), enquanto os valores da dosagem de TAC baixos resultaram em baixo risco para nefrotoxicidade (OR 0,15 p=0,00). Tipo de doador, edema, tremores, aumento da creatinina e aumento dos linfócitos sanguíneos apresentaram elevados riscos para nefrotoxicidade, estatisticamente significativos (OR 1,49 p=0,04; OR 5,27 p=0,04; OR 2,88 p=0,02; OR 1,66 p=0,01 e OR 1,88 p=0,00, respectivamente).

Utilizando o *software* WEKA, foram construídos classificadores para as duas fases de classificação, empregando as técnicas: RNAs, SVM, Árvore de Decisão, Inferência Bayesiana e Vizinhos Mais Próximos.

Como critério de avaliação dos modelos, foi utilizado o maior valor de AUC. A Tabela 4 apresenta a avaliação dos modelos da Fase 1. São apresentados os valores de taxa de acerto (TA), sensibilidade (SEN), especificidade (ESP) e área sob a curva ROC (AUC). Sensibilidade está relacionada à ocorrência de RCA e/ou nefrotoxicidade, e especificidade está relacionada a não ocorrência destas complicações.

**Tabela 4** – Modelos selecionados na Fase 1: os modelos com maior valor de AUC para cada uma das técnicas utilizadas.

<b>Técnica</b>	<b>Algoritmo</b>	<b>TA (%)</b>	<b>SEN</b>	<b>ESP</b>	<b>AUC</b>
Árvore de Decisão	J48	63,94	0,64	0,64	0,69
Inferência bayesiana	Bayesnet	70,47	0,72	0,69	0,79
Inferência bayesiana	NaivesBayes	77,37	0,68	0,88	0,84
RNA	MLP	69,47	0,66	0,73	0,78
SVM	LIBSVM	74,53	0,80	0,7	0,75
Vizinhos mais próximos	IBK	57,98	0,94	0,28	0,74

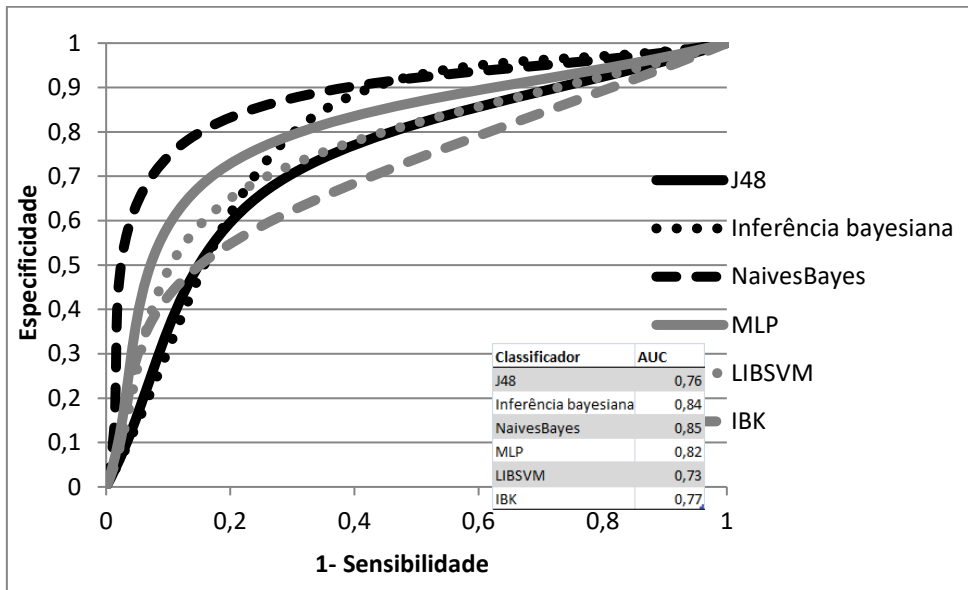
A tabela 5 e figura 2 apresentam a avaliação dos modelos selecionados na Fase 2. São apresentados os valores de taxa de acerto (TA), sensibilidade (SEN), especificidade (ESP) e área sob a curva ROC (AUC). Nesta tabela, sensibilidade está relacionada à ocorrência de nefrotoxicidade e especificidade relacionada à RCA.

**Tabela 5** – Modelos selecionados na Fase 2: os modelos com maior valor de AUC para cada uma das técnicas utilizadas neste trabalho.

<b>Técnica</b>	<b>Algoritmo</b>	<b>TA (%)</b>	<b>SEN</b>	<b>ESP</b>	<b>AUC</b>
Árvore de Decisão	J48	70,60	0,74	0,67	0,76
Inferência bayesiana	Bayesnet	75,25	0,67	0,83	0,84



Inferência bayesiana	NaivesBayes	82,20	0,87	0,78	0,85
RNA	MLP	76,05	0,82	0,71	0,82
SVM	LIBSVM	72,05	0,80	0,65	0,73
Vizinhos mais próximos	IBK	66,60	0,83	0,52	0,77



**Fig. 4** Curva ROC dos Classificadores na Fase 2.

Após essa filtragem foi escolhido o melhor modelo de predição das complicações, utilizando o valor de sensibilidade, e o melhor modelo de predição para nefrotoxicidade-RCA, utilizando o maior valor de AUC.

A técnica que apresentou o melhor valor de sensibilidade para prever a existência das complicações, nefrotoxicidade e RCA, foi a SVM, utilizando o algoritmo LIBSVM, que apresentou taxas de sensibilidade 0,80 (taxa de acerto 79,86%; especificidade 0,70; AUC 0,75). Excluiu-se a técnica de vizinhos mais próximos, pois apresentou a especificidade muito baixa. A técnica que apresentou o melhor valor de AUC para prever nefrotoxicidade ou RCA foi a de inferência bayesiana, utilizando o algoritmo NaiveBayes, que apresentou AUC 0,85 (taxa de acerto 82,2%).

## DISCUSSÃO

Atualmente, o transplante renal é o melhor tratamento para a IRC, pois resgata a cidadania, melhora significativamente a qualidade de vida do paciente e de seus familiares, além de proporcionar grande economia para o sistema de atendimento à saúde pública do Estado. Porém, não é um tratamento definitivo e desprovido de complicações. A RCA e a nefrotoxicidade provocadas pelos imunossupressores estão entre as principais complicações no primeiro ano de transplante renal, que podem levar à falência do órgão transplantado.

A comunidade científica busca um marcador que possa identificar os episódios de RCA antes da expressão clínica, ou seja, antes que haja disfunção do rim transplantado. Recentemente, a aplicação de técnicas mais precisas e reprodutíveis, como a da PCR (*polymerase chain reaction*) em tempo real e a técnica de DNA *microarrays*, permitiram avanços significativos respectivamente, na quantificação e identificação dos genes envolvidos na resposta aloimune, e poderão vir a se constituir em importantes métodos auxiliares na identificação e monitoramento dos eventos inflamatórios relacionados aos enxertos renais (Sabek O., 2002). Porém, atualmente, a biópsia do enxerto é o método diagnóstico mais sensível e específico para a rejeição aguda do enxerto. Da mesma forma, o diagnóstico diferencial entre RCA e nefrotoxicidade através da biópsia do enxerto constitui o padrão ouro.

Apesar da biópsia percutânea ser um método seguro e de baixa morbidade, podendo ser realizado em nível ambulatorial, são poucos os serviços que realizam estudos seriados para monitoramento imunológico do rim transplantado, ou seja, biópsia protocolar em pacientes estáveis. A maioria dos centros de transplante busca a análise histológica somente quando há evidência clínica de disfunção do enxerto renal.

Um ponto a ser considerado é a quase que completa ausência de sinais e de sintomas nos atuais processos de rejeição, devido a uma potente imunossupressão, o que torna

indispensável o uso de técnicas invasivas para o diagnóstico da rejeição aguda (Dias ECA, 2005).

Este estudo procurou, através de análises estatísticas, identificar classificadores de RCA e nefrotoxicidade através de Redes Neurais, que auxiliassem ou indicassem precocemente a punção percutânea para análise histológica, principalmente para aqueles serviços que não realizam biópsias seriadamente.

A abordagem metodológica inicial (de prever se um paciente possui nefrotoxicidade, RCA ou se não há ocorrência de nenhuma das duas complicações), não apresentou resultados satisfatórios. Acredita-se que este resultado esteja relacionado ao fato de que as manifestações clínicas das duas complicações são, na maioria das vezes, similares e os atributos utilizados para a predição não possibilitam uma distinção tão precisa em termos numéricos, evidenciando uma difícil separação dessas complicações em classes bem definidas.

A segunda abordagem metodológica utilizada (verificar se o paciente possui ou não possui uma das duas complicações, e depois prever qual das duas complicações) mostrou-se eficiente. Os resultados do classificador para presença de uma das duas complicações, apresentados na Tabela 4, mostram um resultado satisfatório, facilitando a indicação de enviar o paciente para realizar biópsia do enxerto, porque a taxa de erro da predição em não enviar o paciente para a biópsia, e ele ainda apresentar a complicação, é baixa, sendo 20 em cada 100 pacientes doentes. Em contrapartida, a taxa de acerto na predição de nefrotoxicidade e RCA foi considerada excelente (AUC 0,85). É interessante destacar que as técnicas de indução de árvore de decisão e de aprendizagem bayesiano utilizaram os mesmos atributos que possuem alta significância estatística para inferir o classificador.

## CONCLUSÃO

Ao analisar estes resultados, observou-se que os objetivos iniciais da pesquisa de classificar as complicações nefrotoxicidade e rejeição celular aguda em transplantes renais foram alcançados, com taxas de erro e acerto condizentes com a finalidade de efetuar a triagem dos pacientes que possam ter as complicações estudadas. Este estudo possui caráter exploratório. Para que os resultados sejam utilizados na prática clínica, é necessária a condução de novos estudos. Como exemplo, um estudo para validar o conjunto de classificadores escolhidos em um conjunto diferente de pacientes.

## REFERÊNCIAS

Aha DW, Kibler D, Albert MK. Instance-Based Learning Algorithms. *Machine Learning* 1991; 6:37-66.

Banerjee R, Das A, Ghoshal UC, Sinha M. Predicting mortality in patients with cirrhosis of liver with application of neural network technology. *J Gastroenterol Hepatol.* 2003; 18:1054–60.

Berner, ES *Clinical Decision Support Systems: Theory and Practice.* New York: Springer, 1999.

Bouchon-Meunier B. Uncertainty Management in Medical Applications. In: Akay, M. (Ed.), *Nonlinear Biomedical Signal Processing: Fuzzy Logic, Neural Networks, and New Algorithms.* New York: IEEE Press, 2000; 1:1-26.

Burnham KP, Anderson D. *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach.* Nova York: Springer, 2003.

Camps-Valls G, Porta-Oltra B, Soria-Olivas E, Martín-Guerrero JD, Serrano-López AJ, Pérez-Ruixo JJ, et al. Prediction of cyclosporine dosage in patients after kidney transplantation using neural networks. *IEEE Trans Biomed Eng* 2003; 50:442-8.

Carvalho, Luis Alfredo Vidal de. *Data Mining: A Mineração de Dados no Marketing, Medicina, Economia, Engenharia e Administração.* Rio de Janeiro: Ciência Moderna, 2005

Ciancio G, Burke GW, Gaynor JJ, Mattiazzzi A, Roth D, Kupin W, et al. A randomized long-term trial of tacrolimus and sirolimus versus tacrolimus and mycophenolate mofetil versus cyclosporine (NEORAL) and sirolimus in renal transplantation. I. Drug interactions and rejection at one year. *Transplantation* 2004; 77:244-51.

Cristianini N, Shawe-Taylor J. An introduction to support Vector Machines: and other kernel-based learning methods, Cambridge: MIT Press, 2003.

Dias ECA, Camara NOS, Silva Filho, A P, Manfro RC. Monitorização Molecular da Rejeição de Transplantes Renais. *J Bras Nefrol* 2005;27(2):76-83.

Duda, R. O.; Hart, P. E.; Pattern Classification and Scene Analysis. New York:

John Wiley and Sons, 1973.

Fryer J, Pellar S, Ormond D, Koffron A, Abecassis M. Mortality in Candidates Waiting for Combined Liver– Intestine Transplants Exceeds That for Other Candidates Waiting for Liver Transplants Liver Transplantation. 2003 Jul; 9(7):748-753.

Furness PN, Levesley J, Luo Z, Taub N, Kazi JI, Bates WD, et al. A neural network approach to the biopsy diagnosis of early acute renal transplant rejection. *Histopathology* 1999; 35:461-7.

Hariharan S. Recommendations for outpatient monitoring of kidney transplant recipients. *Am J Kidney Dis* 2006; 47:22-36.

Haykin S. *Neural Networks: A Comprehensive Foundation*. Upper Saddle River: Prentice Hall. 1999.

Hummel AD, Maciel RF, Rodregues RGS, Pisa IT. Application of Artificial Neural Networks in Renal Transplantation: Classification of Nephrotoxicity and Acute Cellular Rejection. *Transplantation Proceedings*, 2010; 42:471-472.

Johnson C, Ahsan N, Gonwa T, Halloran P, Stegall M. Randomized trial of tacrolimus (Prograf) in combination with azathioprine or mycophenolate mofetil versus cyclosporine (Neoral) with mycophenolate mofetil after cadaveric kidney transplantation. *Transplantation* 2000; 69:834-41.

Lisboa PJ. A review of evidence of health benefit from artificial neural networks in medical intervention. *Neural Netw* 2002; 15:11-39.

Lin RS, Horn SD, Hurdle JF, Goldfarb-Rumyantzev AS. Single and multiple time-point prediction models in kidney transplant outcomes. *J Biomed Inform* 2008; 41(6):944-52.

- Kasiske BL, Vazquez MA, Harmon WE, Brown RS, Danovitch GM, Gaston RS, Roth D, et al. Recommendations for the outpatient surveillance of renal transplant recipients. American Society of Transplantation. *J Am Soc Nephrol* 2000; 11:86.
- Kordylewski H, Graupe D, Liu K. A novel large-memory neural network as an aid in medical diagnosis applications. *IEEE Trans Inf Technol Biomed* 2001; 5:202-9.
- Massad E, Menezes RX, Silveira PSP, Ortega NRS. Métodos quantitativos em medicina. São Paulo: Manole; 2004.
- Mcgarra K, A survey of interestingness measures for knowledge discovery. *The Knowledge Engineering Review*, 2001 Mar; 20(1):39–61.
- Metz C. Basic principles of ROC analysis. *Semin Nucl Med* 1978; 8:283-98.
- Mierswa I, Wurst M, Klinkenberg R, Scholz M, Euler T. YALE: Rapid Prototyping for Complex Data Mining tasks. In: *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM Press. 2006; 935–940.
- Mohamad HH. *Fundamentals of Artificial Neural Networks*. Cambridge: MIT Press, 1995.
- Neal DK. Goodness of Fit Tests for Normality. *Mathematica Educ Res* 1996; 5:23-30.
- Quinlan R. *C4.5: Programs for Machine Learning*. San Mateo: Morgan Kaufmann Publishers, 1993.
- Rangel EB, Melaragno CS, Gonzalez AM, Linhares MM, Sá JR, Salzedas A, et al. O impacto da função do enxerto renal e da ocorrência de infecção intra-abdominal na sobrevida em 1 ano do paciente após o transplante simultâneo de pâncreas-rim (TSPR). *JBN* 2009; 31:78-88.
- Romão Junior, JE. Doença renal crônica: definição, epidemiologia e classificação. *J.Bras. Nefrol.*2004; 26:1-3.
- Sabek O, Dorak MT, Kotb M, et al. Quantitative detection of T-cell activation markers by real-time PCR in renal transplant rejection and correlation with histopathologic evaluation. *Transplantation* 2002;7:701.
- Simic-Ogrizovic S, Furuncic D, Lezaic V, Radivojevic D, Blagojevic R, Djukanovic L. Using ANN in selection of the most important variables in prediction of chronic renal allograft rejection progression. *Transplant Proc* 1999; 31:368.
- Shoskes DA, Ty R, Barba L, Sender M. Prediction of early graft function in renal transplantation using a computer neural network. *Transplant Proc* 1998; 30:1316-7.

Stuart Russel, Peter Norvig. Artificial Intelligence, a modern Approach. Second edition, Prentice Hall, 2003.

Vapnik, Vladimir N (2000). The Nature of Statistical Learning Theory. Information Science and Statistics. Springer-Verlag. ISBN 978-0-387-98780-4.

Witten IH, Frank E. Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. São Francisco: Morgan Kaufmann, 2005.

Yuan Y, Shaw MJ. Induction of fuzzy decision trees. Fuzzy Sets and Systems 1995 Jan; 69(2): 125 – 139. TEXTO REFERÊNCIA ...

## Anexo 1



Universidade Federal de São Paulo

Comitê de Ética em Pesquisa  
Hospital São Paulo

São Paulo, 24 de outubro de 2008.  
CEP 1677/08

Ilmo(a). Sr(a).  
Pesquisador(a) RAFAEL FABIO MACIEL  
Co-Investigadores: Renato Glaucio de Souza Rodrigues, Ivan Torres Piza  
Disciplina/Departamento: Pós-Graduação em Saúde Coletiva da Universidade Federal de São Paulo/Hospital São Paulo  
Patrocinador: Recursos Próprios.

### PARECER DO COMITÊ DE ÉTICA INSTITUCIONAL

Ref: Projeto de pesquisa intitulado: "Diagnóstico diferencial de rejeição celular aguda e nefrotoxicidade no primeiro ano de transplante renal utilizando redes neurais artificiais".

CARACTERÍSTICA PRINCIPAL DO ESTUDO: Observacional.

RISCOS ADICIONAIS PARA O PACIENTE: Sem risco.

OBJETIVOS: Realizar uma análise exploratória dos dados relativos ao acompanhamento ambulatorial do paciente transplantado renal usando Redes Neurais Artificiais e outras técnicas de Inteligência Artificial. Os resultados desta pesquisa servirão como base para a construção de uma ferramenta para o auxílio ao profissional por meio de um sistema de alerta para as complicações do transplante renal no primeiro ano de acompanhamento.

RESUMO: Será realizado um estudo retrospectivo e prospectivo em todos os prontuários dos pacientes transplantados renais no centro de transplante da Casa de Saúde Santa Efigênia - Caruaru - PE e a partir dos diagnósticos de rejeição celular aguda e nefrotoxicidade, incluindo avaliação clínica e exames laboratoriais, construindo-se uma Rede Neural Artificial (RNA), considerando-se as variáveis: rejeição celular aguda, nefrotoxicidade, avaliação clínica e laboratorial. Será desenvolvida uma interface gráfica para o sistema, para que médicos e profissionais em saúde possam analisar classificações obtidas pela RNA, interagindo e comparando dados do diagnóstico diferencial das complicações no acompanhamento pós-transplante renal de um indivíduo com as respostas geradas pela rede..

FUNDAMENTOS E RACIONAL: Muitas complicações associadas ao transplante de rim e imunossupressão podem ser prevenidas ou tratadas efetivamente quando são diagnosticadas em fases iniciais com o monitoramento pós-transplante. Recentemente, a atenção se voltou para uma grande variedade de técnicas de modelagem computacional, como Redes Neurais Artificiais. Uma aplicação comum é o sistema para alerta automático de riscos de saúde e procedimentos necessários para cada caso, a partir do conhecimento incorporado pela rede de dados armazenados sobre o paciente..

MATERIAL E MÉTODO: Estão descritos os procedimentos a serem realizados, apresentando carta de aprovação do CEP da UNCISAL e carta de autorização para a realização da pesquisa no Setor de Transplantes da Casa de Saúde Santa Efigênia de Caruaru - PE.





Universidade Federal de São Paulo

Comitê de Ética em Pesquisa  
Hospital São Paulo

TCLE: Adequado, de acordo com a resolução 196/96.

DETALHAMENTO FINANCEIRO: Sem financiamento externo.

CRONOGRAMA: 12 meses.

OBJETIVO ACADÊMICO: Mestrado.

ENTREGA DE RELATÓRIOS PARCIAIS AO CEP PREVISTOS PARA: 19/10/2009 e 19/10/2010.

O Comitê de Ética em Pesquisa da Universidade Federal de São Paulo/Hospital São Paulo **ANALISOU e APROVOU** o projeto de pesquisa referenciado.

1. Comunicar toda e qualquer alteração do projeto e termo de consentimento livre e esclarecido. Nestas circunstâncias a inclusão de pacientes deve ser temporariamente interrompida até a resposta do Comitê, após análise das mudanças propostas.
2. Comunicar imediatamente ao Comitê qualquer evento adverso ocorrido durante o desenvolvimento do estudo.
3. Os dados individuais de todas as etapas da pesquisa devem ser mantidos em local seguro por 5 anos para possível auditoria dos órgãos competentes.

Atenciosamente,

**Prof. Dr. José Osmar Medina Pestana**  
Coordenador do Comitê de Ética em Pesquisa da  
Universidade Federal de São Paulo/ Hospital São Paulo

167708

## Anexo 2

**CASA DE SAÚDE SANTA EFIGÊNIA**  
 Departamento de Transplante  
 Rua Gonçalo Coelho, 40, CEP 55.014-020  
 Caruaru – PE  
 Telefone: 081 21038500 – 081 21038580

## TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

---

### I - DADOS DE IDENTIFICAÇÃO DO SUJEITO DA PESQUISA OU RESPONSÁVEL LEGAL

1. NOME DO PACIENTE:.....  
 DOCUMENTO DE IDENTIDADE Nº : ..... SEXO: .M ( ) F ( )  
 DATA NASCIMENTO: ...../...../.....  
 ENDEREÇO: ..... Nº..... APTO: .....  
 BAIRRO: ..... CIDADE .....  
 CEP:..... TELEFONE: DDD (.....).....

2. RESPONSÁVEL LEGAL: .....  
 NATUREZA (grau de parentesco, tutor, curador etc.): .....  
 DOCUMENTO DE IDENTIDADE:.....SEXO: M( ) F( )  
 DATA NASCIMENTO: ...../...../.....  
 ENDEREÇO: ..... Nº ..... APTO: .....  
 BAIRRO: ..... CIDADE: .....  
 CEP: ..... TELEFONE: DDD (.....).....

---

**CASA DE SAÚDE SANTA EFIGÊNIA**  
Departamento de Transplante  
Rua Gonçalo Coelho, 40, CEP 55.014-020  
Caruaru – PE  
Telefone: 081 21038500 – 081 21038580

### **TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO**

Você está sendo convidado a participar da pesquisa intitulada “APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS EM TRANSPLANTES RENAI: CLASSIFICAÇÃO DE NEFROTOXICIDADE E REJEIÇÃO CELULAR AGUDA”.

O objetivo deste estudo é desenvolver um programa de computador a partir de uma rede neural artificial, aplicado ao acompanhamento ambulatorial do paciente transplantado renal, como uma ferramenta para o auxílio ao profissional, através de um sistema de alerta para as complicações do transplante renal no primeiro ano de acompanhamento. Nenhuma informação será capaz de identificar você dentro desta pesquisa, pois a privacidade e confidencialidade serão totalmente controladas pelos pesquisadores.

Os prontuários médicos e os resultados de exames serão também controlados pelos pesquisadores.

O benefício direto para você será a de ter a oportunidade de seus exames analisados por diversos pesquisadores e médicos especialistas com o objetivo de entender melhor o seu transplante, além disso, essa pesquisa trará benefícios para futuros pacientes transplantados. Outro benefício será o de informar a outros médicos e pesquisadores os resultados encontrados para, assim, contribuir para um melhor acompanhamento do transplante renal.

Em qualquer etapa do estudo, você terá acesso ao pesquisador, Dr. Rafael Fabio Maciel para esclarecimentos ou dúvidas eventuais: Rua José Veríssimo, 100. Caruaru – PE, CEP 55.014-250, Telefone 081 37238823 – Fax 081 2103-8580, e-mail [rfmaciel@uol.com.br](mailto:rfmaciel@uol.com.br) ou recorrer à Secretaria da Pós-Graduação em Saúde Coletiva, Universidade Federal de São Paulo (UNIFESP), Telefone: (11) 5571-5000 ramal 232.

Você tem a liberdade de retirar o consentimento a qualquer momento e deixar de participar do estudo, sem qualquer prejuízo. Você não terá nenhum gasto financeiro por participar deste estudo.

O pesquisador se compromete a utilizar os dados somente para esta pesquisa.

Eu, \_\_\_\_\_, discuti com os médicos envolvidos nesta pesquisa sobre a minha decisão em participar nesse estudo. Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que minha participação é isenta de despesas. Concordo voluntariamente em participar deste estudo e poderei retirar o meu consentimento a qualquer momento, antes ou durante esse, sem penalidades ou prejuízo.

\_\_\_\_\_  
Paciente voluntário

Data:

Declaro que obtive de forma apropriada e voluntária o Consentimento Livre e Esclarecido deste voluntário para a participação neste estudo.

\_\_\_\_\_  
Dr. Rafael Fabio Maciel  
Pesquisador responsável

\_\_\_\_\_  
Nome e CRM do médico que obteve a assinatura do TCLE.

### Anexo 3

#### Base de Dados de Treinamento dos Algoritmos de Inteligência Artificial para

#### Complicações/Não Complicações:

```
@RELATION Complicações
```

```
@ATTRIBUTE 'Tempo TX' real
@ATTRIBUTE 'TAC - Nominal' {alta,normal,baixa}
@ATTRIBUTE 'TAC - Numerico' real
@ATTRIBUTE 'Indução' {nao,sim}
@ATTRIBUTE 'Função inicial' {nao,sim}
@ATTRIBUTE 'TipoDoador' {vivo,cadaver}
@ATTRIBUTE 'CMV REC' {+,-}
@ATTRIBUTE 'Diurese' {nao,sim}
@ATTRIBUTE 'Aumt Temp' {nao,sim}
@ATTRIBUTE 'Edema' {nao,sim}
@ATTRIBUTE 'tremores' {nao,sim}
@ATTRIBUTE 'ureia' real
@ATTRIBUTE 'creat' real
@ATTRIBUTE 'glicemia >99' {nao,sim}
@ATTRIBUTE 'Leuc' real
@ATTRIBUTE 'Linf' real
@ATTRIBUTE 'Plaq' real
@ATTRIBUTE 'TA MÉDIA' real
@ATTRIBUTE 'haplo' {sim,nao}
@ATTRIBUTE 'ident' {nao,sim}
@ATTRIBUTE 'distinto' {nao,sim}
@ATTRIBUTE 'haplo/identico/distinto' {haplo,distinto,identico}
@ATTRIBUTE class {nao,biopsia}
```

```
@DATA
```

```
7.0,alta,15.9,nao,nao,vivo,+,nao,nao,nao,nao,30.0,0.8,nao,4400.0,1760.0
,170000.0,100.0,sim,nao,nao,haplo,nao
4.0,normal,7.0,nao,sim,vivo,+,nao,nao,nao,nao,37.0,1.4,nao,7400.0,2368.
0,260000.0,100.0,nao,nao,nao,distinto,biopsia
4.0,alta,14.0,nao,sim,cadaver,+,nao,nao,nao,sim,61.0,0.5,sim,11200.0,22
00.0,290000.0,115.0,nao,nao,sim,distinto,biopsia
6.0,alta,17.8,nao,sim,vivo,+,nao,nao,nao,sim,36.0,1.0,nao,16700.0,1480.
0,249000.0,100.0,sim,nao,nao,haplo,biopsia
5.0,alta,19.8,nao,sim,vivo,+,nao,nao,nao,nao,35.0,0.7,nao,6700.0,1675.0
,292000.0,100.0,sim,nao,nao,haplo,biopsia
2.0,alta,24.3,nao,sim,vivo,-
,nao,nao,nao,nao,64.0,1.6,sim,12900.0,1290.0,305000.0,105.0,sim,nao,nao
,haplo,biopsia
11.0,normal,8.4,nao,sim,vivo,+,nao,nao,nao,sim,71.0,1.7,nao,7800.0,1620
.0,297000.0,105.0,sim,nao,nao,haplo,biopsia
7.0,alta,9.3,nao,sim,vivo,+,nao,nao,nao,sim,184.0,3.7,sim,13100.0,655.0
,297000.0,110.0,sim,nao,nao,haplo,biopsia
8.0,baixa,4.4,nao,sim,vivo,+,nao,nao,nao,nao,34.0,0.9,nao,5000.0,1150.0
,150000.0,105.0,nao,sim,nao,identico,nao
4.0,alta,20.9,nao,sim,vivo,+,nao,nao,sim,sim,39.0,1.4,nao,7700.0,2695.0
,200000.0,85.0,sim,nao,nao,haplo,biopsia
```

10.0,alta,8.9,nao,sim,vivo,+,nao,nao,nao,nao,38.0,1.3,nao,7800.0,3510.0  
 ,220000.0,90.0,sim,nao,nao,haplo,nao  
 7.0,alta,16.8,sim,sim,vivo,+,nao,nao,nao,sim,74.0,0.7,nao,6200.0,1612.0  
 ,390000.0,100.0,nao,nao,sim,distinto,biopsia  
 2.0,alta,15.9,nao,sim,vivo,+,nao,nao,nao,nao,85.0,1.5,nao,11700.0,2457.  
 0,327000.0,115.0,nao,sim,nao,identico,biopsia  
 9.0,normal,8.6,nao,sim,vivo,+,nao,nao,nao,nao,39.0,1.7,nao,9300.0,1139.  
 0,185000.0,105.0,nao,sim,nao,identico,biopsia  
 2.0,normal,11.5,nao,sim,vivo,+,nao,nao,nao,nao,61.0,1.5,nao,16000.0,800  
 .0,352000.0,90.0,nao,nao,sim,distinto,biopsia  
 8.0,normal,6.9,nao,sim,vivo,+,nao,nao,nao,nao,73.0,1.3,nao,16600.0,2656  
 .0,278000.0,105.0,nao,sim,nao,identico,biopsia  
 1.0,normal,9.3,sim,sim,vivo,+,nao,nao,nao,nao,32.0,1.1,nao,7900.0,1343.  
 0,153000.0,105.0,nao,nao,sim,distinto,nao  
 4.0,baixa,3.3,nao,sim,vivo,-  
 ,nao,nao,nao,nao,29.0,1.2,nao,7600.0,2128.0,208000.0,100.0,nao,sim,nao,  
 identico,nao  
 3.0,alta,12.3,nao,sim,cadaver,-  
 ,nao,nao,nao,nao,31.0,1.2,nao,2200.0,330.0,132000.0,100.0,nao,nao,sim,d  
 istinto,nao  
 5.0,alta,13.3,nao,sim,vivo,+,nao,nao,nao,nao,26.0,1.0,nao,9800.0,1372.0  
 ,230000.0,105.0,sim,nao,nao,haplo,nao  
 1.0,baixa,5.8,sim,sim,cadaver,+,nao,nao,nao,nao,119.0,2.8,nao,6400.0,24  
 32.0,250000.0,85.0,nao,nao,sim,distinto,nao  
 10.0,baixa,4.3,nao,sim,vivo,+,nao,nao,nao,nao,29.0,0.9,nao,2400.0,1224.  
 0,160000.0,95.0,nao,nao,sim,distinto,nao  
 3.0,normal,8.1,sim,sim,vivo,+,nao,nao,nao,sim,51.0,1.1,sim,6900.0,3243.  
 0,200000.0,90.0,nao,sim,nao,identico,biopsia  
 5.0,alta,10.1,sim,sim,vivo,+,nao,nao,nao,nao,40.0,1.5,sim,7800.0,2340.0  
 ,220000.0,85.0,nao,sim,nao,identico,nao  
 6.0,normal,9.6,sim,sim,vivo,+,sim,nao,nao,nao,69.0,2.0,nao,19600.0,1176  
 .0,167000.0,105.0,nao,sim,nao,identico,biopsia  
 4.0,baixa,3.5,sim,nao,cadaver,+,nao,nao,nao,nao,34.0,1.0,nao,7100.0,170  
 4.0,294000.0,120.0,nao,nao,sim,distinto,nao  
 5.0,alta,15.6,sim,sim,cadaver,+,nao,nao,nao,nao,54.0,1.2,nao,5300.0,286  
 2.0,189000.0,105.0,nao,nao,sim,distinto,biopsia  
 11.0,alta,11.9,sim,sim,cadaver,+,nao,nao,nao,nao,57.0,1.6,sim,7300.0,19  
 71.0,257000.0,95.0,nao,nao,sim,distinto,biopsia  
 2.0,baixa,5.0,nao,sim,vivo,+,sim,nao,nao,nao,56.0,1.4,nao,10500.0,1050.  
 0,328000.0,100.0,nao,nao,sim,distinto,biopsia  
 10.0,baixa,3.1,nao,sim,vivo,+,nao,nao,nao,nao,40.0,1.3,nao,10000.0,2300  
 .0,200000.0,95.0,nao,sim,sim,identico,biopsia  
 5.0,normal,8.1,nao,sim,vivo,+,nao,nao,nao,nao,31.0,0.8,nao,8200.0,1968.  
 0,232000.0,100.0,nao,sim,nao,identico,biopsia  
 1.0,baixa,3.2,nao,sim,vivo,+,sim,sim,sim,nao,55.0,3.3,nao,8600.0,1204.0  
 ,150000.0,115.0,sim,nao,nao,haplo,biopsia  
 2.0,normal,10.8,nao,sim,vivo,-  
 ,sim,sim,nao,sim,60.0,1.8,nao,3400.0,680.0,228000.0,110.0,nao,sim,nao,i  
 dentico,biopsia  
 7.0,alta,14.2,nao,sim,vivo,+,nao,nao,nao,nao,68.0,1.4,nao,8600.0,1978.0  
 ,157000.0,90.0,sim,nao,nao,haplo,nao  
 8.0,normal,7.1,sim,sim,vivo,+,nao,nao,nao,nao,46.0,1.5,nao,7200.0,2016.  
 0,194000.0,95.0,nao,nao,sim,distinto,nao  
 1.0,baixa,3.1,sim,nao,vivo,+,nao,nao,nao,sim,132.0,6.8,nao,9100.0,1729.  
 0,194000.0,100.0,nao,nao,sim,distinto,biopsia  
 2.0,baixa,6.0,sim,nao,vivo,+,sim,nao,nao,nao,51.0,1.6,nao,13900.0,556.0  
 ,140000.0,90.0,nao,nao,sim,distinto,nao

4.0,baixa,4.7,sim,nao,vivo,+,nao,nao,nao,nao,44.0,2.1,nao,5500.0,723.0,  
 187000.0,90.0,nao,nao,sim,distinto,biopsia  
 6.0,normal,7.0,sim,nao,vivo,+,nao,nao,nao,nao,148.0,4.7,nao,6500.0,1059  
 .0,220000.0,90.0,nao,nao,sim,distinto,biopsia  
 7.0,alta,9.2,nao,sim,vivo,+,nao,nao,nao,nao,38.0,1.0,nao,8900.0,2840.0,  
 312000.0,80.0,sim,nao,nao,haplo,nao  
 1.0,baixa,4.3,nao,sim,vivo,+,sim,nao,nao,nao,101.0,2.6,sim,4200.0,294.0  
 ,104000.0,110.0,sim,nao,nao,haplo,biopsia  
 2.0,normal,8.2,nao,sim,cadaver,-  
 ,nao,nao,sim,nao,116.0,1.7,sim,6900.0,552.0,157000.0,115.0,nao,nao,sim,  
 distinto,biopsia  
 2.0,alta,26.2,nao,sim,vivo,-  
 ,nao,nao,nao,sim,84.0,2.0,nao,8200.0,984.0,213000.0,120.0,sim,nao,nao,h  
 aplo,biopsia  
 1.0,baixa,2.8,sim,sim,vivo,+,sim,nao,nao,nao,106.0,2.3,sim,15500.0,930.  
 0,251000.0,125.0,nao,nao,sim,distinto,biopsia  
 1.0,baixa,7.6,nao,sim,vivo,+,sim,sim,sim,nao,100.0,1.9,nao,11600.0,1027  
 .0,182000.0,120.0,sim,nao,nao,haplo,biopsia  
 1.0,baixa,4.7,sim,sim,cadaver,-  
 ,nao,nao,nao,nao,195.0,4.2,nao,8700.0,1218.0,266000.0,105.0,nao,sim,nao  
 ,identico,biopsia  
 1.0,baixa,4.7,sim,nao,cadaver,+,nao,nao,nao,nao,250.0,4.4,sim,6500.0,52  
 0.0,150000.0,160.0,nao,nao,sim,distinto,biopsia  
 10.0,baixa,5.1,sim,sim,cadaver,-  
 ,nao,nao,nao,nao,50.0,1.8,nao,7200.0,1728.0,225000.0,100.0,nao,nao,sim,  
 distinto,nao  
 6.0,alta,16.4,sim,sim,cadaver,-  
 ,nao,sim,nao,sim,95.0,3.1,sim,4100.0,920.0,161000.0,100.0,nao,nao,sim,d  
 istinto,biopsia  
 7.0,normal,6.6,nao,sim,vivo,-  
 ,nao,nao,sim,sim,120.0,5.3,nao,6400.0,512.0,312000.0,130.0,sim,nao,nao,  
 haplo,biopsia  
 1.0,baixa,2.5,sim,sim,cadaver,-  
 ,nao,nao,sim,nao,260.0,7.4,sim,14600.0,1168.0,154000.0,110.0,nao,nao,si  
 m,distinto,biopsia  
 9.0,baixa,4.4,sim,sim,vivo,-  
 ,nao,nao,nao,nao,30.0,1.0,nao,8600.0,946.0,125000.0,100.0,nao,nao,sim,d  
 istinto,nao  
 8.0,normal,7.5,nao,sim,vivo,-  
 ,nao,nao,nao,nao,68.0,1.6,nao,7600.0,760.0,180000.0,105.0,sim,nao,nao,h  
 aplo,biopsia  
 2.0,baixa,3.2,nao,sim,vivo,-  
 ,nao,nao,nao,nao,36.0,0.9,nao,7300.0,1168.0,308000.0,100.0,nao,nao,sim,  
 distinto,nao  
 5.0,alta,13.7,nao,sim,vivo,-  
 ,nao,nao,nao,nao,70.0,0.8,nao,11500.0,2070.0,210000.0,100.0,sim,nao,nao  
 ,haplo,nao  
 1.0,baixa,3.1,nao,sim,vivo,-  
 ,sim,nao,nao,nao,67.0,1.6,sim,6100.0,618.0,177000.0,110.0,sim,nao,nao,h  
 aplo,biopsia  
 4.0,normal,9.2,nao,sim,vivo,-  
 ,nao,nao,nao,nao,39.0,1.1,nao,5100.0,1224.0,150000.0,100.0,sim,nao,nao,  
 haplo,nao  
 6.0,baixa,3.8,nao,sim,vivo,-  
 ,nao,nao,nao,nao,34.2,1.0,nao,8100.0,1610.0,271000.0,80.0,sim,nao,nao,h  
 aplo,nao

1.0,baixa,3.8,sim,nao,cadaver,-  
 ,sim,nao,nao,sim,133.0,4.6,nao,9200.0,920.0,171000.0,125.0,nao,nao,sim,  
 distinto,biopsia  
 3.0,baixa,2.6,sim,sim,cadaver+,nao,nao,nao,nao,37.4,0.7,nao,5450.0,114  
 5.0,308000.0,100.0,nao,nao,sim,distinto,nao  
 5.0,normal,8.0,nao,sim,vivo,-  
 ,nao,nao,nao,nao,29.0,1.0,nao,11200.0,5376.0,316000.0,125.0,sim,nao,nao  
 ,haplo,nao  
 1.0,baixa,4.9,sim,sim,cadaver+,nao,nao,nao,nao,66.0,1.8,nao,7200.0,136  
 8.0,160000.0,105.0,nao,nao,sim,distinto,nao  
 4.0,baixa,7.5,sim,sim,cadaver+,nao,nao,nao,nao,76.0,1.8,nao,6800.0,183  
 6.0,134000.0,140.0,nao,nao,sim,distinto,biopsia  
 3.0,alta,12.6,nao,sim,vivo+,nao,nao,nao,nao,66.4,1.2,nao,8300.0,3071.0  
 ,201400.0,120.0,sim,nao,nao,haplo,nao  
 1.0,alta,3.5,nao,sim,vivo,-  
 ,nao,nao,nao,nao,64.0,2.7,nao,10800.0,648.0,153000.0,140.0,sim,nao,nao,  
 haplo,biopsia  
 2.0,baixa,2.5,sim,sim,cadaver,-  
 ,nao,nao,nao,nao,71.0,1.5,nao,4500.0,1170.0,246000.0,95.0,nao,nao,sim,d  
 istinto,nao  
 1.0,baixa,4.0,sim,nao,cadaver+,sim,nao,nao,nao,222.0,7.8,sim,7500.0,60  
 0.0,209000.0,105.0,nao,nao,sim,distinto,nao  
 2.0,baixa,4.3,nao,sim,vivo+,nao,nao,nao,nao,38.0,1.3,nao,1800.0,720.0,  
 188000.0,95.0,nao,nao,sim,distinto,nao  
 9.0,alta,13.5,nao,sim,vivo+,nao,nao,nao,nao,47.0,1.3,nao,8000.0,3200.0  
 ,300000.0,95.0,sim,nao,nao,haplo,nao  
 6.0,normal,7.6,sim,sim,cadaver+,nao,nao,nao,nao,54.0,1.4,nao,7300.0,21  
 90.0,304000.0,95.0,nao,nao,sim,distinto,nao  
 5.0,alta,13.9,nao,sim,vivo+,nao,nao,nao,nao,110.0,1.7,nao,11800.0,4600  
 .0,283640.0,90.0,sim,nao,nao,haplo,biopsia  
 1.0,baixa,4.6,sim,nao,cadaver+,sim,nao,sim,nao,50.0,1.8,sim,12200.0,97  
 6.0,171000.0,110.0,nao,nao,sim,distinto,biopsia  
 4.0,normal,1.0,sim,sim,cadaver+,nao,nao,nao,nao,73.0,1.7,sim,14000.0,6  
 225.0,278800.0,125.0,nao,nao,sim,distinto,nao  
 2.0,baixa,6.2,sim,sim,cadaver+,nao,nao,nao,sim,102.0,2.4,nao,10000.0,6  
 00.0,147000.0,85.0,nao,nao,sim,distinto,nao  
 12.0,normal,6.7,sim,sim,cadaver+,nao,nao,nao,nao,31.0,0.9,nao,7900.0,1  
 349.0,336000.0,100.0,nao,nao,sim,distinto,nao  
 4.0,normal,8.3,sim,nao,cadaver+,nao,nao,nao,nao,42.0,1.2,nao,9500.0,23  
 75.0,290000.0,105.0,nao,nao,sim,distinto,nao  
 4.0,normal,7.9,nao,sim,vivo+,nao,nao,nao,nao,43.0,1.3,nao,7600.0,760.0  
 ,380000.0,80.0,sim,nao,nao,haplo,biopsia  
 4.0,normal,8.0,sim,sim,cadaver+,nao,nao,nao,sim,42.0,1.4,nao,3800.0,16  
 95.0,238000.0,85.0,nao,nao,sim,distinto,biopsia  
 3.0,baixa,5.9,sim,sim,cadaver+,nao,nao,nao,nao,45.0,1.2,nao,5100.0,122  
 4.0,205000.0,95.0,nao,nao,sim,distinto,nao  
 3.0,alta,15.3,sim,nao,cadaver+,nao,nao,nao,nao,38.0,1.0,nao,4600.0,133  
 4.0,290000.0,95.0,nao,nao,sim,distinto,biopsia  
 3.0,baixa,5.8,nao,sim,vivo+,nao,nao,nao,nao,29.0,1.1,nao,10900.0,2180.  
 0,281520.0,95.0,sim,nao,nao,haplo,nao  
 1.0,baixa,2.3,nao,sim,vivo+,nao,nao,sim,nao,92.0,2.2,nao,9900.0,1485.0  
 ,252000.0,110.0,sim,nao,nao,haplo,biopsia  
 3.0,normal,9.8,nao,sim,vivo+,nao,nao,nao,nao,56.0,1.7,nao,10300.0,1339  
 .0,268410.0,100.0,sim,nao,nao,haplo,nao  
 1.0,normal,11.8,sim,nao,cadaver+,sim,nao,nao,nao,223.0,11.0,nao,12500.  
 0,1125.0,278000.0,110.0,nao,nao,sim,distinto,biopsia  
 2.0,normal,9.4,sim,nao,cadaver+,sim,sim,sim,nao,130.0,7.6,nao,8600.0,6  
 02.0,164000.0,105.0,nao,nao,sim,distinto,biopsia



2.0,alta,13.2,nao,sim,vivo,+,nao,nao,nao,sim,26.0,1.5,nao,10900.0,3924.  
0,370500.0,100.0,sim,nao,nao,haplo,nao  
2.0,baixa,5.5,sim,sim,cadaver,+,nao,nao,nao,nao,27.0,1.2,nao,5100.0,918  
.0,267260.0,105.0,nao,nao,sim,distinto,nao  
2.0,normal,1.0,nao,sim,vivo,+,nao,nao,nao,nao,43.0,1.1,nao,11700.0,3861  
.0,323700.0,105.0,sim,nao,nao,haplo,nao  
2.0,baixa,1.0,sim,sim,cadaver,+,nao,nao,nao,nao,43.0,1.3,nao,4900.0,588  
.0,211000.0,100.0,nao,nao,sim,distinto,nao  
1.0,baixa,4.7,nao,sim,vivo,+,sim,sim,nao,nao,84.0,1.6,sim,7000.0,770.0,  
142000.0,150.0,nao,nao,sim,distinto,biopsia  
1.0,baixa,7.4,sim,sim,cadaver,+,nao,nao,nao,nao,55.0,1.4,nao,10300.0,24  
72.0,196540.0,100.0,nao,nao,sim,distinto,nao

## Anexo 4

### Base de Dados de Treinamento dos Algoritmos de Inteligência Artificial para Nefrotoxicidade/RCA:

```
@RELATION rcanefro
```

```
@ATTRIBUTE 'Tempo TX' real
@ATTRIBUTE 'TAC - Nominal' {normal,alta,baixa}
@ATTRIBUTE 'TAC - Numerico' real
@ATTRIBUTE 'Indução' {nao,sim}
@ATTRIBUTE 'Função inicial' {sim,nao}
@ATTRIBUTE 'TipoDoador' {vivo,cadaver}
@ATTRIBUTE 'CMV REC' {+,-}
@ATTRIBUTE 'Diurese' {nao,sim}
@ATTRIBUTE 'Aumt Temp' {nao,sim}
@ATTRIBUTE 'Edema' {nao,sim}
@ATTRIBUTE 'tremores' {nao,sim}
@ATTRIBUTE 'ureia' real
@ATTRIBUTE 'creat' real
@ATTRIBUTE 'glicemia >99' {nao,sim}
@ATTRIBUTE 'Leuc' real
@ATTRIBUTE 'Linf' real
@ATTRIBUTE 'Plaq' real
@ATTRIBUTE 'TA MÉDIA' real
@ATTRIBUTE haplo {nao,sim}
@ATTRIBUTE 'ident' {nao,sim}
@ATTRIBUTE distinto {nao,sim}
@ATTRIBUTE 'haplo/identico/distinto' {distinto,haplo,identico}
@ATTRIBUTE class {nefrotoxicidade,rca}
```

```
@DATA
```

```
4.0,normal,7.0,nao,sim,vivo,+,nao,nao,nao,nao,37.0,1.4,nao,7400.0,2368.0,260000.0,100.0,nao,nao,nao,distinto,nefrotoxicidade
4.0,alta,14.0,nao,sim,cadaver,+,nao,nao,nao,sim,61.0,0.5,sim,11200.0,2200.0,290000.0,115.0,nao,nao,sim,distinto,nefrotoxicidade
6.0,alta,17.8,nao,sim,vivo,+,nao,nao,nao,sim,36.0,1.0,nao,16700.0,1480.0,249000.0,100.0,sim,nao,nao,haplo,nefrotoxicidade
5.0,alta,19.8,nao,sim,vivo,+,nao,nao,nao,nao,35.0,0.7,nao,6700.0,1675.0,292000.0,100.0,sim,nao,nao,haplo,rca
2.0,alta,24.3,nao,sim,vivo,-,nao,nao,nao,nao,64.0,1.6,sim,12900.0,1290.0,305000.0,105.0,sim,nao,nao,haplo,nefrotoxicidade
11.0,normal,8.4,nao,sim,vivo,+,nao,nao,nao,sim,71.0,1.7,nao,7800.0,1620.0,297000.0,105.0,sim,nao,nao,haplo,rca
7.0,alta,9.3,nao,sim,vivo,+,nao,nao,nao,sim,184.0,3.7,sim,13100.0,655.0,297000.0,110.0,sim,nao,nao,haplo,rca
4.0,alta,20.9,nao,sim,vivo,+,nao,nao,sim,sim,39.0,1.4,nao,7700.0,2695.0,200000.0,85.0,sim,nao,nao,haplo,nefrotoxicidade
7.0,alta,16.8,sim,sim,vivo,+,nao,nao,nao,sim,74.0,0.7,nao,6200.0,1612.0,390000.0,100.0,nao,nao,sim,distinto,nefrotoxicidade
2.0,alta,15.9,nao,sim,vivo,+,nao,nao,nao,nao,85.0,1.5,nao,11700.0,2457.0,327000.0,115.0,nao,sim,nao,identico,nefrotoxicidade
```

9.0, normal, 8.6, nao, sim, vivo, +, nao, nao, nao, nao, 39.0, 1.7, nao, 9300.0, 1139.0, 185000.0, 105.0, nao, sim, nao, identico, nefrotoxicidade  
 2.0, normal, 11.5, nao, sim, vivo, +, nao, nao, nao, nao, 61.0, 1.5, nao, 16000.0, 800.0, 352000.0, 90.0, nao, nao, sim, distinto, rca  
 8.0, normal, 6.9, nao, sim, vivo, +, nao, nao, nao, nao, 73.0, 1.3, nao, 16600.0, 2656.0, 278000.0, 105.0, nao, sim, nao, identico, rca  
 3.0, normal, 8.1, sim, sim, vivo, +, nao, nao, nao, sim, 51.0, 1.1, sim, 6900.0, 3243.0, 200000.0, 90.0, nao, sim, nao, identico, nefrotoxicidade  
 6.0, normal, 9.6, sim, sim, vivo, +, sim, nao, nao, nao, 69.0, 2.0, nao, 19600.0, 1176.0, 167000.0, 105.0, nao, sim, nao, identico, rca  
 5.0, alta, 15.6, sim, sim, cadaver, +, nao, nao, nao, nao, 54.0, 1.2, nao, 5300.0, 2862.0, 189000.0, 105.0, nao, nao, sim, distinto, nefrotoxicidade  
 11.0, alta, 11.9, sim, sim, cadaver, +, nao, nao, nao, nao, 57.0, 1.6, sim, 7300.0, 1971.0, 257000.0, 95.0, nao, nao, sim, distinto, nefrotoxicidade  
 2.0, baixa, 5.0, nao, sim, vivo, +, sim, nao, nao, nao, 56.0, 1.4, nao, 10500.0, 1050.0, 328000.0, 100.0, nao, nao, sim, distinto, rca  
 10.0, baixa, 3.1, nao, sim, vivo, +, nao, nao, nao, nao, 40.0, 1.3, nao, 10000.0, 2300.0, 200000.0, 95.0, nao, sim, sim, identico, rca  
 5.0, normal, 8.1, nao, sim, vivo, +, nao, nao, nao, nao, 31.0, 0.8, nao, 8200.0, 1968.0, 232000.0, 100.0, nao, sim, nao, identico, nefrotoxicidade  
 1.0, baixa, 3.2, nao, sim, vivo, +, sim, sim, sim, nao, 55.0, 3.3, nao, 8600.0, 1204.0, 150000.0, 115.0, sim, nao, nao, haplo, rca  
 2.0, normal, 10.8, nao, sim, vivo, -, sim, sim, nao, sim, 60.0, 1.8, nao, 3400.0, 680.0, 228000.0, 110.0, nao, sim, nao, identico, nefrotoxicidade  
 1.0, baixa, 3.1, sim, nao, vivo, +, nao, nao, nao, sim, 132.0, 6.8, nao, 9100.0, 1729.0, 194000.0, 100.0, nao, nao, sim, distinto, rca  
 4.0, baixa, 4.7, sim, nao, vivo, +, nao, nao, nao, nao, 44.0, 2.1, nao, 5500.0, 723.0, 187000.0, 90.0, nao, nao, sim, distinto, nefrotoxicidade  
 1.0, baixa, 4.3, nao, sim, vivo, +, sim, nao, nao, nao, 101.0, 2.6, sim, 4200.0, 294.0, 104000.0, 110.0, sim, nao, nao, haplo, rca  
 2.0, normal, 8.2, nao, sim, cadaver, -, nao, nao, sim, nao, 116.0, 1.7, sim, 6900.0, 552.0, 157000.0, 115.0, nao, nao, sim, distinto, rca  
 2.0, alta, 26.2, nao, sim, vivo, -, nao, nao, nao, sim, 84.0, 2.0, nao, 8200.0, 984.0, 213000.0, 120.0, sim, nao, nao, haplo, nefrotoxicidade  
 1.0, baixa, 2.8, sim, sim, vivo, +, sim, nao, nao, nao, 106.0, 2.3, sim, 15500.0, 930.0, 251000.0, 125.0, nao, nao, sim, distinto, rca  
 1.0, baixa, 7.6, nao, sim, vivo, +, sim, sim, sim, nao, 100.0, 1.9, nao, 11600.0, 1027.0, 182000.0, 120.0, sim, nao, nao, haplo, rca  
 1.0, baixa, 4.7, sim, sim, cadaver, -, nao, nao, nao, nao, 195.0, 4.2, nao, 8700.0, 1218.0, 266000.0, 105.0, nao, sim, nao, identico, rca  
 1.0, baixa, 4.7, sim, nao, cadaver, +, nao, nao, nao, nao, 250.0, 4.4, sim, 6500.0, 520.0, 150000.0, 160.0, nao, nao, sim, distinto, rca  
 6.0, alta, 16.4, sim, sim, cadaver, -, nao, sim, nao, sim, 95.0, 3.1, sim, 4100.0, 920.0, 161000.0, 100.0, nao, nao, sim, distinto, nefrotoxicidade  
 7.0, normal, 6.6, nao, sim, vivo, -, nao, nao, sim, sim, 120.0, 5.3, nao, 6400.0, 512.0, 312000.0, 130.0, sim, nao, nao, haplo, nefrotoxicidade  
 1.0, baixa, 2.5, sim, sim, cadaver, -, nao, nao, sim, nao, 260.0, 7.4, sim, 14600.0, 1168.0, 154000.0, 110.0, nao, nao, sim, distinto, rca  
 8.0, normal, 7.5, nao, sim, vivo, -, nao, nao, nao, nao, 68.0, 1.6, nao, 7600.0, 760.0, 180000.0, 105.0, sim, nao, nao, haplo, nefrotoxicidade

1.0,baixa,3.1,nao,sim,vivo,-  
 ,sim,nao,nao,nao,67.0,1.6,sim,6100.0,618.0,177000.0,110.0,sim,nao,nao,h  
 aplo,rca  
 1.0,baixa,3.8,sim,nao,cadaver,-  
 ,sim,nao,nao,sim,133.0,4.6,nao,9200.0,920.0,171000.0,125.0,nao,nao,sim,  
 distinto,rca  
 4.0,baixa,7.5,sim,sim,cadaver,+,nao,nao,nao,nao,76.0,1.8,nao,6800.0,183  
 6.0,134000.0,140.0,nao,nao,sim,distinto,nefrotoxicidade  
 1.0,alta,3.5,nao,sim,vivo,-  
 ,nao,nao,nao,nao,64.0,2.7,nao,10800.0,648.0,153000.0,140.0,sim,nao,nao,  
 haplo,rca  
 5.0,alta,13.9,nao,sim,vivo,+,nao,nao,nao,nao,110.0,1.7,nao,11800.0,4600  
 .0,283640.0,90.0,sim,nao,nao,haplo,nefrotoxicidade  
 1.0,baixa,4.6,sim,nao,cadaver,+,sim,nao,sim,nao,50.0,1.8,sim,12200.0,97  
 6.0,171000.0,110.0,nao,nao,sim,distinto,rca  
 4.0,normal,7.9,nao,sim,vivo,+,nao,nao,nao,nao,43.0,1.3,nao,7600.0,760.0  
 ,380000.0,80.0,sim,nao,nao,haplo,nefrotoxicidade  
 4.0,normal,8.0,sim,sim,cadaver,+,nao,nao,nao,sim,42.0,1.4,nao,3800.0,16  
 95.0,238000.0,85.0,nao,nao,sim,distinto,nefrotoxicidade  
 3.0,alta,15.3,sim,nao,cadaver,+,nao,nao,nao,nao,38.0,1.0,nao,4600.0,133  
 4.0,290000.0,95.0,nao,nao,sim,distinto,nefrotoxicidade  
 1.0,baixa,2.3,nao,sim,vivo,+,nao,nao,sim,nao,92.0,2.2,nao,9900.0,1485.0  
 ,252000.0,110.0,sim,nao,nao,haplo,rca  
 1.0,normal,11.8,sim,nao,cadaver,+,sim,nao,nao,nao,223.0,11.0,nao,12500.  
 0,1125.0,278000.0,110.0,nao,nao,sim,distinto,rca  
 2.0,normal,9.4,sim,nao,cadaver,+,sim,sim,sim,nao,130.0,7.6,nao,8600.0,6  
 02.0,164000.0,105.0,nao,nao,sim,distinto,rca  
 1.0,baixa,4.7,nao,sim,vivo,+,sim,sim,nao,nao,84.0,1.6,sim,7000.0,770.0,  
 142000.0,150.0,nao,nao,sim,distinto,rca