# We are IntechOpen, the world's leading publisher of Open Access books
# Built by scientists, for scientists

**6,900**
Open access books available

**186,000**
International authors and editors

**200M**
Downloads

**154**
Countries delivered to

Our authors are among the

**TOP 1%**
most cited scientists

**12.2%**
Contributors from top 500 universities

CLARIVATE ANALYTICS
**BOOK CITATION INDEX**
INDEXED

**WEB OF SCIENCE**™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

## Interested in publishing with us?
## Contact book.department@intechopen.com

# The Usefulness of Artificial Neural Networks in Predicting the Outcome of Hematopoietic Stem Cell Transplantation

Giovanni Caocci[1], Roberto Baccoli[2] and Giorgio La Nasa[1]
*[1]Division of Hematology and Hematopoietic Stem Cell Transplantation,
Department of Internal Medical Sciences, University of Cagliari
[2]Technical Physics Division, Faculty of Engineering,
Department of Engineering of the Territory, University of Cagliari
Italy*

## 1. Introduction

### 1.1 Artificial neural networks in clinical medicine

In Medicine, several tools have been developed for the prediction of clinical outcomes following drug treatment and other medical interventions. The standard approach for a binary outcome is to use logistic regression (LR), however, this method requires formal training and a profound knowledge of statistics (Royston, 2000; Harrel et al., 1996). LR is used to predict a categorical (usually dichotomous) variable from e set of predictor variables; it has been especially popular with medical research in which the dependent variable is whether or not a patient has a disease.

Over the past years, artificial neural networks (ANNs) have increasingly been used as an alternative to LR analysis for prognostic and diagnostic classification in clinical medicine (Schwarzer et al., 2000). ANNs are composed of simple elements operating in parallel inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. After training with retrospective data ANNs are capable of making intelligent predictions given new, limited information. The growing interest in ANNs has mainly been triggered by their ability to mimic the learning processes of the human brain. However, the issue remains as to how these ANNs actually succeed in recognizing patterns within data that are too complex for the human brain. From here derives the so-called "black-box" aspect of ANNs. The network operates in a feed-forward mode from the input layer through the hidden layers (like in a black box) to the output layer. Exactly what interactions are modeled in the hidden layers is still a knot that remains untied. Each layer within the network is made up of computing nodes with remarkable data processing abilities. Each node is connected to other nodes of a previous layer through adaptable inter-neuron connection strengths known as synaptic weights. ANNs are trained for specific applications, such as pattern recognition or data classification, through a learning process and knowledge is usually retained as a set of connection weights. The backpropagation algorithm and its variants are learning algorithms that are widely used in neural networks. With backpropagation, the input data is repeatedly presented to the

network. Each time, the output is compared to the desired output and an error is computed. The error is then fed back through the network and used to adjust the weights in such a way that with each iteration it gradually declines until the neural model produces the desired output. ANNs have been successfully applied in the fields of mathematics, engineering, medicine, economics, meteorology, psychology, neurology, and many others. Indeed, in medicine, they offer a tantalizing alternative to multivariate analysis, although their role remains advisory since no convincing evidence of any real progress in clinical prognosis has yet been produced (Linder et al., 2006).

A systematic review on the use of artificial neural networks in decision support in cancer by Lisboa et al. showed that the number of clinical trials (CTs) and randomised controlled trials (RCTs) involving the use of ANNs in diagnosis and prognosis has increased from 1 to 38 over the past decade. However, out of 396 studies involving the use of ANNs in cancer, only 27 were either CTs or RCTs. Out of these trials, 21 showed an increase in benefit to healthcare provision and 6 did not. None of these studies however showed a decrease in benefit. Overall, the reviewed publications support the neural network approach but while on the one hand they identify trends in areas of clinical promise (particularly diagnosis, prognosis and therapeutic guidance for cancer), on the other they highlight the need for more extensive application of rigorous methodologies (Lisboa & Taktak, 2006).

Interesting, a review on the use of ANNs in the field of Gastroenterology over the last 10 years (their application in the field of gastroenterology has now entered the second decade) underlines that the increasing complexity of clinical data requires the use of mathematical models that are able to capture the key properties of entire ensembles, including their linkages and their hubs, abandoning the traditional statistical reductionistic approach, which tends to 'see' things individually, to simplify and to look at one single element at a time (Pace & Savarino, 2007). Some authors, for example, assessed the performance of ANNs in recognizing patients with chronic atrophic gastritis, a state of chronic inflammation that can eventually progress to gastric carcinoma, by using only clinical and biochemical variables (Annibale & Lahner, 2007).

In the field of urology, several papers have addressed the predictive efficacy of ANNs. In urological cancer, ANNs appear to be accurate and more explorative than traditional regression statistics artificial intelligence methods when used to analyze large data cohorts. Furthermore, they allow individualized prediction of disease behaviour. Each artificial intelligence method has characteristics that make it suitable for different tasks. The lack of transparency of ANNs hinders global scientific community acceptance, but this can be overcome by neuro-fuzzy modeling systems (Abbod et al., 2007).

New biomarkers within multivariate models have been analyzed with ANNs to improve early detection of prostate cancer (Stephan et al., 2007). Another field of application is the management of urolithiasis, a worldwide clinical challenge embracing a multitude of difficulties in diagnosis, treatment and prevention of recurrence. Recent reports have examined the role of ANNs in prediction of stone presence and composition, spontaneous passage, clearance and re-growth after treatment (Rajan & Tolley, 2005). The results suggest that ANNs may prove useful in clinician-led decision-making processes.

ANNs can identify important predictive variables and accurately predict treatment outcomes in clinical medicine but although the initial results appear promising, further prospective studies of larger patient cohorts will need to be performed in order to determine whether this mode of analysis can surpass standard statistical predictive methods, not only when solving problems concerning diagnosis and its classification into subtypes but also when predicting clinical outcomes of patients affected by diverse pathologies.

**1.2 Artificial neural networks in hematology and bone marrow transplantation**

Very few reports on this topic have been published in the field of hematology. The first computing devices based on artificial intelligence (AI) have been applied to routine laboratory data management whereas new innovative tools, based on neural networks trained with data from peripheral blood analysis, have been used for differential diagnosis in diseases such as anemias, thalassemias and leukemias. The introduction of the first microarray based and bio-informatic approach for molecular diagnosis of hematological malignancies can be considered a major step ahead. This approach is based on the monitoring of simultaneous expression of thousands of genes using DNA microarray, independently of previous biological knowledge, analyzed using AI devices (Zini, 2005).

In an attempt to create an application for discriminating different types of anemia, simply using data from peripheral blood, Zini & d'Onofrio (2003) collected data from peripheral blood of 1000 patients diagnosed mainly with hematopoietic disorders in 22 Italian Hematology Centers. The ANNs were trained with labeled samples and showed high capability of clustering signals according to the predefined normal as well as pathological profiles.

In 2002, Amendolia et al. investigated the use of ANNs for the classification of thalassemic pathologies, exclusively using the hematologic parameters resulting from hemochromocytometric analysis. Different combinations of ANNs made it possible to discriminate thalassemia carriers from normals with 94% classification accuracy, 92% sensitivity, and 95% specificity. Based on these results, an automated system for real-time support in diagnoses was proposed (Amendolia et al., 2002).

All these intriguing reports of ANNs in the field of hematology kindled our curiosity to discover whether ANNs were capable of predicting the outcome of hematopoietic stem cell transplantation (HSCT) after analyzing donor and recipient pre-transplantation clinical and immunogenetic variables.

**1.3 The difficult setting of unrelated bone marrow transplantation in thalassemia.**

Patients with chronic non-malignant genetic disorders, such as thalassemia, are faced with a dramatic decision: they can either undergo HSCT with a good possibility of cure but a high chance of death or continue the more conventional treatment with blood transfusions and iron chelation therapy. The important advances made in conventional treatment now allow transfusion-dependent thalassemia patients to live much longer (Caocci et al., 2006; Borgna Pignatti et al., 2004) but as a result these patients must cope with complications that occur over time. Treatment may be required for heart or liver diseases, infections, osteoporosis and other serious health problems.

On the other hand, although HSCT from an HLA-identical sibling can offer thalassemia patients a probability of cure that is close to 90% in children and adults in good clinical conditions (Lucarelli et al., 1990), this procedure is associated with a significant risk of mortality (Lucarelli et al., 1997), especially in patients with advanced age or poor clinical conditions. Moreover, the chance that any given sibling will be HLA matched with a potential recipient is one out of four which means that most patients will need to search for a compatible donor in the registries of voluntary donors worldwide. Transplantation from unrelated donors (UD) is burdened by an increased risk of acute and chronic graft-versus host disease (GVHD) with a consequent negative impact on overall survival (La Nasa et al., 2006). Therefore, every effort should be made to carefully evaluate the risk of GVHD before performing UD-HSCT (Hansen et al., 1998).

Graft-versus-host disease remains the major barrier to the successful outcome of HSCT in thalassemia. HLA disparity between the donor and the recipient is clearly the most powerful risk factor but also older age, gender mismatch, Pesaro risk class, cytomegalovirus (CMV) positivity as well as higher median infused hematopoietic stem cell doses have been shown to increase the risk for GVHD (Lucarelli et al., 1996). Evidence emerging from recent reports indicates a correlation between certain immunogenetic variables and the occurrence of GVHD: donor-recipient HLA-Cw ligand groups for killer immunoglobulin-like receptors (KIRs), KIR genotypes, the HLA-G 14-basepair (bp) polymorphism and HLA-DPB1 disparity (La Nasa et al., 2007; Littera et al., 2010; Fleischhauer et al., 2006). Although this information may contribute to our understanding of the pathogenesis of GVHD, it is difficult to apply in clinical practice. What we need is a simple prognostic tool capable of analyzing the most relevant predictive variables.



Fig. 1. To gaze into a crystal ball for a glimpse of the future has always been the dream of every doctor. Reliable assessment of the acute GVHD risk is crucial for making rational treatment decisions. During the process of donor selection and before discussing the choice of treatment with patients and their relatives, it is essential for physicians to integrate their knowledge with statistical or algorithmic tools capable of accurately predicting the likely incidence of GVHD. A more accurate prediction of acute and chronic GVHD would not only improve GVHD prophylaxis and conditioning regimens, but would also allow physicians to adapt their communication practices appropriately and to ensure that patients are supplied with effective

and comprehensive information on the pros and cons of transplantation, including the possibility of dying. This is particularly relevant for patients with chronic non malignant disorders, such as thalassemia. Maybe ANNs represent the crystal ball we are all looking for.

## 2. Patients and methods

### 2.1 Patients

We compared the prognostic performance of ANNs versus LR for predicting acute GVHD in a group of 78 beta-thalasssemia major patients given UD-HSCT (Caocci et al, 2010). The following clinical and immunogenetic paramenters were considered: recipient gender, recipient age, donor gender, donor age, the combination male recipient/female donor versus the other possible combinations, recipients and/or donors with positive CMV serology versus donor and recipient pairs with negative CMV serology, the Pesaro risk class at transplantation, HCV-RNA positivity, median infused CD34 cell dose, Treosulphan-containing conditioning regimen versus other regimens, HLA Class I mismatch, presence of HLA-A11, non permissive HLA-DPB1 disparity in GVHD direction, presence of the HLA-G 14-basepair deletion/deletion polymorphism in recipients, presence of the HLA-G 14-basepair deletion/deletion polymorphism in donors, heterozygosity for HLA-Cw ligand groups 1 and 2 in patients, recipient KIR ligand/donor activating KIR (recipient C1 absent/donor KIR2DS2 present versus the other 3 combinations; recipient C2 absent/donor KIR2DS1 present versus the other 3 combinations), recipient KIR ligand/donor inibitory KIR (patient C1 absent/donor KIR2DL2 present versus the other 3 combinations; recipient C1 absent/donor KIR2DL3 present versus the other 3 combinations; patient C2 absent/donor KIR2DL1 present versus the other 3 combinations), donor homozygosity for KIR haplotype A (Uhrberg, 2002; Colonna, 1995; Bassi, 2007; Cook, 2004; Harrison, 1993).

### 2.2 Statistical analysis

Patient, disease, and transplantation-related variables were expressed as median and range or percentage, as appropriate. For the HSCT outcome, patients were censored at the time of rejection, death, or last follow-up. Probabilities of overall survival (OS) and survival with transfusion independence (thalassemia-free survival) were estimated by the Kaplan-Meier method.

The following 24 independent variables were analyzed for their potential impact on aGVHD: recipient gender, recipient age, donor gender, donor age, the combination male recipient/female donor versus the other possible combinations, recipients and/or donors with positive CMV serology versus donor and recipient pairs with negative CMV serology, the Pesaro risk class at HSCT, HCV-RNA positivity, median infused CD34 cell dose, Treosulphan conditioning regimen versus other regimens, HLA Class I mismatch, presence of HLA-A11, non permissive HLA-DPB1 disparity in GVHD direction, presence of the HLA-G 14-basepair deletion/deletion polymorphism in recipients, presence of the HLA-G 14-basepair deletion/deletion polymorphism in donors, heterozygosity for HLA-Cw ligand groups 1 and 2 in patients, recipient KIR ligand/donor activatory KIR (recipient C1 absent/donor KIR2DS2 present versus the other 3 combinations; recipient C2 absent/donor KIR2DS1 present versus the other 3 combinations), recipient KIR ligand/donor inibitory KIR (patient C1 absent/donor KIR2DL2 present versus the other 3 combinations; recipient

C1 absent/donor KIR2DL3 present versus the other 3 combinations; patient C2 absent/donor KIR2DL1 present versus the other 3 combinations), donor homozygosity for KIR haplotype A.

## 2.3 Logistic regression

A binomial LR model with 24 independent variables (3 continuous and 21 categorical) was developed (Table 2). Acute GVHD was considered as a dichotomous dependent variable. Five consecutive random extractions were performed. For each extraction, patients were split into a learning data set (68 patients) and a test data set (10 patients). The independent variables were fitted into LR models via forward likelihood ratio test (chi-square difference) and stepwise selection. The chi-square test proposed by Hosmer and Lemeshow was used to analyze the goodness of fit: a finding of non-significance corresponds to the conclusion that the model adequately fits the data. Variables were retained only if their resulting p-value was ≤0.05. The final equation, developed through parameter estimates with standard errors, odds ratios and asymptotic 95% confidence intervals for all significant variables calculated, was applied to each case of the data test. A cut-off value of 0.5 was established for assigning the probability of GVHD: "GVHD yes" (1) or "GVHD no" (0). Sensitivity and specificity were determined in the learning and test data sets of each random extraction, sensitivity being the ratio between true positive and true negative plus false negative and specificity the ratio between true negative and true negative plus false positive. Mean sensitivity and specificity of LR obtained in five consecutive extractions were compared to ANN, using the chi-square test with Yate's correction. Statistical analysis was performed using SPSS® software, version 12 (SPSS Inc., Chicago, IL, USA)

## 2.4 Artificial neural networks

ANNs are capable of learning from observed data or examples and under certain conditions are able to approximate nonlinear functions with arbitrary precision. The technique was originally inspired by perceptions of how the human brain learns and processes information and since then has successfully been applied in many different fields, including mathematics, engineering, medicine, economics, meteorology, psychology, neurology, and many others. Although the predictive power of ANNs is often superior to that of other more traditional methods, they are still regarded as black-boxes where it is difficult for the user to gain insight into the influence of the independent variables in the prediction process. While ANNs are capable of learning the relationship between the input parameters and the controlled and uncontrolled variables, they do not generate information on the causal relationship between the input and output patterns. Several studies are currently underway to overcome this problem.

The structure of ANN usually consists of three layers (Fig. 2). The input layer accepts data sets from an external source that constitute inputs to the next layer of neurons. The next layer is called the hidden layer because its neuron values are not visible outside the net. The use of one or more hidden layers increases the net's learning abilities. The final layer is the output layer. Each single neuron is connected to the neurons of the previous layer through adaptable synaptic weights. Knowledge is generally stored as distributed patterns of activation in weight matrices.

The key feature of neural networks is that they learn the input/output relationship through training. The training data set includes a number of cases, each containing values for a range
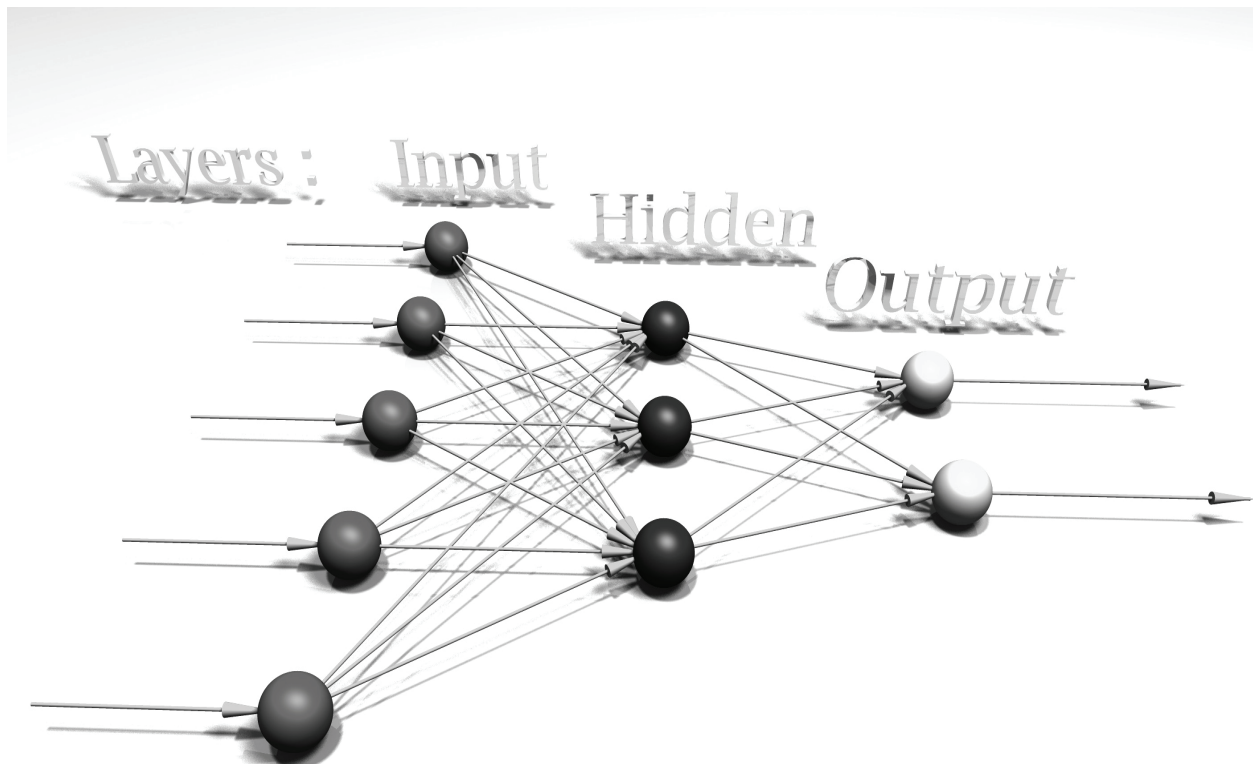
Fig. 2. The three layers structure of ANN

of well-matched input and output variables. The weighted connections between neurons in each layer are adjusted by a training algorithm to minimize error and provide accurate predictions on the training set. The outputs are the dependent variables that the network produces for the corresponding input. It is important that all learning data is supplied to the network as a data set. Once the input is propagated to the output neuron, this neuron compares its activation with the expected training output. If there is an error, the output neuron adjusts the connection weights to compensate the error by going backwards through the network. This step by step process is called backpropagation. The backpropagation (BP) algorithm and its variants are the most powerful learning algorithms in neural networks. By calculating the gradient vector of the error surface, the error gradually declines until all the expected outputs are correctly displayed.

The Neural Network ToolboxTM 6 of the software Matlab® 2008, version 7.6 (MathWorks, inc.) was used to develop a three layer feed forward neural network with the default tansigmoid transfer function in the hidden layer and linear transfer function in the output layer (Schwarzer et al., Demuth, 2008). The input layer of 24 neurons receives data that are processed in the hidden layer (30 neurons) and output layer (1 neuron). The output neuron predicts a number between 1 and 0 (goal), representing the event "GVHD yes" (1) or "GVHD no" (0), respectively. A cut-off value of 0.5 was established for assigning probability 1 or 0.

The architecture of ANN is schematized in Fig. 3. Input neurons receive data represented by the values of 24 independent variables processed in the hidden layer. The meaning of this process is to calculate interconnection weights between variables with the purpose of predicting an outcome and to calculate an error value by comparing this output value with the known outcome. The ANN attempts to minimize the error by adjusting the weights according to a learning algorithm (Linder et al., 2006). For the training procedure, we applied the 'on-line back-propagation' method on the same 5 sets of 68 patients previously analyzed by LR.

The 5 test phases utilized 10 patients randomly extracted from the entire cohort and not used in the training phase. A standard error of less than 10-2 was required. Mean sensitivity and specificity of the 5 consecutive data sets were determined in the data test and compared to LR. Because sensitivity and specificity in the 5 learning tests always resulted to be 100%, they were considered not comparable to LR.

| Input layer | Hidden Layer | Output layer |

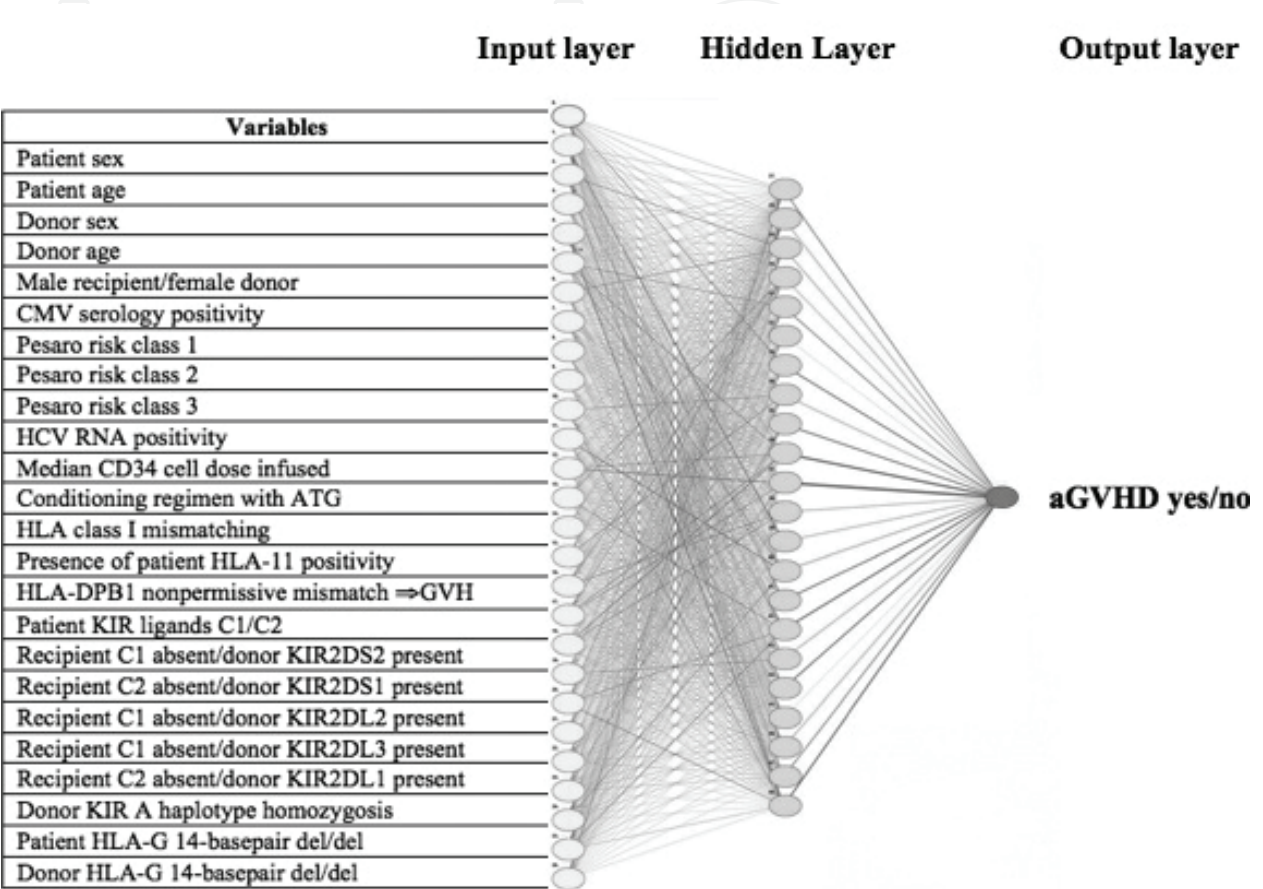| Variables |
| --- |
| Patient sex |
| Patient age |
| Donor sex |
| Donor age |
| Male recipient/female donor |
| CMV serology positivity |
| Pesaro risk class 1 |
| Pesaro risk class 2 |
| Pesaro risk class 3 |
| HCV RNA positivity |
| Median CD34 cell dose infused |
| Conditioning regimen with ATG |
| HLA class I mismatching |
| Presence of patient HLA-11 positivity |
| HLA-DPB1 nonpermissive mismatch ⇒GVH |
| Patient KIR ligands C1/C2 |
| Recipient C1 absent/donor KIR2DS2 present |
| Recipient C2 absent/donor KIR2DS1 present |
| Recipient C1 absent/donor KIR2DL2 present |
| Recipient C1 absent/donor KIR2DL3 present |
| Recipient C2 absent/donor KIR2DL1 present |
| Donor KIR A haplotype homozygosis |
| Patient HLA-G 14-basepair del/del |
| Donor HLA-G 14-basepair del/del |

aGVHD yes/no

Fig. 3. Architecture of the three-layer artificial neural network. The input layer of 24 neurons (independent variables) receives data that are processed in the hidden layer and output layer (1 neuron). The output neuron predicts a number between 1 and 0 (goal), representing the event "GVHD yes" (1) or "GVHD no" (0), respectively.

## 3. Results

Three-year Kaplan-Meier estimates for the 78 patients studied were 89.7% for survival, 76.9% for thalassemia-free survival, 11.5% for the cumulative incidence of rejection and 10.3% for TRM. Nine patients rejected the allograft and 7 died of transplantation-related complications (Figure 4).
Twenty-six patients (33.3%) developed grade II-IV acute GVHD (Figure 5).
In multivariate analysis, only donor KIR AA haplotypes were independently significantly correlated to acute GVHD in our cohort of 78 patients (p=0.037). However, we found a

positive trend for donor age (p=0.51), patient heterozygosity (C1/C2) for the HLA-Cw KIR ligands (p=0.56) and donor homozygosity (deletion/deletion) for the HLA-G 14-bp polymorphism (p=0.57) (Table 2).
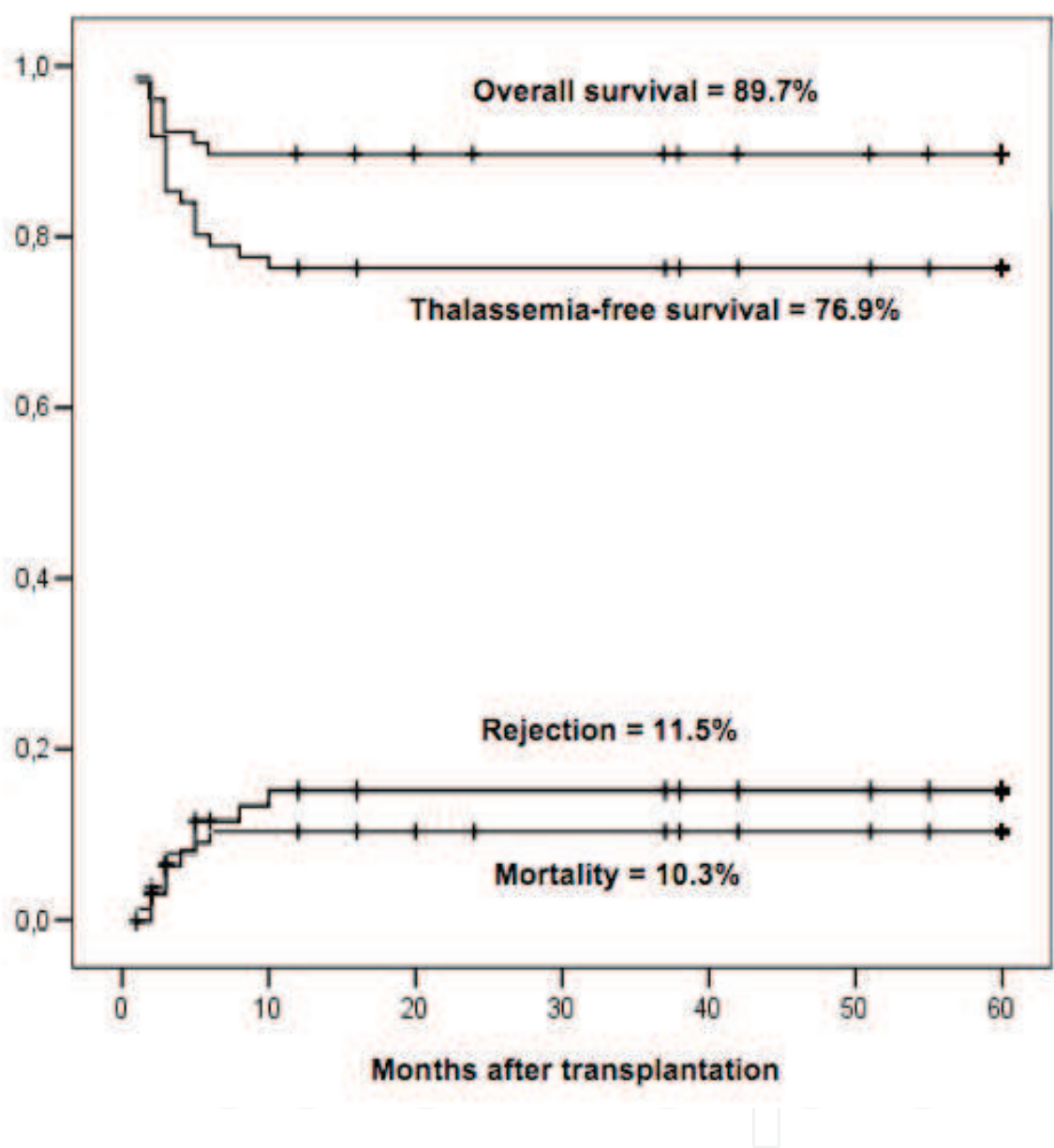


Fig. 4. Kaplan-Meier probabilities of overall survival, thalassemia-free survival, cumulative incidence of mortality and rejection in 78 thalassemia patients transplanted from an unrelated donor.

Table 3 shows the prognostic performance of LR and ANN in predicting acute GVHD in 5 consecutive randomly extracted training and test data sets. Sensitivity and specificity were determined in the learning and test data sets of each random extraction. Comparisons between LR and ANN on training data sets (5 consecutive extractions each composed of 68 patients) were not considered since ANN was able to recognize 100% of correct events by

means of its peculiar learning algorithm, whereas LR showed a mean value of 88.5% for specificity and 36.4% for sensitivity.
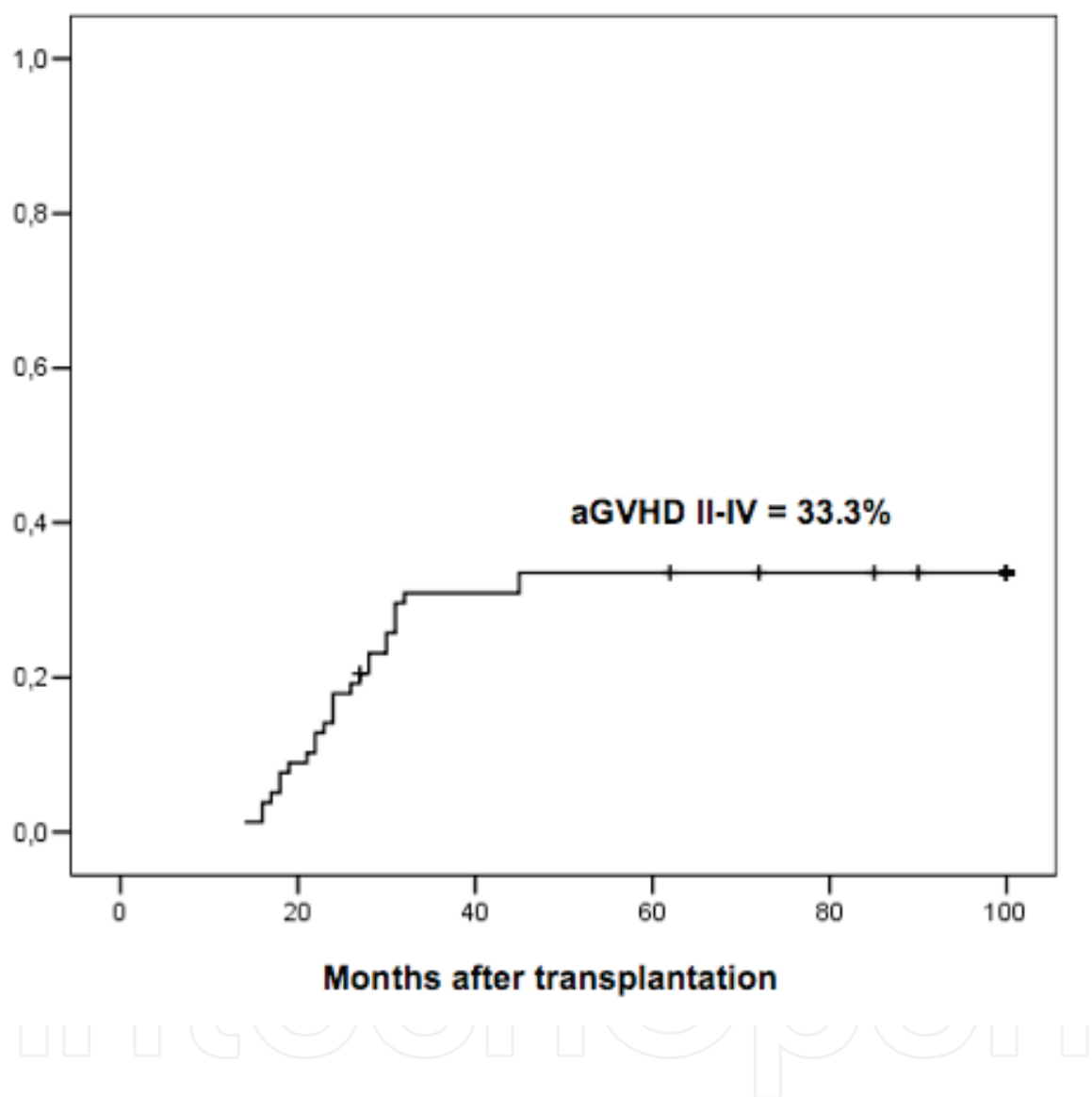


Fig. 5. Kaplan-Meier probabilities of cumulative incidence of acute GVHD in 78 thalassemia patients transplanted from an unrelated donor

In test data sets (5 extractions each composed of 10 patients), the mean specificity of LR was 80.5% compared to 90.1% of ANN (capability of predicting the absence of acute GVHD in patients who did not experience acute GVHD); this difference was not statistically significant. The mean sensitivity of LR was 21.7% compared to 83.3% of ANN (capability of predicting acute GVHD in patients who developed acute GVHD after HSCT). This difference was statistically significant (p<0.001).

| Variables | *p* value |
|---|---|
| Patient sex | .787 |
| Patient age | .110 |
| Donor sex | .496 |
| Donor age | .051 |
| Male recipient/female donor | .157 |
| CMV serology, positivity | .834 |
| Pesaro risk class 1 | .907 |
| Pesaro risk class 2 | .721 |
| Pesaro risk class 3 | .799 |
| HCV RNA positivity | .955 |
| Median CD34 cell dose infused | .315 |
| Conditioning regimen with ATG | .512 |
| HLA class I mismatching | 1.000 |
| Presence of patient HLA-11 positivity | .067 |
| HLA-DPB1 nonpermissive mismatch $\Rightarrow$GVH | .510 |
| Patient KIR ligands C1/C2 | .056 |
| Recipient C1 absent/donor KIR2DS2 present | .844 |
| Recipient C2 absent/donor KIR2DS1 present | .127 |
| Recipient C1 absent/donor KIR2DL2 present | .844 |
| Recipient C1 absent/donor KIR2DL3 present | .799 |
| Recipient C2 absent/donor KIR2DL1 present | .098 |
| Donor homozygosity for KIR A haplotype | .037 |
| Patient HLA-G 14-basepair del/del | .116 |
| Donor HLA-G 14-basepair del/del | .057 |

Table 2. Multivariate analysis of 24 independent variables and onset of acute GVHD in 78 thalassemia patients transplanted from an unrelated donor.

## 4. Conclusion and future research

Artificial intelligence (AI) is a field in which computers or software perform complex tasks by means of programming techniques or programs called expert systems that allow computers to "make decisions" by interpreting data and selecting among alternatives. ANNs are considered to be a branch of AI but what distinguishes them from classical AI are their learning and adaptive capabilities. ANNs first emerged as a collection of small individual interconnected artificial neurons or nodes constructed to mimic the processing properties of the biological neurons of the human brain. The information flows through the network in one direction, from the input nodes, through the hidden layers, to the output nodes. ANNs operate like a "black box" model, requiring no detailed information about the system. In neural net architectures, the connection strengths between nodes are the storehouses of knowledge and the learning process is primarily a process of adjusting these connection strengths. The network learns by adjusting its weights until it identifies a set of weights that produce the correct output for every sample input. ANNs have been applied with success in many different sectors and recent years have registered a growing interest also in the field of medicine.

| | | Acute GVHD | Observed Cases | LR Expected cases (%) | ANN Expected cases (%) |
|---|---|---|---|---|---|
| Extraction_1 | Training N=68 | No | 48 | 43 (89.6) | n.a.* |
| | | Yes | 20 | 8 (40) | |
| | Test N=10 | No | 6 | 5 (83.3) | 5 (83.3) |
| | | Yes | 4 | 1 (25) | 2 (50) |
| Extraction_2 | Training N=68 | No | 48 | 44 (91.7) | n.a* |
| | | Yes | 20 | 6 (30) | |
| | Test N=10 | No | 6 | 2 (66.7) | 6 (100) |
| | | Yes | 4 | 2 (50) | 4 (100) |
| Extraction_3 | Training N=68 | No | 45 | 40 (88.9) | n.a.* |
| | | Yes | 23 | 6 (26.1) | |
| | Test N=10 | No | 9 | 3 (66.7) | 9 (100) |
| | | Yes | 1 | 0 (0) | 1 (100) |
| Extraction_4 | Training N=68 | No | 47 | 41 (87.2) | n.a.* |
| | | Yes | 21 | 9 (42.9) | |
| | Test N=10 | No | 7 | 6 (85.7) | 5 (71.4) |
| | | Yes | 3 | 0 (0) | 3 (100) |
| Extraction_5 | Training N=68 | No | 47 | 40 (85.1) | n.a.* |
| | | Yes | 21 | 9 (42.9) | |
| | Test N=10 | No | 7 | 7 (100) | 7 (100) |
| | | Yes | 3 | 1 (33.3) | 2 (66.7) |
| Specificity % (mean) | No aGVHD | Training | | 88.5 | n.a.* |
| | | Test | | 80.5 | 90.1 (NS) |
| Sensitivity % (mean) | YES aGVHD | Training | | 36.4 | n.a.* |
| | | Test | | 21.7 | 83.3 (p<0.001) |

Table 3. Classification table and correct percentage of acute GVHD prediction for LR and ANN in five consecutive random extractions, splitting 78 thalassemia patients into a learning data set (68 patients) and a test data set (10 patients);

*n.a. = not applicable: comparisons between LR and ANN on training data sets were not considered since ANN was able to recognize 100% of correct events.

The process of clinical decision making that surrounds the choice of unrelated HSCT in a thalassemia patient is particularly difficult. The life span of patients with thalassemia is increasing, mainly attributable to a better control of iron overload.
HSCT is the only definitive curative approach to thalassemia. When the donor is an HLA-identical sibling, the probability of disease-free survival is between 80% and 90%. The worst results are obtained in the highest-risk category, particularly adult thalassemia patients, and in unrelated transplantation. Acute GVHD is the main cause of transplantation-related mortality in the setting of unrelated HSCT for thalassemia. Therefore, it becomes of fundamental importance to carefully evaluate the risk for the development of aGVHD before performing transplantation.
Several studies of unrelated HSCT in thalassemia patients have recently investigated a relatively large number of variables for their possible influence on the outcome and the

onset of acute GVHD. One of our recent reports shows that the HLA-G 14-bp polymorphism could be an important predictive factor for aGvHD; the HLA-CwAsn80 and HLA-CwLys80 molecules expressed by donor/recipient pairs as well as donor homozygosity for KIR haplotype A (AA) also seem to have a significant impact on transplantation outcome and acute GVHD incidence (Littera et al., 2010). In this study, multivariate analysis confirmed the impact of donor KIR haplotype AA on the incidence of acute GVHD (p=0.037). We also found a positive trend for donor age (p=0.51), patient heterozygosity (C1/C2), the HLA-Cw KIR ligands (p=0.56) and donor homozygosity (deletion/deletion) for the HLA-G 14-bp polymorphism (p=0.57).



Fig. 6. Researchers are currently involved in the study of several variables that seem to have a role in GVHD onset.  Unfortunately, a simple prognostic tool capable of analyzing the most relevant predictive variables and predicting GVHD is still missing. ANNs could represent this tool, capable of providing us with a more holistic vision of GVHD-phenomena.

In the field of medicine, several prognostic models have been developed for the prediction of outcome. The complicated clinical scenario of HSCT for chronic non malignant diseases would certainly benefit from a prognostic model based on a set of specific variables relevant to the development of GVHD. This report compared ANNs to LR in a cohort of 78 thalassemia patients transplanted from an unrelated donor. When comparing the prognostic performance of LR to ANN, the ability of predicting acute GVHD when the patient experienced acute GVHD (sensitivity) was 21.7% for LR versus 83.3% for ANN. This

difference was statistically significant (p<0.001). Moreover, the mean specificity of LR was 80.5% compared to 90.1% of ANN (capability of predicting the absence of acute GVHD in patients who did not experience acute GVHD) but this difference was not significant.

The advantage of ANNs can theoretically be explained by their ability to recognize complex relationships that possibly exist between independent and dependent variables, a typical challenge when dealing with medical data. By contrast, ANNs are considered as "black boxes" because of their hidden layer, which remains an obstacle to their acceptance. Moreover, ANNs are unable to calculate the weight of a single variable on the outcome, while LR determines a relative risk for each variable, building a complex equation of outcome prediction. And finally, LR is a widely used statistical method while ANNs are still being developed in the medical field.

There are some limitations to this study: the results are based on a series of information obtained from a relatively small, albeit homogeneous group of patients, selected according to the diagnosis, clinical characteristics and transplantation procedure. Moreover, data were analyzed retrospectively and could have been biased by the small number of cases assigned to the test data set, despite the 5 consecutive random extractions performed to increase this number. Therefore, the results obtained here need to be verified in larger prospective studies of transplanted patients. It should be of major interest to extend the application of neural networks to patients transplanted for pathologies other than talassemia and to include ulterior clinical parameters. Studies are currently underway to setup a neural network model capable of handling missing data in patient samplings.

In conclusion, ANN had a better prognostic performance than LR in predicting acute GVHD in our cohort of patients. This result is particularly important if we consider that GVHD is the major causative factor for TRM. Nevertheless, LR remains the "gold standard" of statistical predicting models in clinical settings.
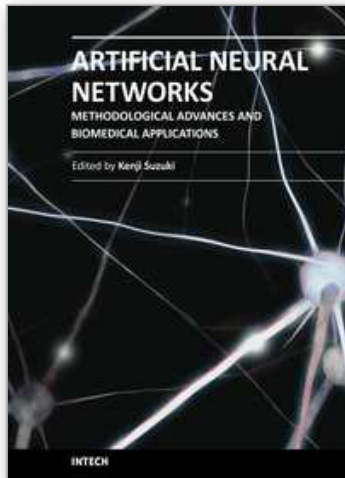
A combination of the two approaches so that each method complements the other has the potential to significantly improve the clinical decision-making process and the overall outcome of HSCT.

## 5. References

Abbod, MF.; Catto, JW.; Linkens, DA.; Hamdy, FC. (2007). Application of artificial intelligence to the management of urological cancer. *Journal of Urology*, 178, 4, (Oct 2007) 1150-1156, 00225347

Amendolia, SR.; Brunetti, A.; Carta, P.; Cossu, G.; Ganadu, ML.; Golosio, B.; Mura, GM.; Pirastru, MG. (2002). A real-time classification system of thalassemic pathologies based on artificial neural networks. *Medical Decision Making*, 22, 1, (Jan 2002) 18-26, 0272-989X

Annibale, B., Lahner, E. (2007). Assessing the severity of atrophic gastritis. *European Journal of Gastroenterology and Hepatology*, 2007; 19, 12, (Dec 2007) 1059–1063, 0954-691X

Bassi, P.; Sacco, E.; De Marco, V.; Aragona, M.; Volpe, A. (2007). Prognostic accuracy of an artificial neural network in patients undergoing radical cystectomy for bladder cancer: a comparison with logistic regression analysis. *British Journal of Urology*, 99, 5, (May 2007), 1007-1012, 1464-4096

Borgna-Pignatti, C.; Rugolotto, S.; De Stefano, P.; Zhao, H.; Cappellini, MD.; Del Vecchio, GC.; Romeo, MA.; Forni, GL.; Gamberini, MR.; Ghilardi, R.; Piga, A.; Cnaan A. (2004) Survival and complications in patients with thalassemia major treated with

transfusion and deferoxamine. *Haematologica*, 89, 10, (Oct 2004) 1187-1193, 1592-8721

Caocci, G.; Pisu, S.; Argiolu, F.; Giardini, C.; Locatelli, F.; Vacca, A.; Orofino, MG.; Piras, E.; De Stefano, P.; Addari, MC.; Ledda, A.; La Nasa, G. (2006). Decision-making in adult thalassemia patients undergoing unrelated bone marrow transplantation: quality of life, communication and ethical issues. *Bone Marrow Transplantation*, 37, 2, (Jan 2006) 165-169, 0268-3369

Caocci, G.; Baccoli, R.; Vacca, A.; Mastronuzzi, A.; Bertaina, A.; Piras, E.; Littera, R.; Locatelli, F.; Carcassi, C.; La Nasa, G. (2010). Comparison between an artificial neural network and logistic regression in predicting acute graft-vs-host disease after unrelated donor hematopoietic stem cell transplantation in thalassemia patients. *Experimental Hematology*, 38, 5, (May 2010) 426-433, 1873-2399

Colonna, M.; Samaridis, J. Cloning of immunoglobulin-superfamily members associated with HLA-C and HLA-B recognition by human natural killer cells. (1995). *Science,* 268, 5209, (Apr 1995) 405-408, 0036-8075

Cook, MA.; Milligan, DW.; Fegan, CD.; Darbyshire, PJ.; Mahendra, P.; Craddock, CF.; Moss, PA.; Briggs, DC. (2004). The impact of donor KIR and patient HLA-C genotypes on outcome following HLA-identical sibling hematopoietic stem cell transplantation for myeloid leukemia. *Blood*, 103, 4, (Feb 2004) 1521-1526, 0006-4971

Demuth, H.; Beale, M.; Hagan, M. (2008). *Neural Network Toolbox™ 6. User's Guide,* The MathWorks, Inc., 0-534-95049-3, Natick, MA.

Fleischhauer, K.; Locatelli, F.; Zecca, M.; Orofino, MG.; Giardini, C.; De Stefano, P.; Pession, A.; Iannone, AM.; Carcassi, C.; Zino, E.; La Nasa, G. (2006). Graft rejection after unrelated donor hematopoietic stem cell transplantation for thalassemia is associated with nonpermissive HLA-DPB1 disparity in host-versus-graft direction. *Blood*, 107, 7, (Apr 2006), 2984-2992, 0006-4971

Hansen, JA.; Gooley, TA.; Martin, PJ.; Appelbaum, F.; Chauncey, TR.; Clift, RA.; Petersdorf, EW.; Radich, J.; Sanders, JE.; Storb, RF.; Sullivan, KM.; Anasetti, C. (1998). Bone marrow transplants from unrelated donors for patients with chronic myeloid leukaemia. *New England Journal of Medicine*, 228, 14, (Apr 1998) 962-968, 0028-4793

Harrell, FE.; Lee, KL.; Mark DB. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, 15, 4, (Feb 1996) 361-387, 0277-6715

Harrison, GA.; Humphrey, KE.; Jakobsen, IB.; Cooper, DW. (1993). A 14 bp deletion polymorphism in the HLA-G gene. *Human Molecular Genetics*, 2, 12 (Dec 1993) 2200, 0964-6906

La Nasa, G.; Caocci, G.; Argiolu, F.; Giardini, C.; Locatelli, F.; Vacca, A.; Orofino, MG.; Piras, E.; Addari, MC.; Ledda, A.; Contu, L.; (2005) Unrelated donor stem cell transplantation in adult patients with thalassemia. *Bone Marrow Transplantation*, 36, 11, (Dec 2005) 971-975, 0268-3369

La Nasa, G.; Littera, R.; Locatelli, F.; Giardini, C.; Ventrella, A.; Mulargia, M.; Vacca, A.; Orrù, N.; Orrù, S.; Piras, E.; Giustolisi, G.; Lisini, D.; Nesci, S., Caocci, G., Carcassi, C. (2007). Status of donor-recipient HLA class I ligands and not the KIR genotype is predictive for the outcome of unrelated hematopoietic stem cell transplantation in beta-thalassemia patients. *Biology of Blood and Marrow Transplantation*, 13, 11, (Nov 2007) 1358-1368, 1083-8791

La Nasa, G.; Littera, R.; Locatelli, F.; Lai, S.; Alba, F.; Caocci, G.; Lisini, D.; Nesci, S.; Vacca, A.; Piras, E.; Bernardo, ME.; Di Cesare-Merlone, A.; Orrù, S.; Carcassi, C. (2007). The human leucocyte antigen-G 14-basepair polymorphism correlates with graft-versus-host disease in unrelated bone marrow transplantation for thalassaemia. *British Journal of Haematology*, 139, 2, (Oct 2007) 284-288, 0007-1048

Linder, R.; König, IR.; Weimar, C.; Diener, HC.; Pöppl, SJ.; Ziegler, A. (2006). Two models for outcome prediction - a comparison of logistic regression and neural networks. *Methods of Information in Medicine*, 45, 5, (May 2006) 536-540, 0026-1270

Lisboa, PJ.; Taktak, AF. The use of artificial neural networks in decision support in cancer: a systematic review. *Neural Networks*, 19, 4, (May 2006) 408-415, 0893-6080

Littera, R.; Orrù, N.; Vacca, A.; Bertaina, A.; Caocci, G.; Mulargia, M.; Giardini, C.; Piras, E.; Mastronuzzi, A.; Vinti, L.; Orrù, S.; Locatelli, F.; Carcassi, C.; La Nasa, G. (2010). The role of killer immunoglobulin-like receptor haplotypes on the outcome of unrelated donor haematopoietic SCT for thalassaemia. *Bone Marrow Transplantation*, Feb 2010, (Epub ahead of print) 1476-5365

Lucarelli, G.; Galimberti, M.; Polchi, P.; Angelucci, E.; Baronciani, D.; Giardini, C.; Politi, P.; Durazzi, SM.; Muretto, P.; Albertini, F. (1990) Bone marrow transplantation in patients with talassemia. *New England Journal of Medicine*, 322, 7, (Feb 1990) 417–421, 0028-4793

Lucarelli, G.; Giardini, C.; Angelucci, E. (1997) Bone marrow transplantation in talassemia, In: *Blood Stem Cell Transplantation*, Winter JN, (Kluwer Academic), 305-315, 0-7923-4260-7, Boston, MA

Pace, F.; Savarino, V. (2007). The use of artificial neural network in gastroenterology: the experience of the first 10 years. *European Journal of Gastroenterology and Hepatology*, 19, 12, (Dec 2007) 1043-1045, 0954-691X

Rajan, P.; Tolley, DA. (2005). Artificial neural networks in urolithiasis. *Current Opinion in Urology*, 15, 2, (Mar 2005) 133-137, 0963-0643

Royston, P. (2000). A strategy for modelling the effect of a continuous covariate in medicine and epidemiology. *Statistics in Medicine*, 19, 14, (Jul 2000) 1831-1847, 0277-6715

Schwarzer, G.; Vach, W.; Schumacher, M. (2000). On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology. *Statistics in Medicine,* 19, 4, (Feb 2000) 541-561, 0277-6715

Stephan, C.; Cammann, H.; Meyer, H.; Lein, M.; Jung, K. (2007). PSA and new biomarkers within multivariate models to improve early detection of prostate cancer. *Cancer Letter*, 249, 1, (Apr 2007) 18-29, 0304-3835

Uhrberg, M.; Parham.; Wernet, P. (2002). Definition of gene content for nine common group B haplotypes of the Caucasoid population: KIR haplotypes contain between seven and eleven KIR genes. *Immunogenetics*, 54, 4, (Jul 2002) 221-229, 0093-7711

Zini G.; d'Onofrio, G. (2003). Neural network in hematopoietic malignancies. *Clinica Chimica Acta*, 333, 2 (Jul 2003) 195-201, 0009-8981

Zini, G. (2005) Artificial intelligence in hematology. Hematology, 10, 5, (Oct 2005), 393-400, 1024-5332

**Artificial Neural Networks - Methodological Advances and Biomedical Applications**

Edited by Prof. Kenji Suzuki

Artificial neural networks may probably be the single most successful technology in the last two decades which has been widely used in a large variety of applications in various areas. The purpose of this book is to provide recent advances of artificial neural networks in biomedical applications. The book begins with fundamentals of artificial neural networks, which cover an introduction, design, and optimization. Advanced architectures for biomedical applications, which offer improved performance and desirable properties, follow. Parts continue with biological applications such as gene, plant biology, and stem cell, medical applications such as skin diseases, sclerosis, anesthesia, and physiotherapy, and clinical and other applications such as clinical outcome, telecare, and pre-med student failure prediction. Thus, this book will be a fundamental source of recent advances and applications of artificial neural networks in biomedical areas. The target audience includes professors and students in engineering and medical schools, researchers and engineers in biomedical industries, medical doctors, and healthcare professionals.

**How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Giovanni Caocci, Roberto Baccoli and Giorgio La Nasa (2011). The Usefulness of Artificial Neural Networks in Predicting the Outcome of Hematopoietic Stem Cell Transplantation, Artificial Neural Networks - Methodological Advances and Biomedical Applications, Prof. Kenji Suzuki (Ed.), ISBN: 978-953-307-243-2, InTech, Available from: http://www.intechopen.com/books/artificial-neural-networks-methodological-advances-and-biomedical-applications/the-usefulness-of-artificial-neural-networks-in-predicting-the-outcome-of-hematopoietic-stem-cell-tr

# INTECH
open science | open minds