Evaluation of prognostic factors and prediction of chronic wound healing rate by machine learning tools

Marko Robnik-Šikonja[1], David Cukjati[2], and Igor Kononenko[1]

[1] University of Ljubljana, Faculty of Computer and Information Science,
Tržaška 25, 1001 Ljubljana, Slovenia
Marko.Robnik@fri.uni-lj.si,
[2] University of Ljubljana, Faculty of Electrical Engineering,
Tržaška 25, 1001 Ljubljana, Slovenia
David.Cukjati@fe.uni-lj.si,

Abstract. In more than a decade of clinical use of electrical stimulation to accelerate the chronic wound healing each patient and wound were registered and a wound healing process was weekly followed. The controlled study involved a conventional conservative treatment, sham treatment, biphasic pulsed current, and direct current electrical stimulation. A quantity of available data suffices for an analysis with machine learning methods.
So far only a limited number of studies have investigated the wound and patient attributes which affect the chronic wound healing. There is none to our knowledge to include the treatment attributes. The aims of our study are to determine effects of the wound, patient and treatment attributes on the wound healing process and to propose a system for prediction of the wound healing rate.
In the first step of our analysis we determined which wound and patient attributes play a predominant role in the wound healing process. Then we investigated a possibility to predict the wound healing rate at the beginning of the treatment based on the initial wound, patient and treatment attributes. Finally we discussed the possibility to enhance the wound healing rate prediction accuracy by predicting it after a few weeks of the wound healing follow-up.
By using the attribute estimation algorithms ReliefF and RReliefF we obtained a ranking of the prognostic factors which was comprehensible to field experts. We also used regression and classification trees to build models for prediction of the wound healing rate. The obtained results are encouraging and may form a basis of an expert system for the chronic wound healing rate prediction. If the wound healing rate is known, then the provided information can help to formulate the appropriate treatment decisions and orient resources to those individuals with poor prognosis.

1 Introduction

Skin is vital organ in a sense that loss of the substantial fraction of its mass immediately threatens life of the individual. Such a loss can result suddenly, either from a fire or mechanical accident, but can also occur in a chronic manner, as in skin ulcers.

In more than a decade of clinical use of an electrical stimulation to accelerate the chronic wound healing at the Institute of the Republic Slovenia for Rehabilitation in

Ljubljana each patient and wound were registered and the wound healing process was weekly followed. At the beginning of the study in 1989, wounds were randomly assigned into four treatment groups: conservative treatment, sham treatment, biphasic current stimulation and direct current stimulation. Jerčinović et al. (1994) proved that stimulated wounds are healing significantly faster than only conservatively or sham treated wounds and Karba et al. (1997) proved that electrical stimulation with direct current is effective only if the positive electrode is placed on the wound surface, which is an invasive method, therefore only stimulation with biphasic current pulses has been in regular use ever since. However, dynamics of the wound healing process does not depend only on the type of the treatment, but depends also on the wound and patient attributes.

The aims of our study are to determine effects of the wound, patient and treatment attributes on the wound healing process and to propose a system for prediction of the wound healing rate.

So far only a limited number of studies have investigated the wound and patient attributes which affect the chronic wound healing. Skene et al. (1992) found that with the presence of the graduated compression the healing occurred more rapidly for patients with smaller initial ulcer area, shorter duration of ulceration, younger age and when no deep vein involvement was detected on photoplethysmography. The measurement of the ulcer area was found to be the strongest predictor of the ulcer healing. Birke et al. (1992) found that a time to complete the wound closure is related to the wound depth and wound diameter. Johnson (1997) found four factors influencing the vascular ulcer healing: ABpI (ankle/brachial pressure index), liposclerosis, edema, wound status and ulcer area. Lyman et al. (1970) found significant relationship between the wound healing rate and bacterial load.

None of listed studies included the treatment attributes. Presently, the quantity of available data permits the use of statistical tools and artificial intelligence methods for analysis of the healing process, as well as of the effects of different therapeutic modalities. In the first step of our analysis we determine which wound and patient attributes play a predominant role in the wound healing process. Then we investigate a possibility to predict the wound healing rate at the beginning of treatment based on the initial wound, patient and treatment attributes. Finally we discuss a possibility to enhance the wound healing rate prediction accuracy by predicting it after a few weeks of the wound healing follow-up.

The paper is organized into 5 Sections. Section 2 discusses the problem and the collected data. Section 3 describes machine learning tools and algorithms used. In Section 4 we present our findings. Final Section contains discussion.

2 The problem and the data set

During more than a decade of clinical use of electrical stimulation, data concerning patients, wounds, and their treatment were collected. 266 patients with 390 wounds were recorded in computer database up to date. Unfortunately many patient and wound data are missing and not all wounds were followed regularly or until the complete wound closure which is a common problem of clinical trials. The wound case inclusion criteria was the initial wound area larger than $1\mathrm{cm}^2$ and at least four weeks (or until the complete wound closure) follow up of the wound healing process. It was were fulfilled in 300 wound cases (214 patients).

Among these 300 wound cases the observation period in 174 cases lasted until the complete wound closure and was shorter in 126 cases. In these cases the time to the complete wound closure was estimated from the wound area measurements obtained during the observation period (Cukjati et al., 2000; 2001a). No significant difference between the actual time to complete the wound closure and the estimated one (from four or more weeks of wound healing observation) was observed.

Wounds in the database are described with length, width, depth and grade. Because the time to the complete wound closure was highly dependent on the initial wound extent, a measure of the wound healing rate was defined as an average advance of the wound margin towards the wound centre and was calculated as the average wound radius (the initial wound area divided by the initial perimeter and multiplied by 2) divided by the time to the complete wound closure (Cukjati et al., 2001a):

$$\Theta = 2\frac{S_0}{p_0 T} \quad \left[\frac{\mathrm{mm}}{\mathrm{day}}\right] \tag{1}$$

Distribution of the wound healing rate was not normal and could not be transformed to normal distribution; non-parametric statistical analysis was therefore used. The attributes we used in our study are briefly described below, further details can be found in (Cukjati et al., 2001b).

2.1 Wound attributes

For an evaluation of the efficacy of a particular treatment modality or for evaluation of influence of the wound and patient attributes on the wound healing it is necessary to periodically follow the wound healing process. It was demonstrated (Cukjati et al., 2001a) that following wound area is sufficient to determine wound healing process dynamics. Further it was shown that wound shape can be approximated with an ellipse and it is thus enough to periodically follow mutually perpendicular diameters (largest wound diameter and diameter perpendicular to it) of the wound. From the measured diameters the wound area, the perimeter and the width to length ratio were calculated.

The wound depth measurement is invasive, because we have to enter our measuring device into the wound. As an alternative measure of the wound, grading systems were presented. We used the four stage Shea grading system (Shea, 1975). The wound depth and grade were collected only at the beginning of the treatment. The wound depth was measured in 43% of cases and the wound grade determined in 94%. Since the wound depth was strongly correlated to the wound grade and the wound depth values were often missing, the depth was omitted from the further analysis. Also due to a strong correlation between the initial wound area and the perimeter, the perimeter was omitted from the further analysis.

Other collected wound attributes were the wound type, location (trochanter, sacrum, gluteus and other), elapsed time from spinal cord injury to wound appearance (InjuryAppear), and elapsed time from the wound appearance to the beginning of the treatment (AppearStart). The major wound aetiology was the pressure ulceration (82.7% of the cases). Other aetiologies were the arterial ulceration (1.0%), neurotrophic ulceration (6.3%), traumatic ulceration (6.0%), and vascular ulceration (3.7%).

2.2 Patient attributes

Recorded patient attributes were age, sex, total number of wounds, diagnosis and, in the case of the spinal cord injury, the degree of spasticity. The most frequent diagnosis was the spinal cord injury (71.7%). The trauma appeared in 11.3% of cases, diabetes mellitus in 7.3%, geriatrics in 3.3%, multiple sclerosis in 3.0%, and venous diseases in 3.0% of wound cases.

2.3 Treatment attributes

The wounds were randomly assigned into four treatment groups. 54 (18.0%) wounds received only the conservative treatment (Feedar and Kloth, 1990), while 23 (7.7%) wounds received also the sham treatment, 42 (14%) wounds received direct current and 181 (60.3%) wounds received biphasic current stimulation. In the sham treated group electrodes were applied to the intact skin on both sides of the wound for two hours daily and connected to stimulators, in which the power source was disconnected and delivered no current. The wounds treated with the direct current were stimulated with 0.6mA (Karba et al., 1997) for half an hour, an hour or two hours daily. Biphasic current stimulated wounds received biphasic, charge-balanced current pulses (Karba et al., 1991) for half an hour, an hour or two hours daily with electrodes placed on both sides of the wound.

The treatment attributes were the type of the treatment and daily duration of the electrical stimulation. The electrically stimulated wounds healed at higher rate and extent then other wounds. Over 90% of electrically stimulated wounds healed within 6 weeks. Only 70% of sham treated wounds and 72% of only conservative treated wounds healed within the same period.

2.4 Mathematical model of wound healing rate

We observed that the dynamics of changing of the wound area over time has a delayed exponential behaviour. Delayed exponential equation is therefore the structure of mathematical model of the wound healing process and by fitting this model to a particular chronic wound case, parameters of the model are calculated. We need at least four measurements (performed in at least three weeks) of the wound area before we can estimate the parameters of the mathematical model. We estimated the time to the complete wound closure from the parameters of the mathematical model (Cukjati et al., 2001a). We considered the wound healed when its estimated area was smaller than five percent of the initial value and smaller than $1\mathrm{cm}^2$. The estimated wound healing rate was calculated with Equation (1) .

The estimated wound healing rates for all wound cases were then compared to the actual ones which were calculated from the observed times to the complete wound closure. We found that the estimated wound healing rate after at least four weeks of wound follow-up did not differ significantly from the actual one (Cukjati et al., 2001b). If a wound was followed only three weeks or less the difference was significant. From the known structure of the mathematical model the wound healing rate can be predicted after at least four weeks of follow-up. In clinical trials four weeks is a short period, but in clinical practice shorter time for the treatment outcome prediction would be desirable.

3 Machine learning algorithms and results

Presently, the quantity of the available data permits the use of machine learning methods. Results of the statistical analysis (Cukjati et al., 2001b) showed that the wound healing rate directly depends on the wound treatment and wound grade, while interactions of the other wound and patient attributes with the wound healing rate are not easy to determine.

We employed machine learning algorithms and built tree based models (regression and classification trees) to predict the wound healing rate based on the initial wound, patient and treatment data. We also considered the estimated wound healing rate based on the mathematical model (Cukjati et al., 2001b) and built trees for prediction of the wound healing rate after one, two, three, four, five and six weeks of follow-up. We tested several algorithms for attribute estimation which were used in a feature subset selection and for selection of splits in interior nodes of the tree based models. In the earlier stages of our data exploration we found that Relief algorithms, RReliefF (Robnik Šikonja and Kononenko, 1997) for regression and ReliefF (Kononenko et al., 1997) for classification problems, were the most effective concerning error and comprehensiveness of the learned models, so we present results only for them. The results for mean squared error and mean absolute error as attribute estimators in regression (Breiman et al., 1984) and Gain ratio (Quinlan, 1993) in classification are not reported because they are inferior to the presented ones.

Relief algorithms are one of the most successful feature subset selection algorithms (Dietterich, 1997). The idea of RReliefF and ReliefF algorithms is to evaluate partitioning power of attributes according to how well their values distinguish between similar observations. An attribute is given a high score if it separates similar observations with different prediction values and does not separate similar observations with similar prediction values. RReliefF and ReliefF sample the space of observations, compute the differences between predictions and values of the attributes and form a kind of statistical measure for the proximity of the probability densities of the attribute and the predicted value. Assigned quality estimates are in the range $[-1, 1]$, however, values below zero are assigned to completely irrelevant (random) attributes. The quality estimate $W[A]$ of attribute A assigned by ReliefF or RReliefF can be interpreted in two ways: as the difference of probabilities and as the proportion of explanation of the concept.

$$\begin{aligned} W[A] = {} & P(\text{different value of A}|\text{near instance with different prediction}) \\ & - P(\text{different value of A}|\text{near instance with same prediction}) \end{aligned} \tag{2}$$

Equation (2) forms the basis for the difference of probabilities interpretation of the quality estimations of the Relief algorithms: the difference of the probability that two instances have different value of the attribute if they have different prediction value and the probability that two instances have different value of the attribute if they have similar prediction values. These two probabilities contain additional condition that the instances are close in the problem space and form an estimate of how well the values of the attribute distinguish between the instances that are near to each other. As it turned out this interpretation was quite difficult for human comprehension. Negated similarity (different values) and subtraction of the probabilities are difficult to comprehend for human experts because they do not contain clear mental image.

Another view on attributes' estimates of ReliefF and RReliefF is possible with the proportion of prediction values which the attribute helps to determine (Robnik-Šikonja and Kononenko, 2001). As the number of examples n goes to infinity the quality estimates $W[A]$ computed from m cases for each attribute converge to the number of changes in the predicted values the attribute is responsible for (R_A):

$$\lim_{n \to \infty} W[A] = \frac{R_A}{m} \tag{3}$$

The interpretation of Relief's weights as the proportion of explained concept is more comprehensible than the interpretation with the difference of two probabilities as confirmed by human experts. Equation (3) has only one non probabilistic term (a simple ratio), which can be understand taking the unlimited number of examples into account. The actual quality estimates for attributes in given problem are approximations of these ideal estimates which occur only with abundance of data.

3.1 Estimation of the prognostic factors

The attributes' quality estimates (see Table 1) calculated using RReliefF revealed that the initial wound area, followed by the patients' age and time from wound appearance to treatment beginning are the most prognostic attributes. Important prognostic attributes are also wound shape (width to length ratio), location and type of treatment. The attribute "Model estimation" represents the estimated wound healing rate calculated from a model of wound healing dynamics. We see (as expected) that its quality estimate increases as we add more and more observations (from 0.0 to 0.67).

Table 1. The quality of wound, patient and treatment attributes assigned by RReliefF.

Attribute	0 weeks	1 week	2 weeks	3 weeks	4 weeks	5 weeks	6 weeks
Area (mm^2)	0.135	0.168	0.171	0.161	0.127	0.123	0.122
Age (year)	0.123	0.114	0.094	0.095	0.096	0.092	0.094
AppearStart (week)	0.119	0.121	0.104	0.131	0.121	0.114	0.115
Width to length ratio	0.096	0.098	0.099	0.095	0.103	0.108	0.113
Location	0.085	0.084	0.085	0.081	0.081	0.081	0.081
Treatment	0.066	0.058	0.051	0.052	0.050	0.051	0.051
InjuryAppear (month)	0.062	0.065	0.044	0.050	0.035	0.040	0.039
Daily duration of treatment (min)	0.046	0.039	0.031	0.035	0.025	0.025	0.026
Grade	0.046	0.039	0.057	0.048	0.048	0.047	0.043
Diagnosis	0.039	0.039	0.038	0.038	0.038	0.038	0.037
Aetiology	0.027	0.025	0.026	0.024	0.024	0.0239	0.024
Model estimation	0.000	0.399	0.602	0.626	0.663	0.659	0.670

3.2 Regression trees

We built regression trees with the improved CORE learning system (Robnik Šikonja, 1997). We used linear equations in the leaves of the tree to model the wound healing rate (we also tried k-NN, median value, mean value, and kernel regression) and stopping rule of minimal five cases in a leaf. To obtain smaller and more comprehensible trees and to get better generalization the trees were pruned. Since the sample size (n=300) was moderate, we could not afford a separate testing set but we rather used the 10-fold cross-validation as an error estimation method. An error of the regression trees was measured with the relative squared error (relative error, RE) (Breiman et al., 1984). The relative error is always nonnegative and should be less than 1 for models of reasonable quality. Trees with the relative error close to 0 produce excellent prediction of the wound healing rate and trees with the relative error around 1 or even greater than 1 produce poor prediction. Some authors are using proportion of the variance explained by the regression tree as a measure of the error. It is calculated as (1 - RE). Although the terminology is not quite appropriate (Breiman et al., 1984), we also used this measure to compare results.

Left hand side of Figure 1 summarizes predictive accuracy of the learned trees. The relative error of the prediction at the beginning of the treatment had relative squared error greater than one, i.e., the resulting regression trees were not usable. By adding the model estimate of the wound healing rate after one week of follow-up we reduced the relative squared error to 0.64, (36% of variance was explained by the tree). After two weeks 65% and after three weeks 82% of variance was explained. Accuracy was slowly approaching 94% of explained variance with six weeks of follow-up.

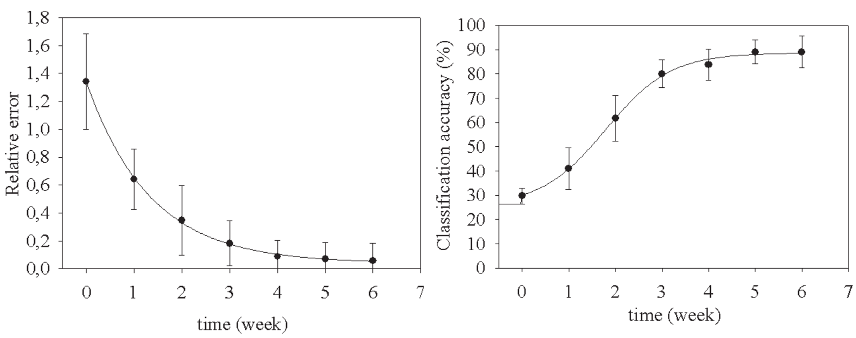

Fig. 1. The relative error of the regression trees (left) and classification accuracy of the classification trees (right) for the wound healing rate prediction as a function of observation time.

Regression trees are more useful than classification trees because the wound healing rate was estimated as continuous variable and can be directly observed in the tree. The minimal follow-up period is two weeks. After five weeks the predicted wound healing rate is equal to the healing rate estimated by the model. However, the predicted wound healing rate with shorter period of follow-up depends also on the wound, patient and treatment attributes. The regression tree built after two weeks of follow-up is presented in Figure 2. The type of the treatment is indirectly included in the regression trees as daily duration of the treatment, which was zero in case of conservative treated wounds. Important prognostic attributes seem to be the wound area, grade, shape (width to length ratio), patients age, elapsed time from the spinal cord injury to the wound appearance and elapsed time from the wound appearance to the beginning of the treatment.

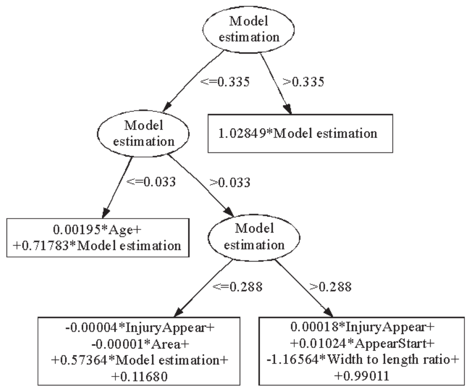

Fig. 2. The regression tree with linear equations in the leaves for prediction of the wound healing rate after two weeks of treatment.

3.3 Classification trees

Sometimes we do not need the exact quantity of the wound healing rate so we decided to divide it into four categories (classes) according to Table 2 and then perform analysis on such classification problem.

Table 2. Dividing 300 wound cases into four classes according to their wound healing rate.

class	condition[mm/day]	no. of cases	a priori
NON HEALING WOUNDS	$\Theta \leq 0.095$	77	0.257
SLOW HEALING WOUNDS	$0.095 < \Theta \leq 0.180$	77	0.257
MEDIUM HEALING WOUNDS	$0.180 < \Theta \leq 0.300$	67	0.223
FAST HEALING WOUNDS	$\Theta > 0.300$	79	0.263

We built classification trees with ReliefF as attribute estimation measure. We used a median value of the cases as the prediction model in the leaves of the tree (we also tried k-NN and mean value) and a stopping rule of at least five wound cases in a leaf. To obtain smaller and more comprehensible trees and to get better generalization the trees were pruned. We used the 10-fold cross-validation as an error estimation method. The accuracy of the trees was measured as a proportion of the correctly classified test samples.

At the beginning of the wound treatment, only initial wound, patient and treatment data are available. The resulting classification tree accuracy at the beginning of the treatment was 30%, which is not much above a priori probability of the most probable class (26%). Adding model estimate of the wound healing rate after one week of follow-up to the data set improved the classification accuracy to 41%. With data available for two weeks the classification accuracy was 62% and with three weeks 80%. Afterwards it is slowly approaching 90% with six weeks of follow-up. Right hand side of Figure 1 summarizes the results. In the trees built after two weeks of follow-up only the model estimate of the wound healing rate can be found in the tree nodes.

We found out that accurate prediction of the wound healing rate is possible with data available for at least three weeks of follow-up. Therefore, with classification trees we also managed to shorten the time of follow-up for one week compared to the mathematical model. Only rough estimate of the wound healing rate is possible after two weeks.

4 Discussion

We estimated the prognostic power of the wound, patient and treatment attributes with RReliefF algorithm. The obtained results revealed that the initial wound area, followed by the patients' age and time from the wound appearance to the treatment beginning are the most important prognostic attributes. Important prognostic attributes are also the wound shape (width to length ratio), location, and type of the treatment. The filed experts agree with these findings and accept this ranking because they understood the meaning of these quality estimates.

The dynamics of wound healing can be accurately predicted after at least four weeks of the wound healing process follow-up. Therefore for accurate wound healing rate estimation, wounds should be followed at least four weeks. In clinical practice the wound healing rate or the time to complete wound closure should be estimated as soon as possible to select a proper treatment and thus to improve patient care. Prediction of the wound healing rate from the initial wound, patient and treatment data collected in our database was not possible. The best prognostic factors are weekly follow-up measurements of wound area. We determined that the minimal follow-up period is two weeks. After three weeks we were able to predict the wound healing rate with 80% classification accuracy (using the classification trees), and explain 82% of the variance with the regression trees. The best results were obtained using regression trees with linear equations in the leaves. In the literature also other wound and patient attributes were reported to have prognostic value. By considering those additional attributes our prediction might be even more accurate.

The presented regression trees in combination with the mathematical model of the wound healing process dynamics possibly present a basis of an expert system for the chronic wound healing rate prediction. If the wound healing rate is known, the provided information can help to formulate the appropriate management decisions, reduce the cost, and orient resources to those individuals with poor prognosis.

Bibliography

J. A. Birke, A. Novick, C. A. Patout, and W. C. Coleman. Healing rates of plantar ulcers in leprosy and diabetes. *Leprosy Rev.*, 63:365–374, 1992.

L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and regression trees*. Wadsworth Inc., Belmont, California, 1984.

D. Cukjati, R. Karba, S. Reberšek, and D. Miklavčič. Modeling of chronic wound healing dynamics. *Med.Biol.Eng.Comput.*, 38:339–347, 2000.

D. Cukjati, S. Reberšek, and D. Miklavčič. A reliable method of determining wound healing rate. *Med.Biol.Eng.Comput.*, 2001a. (in press).

D. Cukjati, M. Robnik-Šikonja, S. Reberšek, I. Kononenko, and D. Miklavčič. Prognostic factors, prediction of chronic wound healing and electrical stimulation. *Medical & Biological Engineering & Computing*, 2001b. (submitted).

T. G. Dietterich. Machine learning research: Four current directions. *AI Magazine*, 18 (4):97–136, 1997.

J. A. Feedar and L. C. Kloth. Conservative management of chronic wounds. In L. C. Kloth and J.M. McCulloch, editors, *Wound Healing: Alternatives in Management*, pages 135–172. F. A. Davis Co., Philadelphia, 1990.

A. Jerčinović, R. Karba, L. Vodovnik, A. Stefanovska, P. Krošelj, R. Turk, I. Džidić, H. Benko, and R. Šavrin. Low frequency pulsed current and pressure ulcer healing. *IEEE Trans.Rehab.Eng.*, 2(4):225–233, 1994.

M. Johnson. Using cluster analysis to develop a healing typology in vascular ulcers. *J.Vasc.Nurs.*, 15:45–49, 1997.

R. Karba, D. Šemrov, L. Vodovnik, H. Benko, and R. Šavrin. DC electrical stimulation for chronic wound healing enhancement. Part 1. Clinical study and determination of electrical field distribution in the numerical wound model. *Bioelectrochemistry and Bioenergetics*, 43:265–270, 1997.

I. Kononenko, E. Šimec, and M. Robnik-Šikonja. Overcoming the myopia of inductive learning algorithms with RELIEFF. *Applied Intelligence*, 7:39–55, 1997.

I. R. Lyman, J. H. Tenery, and R. P. Basson. Corelation between decrease in bacterial load and rate of wound healing. *Surg.Gynecol.Obstet.*, 130(4):616–620, 1970.

J. R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, 1993.

M. Robnik Šikonja. CORE - a system that predicts continuous variables. In *Proceedings of Electrotehnical and Computer Science Conference*, pages B145–148, 1997.

M. Robnik Šikonja and I. Kononenko. An adaptation of Relief for attribute estimation in regression. In Douglas H. Fisher, editor, *Machine Learning: Proceedings of the Fourteenth International Conference (ICML'97)*, pages 296–304. Morgan Kaufmann, 1997.

Marko Robnik-Šikonja and Igor Kononenko. Comprehensible interpretation of Relief's estimates. 2001. (submitted).

D. J. Shea. Pressure sores classification and management. *Clin.Orthop.Rel.Res.*, 112: 89–100, 1975.

A. I. Skene, J. M. Smith, C. J. Doré, A. Charlett, and J. D. Lewis. Venous leg ulcers: a prognostic index to predict time to healing. *BMJ*, 305:1119–1121, 1992.