

Prediction of Clinical Response to Excimer Laser Treatment in Vitiligo by Using Neural Network Models

Simone Cazzaniga^a Fabrizia Sassi^a Santo Raffaele Mercuri^b Luigi Naldi^{a, c}

^aCentro Studi GISED, Fondazione per la Ricerca Ospedale Maggiore (FROM), presidio ospedaliero Matteo Rota, Bergamo, ^bDepartment of Dermatology, Istituto di ricovero e cura a carattere scientifico San Raffaele, Milano, ^cDepartment of Dermatology, Ospedali Riuniti, Bergamo, Italy

Key Words

Vitiligo · Laser treatment, clinical response · Neural network

Abstract

Background: A predictive model may help to select likely responders and to anticipate treatment duration in vitiligo.

Methods: We aimed to develop a predictive rule based on data from a randomized trial of excimer laser in vitiligo. Information on 325 treated patches was available. The degree of repigmentation was assessed by digital image analysis of UVB-reflected photographs. Since no strong relationship between any single predictive parameter and outcome was initially documented, we relied on artificial neural networks.

Results: Using a time-response optimal threshold model, data were divided into 2 groups of responders and nonresponders. A discriminant network was trained in order to detect responders versus nonresponders. A regression network was subsequently used to compute repigmentation time in responders. The neural network discriminator achieved $66.46 \pm 5.37\%$ (95% CI) overall accuracy. The mean absolute error of the neural network regressor was 19.5843 ± 2.0930 with a root mean square error of 23.7156 ± 2.2225 .

Conclusion: Our study offers insight into the difficulty of clinical prediction in vitiligo and presents a way to develop an instrument with which to predict the clinical time response in patients treated by excimer laser.

Copyright © 2009 S. Karger AG, Basel

Introduction

Vitiligo is a pigmentary disorder of the skin that may be disfiguring and may have disabling social consequences, particularly in dark-skinned populations and for lesions in usually uncovered areas such as the face and the dorsa of the hands [1]. The natural course of the disease is highly variable but usually slowly progressive. Several treatment modalities have been suggested, but no single therapy is optimally suited to every patient [2–4].

The introduction of the xenon chloride excimer laser, which generates monochromatic light at 308 nm, allows for more selective phototherapy of vitiligo lesions than other ultraviolet-light (UV)-based treatments [5]. In a recently completed randomized clinical trial, up to 42% of patients who were previously unresponsive to conventional treatment modalities attained a 75% or greater reduction of the lesional areas of the face and neck when treated with a 12-week course of excimer laser therapy combined with topical corticosteroids [6]. Because the treatment is time consuming and highly demanding in terms of personal commitment, a predictive model for clinical response may have the obvious advantages of helping to select likely responders and anticipating treatment duration in individual patients. The aim of the present study was to develop a predictive rule based on a neural network model and on data from vitiligo patients enrolled in the above-mentioned randomized trial of excimer laser treatment.

Patients and Methods

Data Source

To develop the predictive model, we used data from a recently completed randomized clinical trial to estimate the efficacy of the 308-nm excimer laser, according to an incremental dose strategy, alone or in combination with hydrocortisone cream, for the treatment of vitiligo of the face and/or neck. Details regarding the trial design and main clinical results have been published elsewhere [6]. A total of 120 patients with different varieties of vitiligo (focal vs. generalized) were initially screened, and 84 underwent randomization. Data were collected regarding demographics, personal characteristics and habits, and relevant medical histories. These included age, gender, weight, height, eye color, hair color, skin pigmentation, alcohol consumption, smoking habits, family history of vitiligo in first-degree relatives, age at onset of vitiligo, age at first appearance of lesions on the face and/or neck, number of previous treatments with any kind of UVB-based therapy, clinical variety of vitiligo (focal or generalized), presence of autoimmune disease and history of other skin diseases (particularly psoriasis). Treatment consisted of phototherapy with a 308-nm excimer laser twice a week, alone or in combination with twice daily application of hydrocortisone 17-butyrate cream. The treatment course lasted 12 weeks. Of 84 patients initially randomized, 76 completed the treatment period; repigmentation data are fully available for a total of 325 vitiligo lesions in these patients. The degree of repigmentation was calculated separately for each treated lesion by comparing, through digital image analysis [7], standard UVB-reflected photographs taken at the beginning and at the end of the treatment period. Additional information concerning the 325 vitiligo lesions assessed included location (forehead, perioral region and chin area, periocular areas, cheeks, anterior neck and posterior neck) and size (categorized as <3, 3–15 and >15% of the total face and neck areas).

Statistical Model

Our analyses dealt with the time needed to reach a complete repigmentation and its relation with a number of predictive variables, including patient demographics and skin phenotype, features of vitiligo lesions (e.g. location, duration and size) and treatment modalities (i.e. laser treatment with or without application of a topical steroid). Only information about the degree of repigmentation reached at the end of a 12-week study period was available, and we assumed a linear relationship between the repigmentation rate and time afterwards. In order to obtain a single set of variables, we merged together the two originally randomized groups. We also considered individual lesions rather than whole-body patient responses. This was particularly useful in order to retain the lesion-intrinsic characteristics localization and size. When considering individual lesions, we replicated for each individual lesion the corresponding patient's clinical data. We knew this is a limitation of the study because data were not all independent and identically distributed but we decided to treat the problem afterwards adjusting final performance tests for patient membership. All statistical analyses were performed with Matlab (Mathworks Inc.). A preliminary analysis demonstrated that there was no strong relationship between any of the individual predictive parameters and outcome (maximum significant Spearman's $\rho = -0.27$ for skin pigmentation parameter), while a simple stepwise regression provided poor results (adjusted $R^2 = 0.17$). As

a consequence, we used an approach based on an artificial neural network models of which are able to automatically find nonlinear relationships upon multivariate analysis; their use in pattern recognition problems is well known in medicine [8]. Since the data distribution was widely dispersed (88.11 ± 95.33 , mean \pm SD), the first operation was to reduce this spreading. A k-means clustering [9] with 2 groups produced a consistent partition of data (silhouette mean 0.93 ± 0.1). Using a cutoff value of 100 treatment sessions, the repigmentation time was divided into 2 groups, distinguishing patients with no or poor responses from patients with more satisfactory responses. It was therefore necessary that we compiled 2 models, a discriminant network to distinguish responders from nonresponders and a regression network model to compute the time needed to achieve complete repigmentation in responders.

Bayesian Regularization

One of the main problems in model building is to optimize predictor performance without overfitting of training data. In neural networks, a valid method for this purpose is called regularization and involves the performance function used in the training stage. Normally the performance function is chosen to be the sum of squares of the network errors on the training set. It is possible to improve generalization adding a term that consists of the mean of the sum of squares of the network weights and biases. The modified performance function is a linear combination of the two terms above with unknown performance parameters α and β . The problem with regularization is that it is difficult to arbitrarily determine the optimum value for these parameters. One approach to automatically find α and β is the bayesian framework of MacKay [10, 11]. In this algorithm, the weights and biases of the network are assumed to be random variables with gaussian distributions. The regularization parameters are related to the unknown variances associated with these distributions and are estimated using statistical techniques. Bayesian regularization causes the network to adjust automatically the number of weights and biases effectively used, preventing overfitting and enforcing the generalization capability on new data.

Discriminant Network

The target was a dummy variable assuming a value of zero for negative response, and a value of 1 for positive response. The neural network we used was a feed-forward backpropagation network [12] with 6 neurons and a log-sigmoid transfer function on the hidden layer. Bayesian regularization with an epoch limit number of 40 was utilized for the training step that was performed on 325 cases. Tenfold cross-validation was used to test prediction accuracy on the fully trained classifier. In addition, to deal with nonindependent observations, we performed a patient-adjusted k-fold cross-validation (in each step we used data from 1 patient only as validation and built model with the remaining subsets).

Regression Network

The target was the time needed to achieve 100% repigmentation, assuming complete repigmentation to occur within 100 sessions. We used principal-component analysis to reduce the number of variables [13]. With this procedure, after a normalization step, we transformed and decorrelated the starting variables, retaining only those components that contributed more than 5% to total variation. The operation ensured that there were no redun-

Table 1. Correlations among parameters and response variables

Parameters	Sample values	Correlation and p value for discriminant network	Correlation and p value for regression network
Treatment	laser = 47.37% laser combination = 52.63%	0.2589 (≈ 0)	-0.1304 (0.0664) ✓
Gender	male = 47.37% female = 52.63%	0.0017 (0.976)	0.0804 (0.2592) ✓
Age	43.47 \pm 12.27 years	-0.1594 (0.004) ✓	0.1864 (0.0091) ✓
Weight	68.42 \pm 11.64 kg	0.0224 (0.6872)	-0.0495 (0.4876)
Height	170.83 \pm 7.44 cm	0.0376 (0.4989)	-0.1679 (0.0177) ✓
Eye color	light = 30.26% dark = 69.74%	-0.1252 (0.024) ✓	0.0371 (0.6030)
Hair color	light = 28.95% dark = 71.05%	-0.1434 (0.0096) ✓	-0.2108 (0.0028) ✓
Skin pigmentation	very pale = 15.79% pale = 47.37% quite dark = 31.58% dark = 5.26%	0.1564 (0.0047) ✓	-0.2188 (0.0019) ✓
Alcohol consumption	yes = 14.47% no = 85.53%	-0.1056 (0.0573) ✓	0.1018 (0.1525) ✓
Smoking habits	yes = 30.26% no = 69.74%	-0.0831 (0.1351) ✓	-0.0281 (0.6941)
Family history	yes = 44.74% no = 55.26%	0.0309 (0.5784)	-0.0913 (0.1998) ✓
Age at onset of vitiligo	23.08 \pm 13.07 years	-0.1232 (0.0263) ✓	0.1043 (0.1426) ✓
Age at first appearance of lesions on the face and/or neck	32.68 \pm 12.21 years	-0.0774 (0.1637)	0.0775 (0.2763) ✓
Number of previous treatments	<2 = 69.74% \geq 2 = 30.26%	-0.0485 (0.3831)	-0.0404 (0.5711)
Clinical variety of vitiligo	focal = 22.37% generalized = 77.63%	-0.1177 (0.034) ✓	0.111 (0.1186) ✓
Presence of autoimmune diseases	yes = 17.11% no = 82.89%	0.0119 (0.8310)	-0.1135 (0.1104) ✓
Presence of psoriasis	yes = 5.26% no = 94.74%	0.0755 (0.1745)	0.0501 (0.4818)
Location	forehead = 8% perioral region = 30% eyelids = 30.86% other face areas = 8% anterior neck = 20.86% posterior neck = 4.29%	0.1547 (0.0052) ✓ face/neck	0.0625 (0.3807) ✓ -0.0878 (0.2173) ✓ 0.1276 (0.0724) ✓ -0.1785 (0.0117) ✓ 0.0401 (0.5736) -0.1073 (0.1314) ✓
Lesion size	<3% = 49.43% 3% -15% = 34.84% \geq 15% = 15.71%	-0.0259 (0.6420)	0.0892 (0.2103) ✓

Spearman's ρ was used to determine correlation values; p values are given in parentheses. Parameters effectively used in each network are marked with a tick (✓).

Table 2. Discriminant network performance

	Model (n = 325)	10-fold cross- validation (95% CI)	Adjusted k-fold cross- validation (95% CI)
Accuracy, %	85.85	80.92 ± 3.01	66.46 ± 5.37
Sensitivity, %	88.94	85.44 ± 5.11	71.86 ± 6.53
Specificity, %	80.95	73.64 ± 5.94	57.94 ± 9.07
AUROC	0.9319	0.8426 ± 0.0379	0.6880 ± 0.1213

k = 76. AUROC = Area under the receiver-operating characteristic curve.

dant factors among retained variables, improving the generalization power of the network. The neural network we used was a feed-forward backpropagation network with 6 neurons and a tan-sigmoid transfer function on the hidden layer. Bayesian regularization with an epoch limit number of 40 was utilized for the training step that was performed using 199 cases. Tenfold cross-validation and patient-adjusted k-fold cross-validation were used to test prediction error on the fully trained regressor. Mean absolute error, root mean square error and accuracy with a cutoff of 12 were used as performance measures.

Results

The selected variables are presented in table 1. Those considered for inclusion in neural network models are marked with a tick (✓). As a preliminary step to address the discrimination problem, we removed the input variables with a low degree of relation with the target (Spearman's $\rho < 0.05$, $p > 0.15$) or with mutual redundancy with other variables related with the target (Pearson $r > 0.55$, $p < 0.05$). In order to increase significance, we also grouped location dummy variables into 2 categories (face or neck). Out of the initial 24 variables, 10 were taken into the model. The selected inputs and target were then passed to the network training routine, which automatically normalized data. Results of trained model and cross-validation test in the discriminant neural network are shown in table 2. The implemented neural network was able to achieve 85.85% overall accuracy, with 88.94% sensitivity and 80.95% specificity (AUROC = 0.93). The 10-fold cross-validation performed showed a little overestimation of results (accuracy = $80.92 \pm 3.01\%$, 95% CI) but the adjusted cross-validation showed a marked overestimation (accuracy = $66.46 \pm 5.37\%$). The unbiased estimate of prediction performance suggested an overfit of the model on the replicated data. Once a responder

group was identified, we moved to the definition of a predictive rule for repigmentation. Based on a preliminary selection, we retained 18 of the initial 24 variables (table 1). The selection was particularly accurate for dummy inputs but was also reliable for the other factors because we used only conservative selection criteria to remove variables (Spearman's $\rho < 0.05$, $p > 0.3$ for relation with response). We employed principal-component analysis to further reduce this number. The final number of parameters effectively passed to the network was 8. Results of the trained model and cross-validation test are summarized in table 3. The generated neural network regressor achieved an overall accuracy of 69.85% (tolerance = 12 sessions), with a mean absolute error of 9.5468 and root mean square error of 11.9158 on the whole dataset. Again the 10-fold cross-validation performed showed an overestimation of results (root mean square error = 19.1048 ± 2.2331 , 95% CI) that was confirmed by the adjusted cross-validation (root mean square error = 23.7156 ± 2.2225). The unbiased estimate of regressor performance showed that the model overfit the replicated data.

Discussion

Our purpose was to develop a predictive rule for the time needed to reach a given degree of repigmentation in vitiligo patients treated by excimer laser with an incremental dose strategy. A preliminary hypothesis of a linear relationship between time and repigmentation rate was necessary in order to develop a mathematical model using the data available from a recent randomized clinical trial. In addition the need to deal with lesion-intrinsic features, like localization and size, forced the replication of other variables. This limitation turned out to be an issue in the final tests when we adjusted performance measures for patient membership, even if we used a regularization technique in order to improve generalization capability. Results of tests (tables 2, 3) showed a marked difference of error estimates between model and cross-validation stage, for both discriminant and regression networks. Results suggested that the treatment response of multiple vitiligo lesions on the face and/or neck in the same patient is not independent. We acknowledge that the study has other limitations. First, we made the assumption that the repigmentation rate followed a linear trend. Beyond the 12-week duration of our randomized trial, such an assumption remains to be proven. Second, our choice of the variables to be assessed may be arbitrary.

Table 3. Regression network performance

	Model (n = 199)	10-fold cross-validation (95% CI)	Adjusted k-fold cross-validation (95% CI)
Mean absolute error	9.5468	15.1527 ± 1.8726	19.5843 ± 2.0930
Root mean square error	11.9158	19.1048 ± 2.2331	23.7156 ± 2.2225
Accuracy (cutoff = 12 sessions), %	69.85	50.26 ± 4.87	30.65 ± 6.53
R ²	0.6684	0.2143 ± 0.0697	0.1046 ± 0.0852

k = 59.

trary, as it is based on the variables collected in the context of the randomized trial. Finally, the training of neural networks was obtained from different random starting weights. However, there are multiple spatial solutions to the neural network problem, and we found more than one optimal result. There were similar problems at the selection stage. This was a critical step, and different subsets of optimal predictors could lead to different solutions [14, 15]. The availability of a larger database of vitiligo lesion features and responses to excimer laser treatment could help to find a more predictive model.

Vitiligo usually requires a long treatment course to obtain some degree of repigmentation [4]. The availability of a predictive rule for clinical response offers obvious advantages in the management of vitiligo patients. To the best of our knowledge, only a few studies are available that assess variables associated with clinical response in vitiligo [16, 17], and no formal development of a predictive rule has been attempted. Our study offers insight into the difficulty of clinical prediction in vitiligo and presents a way to develop an instrument with which to predict the number of treatment sessions required to obtain repigmentation in patients treated by excimer laser.

References

- Kent G, Al'Abadie M: Psychologic effects of vitiligo: a critical incident analysis. *J Am Acad Dermatol* 1996;35:895–898.
- Whitton ME, Ashcroft DM, Barrett CW, Gonzalez U: Interventions for vitiligo. *Cochrane Database Syst Rev* 2006;1:CD003263.
- Njoo MD, Westerhof W, Bos JD, Bossuyt PM: The development of guidelines for the treatment of vitiligo. *Arch Dermatol* 1999;135:1514–1521.
- Lim HW, Hexsel CL: Vitiligo: to treat or not to treat. *Arch Dermatol* 2007;143:643–646.
- Baltas E, Csoma Z, Ignacz F, et al: Treatment of vitiligo with the 308-nm xenon chloride excimer laser. *Arch Dermatol* 2002;138:1619–1620.
- Sassi F, Cazzaniga S, Tessari GP, Chatenoud L, et al: Randomised clinical trial comparing the effectiveness of 308 nm excimer laser alone or in combination with topical hydrocortisone 17-butyrate cream in the treatment of vitiligo of the face and neck. *Br J Dermatol* 2008;159:1186–1191.
- Gonzalez R, Woods R: *Digital Image Processing*, ed 2. Englewood Cliffs, Prentice Hall, 2002.
- Lisboa PJG, Vellido A, Wong H: Outstanding Issues for Clinical Decision Support with Neural Networks. London, Springer, 2002, pp 63–71.
- Kanungo T, Mount DM, Netanyahu NS, Piatko CD, Silverman R, Wu AY: An efficient k-means clustering algorithm: analysis and implementation. *PAMI* 2002;7:881–892.
- MacKay DJC: Bayesian interpolation. *Neural Comput* 1992;4:415–447.
- MacKay DJC: A practical bayesian framework for backpropagation networks. *Neural Comput* 1992;4:448–472.
- Ganesh A, Abdesselam B: A generalized feed-forward neural network architecture for classification and regression. *Neural Netw* 2003;16:561–568.
- Jolliffe IT: *Principal Component Analysis*, ed 2. New York, Springer, 2002.
- Lawless JF, Fredette M: Frequentist prediction intervals and predictive distributions. *Biometrika* 2005;92:529–542.
- Kwasnicka H, Michalak K: Correlation-based feature selection strategy in neural classification. *ISDA Proc* 2006;1:741–746.
- Nicolaidou E, Antoniou C, Alexander JS, Stefanaki C, Katsambas AD: Efficacy, predictors of response, and long-term follow-up in patients with vitiligo treated with narrow-band UVB phototherapy. *J Am Acad Dermatol* 2007;56:274–278.
- Dave S, Thappa DM, Dsouza M: Clinical predictors of outcome in vitiligo. *Indian J Dermatol Venereol Leprol* 2002;68:323–325.