ORIGINAL RESEARCH

A Priori Estimation of the Narrow-Band UVB Phototherapy Outcome for Moderate-to-Severe Psoriasis Based on the Patients' Questionnaire and Blood Tests Using Random Forest Classifier

Joanna Narbutt[1]
Janusz Krzyścin[2]
Piotr Sobolewski[2]
Małgorzata Skibińska[1]
Marcin Noweta[1]
Witold Owczarek [ID][3]
Bonawentura Rajewska-Więch[2]
Aleksandra Lesiak[1]

[1]Department of Dermatology, Pediatric Dermatology and Oncology, Medical University of Łódź, Łódź, Poland; [2]Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland; [3]Department of Dermatology, Military Institute of Medicine, Warsaw, Poland

Background: Nowadays, patients with moderate-to-severe psoriasis are treated with conventional immunosuppressants or with new biological agents. Phototherapy is the first-line treatment for patients in whom topical therapy is insufficient. Although numerous studies have been carried out, it is still difficult to predict the outcome of phototherapy in individual patients.

Methods: Prior to standard narrow band (NB) ultraviolet B (UVB) phototherapy, the patients filled out a questionnaire about personal life and health status. Several standard blood tests, including selected cytokine levels, were performed before and after a course of 20 NB-UVB treatments. The questionnaire answers, results of the blood tests, and treatment outcomes were analyzed using an artificial intelligence approach—the random forest (RF) classification tool.

Results: A total of 82 participants with moderate-to-severe psoriasis were enrolled. Prior to starting phototherapy, the patients with expected good outcome from the phototherapy, shorter remission, and quitting a possible second course of the NB-UVB treatment could be identified by the RF classifier with sensitivity over 84%, and accuracy of 75%, 85%, and 79%, respectively. The inclusion of cytokine data did not improve the performance of the RF classifier.

Conclusion: This approach offers help in making clinical decisions by identifying psoriatic patients in whom phototherapy will significantly improve their skin, or those in whom other therapies should be recommended beforehand.

Keywords: psoriasis, phototherapy, outcome prediction, artificial intelligence

Introduction

Psoriasis is a chronic, inflammatory, recurrent, and noninfectious skin disease affecting 2–4% of the global population. It can occur in both sexes at any age, and the development of the disease is stimulated by various genetic, immunological and environmental factors, which could be interrelated to some degree. Psoriasis is manifested by increased proliferation of the epidermis, which, along with inflammation, leads to skin lesions of varying severity.[1,2] Nowadays, patients with moderate-to-severe psoriasis are treated with conventional immunosuppressants or with new biological agents, with milder cases being most often treated topically. Phototherapy is the first-line treatment for patients for whom topical therapy is insufficient.[3]

Correspondence: Janusz Krzyścin
Institute of Geophysics, Polish Academy of Sciences, Warsaw, 01-452, Poland
Tel +48-2269-15950
Email jkrzys@igf.edu.pl

Various types of phototherapy have been developed. Among them broadband ultraviolet (UV) B light (UVB, range: 290–315 nm) was the first to be used. However, it was later replaced by more effective narrowband (NB)-UVB (311–313 nm) phototherapy. Nowadays, NB-UVB treatment is still extensively used for plaque psoriasis.[4,5] Moreover, NB-UVB is a highly economical and safe treatment that could avoid or delay starting more expensive third-line treatments for psoriasis.[6]

Although numerous studies have been carried out, it is difficult to predict the outcome of phototherapy in individual patients. There are numerous suggestions that machine learning models developed in recent years could estimate the outcome of treatment without predetermining medical pathways for psoriasis. Random forest (RF) methodology and the related free access RF software appear very promising when searching for complex heterogeneous and high dimensionless data when a link between input and output variables could not be revealed due to unrecognized relationships among the models' predictors and their complicated influences on output variables.[7,8] RF could deal with a large number of input parameters, even larger than the number of output samples, without the need for deleting some input variables to reduce dimensionality. RF models have been used to solve various biological/medical problems and have also found applications in dermatology.[9–12]

The objectives of this study are prediction of outcome of the phototherapy in terms of "good" (or "bad") response to NB-UVB phototherapy, short (or long) duration of remission, and identifying patients who will resign from further follow-up visits and possible treatments after completing the first phototherapy course. These objectives are important for making clinical decisions based on output of the RF classifier.

Patients and Methods
The Subjects

Eighty-two adult patients with plaque-type psoriasis, for whom the NB-UVB phototherapy was recommended, were recruited for the study. There were 42 men and 40 women with a mean age of 55 years and seven months (range: 24–76 years). Patients with a moderate-to-severe psoriasis, ie having Psoriasis Area and Severity Index (PASI) >10 and awaiting NB UVB therapy, were selected for the study. The patients voluntarily agreed to take part in additional blood tests. All the participants were advised not to apply active topical treatment during the phototherapy sessions, not to use sunbeds, and to avoid sunbathing. The study was conducted in accordance with the Declaration of Helsinki. Data acquisition and analysis were performed in compliance with protocols approved by the Ethical Committee at the Medical University of Łódź (MUL), ethical approval number RNN/219/18/KE. Written informed consent was obtained from all participants prior to the study.

Phototherapy Procedure

Outcomes of psoriatic patients treated with NB-UVB phototherapy at the Department of Dermatology MUL were examined. Fluorescent tubes (Phillips TL01) were used to expose patients in a medical cabinet. All patients underwent the same standard phototherapy procedure with 20 treatments.

The mean initial dose was ~0.15 Jul cm^{-2} (depending on the patient's phototype) at the start of phototherapy and was gradually increased by approximately 10% with each subsequent treatment to avoid the development of erythema depending on the patient's skin reaction. The maximum dose reached ~1.5 Jul cm^{-2}. NB-UVB phototherapy was administrated three times weekly, typically for six to eight weeks.

The MUL phototherapy cabinet needed to be calibrated due to the expected aging of the discharge tubes. It was done twice a year by the outer calibration campaigns. Independent occasional measurements by the handheld radiometer ILT 1400, Quantum Design, confirmed the accuracy of the calibration procedure.

Database

The database includes all psoriatic patients (n=82) attending NB-UVB phototherapy. It consists of three parts: the first—answers to a questionnaire, the second—results of medical tests carried out before and after each round of the cabinet exposures, the third—values of Psoriatic Severity Measures (PSM): PASI, body surface area (BSA), and Physician Global Assessment (PGA), obtained before and after each course of the exposures. Data can be shared with other users upon request.

The first part of the database (n=82) contains the answers from the questionnaire grouped into following categories:

- Continuous: age, weight, body mass index (BMI), the year of psoriasis diagnosis, the year of starting phototherapy, number of previous phototherapy treatments, alcohol units per week, cigarettes per week,

years of smoking, duration of time spend outdoors (in hours per week), the Dermatology Life Quality Index (DLQI)

- Binary: male/female—gender, Yes/No—feeling stressed, psoriasis in ancestors, psoriasis in descendants, psoriatic lesions present on head, on trunk, on upper and/or lower extremities, previous phototherapy treatments of psoriasis, hospitalization due to skin diseases, nonphototherapy treatment of psoriasis, psoriatic arthritis, other diseases, internet user, solarium user, frequent staying on the sun when weather permits, use of emollients.
- Categorical: material status (poor, middle, high), marital status (single, married, divorced), education level (primary school, secondary/high school, university level), place of residence (village, small city, large city (over 100,000 citizens), used medications (circulatory diseases, diabetes, depression, steroids, others).

The second part of database (n=71) contains numerical values obtained from various medical blood tests collected during the study. These were as follows:

- Glucose, total cholesterol, LDL, HDL, TG, AspAT, ALAT, CRP, RF, anti-CCP (taken at the beginning of the first phototherapy round).
- Vitamin D_3, cytokine level (IL-12, IL-17, IL-23, TNF-α, and INF-γ) taken at the beginning and end of the first phototherapy round using commercial ELISA kits.

The third part of the database contains PSM scores and their changes due to the therapy. The following values are stored:

- Baseline value (at the start of the exposures) of PASI, BSA, and PGA, and corresponding values at the end of the phototherapy round (n=71 for the first round, n=21 for the second round).
- Outcome of the first and second rounds of NB-UVB phototherapy in terms of relative PSM changes (as a percentage of the baseline value at the start).
- Duration of remission (n=27).

The Study Design
Comparative Statistics
The effectiveness of NB-UVB phototherapy is determined from changes of PASI, BSA, and PGA, after the first

(n=71) and second (n=21) course of the NB-UVB treatment. For all these metrics, the percentage change relative to the baseline value at the start of new course of the 20 exposures is calculated (denoted as ΔPASI, ΔBSA, and ΔPGA). The phototherapy outcome is compared after each course of the exposures and separately for two subgroups: present in both successive courses or only in the first course.

Standard statistical parameters (mean value, standard deviation, median, and the minimum-maximum range) are calculated to find out the metrics differences between two successive rounds of NB-UVB exposure. Moreover, the statistical significance of the differences between the outcomes are calculated using:

- The Wilcoxon rank sum test for comparison of the outcomes of the patients participating in both courses with those who missed the second possible round.
- The Wilcoxon signed-rank test for the comparison of the outcome after the first round with that after the second round for the same group of the patients.

Descriptive statistics and testing are from analytic software package, Statistica, TIBCO Software Inc. (https://www.tibco.com/products/data-science).

Random Forest Classifier
Answers to the questionnaire prior to the first course of phototherapy and results of various medical tests at the beginning and end of each course of NB-UVB treatments provide a set of potential predictors to be used by the RF classifier. PASI is the best rating scale for assessing the psoriasis severity and it is well correlated with other scales.[13] Thus, for the purpose of the RF classification of the NB-UVB phototherapy outcome, only the PASI scores are considered.

All ΔPASI values after the first exposure round and duration of remission were ordered (from the poorest to the best outcome). The ordered ΔPASI data (n=71) were divided into two groups below (n=25) and above (n=46) the threshold of ΔPASI=50%. For examination of the duration of remission, the threshold was set as the median value (eight weeks), so there were two sub-groups with short (below the median with the median inclusive) and long duration of remission (above the median). Some patients missed the second course of the NB-UVB exposures after completing the first course. Two sub-groups thus emerged, absent (n=44) and present (n=27) in the second course of the therapy. RF classifier is used to

predict in advance (before phototherapy) which patients belong to the above mentioned sub-groups.

The RF classification was optimized by examining different sets of the initial setup parameters. We decided to use the model with the greatest accuracy and with the smallest possible number of the explanatory variables. The classification of the psoriatic patients is carried out implementing RF source programs available on public web page: https://www.stat.berkeley.edu/~breiman/RandomForests/cc_software.htm

Results
Effectiveness of NB-UVB Phototherapy

Figure 1 and Table 1 show results of statistical analyses of relative changes (in % of values at the beginning of phototherapy) of PASI, BSA, and PGA, due to the phototherapy. Improvement in psoriasis could be seen in all considered indices. The mean values of the relative changes are: ~60% (PASI), ~40–50% (BSA and PGA).

The statistical tests confirm the hypothesis that there is no change of the statistical distribution for each PSM when comparing outputs of the first and the second rounds of the NB-UVB exposures (see Table 1 for all probability values >0.05).

The patients, who attended only the first course of phototherapy, had almost a similar outcome of phototherapy (for all psoriatic severity measures) when compared to patients continuing phototherapy after relapse. There was no significant difference in therapy outcomes after the first and second courses.
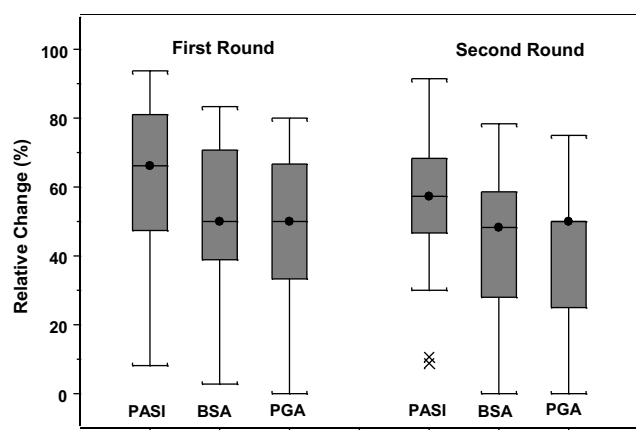

Figure 1 Box and whisker plot for the relative change in psoriatic severity measure (as a percentage of the value at the start of phototherapy), after the first and second rounds of 20 narrowband UVB exposures.
Abbreviations: BSA, body surface area; PASI, Psoriasis Area and Severity Index; PGA, Physician Global Assessment.

The mean duration of remission was about two months. The first patient returns for the next round of phototherapy after two weeks, but the last one after eight months. It is unlikely that duration of remission was longer than the project duration of three years.

RF Classification

Two versions of the RF model, ie with all possible and limited number of the predictors (excluding the cytokine data), are considered. The latter model is a low-cost option for the classification as only standard blood tests are examined. Its performance (Table 2) is shown for the following groups: the "bad" (ΔPASI ≤50) or "good" (ΔPASI >50) effectiveness of the therapy, "short" (≤8 weeks) or "long" duration (>8 weeks) of remission, and presence or absence in the second course of phototherapy. Statistics used to quantify the results of the RF model are in line with the 2×2 contingency table analysis standard, including: sensitivity, specificity, predictive values, likelihood ratios, and overall accuracy.[14]

RF classifier appears as a useful tool for identifying in advance the patients' sub-groups. Its sensitivity is over 84%, and accuracy is of 75% (for "good" and "bad" outcome), 85% ("short" and "long" remission), and 79% ("absence" and "presence" in the second course). Positive and negative likelihood ratios are significantly larger and lower than one, respectively. If a value of 1.0 is within the 95%CI, the prediction is no better than a coin toss. Moreover, for all cases, positive prediction value (PPV) is of ~80% and the 95%CI is over the 50% level of random forecast. This confirms the usefulness of the prediction of the "positive" effect by RF classifier. In only one case, ie prediction of a "bad" result of the phototherapy, the 95%CI of negative predictive value (NPV) contains 50%, ie the prediction can be like a coin toss. Positive effects are better predicted then negative ones in terms of smaller range of 95%CI and larger sensitivity (>84%) compared to specificity (~55–70%). Using a larger number of possible predictors (ie blood tests including cytokine data) at the start of the RF classification surprisingly did not improve its accuracy.

Discussion

Improvement in psoriasis due to NB-UVB phototherapy was found. However, the vast majority of the patients (~70%) attending the first course of therapy did not return

Table 1 Statistical Characteristics of the Phototherapy After the First and Second Courses of NB-UVB Phototherapy for Various Groups of Patients

Course/Patient Group	ΔPASI	ΔBSA	ΔPGA	n
Statistics				
After first/all patients	63±20 66 (8, 94)	51±21 50 (3, 83)	47±20 50 (0, 80)	71
After first/present in first and absent in second	65±20 66 (19, 94)	54±20 53 (5, 83)	49±22 50 (0, 80)	50
After first/present in first and present in second	60±21 63 (8, 89)	45±20 45 (3, 82)	42±17 50 (0, 67)	21
After second/present in first and present in second	56±23 57 (8, 91)	43±22 48 (0, 78)	40±22 50 (0, 75)	21
Probability				
After first/present in second vs absent in second	0.42	0.12	0.11	21/50
After second/present in first vs present in second	0.54	0.87	0.33	21/21

Notes: The mean value and ±SD for the relative changes of the psoriatic metrics for each round of the phototherapy. The corresponding values of the median and the min-max range (in parentheses). n, number of patients in each group. Probability of the similar statistical distribution of outcomes. ΔNormalized difference (end value minus start value as a percentage of the starting value).
Abbreviations: BSA, body surface area; PASI, Psoriasis Area and Severity Index; PGA, Physician Global Assessment.

to the department for further appointments. The comparison of the psoriasis severity scores in patients who participated in two successive phototherapy courses with those completing only the first course of treatment showed that their worse outcome of phototherapy was not a reason to quit follow-up appointments.

The RF classifier is able, ie both PPV and NPV are significantly >50%, to identify in advance the patients who will attend phototherapy courses or quit after the first course. The classification by RF can also effectively identify those who will have a short or long duration of relapse. The RF classification only fails (ie 95%CI of NPV contains 50%) to identify patients with a "bad" outcome of the phototherapy. However, in this case, the RF tool helps to select the patients to be closely supervised during phototherapy as they may need a different form of therapy.

RF classifier selects the best set of predictors (Table 3) from initial large set of potential predictors and provides the ranking of the predictors (from the best to the worst). The accuracy of RF classifier using a limited number of starting predictors (without cytokine data) is even better

Table 2 Statistical Characteristics of the RF Classification, without Cytokine Input, for Various Sub-Groups of the Psoriatic Patients

Statistic	"Positive": ΔPASI ≥50 "Negative": ΔPASI <50	"Positive": Remission ≤8 months "Negative": Remission >8 months	"Positive": Out of second Course "Negative": in second Course
	Value (95%CI)	Value (95%CI)	Value (95%CI)
Sensitivity	83.7 (70.3–92.7)	94.1 (71.3–99.9)	86.3 (72.7–94.8)
Specificity	54.6 (32.2–75.6)	70.0 (34.8–93.3)	66.7 (46.0–83.5)
Positive likelihood ratio	1.84 (1.15–2.96)	3.14 (1.21–8.15)	2.59 (1.50–4.47)
Negative likelihood ratio	0.30 (0.14–0.63)	0.08 (0.01–0.59)	0.20 (0.09–0.45)
Positive predictive value	80.4 (71.9–86.8)	84.2 (67.3–93.3)	80.9 (71.0–87.9)
Negative predictive value	60.0 (41.7–75.9)	87.5 (50.1–98.0)	75.0 (57.7–86.9)
Accuracy	74.7 (62.9–84.2)	85.2 (66.3–95.8)	78.9 (67.6–87.7)
Accuracy[a]	71.8 (59.9–81.9)	85.2 (66.3–95.8)	76.1 (64.5–85.4)

Note: [a]Accuracy of the classification with cytokine input, ΔNormalized difference (end value minus start value as a percentage of the starting value).
Abbreviation: PASI, Psoriasis Area and Severity Index.

Table 3 Ranking of Explanatory Variables from Most Important to Least Important

Output	Explanatory Variables
ΔPASI	Psoriasis in ancestors, number of previous cycles of exposures, LDL, glucose level, the education degree, use of emollients, psoriatic arthritis, difference between the year of the first psoriasis phototherapy and the year of psoriasis diagnosis, taking antidepressants, total cholesterol
Duration of remission	Material status, difference between the year of the first psoriasis phototherapy and the year of psoriasis diagnosis, number of earlier phototherapy treatments, psoriatic arthritis, taking diabetic drugs, glucose level, years of smoking, psoriatic arthritis, frequency of psoriasis exacerbations per year
Phototherapy continuation	BSA at the start of the first round, PGA at the start, number of cigarettes smoked per day, PASI at start, years of life, psoriasis in ancestors

Note: ΔNormalized difference (end value minus start value as a percentage of the starting value).
Abbreviations: BSA, body surface area; LDL, low-density lipoprotein; PASI, Psoriasis Area and Severity Index; PGA, Physician Global Assessment.

than the RF classifier with more potential predictors (with cytokine data) at the start. Thus, it is enough to use a few standard medical tests and a questionnaire for building a limited database to serve as RF input. However, the model response to the selected single predictor cannot be quantified. This requires the use of other models (eg logistic regression for modelling binary outcome).

The list of potential variables explaining outcome of the NB-UVB therapy is not fixed and could be extended and adapted to local circumstances and the patients' profile (by entering additional medical tests and questions to the questionnaire). Therefore, before any psoriatic therapy, the artificial intelligence tool will be able to identify patients for whom NB-UVB phototherapy should be the first choice. This method requires further validation in a larger and a more diverse cohort but seems to be very promising. Despite the small number of patients enrolled in this study, the RF classifier shows that many predictions are much better than a coin toss.

Conclusion

We present the method that helps to identify psoriatic patients in whom phototherapy will significantly improve their skin, or those in whom other therapies should be recommended beforehand. For patients who may avoid

subsequent follow-up visits, some preventive actions can be taken to change their attitude toward the treatment.

RF classification tool does not yet have practical applications in today's standard of the psoriasis treatment. The article presents preliminary results that individual patients' data (from low-cost blood tests and patients' questionnaire) can be transformed, using the RF tool, into valuable information-facilitating decision-making by doctors.

Abbreviations

BMI, body mass index; BSA, body surface area; DLQI, dermatology life quality index; NB-UVB, narrow band ultraviolet B; NPV, negative predictive value; PASI, Psoriasis Area and Severity Index; PGA, Physician Global Assessment; PPV, positive prediction value; PSM, psoriatic severity measure; RF, random forest.

Acknowledgments

Disclosure

The authors report no conflicts of interest in this work.

References

1. Nestle FO, Kaplan DH, Barker J. Psoriasis. *N Engl J Med*. 2009;361:496–509. doi:10.1056/NEJMra0804595
2. World Health Organization (WHO). Global report on psoriasis. 2016. Available from: https://apps.who.int/iris/handle/10665/204417. Accessed February 1, 2021.
3. Wan J, Abuabara K, Troxel AB, et al. Dermatologist preferences for first-line therapy of moderate to severe psoriasis in healthy adult patients. *J Am Acad Dermatol*. 2012;66(3):376–386. doi:10.1016/j.jaad.2011.03.012
4. Higgins E. Psoriasis. *Medicine*. 2017;45(6):368–378. doi:10.1016/j.mpmed.2017.03.010
5. Zhang P, Wu MX. A clinical review of phototherapy for psoriasis. *Lasers Med Sci*. 2018;33(1):173–180. doi:10.1007/s10103-017-2360-1
6. Boswell K, Cameron H, West J, et al. Narrowband ultraviolet B treatment for psoriasis is highly economical and causes significant savings in cost for topical treatments. *Br J Dermatol*. 2018;179(5):1148–1156. doi:10.1111/bjd.16716
7. Breiman L. Random forest. *Mach Learn*. 2001;45:5–32. doi:10.1023/A:1010933404324
8. Breiman L, Cutler A. Random forests. Available from: https://www.stat.berkeley.edu/users/breiman/RandomForests/. Accessed February 1, 2021.
9. Inza I, Calvo B, Armañanzas R, Bengoetxea E, Larrañaga P, Lozano JA. Machine learning: an indispensable tool in bioinformatics. *Methods Mol Biol*. 2010;593:25–48. doi:10.1007/978-1-60327-194-3_2

10. Lebedev AV, Westman E, Van Westen GJ, et al. Random forest ensembles for detection and prediction of alzheimer's disease with a good between-cohort robustness. *NeuroImage Clin*. 2014;6:115–125. doi:10.1016/j.nicl.2014.08.023

11. Denisko D, Hoffman MD. Classification and interaction in random forests. *Proc Natl Acad Sci U S A*. 2018;115(8):1690–1692. doi:10.1073/pnas.1800256115

12. Janney JB, Roslin S. Classification of melanoma from dermoscopic data using machine learning techniques. *Multimedia Tools Appl*. 2020;79:3713–3728. doi:10.1007/s11042-018-6927-z

13. Bachmann F. Evaluation of disease severity: clinical scores and questionnaires. In: Sterry W, Sabat R, Phillipp S, editors. *Psoriasis, Diagnosis and Management*. UK: Publishers John Wiley & Sons Ltd; 2014:129–139. doi:10.1002/9781118661796.ch14

14. Parshall MB. Unpacking the 2 x 2 table. *Heart Lung*. 2013;42 (3):221–226. doi:10.1016/j.hrtlng.2013.01.006