

Original Article

Construction of risk prediction models for psoriasis based on 45 dietary nutrients using machine learning and SHAP analysis

Xin Zhang MM[†], Zhe Gao MM[†], Jiang-Feng Feng MM, Xin-Gang Wu MM

Department of Dermatology, Hangzhou Third People's Hospital, Hangzhou, China

[†]Both authors contributed equally to this manuscript

Background and Objectives: To develop and validate a machine learning model to predict the risk of psoriasis based on 45 dietary nutrients. **Methods and Study Design:** 12,749 participants from the National Health and Nutrition Examination Survey from 2009-2014 were included and their demographic, lifestyle, health status, and dietary nutrient-related information is collected. Imbalanced data were processed using the synthetic minority over-sampling technique (SMOTE). After removing the covariate features, important features were further screened using the Boruta algorithm and six machine learning models were constructed including Random Forest (RF), Light Gradient Boosting Machine (Light GBM), Kernel K-Nearest Neighbor (K-KNN), L Naive Bayes, Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost). The performance of the models was evaluated using benchmarking and the area under the ROC curve (AUC) was the main evaluation metric to choose the optimal machine learning model. Shapley additive explanation (SHAP) values were computed to evaluate each feature's prediction role in the model. **Results:** The Boruta algorithm screened 10 baseline features and 23 dietary nutrient features, and six machine learning models were developed based on them. Compared with other machine learning models, XGBoost demonstrated superior prediction ability. SHAP analysis showed that theobromine, lycopene, caffeine, vitamin D, dietary fiber and vitamin E were the key features that influenced the prediction results. **Conclusions:** The machine learning (ML) algorithm constructed a prediction model for psoriasis by incorporating baseline features and dietary nutrient features. The SHAP values indicate the dominant role of dietary nutrients in the model.

Key Words: psoriasis, dietary nutrient, machine learning, XGBoost, predictive model

INTRODUCTION

Psoriasis is a common chronic inflammatory skin disease characterized primarily by hyperproliferation and parakeratosis of keratinocytes. Statistically, psoriasis affects approximately 2-3% of the U.S. population (about 7.5 million people),^{1,2} and costs over \$110 billion annually,³ posing a significant economic challenge to society. Psoriasis not only causes psychological and physical burdens to patients, but also leads to disfigurement, disability and other related comorbidities such as psoriatic arthritis if not controlled.² The pathogenesis of psoriasis is complex, involving immunoinflammatory, genetic and environmental factors (including diet).⁴ Elevated inflammatory factors put patients with psoriasis at significantly increased risk for other chronic serious diseases, including metabolic syndrome, diabetes mellitus, cardiovascular disease, Crohn's disease, and lymphoma.⁵⁻⁸ Conversely, these diseases exacerbate the systemic and cutaneous inflammatory response in psoriasis.^{9,10} Controlling inflammation has important implications for the prevention and treatment of psoriasis.

Although the inflammatory response caused by immune imbalance plays a dominant role in the pathogenesis of psoriasis, diet is also an important source

of inflammation in the body. An increasing amount of evidence points to a potential close connection between different dietary structures and the pathophysiology and etiology of psoriasis.^{9,11} A high-fat diet (HFD) and excessive saturated fatty acid intake may lead to an elevated risk of psoriasis exacerbation,¹² whereas fasting, low-energy diets, vegetarian diets, and the consumption of n-3 polyunsaturated fatty acid-rich foods attenuate skin damage and reduce the risk of comorbidities, improving the prognosis of psoriasis patients.^{9,13} In addition, some researchers have suggested that a gluten-free diet has a positive impact on the course of psoriasis,¹⁴ although only in patients with gluten sensitivity or serological results of elevated IgA/IgG gliadin antibodies,¹⁵ and the Medical

Corresponding Author: Xin-Gang Wu, Department of Dermatology, Hangzhou Third People's Hospital, Hangzhou, China, 310000

Tel: +18958013054

Email: wuxingang1986@126.com

Electronic supplementary information available. See apjcn.qdu.edu.cn/35_5_800_supp.pdf

Manuscript received 12 March 2026. Initial review completed 11 May 2026. Revision accepted 07 July 2026.

doi: 10.6133/apjcn.202610_35(5).0006

Board of the National Psoriasis Foundation does not strongly recommend a gluten-free diet.¹⁶ Previous studies have been limited to individual or specific nutrients to explore the potential impact of diet on psoriasis and have not taken into account the interactions between various dietary components. Therefore, it is crucial to investigate the association between dietary nutrients and psoriasis in terms of overall dietary patterns.

In recent years, the construction of risk prediction models for diseases based on ML algorithms has significant advantages, bringing breakthroughs to medical practice.¹⁷ ML algorithms are able to deal with high-dimensional data, mine the data for complex patterns and hidden non-linear relationships, increase prediction accuracy, enable individualized assessment, and optimize clinical decision-making.^{18,19} SHAP analysis is a powerful technique to improve the interpretability of ML models. It is based on the classical concepts of game theory (Shapley values) and provides a consistent and interpretable assignment of each feature's contribution to the model's prediction, thus explaining the internal decision logic of ML models.^{20,21}

Therefore, this study intends to use adult data from the National Health and Nutrition Examination Survey (NHANES) to identify potential associations between dietary nutrient intake and psoriasis risk using ML algorithms. It also aims to elucidate the precise role of each dietary nutrient in the model and its clinical significance through SHAP values, as well as to develop effective dietary interventions.

METHODS

Study population

This study used data from the 2009-2014 NHANES, which collected demographic, socioeconomic, and dietary-related data. The study population for this paper consisted primarily of adults (≥ 20 years old) from the United States from 2009 through 2014. Populations with missing important variables were excluded. The detailed inclusion and exclusion criteria are shown in Supplementary Figure 1. Ultimately, 12,749 subjects were included.

Data sources for dietary nutrients

Forty-five foods or nutrients were obtained from NHANES dietary recall data, including added vitamin B-12, alcohol, alpha-carotene, alpha-tocopherol, beta-carotene, beta-cryptoxanthin, caffeine, calcium, carbohydrate, cholesterol, copper, dietary fiber, folate, folic acid, food folate, iron, lutein+zeaxanthin, lycopene, magnesium, moisture, monounsaturated fatty acids, niacin, phosphorus, polyunsaturated fatty acids, potassium, protein, retinol, saturated fatty acids, selenium, sodium, theobromine, total choline, total fat, total folate, total sugar, Vitamin A/B1/B12/B2/B6 /C/D/E/K, Zinc. The NHANES database provided the subjects with two 24-h dietary recall interviews and the average daily intake of dietary nutrients was calculated using the information above.

Diagnosis of psoriasis

A questionnaire was used to collect the prevalence of psoriasis among participants. Participants were categorized as having psoriasis if they answered positively to the question, "Have you ever been told by a doctor or

other health care professional that you had psoriasis?". This definition is consistent with previous studies.^{22,23}

Variables of the study and their pretreatment

The study initially included 10 baseline characteristics related to demographics, lifestyle, and health status and 46 characteristics related to dietary nutrients (including energy). The baseline characteristics mainly included Age, Sex, Race (Mexican, Other Hispanic, Non-Hispanic White, Non-Hispanic Black, Other Race), Education (< 9 th grade, 9-11th grade, High School, Some College, College graduate), Family income to poverty ratio, body mass index (BMI), drinking status, smoking status, hypertension, diabetes. Hypertension is defined as "Ever told you had high blood pressure" and "Now taking prescribed medicine for HBP".²⁴ Diabetes was defined as "Doctor told you had diabetes", "Two Hour Glucose (OGTT) (mmol/L) ≥ 11.1 ", and "Fasting Glucose (mmol/L) ≥ 7.0 ".²⁵ Smoking status was categorized as non-current smokers (never smoked or quit smoking for more than 1 year) or current smokers (current smokers or smokers who smoked more than 1 day in the past 30 days or smoked more than 2 cigarettes per day after waking up or quitting smoking).²⁶ Drinking status was classified as either never drinking (less than 12 drinks in lifetime) or drinking (at least 12 drinks per year or more than 6 drinks in the past 12 months).²⁷

Furthermore, in order to reduce multicollinearity between features, variance inflation factor (VIF) was used to test for collinearity and features with a VIF greater than 3 were eliminated.

Boruta algorithm

The Boruta algorithm (R package version 8.0.0) is a wrapper method for determining important features around the random forest classification algorithm.^{28,29} The Boruta algorithm first uses the random forest model to calculate the importance of each feature and creates the shadow features. The importance of the original feature is compared to the shadow feature, and if the importance of the original feature is significantly higher than the maximum importance of the shadow feature, then it is considered "important", and vice versa for "irrelevant". For those whose importance cannot be determined, the algorithm marks them as "TBD" (to be determined). The process is repeated 100 times until all important variables are selected.

Construction of machine learning models

A total of six machine learning methods were constructed, namely RF, Light GBM, K-KNN, L-Naive Bayes, SVM, and XGBoost. RF is an ensemble learning algorithm that generates multiple decision trees and classifies or regresses them to improve the accuracy and performance of the model, which effectively prevents overfitting.³⁰ Light GBM is based on gradient boosting, which improves each new tree based on the error of the previous one and gradually improves the accuracy of the model, and it is more speedy when dealing with large-scale data.³¹ XGBoost is similar to Light GBM and its core is also based on the gradient boosting decision tree (GBDT) to construct a prediction model. Unlike Light GBM, the

former adopts a level-wise tree growth strategy to make the tree structure relatively balanced, while the latter adopts a leaf-wise tree growth strategy to make the tree depth deeper and the model fit stronger.³² SVM is a feature selection method that combines supervised learning and iterative optimization to filter the optimal features through weight ranking and iterative culling.³³ K-KNN is a classification method for screening similar samples, which is relatively simple to compute.³⁴ L-Naive Bayes is a probabilistic statistical classification algorithm that predicts the class of a sample by calculating the probability of the sample containing a certain feature under each class.³⁵

SHAP interpretability analysis

SHAP analysis is a method for interpreting the prediction results of machine learning models that quantifies the importance of features by calculating the marginal contribution of each feature to the model prediction and weighting the average of the marginal contributions to obtain the SHAP value of each feature. SHAP analysis can interpret the degree of each feature's contribution to each individual predicted probability and identify the key features that have the greatest impact on the model prediction results, making a global interpretation.³⁶

Statistical analysis

The R package (version 4.4.1), Matlab (R2021 a), and Python (3.7.3) were used for all data analyses. Survey-weighted statistical analyses were first used to characterize participants, with continuous variables expressed as means and standard deviations (SD) and categorical variables expressed as frequencies and percentages. Independent t-tests were used for continuous variables and chi-square tests were used for categorical variables. The SMOTE algorithm implemented in Matlab software was used to solve the sample imbalance problem, and six machine learning models were constructed, with a 6:4 ratio dividing the training and validation sets, and 10 internal cross-validations to prevent overfitting. The models are then benchmarked on metrics including accuracy, F-beta, AUC, sensitivity (recall), specificity, and precision-recall area under the curve (PR-AUC). AUC is the primary assessment parameter utilized to choose the best model, and the SHAP algorithm was used to interpret these results. The ANOVA test and the Kruskal-Wallis test were used to test the differences in the metrics between the different models. To evaluate the robustness of the findings, a sensitivity analysis was additionally conducted using the original imbalanced dataset without SMOTE resampling. The same feature selection procedures and machine learning algorithms were applied, and the resulting model performance and feature importance rankings were compared with those obtained from the SMOTE-balanced dataset. All statistical tests were two-sided and $p < 0.05$ was considered statistically significant.

RESULTS

Patient characteristics

A total of 12,749 participants were included in this research, of whom 370 (2.9%) were diagnosed with psoriasis. Compared to participants without psoriasis, those

with psoriasis were typically older (52.03 ± 16.63 vs. 48.50 ± 17.67 , $p < 0.001$), predominantly Non-Hispanic White ($p < 0.001$), had a higher BMI (30.10 ± 6.93 vs. 29.13 ± 6.99 , $p = 0.003$), higher blood pressure (47.30% vs. 35.54% , $p < 0.001$), higher likelihood of comorbid diabetes (17.03% vs. 12.51% , $p = 0.016$), and lower intake of Vitamin B12 (0.87 ± 2.54 vs. 0.97 ± 2.62 , $p = 0.013$), higher intakes of Selenium (184.77 ± 227.19 vs. 152.78 ± 202.70 , $p = 0.002$) and Caffeine (47.88 ± 100.97 vs. 35.68 ± 77.68 , $p = 0.020$). See Table 1 for details.

Screening of predictors

Due to the substantial correlation between some of the variables, correlation coefficients between dietary nutrient characteristics were first calculated using VIF (Figure 1A). Twenty-three variables with high correlation, such as alcohol and alpha-carotene, were eliminated when the VIF was more than three, indicating that the variables were multicollinear. The importance of the remaining 33 features was then assessed using the Boruta algorithm to identify the most crucial features associated with the risk of developing psoriasis. Eventually, 33 key features such as Lutein, zeaxanthin, Hypertension, Lycopene, Sodium, Theobromine were incorporated, including 10 baseline features and 23 dietary nutritional features (Figure 1B).

Model development and validation

As the number of control group was substantially greater than that of psoriasis group, we employed the Smote algorithm to address the imbalanced dataset. Ultimately, 12,379 non-psoriasis patients and 8,665 psoriasis patients were included in this paper (Supplementary Figure 2) and split into training and validation sets in a 6:4 ratio. Six different techniques of ML were used to build predictive models based on the 33 features that were screened above and the predictive performance of these models was evaluated in terms of accuracy, F-beta, AUC, sensitivity, specificity, and PR-AUC. The XGBoost model demonstrates excellent predictive performance in both the training and validation sets, with the largest area under the ROC curve and maximum recall (Figure 2), indicating that the model has a high diagnostic value and is able to identify more patients who actually have psoriasis. The performance of the six ML models in the training and validation sets is detailed in Tables 2 and 3, respectively. In the training set (Table 2), the AUC values of the six models, which use AUC as the primary evaluation criterion, are 0.982 for RF, 0.973 for Light GBM, 0.963 for K-KNN, 0.602 for Naive Bayes, 0.845 for SVM, and 0.991 for XGBoost. In the validation set (Table 3), the corresponding AUC values of the above models are 0.992, 0.982, 0.966, 0.594, 0.868, and 0.997, respectively. In addition, the XGBoost model exhibits higher accuracy, F-beta value, specificity, and PR-AUC. Following a thorough assessment, XGBoost was selected as the best ML model.

Table 1. Baseline characteristics of the participants

Characteristic	N	Overall N = 12,749	No psoriasis N = 12,379	Psoriasis N = 370	<i>p</i>
Age (year) [†] , Mean ± SD	12,749	48.6 ± 17.7	48.5 ± 17.7	52.0 ± 16.6	<0.001
Sex [‡] , n (%)	12,749				0.973
Female		6,351 (49.8%)	6,167 (49.8%)	184 (49.7%)	
Male		6,398 (50.2%)	6,212 (50.2%)	186 (50.3%)	
Race/ethnicity [‡] , n (%)	12,749				<0.001
Mexican		1,701 (13.3%)	1,670 (13.5%)	31 (8.38%)	
Other Hispanic		1,150 (9.02%)	1,114 (9.00%)	36 (9.73%)	
Non-Hispanic White		5,927 (46.5%)	5,703 (46.1%)	224 (60.5%)	
Non-Hispanic Black		2,632 (20.6%)	2,590 (20.9%)	42 (11.4%)	
Other Race		1,339 (10.5%)	1,302 (10.5%)	37 (10.0%)	
Education [‡] , n (%)	12,749				0.589
Less Than 9th		1,052 (8.3%)	1,028 (8.30%)	24 (6.49%)	
9-11th		1,748 (13.7%)	1,702 (13.8%)	46 (12.4%)	
High School		2,846 (22.3%)	2,755 (22.3%)	91 (24.6%)	
Some College		3,936 (30.9%)	3,820 (30.9%)	116 (31.4%)	
College Graduate		3,167 (24.8%)	3,074 (24.8%)	93 (25.1%)	
Family income to poverty ratio [†] , Mean ± SD	12,749	2.51 ± 1.65	2.51 ± 1.65	2.52 ± 1.70	0.903
BMI [†] , Mean ± SD	12,749	29.2 ± 6.99	29.1 ± 6.99	30.1 ± 6.93	0.003
Hypertension [‡] , n (%)	12,749				<0.001
No		8,175 (64.1%)	7,980 (64.5%)	195 (52.7%)	
Yes		4,574 (35.9%)	4,399 (35.5%)	175 (47.3%)	
Diabetes [‡] , n (%)	12,749				0.016
No		10,237 (80.3%)	9,961 (80.5%)	276 (74.6%)	
Yes		1,612 (12.6%)	1,549 (12.5%)	63 (17.0%)	
Borderline		900 (7.06%)	869 (7.02%)	31 (8.38%)	
Smoking status [‡] , n (%)	12,749				0.239
No		10,065 (79.0%)	9,782 (79.0%)	283 (76.5%)	
Yes		2,684 (21.1%)	2,597 (21.0%)	87 (23.5%)	
Drinking status [‡] , n (%)	12,749				0.386
No		1,745 (13.7%)	1,700 (13.7%)	45 (12.2%)	
Yes		11,004 (86.3%)	10,679 (86.3%)	325 (87.8%)	
Energy [†] , Mean ± SD	12,749	2,134 ± 998	2,134 ± 999	2,141 ± 948	0.552
Protein [†] , Mean ± SD	12,749	82.3 ± 43.2	82.4 ± 43.4	80.5 ± 35.7	0.527
Carbohydrate [†] , Mean ± SD	12,749	259 ± 127	259 ± 127	257 ± 126	0.894
Total Sugar [†] , Mean ± SD	12,749	115 ± 79.3	115 ± 79.0	118 ± 87.0	0.862
Dietary fiber [†] , Mean ± SD	12,749	17.0 ± 10.6	17.0 ± 10.6	17.0 ± 10.6	0.874
Total Fat [†] , Mean ± SD	12,749	79.9 ± 46.2	79.8 ± 46.3	82.9 ± 44.6	0.085
Saturated fatty acids [†] , Mean ± SD	12,749	25.7 ± 16.4	25.7 ± 16.4	26.7 ± 16.6	0.163

[†]Student t-test; [‡]Chi-square test; SD: standard deviation.

Table 1. Baseline characteristics of the participants

Characteristic	N	Overall N = 12,749	No psoriasis N = 12,379	Psoriasis N = 370	<i>p</i>
Monounsaturated fatty acids [†] , Mean ± SD	12,749	28.5 ± 17.5	28.5 ± 17.5	29.5 ± 16.7	0.090
Polyunsaturated fatty acids [†] , Mean ± SD	12,749	18.5 ± 12.5	18.5 ± 12.5	19.3 ± 11.9	0.058
Cholesterol [†] , Mean ± SD	12,749	291.5 ± 238	292 ± 239	273 ± 213	0.226
Vitamin E as alpha-tocopherol [†] , Mean ± SD	12,749	8.36 ± 6.88	8.35 ± 6.89	8.77 ± 6.37	0.065
Alpha-tocopherol [†] , Mean ± SD	12,749	0.62 ± 3.16	0.63 ± 3.19	0.47 ± 2.01	0.474
Retinol [†] , Mean ± SD	12,749	411 ± 527	410 ± 528	426 ± 492	0.991
Vitamin A [†] , Mean ± SD	12,749	613 ± 682	613 ± 684	625 ± 594	0.782
Alpha-carotene [†] , Mean ± SD	12,749	397 ± 1,267	397 ± 1,274	412 ± 1,027	0.391
Beta-carotene [†] , Mean ± SD	12,749	2,194 ± 4,604	2,195 ± 4,627	2,151 ± 3,739	0.250
Beta-cryptoxanthin [†] , Mean ± SD	12,749	88.8 ± 322	89.2 ± 325	75.7 ± 154	0.076
Lycopene [†] , Mean ± SD	12,749	4,998 ± 9,118	4,978 ± 9,049	5,654 ± 11,185	0.934
Lutein+zeaxanthin [†] , Mean ± SD	12,749	1,574 ± 3,526	1,571 ± 3,517	1,703 ± 3,804	0.913
Thiamin (Vitamin B1) [†] , Mean ± SD	12,749	1.60 ± 0.90	1.60 ± 0.90	1.61 ± 0.93	0.856
Riboflavin (Vitamin B2) [†] , Mean ± SD	12,749	2.07 ± 1.29	2.07 ± 1.29	2.13 ± 1.33	0.417
Niacin [†] , Mean ± SD	12,749	25.7 ± 15.6	25.7 ± 15.6	24.6 ± 14.4	0.238
Vitamin B6 [†] , Mean ± SD	12,749	2.11 ± 1.65	2.11 ± 1.66	2.02 ± 1.45	0.268
Total folate [†] , Mean ± SD	12,749	181 ± 177	181 ± 176	186 ± 224	0.426
Folic acid [†] , Mean ± SD	12,749	222 ± 147	222 ± 147	230 ± 151	0.451
Food folate [†] , Mean ± SD	12,749	530 ± 359	529 ± 356	546 ± 437	0.961
Folate (DFE) [†] , Mean ± SD	12,749	339 ± 206	339 ± 207	326 ± 176	0.608
Total choline [†] , Mean ± SD	12,749	5.09 ± 6.25	5.10 ± 6.30	4.90 ± 4.40	0.568
Vitamin B12 [†] , Mean ± SD	12,749	0.97 ± 2.61	0.97 ± 2.62	0.87 ± 2.54	0.013
Added vitamin B12 [†] , Mean ± SD	12,749	83.4 ± 98.7	83.4 ± 98.0	84.6 ± 122	0.593
Vitamin C [†] , Mean ± SD	12,749	112 ± 216	112 ± 217	115 ± 164	0.466
Vitamin K [†] , Mean ± SD	12,749	948 ± 600	946 ± 599	997 ± 642	0.145
Calcium [†] , Mean ± SD	12,749	1,377 ± 693	1,377 ± 695	1,390 ± 647	0.284
Phosphorus [†] , Mean ± SD	12,749	300 ± 153	300 ± 153	306 ± 149	0.233
Magnesium [†] , Mean ± SD	12,749	14.8 ± 8.62	14.8 ± 8.63	14.6 ± 8.27	0.778
Iron [†] , Mean ± SD	12,749	11.3 ± 7.06	11.3 ± 7.08	11.3 ± 6.49	0.643
Zinc [†] , Mean ± SD	12,749	1.27 ± 1.02	1.27 ± 1.03	1.26 ± 0.65	0.392
Copper [†] , Mean ± SD	12,749	3,528 ± 1,819	3,526 ± 1,820	3,587 ± 1,785	0.433
Sodium [†] , Mean ± SD	12,749	2,663 ± 1,280	2,660 ± 1,281	2,736 ± 1,263	0.132
Potassium [†] , Mean ± SD	12,749	115 ± 65.7	115 ± 66.0	110 ± 54.5	0.855
Selenium [†] , Mean ± SD	12,749	154 ± 204	153 ± 203	185 ± 227	0.002
Caffeine [†] , Mean ± SD	12,749	36.0 ± 78.5	35.7 ± 77.7	47.9 ± 101	0.020
Theobromine [†] , Mean ± SD	12,749	11.1 ± 29.9	11.2 ± 29.9	10.4 ± 30.5	0.063
Alcohol [†] , Mean ± SD	12,749	2,936 ± 1,516	2,933 ± 1,518	3,028 ± 1,471	0.084
Moisture [†] , Mean ± SD	12,749	4.73 ± 5.68	4.72 ± 5.66	4.99 ± 6.08	0.910
Vitamin D [†] , Mean ± SD	12,749	2,134 ± 998	2,134 ± 999	2,141 ± 948	0.552

[†]Student t-test; [‡]Chi-square test; SD: standard deviation.

Table 2. Six ML models' predictive performance for psoriasis in training set

Model	Accuracy	F Beta	Area under the ROC curve	Sensitivity	Specificity	Area under the PR curve
Random Forest	0.937	0.949	0.982	0.973	0.885	0.986
Light GBM	0.918	0.932	0.973	0.937	0.889	0.979
K-KNN	0.850	0.857	0.963	0.751	0.997	0.981
Naive Bayes	0.489	0.387	0.602	0.270	0.814	0.676
SVM	0.760	0.803	0.845	0.821	0.670	0.890
XGBoost	0.957	0.964	0.991	0.958	0.955	0.992
<i>p</i>	<.001 ^a	<.001 ^a	<.001 ^b	<.001 ^a	<.001 ^a	<.001 ^a

[†]ANOVA test; [‡]Kruskal-Wallis.

Table 3. Six ML models' predictive performance for psoriasis in validation set

Model	Accuracy	F Beta	Area under the ROC curve	Sensitivity	Specificity	Area under the PR curve
Random Forest	0.958	0.965	0.992	0.977	0.932	0.994
Light GBM	0.933	0.943	0.982	0.948	0.913	0.986
K-KNN	0.868	0.872	0.966	0.773	1.000	0.983
Naive Bayes	0.483	0.336	0.594	0.225	0.843	0.651
SVM	0.787	0.818	0.868	0.819	0.743	0.898
XGBoost	0.976	0.979	0.997	0.972	0.981	0.998
<i>p</i>	<.001 ^a	<.001 ^a	<.001 ^b	<.001 ^a	<.001 ^a	<.001 ^a

[†]ANOVA test; [‡]Kruskal-Wallis.

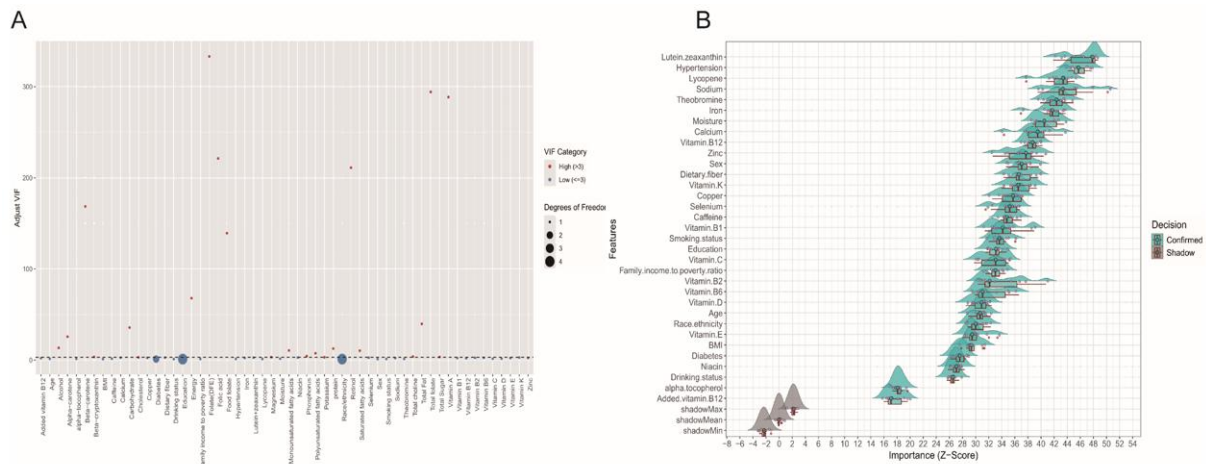


Figure 1. Results of important predictor screening. (A) VIF was used to check for multicollinearity among variables. When VIF >3, multicollinearity was present. When VIF ≤3, multicollinearity was low. (B) The Boruta algorithm was used to screen for important variables

Screening important dietary features based on SHAP values

SHAP analysis was used to identify the key features that had the greatest impact on the model's prediction results. The bee swarm plot (Figure 3A) shows the features that, taken as a whole, have the greatest impact on the model's anticipated outcomes, along with the effects of these features on the model's output. The vertical axis on the left side of the image visualizes the ranking of the importance of the features, from highest to lowest the top 15 features are theobromine (0.0500), lycopene (0.0485), caffeine (0.0463), vitamin D (0.0428), dietary fiber (0.0354), vitamin E (0.0353), vitamin B2 (0.0344), moisture (0.0329), copper (0.0323), calcium (0.0323), niacin (0.0322), vitamin B1 (0.0316), lutein+zeaxanthin (0.0309) and vitamin B6 (0.0290). All above characteristics had a negative impact on model output. The distinct effects of the different dietary nutrient features on the model outputs in the psoriasis

versus non-psoriasis groups are displayed in the SHAP force diagrams and waterfall plots (Figure 3B-C). The waterfall diagram demonstrates the features' incremental contribution to the model's final anticipated values, whereas the force diagram concentrates on breaking down the effects of the features on the model output. The final results suggested that the dietary nutrients described above were associated with lower model-predicted psoriasis probability and contributed negatively to the model output. In addition, we plotted a SHAP dependency plot (Figure S3) showing the nonlinear relationship between dietary components and SHAP values. The results revealed that higher levels of theobromine, lycopene, caffeine, vitamin D, dietary fiber, and vitamin E within specific ranges were associated with lower SHAP values and lower model-predicted psoriasis probability.

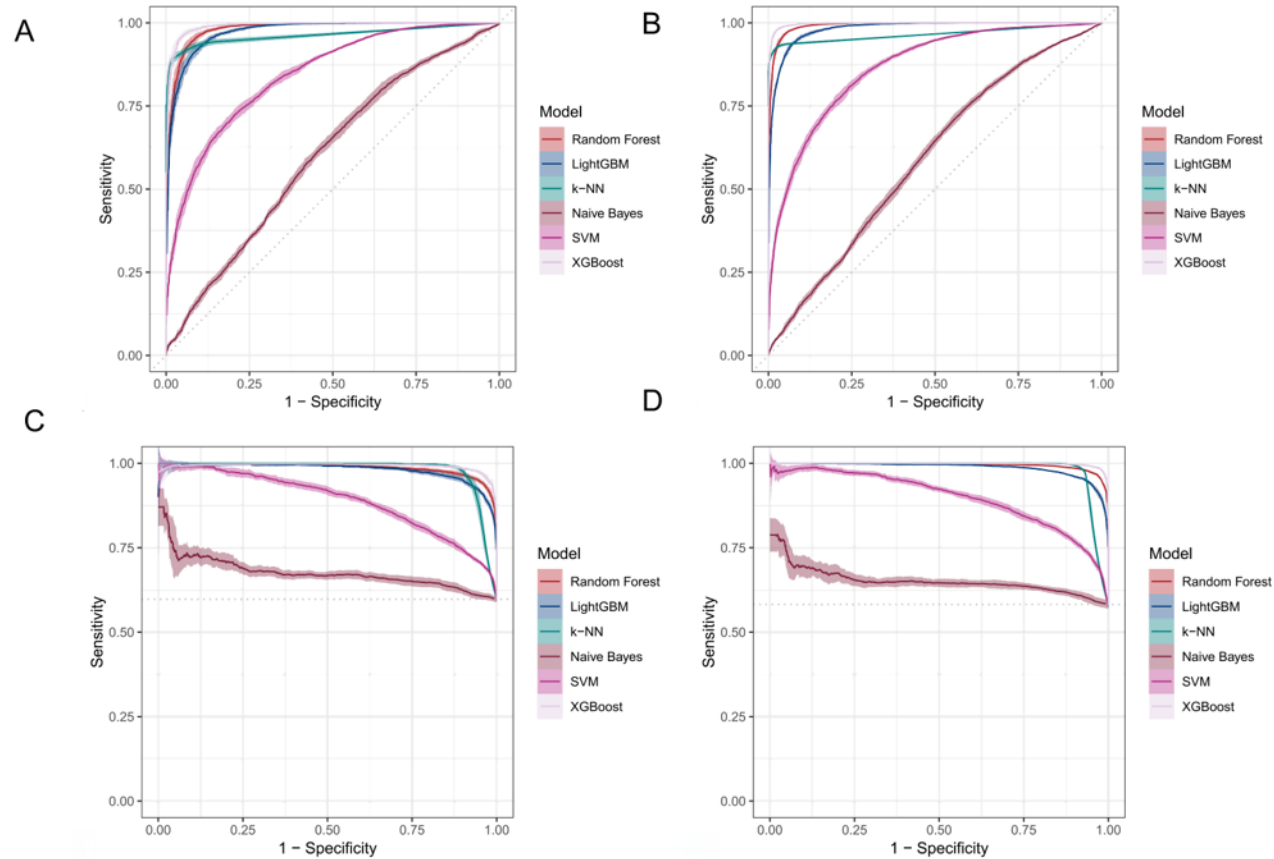


Figure 2. Performance comparison of six different prediction models. (A) Training set ROC curves. (B) Validation set ROC curves. (C) Training set recall rates. (D) Validation set recall rates.

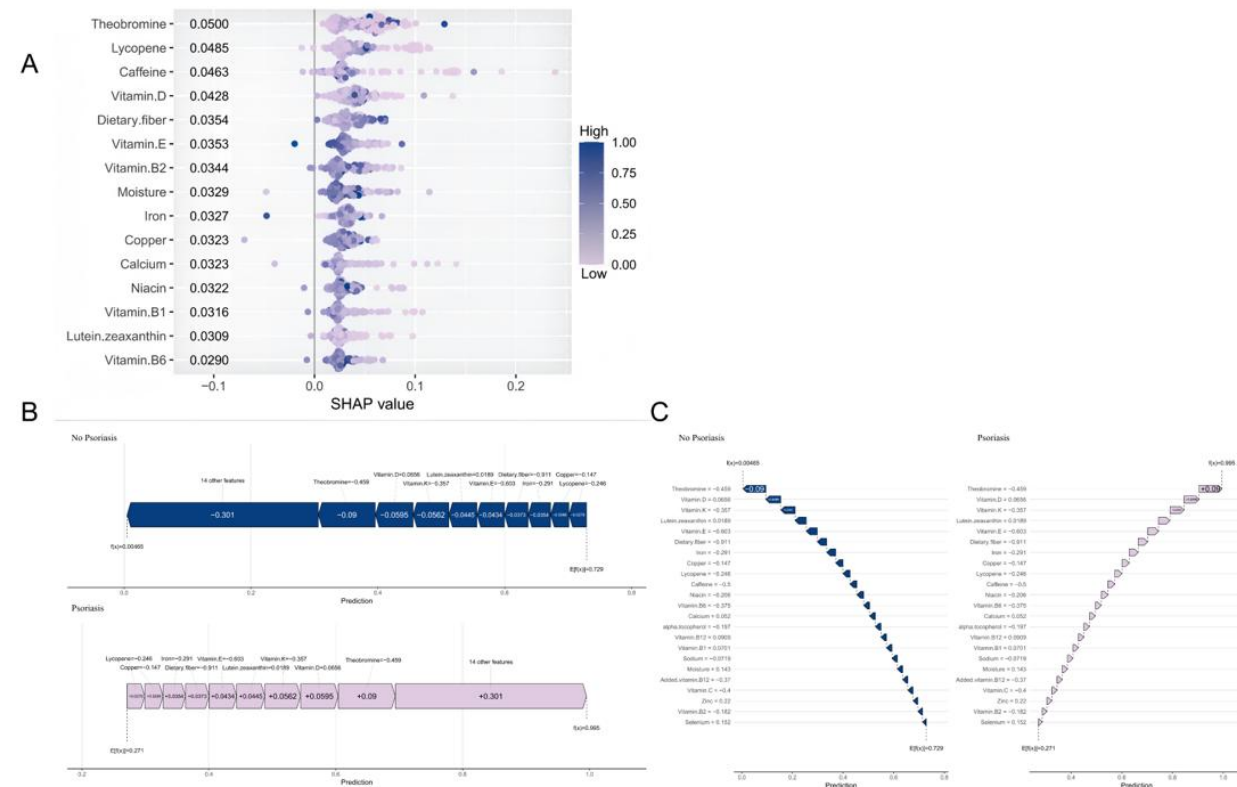


Figure 3. SHAP values of dietary features in the XGBoost model. (A) SHAP beeswarm plot. Each point represents a feature value, and the color represents the change in that feature value. Blue represents larger values, and pink-purple represents smaller values. (B) SHAP force plot and (C) SHAP waterfall plot show the contribution and influence of dietary features on the model output

Sensitivity analysis

A sensitivity analysis was performed using the original imbalanced dataset without SMOTE resampling. Although the predictive performance of the models was slightly reduced compared with the SMOTE-balanced dataset, the major dietary features identified by SHAP remained generally consistent (Supplementary Figure 4A–G). In particular, theobromine, caffeine, lycopene, vitamin D, vitamin E, and moisture continued to rank among the most influential predictors. These findings suggest that the observed feature importance patterns were relatively robust and not primarily driven by the SMOTE procedure.

DISCUSSION

In this paper, a dietary nutrient-based psoriasis risk prediction model was constructed using the ML algorithm, and the performance and stability of the model were evaluated. The results demonstrated that the XGBoost model had the best predictive performance with an AUC of 0.991. SHAP analysis identified theobromine, lycopene, caffeine, vitamin D, dietary fiber, and vitamin E as the most influential features contributing to model predictions.

To our knowledge, this is the first study to use ML algorithms to build and validate a psoriasis risk prediction model based on dietary nutrients, and the model also incorporates the baseline characteristics of the participants to enhance the predictive performance and accuracy of the model. In a previous study, the authors used the dietary inflammation index (DII) to explore the relationship between dietary inflammation and psoriasis risk, but did not find any association.³⁷ There are two possible explanations for this: first, the DII was only calculated based on a few dietary factors; second, there were less psoriasis patients than in the control group, which decreased the statistical effectiveness. In order to address the issue of dataset imbalance and enhance the built ML model's capacity to identify categories with few samples, we employed the Smote algorithm in this research.

In recent years, an increasing number of studies have used ML algorithms to construct risk prediction models for psoriasis (PsO) or psoriatic arthritis (PsA), including, but not limited to, diagnosing or differentiating between PsO and PsA;^{38,39} assessing the severity of psoriatic skin lesions;^{40,41} forecasting the therapeutic efficacy of biologics.^{42,43} For instance, Tsung-Ju Lee et al.⁴⁴ developed a risk prediction model for PsA using convolutional neural networks based on a large-scale and multidimensional electronic medical record in a chronological order, helping physicians to identify high-risk populations at an early stage. Muntyanu et al.,⁴⁵ based on demographic data from the Canadian province of Quebec, and gradient-boosting ML techniques identified the main risk factors for psoriasis incidence at the community level, including climate, vegetation, urbanization, and socio-economic characteristics of the community. Li et al.⁴⁶ investigated the association between probiotics and the prevalence of psoriasis using a range of machine learning algorithms, including multivariate logistic regression, decision tree (DT), AdaBoost, artificial neural network, RF, XGBoost, and K-KNN and found a non-significant negative correla-

tion. A subsequent meta-analysis showed that probiotic supplementation considerably lessened the severity of psoriasis. The article emphasizes that while probiotics help modulate the immune response and reduce inflammation, they are insufficient to lower the risk of psoriasis.⁴⁶ This is similar to the view of Liu, who suggested that while pro-inflammatory diets have the potential to contribute to the development of psoriasis, this role may remain subtle compared to other risk factors.³⁷

Although psoriasis is a genetically predisposed disease with a complex etiology, accumulating evidence suggests that both genetic susceptibility and environmental factors, including nutrition and metabolism, may contribute to disease development and progression.^{15,47} It is well known that psoriasis is an immune-inflammatory skin disease mediated by T cells, in which increased activity of Th17 lymphocytes and elevated pro-inflammatory cytokines such as IL-6 and TNF- α play a key role.⁹ Nakamizo et al.¹² found that high-fat diet (HFD) fed mice exhibited exacerbated psoriasis-like dermatitis, mainly as a result of an increased accumulation of V γ 4+ $\gamma\delta$ T cells in the skin which produced IL-17A and cytokines chemotactic for V γ 4+ $\gamma\delta$ T. In another study, researchers found that HFD was able to induce the accumulation of specific types of CD11c+ macrophages in the skin by triggering the NLRP3 inflammasome pathway and initiate IL-1 β and IL-18 signaling pathways, which in turn stimulate the activation of effector T cells and the proliferation of keratinocytes in skin tissues, thereby intensifying skin inflammation.⁴⁸ In addition, HFD may activate the immune system by altering the gut microbiota, resulting in a subclinical inflammatory state.⁴⁹ In contrast, a study conducted by Jensen et al. confirmed that a low-energy diet (800–1000kcal per day) reduced psoriasis lesions and improved the Dermatology Life Quality Index (DLQI).¹³ However, the aforementioned studies have emphasized more on the effect of diet on the severity of psoriasis, and the relationship between diet and the risk of psoriasis remains to be investigated.

Therefore, this paper attempted to construct psoriasis risk prediction models based on dietary nutrient features using six ML algorithms and evaluated important variables by SHAP analysis, including theobromine, lycopene, caffeine, vitamin D, dietary fiber, and vitamin E. Previous studies have reported associations between these dietary characteristics and psoriasis-related inflammatory pathways or disease activity. Vitamin E is a common antioxidant that reduces peroxyl radicals and forms tocolytic oxyl radicals,⁵⁰ attenuating oxidative stress and antioxidant defense imbalances associated with skin inflammation in psoriasis. Dietary fiber has systemic and intestinal anti-inflammatory effects, modulating intestinal inflammation by metabolizing short-chain fatty acids (SCFA), such as butyrate,⁵¹ and exerting a similar effect in the skin.⁵² SCFA ameliorates psoriasis skin lesions by promoting the activity of regulatory T cells (Treg). Previous findings have shown that sodium butyrate restores reduced regulatory T-cell (Treg), IL-10, and Foxp3 expression in the blood of patients with psoriasis and elevated IL-17 and IL-6 expression in skin lesions.⁵³ Vitamin D is an inflammatory regulator with an important role in the regulation of immune cells and inflammatory factors.

Vitamin D, on the one hand, acts on monocyte macrophages and down-regulates their production of inflammatory factors such as TNF- α , IL-1 β and IL-6.⁵⁴ Conversely, vitamin D regulates the proliferation and differentiation of T-cells, inhibits the production of IL-17A and IL-22 in Th17 cells⁵⁵ and induces the production of Tregs.⁵⁶ In addition, vitamin D has an inhibitory effect on the proliferation of keratinocytes and reduces the synthesis of psoriasis.⁹ Both Finamor et al.⁵⁷ and Gaal et al.⁵⁸ observed a decrease in disease activity in patients who were orally supplemented with vitamin D3. Lycopene is a carotenoid that scavenges oxygen free radicals,⁵⁹ protects against UV damage, and improves endothelial cell function.⁶⁰ Results of a previous cross-sectional study showed low serum levels of lycopene in patients with psoriasis.⁶¹ Shih et al.⁶² discovered that topical and oral administration of lycopene attenuated psoriasiform dermatitis in mice. This may be due to lycopene's ability to inhibit epidermal cell proliferation and reduce adhesion of monocyte molecules, which is extremely important in the pathogenesis of psoriasis.⁶³ Theobromine and caffeine are important bioactive constituents of coffee and belong to the group of methylxanthines, but their mechanisms of action are not identical. Theobromine possesses potent antioxidant and anti-inflammatory properties⁶⁴ and stimulates the production of nitric oxide to dilate blood vessels and improve blood circulation.⁶⁵ Caffeine inhibits Th1/Th2 cell proliferation and reduces the release of pro-inflammatory cytokines (IL-1, IL-6, TNF- α).⁶⁶ However, the effect of coffee appears to be dose-dependent, and excessive coffee intake has been associated with more severe psoriasis manifestations,⁹ which is also consistent with the findings of this paper.

Interestingly, the SHAP dependency plots demonstrated nonlinear relationships between several dietary nutrients and model-predicted psoriasis probability. Several explanations may account for these findings. First, threshold effects may exist, whereby dietary nutrients exert beneficial biological effects only after reaching a certain intake level. For example, antioxidant nutrients such as vitamin E, lycopene, and vitamin D may not substantially influence oxidative stress or immune regulation when intake is insufficient, whereas moderate intake levels may be associated with more favorable biological responses. Similar threshold-like dose-response relationships have been reported in nutritional epidemiology and dietary exposure studies.^{67,68} Second, saturation effects may occur because the physiological pathways involved in antioxidant defense, inflammatory regulation, and immune homeostasis have finite capacities. Consequently, increases in nutrient intake beyond a specific threshold may provide diminishing biological benefits and result in nonlinear exposure-response patterns.^{69,70} Third, residual confounding cannot be excluded. Individuals with higher intakes of antioxidant-rich nutrients may also engage in other healthy lifestyle behaviors that were not fully captured in the NHANES database, potentially contributing to the observed nonlinear associations. Furthermore, dietary intake in NHANES was assessed using 24-hour dietary recalls, which may introduce measurement error and increase variability at extreme intake levels. Similar nonlinear and dose-dependent relationships have been reported

for coffee consumption and psoriasis-related outcomes, suggesting that moderate intake may be associated with more favorable outcomes than either very low or excessive consumption.⁷¹ Therefore, the nonlinear patterns identified by SHAP may reflect a combination of biological mechanisms, exposure thresholds, saturation effects, and residual confounding, and should be further evaluated in prospective studies using explainable machine learning approaches.⁷²

This study successfully developed a prediction model with good predictive performance, suggesting that dietary nutrient information may improve the identification of individuals with a higher predicted probability of psoriasis. Meanwhile, SHAP analysis indicated that several dietary antioxidant-related nutrients were associated with lower model-predicted psoriasis probability.⁷³ However, these findings should not be interpreted as evidence of a causal protective effect. Rather, they highlight potentially relevant dietary features that warrant further investigation in prospective cohort studies and randomized controlled trials. In the future, greater awareness of anti-inflammatory and pro-inflammatory dietary patterns may support public health strategies aimed at promoting healthier dietary behaviors.

Because the prevalence of psoriasis in NHANES was relatively low, SMOTE was used to address class imbalance during model development. To assess whether the identified associations were driven by synthetic oversampling, we conducted a sensitivity analysis using the original dataset without SMOTE. The major dietary features identified by SHAP remained largely consistent across analyses, supporting the robustness of the findings. Nevertheless, synthetic oversampling may alter the underlying data distribution and potentially affect model estimates. Therefore, the identified dietary features should be interpreted as predictors contributing to model performance rather than definitive causal risk or protective factors.

From a clinical and public health perspective, the present findings may have several important implications. First, dietary nutrient profiles identified through explainable machine learning models may serve as complementary tools for identifying individuals with a higher predicted probability of psoriasis, consistent with the growing application of machine learning models in nutritional risk stratification and clinical prediction. Second, although causal relationships cannot be established from this cross-sectional study, the consistent identification of antioxidant-related nutrients, including vitamin D, vitamin E, dietary fiber, lycopene, caffeine, and theobromine, suggests that these dietary components may warrant further consideration in future nutritional recommendations for individuals at risk of psoriasis. Third, our findings provide potential targets for future intervention studies. Prospective cohort studies could evaluate whether long-term adherence to dietary patterns enriched in these nutrients is associated with psoriasis incidence, while randomized controlled trials could investigate whether modifying dietary intake influences disease onset, progression, or treatment response. Furthermore, the nonlinear associations observed in the SHAP analyses suggest that future intervention studies should consider potential threshold

and saturation effects when determining optimal intake levels. Collectively, these findings may contribute to the development of more precise nutrition-based prevention and management strategies for psoriasis in the future.^{73,74}

However, this study has several limitations. First, this paper utilized a cross-sectional design that did not allow for the determination of a causal relationship between dietary nutrients and psoriasis risk. Therefore, the associations identified by the machine learning models and SHAP analysis should be interpreted as predictive rather than causal. Second, there may have been recall bias because the diagnosis of psoriasis was partially relied on self-reported information from the NHANES interview questionnaire. Third, confounding factors could not be totally eliminated, even though covariates were as closely controlled as feasible throughout the investigation. In addition, this paper is a study based on the NHANES database, which mainly represents the dietary nutrient intake habits of the adult population in the U.S. The model's predictions could be impacted by dietary differences among various groups and geographical areas.

Therefore, further large-scale prospective cohort studies and multicenter RCTs are required to further explore the association in depth.

Conclusion

In this study, risk prediction models for psoriasis were developed using six machine learning models including RF, Light GBM, K-KNN, L-Naive Bayes, SVM, and XGBoost, of which XGBoost had the best prediction performance. Theobromine, lycopene, caffeine, vitamin D, dietary fiber, and vitamin E were identified as the main nutritional characteristics by SHAP analysis. The validity of the model will be further confirmed in the future and the model will be utilized in clinical practice to improve disease prevention and management.

ACKNOWLEDGEMENTS

This research was supported by Wonkwang University in 2025.

DISCLOSURE ON THE USE OF AI AND AI-ASSISTED TECHNOLOGIES

The authors reviewed and edited the content and takes full responsibility for the content of the publication.

CONFLICT OF INTEREST AND FUNDING DISCLOSURES

The authors declare that they have no competing interests.

The work was supported by Medical and Health Science Program of Zhejiang Province (2025KY144), Preclinical study of novel EGCG lead compounds for the treatment of vitiligo (LTGY23H110001) and Mechanistic study on the regulation of exosomal microRNA by Chai Hu Shuo Liver Tang in vitiligo treatment (2023ZL124).

REFERENCES

- Vanderpuye-Orgle J, Zhao Y, Lu J, et al. Evaluating the economic burden of psoriasis in the United States. *J Am Acad Dermatol*. 2015;72:961-7.e965.doi:10.1016/j.jaad.2015.02.1099.
- Boehncke WH, Schön MP. Psoriasis. *Lancet* (London, England). 2015;386:983-94.doi:10.1016/s0140-6736(14)61909-7.
- Strober B, Pariser D, Deren-Lewis A, Dickerson TJ, Lebwohl M, Menter A. A Survey of Community Dermatologists Reveals the Unnecessary Impact of Trial-and-Error Behavior on the Psoriasis Biologic Treatment Paradigm. *Dermatology and therapy*. 2021;11:1851-60.doi:10.1007/s13555-021-00573-1.
- Liu Y, Helms C, Liao W, et al. A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. *PLoS Genet*. 2008;4:e1000041.doi:10.1371/journal.pgen.1000041.
- Griffiths CE, Barker JN. Pathogenesis and clinical features of psoriasis. *Lancet* (London, England). 2007;370:263-71.doi:10.1016/s0140-6736(07)61128-3.
- Nestle FO, Kaplan DH, Barker J. Psoriasis. *N Engl J Med*. 2009;361:496-509.doi:10.1056/NEJMr0804595.
- Fu Y, Lee CH, Chi CC. Association of Psoriasis With Inflammatory Bowel Disease: A Systematic Review and Meta-analysis. *JAMA dermatology*. 2018;154:1417-23.doi:10.1001/jamadermatol.2018.3631.
- Dauden E, Blasco AJ, Bonanad C, et al. Position statement for the management of comorbidities in psoriasis. *J Eur Acad Dermatol Venereol*. 2018;32:2058-73.doi:10.1111/jdv.15177.
- Garbicz J, Całyniuk B, Górski M, et al. Nutritional Therapy in Persons Suffering from Psoriasis. *Nutrients*. 2021;14.doi:10.3390/nu14010119.
- Ni C, Chiu MW. Psoriasis and comorbidities: links and risks. *Clinical, cosmetic and investigational dermatology*. 2014;7:119-32.doi:10.2147/ccid.S44843.
- Murzaku EC, Bronsnick T, Rao BK. Diet in dermatology: Part II. Melanoma, chronic urticaria, and psoriasis. *J Am Acad Dermatol*. 2014;71:1053.e1051-3.e1016.doi:10.1016/j.jaad.2014.06.016.
- Nakamizo S, Honda T, Adachi A, et al. High fat diet exacerbates murine psoriatic dermatitis by increasing the number of IL-17-producing $\gamma\delta$ T cells. *Sci Rep*. 2017;7:14076.doi:10.1038/s41598-017-14292-1.
- Jensen P, Zachariae C, Christensen R, et al. Effect of weight loss on the severity of psoriasis: a randomized clinical study. *JAMA dermatology*. 2013;149:795-801.doi:10.1001/jamadermatol.2013.722.
- Barbieri JS, Beidas RS, Gondo GC, et al. Analysis of Specialist and Patient Perspectives on Strategies to Improve Cardiovascular Disease Prevention Among Persons With Psoriatic Disease. *JAMA dermatology*. 2022;158:252-9.doi:10.1001/jamadermatol.2021.4467.
- Wolters M. Diet and psoriasis: experimental data and clinical evidence. *Br J Dermatol*. 2005;153:706-14.doi:10.1111/j.1365-2133.2005.06781.x.
- Ford AR, Siegel M, Bagel J, et al. Dietary Recommendations for Adults With Psoriasis or Psoriatic Arthritis From the Medical Board of the National Psoriasis Foundation: A Systematic Review. *JAMA dermatology*. 2018;154:934-50.doi:10.1001/jamadermatol.2018.1412.
- Guo J, Li B. The Application of Medical Artificial Intelligence Technology in Rural Areas of Developing Countries. *Health equity*. 2018;2:174-81.doi:10.1089/heq.2018.0037.
- Schwalbe N, Wahl B. Artificial intelligence and the future of global health. *Lancet* (London, England). 2020;395:1579-86.doi:10.1016/s0140-6736(20)30226-9.
- May M. Eight ways machine learning is assisting medicine. *Nat Med*. 2021;27:2-3.doi:10.1038/s41591-020-01197-2.
- Lundberg SM, Erion G, Chen H, et al. From Local Explanations to Global Understanding with Explainable AI for Trees. *Nature Machine Intelligence*. 2020;2:56-67.doi:10.1038/s42256-019-0138-9.

21. Zhao W, Joshi, T., Nair, V. N. & Sudjianto. SHAP values for Explaining CNN-based Text Classification Models. 2020/8/26 [cited 2021/7/09]; Available from: <https://arxiv.org/abs/2008.11825>.
22. Ruan Z, Lu T, Chen Y, et al. Association Between Psoriasis and Nonalcoholic Fatty Liver Disease Among Outpatient US Adults. *JAMA dermatology*. 2022;158:745-53.doi:10.1001/jamadermatol.2022.1609.
23. Ma R, Cui L, Cai J, et al. Association between systemic immune inflammation index, systemic inflammation response index and adult psoriasis: evidence from NHANES. *Front Immunol*. 2024;15:1323174.doi:10.3389/fimmu.2024.1323174.
24. Cheng K, Liu X. Relationship between hysterectomy and stroke in National Health and Nutrition Examination Survey (NHANES) 2007-2018: a cross-sectional study. *Annals of translational medicine*. 2022;10:1209.doi:10.21037/atm-22-4681.
25. Casagrande SS, Lee C, Stoeckel LE, Menke A, Cowie CC. Cognitive function among older adults with diabetes and prediabetes, NHANES 2011-2014. *Diabetes Res Clin Pract*. 2021;178:108939.doi:10.1016/j.diabres.2021.108939.
26. Liu Y, Gong R, Luo G, et al. Associations of Triglycerides/High-Density Lipoprotein Cholesterol Ratio With Insulin Resistance, Impaired Glucose Tolerance, and Diabetes in American Adults at Different Vitamin D3 Levels. *Front Endocrinol (Lausanne)*. 2021;12:735736.doi:10.3389/fendo.2021.735736.
27. Huang L, Wang H, Wang Z, Wang Y, Zhang B, Ding G. Associations of Dietary Sodium, Potassium, and Sodium to Potassium Ratio with Blood Pressure- Regional Disparities in China. *Nutrients*. 2020;12.doi:10.3390/nul2020366.
28. Speiser JL, Miller ME, Tooze J, Ip E. A Comparison of Random Forest Variable Selection Methods for Classification Prediction Modeling. *Expert systems with applications*. 2019;134:93-101.doi:10.1016/j.eswa.2019.05.028.
29. Degenhardt F, Seifert S, Szymczak S. Evaluation of variable selection methods for random forests and omics data sets. *Brief Bioinform*. 2019;20:492-503.doi:10.1093/bib/bbx124.
30. Chen X, Ishwaran H. Random forests for genomic data analysis. *Genomics*. 2012;99:323-9.doi:10.1016/j.ygeno.2012.04.003.
31. Michael E, Ma H, Li H, Qi S. An Optimized Framework for Breast Cancer Classification Using Machine Learning. *Bio-med Res Int*. 2022;2022:8482022.doi:10.1155/2022/8482022.
32. Xia Y, Li X, Chen X, Lu C, Yu X. Inferring Retinal Degeneration-Related Genes Based on Xgboost. *Front Mol Biosci*. 2022;9:843150.doi:10.3389/fmolb.2022.843150.
33. Rodríguez-Pérez R, Vogt M, Bajorath J. Support Vector Machine Classification and Regression Prioritize Different Structural Features for Binary Compound Activity and Potency Value Prediction. *ACS omega*. 2017;2:6371-9.doi:10.1021/acsomega.7b01079.
34. Yao Z, Ruzzo WL. A regression-based K nearest neighbor algorithm for gene function prediction from heterogeneous data. *BMC Bioinformatics*. 2006;7:S11.doi:10.1186/1471-2105-7-s1-s11.
35. Irina R. An empirical study of the naive bayes classifier. study of the naive bayes classifier; 2001; ijcai 2001 workshop on empirical methods in artificial intelligence.
36. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. 2017/5/22 [cited 2017/11/25]; Available from: <https://arxiv.org/abs/1705.07874>.
37. Liu N, Zhang C, Hua W. Dietary inflammatory potential and psoriasis: A cross-sectional study. *The Journal of dermatology*. 2023;50:692-9.doi:10.1111/1346-8138.16731.
38. Wang Y, Zhang L, Yang M, et al. Development of a Predictive Model for Screening Patients with Psoriasis at Increased Risk of Psoriatic Arthritis. *Dermatol Ther (Heidelb)*. 2022;12:419-33.doi:10.1007/s13555-021-00663-0.
39. Li H, Chen G, Zhang L, Xu C, Wen J. A review of psoriasis image analysis based on machine learning. *Front Med (Lausanne)*. 2024;11:1414582.doi:10.3389/fmed.2024.1414582.
40. Meienberger N, Anzengruber F, Amruthalingam L, et al. Observer-independent assessment of psoriasis-affected area using machine learning. *J Eur Acad Dermatol Venereol*. 2020;34:1362-8.doi:10.1111/jdv.16002.
41. McMullen EP, Al Naser YA, Maazi M, et al. Predicting psoriasis severity using machine learning: a systematic review. *Clin Exp Dermatol*. 2025;50:520-8.doi:10.1093/ced/llae348.
42. Du AX, Ali Z, Ajgeiy KK, et al. Machine learning model for predicting outcomes of biologic therapy in psoriasis. *J Am Acad Dermatol*. 2023;88:1364-7.doi:10.1016/j.jaad.2022.12.046.
43. Emam S, Du AX, Surmanowicz P, Thomsen SF, Greiner R, Gniadecki R. Predicting the long-term outcomes of biologics in patients with psoriasis using machine learning. *Br J Dermatol*. 2020;182:1305-7.doi:10.1111/bjd.18741.
44. Lee LT, Yang HC, Nguyen PA, Muhtar MS, Li YJ. Machine Learning Approaches for Predicting Psoriatic Arthritis Risk Using Electronic Medical Records: Population-Based Study. *J Med Internet Res*. 2023;25:e39972.doi:10.2196/39972.
45. Muntyanu A, Milan R, Kaouache M, et al. Tree-Based Machine Learning to Identify Predictors of Psoriasis Incidence at the Neighborhood Level: A Populational Study from Quebec, Canada. *J Am Acad Dermatol*. 2024;25:497-508.doi:10.1007/s40257-024-00854-3.
46. Li Y, Liu X, Cao Z, Wu J. Probiotic Supplements Benefit Psoriasis Therapy Rather Than Affecting Disease Risk: Evidence from NHANES Machine-Learning and Meta-Analysis Study. *Probiotics Antimicrob Proteins*. 2024.doi:10.1007/s12602-024-10418-w.
47. Yang Z, Xiao W, Zhuang Z, et al. Genetically mimicked effects of evinacumab on psoriasis: a drug target Mendelian randomization study. *Asia Pac J Clin Nutr*. 2025;34:46-56.doi:10.6133/apjcn.202502_34(1).0004.
48. Zhang Y, Li Q, Rao E, et al. Epidermal Fatty Acid binding protein promotes skin inflammation induced by high-fat diet. *Immunity*. 2015;42:953-64.doi:10.1016/j.immuni.2015.04.016.
49. Laugerette F, Vors C, Gélouën A, et al. Emulsified lipids increase endotoxemia: possible role in early postprandial low-grade inflammation. *J Nutr Biochem*. 2011;22:53-9.doi:10.1016/j.jnutbio.2009.11.011.
50. Niki E. Role of vitamin E as a lipid-soluble peroxy radical scavenger: in vitro and in vivo evidence. *Free Radic Biol Med*. 2014;66:312.doi:10.1016/j.freeradbiomed.2013.03.022.
51. Ezzine C, Loison L, Montbrion N, et al. Fatty acids produced by the gut microbiota dampen host inflammatory responses by modulating intestinal SUMOylation. *Gut microbes*. 2022;14:2108280.doi:10.1080/19490976.2022.2108280.
52. Keshari S, Wang Y, Herr DR, et al. Skin Cutibacterium acnes Mediates Fermentation to Suppress the Calcium Phosphate-Induced Itching: A Butyric Acid Derivative with Potential for Uremic Pruritus. *J Clin Med*. 2020;9.doi:10.3390/jcm9020312.
53. Schwarz A, Philippsen R, Schwarz T. Induction of Regulatory T Cells and Correction of Cytokine Disbalance by Short-Chain Fatty Acids: Implications for Psoriasis Therapy.

- J Invest Dermatol. 2021;141:95-104.e102.doi:10.1016/j.jid.2020.04.031.
54. Umar M, Sastry KS, Al Ali F, Al-Khulaifi M, Wang E, Chouchane AI. Vitamin D and the Pathophysiology of Inflammatory Skin Diseases. *Skin Pharmacol Physiol*. 2018;31:74-86.doi:10.1159/000485132.
 55. Anderson J, Do LAH, Toh ZQ, et al. Vitamin D Induces Differential Effects on Inflammatory Responses During Bacterial and/or Viral Stimulation of Human Peripheral Blood Mononuclear Cells. *Front Immunol*. 2020;11:602.doi:10.3389/fimmu.2020.00602.
 56. Lamikanra AA, Tsang HP, Elsiddig S, et al. The Migratory Properties and Numbers of T Regulatory Cell Subsets in Circulation Are Differentially Influenced by Season and Are Associated With Vitamin D Status. *Front Immunol*. 2020;11:685.doi:10.3389/fimmu.2020.00685.
 57. Finamor DC, Sinigaglia-Coimbra R, Neves LC, et al. A pilot study assessing the effect of prolonged administration of high daily doses of vitamin D on the clinical course of vitiligo and psoriasis. *Dermato-endocrinology*. 2013;5:222-34.doi:10.4161/derm.24808.
 58. Gaál J, Lakos G, Szodoray P, et al. Immunological and clinical effects of alphacalcidol in patients with psoriatic arthropathy: results of an open, follow-up pilot study. *Acta dermato-venereologica*. 2009;89:140-4.doi:10.2340/00015555-0555.
 59. Khan N, Afaq F, Mukhtar H. Cancer chemoprevention through dietary antioxidants: progress and promise. *Antioxidants & redox signaling*. 2008;10:475-510.doi:10.1089/ars.2007.1740.
 60. Kim JY, Paik JK, Kim OY, et al. Effects of lycopene supplementation on oxidative stress and markers of endothelial function in healthy men. *Atherosclerosis*. 2011;215:189-95.doi:10.1016/j.atherosclerosis.2010.11.036.
 61. Rocha ACL, Bortoletto MC, da Costa AC, Oyafuso LKM, Sanudo A, Tomita LY. Low serum lycopene, and adequate α -tocopherol levels in patients with psoriasis: A cross-sectional study. *Nutrition and health*. 2022;28:239-48.doi:10.1177/02601060211014127.
 62. Shih CM, Hsieh CK, Huang CY, et al. Lycopene Inhibit IMQ-Induced Psoriasis-Like Inflammation by Inhibiting ICAM-1 Production in Mice. *Polymers*. 2020;12.doi:10.3390/polym12071521.
 63. Puig L. The role of IL 23 in the treatment of psoriasis. *Expert review of clinical immunology*. 2017;13:525-34.doi:10.1080/1744666x.2017.1292137.
 64. Navajas-Porras B, Pérez-Burillo S, Hinojosa-Nogueira D, Douros K, Pastoriza S, Rufián-Henares J. The Gut Microbiota of Obese Children Releases Lower Antioxidant Capacity from Food than That of Lean Children. *Nutrients*. 2022;14.doi:10.3390/nu14142829.
 65. Fisher ND, Hughes M, Gerhard-Herman M, Hollenberg NK. Flavanol-rich cocoa induces nitric-oxide-dependent vasodilation in healthy humans. *J Hypertens*. 2003;21:2281-6.doi:10.1097/00004872-200312000-00016.
 66. Sharif K, Watad A, Bragazzi NL, Adawi M, Amital H, Shoenfeld Y. Coffee and autoimmunity: More than a mere hot beverage! *Autoimmunity reviews*. 2017;16:712-21.doi:http://doi.org/10.1016/j.autrev.2017.05.007.
 67. Schwingshackl L, Schwedhelm C, Hoffmann G, et al. Food groups and risk of all-cause mortality: a systematic review and meta-analysis of prospective studies. *J Am Acad Dermatol*. 2017;105:1462-73.doi:10.3945/ajcn.117.153148.
 68. Aune D, Giovannucci E, Boffetta P, et al. Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality-a systematic review and dose-response meta-analysis of prospective studies. *Int J Epidemiol*. 2017;46:1029-56.doi:10.1093/ije/dyw319.
 69. Halliwell B. The antioxidant paradox: less paradoxical now? *British journal of clinical pharmacology*. 2013;75:637-44.doi:10.1111/j.1365-2125.2012.04272.x.
 70. Poljsak B, Šuput D, Milisav I. Achieving the balance between ROS and antioxidants: when to use the synthetic antioxidants. *Oxidative medicine and cellular longevity*. 2013;2013:956792.doi:10.1155/2013/956792.
 71. Poole R, Kennedy OJ, Roderick P, Fallowfield JA, Hayes PC, Parkes J. Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes. *BMJ (Clinical research ed)*. 2017;359:j5024.doi:10.1136/bmj.j5024.
 72. Chen C, Liu W, Lan Y, Li F, Li X. Exploring key genes in NAFLD linked to glutamine metabolism: A comprehensive analysis combining multi-omics, machine learning and SHAP. *Asia Pac J Clin Nutr*. 2026;35:81-90.doi:10.6133/apjcn.202602_35(1).0008.
 73. Chen CH, Pai KC, Hsieh HM, Chan YJ, Hsu HL, Wang CY. Artificial intelligence assisted nutritional risk evaluation model for critically ill patients: Integration of explainable machine learning in intensive care nutrition. *Asia Pac J Clin Nutr*. 2025;34:343-52.doi:10.6133/apjcn.202506_34(3).0009.
 74. Sun Q, Che J. Establishment and validation of a machine learning model to stratify malnutrition risk in hospitalized older patients with chronic heart failure. *Asia Pac J Clin Nutr*. 2026;35:143-57.doi:10.6133/apjcn.202602_35(1).0014.