Psoriasis: Targets and Therapy

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Machine Learning-Driven Prediction Models for Brodalumab Therapeutic Effect and Response Speed in Plaque Psoriasis

Lu Peng[1], Liyang Wang[2], Ling Chen[3], Zhu Shen[1]

[1]Department of Dermatology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, 510080, People's Republic of China; [2]School of Biomedical Engineering, Tsinghua University, Beijing, 100084, People's Republic of China; [3]Department of Dermatology, Daping Hospital, Army Medical University, Chongqing, 400042, People's Republic of China

Correspondence: Zhu Shen, Department of Dermatology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, No. 106, Zhongshan 2nd Road, Guangzhou, Guangdong, 510080, People's Republic of China, Email zhushencq@hotmail.com

Purpose: Biologic therapies have transformed plaque psoriasis treatment, but patient responses remain variable, neces+sitating machine prediction model for personalized therapy.

Patients and Methods: Transcriptomic and clinical data from moderate-to-severe psoriatic patient biopsies were sourced from GSE117468. Differential gene analysis identified Brodalumab treatment-associated genes. Lasso regression selected response-related genes, and LightGBM was used to build machine learning models. Model robustness was assessed using five-fold cross-validation.

Results: Biopsies (n=491) from 116 patients' lesional (LS) and non-lesional (NL) tissues were analyzed, divided into Brodalumab (140 mg or 210 mg) and placebo groups. Responders were defined as achieving ≥75% improvement in Psoriasis Area and Severity Index at week 12. Lasso identified genes from classical psoriasis pathways (IL-17, PPAR signaling, HLA-D alleles) and novel targets (WIF1, SLC44A5, LOC441528, SAA1). Six LightGBM models were trained to predict 12-week treatment response and 4-week response speed using LS, NL, and combined (LS_&_NL) data. LS_&_NL models showed superior performance, achieving AUC-ROC values of 95.14% (140 mg) and 92.83% (210 mg) for 12-week response prediction and 98.70% (140 mg) and 97.51% (210 mg) for 4-week response speed prediction.

Conclusion: These models provide robust tools for predicting Brodalumab response, supporting precision medicine and optimizing resource allocation in plaque psoriasis management.

Plain Language Summary:

Why was this study done?

Biologic drugs have improved the management of with moderate-to-severe plaque psoriasis, but responses vary. Predicting who will benefit most could aid clinicians in selecting optimal treatments, reduce trial-and-error approaches, and save healthcare resources.

What did the researchers do?

Scientists analyzed lesional and non-lesional skin biopsies from 116 psoriatic patients treated with Brodalumab (an IL-17RA inhibitor) or a placebo. Utilizing transcriptomic data and computer models, they:

Identify genes linked to treatment success.

Built tools to predict patients achieving significant improvement (≥75% clearer lesions) by week 12, or early respond by week 4.

What did they find?

Integration of data from both lesional (LS) and non-lesional (NL) specimens yielded the highest predictive accuracy. Key genes included known inflammation markers (eg, IL-17) and novel targets (eg, WIF1).

Models developed using LS_&_NL gene signatures predicted outcomes with high AUC-ROC:

12-week treatment success: 95.14% value for Brodalumab 140 mg dose, and 92.83% for 210 mg.

4-week early response: 98.70% value for Brodalumab 140 mg, and 97.51% for 210 mg.

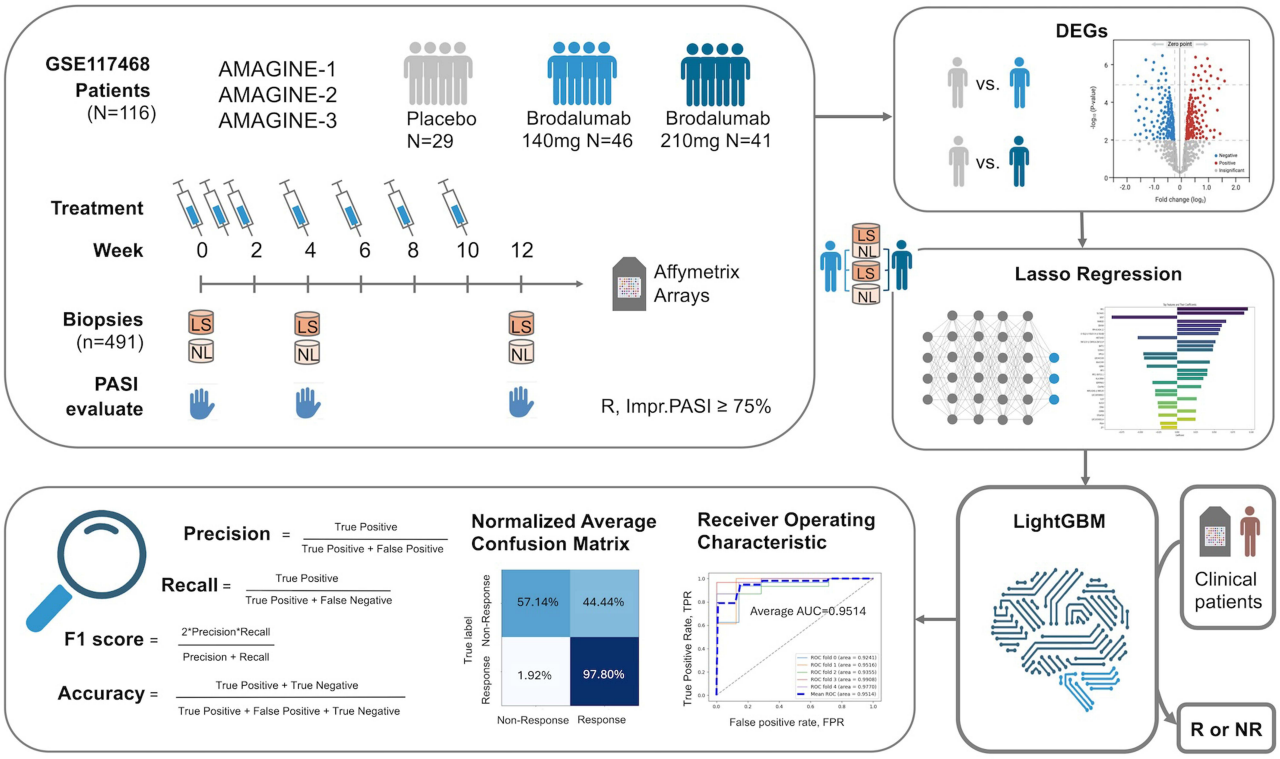
Received: 8 May 2025
Accepted: 12 August 2025
Published: 21 August 2025

What do these results mean?

These tools could aid clinicians in personalizing treatment regimens, giving patients faster relief and avoiding ineffective therapies. This approach may improve care quality and reduce healthcare costs. The study also uncovers new genetic clues that could inspire future psoriasis treatments.

Keywords: machine learning, treatment response prediction, plaque psoriasis

Graphical Abstract



Introduction

Plaque psoriasis is a complex chronic inflammatory skin disorder affecting 2–3% of the global population. Characterized by erythematous, scaly plaques, it causes both physical discomfort and psychological distress, significantly reducing patients' quality of life.[1,2] Despite advancements in therapy, treatment response variability continues to pose challenges in its management.[3,4]

The introduction of biologic therapies has significantly advanced the treatment of moderate-to-severe plaque psoriasis by targeting specific immune system components.[1] Brodalumab, a human monoclonal antibody that blocks the interleukin-17 (IL-17) receptor, has shown rapid efficacy in clinical studies. The median time to achieve ≥ 75% improvement in the Psoriasis Area and Severity Index (PASI 75) of approximately 4.2 weeks.[4,5] However, treatment responses to Brodalumab, like other biologics, vary among patients. Some fail to achieve optimal outcomes or experience diminished efficacy over time.[3,4] Subgroup analyses from the AMAGINE-2 and −3 Phase III trials revealed treatment failure in 26.84% (91/339) of patients receiving Brodalumab 210 mg Q2W and in 46.3% (273/590) of those treated with Ustekinumab at week 16.[4] Nevertheless, the substantial non-responders persist underscores the challenge of identifying patients who will benefit most, especially given the high costs and potential adverse effects of biologic therapies. We

therefore leveraged Brodalumab—the biologic with the highest documented response rate presently—to develop a predictive paradigm. This foundational model may extend to other biologics, optimizing precision medicine across psoriasis therapeutics.

In this context, precision medicine, which tailors treatments to individual patient characteristics, is becoming increasingly important.[6] The ability to predict a patient's response to specific therapies would not only improve clinical outcomes but also enhance the cost-effectiveness of psoriasis management by avoiding ineffective treatments.[7–9]

Artificial Intelligence (AI), particularly neural network-based approaches, offers promising opportunities to utilize large-scale genomic data from transcriptomic studies to predict treatment outcomes. AI models can uncover patterns and correlations in complex datasets, helping to identify patients most likely to respond favorably to specific therapies.[10,11] While earlier predictive models for psoriasis treatment response primarily utilized clinical variables,[8] systemic biomarkers (eg, serum cytokines)[9] or genetic polymorphisms (SNPs),[12–15] their clinical utility has been constrained by some key limitations: Blood-based biomarkers achieved suboptimal accuracy (AUC-ROC: 71–78% for anti-TNF/ JAK inhibitors) due to lack of tissue specificity;[9] Genetic polymorphisms (eg, HLA-Cw6) showed restricted utility,[12–15] only predicting response to specific biologics (ustekinumab: aHR=0.56),[14] unable to reflect dynamic transcriptomic changes during therapy; Systemic AI models which based on clinical variables yielded modest top-1 just accuracy 63.6% for long-term treatment success.[8]

In contrast, transcriptome-based models leveraging disease-relevant tissue transcriptomics directly interrogate the pathological microenvironment. Based on massive genetic and epigenetically influenced data, this approach is expected to overcome previous limitations and improve the accuracy of precision dermatology. In this study, we utilized the AI ensemble learning framework Light Gradient Boosting Machine (LightGBM) to analyze transcriptomic data from 491 specimens of 116 moderate-to-severe plaque psoriasis patients who participated in three Phase 3 clinical trials (AMAGINE-1, -2, -3) conducted at 155 centers across the United States from 2012 to 2015.[2,16] Clinical response data to Brodalumab treatment at weeks 4 and 12 (Q2W, subcutaneous injections on day 1 and weeks 1, 2, 4, 6, 8, and 10) were incorporated. Responders (R) were defined as patients achieving PASI 75 at week 12. This endpoint was selected based on: 1). Clinical relevance: PASI 75 is globally endorsed as the minimum threshold for meaningful therapeutic efficacy in moderate-to-severe psoriasis, distinguishing significant improvement from placebo responses; 2). Regulatory alignment: It serves as the primary endpoint in pivotal phase III trials of biologics (including AMAGINE-1, -2, -3 for Brodalumab);[1–3] 3). Patient-centricity: Achieving PASI 75 correlates with substantial improvements in dermatology-specific quality of life (DLQI) and treatment satisfaction.[17] Non-responders (NR) were those failing to attain PASI 75. Early response speed was defined as achieving PASI 75 by week 4.

Our objective is to develop robust predictive models capable of assessing a patient's potential response to Brodalumab (140mg/210mg) prior to treatment, including early (4-week) response probability. This framework holds transformative potential: 1) Personalizing drug selection by avoiding ineffective therapies; 2) Optimizing dosing (eg, prioritizing 140mg for predicted responders); 3) Empowering adherence through early-response confidence. Ultimately, this shifts plaque psoriasis management from chronic trial-and-error to targeted intervention, advancing precision dermatology practice. Furthermore, these tools can promote healthcare sustainability by optimizing resource allocation and reducing the economic burden of ineffective treatments.

Materials and Methods

Data Collection and Processing

Microarray transcriptomic data (GSE117468) were downloaded from the Gene Expression Omnibus (GEO) database using the getGEO function of the R package GEOquery. The corresponding clinical information for these patients was provided by Professor Mayte Suarez-Farinas from the Icahn School of Medicine at Mount Sinai. The data originating from three Phase III clinical trials (AMAGINE-1, -2, -3).[1–3] These trials aimed to evaluate the efficacy and safety of Brodalumab, a monoclonal antibody targeting the interleukin-17 (IL-17) receptor, in the treatment of moderate-to-severe plaque psoriasis. Lesional (LS) and non-lesional (NL) skin punch biopsies were collected for transcriptomic from patients at baseline (BL), as well as at week 4 and week 12 after Brodalumab 140mg or 210 mg treatment. Dimensionality reduction was conducted based on gene expression levels using the prcomp function in R. This study

used exclusively de-identified public data. Ethical exemption was approved by Guangdong Provincial People's Hospital Ethics Committee (KY2025-680-01) in accordance with the Declaration of Helsinki and China's ethical guidelines for secondary use of biomedical data.

Differential Gene Expression Analysis

We performed differential gene expression analysis using the limma package in R, comparing gene expression between different dosage groups and the placebo group. Kyoto Encyclopedia of Genes and Genomes (KEGG) Functional enrichment analysis was conducted on the selected Differential expression genes (DEGs).

Screening of Key Feature Genes

Lasso regression was employed to identify feature genes associated with Brodalumab response using the scikit-learn library on python pytorch.[18,19] The categorical criteria were the patient's response to Brodalumab treatment at 4 or 12 weeks (Patients who achieved ≥75% improvement in their Psoriasis Area and Severity Index [PASI] score were categorized as responders [R], while those who did not meet this threshold were categorized as non-responders [NR]). Optimal L1 regularization coefficients (α, alpha) were determined through 5-fold cross-validation to ensure that only the best classification performance feature gene sets were retained for further analysis.[20]

Machine Learning Model Construction

The feature gene sets were inputted into LightGBM on python for training prediction model.[21] Model parameters, including learning rate, maximum tree depth, and minimum data in leaf nodes, were optimized based on the characteristics and scale of each dataset to prevent overfitting while ensuring the model effectively learned from the data.[19]

Model Validation and Evaluation

Several evaluation metrics were utilized to assess the predictive performance of our models, including the Area under the Receiver Operating Characteristic Curve (AUC-ROC), confusion matrix, precision, recall, F1 score, and accuracy.

- *AUC-ROC*: The ROC curve plots the corresponding coordinate points of true positive rate (TPR) and the false positive rate (FPR) in the plane. AUC-ROC can intuitively evaluate the performance of classification.
- *Precision*: This measures the proportion of true positive (TP) predictions out of all positive predictions, calculated as:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

- *Recall (Sensitivity)*: This measures the proportion of true positives identified from the total actual positive cases, expressed as:

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

- *F1 Score*: This is the harmonic mean of precision and recall, calculated as:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

- *Accuracy*: This metric reflects the proportion of correctly predicted cases out of the total number of cases, calculated as:

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total\ Population}$$

Results

Analyzing Brodalumab Treatment-Associated Genes

After extracting data from GSE117468,[3] rigorous filtering processes were applied to remove poorly annotated genes, those with missing data (NA), and duplicates, yielding an expression matrix of 23,331 genes. Samples from the Ustekinumab group were excluded, leaving 491 lesional (LS) and non-lesional (NL) samples (n) from 116 patients (N) with moderate-to-severe plaque psoriasis, which were then grouped by treatment: Brodalumab 140 mg (n=207, N=46), Brodalumab 210 mg (n=167, N=41) and Placebo (n=117, N=29). No significant differences in gender, age, race, or PASI scores were observed between the groups ($P > 0.05$).[3] These samples were collected at baseline, week 4, and week 12 (Figure S1A).

DEGs between Brodalumab-treated groups (140 mg, 210 mg) and the placebo group were identified as Brodalumab treatment-associated genes. The top 1500 upregulated and top 1500 downregulated DEGs were gathered for Gene Set Enrichment Analysis (GSEA) to investigate their biological functions. The top DEGs for Brodalumab 140 mg were linked to the upregulation of pathways related to focal adhesion and actin cytoskeleton regulation (Figure S1B). This suggests 140 mg Brodalumab may primarily influence keratinocyte adhesion and cytoskeletal dynamics. In contrast, the top DEGs for Brodalumab 210 mg exhibited enrichment in pathways associated with calcium signaling, cytochrome P450 metabolism, neuroactive ligand-receptor interaction, and drug metabolism (Figure S1C), indicating a broader impact on cellular signaling and metabolic processes at the higher dose. Both dosage groups demonstrated significant inhibition of pathways including Toll-like receptor signaling, cytokine-cytokine receptor interaction, p53 signaling, and the cell cycle (Figure S2A, B and Appendix 1A, B), consistent with the known anti-psoriatic mechanism of IL-17 receptor blockade.

Models Predicting 12-Week Treatment Response

Screening Optimal 12-Week Treatment Response-Related Feature Genes

To identify feature genes associated with 12-week treatment responses, patients who were lost to follow-up before week 12 (N=15, n=32) were excluded from subsequent analysis. Clinical response classification at week 12, based on PASI 75, together with the top Brodalumab treatment-associated genes for each dosage, was used to perform Lasso (Least absolute shrinkage and selection operator) regression analysis. The analysis was conducted separately for combined LS and NL (LS_&_NL) samples, LS samples, and NL samples.

For the Brodalumab 140mg group, the optimal -log(α) values corresponding to the maximum AUC-ROC were 3.011 for LS_&_NL, 3.224 for LS, and 0.972 for NL samples. Similarly, for the Brodalumab 210mg group, the optimal -log(α) values were 3.588 for LS_&_NL, 1.649 for LS, and 0.838 for NL samples (Figures 1A, 1B and S3A).

The convergence plot of the model during 5-fold cross-validation illustrates how the Lasso model stabilizes under the selected optimal α coefficient. At this point, the mean squared error of feature gene weights across different validation sets was minimized, indicating improved classification performance (Figures 1C, 1D and S3B). Ultimately, six sets of feature genes were identified (Appendix 2). Under Brodalumab 140 mg treatment, 151 and 107 feature genes were identified for the LS_&_NL and LS groups, respectively, while 143 and 36 feature genes were identified for the corresponding groups under Brodalumab 210 mg. This demonstrates that integrating inflammatory (LS) and genetic (NL) signals harbors a richer signature predictive of response. In contrast, only 11 and 3 feature genes were selected for the NL groups, respectively (Appendix 2), highlighting the limited predictive information derived solely from non-lesional tissue.

Multidimensional Characteristics of 12-Week Treatment Response-Related Feature Genes

Heatmaps (Figures 1E, 1F and S3C) revealed clusters of highly correlated genes and independent genes, indicating that the selected features capture both co-regulated pathways and unique biological signals relevant to Brodalumab response.

Importance coefficients were calculated for each group, and the top 30 feature genes were visualized using histograms (Figures 1G, 1H and S3D). Notably, key genes overlapped across groups (eg, HLA-DQA1, HLA-DRB4, HLA-

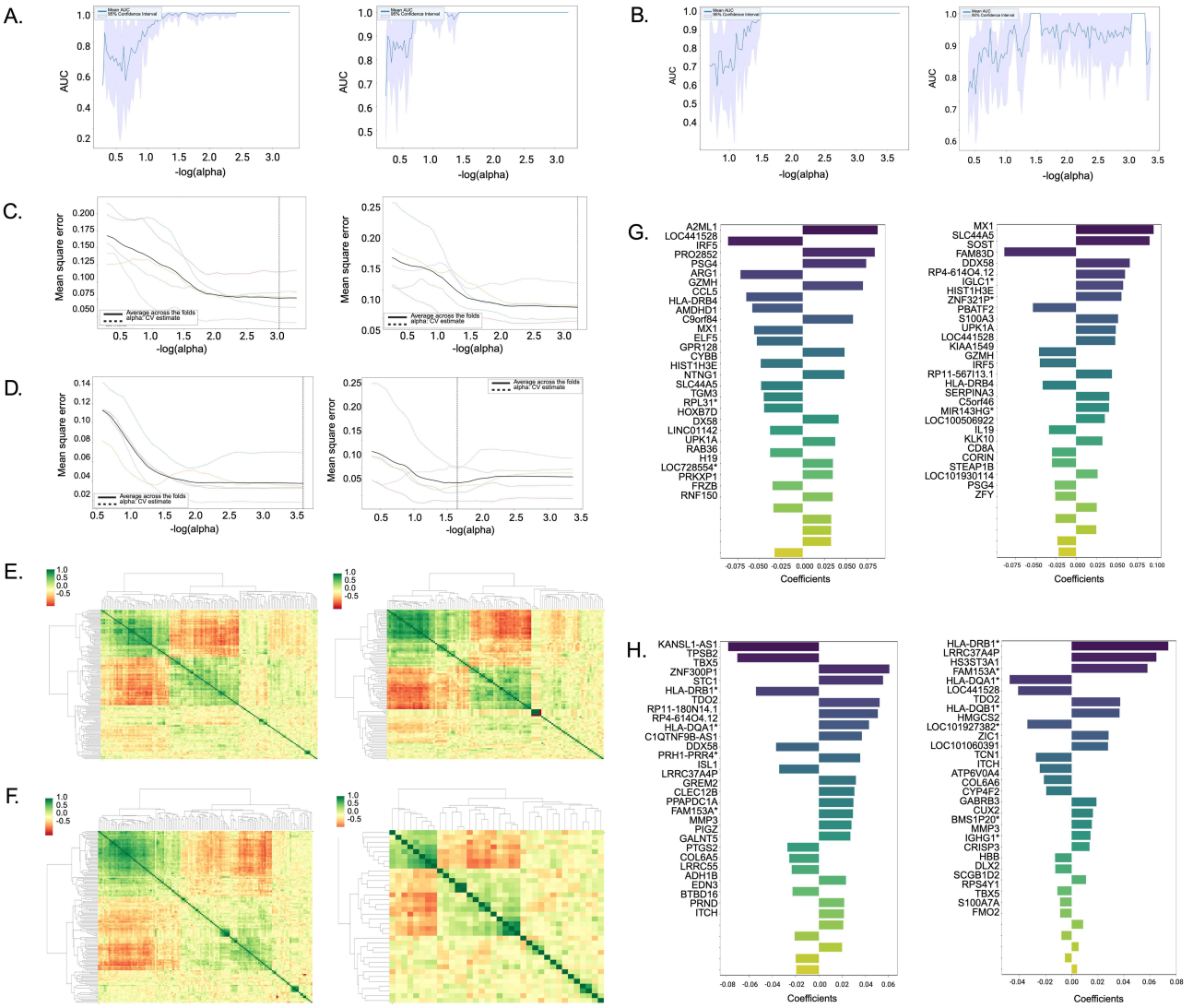


Figure 1 Feature Selection for 12-Week Response Prediction Models. (**A**) AUC-ROC variation plots for feature selection using Lasso regression, for the 12-week response prediction model in the LS_and_NL (left) and LS (right) groups under Brodalumab 140 mg treatment dosages. (**B**) AUC-ROC variation plots for feature selection using Lasso regression, for the 12-week response prediction model in the LS_and_NL (left) and LS (right) groups under Brodalumab 10 mg treatment dosages. (**C**) Convergence plots of the Lasso model at different α coefficients, for the 12-week response prediction model in the LS_and_NL (left) and LS (right) groups under Brodalumab 140 mg treatment dosages. The y-axis represents mean squared error (MSE), and the dashed line corresponds to the α coefficient with the minimum MSE. (**D**) Convergence plots of the Lasso model at different α coefficients, for the 12-week response prediction model in the LS_&_NL (left) and LS (right) groups under Brodalumab 210 mg treatment dosages. The y-axis represents mean squared error (MSE), and the dashed line corresponds to the α coefficient with the minimum MSE. (**E**) Heatmaps displaying the correlations between selected feature genes, for the 12-week response prediction model in the LS_and_NL (left) and LS (right) groups under Brodalumab 140 mg treatment dosages. (**F**) Heatmaps displaying the correlations between selected feature genes, for the 12-week response prediction model in the LS_and_NL (left) and LS (right) groups under Brodalumab 210 mg treatment dosages. (**G**) Histograms of the top 30 feature genes and their respective coefficients, for the 12-week response prediction model in the LS_and_NL (left) and LS (right) groups under Brodalumab 140 mg treatment dosages. (**H**) Histograms of the top 30 feature genes and their respective coefficients, for the 12-week response prediction model in the LS_and_NL (left) and LS (right) groups under Brodalumab 210 mg treatment dosages. *Phase ambiguity in chip sequencing.
Abbreviations: AUC-ROC, Area under the Receiver Operating Characteristic Curve; LS_and_NL, combined lesional and non-lesional samples; LS, lesional samples.

DQB1;[21,22] Figure 2A and B), Other significant genes included FMO2 (drug metabolism), MSMB (immunoglobulin-binding factor), SLC44A5 (regulator of transmembrane protein activity), WIF1(involved in mesoderm segmentation), BTC (epidermal growth factor family), PRR9 (associated with Psoriasis 4), and ADIPOQ (metabolic processes) (Figure 2A and B), suggesting a multi-faceted biological basis for Brodalumab response.

KEGG pathway analysis highlighted core immunological processes of psoriasis, including antigen presentation, cell adhesion molecules, and IL-17 signaling under both dosages. Interestingly, feature genes identified under the Brodalumab 140 mg LS_&_NL group did not show significant enrichment in the IL-17 pathway, potentially reflecting dose-specific modulation of alternative pathways. Whereas those under the 210 mg dosage were enriched in pathways such as

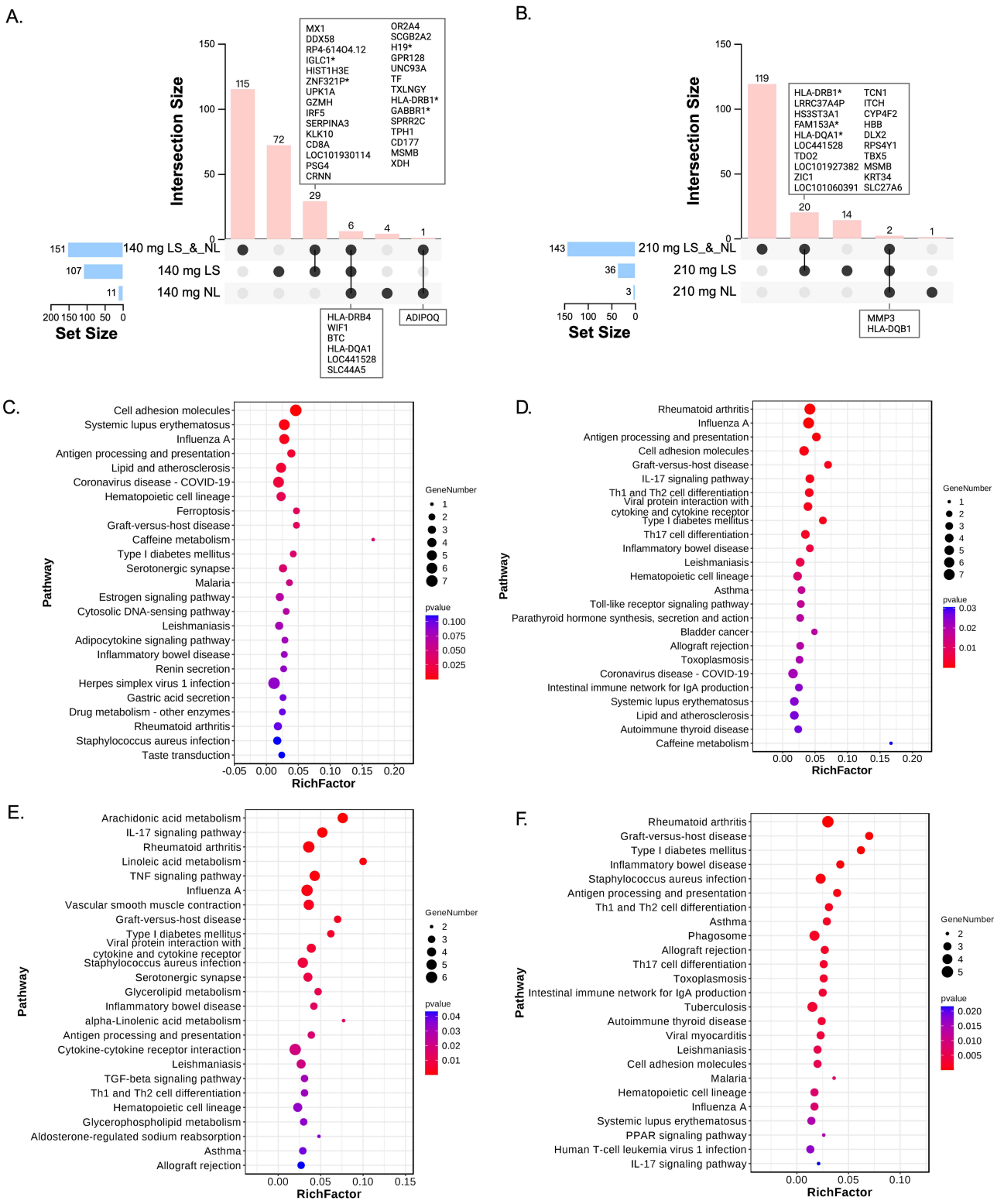


Figure 2 Feature Gene Pathway Enrichment for 12-Week Response Prediction Models. (**A**) Venn diagrams illustrating overlapping feature genes related to the 12-week response prediction models, identified in different lesional states under Brodalumab 140 mg dosages. (**B**). Venn diagrams illustrating overlapping feature genes related to the 12-week response prediction models, identified in different lesional states under Brodalumab 210 mg dosages. (**C**) KEGG pathway enrichment analysis of feature genes selected for the LS_&_NL group under Brodalumab 140 mg treatment dosages. (**D**) KEGG pathway enrichment analysis of feature genes selected for the LS group under Brodalumab 140 mg treatment dosages. (**E**) KEGG pathway enrichment analysis of feature genes related to the 12-week response prediction model, in the LS_and_NL group under Brodalumab 210 mg treatment dosages. (**F**) KEGG pathway enrichment analysis of feature genes the 12-week response prediction model, in the LS group under Brodalumab 210 mg treatment dosages. * Indicating phase ambiguity in chip sequencing.

Abbreviations: LS_and_NL, combined lesional and non-lesional samples; LS, lesional samples; NL, non-lesional samples.

https://doi.org/10.2147/PTT.S531925

arachidonic acid metabolism, linoleic acid metabolism, and TNF signaling (Figure 2C–F and Appendix 3). The finding emphasize that the carefully selected feature genes are not confined to identifying the IL-17 pathway but also reflect recognition of metabolic processes and immune-type preferences at the higher dose.

Protein interaction analyses were conducted on feature gene sets for the LS_&_NL and LS groups under Brodalumab 140 mg and 210 mg dosages. These analyses revealed interactions among 12-week treatment-response feature genes, while some genes remained functionally independent. The results were consistent with the patterns observed in the heatmaps (Figure S4A–D). NL groups under Brodalumab 140 mg and 210 mg dosages.

Exceeding 90% Mean AUC-ROC of 12-Week Response Prediction Models

Following Lasso-based feature selection, the feature gene sets were used to train 12-week response prediction models via LightGBM. To prevent overfitting, the models were tuned and validated using five-fold cross-validation, with AUC-ROC quantifying classification accuracy. For the LS_&_NL and LS groups, the five-fold average AUCs under Brodalumab 140 mg dosage were 95.14% and 91.13%, respectively (Figure 3A). Under the 210 mg dosage, the five-fold average AUC-ROC were 92.83% and 93.36%, respectively (Figure 3B). In contrast, the NL models showed weaker performance, with AUCs of 79.90% and 61.70% (Figure S5A and B).

Confusion matrices revealed that the LS_&_NL and LS models for both dosages achieved high true positive (TP) rates, exceeding 90% (98.08% and 97.80% (Figure 3C), 96.10% and 94.52% (Figure 3D), respectively). However, true negative (TN) rates were comparatively lower, at 55.56% and 57.14% (Figure 3C), 50.00% and 36.36%, respectively (Figure 3D). This pattern indicates the models are highly effective at identifying responders but slightly less accurate at

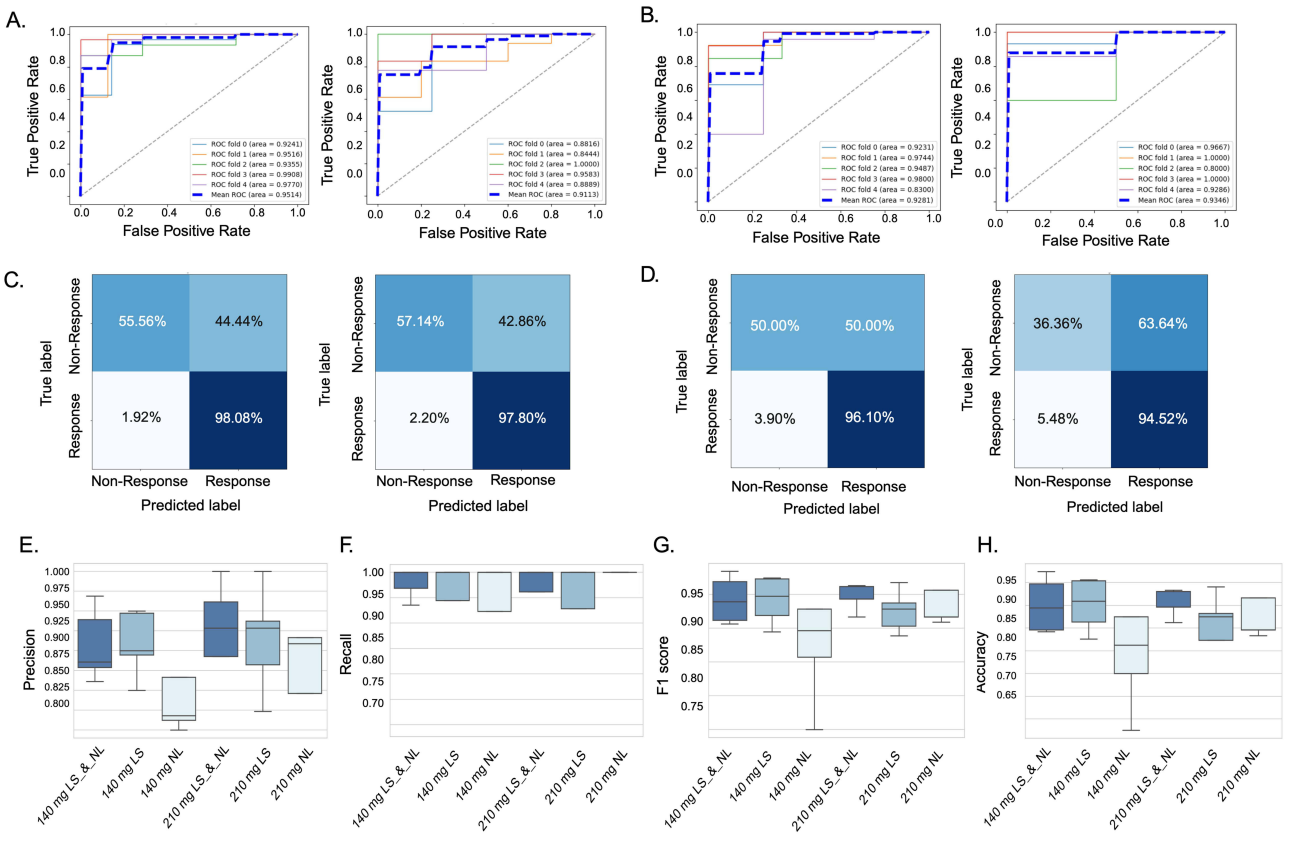


Figure 3 Evaluation of the 12-Week Response Prediction Model. (**A**) Five-fold cross-validation AUC-ROC for the 12-week response prediction models in the LS_and_NL (left) and LS (right) groups under Brodalumab 140 mg dosages. (**B**) Five-fold cross-validation AUC-ROC f for the 12-week response prediction models in the LS_&_NL (left) and LS (right) groups under Brodalumab 210 mg dosages. (**C**) Confusion matrices for the 12-week response prediction models in the LS_&_NL (left) and LS (right) groups under Brodalumab 140 mg dosages. (**D**) Confusion matrices for the 12-week response prediction models in the LS_&_NL (left) and LS (right) groups under Brodalumab 210 mg dosages. (**E–H**) Box plots of evaluation metrics for the 12-week response prediction models, including precision (**E**), recall (**F**), F1 score (**G**), and accuracy (**H**). **Abbreviations**: AUC-ROC, Area under the Receiver Operating Characteristic Curve; LS_and_NL, combined lesional and non-lesional samples; LS, lesional samples.

definitively ruling out non-responders based on these transcriptomic signatures. For NL models, TP and TN rates under the 140 mg dosage were 92.31% and 20.00%, respectively, while under the 210 mg dosage, these rates were 99.18% and 0.00% (Figure S5C and D). The low TN rates may be attributed to the imbalance between responder and non-responder cases in the NL samples.

Box plots of accuracy, precision, recall, and F1 scores provided a comprehensive evaluation of model performance (Figure 3E–H and Appendix 4), demonstrating robust and consistent predictive performance across evaluation metrics. For the LS_&_NL and LS models, precision and recall exceeded 80%, with F1 scores and accuracy generally above 90%. By contrast, the NL models exhibited greater variability in performance, underscoring their inferior reliability for clinical prediction.

Models Predicting 4-Week Response Speed

Selection and Characteristics of Response Speed-Related Feature Genes

To predict the 4-week response speed, we developed models using the 4-week response as the target variable. Feature gene sets for classifying the rapid response phenotype at week 4 were identified from treatment response-related genes using Lasso regression. During this process, the optimal -log(α) values corresponding to the maximum AUC-ROC for the Brodalumab 140 mg LS_&_NL, LS, and NL groups were 2.39, 1.32, and 0.83, respectively, as determined by AUC-ROC variation. Similarly, the optimal -log(α) values for the Brodalumab 210 mg LS_&_NL, LS, and NL groups were 2.28, 1.50, and 0.83 (Figures 4A, B and S6A), respectively. The optimal α coefficients minimized the mean square error during the convergence plot in 5-fold cross-validation, indicating optimal classification performance (Figures 4C, D and S6B).

For the LS_&_NL groups, 124 and 122 response speed-related feature genes were identified for the 140 mg and 210 mg dosages, respectively, while the LS groups identified 34 and 42 feature genes (Appendix 5). Bar charts illustrated the top 30 feature genes across treatment groups (Figure 4E and F). Key feature genes with high classification relevance and overlap between groups included HLA-D alleles, WIF1, SLC44A5, LOC441528, SAA1, and SYNPR. In contrast, the NL groups identified only five feature genes for each dosage, such as SAA1, HLA-DRB4 (both dosages), IGKV1-17, IGKC, and ADCYAP1 (140 mg), as well as IGHG1, IGK, and GSTT1 (210 mg) (Figure S6C and Appendix 5).

For the Brodalumab 140 mg dosage, response speed-related feature genes in the LS_&_NL group were enriched primarily in pathways such as cell adhesion, IL-17 signaling, graft-versus-host disease, and T helper cell differentiation. Under the 210 mg dosage, the LS_&_NL group's feature genes were mainly associated with the PPAR signaling pathway, cell adhesion molecules, and ECM-receptor interactions. In the LS groups, feature genes under both dosages largely overlapped with those in the LS_&_NL groups but also exhibited enrichment in pathways linked to Staphylococcus aureus infection and influenza A signaling (Figure 4G, H and Appendix 6).

Response Speed Prediction Models Achieving Over 97% True Positive Rates Using LS_&_NL Transcriptome

LightGBM models were trained using the selected features. The ROC curves for each model were plotted during five-fold cross-validation, and the average AUC-ROC values for the LS_&_NL models under 140 mg and 210 mg dosages were 98.70% and 97.51%, respectively (Figure 5A and B). For the LS models, the average AUC-ROC values were 93.54% and 84.88% for the two dosages (Figure 5A and B). In comparison, the AUC-ROC values for the NL models were 69.96% and 71.87%, respectively (Figure S6D). Confusion matrix analysis showed exceptional true positive rates for the LS_&_NL models, achieving 99.41% and 98.82% under the 140 mg and 210 mg dosages, respectively (Figure 5C and D). In contrast, the true negative rates were moderate, at 86.76% and 89.71%, respectively (Figure S6E). This pattern indicates the models are both highly effective at identifying 140 mg and 210 mg dosages responders and definitively ruling out non-responders based on LS_&_NL transcriptomic signatures.

Box plots of accuracy, precision, recall, and F1 scores further validated the strong and consistent performance of the LS_&_NL models. The median values for all dosages in the LS_&_NL and LS models exceeded 85% (Figure 5E–H). Notably, the LS_&_NL models under 140 mg achieved a recall close to 1, and F1 scores approaching 95%, reflecting their stability across all metrics (Appendix 7).

https://doi.org/10.2147/PTT.S531925

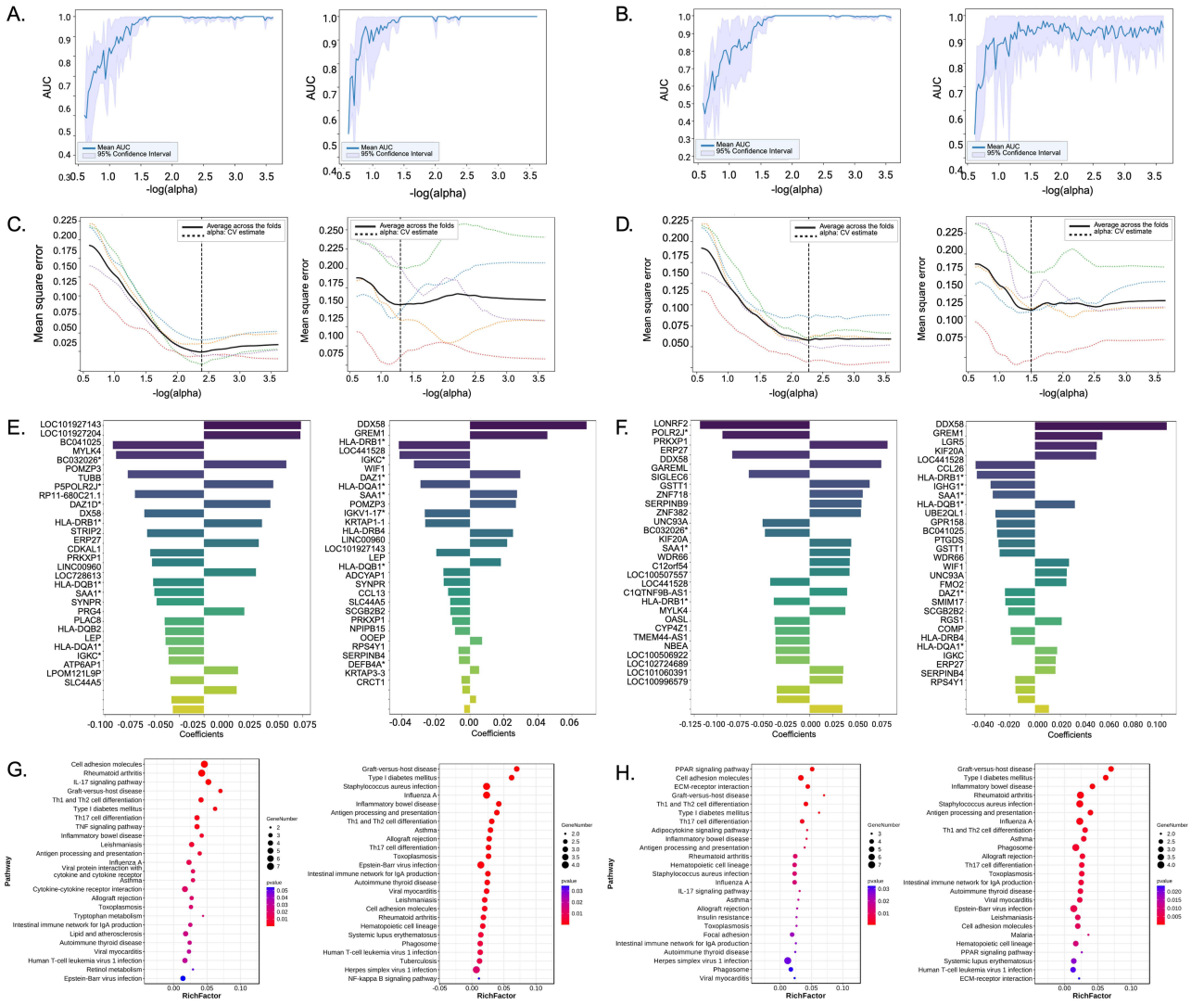


Figure 4 Feature Selection for 4-week Response Speed Prediction Models. (**A**) AUC-ROC variation plots for α coefficient selection in Lasso regression, for 4-week response speed prediction models in the LS_and_NL (left) and LS (right) groups under Brodalumab 140 mg dosages. (**B**) AUC-ROC variation plots for α coefficient selection in Lasso regression, for 4-week response speed prediction models in the LS_&_NL (left) and LS (right) groups under Brodalumab 210 mg dosages. (**C**) Convergence plots of Lasso models showing the learning process of the models as iterations increase, for 4-week response speed prediction models in the LS_&_NL (left) and LS (right) groups under Brodalumab 140 mg dosages. The y-axis represents mean squared error (MSE), and the dashed line corresponds to the α coefficient with the minimum MSE. (**D**) Convergence plots of Lasso models showing the learning process of the models as iterations increase, for 4-week response speed prediction models in the LS_&_NL (left) and LS (right) groups under Brodalumab 210 mg dosages. The y-axis represents mean squared error (MSE), and the dashed line corresponds to the α coefficient with the minimum MSE. (**E**) Bar charts displaying the contributions of the top 30 feature genes related to response speed, for 4-week response speed prediction models, in the LS_and_NL (left) and LS (right) groups under Brodalumab 140 mg dosages. (**F**) Bar charts displaying the contributions of the top 30 feature genes related to response speed, for 4-week response speed prediction models in the LS_&_NL (left) and LS (right) groups under Brodalumab 210 mg dosages. (**G**) KEGG functional enrichment analysis of selected feature genes related to 4-week response speed prediction models in the LS_and_NL (left) and LS (right) groups under Brodalumab 140 mg dosages. (**H**) KEGG functional enrichment analysis of selected feature genes related to 4-week response speed prediction models, in the LS_&_NL (left) and LS (right) groups under Brodalumab 210 mg dosages. *Phase ambiguity in chip sequencing. LS_and_NL, combined lesional and non-lesional samples. LS, lesional samples.

Abbreviation: AUC-ROC, Area under the Receiver Operating Characteristic Curve.

Discussion

Biologics are highly effective for moderate-to-severe plaque psoriasis, with randomized controlled trials showing superior skin clearance for the IL-17 receptor antagonist Brodalumab compared to TNF inhibitors and IL-17A or IL-12/IL-23 antibodies.[3,4,23] However, some patients remain non-responsive or develop resistance.[4] Biologics are costly and require long-term use due to the chronic nature of psoriasis, posing significant financial burdens even with insurance coverage. The development of predictive models for biologic response is thus essential, as it enables personalized treatment strategies while reducing economic strain on both patients and public healthcare.

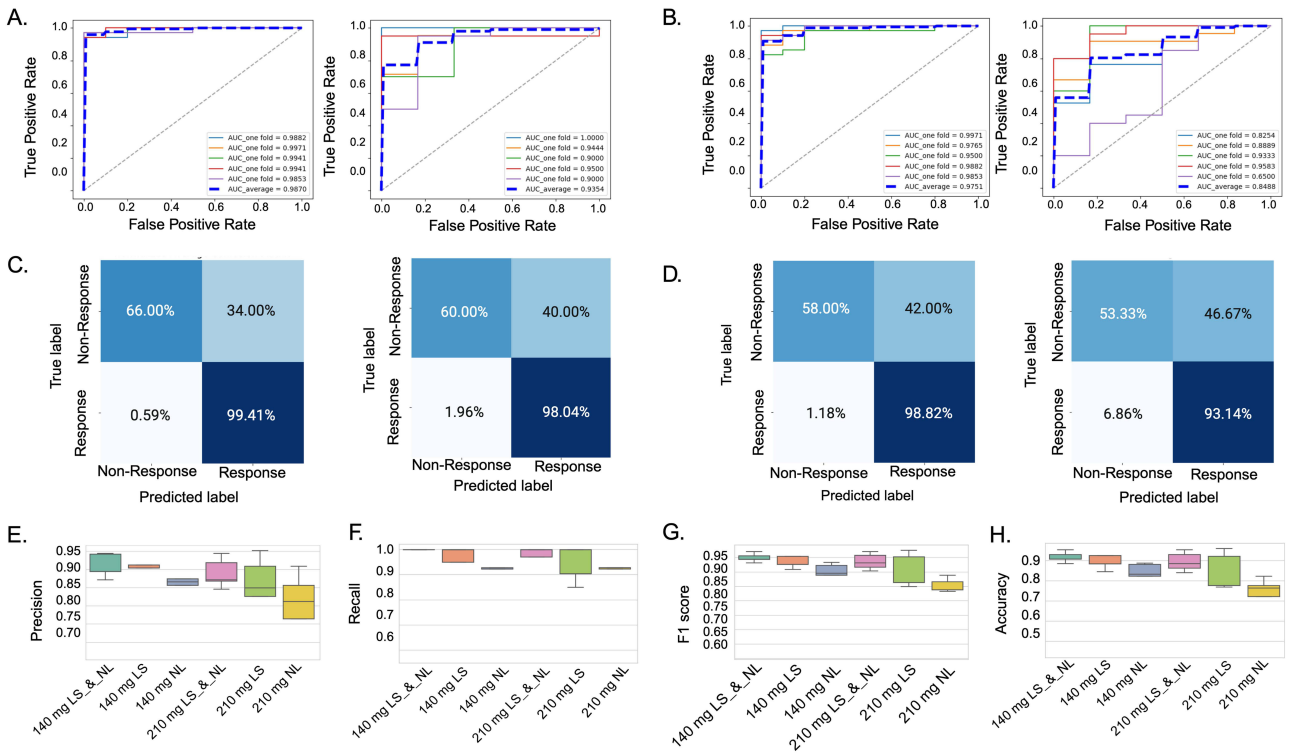


Figure 5 Validation of the 4-week Response Speed Prediction Models. (**A**) Five-fold cross-validation AUC-ROC for 4-week response speed prediction models in the LS_and_NL (left) and LS (right) groups under Brodalumab 140 mg dosages. (**B**). Five-fold cross-validation AUC-ROC for 4-week response speed prediction models in the LS_and_NL (left) and LS (right) groups under Brodalumab 210 mg dosages. (**C**). Confusion matrices evaluating prediction performance for 4-week response speed prediction models in the LS_and_NL (left) and LS (right) groups under Brodalumab 140 mg dosages. (**D**). Confusion matrices evaluating prediction performance for 4-week response speed prediction models in the LS_and_NL (left) and LS (right) groups under Brodalumab 210 mg dosages. (**E–H**). Box plots of evaluation metrics for 4-week response speed prediction models, including precision (**E**), recall (**F**), F1 score (**G**), and accuracy (**H**).

Abbreviations: AUC-ROC, Area under the Receiver Operating Characteristic Curve; LS_and_NL, combined lesional and non-lesional samples; LS, lesional samples.

Early predictive models for treatment response in plaque psoriasis initially centered on single nucleotide polymorphisms (SNPs) in susceptibility genes and limited clinical data.[12] Subsequent developments incorporated clinical markers, specific proteins,[8,9,13] and research on psoriasis susceptibility alleles.[14,15] However, disease pathogenesis typically involves the interaction of multiple key genes rather than the influence of a single gene. With advancements in artificial intelligence and neural networks, omics-based models have emerged as a promising approach, offering significant advantages in predictive accuracy.[10,11] In this study, we utilized lesional transcriptomic data collected from multiple time points (baseline, week 4, and week 12), skin conditions (lesional and non-lesional), and treatment dosages (Brodalumab 140 mg or 210 mg). Through machine learning, we identified complex genetic relationships influencing Brodalumab sensitivity and developed high-performance predictive models for 12-week treatment response and 4-week response speed.

Treatment satisfaction varies among individuals, and decisions regarding higher skin clearance levels (eg, PASI 90 or PASI 100) should be personalized. Optimal treatment outcomes should reflect patients' needs and priorities rather than being solely guided by physicians' objectives. Studies indicate that misalignment between patient and physician goals can affect satisfaction. For instance, nearly 70% of treatment goals between psoriasis patients and physicians in a Japanese study were misaligned.[17] Many patients weigh the benefits and risks of long-term Brodalumab use, including potential mental health concerns. Considering these factors, we selected PASI 75, rather than PASI 90 or PASI 100, as the benchmark for effective and satisfactory skin lesion treatment.

This study presents the first multi-dimensional model for predicting Brodalumab response. Transcriptomic data provide a comprehensive view of gene expression in disease-targeted tissues, alongside insights into a patient's genetic background, environmental factors, and physiological state. Our models predict the likelihood of response at week 12 for two Brodalumab dosages (140 mg and 210 mg), estimate response speed (week 4), and compare response probabilities

https://doi.org/10.2147/PTT.S531925

between dosages. Identifying responders to lower dosages helps minimize the risk of side effects associated with higher dosages, while predicting response speed enhances patient confidence in treatment adherence. We also developed models using transcriptomic data from non-lesional skin. This enables clinical patients to choose non-lesional biopsies for prediction, avoiding biopsy from cosmetically or functionally sensitive areas (eg, face or joints). It advances precision medicine by offering a more personalized and less invasive approach.

We applied Lasso regression for feature selection, combined with LightGBM for machine learning model construction,[18,19] and utilized five-fold cross-validation to assess model performance. This approach greatly improved model robustness in handling complex transcriptomic data. Lasso demonstrated distinct advantages in feature selection,[18] as increasing the penalty coefficient (α) systematically reduced irrelevant feature coefficients to zero, enabling the identification of genes and biomarkers most relevant to drug response. In the practical applications, this feature selection approach not only identified core biological pathways related to treatment but also revealed less-studied, independent signals from different dimensions, such as SLC44A5, BTC, WIF1, LOC441528, and MMP3. This provides strong support for precision medicine and disease research. LightGBM, a gradient-boosting framework leveraging tree-based models as weak learners, is renowned for its computational efficiency and high performance, making it particularly effective for large-scale datasets like transcriptomic data.[21]

Comprehensive performance evaluation metrics included ROC curves, confusion matrices, and boxplots from precision, recall, F1 score, and accuracy. Notably, LS_&_NL models predicting 12-week treatment responses and 4-week response speed for both Brodalumab 140 mg and 210 mg dosages performed exceptionally well, achieving AUC-ROC exceeding 90%. This performance, derived from skin tissue transcriptomic data, is notably superior to previous prediction models based on systemic inflammatory factors from circulating blood, which reported an AUC-ROC of 78% for tofacitinib prediction, and 71% for etanercept prediction.[9] Our skin biopsy- relied approach does add clinical trauma. Fortunately, the small scars of micro-punch biopsies (1–4mm) and the decreasing cost of NGS has improved feasibility. The risk of a higher financial burden of ineffective biologics makes the investment in biopsy for precision treatment of psoriasis cost-effective. Furthermore, our identification of core predictive genes (eg, HLA-D alleles, WIF1, SLC44A5) enables future development of targeted PCR- or NGS-based sequence panels. Such assays could be deployed routinely in clinical labs, drastically reducing costs and turnaround times versus full transcriptome analysis.

We analyzed why models of combined lesional and non-lesional (LS_&_NL) outperform those using LS or NL data alone. LS data models may mainly reflect the amplified inflammatory cascade, while NL data models may emphasize the patient's genetic background. Combining both integrates these factors, leading to a more effective predictive model.

This study has certain limitations. While our training data, sourced from 155 independent hospitals across the United States, is diverse and suitable for robust internal validation, external validation using an independent dataset would further strengthen the findings. Additionally, we are developing a web-based platform for clinical application. Moving forward, we plan to continuously collect external validation data and refine the model alongside its practical implementation.

Conclusion

Our study highlights the synergistic potential of advanced AI technologies and molecular data in transforming patient-centered healthcare. By providing novel insights into Brodalumab response in plaque psoriasis, this research lays the foundation for more personalized treatments. While AI shows great promise in enhancing treatment precision and patient outcomes, further efforts are needed. The identified features and models serve as a foundation for future studies, with the goal of optimizing clinical practice and improving patients' quality of life. Future work will focus on expanding sample sizes and integrating multi-omics data, such as proteomics and metabolomics, to refine prediction models and improve chronic plaque psoriasis management.

Abbreviations

AI, Artificial Intelligence; AUC-ROC, Area under the Receiver Operating; Characteristic Curve; BL, baseline; DEGs, differential expression genes; DLQI, dermatology-specific quality of life; FPR, false positive rate; GSEA, Gene Set Enrichment Analysis; IL-17, interleukin-17; KEGG, Kyoto Encyclopedia of Genes and Genomes; LightGBM, Light

Gradient Boosting Machine; LS, lesional; LS_&_NL, lesional and non-lesional; MSE, mean squared error; NL, non-lesional; NR, non-responders; PASI, Psoriasis Area and Severity Index; R, responders; SNPs, single nucleotide polymorphisms; TN, true negative; TP, true positive; TPR, true positive rate.

Data Sharing Statement

The skin transcriptome data used for model training and validation is available in the GEO database (GSE117468). Clinical information for the patients was provided by Professor Mayte Suarez-Farinas from the Icahn School of Medicine at Mount Sinai. Data generated in this study can be accessed upon reasonable request to the corresponding author, Dr. Zhu Shen. A web portal is currently under development. Researchers interested in making predictions or collaborating can Email deidentified transcriptomic or microarray data from psoriatic patients to the corresponding author. We will process the data using our models at no cost and provide feedback on the predicted outcomes.

Ethical Compliance

This study utilized publicly archived, fully de-identified transcriptomic data (GSE117468) and anonymized clinical metadata. Per Article 32 of the Declaration of Helsinki and China's Ethical Guidelines for Biomedical Research Involving Human Subjects (2016), secondary analysis of such data is exempt from additional ethics review. Formal exemption was confirmed by the Institutional Review Board of Guangdong Provincial People's Hospital (Ethics Ref: KY2025-680-01). All original trials (AMAGINE-1, -2, -3) obtained written informed consent and were approved by local ethics committees at participating centers.

Acknowledgments

We appreciated the bio-sample donations of participants of validation cohort in the study. We thank Professor Mayte Suarez-Farinas, Icahn School of Medicine at Mount Sinai, for providing clinical information. We express our gratitude to the contributors of GEO transcriptome dataset GSE117468.

Funding

Funding for this research was partially provided by the National Natural Science Foundation of China (82073444 and 82273537), Supporting scientific funds for talent introduction of Guangdong Provincial People's Hospital (No. KJ0120220181).

Disclosure

The authors declare no conflicts of interest or competing financial interests. ChatGPT 4o was used to assist in translating this manuscript from Chinese to English expression. The authors have carefully checked the full text and data. The authors take full responsibility for the use of generative AI in the preparation of this manuscript.

References

1. Griffiths CEM, Armstrong AW, Gudjonsson JE, Barker JNWN. Psoriasis. *Lancet*. 2021;397(10281):1301–1315. doi:10.1016/S0140-6736(20)32549-6
2. Boehncke W-H, Schön MP. Psoriasis. *Lancet*. 2015;386(9997):983–994. doi:10.1016/S0140-6736(14)61909-7
3. Tomalin LE, Russell CB, Garcet S, et al. Short-term transcriptional response to IL-17 receptor-A antagonism in the treatment of psoriasis. *J Allergy Clin Immunol*. 2020;145(3):922–932. doi:10.1016/j.jaci.2019.10.041
4. Langley RG, Armstrong AW, Lebwohl MG, et al. Efficacy and safety of brodalumab in patients with psoriasis who had inadequate responses to ustekinumab: subgroup analysis of two randomized phase III trials. *Br J Dermatol*. 2019;180(2):306–314. doi:10.1111/bjd.17318
5. Blauvelt A, Papp KA, Lebwohl MG, et al. Rapid onset of action in patients with moderate-to-severe psoriasis treated with brodalumab: a pooled analysis of data from two phase 3 randomized clinical trials (AMAGINE-2 and AMAGINE-3). *J Am Acad Dermatol*. 2017;77(2):372–374. doi:10.1016/j.jaad.2017.03.026
6. Bhinder B, Gilvary C, Madhukar NS, Elemento O. Artificial intelligence in cancer research and precision medicine. *Cancer Discov*. 2021;11(4):900–915. doi:10.1158/2159-8290.CD-21-0090
7. Correa da Rosa J, Kim J, Tian S, Tomalin LE, Krueger JG, Suárez-Fariñas M. Shrinking the psoriasis assessment gap: early gene-expression profiling accurately predicts response to long-term treatment. *J Invest Dermatol*. 2017;137(2):305–312. doi:10.1016/j.jid.2016.09.015
8. Nielsen ML, Petersen TC, Maul JT, et al. Multivariable predictive models to identify the optimal biologic therapy for treatment of patients with psoriasis at the individual level. *JAMA Dermatol*. 2022;158(10):1149–1156. doi:10.1001/jamadermatol.2022.3171

9. Tomalin LE, Kim J, Correa da Rosa J, et al. Early quantification of systemic inflammatory proteins predicts long-term treatment response to tofacitinib and etanercept. *J Invest Dermatol*. 2020;140(5):1026–1034. doi:10.1016/j.jid.2019.09.023
10. Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. *N Engl J Med*. 2023;388(13):1201–1208. doi:10.1056/NEJMra2302038
11. Yang Y, Zhao Y, Liu X, Huang J. Artificial intelligence for prediction of response to cancer immunotherapy. *Semin Cancer Biol*. 2022;87:137–147. doi:10.1016/j.semcancer.2022.11.008
12. Prieto-Pérez R, Cabaleiro T, Daudén E, Abad-Santos F. Gene polymorphisms that can predict response to anti-TNF therapy in patients with psoriasis and related autoimmune diseases. *Pharmacogenomics J*. 2013;13(4):297–305. doi:10.1038/tpj.2012.53
13. García-Doval I, Pérez-Zafrilla B, Ferrandiz C, et al. Development of clinical prediction models for good or bad response to classic systemic drugs, anti-TNFs, and ustekinumab in psoriasis, based on the BIOBADADERM cohort. *J DermatolTreat*. 2016;27(3):203–209. doi:10.3109/09546634.2015.1088130
14. Papini M, Cusano F, Romanelli M, et al. Secukinumab shows high efficacy irrespective of HLA-Cw6 status in patients with moderate-to-severe plaque-type psoriasis: results from extension phase of the SUPREME study. *Br J Dermatol*. 2019;181(2):413–414. doi:10.1111/bjd.18013
15. Alabas OA, Mason KJ, Yiu ZZN, et al. The association of age at psoriasis onset and HLA-C*06:02 with biologic survival in patients with moderate-to-severe psoriasis: a cohort study from the British Association of Dermatologists Biologics and Immunomodulators Register (BADBIR). *Br J Dermatol*. 2024;190(5):689–700. doi:10.1093/bjd/ljad481
16. Papp K, Menter A, Leonardi C, et al. Long-term efficacy and safety of brodalumab in psoriasis through 120 weeks and after withdrawal and retreatment: subgroup analysis of a randomized phase III trial (AMAGINE-1). *Br J Dermatol*. 2020;183(6):1037–1048. doi:10.1111/bjd.19132
17. Svoboda SA, Ghamrawi RI, Owusu DA, Feldman SR. Treatment goals in psoriasis: which outcomes matter most? *Am J Clin Dermatol*. 2020;21(4):505–511. doi:10.1007/s40257-020-00521-3
18. Nath A, Cohen AL, Bild AH. ENDORSE: a prognostic model for endocrine therapy in estrogen-receptor-positive breast cancers. *Mol Syst Biol*. 2022;18(6). doi:10.15252/msb.202110558
19. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ*. 2024;384:e074819. doi:10.1136/bmj-2023-074819
20. Emmert-Streib F, Dehmer M. High-dimensional LASSO-based computational regression models: regularization, shrinkage, and selection. *Mach Learn Knowl Extr*. 2019;1(1):359–383. doi:10.3390/make1010021
21. Ke G, Meng Q, Finley T, et al. Lightgbm: a highly efficient gradient boosting decision tree. *Adv Neural Inf Process Syst*. 2017;30:3149–3157.
22. Yap CF, Nair N, de Vries A, et al. HLA-DRB1 and HLA-DQA1 associated with immunogenicity to Adalimumab therapy in patients with rheumatoid arthritis. *Ann Rheum Dis*. 2024;83(2):263–265. doi:10.1136/ard-2023-223955
23. Lebwohl M, Strober B, Menter A, et al. Phase 3 studies comparing brodalumab with ustekinumab in psoriasis. *N Engl J Med*. 2015;373(14):1318–1328. doi:10.1056/NEJMoa1503824