Paper 29:

**Title**

* **Therapeutic application of machine learning in psoriasis: A Prisma systematic review**

**Authors and Affiliations**

* **Authors**: Snehal Balvant Lunge, Nandini Sundar Shetty, Vidyadhar R. Sardesai, Priyanka Karagaiah, Paul S. Yamauchi, Jeffrey M. Weinberg, Leon Kircik, Mario Giulini, Mohamad Goldust
* **Affiliations**: Authors are affiliated with institutions in dermatology and dermatologic research, spanning universities and dermatology institutes in India, USA, and Germany.

**Year of Publication**

* **2023**

**Venue**

* **Published in**: Journal of Cosmetic Dermatology

**Summary**

* This systematic review evaluates machine learning (ML) applications in the therapeutic management of psoriasis. The focus is on non-invasive techniques for predicting treatment outcomes, optimizing patient management, and standardizing diagnosis for greater consistency across practitioners. The study underscores the increasing integration of ML in dermatology, particularly for diseases like psoriasis, where visual and quantitative assessment methods can benefit significantly from machine learning-based objectivity and precision.

**Contribution**

* **Key Contribution**: Consolidates and critically reviews existing ML models and methods for therapeutic applications in psoriasis. This work highlights advances in ML algorithms for analyzing treatment response, determining treatment efficacy, and improving patient outcome prediction. The review acts as a reference for future applications of ML in clinical dermatology, with an emphasis on treatment management rather than diagnosis alone.

**Domain**

* **Field**: Dermatology and Artificial Intelligence (AI), specifically focusing on applications of Machine Learning (ML) for managing dermatological conditions such as psoriasis.

**Motivations**

* Psoriasis, an immune-mediated condition, varies significantly in how it affects patients, leading to disparate treatment responses. Dermatological assessments, often highly subjective, can vary by practitioner and assessment method. This review is motivated by the need for more consistent, objective tools to enhance diagnostic accuracy, manage treatment efficacy, and predict patient-specific responses. ML-based tools offer an opportunity to mitigate these variances and improve treatment standardization, accuracy, and efficiency.

**Problems Addressed**

* **Inconsistencies in Clinical Evaluation**: Variations in dermatological assessments result in inconsistencies in diagnoses and treatment responses.
* **Difficulty in Predicting Treatment Outcomes**: The complexity of individual patient response to psoriasis treatments requires advanced tools to predict efficacy accurately.
* **Manual and Subjective Assessments**: Current evaluations, like PASI scoring, are limited by subjective, human-dependent interpretation.

**Challenges Identified**

* **Data Diversity and Generalization**: ML models must generalize across diverse populations with varying treatment responses, making it challenging to develop a one-size-fits-all model.
* **Data Quality and Quantity**: Limited datasets, lack of uniformity in patient data, and incomplete information across studies challenge the robustness and accuracy of ML applications.
* **Computational Complexity**: Developing models that balance predictive accuracy with computational feasibility, particularly when using deep learning methods that require significant data and resources.

**Related Work**

* This study reviews ML applications in dermatology, specifically focused on psoriasis. Key related studies include:
  + **Savolainen et al.**: Compared ML-based lesion detection with human assessment to evaluate body surface involvement in psoriasis.
  + **Meienberger et al.**: Developed a neural network model with a penalty factor to improve predictions of psoriatic lesions.
  + **Damiani et al.**: Created predictive models for fast-response to secukinumab treatment using Artificial Neural Networks (ANNs).
  + **Tomalin et al.**: Utilized ML to predict treatment response using biochemical markers of inflammation and cardiovascular risk.
  + **Zang et al.**: Developed a semantic prediction-based system using **SemMedDB**, a structured knowledge database from MEDLINE, to identify potential drug-drug interactions relevant to psoriasis treatment. By analyzing semantic relationships, the system could identify interactions that might go unnoticed with conventional methods. For example, it flagged the likelihood of severe psoriasis exacerbation when **Lisinopril** is used in conjunction with **Sertraline**. This approach highlights the potential of ML to uncover complex drug interactions and improve treatment safety for psoriasis patients.

**Techniques Used**

* Artificial Neural Networks (ANN): Applied for identifying fast-responders among patients, creating semantic maps of relevant biochemical indicators.
* Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA): Used for classification and reduction of spectra data, enabling ML-based lesion classification and treatment prediction.
* Bagging and Ensemble Methods: Employed to aggregate various models, enhancing the predictive accuracy over standalone models.
* Semantic Prediction with MEDLINE Data (SemMedDB): Leveraging the SemMedDB database, this technique uses data from biomedical literature to identify possible drug interactions by mapping relationships among drugs, genes, and proteins. This method was effective in predicting critical drug interactions, such as the interaction between Lisinopril and Sertraline, which was shown to potentially worsen psoriasis severity.

**Datasets**

* **Data Types**: Patient records, treatment response markers, protein biomarkers, spectroscopy data, and psoriasis lesion images.
* **Specific Datasets**: Most reviewed studies rely on proprietary datasets, such as longitudinal serum profiles or national registries, for model training and validation. For example, one study used inflammatory and cardiovascular biomarkers to predict responses to biologic therapies like tofacitinib.

**Methodology**

* **Literature Search**: Conducted a comprehensive search on PUBMED, using a combination of terms such as “Psoriasis,” “Therapeutics,” “Treatment,” and other machine learning-related terms (e.g., “artificial intelligence,” “deep learning,” “neural network”).
* **Selection Process**: Applied the PRISMA framework to filter relevant studies, ultimately narrowing down to eight studies that best demonstrated therapeutic applications of ML for psoriasis treatment.

**Proposed Model**

* **Model Design**: No specific ML model was proposed by the authors. Instead, the paper reviews multiple models developed by different studies, such as neural networks for response prediction and regression models for biomarker analysis.

**Results**

* **Model Performance**: ML models show promising results in achieving consistency and accuracy for treatment prediction, surpassing traditional manual assessments in some cases. For example, Savolainen et al. reported that ML models could more accurately assess the affected body surface area than dermatologists' manual estimates.

**Evaluation Metrics**

* **Metrics Used**: Accuracy, F1 score, PASI score comparisons, and specific area differences for lesion assessments.
* **Study-Specific Metrics**: Each study employed metrics suited to its dataset, such as precision in lesion identification or recall in treatment response prediction, depending on the application.

**Limitations**

* The paper is limited by its exclusive reliance on PUBMED for data sources, potentially missing out on relevant studies from other databases (e.g., EMBASE, Google Scholar, SCOPUS). The focus on specific search terms might have restricted the study's comprehensiveness.

**Future Work**

* **Expansion of ML Applications**: Suggests further exploration of ML models across diverse datasets and varying psoriasis treatment plans to achieve broader applicability.
* **Interdisciplinary Approaches**: Recommends integrating AI with other technologies, such as wearable devices, to gather real-time patient data for enhanced prediction accuracy.
* **Enhanced Generalization**: Emphasizes the need for models that can generalize to diverse patient demographics and clinical conditions for better real-world utility.

Paper 9 ----diagnosis & treatment

**1. Authors**

Mohammad Yaseliani, Abtin Ijadi Maghsoodi, Erfan Hassannayebi, and Uwe Aickelin.

**2. Year**

Published in 2024, with availability online from November 2023.

**3. Venue**

Published in *Computers & Industrial Engineering*.

**4. Title**

"Diagnostic clinical decision support based on deep learning and knowledge-based systems for psoriasis: From diagnosis to treatment options."

**5. Summary**

The study presents a Decision and Diagnostic Support System (D&DSS) for psoriasis using deep learning-based Computer-Aided Diagnosis (CAD) and treatment suggestions. The system utilizes an ensemble of Convolutional Neural Networks (CNNs) to classify psoriasis images and suggests optimal treatments based on Multi-Criteria Decision Making (MCDM) methods.

**6. Contribution**

The study’s main contributions include:

* Developing an ensemble CNN model using ResNet50V2, ResNet101V2, and ResNet152V2 for detecting and classifying psoriasis images.
* Integrating a treatment recommendation system using MCDM to suggest treatments based on clinical symptoms and severity.

**7. Domain**

Dermatology, particularly focusing on the diagnosis and treatment of psoriasis using AI and machine learning.

**8. Motivations**

Motivated by the need for accurate diagnostic tools and treatment recommendations in dermatology, the study aims to bridge gaps in CAD systems specifically for psoriasis by improving diagnostic accuracy and providing personalized treatment suggestions.

**9. Problems**

The complex symptom overlap in dermatological conditions often complicates traditional diagnostics, leading to errors. Additionally, there is a lack of tailored CAD frameworks for psoriasis classification and treatment guidance.

**10. Challenges**

Key challenges include:

* High visual similarity between psoriasis and other dermatological conditions.
* The need for a reliable system capable of classifying multiple types of psoriasis.
* Implementing effective treatment recommendations based on a diverse symptom profile.

**11. Related Work**

Previous studies used AI and ML for skin disease classification, but none specifically developed a comprehensive psoriasis CAD system or used an ensemble CNN model for this purpose. The paper also mentions existing studies using MCDM for dermatology treatment assessments.

**12. Techniques**

* **CNN Models**: Ensemble of ResNet50V2, ResNet101V2, and ResNet152V2.
* **MCDM Methods**: TRUST (MulTi-noRmalisation mUlti-Distance aSsessmenT) for treatment suggestion based on severity and symptom criteria.

**13. Datasets**

The image dataset used includes 2,100 images from DermNet NZ and DermNet databases, covering both psoriasis and non-psoriasis cases with variations such as eczema, dermatitis, and rosacea.

**14. Methodology**

* **Image Classification**: Binary classification for detecting psoriasis and multi-class classification for identifying specific types.
* **Treatment Recommendation**: TRUST-based MCDM, which considers patient symptoms to rank treatment options.

**15. Proposed Model**

An ensemble CNN model combining ResNet50V2, ResNet101V2, and ResNet152V2 networks, which performs both binary and multi-class classification for identifying psoriasis types.

**16. Results**

The proposed model achieved high accuracy, with 91.90% accuracy in binary classification and 93.29% in multi-class classification for different psoriasis types.

**17. Evaluation Metrics**

* **Classification Metrics**: Accuracy, recall, precision, and F1-score.
* **Loss Curves**: Test and training losses to assess model performance.

**18. Limitations**

The model showed lower classification accuracy for rarer types of psoriasis, such as erythrodermic psoriasis, due to limited data on these types.

**19. Future Work**

Suggested areas include expanding the dataset to improve classification of rarer psoriasis types and exploring additional deep learning architectures for more precise treatment suggestions.

**20. Genetic History (Relevant to Motivations and Problems sections)**

Psoriasis prevalence varies with genetic background, with approximately 36% of psoriasis patients having a family history, highlighting the role of genetics in the disease.

**21. Severity (Relevant to Challenges, Datasets, and Methodology sections)**

Severity is assessed using the PASI scoring system, which evaluates erythema, induration, and desquamation, along with other symptoms such as pain and fatigue. These indicators of severity are crucial for tailored treatment recommendations.

Paper 12------severity

**1. Authors**

Yi Li, Zhe Wu, Shuang Zhao, Xian Wu, Yehong Kuang, Yangtian Yan, Shen Ge, Kai Wang, Wei Fan, Xiang Chen, Yong Wang.

**2. Year**

Presented at the AAAI Conference on Artificial Intelligence in 2020.

**3. Venue**

The 34th AAAI Conference on Artificial Intelligence (AAAI-20).

**4. Title**

“PSENet: Psoriasis Severity Evaluation Network”

**5. Summary**

The paper introduces PSENet, a deep learning model designed to automatically assess psoriasis severity based on clinical images. It aims to address the limitations of the traditional Psoriasis Area and Severity Index (PASI) by providing a faster, more consistent method for evaluating the severity of psoriasis lesions.

**6. Contribution**

* Developed PSENet, a model that evaluates psoriasis severity from clinical images.
* Introduced a score refinement module and siamese network structure to enhance severity assessment.
* Enabled more convenient monitoring of psoriasis progression, particularly in settings with limited dermatologist access.

**7. Domain**

Dermatology, focusing on AI-based evaluation for psoriasis severity.

**8. Motivations**

Motivated by the need for a quicker, consistent, and less labor-intensive method to evaluate psoriasis severity, as PASI scoring is both time-consuming and prone to inconsistencies among dermatologists.

**9. Problems**

* The PASI scoring process requires manual scoring by dermatologists, which is both time-consuming and inconsistent.
* Lack of dermatologist availability in some regions adds to the need for automated solutions.

**10. Challenges**

* Limited availability of annotated data for training deep learning models.
* Variability in skin lesion appearance, requiring robust image analysis to accurately focus on affected areas.

**11. Related Work**

Existing studies have attempted to assess psoriasis severity by evaluating individual aspects like erythema, scaling, and induration but lack an end-to-end solution for full PASI scoring.

**12. Techniques**

* **Score Refine Module**: Focuses on lesion location and severity scoring within an image.
* **Siamese Network**: Aids in training with limited data by using paired images to model severity differences.

**13. Datasets**

A proprietary dataset was developed over years with the assistance of dermatologists, containing 5,205 images from 1,787 patients, with severity scores annotated by experienced dermatologists.

**14. Methodology**

* **Image Processing**: Each image is processed through a series of convolutional layers and score refinement modules to extract and assess features related to psoriasis severity.
* **Training**: Used siamese network structure to handle data insufficiency, training with paired images for severity comparisons.

**15. Proposed Model**

PSENet leverages the score refinement module and siamese structure, processing images at various granularities to output severity scores with higher consistency and efficiency than PASI scoring.

**16. Results**

PSENet achieved a mean absolute error (MAE) of 2.21 and a pair accuracy (PA) of 77.87% in severity assessment, outperforming baseline models.

**17. Evaluation Metrics**

* **Mean Absolute Error (MAE)**: Measures the deviation from dermatologist-assigned scores.
* **Pair Accuracy (PA)**: Assesses the consistency of PSENet’s severity rankings compared to dermatologist rankings.

**18. Limitation**

The model struggles with extreme severity cases due to limited data for such instances, resulting in higher errors when evaluating these cases.

**19. Future Work**

Suggestions include expanding the dataset to cover a broader range of severity cases and improving model accuracy for extreme cases.

**20. Severity of Psoriasis**

Severity in psoriasis is quantified using the PASI system, which assesses erythema, induration, desquamation, and lesion area. PSENet aims to automate this scoring by using clinical images, eliminating the need for manual PASI calculations, which are time-intensive and prone to inconsistencies.

Paper 40----Classification

**1. Authors**

Muhammad Sajid Rashid, Ghulam Gilanie, Saira Naveed, Sana Cheema, Muhammad Sajid.

**2. Year**

Published in 2024.

**3. Venue**

Published in *Signal, Image and Video Processing*.

**4. Title**

"Automated Detection and Classification of Psoriasis Types Using Deep Neural Networks from Dermatology Images."

**5. Summary**

This study proposes a lightweight Deep Neural Network (DNN)-based model for the automated detection and classification of five types of psoriasis from dermatology images. The model’s effectiveness is compared with pre-trained models such as GoogLeNet, InceptionV3, and VGG-19, showing superior classification performance. The research highlights the use of color space transformation (RGB to YCbCr) for better feature extraction.

**6. Contribution**

* Development of a 17-layer DNN for classifying five psoriasis types.
* High classification performance using YCbCr color space.
* Deployment-ready model for smartphone applications to aid dermatological assessment in remote areas.

**7. Domain**

Dermatology, specifically focused on AI-assisted diagnosis for psoriasis classification.

**8. Motivations**

To address the need for accessible diagnostic tools due to a shortage of dermatologists in certain regions, particularly rural areas in Pakistan. The study aims to simplify psoriasis classification using a model that can be deployed on mobile devices.

**9. Problems**

* Shortage of dermatologists and the need for a remote, automated classification solution.
* Existing models trained on single datasets may lack generalizability across diverse skin images.

**10. Challenges**

* Ensuring high classification accuracy across multiple psoriasis types.
* Reducing model complexity while maintaining performance to allow smartphone deployment.

**11. Related Work**

Previous studies in dermatological classification focused on specific skin conditions with limited multi-class categorization capabilities. Compared to existing models, this DNN framework is lightweight and able to classify multiple psoriasis types.

**12. Techniques**

* **DNN Layers**: A 17-layer network incorporating convolutional, pooling, and activation layers.
* **Color Space Transformation**: Conversion of RGB images to YCbCr for enhanced feature extraction and improved classification.

**13. Datasets**

Five publicly available datasets containing images of different psoriasis types:

* **Guttate Psoriasis**: 1,570 images
* **Flexural or Inverse Psoriasis**: 535 images
* **Pustular Psoriasis**: 5,790 images
* **Erythrodermic Psoriasis**: 2,290 images
* **Psoriatic Arthritis**: 1,830 images

**14. Methodology**

* **Image Preprocessing**: RGB images are converted to YCbCr for improved color differentiation in psoriasis classification.
* **Model Architecture**: Convolutional layers for feature extraction, RELU activation functions, and pooling layers to reduce dimensionality.
* **Classification and Training**: Model was trained with 80% of images, tested with 10%, and validated with 10%.

**15. Proposed Model**

A lightweight 17-layer DNN tailored to classify five types of psoriasis efficiently. It uses fewer parameters than conventional models, making it suitable for deployment on mobile devices.

**16. Results**

The model achieved high classification accuracy across all five psoriasis types, with accuracy up to 99.89% for psoriatic arthritis. The YCbCr color space was identified as optimal for classifying psoriasis.

**17. Evaluation Metrics**

* **Accuracy**: Measures overall classification performance.
* **Specificity**: Indicates the model’s ability to correctly identify non-affected cases.
* **Sensitivity**: Evaluates how well the model detects psoriasis cases.
* **AUC (Area Under Curve)**: Used to gauge model accuracy for each psoriasis type.

**18. Limitation**

The model is limited to five psoriasis types and requires noise-free images for optimal performance.

**19. Future Work**

Expanding the model to cover additional skin conditions, increasing dataset diversity, and improving robustness across image quality variations.

**20. Classification of Psoriasis**

This study specifically focuses on classifying five types of psoriasis using DNN, achieving high accuracy and sensitivity for each type, thereby addressing the need for precise categorization.

Paper 4-----severity

1. **Authors**

Yunzhao Xing, Sheng Zhong, Samuel L. Aronson, Francisco M. Rausa, Dan E. Webster, Michelle H. Crouthamel, Li Wang​

1. **Year**

Published in 2024

1. **Venue**

Published in *Digital Biomarkers*​

1. **Title**

"Deep Learning-Based Psoriasis Assessment: Harnessing Clinical Trial Imaging for Accurate Psoriasis Area Severity Index Prediction"​

1. **Summary**

This study develops a "one-step PASI" framework that uses a deep convolutional neural network (DCNN) to predict PASI scores by integrating body region images from a clinical trial setting. The model demonstrates high correlation with dermatologist-assessed PASI scores and can track treatment progress over time.

1. **Contribution**

The study presents a novel image-processing and deep learning workflow that integrates multiple tasks (body and lesion detection, severity classification) into a single model to improve PASI assessment accuracy.

1. **Domain**

Dermatology and machine learning, focusing on psoriasis severity classification.

1. **Motivation**

PASI is a complex, manual, and subjective scoring method for psoriasis severity, leading to inter-rater variability and inefficiencies. Automated PASI scoring could standardize assessments in clinical trials and practice

1. **Problems**

Traditional PASI scoring methods are labor-intensive, prone to variability, and can hinder consistent assessment in clinical settings.

1. **Challenges**

Integrating clinical trial data for machine learning without biases, ensuring model accuracy with limited data, and handling high variability in lesion appearance.

1. **Related Work**

Prior machine learning efforts have applied image-based models to PASI, but few have used interventional clinical trial data for longitudinal assessments

1. **Techniques**

Deep convolutional neural network (DCNN) models with ensemble learning, including 145 models evaluated through cross-validation​

1. **Datasets**

Images from the UltIMMa-2 clinical trial, consisting of 2,700 images from 60 psoriasis patients annotated by dermatologists

1. **Methodology**

The authors created composite images for three body regions, then trained an ensemble of DCNN models to predict PASI scores based on erythema, induration, and scaling. A "One-Step PASI" framework was implemented

1. **Proposed Model**

The "One-Step PASI" framework, an ensemble of DCNN models combining predictions from multiple models for accurate PASI scoring.

1. **Results**

The best model achieved a mean absolute error of 3.3 and a Pearson correlation coefficient of 0.90 with dermatologist PASI score

1. **Evaluation Metrics**

Mean absolute error (MAE), Lin’s concordance correlation coefficient (CCC), and Pearson correlation coefficient (PCC)​

1. **Limitation**

The dataset excluded head and neck images for privacy, limiting model accuracy for patients with psoriasis primarily in these areas. The study used a relatively small dataset.

1. **Future Work**

Further studies are suggested on larger datasets, including head and neck images and exploring smartphone-based image capture for remote PASI assessments

​

Paper 41-----genetic history

**Title**

* **Identification of Potential Biomarkers for Psoriasis by DNA Methylation and Gene Expression Datasets**

**Authors and Affiliations**

* **Authors**: Yong Liu, Shengnan Cui, Jiayi Sun, Xiaoning Yan, Dongran Han
* **Affiliations**:
  + School of Life Sciences, Beijing University of Chinese Medicine, Beijing, China
  + Department of Dermatology, Shaanxi Hospital of Chinese Medicine, Xi’an, China
  + Xiyuan Hospital, China Academy of Chinese Medical Sciences, Beijing, China

**Year of Publication**

* **2021**

**Venue**

* **Published in**: Frontiers in Genetics

**Summary**

* This study investigates the role of DNA methylation in regulating gene expression within psoriasis-affected skin. By integrating DNA methylation and gene expression data, the researchers identified potential biomarkers for psoriasis that help explain its pathogenesis. The study highlights hypermethylated and differentially expressed genes, suggesting that these genes and their associated signaling pathways could be targeted for therapeutic intervention.

**Contribution**

* **Primary Contribution**: The paper identifies hypermethylated genes and their relationship with immune infiltration in psoriatic tissues. It uses machine learning algorithms to pinpoint hub genes that may serve as biomarkers for psoriasis, providing a foundation for future studies on the disease's molecular mechanisms.

**Domain**

* **Field**: Dermatology, Epigenetics, and Machine Learning in disease biomarker discovery for autoimmune conditions.

**Motivations**

* Psoriasis is a complex, chronic inflammatory skin disease with genetic and environmental causes, yet its pathogenesis remains unclear. The authors aimed to identify epigenetic biomarkers, particularly DNA methylation sites, which could aid in understanding the disease mechanism and developing targeted therapies.

**Problems Addressed**

* **Lack of Specific Biomarkers**: Identifying reliable biomarkers for psoriasis remains challenging due to the complex interplay of genetic and environmental factors.
* **Insufficient Understanding of Pathogenesis**: There is a need to understand how genetic expression changes, influenced by DNA methylation, contribute to psoriasis progression.

**Challenges Identified**

* **High Data Heterogeneity**: Integrating datasets from different sources introduces variability, which the study addresses by applying statistical methods to minimize batch effects.
* **Complexity in Methylation-Gene Expression Relationships**: Identifying consistent hypermethylated and differentially expressed genes across samples is challenging due to the genetic diversity in psoriasis patients.

**Related Work**

* The paper discusses prior research linking DNA methylation to psoriasis and its role in autoimmune diseases. Studies have shown the importance of methylation in gene regulation and immune response, contributing to psoriasis pathology by affecting inflammatory pathways and immune cell activation.

**Techniques Used**

* **Differential Gene Expression Analysis**: The study utilized the Limma package in R to identify genes that are differentially expressed between psoriatic and normal skin.
* **DNA Methylation Analysis**: Differentially methylated CpG sites were identified using microarray data from the GEO database and normalized through the ComBat method to reduce batch effects.
* **Machine Learning Algorithms**: LASSO, Random Forest, and Support Vector Machine-Recursive Feature Elimination (SVM-RFE) algorithms were used to identify potential target genes by filtering and cross-validating genes relevant to psoriasis pathology.
* **Immune Infiltration Analysis**: CIBERSORT, a deconvolution algorithm, was used to analyze immune cell composition in psoriatic versus normal skin, providing insights into immune regulation in psoriasis.

**Datasets**

* **Source**: Gene Expression Omnibus (GEO) database
* **Specific Datasets Used**:
  + **GSE115797** and **GSE73894** for DNA methylation profiles from 48 and 82 samples, respectively.
  + **GSE30999** and **GSE41662** for gene expression profiles from 170 and 48 samples, respectively.

**Methodology**

* The study downloaded DNA methylation and gene expression profiles from GEO and processed the data using bioinformatics and statistical methods. The authors analyzed differential methylation, integrated datasets to address batch effects, and used machine learning to filter out key genes. Enrichment analyses for gene ontology (GO) and KEGG pathways were conducted to identify biological functions and pathways involved in psoriasis.

**Proposed Model**

* **Biomarker Identification Framework**: While not a predictive model, the study proposes a biomarker identification framework combining DNA methylation data with gene expression profiles, immune cell infiltration data, and machine learning to identify and validate key genetic biomarkers.

**Results**

* **Key Findings**: The study identified three core genes, **IRS1**, **RAI14**, and **ARHGEF10**, as potential biomarkers for psoriasis. These genes were hypermethylated and showed differential expression in psoriatic lesions, suggesting a role in immune regulation and metabolic pathways linked to psoriasis.
* **Enriched Pathways**: Hyper-downregulated genes were enriched in pathways related to glucose homeostasis and the AMPK signaling pathway, hinting at a connection between psoriasis and metabolic dysfunction.

**Evaluation Metrics**

* **Metrics Used**: Receiver Operating Characteristic (ROC) curve analysis was performed for selected biomarkers, with high Area Under the Curve (AUC) values indicating strong diagnostic potential. For example, **IRS1**, **RAI14**, and **ARHGEF10** showed AUC values around 0.95, demonstrating high reliability as biomarkers.

**Limitations**

* **Data Limitation**: The study was limited to publicly available data, which may lack diversity in patient demographics and environmental factors.
* **Heterogeneity and Potential Confounding Factors**: Although batch effects were reduced, factors like infections, diet, and obesity were not controlled, potentially affecting methylation and gene expression results.

**Future Work**

* **Functional Validation**: Future studies should validate the biological functions of identified target genes experimentally.
* **Broader Dataset Integration**: Including a wider variety of clinical and demographic data could improve biomarker reliability and generalizability.
* **Exploration of Additional Epigenetic Mechanisms**: Future research could explore histone modification and non-coding RNA interactions to further understand psoriasis pathogenesis.

Paper 42-----genetic history

**Title**

* **Identification of gene expression signatures for psoriasis classification using machine learning techniques**

**Authors and Affiliations**

* **Authors**: Nguyen Quoc Khanh Le, Duyen Thi Do, Trinh-Trung-Duong Nguyen, Ngan Thi Kim Nguyen, Truong Nguyen Khanh Hung, Nguyen Thi Thu Trang
* **Affiliations**: Taipei Medical University, Taiwan; Yuan Ze University, Taiwan; Cho Ray Hospital, Vietnam

**Year of Publication**

* **2021**

**Venue**

* **Published in**: Medicine in Omics

**Summary**

* This study aims to classify psoriasis using gene expression data to identify differentially expressed genes that could act as biomarkers for psoriasis. Using machine learning techniques like Random Forest and Support Vector Machines, the study identifies 35 genes, including FABP5, TGM1, and BCAR3, as potential biomarkers for distinguishing lesional from non-lesional skin. The study achieved high classification accuracy, suggesting that these gene signatures may be valuable for understanding and diagnosing psoriasis.

**Contribution**

* **Primary Contribution**: This study develops a machine learning-based classification model using gene expression data to identify gene signatures for psoriasis. The proposed model demonstrates high performance and identifies novel biomarkers for psoriasis classification, which could facilitate early and accurate diagnosis.

**Domain**

* **Field**: Dermatology, Machine Learning in Biomedical Data, Psoriasis Classification

**Motivations**

* Psoriasis diagnosis relies on visual examination, which is prone to inaccuracies. There is a need for accurate biomarkers to differentiate psoriasis lesions from normal tissue, enabling early intervention and improved treatment outcomes.

**Problems Addressed**

* **Lack of Reliable Biomarkers**: Current diagnostic methods lack specific biomarkers for psoriasis classification.
* **High Dimensionality of Genomic Data**: Gene expression datasets contain many features, posing challenges for effective classification without overfitting.

**Challenges Identified**

* **Data Complexity and Dimensionality**: Processing large sets of gene expression data requires dimensionality reduction to avoid overfitting.
* **Consistency Across Studies**: Variation in gene expression patterns makes it difficult to identify a universally applicable set of biomarkers.

**Related Work**

* The study references previous research that identified genes associated with psoriasis. Key comparisons are made with studies using neural networks and Support Vector Machines (SVMs) for psoriasis classification, with the Random Forest model shown to outperform these approaches in accuracy and stability.

**Techniques Used**

* **Machine Learning Models**: Four models were tested—k-Nearest Neighbors (kNN), Naïve Bayes, Random Forest, and Support Vector Machine (SVM)—to classify lesional and non-lesional skin samples.
* **Feature Selection with Random Forest**: Random Forest was used to rank genes by importance, reducing the feature set to 35 genes.
* **Cross-Validation**: 5-fold cross-validation was employed to validate model performance.
* **Gene Enrichment Analysis**: Conducted to identify biological processes associated with the selected gene signatures.

**Datasets**

* **Source**: Gene Expression Omnibus (GEO) database.
* **Specific Datasets Used**:
  + Training: GDS4602 and GDS4600 datasets (286 samples total).
  + Independent Testing: GDS3539 and GDS4891 datasets.

**Methodology**

* Gene expression data from the GEO database were preprocessed using log transformation. The Random Forest model was then applied to select the most informative genes. Performance was assessed with cross-validation and evaluated using independent datasets. The study also employed gene ontology (GO) analysis to understand the biological roles of identified genes.

**Proposed Model**

* **Classification Model**: The study proposes a Random Forest-based model for psoriasis classification, using 35 key gene expression signatures. This model achieved high accuracy and provides a framework for future psoriasis biomarker studies.

**Results**

* **Performance**: The Random Forest model achieved 98.3% accuracy, 98.6% recall, and 98% precision on training data, and 96.7% and 100% accuracy on two independent test datasets, respectively.
* **Key Biomarkers**: Genes FABP5, TGM1, and BCAR3 were identified as significant in differentiating psoriasis lesions from normal skin.

**Evaluation Metrics**

* **Metrics Used**: Accuracy, recall, and precision were calculated. Performance was validated with cross-validation and independent datasets, with metrics indicating high accuracy and reliability.

**Limitations**

* **Dataset Limitations**: The study relies on a limited set of publicly available data, which may not capture all variations in gene expression patterns across diverse populations.
* **Feature Dimensionality**: Reducing the feature set to a small number of genes may overlook potentially relevant biomarkers.

**Future Work**

* **Enhanced Biomarker Panels**: Future research could expand on these findings by including additional biomarkers and verifying their relevance in clinical settings.
* **Hybrid Optimization Models**: Suggests combining machine learning models with optimization techniques to further improve classification performance.

Paper 44