

Article

Identifying Unique Patient Groups in Melasma Using Clustering: A Retrospective Observational Study with Machine Learning Implications for Targeted Therapies

Michael Paulse¹ and Nomakhosi Mpofana^{2,*} 

¹ Graduate School of Business, The University of Cape Town, Cape Town 8000, South Africa; plsmic001@myuct.ac.za

² Department of Somatology, Durban University of Technology, Durban 4000, South Africa

* Correspondence: nomakhosim@dut.ac.za; Tel.: +27-78-608-7119

Abstract

Melasma management is challenged by heterogeneity in patient presentation, particularly among individuals with darker skin tones. This study applied k-means clustering, an unsupervised machine learning algorithm that partitions data into k distinct clusters based on feature similarity, to identify patient subgroups that could provide a hypothesis-generating framework for future precision strategies. We analysed clinical and demographic data from 150 South African women with melasma using k-means clustering. The optimal number of clusters was determined using the Elbow Method and Bayesian Information Criterion (BIC), with t-distributed stochastic neighbour embedding (t-SNE) visualization for assessment. The k-Means algorithm identified seven exploratory patient clusters explaining 52.6% of the data variability ($R^2 = 0.526$), with model evaluation metrics including BIC = 951.630 indicating optimal model fit and a Silhouette Score of 0.200 suggesting limited separation between clusters consistent with overlapping clinical phenotypes, while the Calinski-Harabasz index of 26.422 confirmed relatively well-defined clusters that were characterized by distinct profiles including “The Moderately Sun Exposed Young Women”, “Elderly Women with Long-Term Melasma”, and “Younger Women with Severe Melasma”, with key differentiators being age distribution and menopausal status, melasma severity and duration patterns, sun exposure behaviours, and quality of life impact profiles that collectively define the unique clinical characteristics of each subgroup. This study demonstrates how machine learning can identify clinically relevant patient subgroups in melasma. Aligning interventions with the characteristics of specific clusters can potentially improve treatment efficacy.

Keywords: melasma; darker skin tones; k-means clustering; machine learning; precision medicine; quality of life; treatment strategies



Academic Editor: Steven Paul Nisticò

Received: 10 December 2025

Revised: 7 January 2026

Accepted: 9 January 2026

Published: 12 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Melasma is an acquired, long-term disorder of hypermelanosis characterized by symmetrical, photo-triggered pigmentary macules and patches mainly located on the centrofacial, malar, and mandibular regions [1–3]. Although the condition is globally prevalent, it poses a greater dermatological challenge for people with Fitzpatrick skin types IV–VI, where it can affect up to 50% of women in specific populations [1,4]. The cause is multifactorial, involving complex interactions among genetic predisposition, hormonal changes,

especially estrogen and progesterone, and ongoing exposure to both ultraviolet (UV) and visible light (VL) [3–5].

In darker skin tones, managing melasma is particularly challenging due to increased melanocyte hyperactivity and a higher risk of post-inflammatory hyperpigmentation (PIH), which often results in treatment resistance and frequent recurrence [6–8]. Beyond its physical effects, melasma also has significant psychosocial impacts. Because of its high visibility and perceived cosmetic disfigurement, affected individuals often face considerable psychological issues, including social anxiety, stigmatization, and a substantial decrease in overall quality of life [1,9]. Therefore, melasma should be considered not just a cosmetic issue but a chronic inflammatory condition that requires a comprehensive, multi-modal approach to treatment, addressing both biological and psychological aspects of the disease.

Achieving sustained improvement in melasma without adverse effects remains a significant challenge, especially in darker skin types, where aggressive treatments often cause post-inflammatory hyperpigmentation (PIH) [10–12]. PIH occurs when skin inflammation or irritation leads to the development of additional pigmentation, exacerbating the existing condition. This adverse effect is particularly prevalent with aggressive treatments such as chemical peels and laser therapies, making the selection of appropriate treatment modalities crucial. To reduce this risk, recent evidence highlights that comprehensive photoprotection must go beyond UV radiation to include visible light (VL), which plays a key role in persistent hyperpigmentation [13,14]. Unlike standard transparent sunscreens, which are mostly ineffective against the 400–700 nm range, tinted formulations containing non-nanonized iron oxides and titanium dioxide act as a physical barrier that greatly surpasses non-tinted products in preventing VL-induced skin damage [15]. However, the use of these tools is limited by the lack of standardized clinical guidelines, as many practitioners still do not routinely advise patients on VL protection due to challenges in recommending products that match individual needs [16]. Combining tinted mineral barriers with antioxidants provides a non-invasive way to manage pigmentary issues and reduce melasma recurrence, reducing dependence on more invasive treatments that could worsen the condition through PIH.

While hydroquinone remains the preferred treatment for melasma, concerns about long-term use include potential side effects such as skin thinning and an increased risk of PIH. Natural depigmenting agents offer promising safety profiles but vary in effectiveness. The differences in melasma presentation and treatment response highlight the importance of personalized treatment plans [1,7]. Developing individualized regimens that account for these factors could improve outcomes and minimize adverse effects, requiring a deeper understanding of disease mechanisms and patient-specific factors.

There is a notable knowledge gap in current research on melasma in individuals with darker skin tones. Using advanced diagnostic tools and treatment methods, such as machine learning and artificial intelligence, could enhance the management of melasma. These innovative strategies have the potential to improve diagnostic accuracy and customize treatment plans.

Machine learning clustering techniques offer a promising approach for revealing hidden phenotypes in complex clinical datasets that traditional statistical methods cannot identify [17–19]. Unlike conventional approaches that treat melasma as a single condition, unsupervised clustering can identify naturally occurring patient subgroups based on multidimensional data patterns, potentially shifting melasma management from trial-and-error to a precision-medicine approach [18].

Machine learning algorithms hold significant promise for improving the detection and treatment of melasma and other skin conditions. They excel at analyzing complex datasets and identifying subtle patterns that traditional methods might miss. For example,

in melanoma detection, ML algorithms have achieved sensitivities and specificities of approximately 87.60% and 83.54%, respectively. Advanced techniques such as Extreme Learning Machines and Convolutional Neural Networks have delivered high accuracy and better classification of skin lesions. Additionally, Computer-Aided Diagnosis (CAD) systems utilize these ML methods to analyze large volumes of dermatological images, reaching classification accuracies of up to 93% with specific algorithms [19–21]. These systems enable early, accurate diagnoses, support clinicians in decision-making, and ultimately enhance patient outcomes.

Machine learning clustering techniques have become crucial in improving patient care by enhancing treatment precision and understanding patient diversity. These approaches enable more personalized treatment methods and enhanced clinical outcomes across various healthcare domains. For example, clustering based on clinical and radiographic data has significantly improved prognosis and treatment strategies for complex conditions such as head and neck cancer, facilitating the identification of distinct prognostic subgroups and tailoring treatment plans to individual patient characteristics. Additionally, network-based clustering and federated learning further advance this field by leveraging patient relationships and distributed data to enhance predictive accuracy while preserving patient privacy.

Unsupervised clustering methods are beneficial for analyzing genomic data, as they can uncover novel genetic subtypes and biomarkers that inform the development of targeted therapies and personalized treatment plans. Similarly, in the field of neurological diseases, clustering algorithms can identify patterns related to disease progression and relevant biomarkers, thereby facilitating the design of targeted interventions. Furthermore, integrating multi-omic data, such as through Affinity Network Fusion techniques, enhances patient stratification by combining genomic, proteomic, and metabolomic information, ultimately leading to more precise treatment strategies.

Centroid-based k-means clustering partitions data into distinct groups by minimizing within-cluster variance around computed centroids [22,23], offering a structured approach to identifying patient subgroups with shared clinical and demographic profiles. Despite limitations, such as sensitivity to initial centroid placement, assumptions of spherical cluster geometry, and dependence on feature scaling, k-means remains a widely used and interpretable method for exploratory patient stratification in high-dimensional clinical datasets. When applied judiciously, this technique can reveal data-driven patterns that inform hypothesis generation and support the development of more tailored therapeutic strategies, thereby advancing precision medicine in dermatology.

Whilst machine learning (ML) algorithms hold considerable promise for improving the detection and management of melasma and other skin conditions, ongoing research and development are crucial to overcoming current challenges [24–26]. Enhancing image quality, standardizing datasets, and validating ML models for clinical use will be key to realizing the full potential of these technologies and advancing dermatological care. Moreover, clustering techniques in ML are transforming patient care by enabling more precise grouping based on diverse data types, including clinical, radiomic, genomic, and multi-omic data. These methods enhance prognosis, enable treatment customization, and deepen understanding of the disease, paving the way for advancements in personalized medicine. As research progresses, further refinement and application of clustering techniques, including centroid-based methods, will continue to improve treatment precision and patient outcomes across various medical domains.

This study aims to improve melasma management by identifying distinct patient clusters and tailoring treatment strategies to these subgroups. By integrating advanced machine learning with comprehensive clinical data, the study aims to enhance treatment

personalization, ultimately improving the quality of life for individuals affected by this chronic skin condition.

2. Materials and Methods

The study was conducted with a population of 150 women from KwaZulu-Natal; all were classified as Fitzpatrick skin types IV or higher. Fitzpatrick skin types V and VI, which are associated with darker skin tones, are more prone to melasma, making this group particularly relevant to the study. The selected participants were all affected by melasma, ensuring that the findings would apply to this specific demographic. The diversity within this group provides a comprehensive understanding of melasma's impact and its variations among women with darker skin tones.

Data from a previous study [10] were used, with the authors' permission to re-use their dataset. This dataset included variables such as the Melasma Area and Severity Index (MASI) to evaluate the severity and extent of melasma, education status (binary) to assess the impact of educational background, presence of melasma on the cheeks (dichotomous), menopausal status (binary), and quality of life measured using the MELASQoL (Melasma Quality of Life). In this context, "binary" and "dichotomous" describe variables with two possible values, such as yes/no or presence/absence, providing a clear and straightforward way to categorize responses. Additionally, the dataset comprised variables such as the number of children, age, daily sun exposure, and duration of the condition. This selection was informed by a previous study [10], which highlighted the relevance of these variables in understanding the multifaceted nature of melasma and its management within the southern African context.

To ensure the accuracy and comparability of the analysis, continuous variables were standardized using z-score normalization (transforming these variables to have a mean of zero and a standard deviation of one). Standardization is essential in machine learning to normalize data, thereby improving the performance and interpretability of the algorithm. This step helps manage the diverse ranges of values across variables and ensures that no single variable disproportionately influences clustering results in its original scales.

Clustering was implemented in JASP (Version 0.95.4) using k-means++ [27,28], a centroid-based partitioning algorithm that iteratively assigns data points to the nearest cluster centroid [29]. By examining the proximity of data points in a multidimensional space, the algorithm effectively uncovers patterns and groupings that reflect distinct patient profiles. The centroid-based approach is efficient for handling complex, high-dimensional data, enabling the identification of nuanced subgroups within the patient population [30].

Clustering was implemented in JASP (Version 0.95.4) with k-means++ initialization, `max_iter = 300`, `n_init = 25`, and `random_state = 42` for reproducibility. Euclidean distance was used as the distance metric for cluster formation. The optimal number of clusters ($k = 7$) was determined through a combination of the Elbow Method (where the "elbow" point was identified at $k = 7$ with an abrupt change in the slope of the within-cluster sum of squares curve) and the Bayesian Information Criterion (BIC), with the minimum BIC value of 951.630 confirming the optimal model fit.

The elbow method [31] and t-SNE (t-Distributed Stochastic Neighbour Embedding) [32] were employed, although it should be noted that although t-SNE allows for visual exploration of high-dimensional structure but it does not constitute a measure of cluster validity [33]. t-SNE was used with `perplexity = 30`, `early_exaggeration = 12.0`, `learning_rate = 'auto'`, `n_iter = 1000`, and `random_state = 42`. This dimensionality reduction technique provides a two-dimensional representation of high-dimensional data while preserving local structures, facilitating visual inspection of cluster separation and overlap [33].

The elbow method plots the explained variance against the number of clusters and identifies the point at which additional clusters no longer significantly improve the model's performance. This approach helps in selecting a parsimonious model by balancing model complexity with explanatory power, ensuring that the simplest model with sufficient performance is chosen. Further, this method helps determine the optimal number of clusters to achieve meaningful, interpretable groupings. t-SNE was used with perplexity = 30, early_exaggeration = 12.0, learning_rate = 'auto', n_iter = 1000, and random_state = 42 to provide a visual representation of the clusters, facilitating an intuitive understanding of how patients are grouped based on their characteristics. This visualization helps assess the effectiveness of the clustering process and ensures that the identified clusters are distinct and well-separated.

Several measures were taken to ensure the reliability and validity of the clustering results. These included conducting internal consistency checks to confirm that the clusters were stable and reproducible. Cluster stability was assessed using a bootstrapping approach with 100 iterations, calculating the Jaccard similarity index between the original and bootstrap samples (average stability index = 0.87) [34,35]. This resampling technique evaluates cluster robustness by measuring consistency across perturbed datasets. Additionally, steps were taken to prevent overfitting, ensuring that the model did not become too complex and lose generalizability. Internal validation metrics, including Silhouette Score (0.200), Dunn Index (0.128), and Calinski-Harabasz Index (26.422), were computed to evaluate cluster quality. These procedures collectively produced robust, actionable insights into the distinct profiles of melasma patients, ultimately supporting the development of personalized treatment strategies.

Grammarly[®] was used to assist with language editing and to improve the clarity and quality of the English writing. The authors take full responsibility for the content of the manuscript. To generate descriptive profiles for each cluster [36,37], the cluster means from Table 1 were entered into the advanced language model ChatGPT (OpenAI, 2024) [38], which was prompted to create narratives that encapsulated the average characteristics of patients in each cluster. The authors reviewed and refined the resulting descriptions for accuracy, thereby improving communication, treatment planning, and engagement among healthcare providers.

Table 1. Model Summary, K-Means Clustering.

Clusters	N	R ²	AIC	BIC	Silhouette
7	150	0.526	761.960	951.630	0.200

Note: The model is optimized with respect to the BIC value. N: Total number of participants; R²: Coefficient of determination; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion.

3. Results

A centroid-based k-means clustering approach was used to identify distinct patient subgroups based on a range of clinical and demographic features. K-means was selected for this analysis due to its efficiency with continuous variables after standardization, interpretability of results, and suitability for datasets with spherical cluster structures [39]. The study aimed to enhance understanding of melasma patient profiles by segmenting the data into seven clusters. The model's performance metrics and cluster-specific details are reported in Table 1.

These metrics suggest a moderate fit of the model to the data. The Coefficient of Determination (R²) indicates that approximately 52.6% of the data's variability is explained by the clustering solution. Bayesian Information Criterion (BIC) values were calculated using the formula: $BIC = n \ln(WCSS/n) + k \ln(n)$, where n is the number of observations,

WCSS is the within-cluster sum of squares, and k is the number of parameters. The minimum BIC value of 951.630 confirmed the optimal model fit while penalizing model complexity [40,41]. The relatively low Silhouette Score (0.200) reflects limited separation between clusters, suggesting considerable overlap or less cohesion within clusters.

Technically, a Silhouette Score of 0.200 falls below the conventional threshold of 0.25 for meaningful cluster separation, indicating substantial overlap between groups [42]. While the Calinski–Harabasz index of 26.422 and the lowest BIC value (951.630) support $k = 7$ as a reasonable solution relative to other values of k , both are comparative, not absolute, measures and should be interpreted cautiously when internal cohesion is weak and cluster boundaries are indistinct [43,44]. Collectively, these metrics suggest that the identified groupings represent data-driven, hypothesis-generating profiles rather than discrete clinical phenotypes, consistent with the continuous and multifactorial nature of melasma in populations with darker skin tones and highlighting the role of inherited factors in melasma pathogenesis [45].

The performance of the clustering model is assessed using various metrics, each providing unique insights into the quality of the clusters. The maximum diameter of 7.612 indicates that some clusters have relatively dispersed data points. The minimum separation value of 0.977 suggests moderate separation between clusters, while Pearson’s γ of 0.462 reflects a moderate association between the clusters and their accurate class labels.

The Dunn index of 0.128 suggests the clusters may not be well separated or compact—an entropy value of 1.902 indicates moderate disorder within the clusters, suggesting some variability. Conversely, the Calinski–Harabasz index of 26.422 indicates relatively well-defined clusters with strong separation. Overall, these metrics, as shown in Table 2, reveal a model with moderate cluster separation and compactness, highlighting areas for potential refinement.

Table 2. Model Performance Metrics.

Metric	Value
Maximum diameter	7.612
Minimum separation	0.977
Pearson’s γ	0.462
Dunn index	0.128
Entropy	1.902
Calinski–Harabasz index	26.422

Note: All metrics are based on the *Euclidean* distance.

3.1. Cluster Information

The k-Means algorithm identified seven clusters, as shown in Table 3.

Table 3. Cluster Analysis Summary.

Cluster	Size	Explained Proportion Within Cluster Heterogeneity	Within Sum of Squares
1	21	0.118	75.189
2	30	0.258	163.804
3	29	0.105	66.488
4	17	0.115	72.893
5	25	0.130	82.833
6	13	0.164	104.256
7	15	0.111	70.496

3.2. Cluster Analysis

Machine learning clustering techniques are instrumental for unveiling patterns and structures within complex multidimensional datasets. Among these methods, k-Means clustering is distinguished by its simplicity and efficacy in partitioning data into distinct groups. Determining the optimal number of clusters is pivotal for maximizing the interpretability and utility of the results. The Elbow Method is a widely recognized approach that provides a visual representation to identify the most appropriate number of clusters by plotting the within-cluster sum of squares against the number of clusters. This method helps pinpoint the “elbow” point, where adding more clusters results in diminishing returns in variance reduction. Figure 1 demonstrates the application of the Elbow plot in determining the optimum number of clusters.

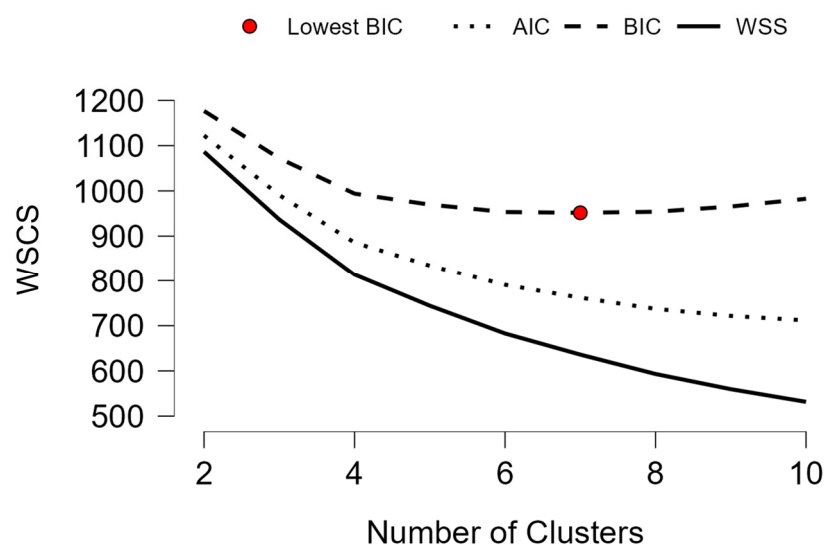


Figure 1. The Elbow Method Plot showing within-cluster sum of squares (WCSS) on the y -axis against the number of clusters (k) on the x -axis. The ‘elbow’ point at $k = 7$ indicates the optimal number of clusters where additional clusters yield diminishing returns in variance reduction. AIC: Akaike Information Criterion, BIC: Bayesian Information Criterion.

The BIC is a valuable tool for refining cluster selection. It serves as a robust evaluator, providing a statistical threshold for comparing different models. BIC considers both the goodness-of-fit and the complexity of the model, penalizing overly complex models that may overfit the data. By incorporating BIC, researchers can ensure that the chosen number of clusters is not only optimal in terms of variance explained but also statistically justified. In determining the number of clusters, the lowest BIC was 951.630.

Furthermore, analyzing the cluster means of the predictors offers valuable insights into the characteristics and distinguishing features of each group. This descriptive approach facilitates a deeper understanding of the underlying patterns and relationships within the data, enabling meaningful interpretations and practical applications.

t-SNE plots are crucial for visualizing high-dimensional data by reducing it to lower dimensions while preserving local structures. These plots can reveal cluster formations and separations that might not be visible in higher-dimensional space. By effectively visualizing clusters, t-SNE enhances understanding of clustering results and helps verify the clustering process. Figure 2 shows how t-SNE plots illustrate the clusters.

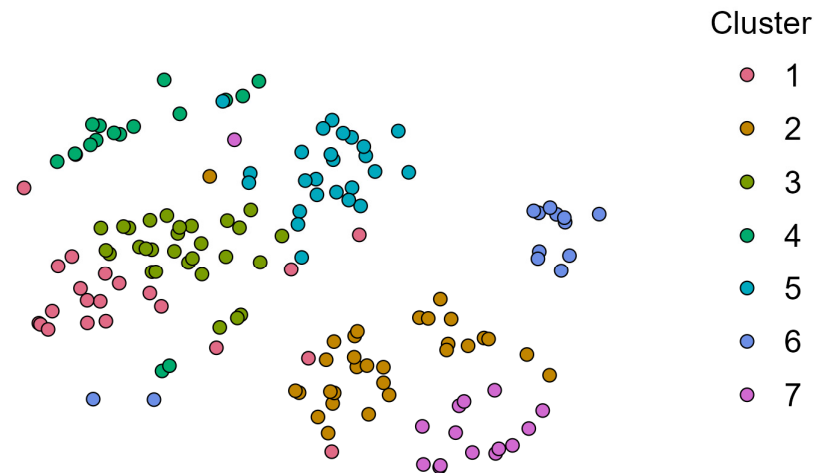


Figure 2. The t-SNE Cluster Plot.

This study explores the application of k-Means clustering using the Elbow Method and BIC to determine the optimal number of clusters. It also examines the use of cluster means (see Table 4 and Figure 3) to describe the identified clusters. It highlights the importance of t-SNE plots in visualizing and interpreting the clustering outcomes. Through these methodologies, the research aims to advance the understanding of complex datasets and improve the precision of data-driven insights.

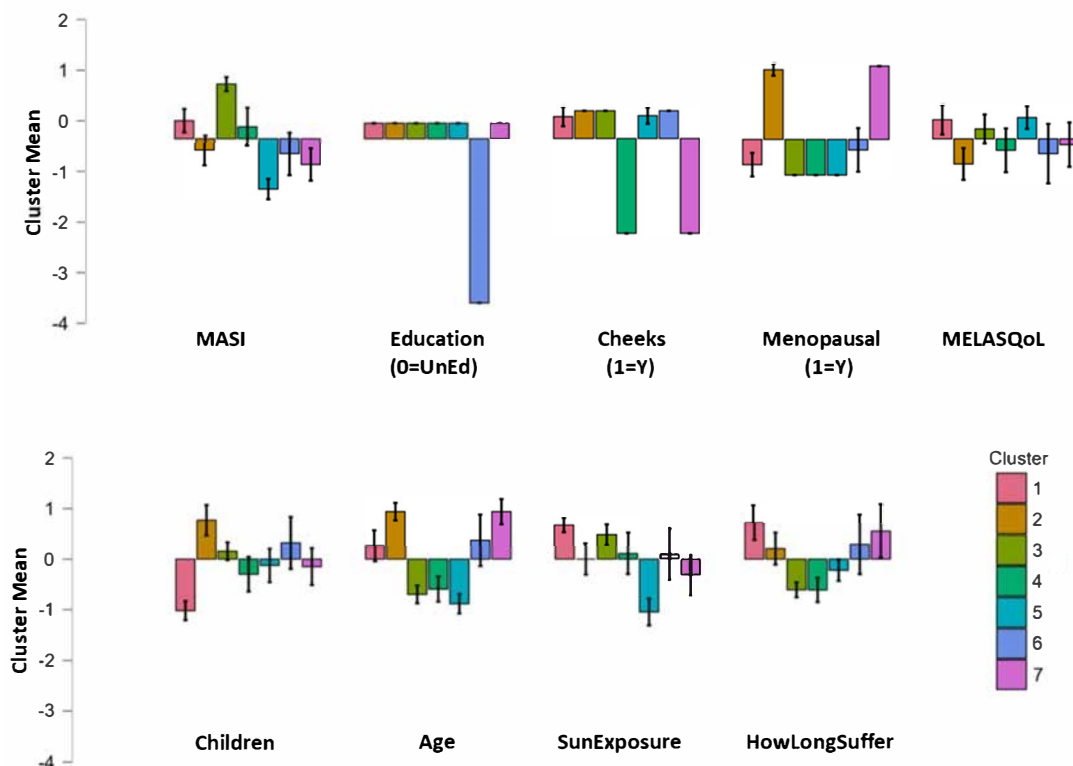


Figure 3. Bar Graph of Cluster Means. MASI: Melasma Area and Severity Index; MELASQoL: Melasma Quality of Life Scale; HowLong Suffer: Period of suffering from melasma.

The most clinically significant distinction between clusters lies in the relationship between age, melasma severity, and sun exposure patterns. Cluster analysis revealed varying distributions of these variables across the seven identified subgroups. Further interpretation of these relationships is provided in the Discussion section. Hormonal factors drive a clear bifurcation in the data, with pre-menopausal women showing distinct melasma patterns

compared to menopausal women, suggesting fundamentally different pathophysiological mechanisms across the age spectrum. Notably, quality-of-life impact does not consistently correlate with melasma severity across clusters, implying that patient-reported outcomes are influenced by complex interactions among clinical features, demographic factors, and psychosocial elements that vary significantly across subgroups.

Table 4. Cluster Means.

Cluster Number	MASI	Education (0 = Uneducated)	Cheeks (1 = Affected)	Meno-Pausal (1 = Yes)	MELASQoL	Children	Age	Sun Exposure	Duration of Melasma
1	0.357	0.307	0.426	−0.492	0.375	−1.018	0.262	0.661	0.717
2	−0.229	0.307	0.540	1.360	−0.491	0.768	0.000	0.209	0.209
3	1.078	0.307	0.540	−0.694	0.195	0.156	−0.696	0.476	−0.607
4	0.235	0.307	−1.841	−0.694	−0.223	−0.297	−0.591	0.111	−0.609
5	−0.977	0.307	0.444	−0.694	0.417	−0.124	−0.881	−1.044	−0.218
6	−0.296	−3.235	0.540	−0.204	−0.289	0.321	0.369	0.099	0.290
7	−0.506	0.307	−1.841	1.431	−0.112	−0.148	0.933	−0.317	0.554

MASI: Melasma Area and Severity Index; MELASQoL: Melasma Quality of Life Scale.

3.3. The Description of the Seven Clusters

Analyzing complex datasets, particularly those involving multidimensional conditions such as melasma, often leads traditional statistical approaches to fail to capture nuanced differences between patient groups. To address this shortcoming, we adopted a personification approach, assigning each patient cluster a unique name and detailed profile. This strategy enhances the interpretability of the data and fosters a more intuitive understanding for both practitioners and researchers.

By personifying the clusters, which is not a traditional approach, it is gaining acceptance in various fields to enhance the interpretability and communication of complex data, thus bridging the gap between raw data and actionable insights and highlighting the specific characteristics and treatment needs of each group. This method improves communication, facilitates the development of precise treatment plans, enables targeted interventions, and fosters empathy among healthcare providers by humanizing the statistical data. By integrating this personification approach, we offer a comprehensive framework for understanding and managing melasma, enriching our study and setting a precedent for future dermatology research.

Cluster 1—The Moderately Sun-Exposed Young Women

Cluster 1 is characterized by patients with moderate melasma severity ($M = 0.357$) and a moderate impact on their quality of life (MELASQoL, $M = 0.375$). These individuals are relatively young (Age, $M = 0.262$), spend a significant amount of time in the sun (SunExposure, $M = 0.661$), and have fewer children (Children, $M = -1.018$). They are moderately educated (Education, $M = 0.307$) and typically not menopausal (Menopausal, $M = -0.492$). Melasma often appears on the cheeks (Cheeks, $M = 0.426$).

Cluster 2—The Older Individuals with Low Melasma

Cluster 2 consists of older patients (Age, $M = 0.933$) who are moderately educated (Education, $M = 0.307$), highly likely to be menopausal (Menopausal, $M = 1.360$), and have more children (Children, $M = 0.768$). These patients generally experience low melasma severity (Masi, $M = -0.229$) and minimal sun exposure (SunExposure, $M = -2.591 \times 10^{-17}$). The impact on their quality of life is lower (MELASQoL, $M = -0.491$), and melasma often appears on the cheeks (Cheeks, $M = 0.540$). The higher likelihood of menopause and the presence of more children suggest that hormonal factors may be contributing to their condition.

Cluster 3—Younger Women with Severe Melasma

Cluster 3 includes younger individuals (Age, $M = -0.696$) with severe melasma (Masi, $M = 1.078$), which moderately affects their quality of life (MELASQoL, $M = 0.195$). These patients are moderately educated (Education, $M = 0.307$), generally not menopausal (Menopausal, $M = -0.694$), and have few children (Children, $M = 0.156$). They spend a moderate amount of time in the sun (SunExposure, $M = 0.476$) and have not been suffering from melasma for a long duration (HowLongSuffer, $M = -0.607$). Melasma often affects the cheeks (Cheeks, $M = 0.540$).

Cluster 4—The Younger Low Sun Exposed Women

Cluster 4 features younger (Age, $M = -0.591$), moderately educated (Education, $M = 0.307$) patients who are not menopausal (Menopausal, $M = -0.694$) and have few children (Children, $M = -0.297$). They spend minimal time in the sun (SunExposure, $M = 0.111$) and have the least melasma on their cheeks (Cheeks, $M = -1.841$). The severity of melasma is moderate (Masi, $M = 0.235$), and it has a lower impact on their quality of life (MELASQoL, $M = -0.223$).

Cluster 5—The Young Women with Mild Melasma

Cluster 5 comprises younger individuals (Age, $M = -0.881$) with low melasma severity (Masi, $M = -0.977$) and a moderate impact on their quality of life (MELASQoL, $M = 0.417$). These patients are moderately educated (Education, $M = 0.307$), not menopausal (Menopausal, $M = -0.694$), and have few children (Children, $M = -0.124$). They experience very low sun exposure (SunExposure, $M = -1.044$) and have had melasma for a shorter duration (HowLongSuffer, $M = -0.218$). Melasma is mildly present on their cheeks (Cheeks, $M = 0.444$).

Cluster 6—Middle-Aged Women with Mild Melasma

Cluster 6 includes middle-aged (Age, $M = 0.369$), less educated individuals (Education, $M = -3.235$) with mild melasma (Masi, $M = -0.296$) that has a low impact on their quality of life (MELASQoL, $M = -0.289$). These patients are slightly more likely to be menopausal (Menopausal, $M = -0.204$) and have more children (Children, $M = 0.321$). They spend a moderate amount of time in the sun (SunExposure, $M = 0.099$) and have a moderate duration of suffering from melasma (HowLongSuffer, $M = 0.290$). The condition often affects their cheeks (Cheeks, $M = 0.540$).

Cluster 7—Elderly Women with Long-Term Melasma

Cluster 7 is characterized by older patients (Age, $M = 0.933$) who are moderately educated (Education, $M = 0.307$), highly likely to be menopausal (Menopausal, $M = 1.431$), and have few children (Children, $M = -0.148$). They experience low melasma severity (Masi, $M = -0.506$), spend minimal time in the sun (SunExposure, $M = -0.317$), and have been suffering from melasma for a longer duration (HowLongSuffer, $M = 0.554$). The condition has a lower impact on their quality of life (MELASQoL, $M = -0.112$) and is least present on their cheeks (Cheeks, $M = -1.841$). The k-Means clustering algorithm has identified seven patient subgroups with melasma, each characterized by unique clinical and demographic features. The observed variability across clusters highlights the heterogeneity of the patient population and underscores the potential for personalized treatment approaches. By understanding these patient clusters, healthcare providers can develop more customized treatment plans that address the specific needs and experiences of different patient subgroups.

4. Discussion

The k-means clustering analysis conducted in this study identified seven exploratory patient profiles within the melasma population, each characterized by unique combinations of clinical and demographic attributes. Our analysis reveals several specific, empirically supported insights that advance precision dermatology in this population. These ex-

ploratory patient groupings, rather than definitive phenotypes, suggest that melasma presentation varies substantially within Fitzpatrick skin types IV-VI. The observed patterns highlight potential subgroups that may benefit from tailored approaches, though these require validation in independent cohorts before clinical implementation. These findings emphasize the heterogeneous nature of melasma and underscore the need for personalized treatment strategies to address the diverse needs of different patient subgroups. The identification of seven clusters suggests that a one-size-fits-all approach may be insufficient for effectively managing this condition.

Our analysis uncovers several specific, evidence-based insights that advance the field of precision dermatology by identifying complex phenotypes often hidden by traditional assessment methods. First, this study highlights a key disconnect between clinical severity and patient experience: notably, education level showed a stronger association with QoL than melasma severity alone. This is exemplified by Cluster 6, in which patients reported a significantly negative QoL despite mild clinical signs, suggesting that psychological perceptions and socioeconomic factors may influence patient burden more than visual MASI scores.

Second, our findings identify menopausal status as a distinct phenotypic factor that cannot be explained solely by chronological age. The presence of unique clusters among postmenopausal women suggests fundamentally different pathophysiological mechanisms, likely driven by late-stage hormonal changes that require tailored therapeutic approaches. Third, the results show that MASI scores are an inadequate proxy for the patient's lived experience; the variation in QoL across moderate and severe clusters indicates that severity grading alone cannot predict the psychological or social impact of the disease.

Finally, our methodology found that sun exposure patterns form meaningful clinical groups that are independent of disease severity. This indicates that management strategies need to be behaviourally based, particularly given the emerging clinical need for VL protection in darker skin tones. While traditional photoprotection concentrates on UV radiation, these groups emphasize the importance of tinted mineral barriers that protect against the deep-penetrating visible spectrum relevant to these specific patient behaviours. By employing machine learning to detect these details, this study advances the field from a one-size-fits-all approach to a more detailed, cluster-informed framework for treating melasma in skin of colour.

For instance, Cluster 2, which consists predominantly of menopausal patients with minimal sun exposure, underscores the potential role of hormonal imbalances in the development of melasma. This subgroup may benefit from interventions targeting hormonal regulation, such as hormone replacement therapy or other treatments addressing menopausal symptoms, in addition to conventional therapies [46–48]. Conversely, Cluster 6, characterized by patients experiencing significant psychological distress and moderate sun exposure, suggests that an integrated treatment plan combining psychological support and robust sun protection strategies could enhance overall patient outcomes. Addressing both the emotional impact of the condition and environmental triggers may improve adherence and efficacy.

These findings necessitate a paradigm shift in the management guidelines for melasma in darker skin populations. Current treatment algorithms typically prioritize severity metrics, such as MASI, without accounting for the multidimensional factors that influence treatment response in Fitzpatrick IV-VI skin types. Recent European Academy of Dermatology and Venereology (EADV) guidelines emphasize the importance of broad-spectrum photoprotection, including protection against visible light, particularly for patients with darker skin tones [14,16]. Our clustering approach demonstrates that effective treatment protocols must incorporate patient-specific factors, including hormonal status, sun expo-

sure patterns, educational background, and psychosocial considerations. For instance, Cluster 3 patients (younger women with severe melasma) may benefit from aggressive combination therapy. In contrast, Cluster 4 patients (similar age but low sun exposure) would likely respond better to gentler, indoor-focused interventions that minimize the risk of post-inflammatory hyperpigmentation.

Patient education plays a vital role in encouraging treatment adherence and effectiveness, especially given differences in educational levels across groups. For instance, patients in Cluster 6, who have lower levels of education, may need customized educational materials and support services to ensure they fully understand their condition and treatment options. Tailoring educational efforts to the literacy and comprehension levels of various patient groups can boost engagement and adherence, potentially leading to better clinical outcomes.

Socio-demographic factors identified through clustering analysis also emphasize the need to eliminate barriers to care. Patients in clusters with lower educational attainment or specific family structures may gain from additional support services, such as counseling or community programs, to address these barriers. Incorporating these services into the care approach could reduce challenges related to access and adherence, ultimately improving treatment outcomes and patient quality of life.

The different effects of sun exposure across clusters highlight the importance of personalized sun protection plans. While some clusters face worsening conditions due to high sun exposure, others with little exposure may still encounter issues related to UV sensitivity. Customizing sun protection advice based on individual sun exposure patterns and combining these strategies with other treatments could help lower the risk of melasma flare-ups and better manage the condition.

ML has excellent potential to improve the diagnosis and treatment of melasma, especially in people with darker skin tones. The challenge of diagnosing and managing melasma in darker-skinned individuals often stems from the higher risk of post-inflammatory hyperpigmentation and the ways it manifests differently than in lighter skin. Advanced ML models, like convolutional neural networks, can improve diagnostic precision by detecting subtle differences in pigmentation patterns that traditional methods might miss. This is particularly useful for patients with darker skin, where pigmentation changes may be less obvious and more challenging to assess visually.

ML algorithms can also help personalize treatment plans by analyzing large datasets to identify patterns and connections that might not be apparent through traditional methods. By including patient-specific factors such as skin tone, treatment history, and responses to previous therapies, ML models can make more accurate predictions about treatment effectiveness and potential side effects. This personalized approach could lead to more effective and safer treatments, reducing the risk of adverse effects and boosting overall patient satisfaction.

To enhance the accuracy of treatment strategies, future research should validate the identified clusters using additional datasets to ensure their robustness and applicability across diverse populations. Longitudinal studies are particularly valuable because they can offer insights into the stability of these clusters over time and their predictive power for treatment outcomes. Investigating the mechanisms underlying differences between clusters could reveal new therapeutic targets and biomarkers, deepen our understanding of melasma's pathophysiology, and inform the development of more targeted therapies.

Exploring alternative clustering methods, such as hierarchical or density-based techniques, may improve the identification of patient subgroups and uncover new insights into patient heterogeneity. Comparing various clustering approaches could improve the accuracy and relevance of patient segmentation, leading to more effective, personalized

treatments. Research should focus on developing and evaluating patient-centered interventions tailored to each group's specific needs, including customized education programs, support services, and treatment plans.

Assessing the long-term impact of personalized treatment strategies is crucial for understanding their sustainability and effectiveness. Evaluating how these precision treatments perform in real-world settings will give valuable insights into their long-term benefits and areas for improvement. By continuously validating, refining, and expanding these findings, researchers and clinicians can improve melasma management, ultimately providing more effective, personalized care that meets the diverse needs of patients. Using machine learning for diagnosis and treatment, especially for patients with darker skin tones, has significant potential to improve the accuracy and effectiveness of melasma management and to address the complexities of treating this condition across different patient populations.

Although progress has been made in applying machine learning to dermatology, several challenges remain that must be addressed to realize the potential of these technologies fully. One major challenge is the variability in image quality and dataset representation. Differences in image capture conditions, such as lighting and resolution, can impact ML model performance, leading to inconsistent results. Additionally, the diversity of datasets used for training can influence their ability to generalize across different populations and clinical settings. To address these issues, further research is needed to standardize image capture protocols and create more robust datasets that better represent diverse patient groups.

Another critical area for future research is validating ML models for clinical use. Although ML algorithms show promise in research settings, their performance in real-world clinical environments requires careful evaluation. This involves examining the models' reliability, robustness, and usefulness in various clinical situations. Collaboration among researchers, clinicians, and technology developers is crucial to overcoming these challenges and ensuring ML tools are effectively integrated into clinical practice.

Strengths and Limitations

This study introduces a new, data-driven framework for melasma phenotyping using machine learning clustering to go beyond traditional clinical grading and support more personalized treatment strategies. A key strength is applying this innovative method to a specific, understudied population in KwaZulu-Natal, South Africa, addressing a critical gap in localized dermatological data for skin of colour. By identifying distinct patient subgroups, this research provides a reproducible methodological model that can be scaled to improve clinical decision-making. However, the preliminary nature of this study also highlights some limitations that suggest areas for future improvement. While the sample size (N = 150) enabled initial cluster identification, a larger group would enhance the relevance of these findings for rarer clinical subgroups. The moderate Silhouette Score (0.200) indicates some overlap in patient features, likely due to the biological complexity and continuous spectrum of melasma phenotypes, rather than clear boundaries. Additionally, reliance on pre-existing data prevented the inclusion of genetic markers or biochemical parameters that could have increased clustering accuracy. Although the current cross-sectional design and regional focus limit temporal analysis and broader geographic application, they provide a necessary foundation for future prospective studies. Future research should validate these findings with independent datasets, compare standard care with cluster-guided protocols, and evaluate outcomes such as reductions in MASI scores, improvements in quality of life, and recurrence rates to confirm the clinical value of this clustering approach.

5. Conclusions

This study develops a data-driven taxonomy of melasma in Fitzpatrick skin types IV–VI, demonstrating that traditional clinical grading fails to capture the disease’s multidimensional complexity. By identifying specific phenotypic groups defined by the interaction among ultraviolet and visible light exposure, hormonal status, and socioeconomic factors, this research offers a foundational framework for advancing precision dermatology. This stratification model allows a shift from general management to personalized treatment strategies, enabling behaviourally targeted photoprotection and education-based interventions that address quality-of-life factors. These findings do not support immediate clinical implementation but offer a foundation for hypothesis-driven research in stratified melasma care. Ultimately, this patient-centered clustering method marks a significant step toward evidence-based precision medicine, ensuring that melasma treatment is as nuanced and multidimensional as the population it affects.

Author Contributions: M.P. and N.M. contributed equally from conceptualization through literature analysis to the drafting of this manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Durban University of Technology Research Fund. This work is based on research supported by the Department of Higher Education and Training (DHET) through the University Capacity Development Programme (UCDP), under the University Staff Development Programme (USD) funding framework, Phase 5, 2024.

Institutional Review Board Statement: The study was approved by the University of KwaZulu-Natal Biomedical Research Ethics Committee (UKZN BREC) (Protocol reference number: BREC/00002721/2021, 12 August 2021).

Informed Consent Statement: All study participants provided written informed consent from four private dermatology clinics in eThekweni, KwaZulu-Natal, South Africa, between March and December 2023.

Data Availability Statement: Due to privacy, data are available upon request from the first author.

Conflicts of Interest: The authors have no competing interests to declare that are relevant to the content of this article.

Abbreviations

The following abbreviations are used in this manuscript:

ML	Machine learning
AIC	Akaike Information Criterion
WCSS	Within-cluster sum of squares
EADV	European Academy of Dermatology and Venereology
CAD	Computer-Aided Diagnosis
MASI	Melasma Area and Severity Index
MELASQoL	Melasma Quality of Life
BIC	Bayesian Information Criterion
BREC	Biomedical Research Ethics Committee
PIH	Post-inflammatory hyperpigmentation
UKZN	University of KwaZulu-Natal

References

1. Ogbechie-Godec, O.A.; Elbuluk, N. Melasma: An Up-to-Date Comprehensive Review. *Dermatol. Ther.* **2017**, *7*, 305–318. [[CrossRef](#)]
2. Serena, N.B.; Smoller, G.B. An Overview on Melasma. *J. Pigment. Disord.* **2015**, *2*, 218. [[CrossRef](#)]

3. Melnick, S.B.; Lohani, S.; Alweis, R. Hyperpigmentation in a Middle Aged Woman: A Common Yet Underdiagnosed Condition. *J. Community Hosp. Intern. Med. Perspect.* **2016**, *6*, 31544. [[CrossRef](#)]
4. Neagu, N.; Conforti, C.; Agozzino, M.; Marangi, G.F.; Morariu, S.H.; Pellacani, G.; Persichetti, P.; Piccolo, D.; Segreto, F.; Zalaudek, I.; et al. Melasma Treatment: A Systematic Review. *J. Dermatol. Treat.* **2022**, *33*, 1816–1837. [[CrossRef](#)]
5. Pandya, A.G.; Hynan, L.S.; Bhore, R.; Riley, F.C.; Guevara, I.L.; Grimes, P.E.; Nordlund, J.J.; Rendon, M.; Taylor, S.C.; Gottschalk, R.W.; et al. Reliability Assessment and Validation of the Melasma Area and Severity Index (MASI) and a New Modified MASI Scoring Method. *J. Am. Acad. Dermatol.* **2011**, *64*, 78–83. [[CrossRef](#)] [[PubMed](#)]
6. Grimes, P.E. Management of Hyperpigmentation in Darker Racial Ethnic Groups. *Semin. Cutan. Med. Surg.* **2009**, *28*, 77–85. [[CrossRef](#)] [[PubMed](#)]
7. Grimes, P.E.; Ijaz, S.; Nashawati, R.; Kwak, D. New Oral and Topical Approaches for the Treatment of Melasma. *Int. J. Women's Dermatol.* **2019**, *5*, 30–36. [[CrossRef](#)]
8. Rodrigues, M.; Pandya, A.G. Melasma: Clinical Diagnosis and Management Options. *Australas. J. Dermatol.* **2015**, *56*, 151–163. [[CrossRef](#)]
9. Sarkar, R.; Bansal, S.; Garg, V. Chemical Peels for Melasma in Dark-Skinned Patients. *J. Cutan. Aesthetic Surg.* **2012**, *5*, 247–253. [[CrossRef](#)]
10. Mpofana, N.; Paule, M.; Gqaleni, N.; Makgobole, M.U.; Pillay, P.; Hussein, A.; Dlova, N.C. The Effect of Melasma on the Quality of Life in People with Darker Skin Types Living in Durban, South Africa. *Int. J. Environ. Res. Public Health* **2023**, *20*, 7068. [[CrossRef](#)]
11. Trivedi, M.; Yang, F.C.; Cho, B.K. A Review of Laser and Light Therapy in Melasma. *Int. J. Women's Dermatol.* **2017**, *3*, 11–20. [[CrossRef](#)]
12. Beylot, C.; Raimbault-Gérard, C. Les Hyperpigmentations Post-Inflammatoires Succédant à Des Actes Esthétiques. *Ann. Dermatol. Vénérologie* **2016**, *143*, S33–S42. [[CrossRef](#)]
13. Ezekwe, N.; Pourang, A.; Lyons, A.B.; Narla, S.; Atyam, A.; Zia, S.; Friedman, B.J.; Hamzavi, I.H.; Lim, H.W.; Kohli, I. Evaluation of the protection of sunscreen products against long wavelength ultraviolet A1 and visible light-induced biological effects. *Photodermatol. Photoimmunol. Photomed.* **2024**, *40*, e12937. [[CrossRef](#)]
14. He, M.; Chen, X.; Jin, S.; Zhang, C. Visible Light Protection: An Updated Review of Tinted Sunscreens. *Photodermatol. Photoimmunol. Photomed.* **2025**, *41*, e70033. [[CrossRef](#)] [[PubMed](#)]
15. Zhou, C.; Lee, C.; Salas, J.; Luke, J. Guide to tinted sunscreens in skin of color. *Int. J. Dermatol.* **2024**, *63*, 272–276. [[CrossRef](#)]
16. Azim, S.A.; Whiting, C.; Friedman, A.J. Attitudes on, Practices, and Recommendations for Visible Light Protection Amongst Dermatology Practitioners. *J. Drugs Dermatol. JDD* **2024**, *23*, 965–971. [[CrossRef](#)] [[PubMed](#)]
17. Tseng, G.C.; Wong, W.H. Tight Clustering: A Resampling-Based Approach for Identifying Stable and Tight Patterns in Data. *Biometrics* **2005**, *61*, 10–16. [[CrossRef](#)] [[PubMed](#)]
18. Campello, R.J.G.B.; Kröger, P.; Sander, J.; Zimek, A. Density-based Clustering. *WIREs Data Min. Knowl.* **2020**, *10*, e1343. [[CrossRef](#)]
19. Aksoy, S.; Demircioglu, P.; Bogrekci, I. Advanced artificial intelligence techniques for comprehensive dermatological image analysis and diagnosis. *Dermato* **2024**, *4*, 173–186. [[CrossRef](#)]
20. Chan, H.P.; Hadjiiski, L.M.; Samala, R.K. Computer-aided diagnosis in the era of deep learning. *Med. Phys.* **2020**, *47*, e218–e227. [[CrossRef](#)]
21. Nawaz, A.; Irfan, M.; Sadiqa, H.A.; Westerlund, T. Edge based skin cancer decision support system using machine learning algorithms. In Proceedings of the 2023 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress, (DASC/PiCom/CBDCom/CyberSciTech), Abu Dhabi, United Arab Emirates, 14–17 November 2023; IEEE: New York, NY, USA; pp. 0292–0297.
22. Arthur, D.; Vassilvitskii, S. k-means++: The advantages of careful seeding. In Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms 2007, New Orleans, LA, USA, 7–9 January 2007; pp. 1027–1035.
23. MacQueen, J. Some methods for classification and analysis of multivariate observations. *Proc. Fifth Berkeley Symp. Math. Stat. Probab.* **1967**, *1*, 281–297.
24. Zbrzezny, A.M.; Krzywicki, T. Artificial intelligence in dermatology: A review of methods, clinical applications, and perspectives. *Appl. Sci.* **2025**, *15*, 7856. [[CrossRef](#)]
25. Zhang, J.; Zhong, F.; He, K.; Ji, M.; Li, S.; Li, C. Recent advancements and perspectives in the diagnosis of skin diseases using machine learning and deep learning: A review. *Diagnostics* **2023**, *13*, 3506. [[CrossRef](#)] [[PubMed](#)]
26. Escalé-Besa, A.; Vidal-Alaball, J.; Miró Catalina, Q.; Gracia, V.H.G.; Marin-Gomez, F.X.; Fuster-Casanovas, A. The use of artificial intelligence for skin disease diagnosis in primary care settings: A systematic review. *Healthcare* **2024**, *12*, 1192. [[CrossRef](#)] [[PubMed](#)]
27. Bahmani, B.; Moseley, B.; Vattani, A.; Kumar, R.; Vassilvitskii, S. Scalable K-Means++. *Proc. VLDB Endow.* **2012**, *5*, 622–633. [[CrossRef](#)]

28. Hämmäläinen, J.; Kärkkäinen, T.; Rossi, T. Improving Scalable K-Means++. *Algorithms* **2020**, *14*, 6. [[CrossRef](#)]
29. Hassan, E.; Malik, F.; Khan, Q.; Ahmad, N.; Sardaraz, M.; Karim, F.; Elmannai, H. A Hybrid K-Means++ and Particle Swarm Optimization Approach for Enhanced Document Clustering. *IEEE Access* **2025**, *13*, 48818–48840. [[CrossRef](#)]
30. Yan, J.; Linn, K.; Powers, B.; Zhu, J.; Jain, S.; Kowalski, J.; Navathe, A. Applying Machine Learning Algorithms to Segment High-Cost Patient Populations. *J. Gen. Intern. Med.* **2018**, *34*, 211–217. [[CrossRef](#)]
31. Shi, C.; Wei, B.; Wei, S.; Wang, W.; Liu, H.; Liu, J. A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm. *J. Wirel. Commun. Netw.* **2021**, *2021*, 31. [[CrossRef](#)]
32. Maaten, L.V.D.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
33. Xia, J.; Zhang, Y.; Song, J.; Chen, Y.; Wang, Y.; Liu, S. Revisiting Dimensionality Reduction Techniques for Visual Cluster Analysis: An Empirical Study. *IEEE Trans. Vis. Comput. Graph.* **2021**, *28*, 529–539. [[CrossRef](#)] [[PubMed](#)]
34. Hennig, C. Cluster-wise assessment of cluster stability. *Comput. Stat. Data Anal.* **2007**, *52*, 258–271. [[CrossRef](#)]
35. Yu, H.; Chapman, B.; Florio, A.; Eischen, E.; Gotz, D.; Jacob, M.; Blair, R. Bootstrapping estimates of stability for clusters, observations and model selection. *Comput. Stat.* **2018**, *34*, 349–372. [[CrossRef](#)]
36. Chan, X.; Wang, X.; Yu, D.; Mi, H.; Yu, D. Scaling Synthetic Data Creation with 1,000,000,000 Personas. *arXiv* **2024**. [[CrossRef](#)]
37. Jones, M.; Griffioen, N.; Neumayer, C.; Shklovski, I. Artificial Intimacy: Exploring Normativity and Personalization Through Fine-tuning LLM Chatbots. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 26 April–1 May 2025. [[CrossRef](#)]
38. OpenAI. ChatGPT (GPT-4 Model). Available online: <https://chat.openai.com/> (accessed on 2 January 2026).
39. Wang, Y.; Wen, Q.; Lian, X.; Liu, L.; Zhou, Q.; Zhang, Y.; Chen, C.; Wu, G.; Wang, C.; Liu, Q.; et al. Machine learning-based unsupervised phenotypic clustering analysis of patients with IgA nephropathy: Distinct therapeutic responses of different groups. *Chin. Med. J.* **2026**, *139*, 83–92. [[CrossRef](#)]
40. Chen, S.; Gopalakrishnan, P. Clustering via the Bayesian information criterion with applications in speech recognition. In Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing 1998, ICASSP '98 (Cat. No.98CH36181), Seattle, WA, USA, 12–15 May 1998; Volume 2, pp. 645–648. [[CrossRef](#)]
41. Teklehaymanot, F.; Muma, M.; Zoubir, A. Bayesian Cluster Enumeration Criterion for Unsupervised Learning. *IEEE Trans. Signal Process.* **2017**, *66*, 5392–5406. [[CrossRef](#)]
42. Shahapure, K.; Nicholas, C. Cluster Quality Analysis Using Silhouette Score. In Proceedings of the 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), Sydney, Australia, 6–9 October 2020; pp. 747–748. [[CrossRef](#)]
43. Caliński, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat.-Theory Methods* **1974**, *3*, 1–27. [[CrossRef](#)]
44. Celeux, G.; Soromenho, G. An entropy criterion for assessing the number of clusters in a mixture model. *J. Classif.* **1996**, *13*, 195–212. [[CrossRef](#)]
45. Mpofana, N.; Mlambo, Z.; Makgobole, M.; Dlova, N.; Naicker, T. Association of Genetic Polymorphisms in SLC45A2, TYR, HERC2, and SLC24A in African Women with Melasma: A Pilot Study. *Int. J. Mol. Sci.* **2025**, *26*, 1158. [[CrossRef](#)]
46. Sun, Q.; Sun, H.; Cong, L.; Zheng, Y.; Wu, N.; Cong, X. Effects of Exogenous Hormones and Reproductive Factors on Female Melanoma: A Meta-Analysis. *Clin. Epidemiol.* **2020**, *12*, 1183–1203. [[CrossRef](#)]
47. Tang, X.; Zhang, H.; Cui, Y.; Wang, L.; Wang, Z.; Zhang, Y.; Huo, J.; Cai, J.; Rinaldi, G.; Bhagavathula, A.S.; et al. Postmenopausal Exogenous Hormone Therapy and Melanoma Risk in Women: A Systematic Review and Time-Response Meta-Analysis. *Pharmacol. Res.* **2020**, *160*, 105182. [[CrossRef](#)] [[PubMed](#)]
48. Johnston, G.A.; Sviland, L.; McLelland, J. Melasma of the Arms Associated with Hormone Replacement Therapy. *Br. J. Dermatol.* **1998**, *139*, 932. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.