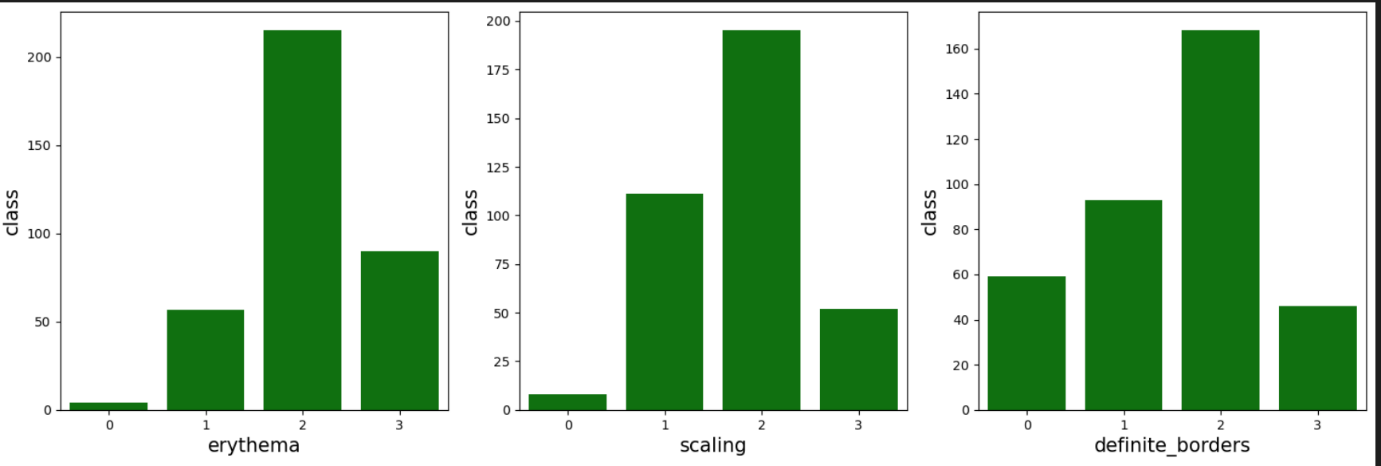


**Age Distribution:** The age distribution is approximately normal but slightly skewed to the right. This suggests that most people are in the middle age range.

**Peak Age Group:** The highest frequency occurs in the 20-40 age range, indicating a larger population in this group.

**Decline in Older Ages:** There is a noticeable decline in count after the age of 50, showing fewer people in the older age groups.

**Young Age Representation:** There is also a smaller representation of very young individuals (below 20).



**Erythema:**

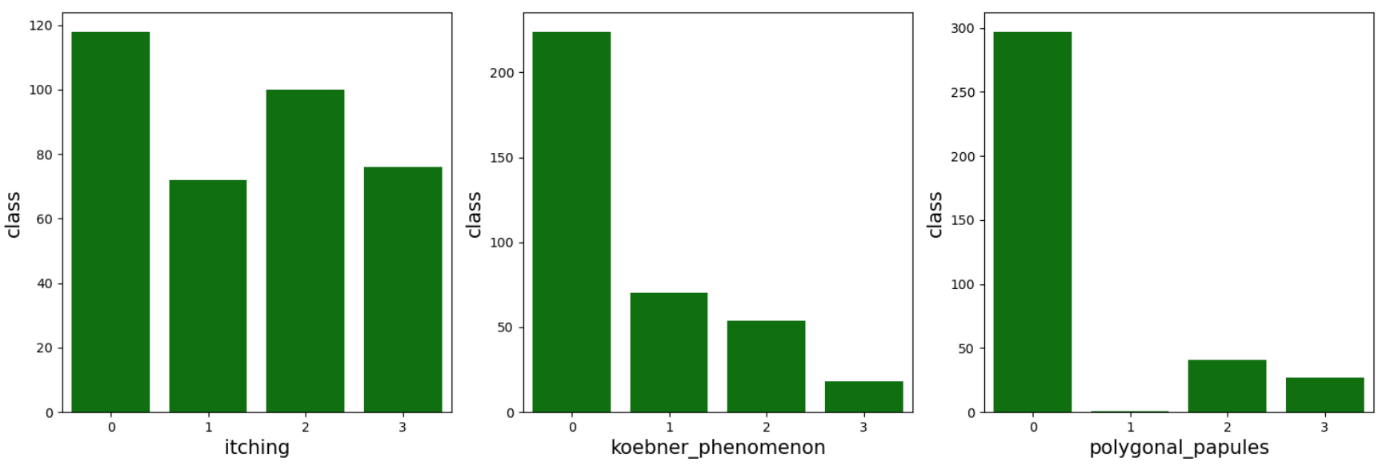
* Most observations are concentrated at level 2, indicating moderate erythema is the most common.
* Level 1 and 3 are less frequent, and level 0 is very rare.

**Scaling:**

* Level 2 has the highest count, suggesting moderate scaling is prevalent.
* Level 1 also has a significant number of cases, while level 3 is less common.
* Level 0 is rarely observed.

**Definite Borders:**

* Level 2 is the most frequent, showing that definite borders are commonly observed.
* Levels 1 and 0 have moderate counts, while level 3 is the least common.



**Itching:**

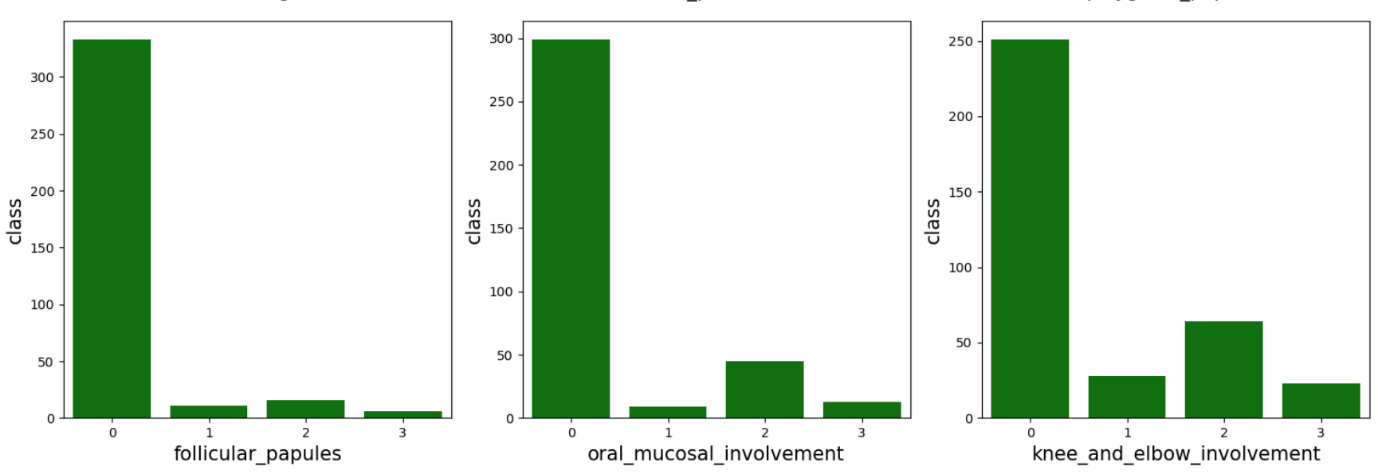
* Level 0 (no itching) is the most common, suggesting many cases do not involve itching.
* Levels 2 and 3 have a noticeable presence, indicating moderate to severe itching in some cases.
* Level 1 is the least frequent, showing mild itching is less common.

**Koebner Phenomenon:**

* Level 0 dominates, showing the majority of cases do not exhibit the Koebner phenomenon.
* Levels 1 and 2 are less frequent, and level 3 is rare.

**Polygonal Papules:**

* Level 0 is overwhelmingly common, indicating that polygonal papules are mostly absent.
* Levels 2 and 3 are rare, showing few cases with severe symptoms.
* Level 1 is almost nonexistent.



**Follicular Papules:**

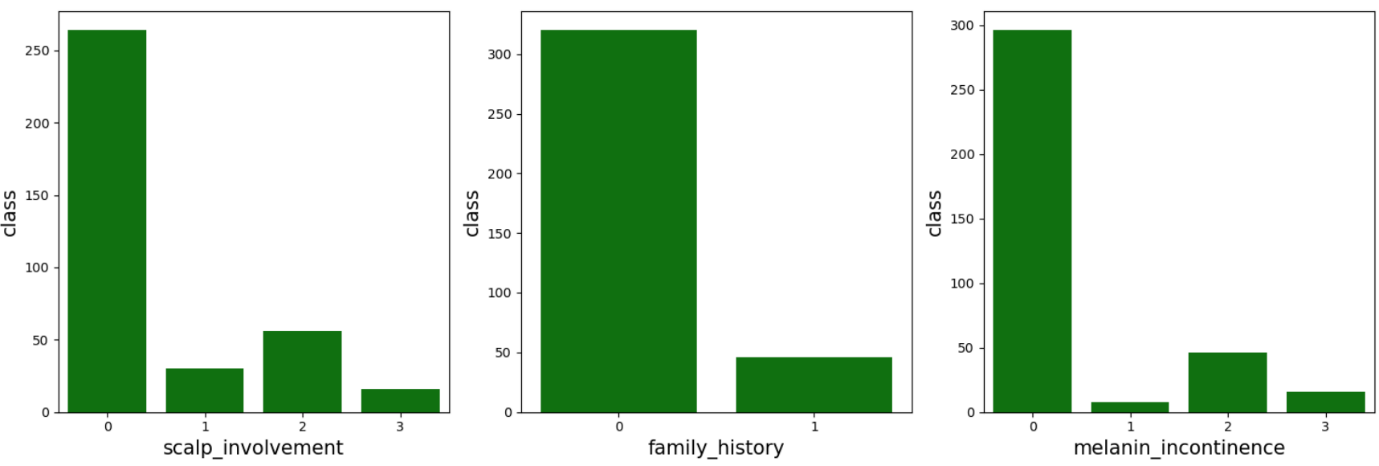
* Level 0 (absence of follicular papules) is the most common, indicating that most cases do not show this symptom.
* Levels 1, 2, and 3 are significantly less frequent, showing that moderate to severe follicular papules are rare.

**Oral Mucosal Involvement:**

* Level 0 is overwhelmingly common, suggesting oral mucosal involvement is mostly absent.
* Level 2 shows a noticeable presence, indicating moderate involvement in some cases.
* Levels 1 and 3 are rare, indicating mild and severe cases are uncommon.

**Knee and Elbow Involvement:**

* Level 0 is the most frequent, showing that the majority of cases do not involve the knees and elbows.
* Level 2 has a notable presence, indicating moderate involvement in some cases.
* Levels 1 and 3 are less frequent, suggesting mild and severe involvement are uncommon.



**Scalp Involvement:**

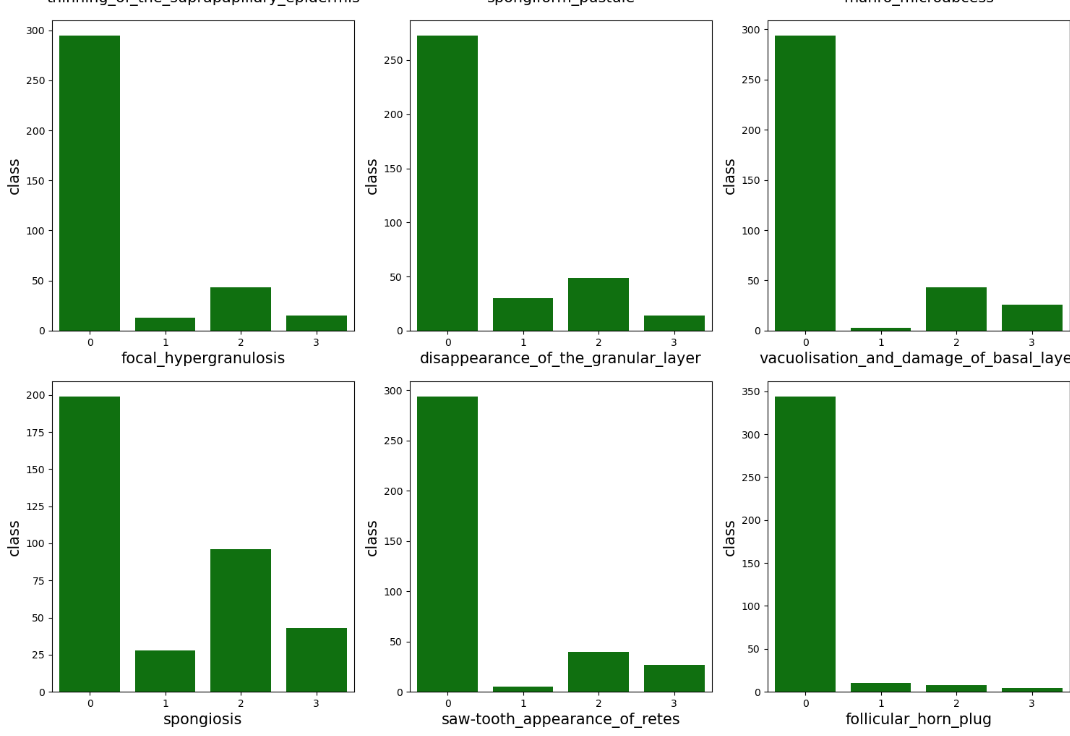
* Level 0 (absence of scalp involvement) is the most common, indicating that most cases do not show this symptom.
* Level 2 has a noticeable presence, showing moderate involvement in some cases.
* Levels 1 and 3 are less frequent, suggesting mild and severe involvement are rare.

**Family History:**

* Level 0 is overwhelmingly common, indicating that most cases have no family history of the condition.
* Level 1 (presence of family history) is much less frequent, suggesting genetic factors are not prevalent in the majority of cases.

**Melanin Incontinence:**

* Level 0 is the most frequent, indicating absence of melanin incontinence in most cases.
* Level 2 shows a moderate presence, suggesting partial involvement in some cases.
* Levels 1 and 3 are rare, indicating mild and severe cases are uncommon.



**Class Imbalance:**

* In all six plots, the majority class (0) has significantly higher counts compared to the other classes (1, 2, and 3).
* This indicates that most instances have a value of 0 for these features, suggesting their rarity in the dataset.

**Patterns Across Features:**

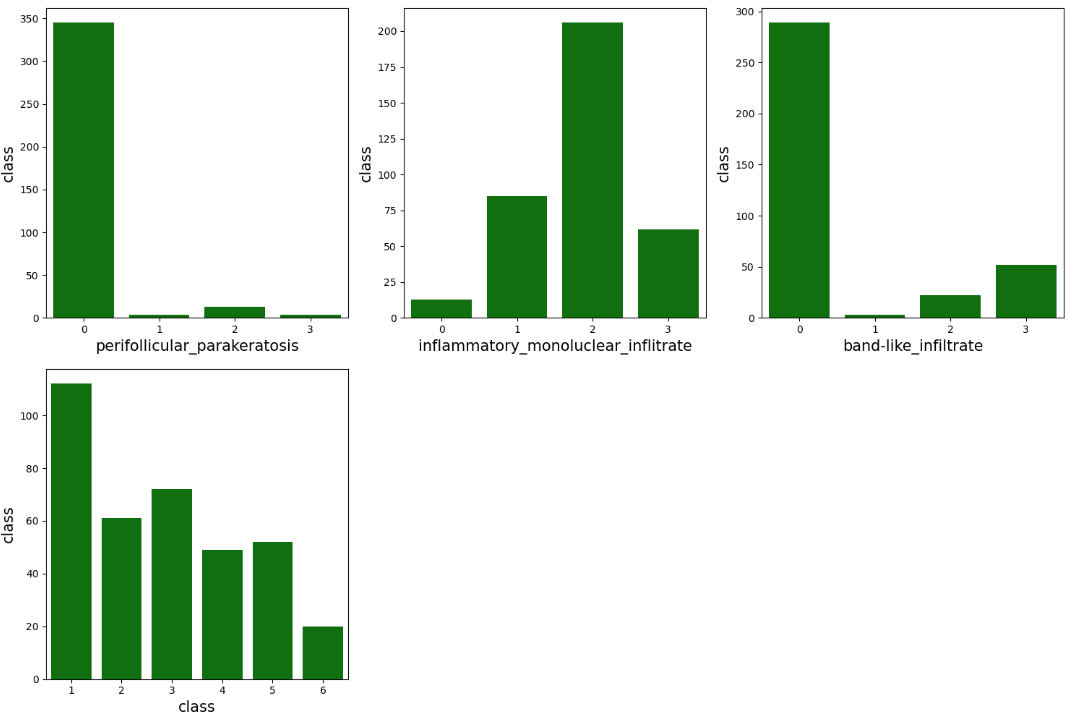
* Features like **focal hypergranulosis, disappearance of the granular layer, vacuolisation and damage of the basal layer, saw-tooth appearance of retes, and follicular horn plug** show highly skewed distributions where class 0 dominates.
* **Spongiosis** appears to have a slightly more balanced distribution compared to the other features, but class 0 is still the highest.

**Potential Challenges for Machine Learning Models:**

* Since most features are heavily imbalanced, ML models might struggle to learn meaningful patterns for the minority classes.
* Techniques like oversampling (SMOTE), undersampling, or weighting class contributions may be needed.

**Clinical Significance:**

* The dominance of class 0 suggests that these features might not be present in a majority of cases.
* Features with slightly higher representation in other classes (e.g., spongiosis) could be more clinically relevant for distinguishing between conditions.



**Perifollicular Parakeratosis:**

* Majority of the instances belong to class **0**, indicating that this feature is rarely present in most cases.
* Very few samples belong to classes **1, 2, and 3**.

**Inflammatory Mononuclear Infiltrate:**

* The distribution is more balanced compared to other features.
* Class **2** has the highest count, followed by **1** and **3**, while class **0** is the lowest.
* This suggests that inflammatory mononuclear infiltrate is commonly observed in cases.

**Band-like Infiltrate:**

* The distribution is heavily skewed towards class **0**, meaning that in most cases, this feature is absent.
* Classes **1, 2, and 3** have very few occurrences.

**Conclusion**

The dataset shows significant class imbalance, with most instances in class **0**, indicating the absence of key features. Highly skewed distributions in features like **focal hypergranulosis and vacuolisation** may hinder model learning. However, **spongiosis and inflammatory mononuclear infiltrate** have a more balanced distribution, making them clinically relevant. The **age distribution** peaks at **20-40 years** and declines in older individuals. To improve model performance, **data balancing techniques** like **SMOTE** or **weight adjustments** are necessary.