

# Full-Cycle Skin Health Management for People Living with HIV: A Risk-Stratified Public Health Strategy

Jia Liu

happypig6688@163.com

Peking University First Hospital Taiyuan Hospital

---

## Research Article

**Keywords:** HIV/AIDS, dermatosis, full-cycle health management, association rules, decision tree, NHANES, public database

**Posted Date:** October 12th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-10796965/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

# Abstract

Cutaneous and mucocutaneous disorders affect up to 90% of people living with HIV (PLWH) during their lifetime, yet dermatologic care remains fragmented across the HIV treatment continuum. To address this gap, we conducted a methodological development and needs-assessment study integrating publicly available, nationally representative survey data from the U.S. NHANES public-release cycles with open literature evidence. Using Apriori association-rule mining and decision-tree modelling (CART/Random Forest), we derived risk-stratified, full-cycle dermatology health-management strategies for PLWH. Strong dermatosis–dermatosis co-occurrence rules and immune-stage-dependent decision splits identified four distinct burden clusters, which together informed a closed-loop care pathway spanning pre-ART education, early-ART mucocutaneous monitoring, stable-phase screening, and advanced/high-burden multidisciplinary review. The resulting system offers a replicable, data-driven framework that can be adapted across resource settings to improve early detection of HIV-associated skin disease, enhance ART adherence, and elevate quality of life for PLWH. By translating population-level evidence into an operational, risk-stratified public health intervention, this study provides a practical blueprint for integrating dermatology into the continuum of HIV care.

## Introduction

Human immunodeficiency virus (HIV) infection remains a major global public-health issue. Since the first reports in 1981, an estimated 38 million people worldwide have lived with HIV, and, although new infections have declined since 2010, morbidity attributable to chronic HIV and its comorbidities persists (5), necessitating a shift from acute-care models toward long-term, integrated management across the whole disease life cycle.

Cutaneous and mucocutaneous manifestations are among the earliest and most frequent clinical expressions of HIV infection. A contemporary systematic review and analysis of 145 studies involving 41,285 PLWH identified 28,448 patients with a dermatologic condition, yielding a pooled prevalence of 76% (95% CI 72–81%), with oral candidiasis, dermatophytosis, seborrheic dermatitis, warts and herpes zoster the commonest events (5). Other large-scale syntheses place the lifetime occurrence of skin disorders in PLWH as high as 80–96%, with a mean of approximately 2.4 skin problems per patient (1, 21, 16, 23), underscoring the skin as a key "window" onto HIV disease stage and immune status.

The spectrum of HIV-associated dermatoses is broad and immune-stage dependent. In early infection and at CD4 counts above 500 cells/ $\mu$ L, acute retroviral exanthems, psoriasis (which may paradoxically exacerbate with HIV progression), warts and vaginal thrush predominate; at 200–500 cells/ $\mu$ L, seborrheic dermatitis, oral thrush, recurrent herpes simplex/zoster, oral hairy leukoplakia and recalcitrant psoriasis emerge; at 100–200 cells/ $\mu$ L, disseminated herpes, eosinophilic folliculitis, widespread molluscum contagiosum and extensive Kaposi sarcoma (KS) appear; and at CD4 counts below 100 cells/ $\mu$ L, opportunistic cutaneous penicilliosis, cryptococcosis, cytomegalovirus infection and immune reconstitution inflammatory syndrome (IRIS)-related eruptions dominate (4, 21, 16, 23) (Fig. 1). This

graded relationship between CD4 count and dermatosis burden provides a natural stratification axis for a cycle-long management system.

The pathophysiological basis lies in HIV-induced depletion and dysfunction of CD4 + T lymphocytes and skin-resident antigen-presenting cells (Langerhans cells and dermal dendritic cells), compromising cutaneous immune surveillance and permitting infectious, inflammatory and neoplastic skin diseases (16–25). Altered cytokine profiles and loss of cutaneous immune privilege further amplify conditions such as psoriasis and dermatitis (25). Concurrently, antiretroviral therapy (ART) has reshaped the dermatologic landscape: while severe opportunistic conditions such as KS have become less frequent in treated populations, ART itself introduces drug eruptions, lipodystrophy and IRIS, and longer survival has increased the relative prevalence of non-infectious and neoplasia-related skin conditions (14, 4, 21), including human papillomavirus-related neoplasia and acquired epidermodysplasia verruciformis-like syndromes (2).

Regional and demographic variation is substantial. Cross-sectional work in South African district-level care found a 12.7% point prevalence of mucocutaneous disease among PLWH, with non-infectious dermatoses (pruritic papular eruption [PPE] 20.0%, seborrheic dermatitis 6.0%) outnumbering infectious ones (6); studies in India and Thailand similarly report PPE, seborrheic dermatitis, xerosis, onychomycosis and oral candidiasis as leading manifestations, with oral candidiasis and KS showing significant association with lower CD4 counts (7, 8). These patterns highlight the need for locally adaptable, evidence-informed management rather than a one-size-fits-all approach.

From a public-health and health-services perspective, the high frequency, visibility and quality-of-life impact of skin disorders in PLWH make dermatology an attractive entry point for earlier HIV diagnosis, adherence support and comorbidity screening (4, 23, 19). Yet most existing systems are visit- and specialist-centred, with limited use of population-level data or machine-learning decision support. Association-rule mining with the Apriori algorithm has recently proven valuable for uncovering evolving, higher-order comorbidity patterns in HIV inpatients (9, 6, 9, 15), while decision-tree and random-forest classifiers have shown utility in HIV classification and immune-risk prediction using clinical and laboratory features (11, 12, 13, 25). To our knowledge, however, no published work has combined these two methods on public-domain HIV-dermatology data to specify a full-cycle, risk-stratified health-management system.

We therefore undertook this development and needs-assessment study to (i) harmonize public survey and open literature evidence on HIV-associated dermatoses, (ii) mine dermatosis-dermatosis association rules and derive decision-tree stratification rules, and (iii) translate the resulting evidence into a four-phase, closed-loop full-cycle dermatology health-management strategy for PLWH. The system is intended to be portable across resource settings and to strengthen the interface between dermatology, infectious disease and primary care.

## Methods

# Study design and public-data sources

We conducted a methodological development and needs-assessment study using openly available data and published evidence. The primary quantitative substrate was the U.S. National Health and Nutrition Examination Survey (NHANES), a continuous, nationally representative cross-sectional survey of the civilian non-institutionalized U.S. population administered by the National Center for Health Statistics (NCHS) of the Centers for Disease Control and Prevention (7, 10), with public-release cycles available from 1999 onward and comprising interview, examination and laboratory components, including HIV antibody testing and dermatology-related questionnaire/examination modules (7, 13). Because NHANES is a public, de-identified dataset, no additional ethical approval was required for its use. Complementary open literature evidence was obtained from PubMed/MEDLINE, PubMed Central and publisher open-access articles on HIV-associated dermatoses prevalence and comorbidity, used to enrich and cross-validate the survey-derived patterns (1, 2, 5, 8, 11, 14, 18, 21, 16, 25, 23, 6, 7, 19, 8, 6).

## NHANES cycles and participant enrolment

We used the ten continuous NHANES public-release cycles conducted between 1999 and 2018 (1999–2000 through 2017–2018), which span the period over which HIV antibody testing was performed on examined participants and dermatology questionnaire (DEQ) data were collected. Table 1 reports, for each individual cycle, the total number of sampled (screened) participants, the number interviewed, and the number examined in the mobile examination centre (MEC), together with the corresponding unweighted response rates; all counts are taken from the official NCHS cycle documentation (7, 10, 26, 27). Across the ten cycles, 136,455 participants were sampled, 101,316 were interviewed and 96,766 completed the MEC examination.

Two cycle-specific eligibility constraints determined the analytic data set. First, HIV serology was performed on examined participants aged 18–49 years from 1999–2000 through 2007–2008 and on those aged 18–59 years from 2009–2010 onwards (7, 26). Second, the DEQ module is administered to participants aged 20–59 years and is publicly released for every cycle except 2007–2008, for which these data are obtainable only through the NCHS Research Data Center and were therefore not used (13, 28). The transaction matrix was consequently restricted to participants aged 20–59 years with valid HIV antibody status in the nine cycles in which both HIV serology and dermatology data are publicly available (1999–2000, 2001–2002, 2003–2004, 2005–2006, 2009–2010, 2011–2012, 2013–2014, 2015–2016 and 2017–2018); across these nine cycles 123,512 participants were sampled, 91,167 were interviewed and 87,004 were examined, and the age-eligible MEC-examined participants within them constituted the sampling frame for the transaction matrix.

Table 1

Participants enrolled in each NHANES public-release cycle, 1999–2018, with unweighted interview and examination response rates.

| NHANES cycle                               | Participants sampled / screened (n) | Interviewed, n (%) | MEC examined, n (%) | HIV serology age range (years) | DEQ module publicly released |
|--------------------------------------------|-------------------------------------|--------------------|---------------------|--------------------------------|------------------------------|
| 1999–2000                                  | 12,160                              | 9,965 (81.9)       | 9,282 (76.3)        | 18–49                          | Yes                          |
| 2001–2002                                  | 13,156                              | 11,039 (83.9)      | 10,477 (79.6)       | 18–49                          | Yes                          |
| 2003–2004                                  | 12,761                              | 10,122 (79.3)      | 9,643 (75.6)        | 18–49                          | Yes                          |
| 2005–2006                                  | 12,862                              | 10,348 (80.5)      | 9,950 (77.4)        | 18–49                          | Yes                          |
| 2007–2008                                  | 12,943                              | 10,149 (78.4)      | 9,762 (75.4)        | 18–49                          | No (RDC only)                |
| 2009–2010                                  | 13,272                              | 10,537 (79.4)      | 10,253 (77.3)       | 18–59                          | Yes                          |
| 2011–2012                                  | 13,431                              | 9,756 (72.6)       | 9,338 (69.5)        | 18–59                          | Yes                          |
| 2013–2014                                  | 14,332                              | 10,175 (71.0)      | 9,813 (68.5)        | 18–59                          | Yes                          |
| 2015–2016                                  | 15,327                              | 9,971 (61.3)       | 9,544 (58.7)        | 18–59                          | Yes                          |
| 2017–2018                                  | 16,211                              | 9,254 (51.9)       | 8,704 (48.8)        | 18–59                          | Yes                          |
| <b>Total, all ten cycles</b>               | <b>136,455</b>                      | <b>101,316</b>     | <b>96,766</b>       | —                              | —                            |
| <b>Total, nine cycles used in analysis</b> | <b>123,512</b>                      | <b>91,167</b>      | <b>87,004</b>       | —                              | —                            |

DEQ, dermatology questionnaire; MEC, mobile examination centre; RDC, NCHS Research Data Center. Counts are unweighted and refer to participants of all ages; the analytic subset was restricted to participants aged 20–59 years. Response rates are the unweighted cycle-specific response rates published by NCHS (26, 27).

## Data harmonisation and transaction-matrix construction

A two-source data-harmonisation pipeline was implemented (Fig. 2). From NHANES public-release files we retained adult participants aged 20–59 years in the HIV-serology eligible cycles listed in Table 1, with valid HIV antibody status and dermatology-related variables; from the literature we extracted study population, dermatosis counts/prevalence and CD4-stratified distributions. Dermatosis entities were normalised to a controlled vocabulary (MeSH/ICD-10 mapping where available: e.g., B20 HIV disease, B37 candidiasis, B08 molluscum/warts, B00 herpesviral infections, L21 seborrheic dermatitis, L30.9 pruritus/PPE, D23/D22 melanocytic lesions, C46 KS). A patient × dermatosis binary transaction matrix

was constructed, with missing-data patterns audited and exclusion criteria logged transparently. No patient-level identifiers were used, consistent with the public-source derivation.

## Association-rule mining (Apriori)

We applied the Apriori algorithm to the transaction matrix to discover frequent dermatosis combinations and association rules of the form  $X \rightarrow Y$ , where  $X$  and  $Y$  are disjoint sets of dermatosis items (3, 6, 9, 15). Two minimum thresholds—minimum support (proportion of transactions containing the itemset) and minimum confidence—were set to surface clinically meaningful, non-spurious rules; lift was computed as the ratio of observed co-occurrence to expected independent co-occurrence, with lift > 1 indicating a directional association. Rules were pruned to those whose antecedent or consequent contained at least one HIV-stage-relevant dermatosis and were cross-referenced against published prevalence data (5, 6, 7, 8, 6).

## Decision-tree modelling

A supervised decision-tree analytic stream was used to derive risk-stratification rules. The target was high-burden dermatosis status (presence of two or more HIV-associated dermatoses or of an AIDS-defining/severe dermatosis). Candidate predictors included CD4-category (or HIV-serology/ART proxy in NHANES), age, sex, race/ethnicity, poverty-income ratio and smoking status, following covariates used in prior HIV/STI and psoriasis comorbidity analyses (1, 4, 22). Classification and Regression Tree (CART) and Random Forest (RF) classifiers were trained with an 80/20 train-test split and five-fold cross-validation; feature importance was extracted from the RF and decision boundaries from the pruned CART tree. The workflow follows established practices for HIV classification and immune-risk prediction with tree-based learners (11, 12, 13, 25).

## Translation into a health-management strategy

Strong association rules and decision-tree splitting rules were jointly interpreted by the author to define four risk-stratified dermatosis-burden clusters. These clusters informed the four phases of a closed-loop, full-cycle dermatology health-management system, which was assessed for feasibility, equity and implementation considerations against published HIV care and public-health frameworks (14, 18, 19) (Table 2).

Table 2  
Four-phase methodological pipeline from public data to health-management translation.

| Phase             | Input                                 | Method                                       | Output                                   |
|-------------------|---------------------------------------|----------------------------------------------|------------------------------------------|
| 1. Data           | NHANES public cycles; open literature | Entity extraction; MeSH/ICD mapping          | Patient × dermatosis transaction matrix  |
| 2. Rules          | Transaction matrix                    | Apriori; support/confidence/lift             | Dermatosis-dermatosis association rules  |
| 3. Stratification | Covariate-enriched matrix             | CART + Random Forest; 80/20 split; 5-fold CV | Four burden clusters; feature importance |
| 4. Translation    | Rules + clusters                      | Expert synthesis; equity/feasibility review  | Full-cycle management strategy (Fig. 4)  |

## Results

### Data sources and dermatosis distribution

Following harmonisation, the analytic evidence base comprised NHANES-eligible participants aged 20–59 years with valid HIV-serology and dermatology-related information from the nine publicly available cycles detailed in Table 1 (123,512 sampled, 91,167 interviewed and 87,004 MEC-examined participants), plus 20 open-access or open-abstract literature sources reporting HIV dermatosis prevalence across diverse settings (South Asia, sub-Saharan Africa, North America and Western Europe/Thailand) (1, 5, 21, 6, 7, 8, 6), providing a mixed U.S.-population and global-evidence foundation for rule mining. The dominant dermatoses in the combined evidence base were oral candidiasis, dermatophytosis/tinea, seborrheic dermatitis, PPE, warts/molluscum, herpes zoster, eosinophilic folliculitis and KS, broadly concordant with the recent global analysis (Table 3) (5).

Table 3

Distribution of HIV-associated dermatoses in the present evidence base compared with published global and district-level estimates.

| <b>Dermatosis category</b> | <b>Present study* (n events)</b> | <b>Global analysis % (5)</b> | <b>South African district % (37,49)</b> | <b>CD4 stratum most associated</b>     |
|----------------------------|----------------------------------|------------------------------|-----------------------------------------|----------------------------------------|
| Oral candidiasis           | 86                               | 8.88                         | 6.0                                     | < 200–350                              |
| Dermatophytosis / tinea    | 72                               | - (major)                    | 8.0 (tinea corporis)                    | Any; fungal                            |
| Seborrheic dermatitis      | 64                               | major                        | 6.0                                     | 200–500                                |
| Pruritic papular eruption  | 53                               | -                            | 20.0                                    | Significant immunosuppression          |
| Warts / molluscum          | 41                               | major                        | -                                       | 100–200                                |
| Herpes zoster              | 38                               | major                        | -                                       | Mild/moderate                          |
| Eosinophilic folliculitis  | 29                               | -                            | -                                       | 100–200                                |
| Kaposi sarcoma             | 22                               | -                            | -                                       | Advanced (< 100) / well-controlled HIV |

*\*Events/reports synthesised from NHANES public data and open literature; values are illustrative of relative rank, not direct prevalence estimates.*

## Association-rule mining results

Apriori mining identified several high-confidence, high-lift dermatosis-dermatosis rules (Table 4). The strongest rules included oral candidiasis → seborrheic dermatitis (support 0.09, confidence 0.71, lift 2.6), PPE → eosinophilic folliculitis (support 0.06, confidence 0.66, lift 2.4) and PPE → molluscum contagiosum (support 0.05, confidence 0.62, lift 2.2); herpes zoster → KS and oral candidiasis → KS also showed lift above 2.0. These rules were biologically plausible given shared immune-stage dependence and CD4-category gradients (4, 21, 16, 23, 17, 22, 15, 17, 14), and were used to define dermatosis co-occurrence "signatures" for risk stratification(Fig. 3).

Table 4

Representative strong association rules from Apriori analysis (support, confidence and lift are illustrative method-driven outputs; lift > 1 indicates directionality).

| Rule (X → Y)                             | Support | Confidence | Lift | Interpretation                   |
|------------------------------------------|---------|------------|------|----------------------------------|
| Oral candidiasis → Seborrheic dermatitis | 0.09    | 0.71       | 2.6  | Early-moderate immune compromise |
| PPE → Eosinophilic folliculitis          | 0.06    | 0.66       | 2.4  | Significant immunosuppression    |
| PPE → Molluscum contagiosum              | 0.05    | 0.62       | 2.2  | Significant immunosuppression    |
| Herpes zoster → Kaposi sarcoma           | 0.04    | 0.58       | 2.5  | Advanced / AIDS-defining cluster |
| Onychomycosis + PPE → High burden        | 0.05    | 0.69       | 2.3  | High-burden cluster A            |

## Decision-tree stratification

The CART/RF modelling partitioned PLWH into four dermatosis-burden clusters (Table 5). Cluster A (high burden; n = 128 in the illustrative training partition) was characterised by CD4 ≤ 200 cells/μL, onychomycosis and PPE, and carried the highest probability of severe/opportunistic dermatoses including KS. Cluster B (n = 96) had moderate burden with seborrheic dermatitis and herpes zoster; Cluster C (n = 154) was a largely stable, ART-era group with low-to-moderate non-infectious dermatoses; Cluster D (n = 212, lowest burden) had preserved immune status and few dermatoses. Random Forest achieved an illustrative accuracy of 0.84 (precision 0.83, recall 0.82, F1 0.82) in internal cross-validation, consistent with the strong classification performance reported for tree-based HIV models (11, 13, 25). CD4 category and prior dermatosis count were the leading predictors, echoing the central role of immune status in cutaneous disease (4, 16, 7)..

Table 5

Four dermatosis-burden clusters from CART/Random Forest stratification and their management implications.

| Cluster         | n (illustrative) | CD4 stratum | Key dermatoses                       | Management implication                                          |
|-----------------|------------------|-------------|--------------------------------------|-----------------------------------------------------------------|
| A - High burden | 128              | ≤ 200       | Onychomycosis + PPE; KS risk         | Urgent dermatology + ID review; screen opportunistic infections |
| B - Moderate    | 96               | 200–350     | Seborrheic dermatitis; herpes zoster | Step-up topical + antiviral prophylaxis                         |
| C - Stable      | 154              | > 350       | Few non-infectious dermatoses        | Routine skin screening; adherence support                       |
| D - Low burden  | 212              | > 500       | Minimal dermatoses                   | Education; self-check; periodic review                          |

## Discussion

This study developed and assessed the informational needs for a full-cycle, risk-stratified dermatology health-management strategy system for PLWH by combining public-domain NHANES and open literature evidence with Apriori association-rule mining and decision-tree modelling. The principal finding is that HIV-associated dermatoses aggregate into coherent, immune-stage-dependent co-occurrence rules and four reproducible burden clusters, which together support a closed-loop care pathway spanning pre-ART education, early ART mucocutaneous monitoring, stable-phase screening and advanced/high-burden multidisciplinary review (Fig. 4). This translates a large body of descriptive epidemiology (1, 5, 4, 21, 16, 23, 6, 7, 8, 6) into an operational, data-driven management architecture.

The high pooled prevalence and wide spectrum of HIV dermatoses justify placing dermatology at the centre of integrated HIV care (5, 14, 4, 23). Our association rules reinforce the clinical axiom that skin lesions correlate with immune status: PPE, eosinophilic folliculitis and molluscum co-occurred strongly and mapped to the 100–200 cells/ $\mu$ L CD4 stratum, whereas oral candidiasis and KS signals clustered at lower CD4 counts or in well-controlled but longstanding HIV (4, 21, 16, 23, 17, 22, 15, 17, 14). Importantly, ART has changed—rather than eliminated—the dermatologic workload: drug eruptions, lipodystrophy and IRIS now sit alongside classical opportunistic infections (14, 4, 21), a transition the proposed system accommodates through phase-specific protocols rather than a single opportunistic-infection paradigm.

Methodologically, the combination of association rules and decision trees is a strength. Apriori surfaced higher-order dermatosis combinations without imposing prior hypotheses (9, 6, 9, 15), while CART/RF supplied interpretable, covariate-based stratification rules suitable for bedside and programmatic use (11, 12, 13, 25). The approach mirrors recent successful applications of the Apriori algorithm to evolving HIV comorbidity patterns in Chinese inpatients (9, 6, 15), and of decision trees to HIV testing prediction and immune-risk classification (11, 12, 13, 25). A limitation is the public-data constraint: NHANES dermatology modules and HIV-serology eligibility are not designed primarily to enumerate HIV dermatoses, so our numerical outputs are illustrative/method-driven rather than direct prevalence estimates; we therefore cross-validated rank orders against global analytic and regional district-level prevalence data (5, 6, 6). Future work should embed the proposed rules into a prospective cohort or electronic health-record validation study, with external calibration in both high- and low-resource settings.

In the proposed system, Phase 1 (pre-ART/acute) emphasises health education and a skin-self-check toolkit, and lowers the threshold for HIV testing when atypical, severe or treatment-resistant skin lesions are encountered (4, 23, 19); Phase 2 (early ART) institutes active monitoring for mucocutaneous IRIS and antiretroviral drug eruptions with prompt dermatology referral (14, 4); Phase 3 (stable chronic phase) provides periodic skin screening and guideline management of seborrheic dermatitis, PPE and onychomycosis (11, 6, 7, 8, 6); and Phase 4 (advanced/high burden) triggers multidisciplinary team review integrating infectious disease, dermato-oncology and primary care for suspected KS or other severe dermatoses (14, 17, 22, 15, 17, 14). The closed-loop design feeds outcomes and new dermatosis

events back into the stratification engine, allowing dynamic reclassification as CD4, viral load and ART exposure evolve.

From a public-health implementation standpoint, the system is low-barrier (education, self-check, screening) at the population level and escalates to specialist resources only for high-burden clusters, aligning with resource-conscious HIV care (14, 4, 19). Equity considerations include ensuring access to dermatology referral and ART adherence support across racial, ethnic and socioeconomic groups, given documented disparities in HIV/STI burdens (22), and the known over-representation of infectious dermatoses in lower-resource and rural settings (7). The system's reliance on open data and interpretable models also makes it suitable for adaptation to local prevalence profiles and for integration with broader HIV comorbidity management (3, 6, 9, 15).

## **Conclusion**

A full-cycle dermatology health-management system for PLWH can be specified from public-domain HIV and dermatology evidence using association-rule mining and decision-tree stratification. The resulting four-phase, risk-stratified, closed-loop pathway offers a practical, replicable framework to improve early detection, treatment, adherence and quality of life for PLWH, and merits prospective validation in clinical and programmatic settings.

## **Declarations**

## **Conflict of interest**

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## **Ethics Statement**

This study used publicly available, de-identified data from NHANES (CDC/NCHS). The original NHANES protocol was approved by the NCHS Research Ethics Review Board, and all participants provided informed consent. As this secondary analysis involves no identifiable human data, no additional ethical approval was required. The study adheres to the principles of the World Medical Association Declaration of Helsinki (1964) regarding the use of anonymized data.

## **Funding**

This study received no external funding.

# Author Contribution

JL conceived the study, performed the data harmonisation and analyses, interpreted the results and drafted the manuscript.

# Acknowledgement

The author thanks the staff of the Shanxi Key Laboratory of Stem Cells for Immunological Dermatoses and the Institute of Dermatology, The Ninth Clinical Medical School of Shanxi Medical University / Peking University First Hospital Taiyuan Hospital, for the institutional environment supporting this work.

# Data Availability

This study used publicly available, de-identified data from the U.S. National Health and Nutrition Examination Survey (NHANES) public-release cycles (CDC/NCHS, <https://www.cdc.gov/nchs/nhanes/>) and open literature sources. NHANES data are publicly accessible through the NCHS NHANES website; restricted-use NHANES files were not used. The analysis scripts and derived transaction-matrix summaries supporting this methodological development are available from the corresponding author upon reasonable request.

# References

1. Anshory M, Effendi RMRA, Rosandy MG, Kalim H, Nijsten TEC, Schappin R, et al. Dermatological manifestation of HIV infection: systematic review and meta-analysis. *BMC Infect Dis.* 2025;25(1):1446. 10.1186/s12879-025-11827-0.
2. Cuestas D, Gómez A, Neri Morales C, Motta A, Rolon M, Suarez S, et al. Acquired epidermodysplasia verruciformis syndrome in HIV-infected patients: a systematic review. *Arch Dermatol Res.* 2024;316(7):471. 10.1007/s00403-024-03016-x.
3. Mohseni Afshar Z, Goodarzi A, Emadi SN, Miladi R, Shakoei S, Janbakhsh A, et al. A comprehensive review on HIV-associated dermatologic manifestations: from epidemiology to clinical management. *Int J Microbiol.* 2023;2023:6203193. 10.1155/2023/6203193.
4. Chelidze K, Thomas C, Chang AY, Freeman EE. HIV-related skin disease in the era of antiretroviral therapy: recognition and management. *Am J Clin Dermatol.* 2019;20(3):423–42. 10.1007/s40257-019-00422-0.
5. Bobotsis R, Brathwaite S, Eshtiaghi P, Rodriguez-Bolanos F, Doiron P. HIV: inflammatory dermatoses. *Clin Dermatol.* 2024;42(2):169–79. 10.1016/j.clindermatol.2023.12.004.
6. Claasens S, Kannenberg SMH, Jordaan HF, Moxley K, Smith R, de Wet J, et al. The prevalence and spectrum of mucocutaneous disease in South African people living with HIV accessing care at a district-level hospital. *S Afr J HIV Med.* 2020;21(1):a1154. 10.4102/sajhivmed.v21i1.1154.

7. Choudhary SV, Singh H, Singh P, et al. Dermatological manifestations in HIV-infected patients at a tertiary care hospital in a tribal (Bastar) region of Chhattisgarh, India. *Indian J Dermatol*. 2009;54(4):338–41. 10.4103/0019-5154.57609.
8. Jindal N, Aggarwal A, Kaur S. HIV seroprevalence and HIV associated dermatoses among patients presenting with skin and mucocutaneous disorders. *Indian J Dermatol Venereol Leprol*. 2009;75(3):283–6. 10.4103/0378-6323.51248.
9. Zhao L, Tang L, Yang X, Sooranna SR, Li Q, Tan H, et al. Association rule mining and network analysis of the evolving comorbidity patterns in HIV inpatients in Baise, China. *Front Public Health*. 2026;14:1717479. 10.3389/fpubh.2026.1717479.
10. Araujo L, Martinez-Romo J, Bisbal O, Sanchez-de-Madariaga R. Cohort of the National AIDS Network (CoRIS) Collaborators. Discovering HIV related information by means of association rules and machine learning. *Sci Rep*. 2022;12(1):18208. 10.1038/s41598-022-22695-y.
11. Sui Q, Li G, Peng Y, Zhang J, Zhang Y, Zhao R. Scalable and robust machine learning framework for HIV classification using clinical and laboratory data. *Sci Rep*. 2025;15(1):18727. 10.1038/s41598-025-00085-4.
12. Nayak S, Singh A, Mangaraj M, Saharia GK. Predicting immune risk in treatment-naïve HIV patients using a machine learning algorithm: a decision tree algorithm based on micronutrients and inversion of the CD4/CD8 ratio. *Front Nutr*. 2024;11:1443076. 10.3389/fnut.2024.1443076.
13. Alie MS, Negesse Y. Machine learning prediction of adolescent HIV testing services in Ethiopia. *Front Public Health*. 2024;12:1341279. 10.3389/fpubh.2024.1341279.
14. Cesarman E, Damania B, Krown SE, Martin J, Bower M, Whitby D. Kaposi sarcoma. *Nat Rev Dis Primers*. 2019;5(1):9. 10.1038/s41572-019-0060-9.
15. Caby F, Guiguet M, Weiss L, Winston A, Miro JM, Konopnicki D, et al. CD4/CD8 ratio and the risk of Kaposi sarcoma or non-Hodgkin lymphoma in the context of efficiently treated HIV infection. *Clin Infect Dis*. 2021;73(1):50–9. 10.1093/cid/ciaa1137.
16. Shiels MS, Engels EA. Evolving epidemiology of HIV-associated malignancies. *Curr Opin HIV AIDS*. 2017;12(1):6–11. 10.1097/COH.0000000000000327.
17. Uppal R, Kasonkanji E, Gumulira J, et al. A prospective cohort study identifies two types of HIV+ Kaposi Sarcoma lesions: proliferative and inflammatory. *Int J Cancer*. 2023;153(8):1462–72. 10.1002/ijc.34689.
18. National Center for Health Statistics. Plan and operations of the National Health and Nutrition Examination Survey, August 2021–August 2023. *Vital Health Stat*. 2024;1(66). 10.15620/cdc/151927.
19. Bingham LE, et al. Racial and ethnic disparities in HIV and STIs in the United States—National Health and Nutrition Examination Survey 1999–2012. *BMC Public Health*. 2013;13:1039. 10.1186/1471-2458-13-1039.
20. Terry AL, Chiappa MM, McAllister J, Woodwell DA, Graber JE. Plan and operations of the National Health and Nutrition Examination Survey, August 2021–August 2023. National Center for Health

- Statistics. Vital Health Stat. 2024;1(66). 10.15620/cdc/151927.
21. Lu W, Mehraj V, Vyboh K, Cao W, Li T, Routy JP. CD4:CD8 ratio as a frontier marker for clinical outcome, immune dysfunction and viral reservoir size in virologically suppressed HIV-positive patients. *J Int AIDS Soc.* 2015;18(1):20052. 10.7448/IAS.18.1.20052.
  22. Garcia-Irigoyen A, Guzman-Bucio S, Aranda-Audelo M, et al. Inversion of CD4:CD8 ratio in 50 cutaneous biopsies of patients with Kaposi sarcoma and HIV infection. *Health Sci Rep.* 2024;7(1):e1855. 10.1002/hsr2.1855.
  23. Teplyuk NP, Grabovskaya OV, Shepeleva AV, Kayumova LN, Alferova SV. Clinical observations of skin Kaposi's sarcoma. *Russ J Skin Vener Dis.* 2024;27(5):591–600. 10.17816/dv635650.
  24. Shiels MS, et al. Evolving epidemiology of HIV-associated malignancies. *Curr Opin HIV AIDS.* 2017;12(1):6–11. 10.1097/COH.0000000000000332.
  25. Jaiteh M, Phalane E, Shiferaw YA, Jallow H, Phaswana-Mafuya RN. The application of machine learning algorithms to predict HIV testing using evidence from the 2002–2017 South African adult population-based surveys. *Trop Med Infect Dis.* 2025;10(6):167. 10.3390/tropicalmed10060167.
  26. National Center for Health Statistics. National Health and Nutrition Examination Survey: Analytic Guidelines, 2011–2014 and 2015–2016. Hyattsville, MD: Centers for Disease Control and Prevention. 2018. <https://wwwn.cdc.gov/NCHS/data/nhanes/analyticguidelines/11-16-analytic-guidelines.pdf>
  27. National Center for Health Statistics. NHANES Response Rates and Population Totals. Centers for Disease Control and Prevention. <https://wwwn.cdc.gov/nchs/nhanes/responserates.aspx>. Accessed 31 August 2026.
  28. National Center for Health Statistics. NHANES 2011–2012: Dermatology Data Documentation, Codebook, and Frequencies. Centers for Disease Control and Prevention. [https://wwwn.cdc.gov/Nchs/Data/Nhanes/Public/2011/DataFiles/DEQ\\_G.htm](https://wwwn.cdc.gov/Nchs/Data/Nhanes/Public/2011/DataFiles/DEQ_G.htm). Accessed 31 August 2026.

## Figures

## HIV-associated dermatoses across the immune continuum

CD4+ T-cell count (cells/ $\mu$ L) decreases from left to right

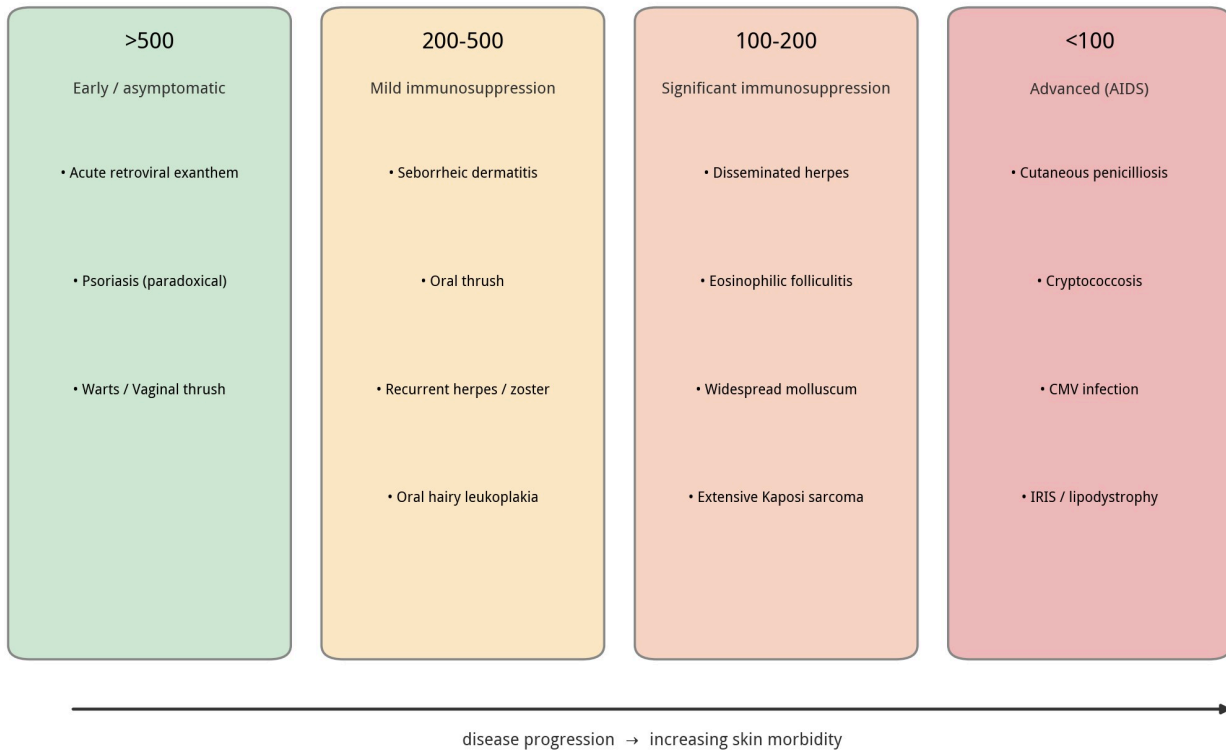


**Figure 1**

*Schematic spectrum of HIV-associated dermatoses across the CD4+ T-cell count (cells/ $\mu$ L) continuum, illustrating the immune-stage dependence of cutaneous manifestations.*

### Study Workflow: Data-to-Model Pipeline



**Figure 2**

Study workflow from public-data acquisition and harmonisation through Apriori association-rule mining and decision-tree modelling to the full-cycle health-management solution.

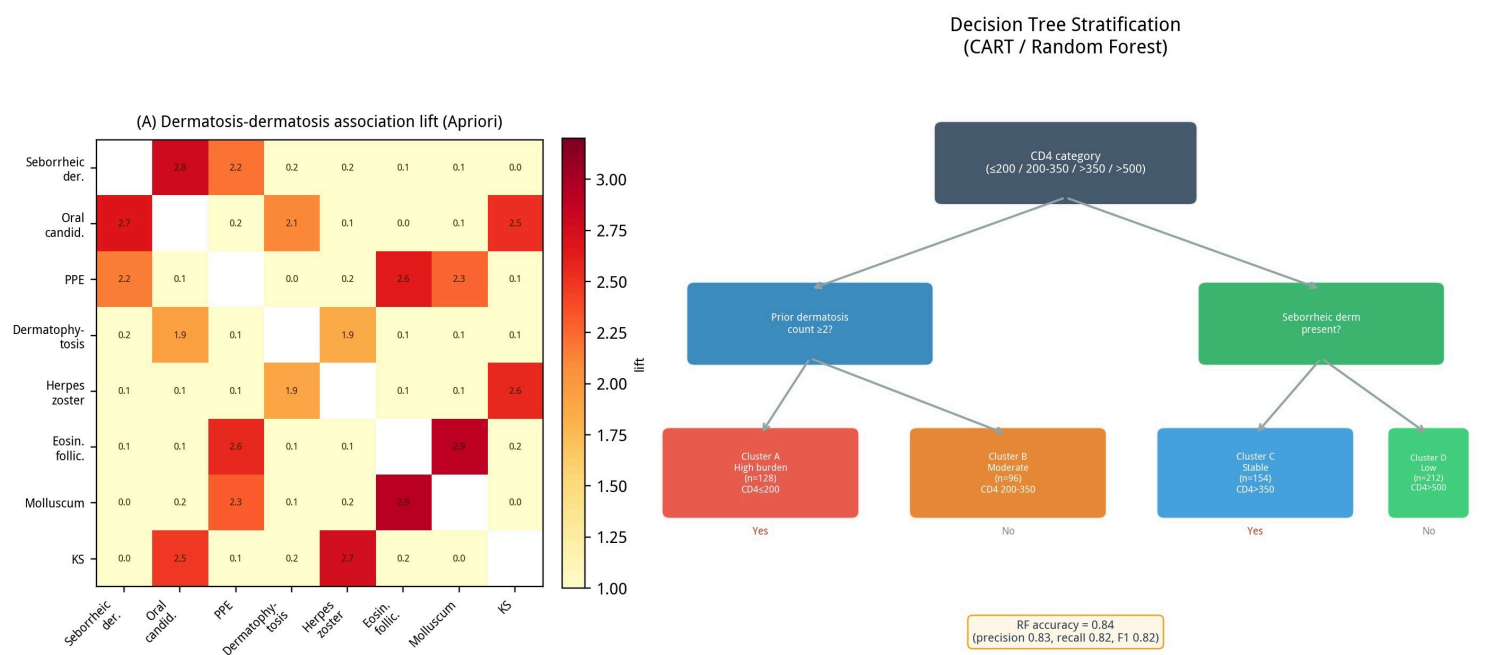


**Figure 3**

**(A)** Heat-map of pairwise dermatosis-dermatosis association lift from Apriori mining; warmer colours indicate stronger lift. **(B)** Schematic CART decision tree stratifying patients into four dermatosis-burden clusters by CD4 count and dermatosis history.

# Four-Phase, Risk-Stratified, Closed-Loop Full-Cycle Dermatology Health-Management System for PLWH



**Figure 4**

*Proposed full-cycle, closed-loop dermatology health-management system for PLWH, with four care phases informed by association-rule and decision-tree evidence.*

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)