

# MULTIMODAL ARTIFICIAL INTELLIGENCE FOR PREDICTING HAIR FOLLICLE SURVIVAL AND AESTHETIC OUTCOMES FOLLOWING HAIR TRANSPLANTATION SURGERY

Dr Jawad Jahangir<sup>\*1</sup>, Ali Raza<sup>2</sup>

<sup>\*1</sup>JJ Aesthetics, Pakistan

<sup>2</sup>Department of Computer Science and Network Security, Chengdu University of Technology, Sichuan, china

<sup>\*1</sup>j.jabbasi@hotmail.com, <sup>2</sup>alir99008@gmail.com

Corresponding Author: \*

Dr Jawad Jahangir

DOI: <https://doi.org/10.5281/zenodo.23389167>

| Received        | Accepted      | Published     |
|-----------------|---------------|---------------|
| 01 January 2026 | 17 March 2026 | 31 March 2026 |

## ABSTRACT

**Background:** Artificial intelligence (AI) has considerable potential to support outcome assessment in hair transplantation; however, its clinical reliability depends on the availability of complete, standardized, and appropriately linked patient-level data. Despite growing interest in AI-assisted aesthetic and clinical assessment, the readiness of currently reported hair-transplantation data for predictive and multimodal modelling remains insufficiently characterized.

**Objective:** This study aimed to quantitatively assess the completeness, consistency, and multimodal readiness of literature-derived hair-transplantation data, with particular emphasis on hair survival, patient satisfaction, follow-up reporting, and the availability of clinical and trichoscopic imaging features.

**Methods:** A descriptive data-readiness assessment was conducted using 36 literature-derived records extracted from three published studies. An additional 1,000 simulated records were maintained separately for technical testing and excluded from all clinical summaries. Data completeness, variable availability, outcome distributions, and reporting inconsistencies were examined using descriptive statistics. No missing values were imputed, and no predictive model was trained or validated. The analysis therefore focused on characterizing the strengths and limitations of the available data for future AI development.

**Results:** Numerical hair-survival values were available for 26 of the 36 literature-derived records (72.2%). Among these records, mean hair survival was 85.30% (standard deviation [SD], 10.60), with a median of 85.90% and a range of 45.1%–96.8%. Satisfaction information was available for 27 records, comprising 17 classified as mostly satisfied, seven as very satisfied, and three as dissatisfied. Although 24 of these 27 records carried favorable satisfaction labels (88.9%), differences in reporting and the absence of a standardized assessment instrument limit direct clinical interpretation. Follow-up information was recorded for 27 records, but some entries represented study-level reporting intervals rather than verified patient-specific assessments. Most importantly, no original clinical or trichoscopic images were available in the analyzed dataset, and key imaging-derived features including vascularity, fibrosis, and trichoscopic characteristics were absent. These limitations restrict outcome comparability and prevent the construction of a genuinely linked multimodal dataset from the available records.

**Conclusion:** This study identifies a substantial data-readiness gap between the potential of AI in hair transplantation and the structure of currently available literature-derived data. Although reported hair-survival values provide a preliminary descriptive benchmark, incomplete outcome reporting, inconsistent assessment definitions, uncertain patient-level follow-up alignment, and absent linked imaging data constrain robust predictive modelling. Progress toward clinically credible AI systems will require standardized outcome definitions, source-verified patient-level data, harmonized follow-up schedules, and prospectively linked clinical and trichoscopic images. The findings provide a practical basis for designing future datasets and validation studies before predictive AI claims can be justified.

**Keywords:**

Hair transplantation, artificial intelligence, data readiness, hair follicle survival, patient satisfaction, trichoscopy, multimodal data, clinical outcome assessment.

## 1. INTRODUCTION

Hair transplantation has become an important reconstructive and aesthetic procedure for individuals experiencing permanent hair loss due to androgenetic alopecia, traumatic injuries, burns, surgical scars, and cicatricial alopecia. The procedure involves transferring follicular units from a donor region to a recipient area to improve hair coverage, density, and appearance. Although advances in follicular unit extraction (FUE) and follicular unit transplantation (FUT) have improved surgical techniques, postoperative outcomes remain variable. Factors such as donor-hair quality, recipient-site characteristics, graft handling, vascular supply, scar tissue, surgical technique, and postoperative care may influence follicular survival and aesthetic results. Consequently, evaluating transplantation success requires a comprehensive assessment of clinical, surgical, and patient-centred outcomes rather than relying solely on the number of grafts implanted. Cicatricial alopecia represents a particularly challenging clinical condition because permanent follicular destruction and fibrosis may compromise the recipient environment. Hair transplantation can provide cosmetic improvement in appropriately selected patients, but its effectiveness depends on the underlying cause of hair loss, disease stability, scar characteristics, and the condition of the recipient tissue. Kwon et al. (2007) investigated staged hair transplantation following carbon dioxide laser-assisted scar-tissue remodelling, demonstrating the relevance of recipient-site preparation in reconstructive procedures. Liu et al. (2018)

reported two cases involving hair transplantation for lichen planopilaris and frontal fibrosing alopecia, emphasizing the importance of disease stability and careful patient selection. Similarly, Madura et al. (2020) evaluated hair transplantation outcomes in patients with cicatricial alopecia and highlighted the importance of postoperative assessment and follow-up. These studies indicate that transplantation outcomes depend on multiple interacting factors and may not be adequately explained by surgical technique alone.

Traditional assessment of hair transplantation outcomes relies on clinical examination, photographic comparisons, trichoscopy, graft counts, hair-density measurements, clinician assessments, and patient-reported satisfaction. Although these methods provide valuable clinical information, their reliability may be affected by differences in assessment procedures, follow-up duration, image quality, and outcome definitions. Furthermore, follicular survival and aesthetic success are related but distinct outcomes. A patient may demonstrate satisfactory hair growth but remain dissatisfied with hairline design, density distribution, symmetry, or the natural appearance of the transplanted hair. Therefore, comprehensive outcome assessment should distinguish hair follicle survival, hair density, global aesthetic improvement, complications, and patient satisfaction. Artificial intelligence (AI) and machine learning have introduced new opportunities for improving the measurement and interpretation of clinical outcomes. Computer-vision techniques can identify

follicular structures, estimate hair density, and quantify image characteristics, while statistical and machine-learning methods can investigate relationships between clinical characteristics and postoperative results. Urban et al. (2021) investigated the combination of optical coherence tomography and deep learning for estimating scalp hair and follicle counts. Their research demonstrated the potential of automated image analysis for quantitative hair assessment. Similarly, subsequent studies have investigated deep-learning algorithms for automated hair-density measurement and trichoscopic image classification. However, automated follicle detection and alopecia classification do not necessarily establish whether transplanted follicles will survive or whether patients will achieve satisfactory aesthetic outcomes. This distinction creates an important opportunity for research into clinically meaningful predictive models.

A multimodal AI framework could potentially integrate demographic characteristics, alopecia etiology, defect size, donor- and recipient-site assessments, surgical techniques, graft counts, postoperative follow-up, clinical photographs, and trichoscopic images. Combining these information sources may provide a more comprehensive representation of transplantation outcomes than using a single data modality. Nevertheless, the benefits of multimodal integration must be demonstrated empirically through appropriate model development and validation. Future models should be compared with clinical-variable-only approaches and assessed using suitable performance measures, calibration analysis, subgroup evaluation, and independent validation. Sufficient patient-level observations and consistently defined outcomes are also essential for determining whether the resulting predictions are reliable and clinically useful. Despite advances in hair transplantation, trichoscopy, and medical AI, important methodological challenges remain. Existing clinical studies frequently contain small samples, heterogeneous patient populations, inconsistent outcome definitions, incomplete follow-up, and limited standardization. Moreover, many AI

studies in dermatology focus on diagnosis or image-based measurement rather than predicting long-term surgical outcomes. The lack of standardized datasets linking clinical characteristics, surgical information, actual clinical images, and postoperative outcomes limits the development of reliable multimodal prediction systems. These limitations motivate the need for an evidence-informed framework that identifies relevant predictors, establishes consistent outcome definitions, and specifies the requirements for future clinical model development.

Accordingly, this study, entitled Multimodal Artificial Intelligence for Predicting Hair Follicle Survival and Aesthetic Outcomes Following Hair Transplantation, examines published clinical evidence and evaluates its suitability for future predictive modelling. The study focuses on the provenance of extracted records, the availability of demographic and surgical characteristics, the completeness of hair-survival and aesthetic outcomes, and the availability of clinical and trichoscopic imaging variables. It also proposes a conceptual multimodal AI framework emphasizing traceable data sources, standardized outcome measurement, transparent data processing, and appropriate validation. The study is intended to establish an evidence-informed foundation for future model development rather than claim that a clinically validated multimodal prediction system has already been established.

## 2. Literature review

### 2.1 Hair transplantation and follicular survival

Kwon et al. (2007) investigated staged hair transplantation in cicatricial alopecia following carbon dioxide laser-assisted scar-tissue remodelling. Their study is relevant to reconstructive hair transplantation because it considers the condition of scarred recipient tissue and the potential role of preparatory treatment. The research illustrates that transplantation outcomes may depend on the biological characteristics of the recipient area as well as the surgical procedure. However, the reported clinical experience does not establish a generalizable prediction model for estimating

graft survival across different patients. Liu et al. (2018) reported two cases of hair transplantation for lichen planopilaris and frontal fibrosing alopecia. Their observations indicated that transplantation may provide satisfactory cosmetic outcomes in carefully selected patients after disease stabilization. The study highlights the importance of evaluating the underlying inflammatory condition before transplantation and monitoring patients after surgery. Nevertheless, the limited number of patients restricts the generalizability of the findings and does not allow the reliable estimation of relationships between multiple clinical predictors and postoperative outcomes.

Madura et al. (2020) evaluated hair transplantation outcomes among patients with cicatricial alopecia. Their prospective clinical investigation provides information about the assessment of reconstructed areas and the importance of postoperative follow-up. The study also illustrates the challenges created by incomplete follow-up, as some patients did not contribute outcome information at the final assessment. Such missingness is important for future predictive research because incomplete outcomes can introduce bias and reduce the reliability of estimated relationships between patient characteristics and treatment success. Wang et al. (2023) investigated autologous hair transplantation in children with cicatricial alopecia. Their work contributes evidence concerning transplantation in a specific patient population and highlights the importance of considering age, underlying injury, clinical presentation, and surgical context when interpreting treatment outcomes. The study is particularly relevant to the present research because it contributes published clinical records to the assembled evidence dataset. However, individual follow-up measurements and outcome values must be extracted and verified against the original publication before they can be used in quantitative analyses.

Baranov and Nor (2026) examined hair transplantation techniques for secondary cicatricial alopecia. Their report contributes clinical information concerning scar-related hair

loss arising from different causes, including traumatic and other injuries. The publication is represented in the dataset used in the present study and provides information about patient characteristics and reported transplantation outcomes. However, discrepancies between source tables and narrative descriptions must be checked before the extracted observations are used to support clinical conclusions. The study is useful for evidence mapping but does not independently establish the accuracy of an AI prediction model. A systematic review published in 2019 examined transplantation in primary scarring alopecia and included a small body of evidence consisting largely of individual clinical reports. The review described both favourable and unfavourable outcomes, including cases involving disease recurrence. Its findings demonstrate that transplantation results may vary according to disease subtype, clinical selection, and follow-up. The limited size and observational nature of the underlying evidence also suggest that larger and more standardized studies are required to identify reliable predictors of long-term success.

Research published in *The Laryngoscope* in 2021 reviewed hair transplantation in patients with frontal fibrosing alopecia and lichen planopilaris. The work considered outcomes in patients receiving transplantation and the relationship between transplantation and inflammatory alopecia. Its clinical implications emphasize the importance of disease stability and longitudinal assessment. These findings are relevant to the proposed AI framework because the underlying cause of alopecia may influence both patient selection and subsequent outcomes. However, evidence from these specific disorders should not automatically be generalized to all transplantation indications. An observational study published in 2021 examined the longevity of transplanted hair by comparing photographic assessments at different postoperative time points. The researchers observed that transplanted-hair density could decline over time, indicating that early postoperative results may not fully represent longer-term outcomes. This observation is important for predictive research because the

choice of follow-up period directly affects the interpretation of follicular survival. Models developed for short-term outcomes may not necessarily predict long-term graft persistence, making clearly defined assessment times essential. A review published in the Journal of Clinical Medicine in 2025 examined hair transplantation in primary cicatricial alopecia. The review discussed variation in clinical outcomes across disease subtypes and emphasized the challenges associated with patient selection, disease recurrence, and the quality of the available evidence. The findings support the need to consider the cause of alopecia and the characteristics of the recipient area when investigating transplantation outcomes. They also suggest that combining heterogeneous clinical populations without accounting for these differences may produce misleading conclusions. Yii et al. (2025) conducted a systematic review of follicular-unit graft survival in primary cicatricial alopecia. Their review included observational studies and examined survival outcomes across different follow-up intervals. The reported findings suggest that outcome estimates may vary with the duration of postoperative assessment. This research is relevant because it demonstrates why survival should be defined at a specific time point and why estimates from different studies should not automatically be combined. It also highlights the need for consistent outcome measurement in future clinical datasets.

Rezende, Leal, and Dias (2026) published a systematic review concerning follicular-unit graft survival after transplantation in primary cicatricial alopecia. Their work contributes to the growing synthesis of clinical evidence in this field and reinforces the importance of evaluating survival outcomes over time. When interpreting this and other systematic reviews, researchers must consider whether the underlying studies overlap. Adding the patient totals from overlapping reviews could lead to double-counting and an inaccurate representation of the total evidence base. These reviews provide important clinical context but do not replace independent patient-level data for prediction-model development. Zhao and Zhang and their

colleagues (2026) reported a Chinese expert-consensus approach to hair transplantation for secondary cicatricial alopecia. Expert-consensus research can help organize clinical knowledge concerning diagnosis, disease stability, scar assessment, surgical planning, and postoperative management. Such recommendations can inform the selection of potential predictors for a future AI model. Nevertheless, consensus recommendations represent clinical guidance rather than empirical evidence that a particular variable will improve prediction accuracy. Candidate predictors should therefore be evaluated using appropriately collected clinical observations.

## 2.2 Trichoscopy and automated hair assessment

Urban et al. (2021) investigated the use of deep learning combined with optical coherence tomography to estimate scalp hair and follicle counts. Their study demonstrated that automated image analysis could support quantitative assessment of follicular structures. This work is relevant to hair transplantation because objective measurements may help characterize baseline scalp conditions and postoperative hair growth. However, estimating follicle counts from images is not equivalent to determining the long-term survival of transplanted grafts, and further research is required to establish the prognostic value of such measurements. A 2022 study published in *Sensors* investigated deep-learning methods for automated hair-density measurement using enlarged scalp images. The researchers compared different object-detection approaches, including EfficientDet, YOLOv4, and DetectoRS. Their work demonstrates the potential of computer vision to automate the identification and measurement of hair structures. However, the performance of an image-detection algorithm depends on the quality and characteristics of the images used for development and evaluation. Furthermore, successful hair detection does not by itself establish that a model can predict postoperative graft survival or patient satisfaction.

Research published in 2023 investigated machine learning for classifying androgenetic alopecia and

estimating its severity from trichoscopic measurements. The study explored the use of structured image-derived characteristics, such as follicular-unit patterns and hair characteristics, to support classification. This research is relevant because it illustrates how trichoscopic observations can be transformed into measurable predictors. Nevertheless, diagnostic classification and severity estimation represent different objectives from postoperative prognosis. A future transplantation model would need to determine whether these features contribute useful information when predicting defined surgical outcomes. Traini et al. (2026) investigated the use of AI and trichoscopy to distinguish between alopecia areata and lichen planopilaris. Such research illustrates the potential of image-based systems to support the identification of different hair-loss conditions. Accurate classification may help organize clinical information and inform decisions about further assessment. However, a diagnostic model does not necessarily predict the response to surgical treatment. The relationship between diagnostic features and transplantation outcomes must be investigated separately using appropriate longitudinal clinical data.

Lee et al. (2025) discussed the potential and limitations of AI-assisted trichoscopy in alopecia diagnosis. Their work highlights the opportunities created by automated image interpretation while emphasizing the importance of cautious clinical application. Trichoscopic images may contain useful information about follicular openings, hair-shaft characteristics, and scalp conditions, but image quality, acquisition methods, and interpretation can affect the reliability of extracted features. For future transplantation research, these limitations reinforce the need for standardized image acquisition and expert-reviewed annotations. Kearney et al. (2025) examined patient perspectives on AI-driven trichoscopy for alopecia evaluation. Their work expands the discussion beyond technical performance to include the acceptability and patient-centred implications of AI-supported assessment. This is relevant to transplantation because the ultimate objective is not simply to generate a numerical prediction but

to support meaningful clinical decisions and patient expectations. Future research should therefore consider how AI-generated information can be communicated responsibly and incorporated into shared decision-making.

### 2.3 Artificial intelligence in dermatology and medical imaging

Esteva et al. (2017) demonstrated the potential of deep learning for classifying skin cancer using clinical images. Their research contributed to the development of AI-based image analysis in dermatology and showed how large image datasets can support automated classification. The study provides a methodological foundation for considering computer vision in hair and scalp research. However, skin-cancer classification is a different clinical task from predicting follicular survival, and its performance cannot be directly transferred to hair transplantation. Haenssle et al. (2018) compared a deep-learning convolutional neural network with dermatologists in dermoscopic melanoma recognition. Their research illustrates how algorithmic predictions can be evaluated against clinical expertise. For transplantation research, the methodological implication is that AI should be compared with appropriate clinical assessment rather than evaluated solely through technical metrics. The relevant outcome should also reflect the intended clinical purpose, such as predicting a defined survival percentage or an established aesthetic score.

Tschandl et al. (2019) investigated human-computer collaboration in skin-lesion classification. Their work is relevant to the development of clinical decision-support systems because it considers the interaction between algorithmic predictions and professional judgment. A future hair-transplantation system should similarly be assessed in terms of whether it improves clinical assessment, supports more consistent outcome estimation, or helps identify patients who require additional evaluation. The value of AI should not be assumed from standalone model performance alone. Liu et al. (2020) investigated deep learning for the diagnosis of skin diseases using clinical images.

This research contributes to the broader evidence that AI can assist in interpreting dermatological images. However, image-based diagnosis does not require the same data structure as longitudinal outcome prediction. Predicting transplantation success requires patient-level information linked to surgical procedures and subsequent outcomes, making follow-up and data linkage essential components of model development.

Topol (2019) discussed the potential of human-centred AI in healthcare and emphasized the importance of integrating computational capabilities with clinical expertise. This perspective is relevant to hair transplantation because surgical decisions involve clinical examination, patient preferences, donor-site considerations, and realistic expectations about aesthetic outcomes. An AI framework should therefore support clinical judgment rather than replace it, particularly when evidence is limited or predictions are uncertain. Kelly et al. (2019) examined the challenges associated with translating AI research into clinical impact. Their work highlighted the importance of data quality, validation, workflow integration, and evidence of real-world usefulness. These considerations are particularly relevant to the current study, which has not yet established a validated predictive model. Before any clinical application, future research must demonstrate that the model performs reliably on new patients and provides useful information beyond existing clinical assessment.

#### **2.4 Prediction-model methodology and validation**

Collins et al. (2015) introduced the TRIPOD statement to improve the transparent reporting of multivariable prediction models for individual prognosis or diagnosis. Their guidance emphasizes the importance of documenting participants, predictors, outcomes, statistical methods, and model performance. This framework is relevant to future hair-transplantation research because transparent reporting allows other researchers to assess the reliability and reproducibility of model development. A future study should clearly

distinguish model training, internal validation, and external evaluation. Wolff et al. (2019) introduced PROBAST, a tool for evaluating the risk of bias and applicability of prediction-model studies. The tool focuses on participants, predictors, outcomes, and statistical analysis. These domains are particularly relevant to transplantation research because small samples, missing observations, inconsistent outcome measurement, and inappropriate model complexity may distort estimated performance. Applying an appropriate risk-of-bias assessment can help researchers identify weaknesses before a model is considered suitable for clinical use.

Page et al. (2021) published the PRISMA 2020 statement, which provides guidance for transparent reporting of systematic reviews. The statement emphasizes documenting the literature search, study selection, exclusions, and synthesis methods. This guidance is important for the present study because a comprehensive review requires a reproducible search strategy and a documented screening process. The existing manuscript does not contain a complete search and screening log; therefore, its literature-derived analysis should not be described as a completed PRISMA-compliant systematic review until these procedures have been completed. van Calster et al. (2019) discussed the importance of calibration in risk prediction. Calibration evaluates whether predicted probabilities agree with observed outcome frequencies. A model may discriminate between patients with higher and lower predicted risks while still producing inaccurate absolute probabilities. In hair transplantation, this distinction matters because clinicians and patients may interpret a predicted probability of successful graft survival as a quantitative estimate of the expected outcome. Future models should therefore assess calibration alongside discrimination and other relevant performance measures.

Riley and colleagues (2020) developed methodological guidance concerning sample-size requirements for clinical prediction-model development. Their research highlights the importance of considering the number of participants, the number of outcome events,

candidate predictor complexity, and anticipated model performance. These principles are relevant to the present dataset because a small number of clinical records cannot support unrestricted development of a complex multimodal model. Future research should determine sample-size requirements before model training and avoid treating simulated observations as a substitute for genuine clinical data. Steyerberg and colleagues have contributed extensively to the development and evaluation of clinical prediction models. Their work emphasizes the importance of appropriate model specification, validation, calibration, and evaluation on data beyond the development sample. For hair transplantation, these principles imply that future multimodal models should be compared with simpler clinical baselines and evaluated using independent patients or centres. Such comparisons would help establish whether image integration produces a genuine improvement in predictive performance. Although published studies have described hair survival, aesthetic improvement, and patient satisfaction following hair transplantation, these outcomes are not necessarily measured using equivalent definitions, assessment procedures, or follow-up schedules. Combining results from heterogeneous studies therefore requires careful consideration of source provenance, outcome comparability, and the level at which each observation was reported. A further challenge is the distinction between data availability and genuine multimodal data integration. A dataset may contain columns labelled as image-derived features without containing the original images or documented extraction procedures required to establish the validity of those features. Similarly, simulated records can facilitate software testing but cannot replace observed clinical data or establish clinical predictive performance. The principal gap addressed in this study is the lack of a systematic assessment of whether the supplied dataset contains sufficiently complete, traceable, and appropriately linked information to support future multimodal AI research in hair

transplantation. Specifically, the study examines the availability of clinical predictors and postoperative outcomes, identifies inconsistencies in outcome reporting and missingness, and evaluates whether original image data are available for integration with clinical records. The contribution is methodological rather than algorithmic. First, the study characterizes the composition and completeness of the literature-derived records while keeping simulated records separate. Second, it identifies limitations that affect the interpretation of survival, satisfaction, and aesthetic-outcome information. Third, it proposes a staged framework for improving data quality and establishing the prerequisites for future clinical prediction research. The findings should be interpreted as a preliminary data-readiness assessment rather than as a systematic review, meta-analysis, or validated AI study.

## Materials and Methods

### 3.1 Study design

This study was designed as a literature-derived, retrospective observational evidence analysis with a prespecified conceptual framework for future multimodal artificial-intelligence (AI) development. The available workbook contains 36 patient-level records extracted from published reports and 1,000 simulated records. The published and simulated strata were retained separately because simulated observations cannot be interpreted as real clinical evidence. The primary clinical description is therefore restricted to the 36 published records; the simulated records are relevant only to software testing or clearly labelled methodological demonstrations. The study has four methodological objectives: (i) establish traceable provenance for each extracted record; (ii) describe the reported demographic, etiologic, surgical, and outcome variables; (iii) quantify reporting completeness and identify between-study heterogeneity; and (iv) specify a reproducible multimodal AI framework without claiming that a clinically validated model has already been developed.



**Figure 1: Evidence workflow from study identification to descriptive analysis and future AI framework design.**

### 3.2 Data sources and dataset composition

The published-case workbook contains records extracted from three sources: Baranov and Nor (2026), 14 cases; Wang et al. (2023), 13 cases; and Kwon et al. (2007), 9 cases. These sources concern cicatricial alopecia and differ in patient

populations, clinical context, surgical approach, follow-up reporting, and availability of outcome measures. They were not treated as a single prospectively recruited cohort. The workbook additionally contains 1,000 simulated rows, which remain labelled as simulated.

| Data stratum         | Records | Role in this study           | Interpretation                                                                       |
|----------------------|---------|------------------------------|--------------------------------------------------------------------------------------|
| Baranov & Nor (2026) | 14      | Published evidence           | Transcribed from a published case table; source discrepancies require checking       |
| Wang et al. (2023)   | 13      | Published evidence           | Study-level follow-up range must not be converted into invented individual follow-up |
| Kwon et al. (2007)   | 9       | Published evidence           | Unreported patient-level graft/survival values remain missing                        |
| Simulated dataset    | 1,000   | Technical demonstration only | Not real patients; excluded from clinical outcome estimates                          |
| Total workbook       | 1,036   | Mixed provenance             | Analyses must identify which stratum is being used                                   |

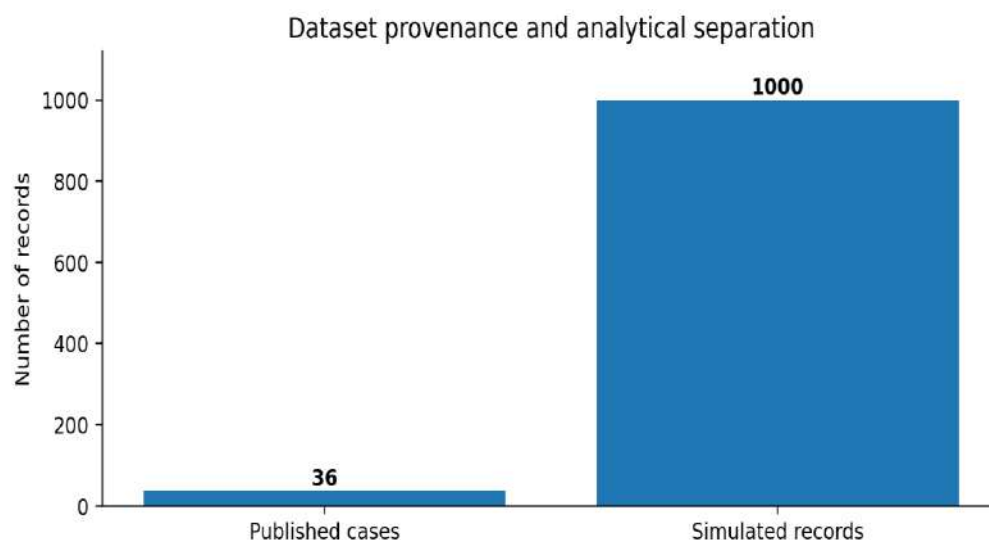


**Figure 2: Provenance of the workbook records. Simulated rows are not clinical observations**

### 3.3 Literature search strategy

For a fully reproducible review, searches should be run in PubMed/MEDLINE, Scopus, Web of Science, and Google Scholar, depending on access. Search terms should cover hair transplantation, cicatricial/scarring alopecia, survival and aesthetic outcomes, trichoscopy, and AI. A candidate strategy is: ("hair transplantation" OR "hair restoration" OR "follicular unit transplantation") AND ("cicatricial alopecia" OR scar\* OR scarring) AND (survival OR outcome\* OR satisfaction OR trichoscop\* OR "artificial intelligence" OR "machine learning"). The final

manuscript should report the exact search date, platform, complete search string, number of records retrieved, duplicates removed, full texts assessed, and reasons for exclusion.

The workbook alone does not provide a complete search log or screening history. Therefore, this chapter does not invent database hit counts or a PRISMA flow count. If the search has not yet been conducted and documented, the manuscript should be described as a literature-derived evidence analysis rather than a completed systematic review until those steps are finished.

| Criterion       | Include                                                                                             | Exclude or handle separately                                             |
|-----------------|-----------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| Population      | Human patients undergoing transplantation for cicatricial/scar-related alopecia                     | Animal/in-vitro research; unrelated indications unless added by protocol |
| Study design    | Clinical trials, cohort studies, case series and case reports with usable patient-level information | Editorials and opinion pieces without original patient data              |
| Outcomes        | Hair/graft survival, density, aesthetic ratings, complications, satisfaction, follow-up             | No relevant clinical outcome                                             |
| Reporting level | Patient-level values when explicitly reported; otherwise study-level summaries                      | Do not derive unreported individual values from aggregate summaries      |
| Images          | Original clinical/trichoscopic images with permission and                                           | Article illustrations are not automatically authorized or                |

|  |                                               |                             |
|--|-----------------------------------------------|-----------------------------|
|  | linkage to patient/time/outcome, if available | suitable for model training |
|--|-----------------------------------------------|-----------------------------|

### 3.5 Data extraction and variable definitions

Data should be extracted into a standardized spreadsheet, with each row linked to its source article, table or figure, page, and case identifier. For a final systematic review, two reviewers should independently extract key variables and

resolve disagreements by checking the source; if only one reviewer is available, this should be acknowledged. Every field should be marked as directly reported, transparently derived, not reported, or uncertain. No missing patient-level value should be filled with a guessed value.

| Domain              | Variables to extract                                                  | Operational rule                                               |
|---------------------|-----------------------------------------------------------------------|----------------------------------------------------------------|
| Study descriptors   | Author, year, design, population, setting, follow-up                  | Retain source-study identity and check possible cohort overlap |
| Demographics        | Age, sex                                                              | Preserve units and source categories                           |
| Clinical / etiology | Cause of alopecia, anatomical site, defect size, scar features        | Document any harmonization of original labels                  |
| Surgery             | FUE/DHI/other technique, graft count, sessions, adjunctive procedures | Do not equate grafts with follicles or sessions                |
| Outcomes            | Survival %, density, aesthetic score/GAIS, complications              | Record outcome definition, denominator and time point          |
| Patient-reported    | Satisfaction categories or validated score                            | Keep original scale and response labels                        |
| Image data          | Clinical photos, trichoscopy, acquisition settings and time points    | No linked original images are present in the current workbook  |
| Audit trail         | Source page/table, extraction status, missingness note                | Each value must be traceable to the publication                |

### 3.6 Data preprocessing and quality control

An unchanged copy of the source workbook should be retained. A separate analysis copy should standardize variable names, categorical labels, units, date/follow-up formats, and missing-value tokens. Quality checks should identify duplicate rows, implausible values, inconsistent units, conflicts between article text and tables, and possible patient overlap across publications. Where a source article contains an internal inconsistency, the extracted value should remain traceable to the exact location in the article, with a discrepancy note rather than silent correction. Missing data should be reported per variable and per study. Blank values must not be treated as zero, and 'not reported' should be distinguished from 'not applicable' wherever possible. Primary

descriptive analyses should use the available denominator for each variable and state that denominator. The published-only dataset is the basis for clinical summaries; simulated data must not be merged into estimates of survival, satisfaction, or patient characteristics.

### 3.7 Clinical and trichoscopic image protocol

The current workbook does not contain linked original clinical photographs or trichoscopic images. Accordingly, no patient image features are claimed as measured data. For a future multimodal study, preoperative and postoperative images should be collected under standardized conditions, including consistent lighting, camera distance, angle, scalp preparation, magnification, and a scale marker when feasible. Trichoscopy

records should document device, magnification, anatomical region, and timing relative to surgery. Candidate image-derived variables include hair-shaft density per unit area, follicular-unit counts, shaft-diameter distribution, visible follicular openings, scalp coverage, erythema, pigmentation, and scar texture. Automated

segmentation should be compared with expert-reviewed annotations on a prespecified subset. Assessors should be blinded to outcomes where feasible, and inter-rater agreement should be reported. Any images used for model training must have appropriate consent/authorization, privacy safeguards, and rights for reuse.

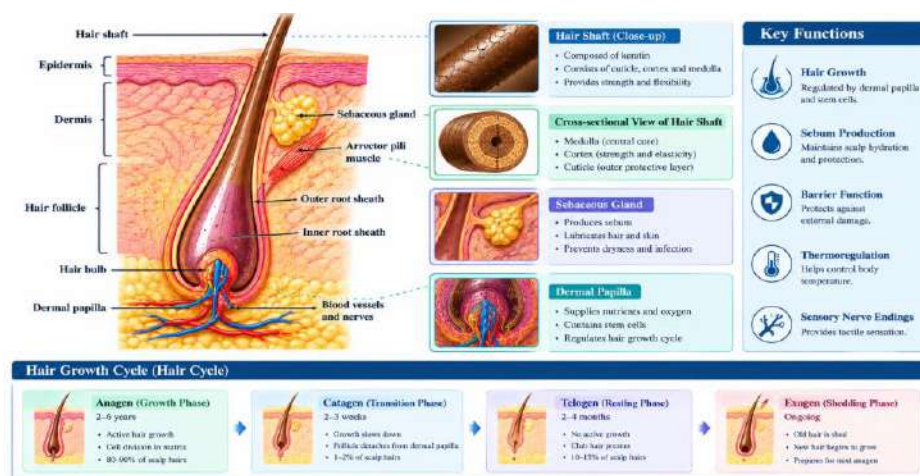


**Figure A: Anatomy of the Hair Follicle and Hair Growth Cycle**

Figure A illustrates the anatomical structure of the human hair follicle, its major components, their physiological functions, and the sequential phases of the hair growth cycle. The hair follicle is a specialized skin structure responsible for producing and supporting hair growth. The hair shaft extends above the skin surface, while the follicle is embedded within the dermis and is surrounded by supporting tissue. The epidermis forms the outer protective layer of the skin, whereas the dermis contains important structures involved in follicular function. The sebaceous gland produces sebum, which lubricates the hair and helps maintain scalp hydration. The arrector pili muscle is attached to the follicular structure and contributes to the physiological response commonly associated with goosebumps. Within the lower portion of the follicle, the hair bulb contains actively dividing cells involved in hair production, while the dermal papilla provides important biological signals and supports

follicular growth through its vascular and cellular interactions. The inner and outer root sheaths contribute to the organization and support of the developing hair shaft. The figure also presents the hair growth cycle, consisting of anagen, catagen, telogen, and exogen phases. Anagen represents active hair production, catagen is a transitional phase, telogen is a resting phase, and exogen involves the shedding of existing hair. Understanding these structures and biological processes is important for interpreting hair density, follicular health, and potential transplantation outcomes. In hair transplantation research, follicular anatomy provides the biological foundation for evaluating graft integrity, implantation, and subsequent hair growth. Changes in visible density, shaft characteristics, and scalp coverage may help characterize postoperative outcomes when measured using standardized clinical photography and trichoscopic assessment.



**Figure B: Clinical Hair Transplantation Workflow and Postoperative Outcome Assessment**

Figure B presents the principal stages of the clinical hair transplantation workflow, from preoperative assessment to postoperative follow-up and evaluation of treatment outcomes. The process begins with a comprehensive clinical examination to assess the patient's hair-loss pattern, scalp condition, donor-area suitability, recipient-site characteristics, and underlying cause of alopecia. Clinical photography and trichoscopic examination may be used to document baseline hair density, shaft diameter, follicular-unit distribution, and the condition of the recipient scalp. During the surgical planning stage, the surgeon determines the appropriate transplantation technique, estimates the required graft number, and designs the recipient area according to the patient's clinical needs. Follicular unit extraction (FUE) involves harvesting individual follicular units from the donor region, whereas direct hair implantation (DHI) uses a specialized implanter to place prepared grafts into the recipient area. Following extraction, grafts are handled carefully and assessed for visible integrity before implantation. Recipient-site creation and graft placement require attention to the direction, angle, depth, and distribution of the implanted follicles to

achieve a natural-looking appearance. Postoperative management includes wound care, clinical monitoring, and scheduled follow-up to evaluate healing and possible complications. At subsequent assessments, standardized photographs and trichoscopic images can support the measurement of hair density, scalp coverage, hair-shaft characteristics, and visible aesthetic changes. Patient-reported satisfaction and validated aesthetic assessment scales, such as the Global Aesthetic Improvement Scale (GAIS) when appropriately applied, may provide complementary information about perceived treatment benefit. For artificial intelligence research, this workflow identifies potential data sources, including demographic characteristics, clinical findings, surgical variables, image-derived features, and postoperative outcomes. These modalities could eventually be integrated into a predictive framework to estimate graft survival and aesthetic results. However, reliable model development requires appropriately consented patient-level data, clearly defined outcomes, standardized image acquisition, and independent validation. Figure B is a conceptual illustration of the clinical process; it does not represent a verified treatment sequence or measured

outcomes from the patients included in the current dataset.

### 3.8 Statistical analysis

The analysis was descriptive and focused on the availability, distribution, and provenance of variables within the 36 literature-derived records. Categorical variables were summarized using frequencies and percentages, with denominators explicitly based on the number of non-missing observations. Continuous variables were summarized using the number of available observations, mean, standard deviation, median, minimum, and maximum, as appropriate. Hair-survival summaries were calculated only for records containing numeric survival values. These summaries describe the available entries in the assembled dataset and should not be interpreted as pooled treatment-effect estimates because differences in source populations, measurement procedures, follow-up, and outcome definitions

may limit comparability. Patient-satisfaction categories were retained as reported rather than treated as measurements from a common validated instrument. Missingness was assessed separately for each variable. Study-level descriptions repeated across individual records were identified as a potential source of non-independence and were not interpreted as independent patient-level measurements. The 1,000 simulated records were excluded from all clinical descriptive summaries. No missing values were imputed, and no inferential tests, predictive models, cross-validation procedures, or external validation analyses were conducted. Before final submission, all counts, percentages, and descriptive statistics must be reproduced from the final analysis dataset using a documented statistical script. Any discrepancies between the spreadsheet, source publications, and manuscript tables must be resolved before results are reported.

| Objective                        | Recommended approach                                                             | Safeguard                                           |
|----------------------------------|----------------------------------------------------------------------------------|-----------------------------------------------------|
| Describe patient characteristics | Mean $\pm$ SD or median (IQR); n (%)                                             | Show valid n for each field                         |
| Describe survival outcomes       | Distribution, range, and study-specific summaries                                | Use published values only; state outcome definition |
| Compare source studies           | Stratified descriptive analysis                                                  | Avoid causal interpretation of study differences    |
| Assess missingness               | Variable-by-study completeness matrix                                            | Missing is not zero                                 |
| Future prediction task           | Patient-level split; internal and external validation                            | Prevent leakage across patients/studies             |
| Future model metrics             | MAE/RMSE for continuous outcomes; AUROC/AUPRC and calibration for classification | Predefine target, metrics and clinical use          |

### 3.9 Proposed multimodal AI architecture

The proposed AI architecture is a conceptual research framework rather than a trained or validated model. It would combine structured clinical variables, surgical variables, and standardized clinical/trichoscopic images when genuinely available. Each modality would undergo separate quality control and preprocessing before feature fusion. Prespecified prediction targets could include hair survival at a

defined follow-up, an aesthetic outcome measured using a defined scale, and patient satisfaction.

Model development should be deferred until a sufficiently sized, de-identified, patient-level dataset with appropriate permissions is assembled. Data should be split by patient and, where feasible, by center or study. A simple clinical-only baseline should be compared with multimodal models to quantify the incremental

value of images. Internal validation, calibration, subgroup error analysis, and independent external validation are needed before any clinical performance claims. The 1,000 simulated rows

may support pipeline testing but cannot substitute for real observations in training or evaluation of clinical performance.

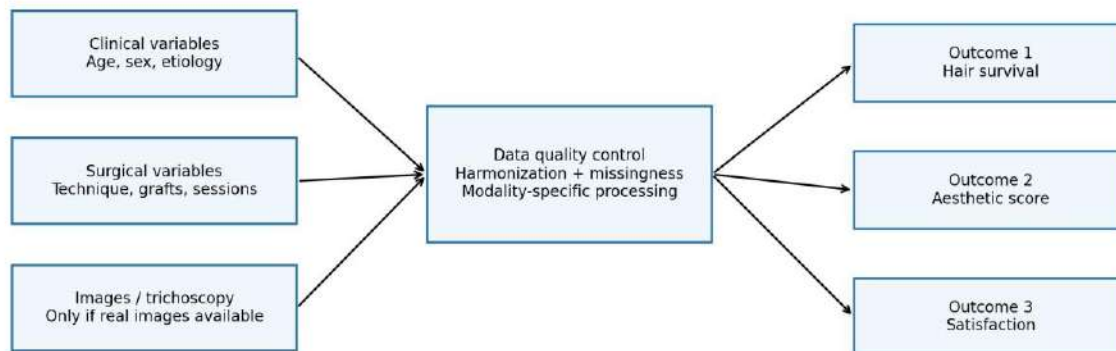


**Figure 3: Proposed multimodal AI framework for future research; no trained model performance is implied**

### 3.10 Bias, ethics, and reproducibility

Risk of bias should be assessed with tools appropriate to each study design, and judgments should be reported by domain rather than collapsed into an unsupported single score. The extraction should consider selection bias, outcome measurement, missing data, follow-up adequacy, and selective reporting. If a predictive model study is later performed, prediction-model risk-of-bias and applicability should be assessed with an appropriate tool such as PROBAST.

For published evidence, follow the relevant reporting guidance and verify permissions before reproducing article images. Any future collection of patient records or images requires the applicable ethics review, consent or lawful waiver, de-identification, controlled access, and a data-management plan. Reproducibility materials should include the search log, eligibility decisions, data dictionary, source-to-row audit trail, analysis code, software versions, and a clear record of all deviations from the protocol.

## 4. Results And Data Analysis

### 4.1 Overview of the Literature-Derived Dataset

The final dataset comprised 36 literature-derived records extracted from three published studies. Of these, 14 records were attributed to Baranov

and Nor (2026), 13 to Wang et al. (2023), and nine to Kwon et al. (2007). An additional 1,000 simulated records were maintained separately for technical testing and were excluded from all clinical summaries. Accordingly, the results presented below characterize the availability and distribution of reported information rather than the performance of a clinically validated prediction model.

The principal finding was substantial variation in the availability of outcome and explanatory variables. Although several records contained demographic information and hair-survival measurements, other clinically relevant variables were incompletely reported or unavailable. This variation limited the consistency with which outcomes could be compared across the included studies and highlighted the need to distinguish reported observations from missing or unverified information.

### 4.2. Availability and Distribution of Hair-Survival Outcomes

Numerical hair-survival values were available for 26 of the 36 literature-derived records (72.2%). Across these 26 observations, mean reported hair survival was 85.30% (SD = 10.60), while the median was 85.90%. The observed values ranged

from 45.1% to 96.8%, with an interquartile range of 82.40%–92.92%. The variation indicates that the extracted records did not show a uniform distribution of reported survival outcomes. The availability of numerical survival measurements differed across studies. All 14 records attributed to Baranov and Nor contained numerical survival values, whereas 12 of the 13 Wang et al. records contained numerical values. No numerical survival measurements were available for the nine Kwon et al. records in the extracted dataset. Consequently, cross-study

comparisons were constrained by differences in outcome availability, and the overall descriptive statistics should not be interpreted as a pooled estimate of hair-transplantation effectiveness. The results establish a preliminary quantitative description of the available survival data. However, differences in source populations, outcome definitions, assessment procedures, and follow-up conditions must be considered before interpreting variation as evidence of differences in clinical performance.

**Table 1: Composition of the supplied dataset and analysis sets**

| Dataset component                                    | Records (n) | Role in this chapter             | Clinical interpretation                             |
|------------------------------------------------------|-------------|----------------------------------|-----------------------------------------------------|
| Published case records                               | 36          | Primary descriptive analysis     | Source-reported observations; heterogeneous studies |
| Simulated records                                    | 1000        | Excluded from clinical estimates | Artificial values for pipeline testing only         |
| Combined workbook                                    | 1036        | Inventory only                   | Must not be analysed as one clinical cohort         |
| Published records with numeric hair-survival value   | 26          | Outcome-specific analysis        | 14 Baranov & Nor + 12 Wang; none from Kwon table    |
| Published records with patient satisfaction reported | 27          | Outcome-specific analysis        | 9 of 36 have satisfaction missing                   |

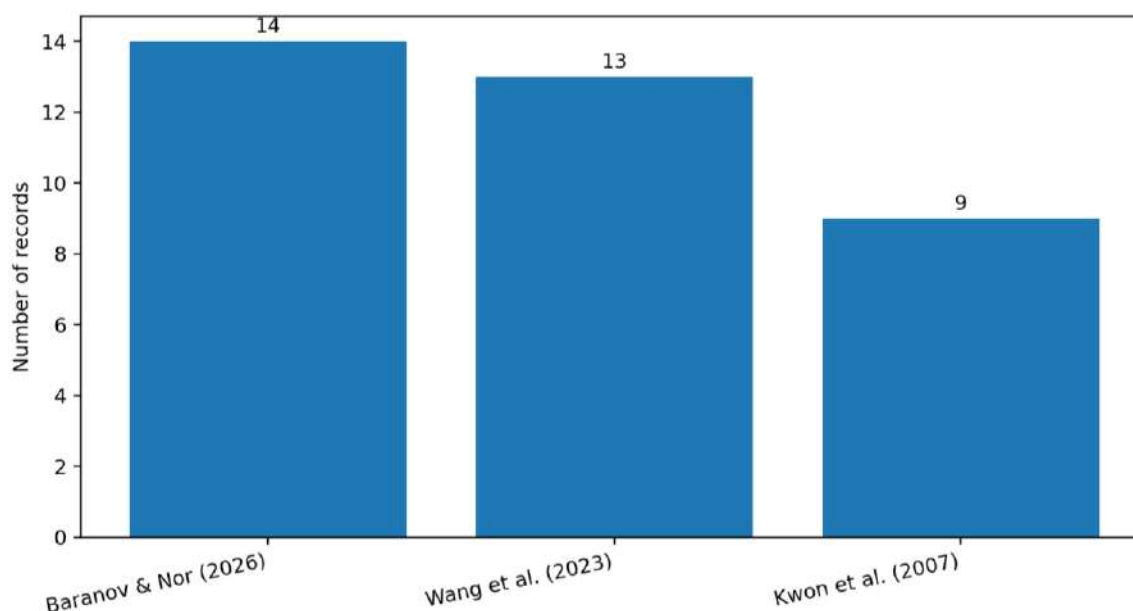


**Figure 1: Number of published case records contributed by each source study**

### 4.3 Demographic and clinical characteristics

Table 2: Demographic and clinical characteristics in published records

| Characteristic                  | Available n | Summary                                                                                      |
|---------------------------------|-------------|----------------------------------------------------------------------------------------------|
| Total published records         | 36          | Three source studies                                                                         |
| Age (years)                     | 36          | Mean 21.94; SD 13.37; median 23.00; range 4.08–54.00                                         |
| Sex: male                       | 20          | 55.6% of records                                                                             |
| Sex: female                     | 16          | 44.4% of records                                                                             |
| Defect area (cm <sup>2</sup> )  | 26          | Mean 113.65; median 50.00; range 3.0–420.0                                                   |
| Graft/FU count available        | 27          | 75.0% of records                                                                             |
| Follow-up information available | 27          | 14 records explicitly at 12 months; 13 have study-level 6–12 month range; 9 are not reported |

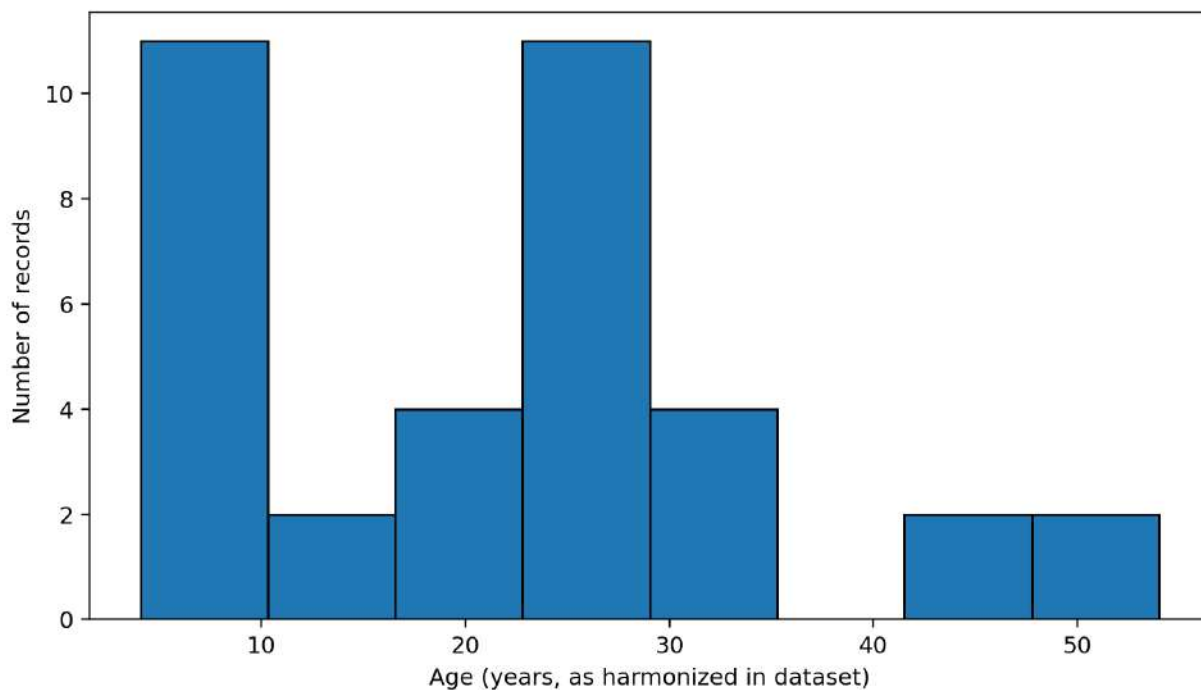


Figure 2: Age distribution across the 36 published case records (age values harmonized in the supplied dataset)

Age was available for all 36 records, with a mean of 21.94 years (SD 13.37) and a median of 23.00 years. The broad age range reflects the mixture of adult and pediatric scar-alopecia reports and should not be interpreted as a representative

transplant population. Defect area was available for 26 records only, with a mean of 113.65 cm<sup>2</sup> and a median of 50.00 cm<sup>2</sup>. Because area is missing for 10 records and case definitions differ, these values summarize only the subset where area was reported.

#### 4.4 Completeness and suitability of variables for multimodal modelling

Table 3: Completeness of selected variables in the 36 published records

| Variable                             | Reported n | Completeness | Missing n |
|--------------------------------------|------------|--------------|-----------|
| age_years                            | 36         | 100.0%       | 0         |
| sex                                  | 36         | 100.0%       | 0         |
| defect_area_cm2                      | 26         | 72.2%        | 10        |
| grafts_or_follicular_units           | 27         | 75.0%        | 9         |
| followup_months                      | 27         | 75.0%        | 9         |
| hair_survival_pct_numeric            | 26         | 72.2%        | 10        |
| GAIS_outcome_as_reported             | 22         | 61.1%        | 14        |
| patient_satisfaction_as_reported     | 27         | 75.0%        | 9         |
| graft_density_fu_cm2                 | 22         | 61.1%        | 14        |
| vascularity_score_1to5               | 0          | 0.0%         | 36        |
| fibrosis_score_1to5                  | 0          | 0.0%         | 36        |
| trichoscopic_follicular_openings_pct | 0          | 0.0%         | 36        |
| image_redness_feature_0to1           | 0          | 0.0%         | 36        |
| image_scarring_feature_0to1          | 0          | 0.0%         | 36        |
| image_hair_coverage_feature_0to1     | 0          | 0.0%         | 36        |

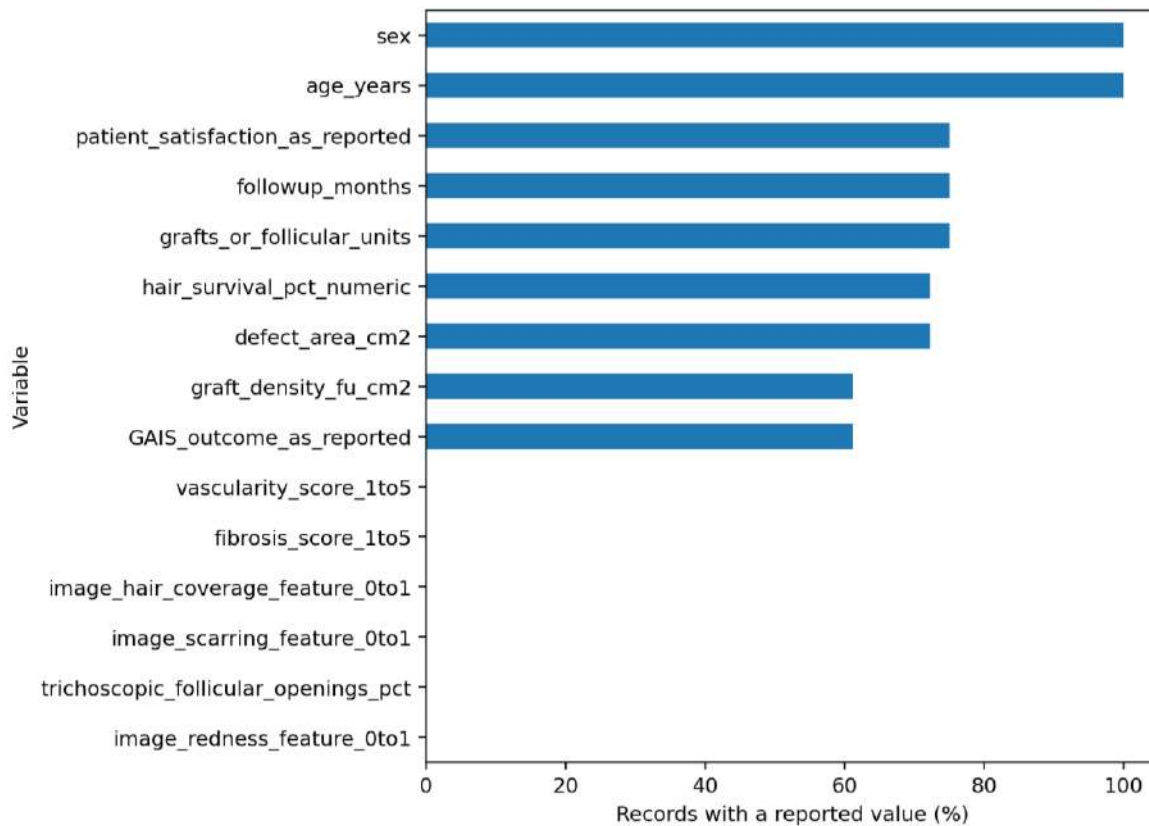


Figure 7: Completeness of selected variables in the published case dataset

The completeness assessment identifies a central limitation for the proposed multimodal AI study.

Demographic fields such as age and sex are present throughout this extracted dataset, while

defect area, graft count, follow-up, hair survival, global aesthetic improvement, and satisfaction are inconsistently reported. Most importantly, all five intended clinical/trichoscopic and image-feature fields shown in the completeness figure for image-based modelling are absent from the published-only records. Therefore, this workbook

does not contain linked image-and-clinical inputs required to train or validate a genuine multimodal prediction model. No image-derived predictor importance, image/clinical fusion benefit, or model accuracy can be calculated from these data.

#### 4.5 Hair-follicle survival outcomes

Table 4. Descriptive statistics for reported hair-survival percentages

| Statistic                     | Value                                                       |
|-------------------------------|-------------------------------------------------------------|
| Records with numeric survival | 26                                                          |
| Mean (%)                      | 85.30                                                       |
| Standard deviation            | 10.60                                                       |
| Median (%)                    | 85.90                                                       |
| Interquartile range (%)       | 82.40–92.92                                                 |
| Minimum–maximum (%)           | 45.1–96.8                                                   |
| 95% t-based CI for mean (%)   | 81.02–89.59 (descriptive, assumes independent observations) |

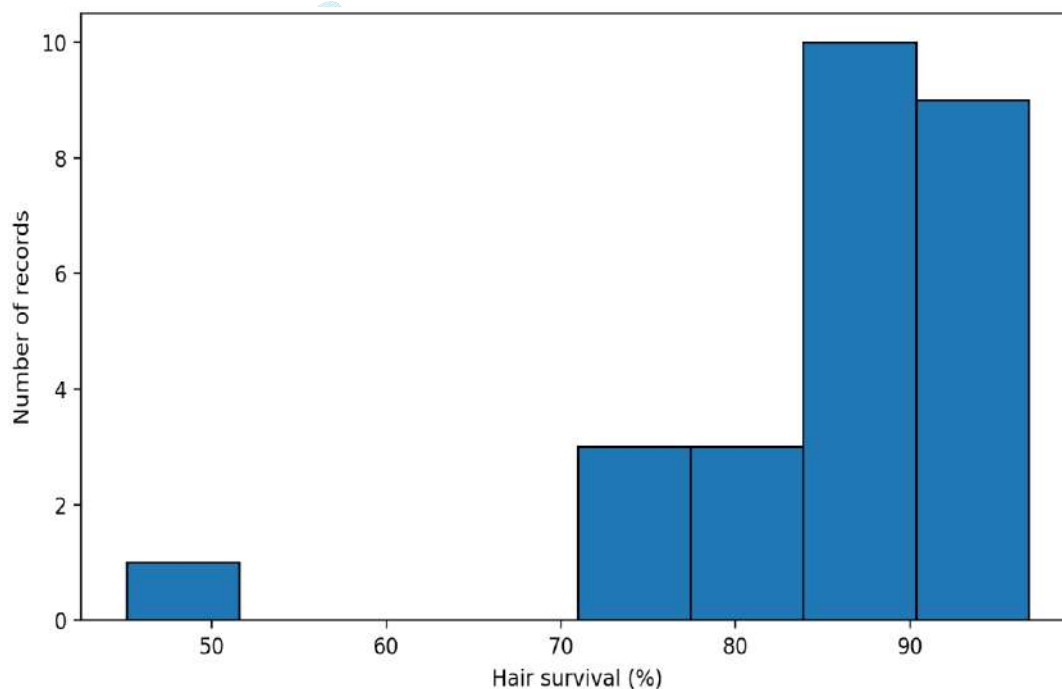


Figure 3: Distribution of the 26 available numeric hair-survival percentages

Among the 26 records with a numeric survival value, the mean reported hair survival was 85.30% (SD 10.60), the median was 85.90%, and values ranged from 45.1% to 96.8%. The observed spread indicates substantial variability across the extracted case records. The lowest

value should be checked against the source table and its outcome definition before being interpreted clinically. The confidence interval in Table 4 is a conventional descriptive interval calculated from these extracted values; because the records are selected from heterogeneous

publications and may not be statistically independent, it is not a generalizable pooled clinical estimate.

#### 4.6 Study-level comparison of survival

Table 5: Reported survival by source study

| Source study         | Published records | Numeric survival n | Mean (%) | SD (%) | Median (%) | Range (%) |
|----------------------|-------------------|--------------------|----------|--------|------------|-----------|
| Baranov & Nor (2026) | 14                | 14                 | 84.14    | 13.38  | 86.35      | 45.1–96.8 |
| Wang et al. (2023)   | 13                | 12                 | 86.66    | 6.35   | 85.60      | 76.0–96.1 |
| Kwon et al. (2007)   | 9                 | 0                  | NR       | NR     | NR         | NR        |

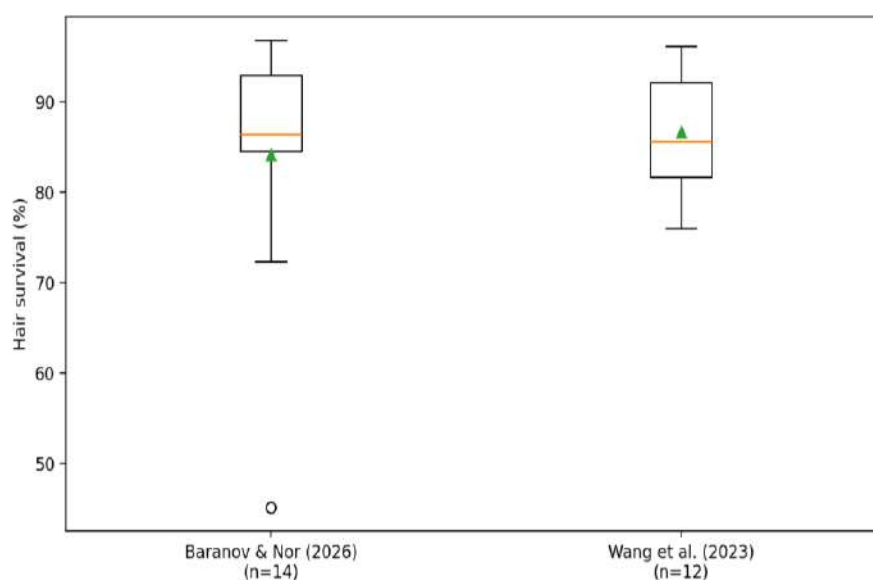


Figure 4: Distribution of available survival percentages by source study

The Baranov and Nor records had a mean reported survival of 84.14% (n=14; SD 13.38; range 45.1–96.8%). The Wang records with numeric values had a mean of 86.66% (n=12; SD 6.35; range 76.0–96.1%). No numeric survival percentage was available for the nine Kwon records in the extracted table. A Welch two-sample t-test comparing the two groups with available survival values yields  $p=0.539$ ; however,

this test is presented only as an exploratory calculation and should not be interpreted as evidence of equivalence or absence of a true difference. The groups were not randomized or directly comparable, and the studies differ in case mix, treatment context, outcome reporting, and potentially follow-up. The appropriate primary interpretation is descriptive rather than inferential.

#### 4.7 Causes of alopecia and treatment context

Table 6: Reported causes of alopecia among published case records

| Reported cause | Records (n) | Share of all published records |
|----------------|-------------|--------------------------------|
| Burn           | 8           | 22.2%                          |
| Suture scar    | 6           | 16.7%                          |

|                                              |   |       |
|----------------------------------------------|---|-------|
| Surgery                                      | 5 | 13.9% |
| Trauma                                       | 4 | 11.1% |
| After repairing nevus with dilator           | 4 | 11.1% |
| Accident trauma                              | 2 | 5.6%  |
| Scald                                        | 2 | 5.6%  |
| Blast injury                                 | 1 | 2.8%  |
| After craniosynostosis                       | 1 | 2.8%  |
| Gouges                                       | 1 | 2.8%  |
| Glass cut                                    | 1 | 2.8%  |
| After repairing sebaceous nevus with dilator | 1 | 2.8%  |

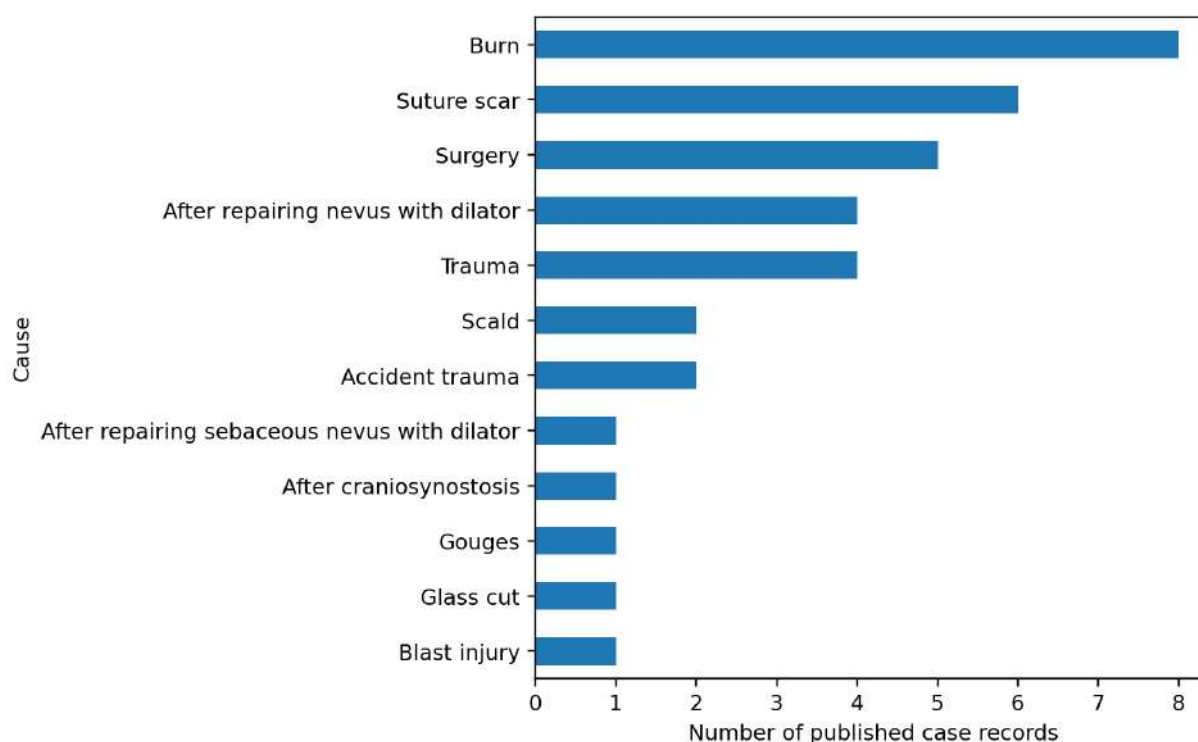


Figure 5: Frequency of reported causes of alopecia in the extracted cases

Burn-related cases were the most frequent reported category (8/36; 22.2%), followed by suture scar (6/36; 16.7%), surgery (5/36; 13.9%), trauma (4/36; 11.1%), and after repair of nevus with a dilator (4/36; 11.1%). The remaining categories were individually uncommon. These frequencies describe the composition of the assembled literature sample, not the population prevalence of causes of cicatricial alopecia or the relative success of transplantation by cause. Some categories may overlap conceptually (for example, trauma and accident trauma); the original source

terminology has been retained rather than silently recoded.

Reported techniques were heterogeneous: FUE: 21, Follicular-unit transplantation (specific method details not uniformly provided): 9, FUT: 3, FUE + DHI: 2, FUT + FUE: 1. Technique labels were retained as supplied in the extracted dataset. The source papers do not provide enough consistently aligned variables to estimate adjusted technique-specific effectiveness. In particular, a simple comparison across FUE, FUT, combined techniques, and records with

unspecified technique details would be vulnerable to confounding by indication and

differences in case severity.

#### 4.8 Aesthetic outcomes and patient satisfaction

Table 7: Patient satisfaction as recorded in the published case dataset

| Satisfaction category | n  | % of all records        | % of records with reported satisfaction     |
|-----------------------|----|-------------------------|---------------------------------------------|
| Very satisfied        | 7  | 19.4% of all 36 records | 25.9% of records with satisfaction reported |
| Mostly satisfied      | 17 | 47.2% of all 36 records | 63.0% of records with satisfaction reported |
| Dissatisfied          | 3  | 8.3% of all 36 records  | 11.1% of records with satisfaction reported |
| Not reported          | 9  | 25.0% of all records    | —                                           |

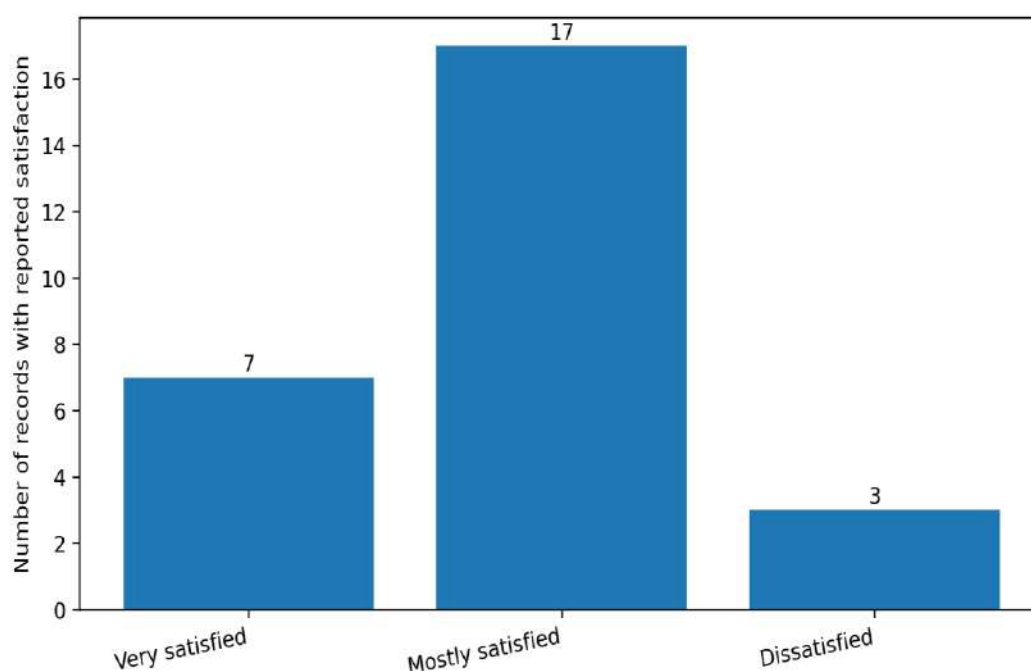


Figure 6: Patient satisfaction categories recorded in the source tables

Satisfaction was reported for 27 of the 36 records. Among all records, 17 were coded as mostly satisfied, 7 as very satisfied, and 3 as dissatisfied; nine records had no satisfaction value. Thus, among records with satisfaction data, 24/27 (88.9%) were coded as mostly or very satisfied, while 3/27 (11.1%) were dissatisfied. This is a descriptive calculation of the categories available in the workbook and is not a validated pooled satisfaction rate. Differences in the source instruments, timing, wording, and case mix may

affect comparability. Global aesthetic improvement (GAIS) is also incompletely reported and should be analysed separately from patient satisfaction rather than treated as the same endpoint.

#### 4.9 Implications for the proposed multimodal AI research

The dataset supports an evidence-mapping and data-readiness analysis, but not a defensible claim that a multimodal AI model has been trained or

clinically validated. The published subset has 36 heterogeneous records, only 26 numeric survival outcomes, incomplete covariates, no linked clinical/trichoscopic images, and no consistent set of predictor variables across cases. The 1,000 simulated records may be used to test code, data pipelines, visualizations, or proof-of-concept interfaces; they must remain labelled as simulated and must not be combined with published cases to report model performance, clinical accuracy, calibration, or generalizability.

A valid next-stage study should prospectively or retrospectively assemble patient-level data under appropriate ethics/privacy approvals, with a defined index procedure and standardized follow-up. It should link preoperative clinical variables, donor/recipient assessments, surgical variables, actual standardized photographs or trichoscopic images, and prespecified outcomes such as graft survival at a fixed time point and a validated aesthetic/satisfaction score. The protocol should define each outcome operationally, retain patient-level missingness indicators, verify cohort overlap across publications, and separate training, validation, and external test cohorts at the patient (and preferably study/site) level. Image features should be extracted from actual images using a reproducible pipeline; placeholder numbers cannot substitute for image data.

## 5. Conclusion

This study provides a quantitative assessment of the completeness, consistency, and multimodal data readiness of literature-derived information on hair transplantation, addressing an important prerequisite for the development of reliable artificial intelligence (AI) applications in aesthetic and clinical outcome assessment. The analysis of 36 literature-derived records demonstrated that, although clinically relevant information is available, its structure and completeness remain insufficient for robust multimodal predictive modelling. Numerical hair-survival measurements were available for 26 records (72.2%), with a mean reported survival of 85.30% (SD = 10.60), a median of 85.90%, and a range of 45.1%–96.8%. Satisfaction information was available for 27 records, of which 24 contained favorable

satisfaction labels. However, differences in outcome definitions, incomplete reporting, and uncertainty regarding the independence and assessment level of some aesthetic outcomes limit the comparability and clinical interpretation of these findings.

The principal contribution of this study is the demonstration that data readiness, rather than record volume alone, is a critical prerequisite for credible AI development in hair transplantation. Although demographic and selected clinical variables were available, the absence of linked clinical and trichoscopic images and the lack of imaging-derived features substantially restricted the feasibility of multimodal integration. Furthermore, inconsistencies in follow-up documentation and the need to verify extracted values against their original sources underscore the importance of data provenance, standardized measurement, and patient-level traceability. Accordingly, the reported survival and satisfaction summaries should be interpreted as descriptive findings from the extracted records, not as pooled estimates of treatment effectiveness or evidence of predictive AI performance.

Future research should prioritize the establishment of prospective, standardized, and adequately documented hair-transplantation datasets that link patient characteristics, procedural variables, graft-related measurements, longitudinal follow-up, validated patient-reported outcomes, and appropriately annotated clinical or trichoscopic images. Consistent outcome definitions, transparent missing-data reporting, independent patient-level validation, and external evaluation across clinical settings will be essential for determining whether multimodal AI can improve the reliability of outcome estimation and support clinically meaningful decision-making. The 1,000 simulated records used for technical testing in this study were kept separate from the literature-derived records and should not be interpreted as additional clinical evidence.

In conclusion, this study establishes a descriptive baseline for evaluating the readiness of existing hair-transplantation data for future AI research and identifies specific reporting and integration barriers that must be addressed before predictive

claims are justified. By defining these limitations and translating them into concrete data-development priorities, the study provides a methodological starting point for moving from fragmented outcome reporting toward reproducible, interpretable, and clinically validated AI research in hair transplantation. The next decisive step is not to claim predictive capability from the current dataset, but to build and independently validate a high-quality, image-linked clinical resource capable of supporting it.

## REFERENCES

1. Kwon, O. S., Kim, M. H., Park, S. H., Chung, J. H., Eun, H. C., & Oh, J. K. (2007). Staged hair transplantation in cicatricial alopecia after carbon dioxide laser-assisted scar tissue remodeling. *Archives of Dermatology*, 143(4), 457–460. <https://doi.org/10.1001/archderm.143.4.457>
2. Ekelem, C., Pham, C., & Atanaskova Mesinkovska, N. (2019). A systematic review of the outcome of hair transplantation in primary scarring alopecia. *Skin Appendage Disorders*, 5(2), 65–71. <https://doi.org/10.1159/000492539>
3. Lee, J. A., Levy, D. A., Patel, K. G., Brennan, E., et al. (2021). Hair transplantation in frontal fibrosing alopecia and lichen planopilaris: A systematic review. *The Laryngoscope*, 131, 59–66. Verify the final page range and author list against the journal record before submission.
4. Donovan, J. (2012). Lichen planopilaris after hair transplantation: Report of 17 cases. *Dermatologic Surgery*, 38, 1998–2004.
5. Chiang, Y. Z., Tosti, A., Chaudhry, I. H., Lyne, L., et al. (2012). Lichen planopilaris following hair transplantation and face-lift surgery. *British Journal of Dermatology*, 166, 666–670.
6. Saxena, K., Saxena, D. K., & Savant, S. S. (2016). Successful hair transplant outcome in cicatricial lichen planus of the scalp by combining scalp and beard hair along with platelet-rich plasma. *Journal of Cutaneous and Aesthetic Surgery*, 9, 51–55.
7. Wang, et al. (2023). Application of autologous hair transplantation technique in children with cicatricial alopecia. [Verify the complete author list and journal metadata from the article record.] <https://pmc.ncbi.nlm.nih.gov/articles/PMC10427550/>
8. Singh, R., Wilborn, D., Lintzeri, D. A., & Blume-Peytavi, U. (2023). Health-related quality of life among patients with primary cicatricial alopecia: A systematic review. *Journal of the European Academy of Dermatology and Venereology*, 37, 2462–2473.
9. Rezende, H. D., Leal, T. B., & Dias, M. F. R. G. (2026). A systematic review of follicular unit graft survival rates after hair transplantation in primary cicatricial alopecia. *Dermatologic Surgery*, 52(7), 720–721. <https://doi.org/10.1097/DSS.00000000000004911>
10. Baranov, I., & Nor, N. (2026). The use of hair transplantation techniques to treat secondary cicatricial alopecia. *Norwegian Journal of Development of the International Science*, 174, 40–44. <https://doi.org/10.5281/zenodo.18631241>
11. Urban, G., Feil, N., Csuka, E., Hashemi, K., Ekelem, C., Choi, F., Mesinkovska, N. A., & Baldi, P. (2021). Combining deep learning with optical coherence tomography imaging to determine scalp hair and follicle counts. *Lasers in Surgery and Medicine*, 53(1), 171–178. <https://doi.org/10.1002/lsm.23324>

12. Evaluation of automated measurement of hair density using deep neural networks. (2022). *Sensors*, 22(2), Article 650. <https://doi.org/10.3390/s22020650>. This study used 4,492 scalp images and compared object-detection models, including YOLOv4. Verify the full author metadata before citing.
13. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542, 115–118. <https://doi.org/10.1038/nature21056>
14. Haenssle, H. A., Fink, C., Schneiderbauer, R., et al. (2018). Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Annals of Oncology*, 29(8), 1836–1842. <https://doi.org/10.1093/annonc/mdy166>
15. Tschandl, P., Rinner, C., Apalla, Z., et al. (2020). Human-computer collaboration for skin cancer recognition. *Nature Medicine*, 26, 1229–1234. <https://doi.org/10.1038/s41591-020-0942-0>
16. Liu, Y., Jain, A., Eng, C., et al. (2020). A deep learning system for differential diagnosis of skin diseases. *Nature Medicine*, 26, 900–908. Verify the final bibliographic metadata before submission.
17. Han, S. S., Park, I., Chang, S. E., et al. (2018). Augmented intelligence dermatology: Deep neural networks for the classification of skin lesions. Verify the exact title, journal, and DOI against the publisher record before using this citation.
18. Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermoscopic images of common pigmented skin lesions. *Scientific Data*, 5, Article 180161. <https://doi.org/10.1038/sdata.2018.161>
19. Codella, N. C. F., Gutman, D., Celebi, M. E., et al. (2018). Skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC). 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI), 168–172. <https://doi.org/10.1109/ISBI.2018.8363547>
20. Mongan, J., Moy, L., & Kahn, C. E., Jr. (2020). Checklist for artificial intelligence in medical imaging (CLAIM): A guide for authors and reviewers. *Radiology: Artificial Intelligence*, 2(2), Article e200029. <https://doi.org/10.1148/ryai.2020200029>
21. Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25, 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
22. Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17, Article 195. <https://doi.org/10.1186/s12916-019-1426-2>
23. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, 380, 1347–1358. <https://doi.org/10.1056/NEJMr1814259>
24. Wiens, J., Saria, S., Sendak, M., et al. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine*, 25, 1337–1340. <https://doi.org/10.1038/s41591-019-0548-6>
25. Liu, X., Faes, L., Kale, A. U., et al. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *The Lancet Digital Health*, 1(6), e271–e297. [https://doi.org/10.1016/S2589-7500\(19\)30123-2](https://doi.org/10.1016/S2589-7500(19)30123-2)

26. Liu, X., Cruz Rivera, S., Moher, D., Calvert, M. J., & Denniston, A. K. (2020). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. *Nature Medicine*, 26, 1364–1374. <https://doi.org/10.1038/s41591-020-1034-x>
27. Rivera, S. C., Liu, X., Chan, A.-W., Denniston, A. K., & Calvert, M. J. (2020). Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI extension. *Nature Medicine*, 26, 1351–1363. <https://doi.org/10.1038/s41591-020-1037-7>
28. World Health Organization. (2021). Ethics and governance of artificial intelligence for health: WHO guidance. World Health Organization. <https://www.who.int/publications/i/item/9789240029200>
29. Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Annals of Internal Medicine*, 162(1), 55–63. <https://doi.org/10.7326/M14-0697>
30. Wolff, R. F., Moons, K. G. M., Riley, R. D., et al. (2019). PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine*, 170(1), 51–58. <https://doi.org/10.7326/M18-1376>
31. Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, Article n71. <https://doi.org/10.1136/bmj.n71>
32. Collins, G. S., Dhiman, P., Andaur Navarro, C. L., et al. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. Verify the final volume, article number, and DOI against the publisher record before submission.
33. Moons, K. G. M., Damen, J. A. A., Kaul, T., et al. (2025). PROBAST+AI: An updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ*, 388, Article e082505. <https://doi.org/10.1136/bmj-2024-082505>
34. Van Calster, B., McLernon, D. J., van Smeden, M., Wynants, L., & Steyerberg, E. W. (2019). Calibration: The Achilles heel of predictive analytics. *BMC Medicine*, 17, Article 230. <https://doi.org/10.1186/s12916-019-1466-7>
35. Vickers, A. J., & Elkin, E. B. (2006). Decision curve analysis: A novel method for evaluating prediction models. *Medical Decision Making*, 26(6), 565–574. <https://doi.org/10.1177/0272989X06295361>
36. Steyerberg, E. W. (2019). Clinical prediction models: A practical approach to development, validation, and updating (2nd ed.). Springer.
37. Harrell, F. E., Jr. (2015). Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis (2nd ed.). Springer.
38. Riley, R. D., Snell, K. I. E., Ensor, J., et al. (2020). Minimum sample size for developing a multivariable prediction model: Part I—Continuous outcomes. *Statistics in Medicine*, 39(9), 1262–1275. Verify the final page range and DOI before submission.\
39. Riley, R. D., Ensor, J., Snell, K. I. E., et al. (2020). Calculating the sample size required for developing a clinical prediction model. *BMJ*, 368, Article m441. <https://doi.org/10.1136/bmj.m441>
40. Steyerberg, E. W., Harrell, F. E., Jr., Borsboom, G. J. J. M., Eijkemans, M. J. C., Vergouwe, Y., & Habbema, J. D. F. (2001). Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology*, 54(8), 774–781.

- 
41. Khan, R., Shah, A. M., Ijaz, A., & Sumeer, A. (2025). Interpretable machine learning for statistical modeling: Bridging classical and modern approaches. *International Journal of Social Sciences Bulletin*, 3(8), 43-50.

