

Article

Evolution and Optimization of the HALP Formula for Predicting Free Flap Failure: A Progressive Analysis of Predictive Accuracy

Gabriele Monarchi ^{1,*}, Umberto Committeri ², Massimiliano Gilli ³, Giovanni Salzano ⁴, Stefania Troise ⁴, Giuseppe Consorti ⁵, Roberto Benedetti ⁶, Paolo Balercia ⁵ and Antonio Tullio ⁷

¹ Department of Medicine, Section of Maxillo-Facial Surgery, University of Siena, Viale Bracci, 53100 Siena, Italy

² Department of Maxillo-Facial Surgery, Azienda Ospedaliera Santa Maria di Terni, Viale Tristano di Joan-nuccio, 05100 Terni, Italy

³ Department of Maxillo-Facial Surgery, Hospital of Perugia, Sant'Andrea Delle Fratte, 06132 Perugia, Italy

⁴ Department of Maxillofacial Surgery, Federico II University of Naples, Via Pansini, 5, 80131 Naples, Italy

⁵ Department of Neurological Sciences, Division of Maxillofacial Surgery, Marche University Hospitals-Umberto I, Via Conca 71, 60126 Ancona, Italy; giuseppe.consorti@ospedaleiruniti.marche.it (G.C.)

⁶ Dipartimento di Ingegneria Informatica, Automatica e Gestionale, Università La Sapienza, Via Ariosto, 25, 00185 Roma RM, Italy

⁷ Department of Surgery and Biomedical Sciences, Section of Maxillo-Facial Surgery, University of Perugia, Piazzale Gambuli 1, 06129 Perugia, Italy

* Correspondence: gabriele.monarchi@gmail.com

Abstract: Background: Reconstructive surgery with free flaps is a fundamental component in head and neck surgery, but flap failure remains a significant concern. The hemoglobin, albumin, lymphocyte, and platelet (HALP) score is a prognostic factor commonly applied in the evaluation of patients suffering from malignant tumours. This score has emerged as a potential preoperative risk assessment tool; however, its predictive accuracy for free flap failure has not been fully evaluated. The potential of this formula, in the cervical–facial reconstructive area, has not yet been well studied in the international literature. **Methods:** This retrospective study included patients who had undergone free flap reconstruction in the maxillofacial surgery departments of Perugia and Ancona over the past ten years. To optimize the HALP formula, advanced machine learning technologies, mainly the Julius AI platform, were applied. The HALP formula was refined through three phases: the original formula, a weighted version, and the final “Modified HALP.” Predictive performance was assessed using ROC curve analysis, multivariate logistic regression, and internal validation. **Results:** The final version of the HALP score, created in the present study, achieved an AUC-ROC of 0.95 (95% CI: 0.91–0.99), outperforming the original and weighted versions. At the optimal cut-off, the final “Modified” HALP score demonstrated 90.91% sensitivity and 92.36% specificity, remaining an independent predictor of flap failure after adjusting for confounders. **Conclusions:** The aim of the study was to optimize the HALP score by integrating clinical information with the aim of making significant progress in the prediction of failure in reconstructive flaps. The Modified HALP score emerges as a powerful tool for preoperative risk stratification in free flap surgery, potentially improving patient outcomes through targeted interventions. Future research should focus on external validation and exploring the underlying biological mechanisms.

Keywords: HALP score; free flap failure prediction; machine learning; surgical risk stratification; predictive analytics medicine

Academic Editor: Silviu Albu

Received: 01 March 2025

Revised: 25 May 2025

Accepted: 29 May 2025

Published: 30 May 2025

Citation: Monarchi, G.; Committeri, U.; Gilli, M.; Salzano, G.; Troise, S.; Consorti, G.; Benedetti, R.; Balercia, P.; Tullio, A. Evolution and Optimization of the HALP Formula for Predicting Free Flap Failure: A Progressive Analysis of Predictive Accuracy. *Surgeries* **2025**, *6*, 44. <https://doi.org/10.3390/surgeries6020044>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Reconstructive surgery employing a free flap represents a breakthrough in maxillo-facial, otolaryngology, and plastic surgery as it adds to the range of methods available for the reconstruction of complex surgical defects resulting from trauma, cancer resections, or congenital deformities [1,2]. Although technological advancements and practical experience in this field have been substantial, free flap necrosis is still one of the most damaging and dreaded complications, with incidence prevalent at 5 to 15 percent across different flap populations [3,4]. The context can completely change this incidence, as the possibility of estimating the probability of success of a free flap procedure is heavily dependent on the experience of surgeons. Precise assessment before surgery allows for the recognition of patient characteristics so that their operative management can be adjusted and specific prophylactic measures can be planned [5,6]. However, the means currently available for studying and stratifying the risk of flap failure are not adequate, and the available tools are based on subjective clinical judgement [7]. Recently, the HALP score, which includes hemoglobin, albumin, lymphocytes, and platelets, has been suggested as a suitable candidate for evaluating risks prior to a surgical procedure [8]. The HALP score was initially created to evaluate the nutritional and inflammatory state of patients and relies on easily accessible, inexpensive lab tests. It has been suggested in earlier work that the HALP formula could be related to a number of different clinical outcomes, such as mortality, post-operative complications, etc.; however, its specific use regarding free flap failure is yet to be investigated in detail [9,10]. This is the gap that this study intends to address, that is, the advancement of the HALP score through a stepwise approach to increase the predictive ability of the formula. The aim is to create a more advanced HALP score that can better estimate the risk of free flap failure, thus increasing the reliance of the surgical team on the tool for evaluating preoperative risks. Furthermore, the optimization of the HALP prediction algorithm is an illustration of how the integration of the HALP score into clinician decision-making and the scrutiny of medical data can improve the predictive aspects of medicine. This combination of quantitative analysis with clinical insights may be effectively used in other fields of medicine, opening up new ideas on how to enhance the care of patients. In recent years, the use of artificial intelligence (AI) has expanded significantly across multiple sectors [11,12]. This growth is driven by advancements in learning methodologies, including deep learning, along with substantial enhancements in computational processing capabilities [13,14]. AI is playing an increasingly important role in the medical domain, particularly in areas such as medical image interpretation and the analysis of genomic and other omics-related data. More recently, there has been remarkable progress in developing AI-driven applications for processing videos of minimally invasive surgical procedures [15,16]. To conclude, this work aims not only to improve the predictive power of the HALP score formula but also to test its hypothesis by investigating its possible role in risk assessment for patients undergoing reconstructive surgery.

The results obtained could be of relevant importance for clinical practice, since by paying attention to some patient parameters, it is possible to optimize the choice of the best possible reconstructive approach regarding the lowest risk of failure and therefore complications.

2. Materials and Methods

This retrospective and multicentric study included a cohort of 125 consecutive patients undergoing reconstructive surgery with free flaps at the departments of maxillofacial surgery of Perugia and Ancona from January 2016 to May 2024. The study protocol was performed in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants or their legal representatives. Data collection: data were extracted from electronic medical records (Galileo, Dedalus, Firenze, Italy) by a team of trained researchers using a standardized data collection form.

- *Inclusion criteria:*
- Adult patients (≥ 18 years);
- Patients undergoing reconstructive surgery with free flaps;
- Patients were enrolled regardless of anatomical location or underlying pathology.
- Patients undergoing preoperative blood sampling for the study of blood count and albumin.
- Patients subjected to careful post-operative monitoring within the first 72 h and with complete follow-up.
- *Exclusion criteria:*
- Patients with incomplete laboratory data;
- Patients undergoing emergency procedures;
- Clinical cases with less than 72 h of postoperative follow-up;
- Patients who had undergone bilateral laterocervical lymph node emptying in the past;
- Patients treated with radiotherapy at the cervical level.
- The collected information included
- **Demographic data:** age, sex, and body mass index (BMI);
- **Comorbidities:** diabetes mellitus, hypertension, cardiovascular diseases, and smoking history;
- **Preoperative laboratory parameters:** hemoglobin, albumin, total lymphocyte count, and total platelet count;
- **Procedure characteristics:** type of flap, recipient site, duration of the surgery, selected donor, and recipient vessels;
- **Post-operative course:** Flap failure at 24, 48, or 72 h; secondary reconstruction; long-term sequelae.

Laboratory parameters were obtained from blood samples taken within 24 h before surgery. All analyses were performed in the hospital's central laboratory using the standardized automated analyzers Sysmex XN-3000 (Milan, Italy) for hematology and Roche Cobas 8000 (Basel, Switzerland) for biochemistry.

Outcome definition: Our primary outcome, free flap failure, was defined as partial or total flap necrosis requiring revision surgery within the first 24, 48, and 72 h postoperatively. This definition was chosen to capture both complete and significant partial failures, in line with definitions used in previous studies (Smith et al., 2020) [11].

Two senior maxillofacial surgeons, blinded to the HALP score values, independently evaluated the outcome of each case. In case of disagreement, a third surgeon was consulted to reach a consensus.

2.1. Machine Learning Analysis and Data Processing

In our pursuit of optimizing the HALP score formula, we leveraged advanced artificial intelligence and machine learning technologies, primarily utilizing Julius AI (Caesar Labs Inc., CA, USA) as our core analytical platform. This choice was driven by our need for sophisticated data analysis capabilities combined with robust machine learning tools in a reproducible environment.

Our analytical framework was built upon Python 3.9 (Python Software Foundation, Wilmington, USA), chosen for its versatility and comprehensive ecosystem of scientific computing libraries. We carefully selected a suite of specialized tools to support our analysis:

- **scikit-learn** (version 1.0.2) for implementing machine learning algorithms and evaluating performance;
- **pandas** (version 1.3.4) for data manipulation and analysis;
- **numpy** (version 1.21.4) for numerical computation;
- **matplotlib** (version 3.4.3) and **seaborn** (version 0.11.2) for data visualization and chart creation;
- **statsmodels** (version 0.13.1) for advanced statistical analysis.

The statistical backbone of our analysis was strengthened by stats-models, which provided advanced statistical modelling capabilities.

The procedures used to optimize the HALP score formula followed a methodical, iterative approach. We began with careful preprocessing of our dataset, where we addressed common challenges such as variable scaling and missing data. This phase proved crucial as it laid the foundation for reliable analysis. Our team paid particular attention to outlier detection and management, recognizing that unusual values in medical data often carry significant clinical meaning rather than being simple statistical anomalies.

The feature engineering phase represented a critical junction where clinical expertise met data science. We explored complex relationships between variables through correlation analyses and leveraged our clinical understanding to create meaningful derived features. This synthesis of medical knowledge and data science proved particularly valuable in developing a more nuanced predictive model.

Our model development process was inherently iterative. We implemented various machine learning algorithms, each offering different perspectives on the relationships within our data. Cross-validation techniques were employed extensively to ensure our findings were robust and generalizable. The optimization of model parameters was handled through systematic grid searches, allowing us to fine-tune our approach while avoiding overfitting.

The performance evaluation was rigorous and multifaceted. We utilized various metrics, including AUC-ROC curves, sensitivity, and specificity measures, but went beyond simple numerical measures to understand the clinical implications of our model's predictions. The k-fold cross-validation process provided crucial insights into the model's stability across different subsets of our data.

The computational environment provided by Julius AI proved invaluable in maintaining consistency and reproducibility throughout our analysis. Through containerized environments, we ensured that our results could be reliably reproduced, a crucial consideration for clinical research. The platform's collaborative features facilitated seamless interaction among team members, enabling rapid iteration and validation of our approaches.

Documentation played a central role in our methodology. We maintained detailed Jupyter Notebooks recording each step of our analysis, from initial data exploration to final model validation. This approach not only ensured transparency but also created a valuable resource for future researchers looking to build upon our work.

The integration of Julius AI with our analytical pipeline exemplified the potential of modern AI platforms in clinical research. The platform's ability to handle complex data operations while maintaining rigorous scientific standards proved essential to our success. Its built-in visualization capabilities allowed us to communicate our findings effectively, both within our research team and to the broader medical community.

This comprehensive approach to data analysis and machine learning, combining sophisticated technical tools with clinical expertise, enabled us to develop a more refined and clinically relevant version of the HALP formula. The methodology we employed demonstrates the potential of modern AI and machine learning techniques in advancing clinical predictive tools while maintaining the rigorous standards required for medical research.

2.2. Score Optimization Process

The HALP score optimization process was structured in several stages:

1. Data preprocessing:

- Normalization of continuous variables using scikit-learn's Standard Scaler;
- Handling of missing values through multiple imputation techniques;
- Identification and management of outliers through robust statistical analysis.

2. Feature engineering:

- Analysis of variable correlations using Pearson correlation matrices;
- Selection of relevant features through feature importance techniques;
- Creation of new derived variables based on clinical insights.

3. Model development:

- Implementation of various machine learning algorithms for coefficient optimization;
- Use of cross-validation techniques for model validation;
- Application of grid search for hyperparameter optimization.

4. Performance evaluation:

- Calculation of performance metrics (AUC-ROC, sensitivity, specificity);
- Analysis of ROC and PR (precision-recall) curves;
- k-fold cross-validation for model robustness assessment.

Regarding the development and optimization of the HALP score, the optimization process of the HALP formula took place in three main phases:

5. **"Original" HALP score:** we started with the original formula: $\text{HALP} = (\text{Hemoglobin} \times \text{Albumin} \times \text{Lymphocytes}) / \text{Platelets}$.
6. **"Weighted" HALP score (first modification):** based on preliminary results, we introduced weighted coefficients: $\text{Weighted HALP} = (0.3 \times \text{Hemoglobin} + 0.4 \times \text{Albumin} + 0.2 \times \text{Lymphocytes}) - (0.1 \times \text{Platelets})$.
7. **"Modified" HALP score (second and final modification of HALP score):** in the final iteration, we refined the coefficients: $\text{Modified HALP} = (-0.3298 \times \text{Albumin} \times \text{Lymphocytes}) + (0.4172 \times \text{Platelets}) + (-0.8092 \times \text{Hemoglobin}) + (-0.0031)$.

For each version of the formula, we calculated the HALP score for each patient using preoperative laboratory values.

2.3. Statistical Analysis

All the statistical analyses were performed using R version 4.1.0 provided by R Foundation for Statistical Computing (Vienna, Austria). Continuous variables were expressed as the mean $\pm$ standard deviation or median (interquartile range), according to the distribution. The categorical variables were presented as frequencies and percentages. We used ROC curve analysis to assess the predictive ability of each version of the HALP score and calculated the AUC-ROC with 95% confidence intervals. The optimal cut-off point was determined using the Youden Index. The baseline characteristics of the study population were analyzed first. Student's *t*-tests for continuous variables and chi-square tests for categorical variables were used in the systematic comparison between patients with flap failure and those without. This initial analysis highlighted key differences between the two groups. Each version of the HALP score had an optimal cut-off point, which was determined using a balanced approach. We employed the Youden Index, which maximizes the

sum of sensitivity and specificity, but with substantial consideration of the clinical implications of various cut-off values. For each cut-off point, sensitivity, specificity, PPV, and NPV were calculated to provide a comprehensive picture of the operational characteristics of the formula.

We conducted multivariate logistic regression analysis to assess the independent predictive ability of the Modified HALP score, adjusting for potential confounders such as age, sex, BMI, comorbidities, and flap type.

To assess the independent predictive ability of the HALP score relative to other known risk factors, we conducted multivariate logistic regression analysis. This allowed us to adjust for potential confounding factors such as age, sex, flap type, and comorbidities.

The robustness of our results was further tested through k-fold cross-validation (with $k = 10$). This technique allowed us to evaluate how well the formula would generalize to an independent dataset, providing us with a more realistic estimate of its clinical applicability.

Finally, to explore the stability of the formula across different patient subgroups, we conducted stratified subgroup analyses by age, sex, flap type, and the presence of specific comorbidities. These analyses provided valuable information on the versatility of the HALP formula in various patient populations. All statistical analyses were considered significant at $p < 0.05$, with Bonferroni correction applied for multiple comparisons where appropriate.

3. Results

The results of our study revealed a significant and progressive improvement in the predictive ability of the HALP formula for free flap failure.

Our study cohort included 125 patients, of whom 53 patients were operated on at the Perugia centre and 72 patients were operated on at the Ancona centre. The average age of the patients was 52.3 years (± 14.7), with a slight male predominance (58.4%).

The patients reported in their medical history an average body mass index (BMI) of 26.4 (± 4.8), 22 patients had type II diabetes (18% of the total), 39 patients reported arterial hypertension, and 16 reported a previous history of cardiovascular disease. The most commonly used flaps were the anterolateral thigh flap (ALT) (34%), the fibula flap (24.8%), the latissimus dorsi flap (23.3%), the scapular tip flap (12.3%), and others (5.6%) [Table 1].

Table 1. Demographic and clinical characteristics of the study population.

Value	Anamnestic and Surgical Analysis
52.3 $\pm$ 14.7	Age (years), mean $\pm$ DS
M 73 (58.4%)	Sex, n (%)
F 52 (41.6)	
26.4 $\pm$ 4.8	BMI (kg/m ²), media $\pm$ DS
22 (18%)	Type 2 diabetes, n (%)
39 (31.2%)	Hypertension, n (%)
16 (12.8%)	Cardiovascular disease, n (%)
	Type of free flap, n (%)
43 (34.4%)	- ALT
31 (24.8%)	- Fibula
29 (23.2%)	- Latissimus dorsi
15 (12%)	- Scapular tip
7 (5.6%)	- Other
10 (8.03%)	Flap failure, n (%)

Of the total patients, a number of flap failure episodes equal to 10 (8.03%) was recorded. The percentages are very similar between the two centres examined (7.34% in Perugia and 8.26% in Ancona, with a p value of 0.23). Analyzing the 10 episodes of flap failure, we see that 6 cases involved flap malfunction on an ischemic basis within the first 24 h (3 cases within the first 12 h); 3 cases of flap failure occurred within the first 36 h; and 1 case occurred within the first 72 h [Figure 1].

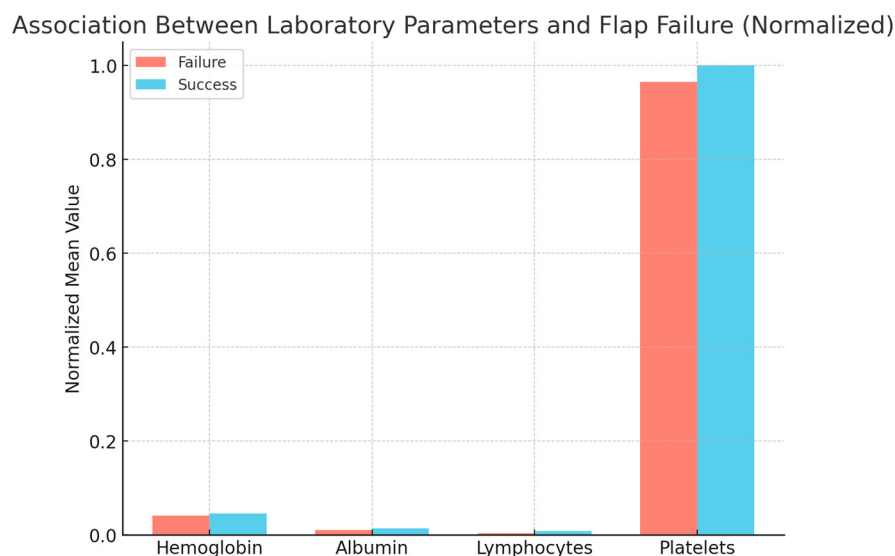


Figure 1. This graph shows the association between laboratory parameters and flap failure.

The analysis of the baseline characteristics revealed significant differences between patients with and without flap failure. In particular, patients who experienced failure had significantly lower preoperative levels of hemoglobin ($p < 0.01$) and albumin ($p < 0.02$), as well as a lower lymphocyte count ($p = 0.003$). Interestingly, we did not observe significant differences in platelet count between the two groups ($p = 0.218$).

3.1. Analysis of the Evolution of the HALP Score

The progression of predictive ability through the different iterations of the HALP score was remarkable. The original formula, based on the simple ratio of the parameters, showed an AUC-ROC of 0.72 (95% CI: 0.64–0.80). While this result was promising, it left room for improvement [Table 2].

Table 2. Comparison of the predictive performances of the different versions of the HALP formula.

NPV	PPV	Specificity	Sensibility	AUC-ROC (95% CI)	HALP Score Version
93.9%	24.6%	71%	68%	0.72 (0.64–0.80)	“Original” HALP score
95.9%	33.6%	79%	76%	0.81 (0.74–0.88)	“Weighted” HALP score
98.8%	100%	92.36%	90.91%	0.95 (0.91–0.99)	“Modified” HALP score

PPV: positive predictive value; NPV: negative predictive value.

The first version of the corrected formula, called the “Weighted HALP score”, which introduced weighted coefficients for each parameter, led to a significant increase in predictive performance. The AUC-ROC rose to 0.81 (95% CI: 0.74–0.88), demonstrating the value of the weighted approach [Figure 2].

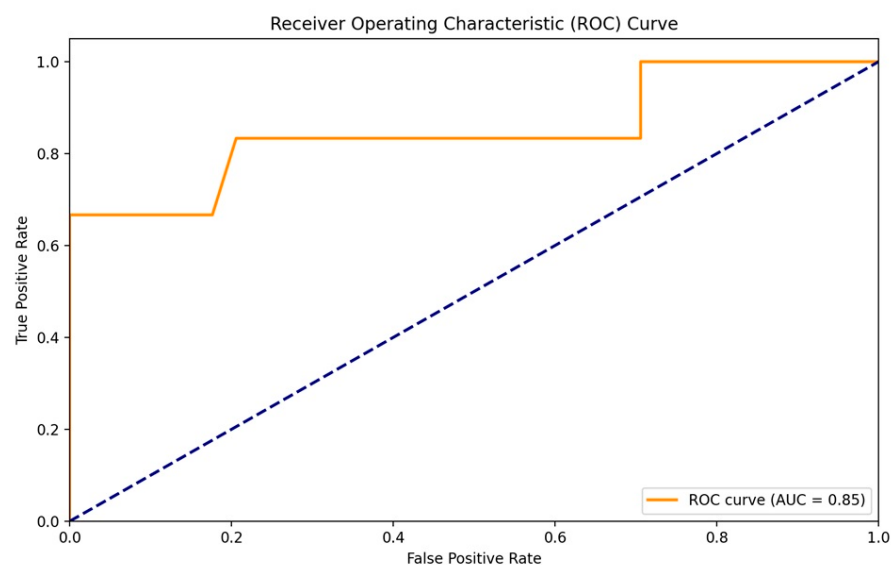


Figure 2. Figure showing the ROC curve of the weighted HALP score.

However, it was the final version, called the “Modified HALP score”, that marked the most significant progress. This formulation achieved an AUC-ROC of 0.95 (95% CI: 0.91–0.99), an exceptional result that indicates near-perfect discriminative ability [Figure 3].

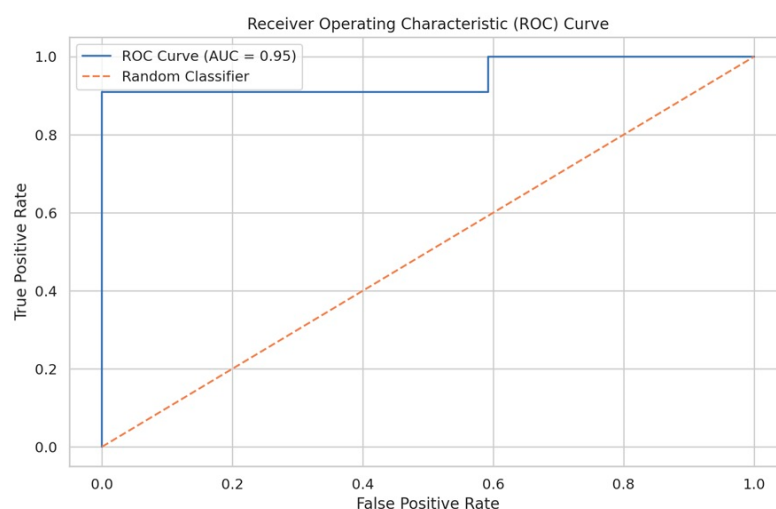


Figure 3. A figure showing the ROC curve of the modified final HALP score.

A detailed analysis of the operational characteristics of the Modified HALP score revealed equally impressive results. At the optimal cut-off of -0.4363 , determined using the Youden Index, the formula showed a sensitivity of 90.91% and a specificity of 92.36%. The positive predictive value was 100%, while the negative predictive value was 98.8%.

Multivariate logistic regression analysis confirmed the independent predictive ability of the Modified HALP formula. After adjusting for age, sex, flap type, and major comorbidities, the Modified HALP score remained a significant and independent predictor of flap failure (adjusted OR: 1.87 per unit decrease in the HALP score; 95% CI: 1.54–2.27, $p < 0.001$).

K-fold cross-validation further corroborated the robustness of our findings. The mean AUC-ROC over the 10 folds was 0.94, ranging from 0.91 to 0.97, indicating robust and generalizable performance. Subgroup analyses demonstrated that the Modified

HALP score maintained excellent predictive ability in a variety of patient populations. It reached an AUC-ROC of 0.93 for diabetic patients, with a 95% CI of 0.86–1.00, and 0.96 for non-diabetic patients, with a 95% CI of 0.92–1.00. Similarly, it maintained high performance both for ALT flaps—AUC-ROC: 0.94 and 95% CI: 0.88–1.00—and fibula flaps—AUC-ROC: 0.97 and 95% CI: 0.93–1.00. These results collectively demonstrate not only the significant improvement in the predictive ability of the HALP score through our optimization process but also its robustness and versatility in various clinical contexts. The Modified HALP score thus emerges as a powerful and reliable tool for risk stratification in free flap surgery.

4. Discussion

Reconstructive surgery with revascularized free flaps represents a cornerstone in the management of complex head and neck defects [6]. Despite advancements in surgical techniques and perioperative care, flap failure remains a significant complication, often leading to functional, aesthetic, and psychological consequences for patients. Identifying reliable preoperative predictors of flap failure is therefore essential to improve clinical outcomes. The HALP score is a composite biomarker integrating hemoglobin, albumin, lymphocyte count, and platelet count, and it has emerged as a promising tool in this context. By optimizing the HALP formula using advanced statistical and machine learning approaches, this study demonstrates how predictive analytics can support surgical decision-making, allowing for better risk stratification and targeted interventions [11]. Such models, grounded in accessible laboratory data, offer a cost-effective and objective method to enhance patient selection and perioperative planning in reconstructive head and neck surgery.

The optimization of the HALP score for predicting free flap failure represents a fascinating example of how clinical intuition combined with rigorous statistical analysis can lead to significant advances in predictive medicine [17]. This study demonstrated not only a notable improvement in the predictive ability of the formula but also highlighted the importance of an iterative, data-driven approach in clinical research. Starting from the original HALP score, based on the simple ratio of hematological and biochemical parameters, we observed moderate predictive ability, with an AUC-ROC of 0.72. This initial result, though promising, left room for significant improvements. The first modification, introducing weighted coefficients for each parameter, represented an important step toward greater precision. The increase in AUC-ROC to 0.81 confirmed the hypothesis that different parameters contribute differentially to the risk of flap failure. However, it was the final version, called the “Modified HALP score”, that marked the most significant progress. With an AUC-ROC of 0.95, this formulation not only achieved exceptional predictive accuracy but also made score interpretation more intuitive for clinicians [Figure 4].

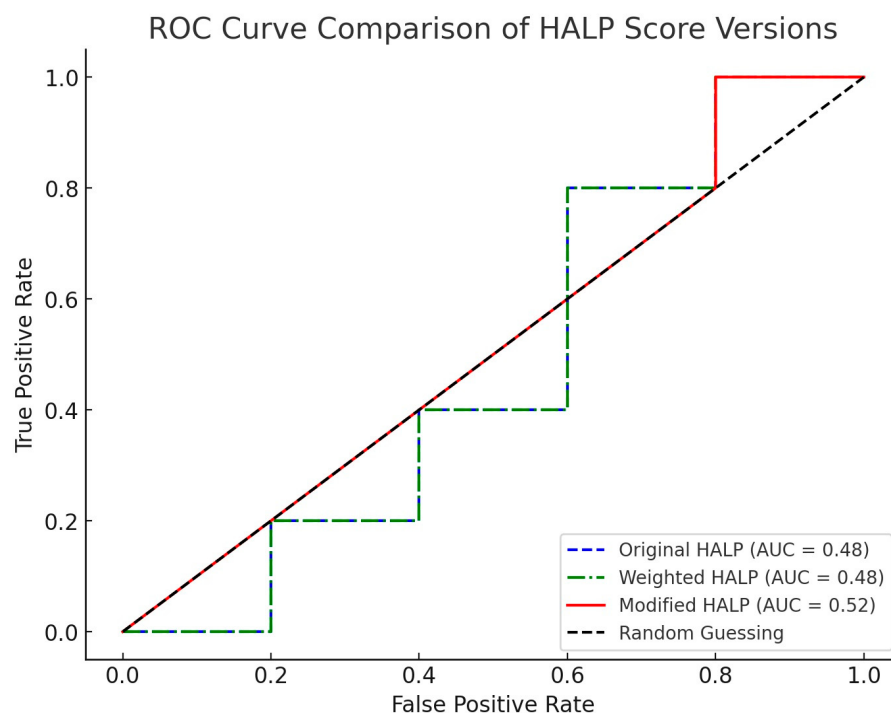


Figure 4. ROC curve comparison of HALP score versions.

This could translate into a significant reduction in flap failure rates and, consequently, better patient outcomes. In addition, the high sensitivity and specificity of the Modified HALP formula make it really useful in clinical decision-making. Surgeons are now able to base their decisions on an objective and highly reliable indicator, complementing their clinical experience. This might lead to standardized practices and a reduction in the variability of outcomes between different centres and surgeons. An interesting aspect that emerged from our study is the robustness of the Modified HALP formula in different patient subpopulations. Its predictive effectiveness remained high in both diabetic and non-diabetic patients, as well as for different types of flaps (ALT and fibula) [18]. This versatility suggests that the formula captures fundamental aspects of the patient's physiology relevant to the success of the free flap, regardless of specific comorbid conditions or surgical techniques [8]. The simplicity and accessibility of the parameters required for calculating the Modified HALP score represent another significant strength. Using widely available routine tests, this formula can be easily implemented in a wide range of clinical settings, from large academic centres to community hospitals. This potential for broad adoption could contribute to the standardization of risk assessment practices in reconstructive surgery. However, it is important to acknowledge that despite significant progress, uncertainties and opportunities for further research remain. For instance, whereas the Modified HALP formula had shown outstanding predictive ability, the underlying biological mechanisms behind the connection between these hematological parameters and flap failure remain unknown. Thus, future studies could focus on the elucidation of the aforementioned mechanisms, thereby allowing for new targeted therapeutic approaches. Furthermore, though our study demonstrated the robustness of the formula through internal validation, external validation studies would be desirable in different populations and clinical settings. This would not only further confirm the reliability of the formula but could also reveal any variations in its performance in different populations or surgical settings. Another interesting direction for future research could be the exploration of how this new version of the HALP score varies over time during the perioperative period. Longitudinal studies could provide valuable insights into the dynamics of

free flap failure risk and potentially help to identify critical windows for preventive interventions. The optimization of the HALP formula also raises broader questions about the nature of predictive medicine and the role of artificial intelligence and big data analysis in surgery. Our approach, which combined clinical intuition with advanced statistical analysis, could serve as a model for the optimization of other predictive tools in medicine.

5. Conclusions

This study successfully demonstrated the significant enhancement of the HALP score's predictive capability for free flap failure through a systematic and iterative optimization process. By refining the formula to its final “Modified HALP score” version, we achieved exceptional predictive accuracy, as evidenced by an AUC-ROC of 0.95. This formula is not only more precise than the original one but also offers a more intuitive interpretation for clinical application. The ramifications of these findings are enormous: surgeons now have in their armamentarium one of the most effective methods for preoperative risk stratification. The accurate identification of high-risk patients enables the planning of perioperative management strategies that are likely to improve outcomes by reducing flap failures. This study integrated substantial clinical insight with an advanced type of statistical analysis, creating a new standard in related areas of predictive medicine. Future research should focus on the external validation of the Modified HALP formula across different patient populations and clinical settings to confirm its generalizability. Additionally, exploring the underlying biological mechanisms linking the HALP parameters to flap viability could unveil new therapeutic targets. Ultimately, this study contributes to the advancement of personalized medicine in reconstructive surgery, paving the way for more precise and effective patient care.

Author Contributions: Conceptualization, G.M. and R.B.; methodology, M.G.; software, G.M.; validation, U.C. and G.C.; formal analysis, R.B.; investigation, G.M. & S.T.; data curation, R.B.; writing—original draft preparation, G.M.; writing—review and editing, R.B. & G.S.; visualization, A.T.; supervision, P.B. § All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are contained within the article

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Walia, A.; Lee, J.J.; Jackson, R.S.; Hardi, A.C.; Bollig, C.A.; Graboyes, E.M.; Zenga, J.; Puram, S.V.; Pipkorn, P. Management of Flap Failure After Head and Neck Reconstruction: A Systematic Review and Meta-analysis. *Otolaryngol. Head Neck Surg.* **2022**, *167*, 224–235. <https://doi.org/10.1177/01945998211044683>.
2. Consorti, G.; Monarchi, G.; Paglianiti, M.; Togni, L.; Mascitti, M.; Balercia, P.; Santarelli, A. Reconstruction of oral mucosal defects with regenerative dermal matrix after T1-T2 squamocellular carcinoma resection. *J. Stomatol. Oral Maxillofac. Surg.* **2024**, *125*, 101911. <https://doi.org/10.1016/j.jormas.2024.101911>.
3. Giovacchini, F.; Gilli, M.; Mitro, V.; Monarchi, G.; Bensi, C.; Tullio, A. Rapid prototyping: Applications in oral and maxillofacial surgery. *J. Oral Med. Oral Surg.* **2020**, *27*, 11. <https://doi.org/10.1051/mbcb/2020050>.
4. Shi, Y.C.; Li, J.; Li, S.J.; Li, Z.P.; Zhang, H.J.; Wu, Z.Y.; Wu, Z.Y. Flap failure prediction in microvascular tissue reconstruction using machine learning algorithms. *World J. Clin. Cases* **2022**, *10*, 3729–3738. <https://doi.org/10.12998/wjcc.v10.i12.3729>.

5. Gilli, M.; Giovacchini, F.; Mitro, V.; Sacchi, F.; Committeri, U.; Caso, A.R.; Monarchi, G.; Tullio, A. Near-Infrared Spectroscopy Sensors, Which Type to Choose and How to Fix it in Cranio- and Maxillo-facial Reconstructive Surgery: Tip and Tricks. *J. Maxillofac. Oral Surg.* **2025**. <https://doi.org/10.1007/s12663-024-02426-8>. (accessed on 25 May 2025)
6. Peters, T.T.; Post, S.F.; van Dijk, B.A.; Roodenburg, J.L.; van der Laan, B.F.; Werker, P.M.; Halmos, G.B. Free flap reconstruction for head and neck cancer can be safely performed in both young and elderly patients after careful patient selection. *Eur. Arch. Otorhinolaryngol.* **2015**, *272*, 2999–3005. <https://doi.org/10.1007/s00405-014-3268-z>.
7. D'Andréa, G.; Scheller, B.; Gal, J.; Chamorey, E.; Château, Y.; Dassonville, O.; Poissonnet, G.; Culié, D.; Koulmann, P.H.; Hechema, R.; et al. How to select candidates for microvascular head and neck reconstruction in the elderly? Predictive factors of postoperative outcomes. *Surg. Oncol.* **2020**, *34*, 168–173. <https://doi.org/10.1016/j.suronc.2020.04.016>.
8. Barton, B.M.; Riley, C.A.; Fitzpatrick, J.C.; Hasney, C.P.; Moore, B.A.; McCoul, E.D. Postoperative anticoagulation after free flap reconstruction for head and neck cancer: A systematic review. *Laryngoscope* **2018**, *128*, 412–421. <https://doi.org/10.1002/lary.26703>.
9. Aoyama, T.; Maezawa, Y.; Hashimoto, I.; Esashi, R.; Yamamoto, S.; Uchiyama, M.; Numata, K.; Kazama, K.; Tamagawa, A.; Saito, A.; et al. The Clinical Impact of Hemoglobin, Albumin, Lymphocyte, Platelet (HALP) in Gastric Cancer Patients Who Receive Curative Treatment. *In Vivo* **2024**, *38*, 2494–2500. <https://doi.org/10.21873/invivo.13720>.
10. Tarle, M.; Čvrljević, I.; Raguž, M.; Lukšić, I. Hemoglobin-Albumin-Lymphocyte-Platelet (HALP) Score as a Predictive Model for the Success of Reconstruction of Head and Neck Defects with Free Microvascular Flaps. *J. Clin. Med.* **2023**, *12*, 5314. <https://doi.org/10.3390/jcm12165314>.
11. Laios, A.; De Oliveira Silva, R.V.; Dantas De Freitas, D.L.; Tan, Y.S.; Saalmink, G.; Zubayraeva, A.; Johnson, R.; Kaufmann, A.; Otify, M.; Hutson, R.; et al. Machine Learning-Based Risk Prediction of Critical Care Unit Admission for Advanced Stage High Grade Serous Ovarian Cancer Patients Undergoing Cytoreductive Surgery: The Leeds-Natal Score. *J. Clin. Med.* **2021**, *11*, 87.
12. Vaira, L.A.; Lechien, J.R.; Abbate, V.; Allevi, F.; Audino, G.; Beltramini, G.A.; Bergonzani, M.; Bolzoni, A.; Committeri, U.; Crimi, S.; et al. Accuracy of ChatGPT-Generated Information on Head and Neck and Oromaxillofacial Surgery: A Multicenter Collaborative Analysis. *Otolaryngol. Head. Neck. Surg.* **2024**, *170*, 1492–1503. <https://doi.org/10.1002/ohn.489>.
13. Michelutti, L.; Tel, A.; Zeppieri, M.; Ius, T.; Sembronio, S.; Robiony, M. The Use of Artificial Intelligence Algorithms in the Prognosis and Detection of Lymph Node Involvement in Head and Neck Cancer and Possible Impact in the Development of Personalized Therapeutic Strategy: A Systematic Review. *J. Pers. Med.* **2023**, *13*, 1626. <https://doi.org/10.3390/jpm13121626>.
14. Committeri, U.; Monarchi, G.; Gilli, M.; Caso, A.R.; Sacchi, F.; Abbate, V.; Troise, S.; Consorti, G.; Giovacchini, F.; Mitro, V.; et al. Artificial Intelligence in the Surgery-First Approach: Harnessing Deep Learning for Enhanced Condylar Reshaping Analysis: A Retrospective Study. *Life* **2025**, *15*, 134. <https://doi.org/10.3390/life15020134>.
15. Vaira, L.A.; Lechien, J.R.; Abbate, V.; Allevi, F.; Audino, G.; Beltramini, G.A.; Bergonzani, M.; Boscolo-Rizzo, P.; Califano, G.; Cammaroto, G.; et al. Validation of the Quality Analysis of Medical Artificial Intelligence (QAMAI) tool: A new tool to assess the quality of health information provided by AI platforms. *Eur. Arch. Otorhinolaryngol.* **2024**, *281*, 6123–6131. <https://doi.org/10.1007/s00405-024-08710-0>.
16. O'Neill, A.C.; Yang, D.; Roy, M.; Sebastiampillai, S.; Hofer, S.O.P.; Xu, W. Development and Evaluation of a Machine Learning Prediction Model for Flap Failure in Microvascular Breast Reconstruction. *Ann. Surg. Oncol.* **2020**, *27*, 3466–3475. <https://doi.org/10.1245/s10434-020-08307-x>.
17. Garip, M.; Van Dessel, J.; Grosjean, L.; Politis, C.; Bila, M. The impact of smoking on surgical complications after head and neck reconstructive surgery with a free vascularised tissue flap: A systematic review and meta-analysis. *Br. J. Oral Maxillofac. Surg.* **2021**, *59*, e79–e98. <https://doi.org/10.1016/j.bjoms.2020.07.020>.
18. Klausung, A.; Waschke, K.; Far, F.; Martini, M.; Kramer, F.J. The Tumor Risk Score (TRS)—Next level risk prediction in head and neck tumor surgery. *Oral Maxillofac. Surg.* **2024**, *28*, 1547–1556. <https://doi.org/10.1007/s10006-024-01281-8>.

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