19. Kurniawan M, Tio P, Abdel AT, et al. 3D analysis of the cranial and facial shape in craniosynostosis patients: a systematic review. *J Craniofac Surg* 2024;35:813–821
20. Crofts H, Mankowski P, Tamber M, et al. Head shape after sagittal craniosynostosis surgery: open versus endoscopic strip craniectomy. *Plast Surg (Oakv)* 2025;33:283–288

Time and Expertise in Open Structural Rhinoplasty: A Task-Based Analysis Using Hierarchical Task Analysis and Machine Learning

*Mehmet Mazhar Celikoyar, MD,**
Oguzhan Topsakal, PhD,† Eric Dobratz, MD,‡ and Doga Demirel, PhD§

Abstract: Rhinoplasty consistsof specific surgical tasks performed in order and executed at specific times. Hierarchical task analysis (HTA) is an essential tool for developing performance metrics to help evaluate surgeries. The authors aimed to determine if there is a correlation with experience and time required for task completion. We developed an HTA for open structural rhinoplasty, then performed a survey to gather surgeons' self-reported time to complete tasks. Surgeons were grouped according to the number of rhinoplasty cases they have performed; those who performed <100 were considered "nonexpert," and those who performed more than 100 cases were considered "expert." Statistical analysis was done. Machine learning (ML) was utilized as well to help evaluate the comparison of two groups. Responses from 25 surgeons were analyzed. The surgical steps that showed statistically significant differences between the two surgeon groups included the elevation of (septal) mucoperichondrial-mucoperiosteal flaps, cephalic trim, septoplasty closure, and rhinoplasty closure, with significantly shorter time required by the expert surgeons. According to ML model, rhinoplasty closure, injection, transcolumellar incisions, dorsal hump reduction, dorsal surgery-lateral osteotomies, assessment of lower lateral cartilage, and dorsal hump bone reduction were the steps where the 2 groups of surgeons had significantly different time frames. These tasks may be accepted as more prone to benefits from time and surgical volume. The number of cases observed had no significant effect, therefore, the benefits from time and surgical volume are most noted with hands-on practice and performing the procedure.

Key Words: Hierarchical task analysis, machine learning, rhinoplasty, surgical education, technology assessment, biomedical

The technical performance of any operation is crucial to its outcomes.[1] Rhinoplasty is among the most challenging cosmetic facial procedures, requiring precise surgical steps to achieve the desired result. Surgeons undertake a variety of tasks, some mandatory and others optional, depending on the surgical goals, intraoperative findings, and their preferences. Each task is performed in an orderly sequence to successfully achieve the goal of creating esthetic harmony with the surrounding facial features while maintaining or enhancing nasal function and structural support.[2]

Developing a successful surgeon takes time. Surgical training needs to be enhanced by identifying knowledge gaps, improving visuospatial conceptualization, and other methods to be developed.[3] Many techniques have been tried to aid in the learning process of the surgery residents.[4] A learning curve is a fact. However, previous research on learning curves has been limited, and it was shown that the operating time differs significantly between inexperienced and experienced surgeons.[5]

The best methods to assess the current surgical knowledge of residents and their gains throughout their training are still debated.[6] As training programs increasingly shift towards competency-based approaches, it is essential to identify specific steps and tasks that can be incorporated into simulation and practice to effectively assess and measure the skills of rhinoplasty surgeons in training.[7] Identifying these steps or tasks not only facilitates data analysis but also informs oversight policies.[8]

Hierarchical task analysis (HTA) breaks down a procedure into a hierarchy of tasks and subtasks, outlining the relationships between them.[9] Hierarchical task analysis serves as an invaluable tool for creating performance metrics, enabling detailed time and performance analyses, and ultimately aiding in the development of surgical evaluation techniques.[10]

The time required to complete specific steps of a procedure often reflects the surgeon's level of expertise.[5] Expert surgeons tend to complete procedures more efficiently, requiring less time than novices. By defining a procedure step by step, it becomes possible to analyze the time a surgeon takes for each step compared with an average and draw inferences about their expertise. For instance, observational studies of intraoperative mastoidectomy recordings have demonstrated this as a feasible method for assessing surgeon skill levels.[11] A similar approach has also been applied to cricothyroidotomy evaluations.[12]

In this study, we developed an HTA for open structural rhinoplasty, detailing all necessary and optional tasks and ac-

From the *Department of Otolaryngology, Demiroğlu Bilim University School of Medicine, Istanbul, Turkey; †Bellini College of Artificial Intelligence, Cybersecurity and Computing, University of South Florida, Tampa, FL; ‡Department of Otolaryngology/Head and Neck Surgery, Eastern Virginia Medical School, Norfolk, VA; and §School of Computer Science, Gallogly College of Engineering, University of Oklahoma, Norman, OK.
Received January 29, 2025.
Accepted for publication August 17, 2025.
Address correspondence and reprint requests to Mehmet Mazhar Çelikoyar, MD, Kucuksu Mahallesi, Yalniz Selvi Caddesi, Cengelkoy'de Mesa Evleri, B7 Blok, Daire 1, Uskudar, Istanbul 34688, Turkey; E-mails: mazhar.celikoyar@demiroglu.bilim.edu.tr; mazhar.celikoyar@gmail.com
The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Demiroglu Bilim University (Protocol no: 44140529/8743; date: October 5, 2021). All participants consented to participate, and the questionnaire was designed to protect the privacy and anonymity of all responses.
The authors report no conflicts of interest.
Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.jcraniofacialsurgery.com.

ISSN: 1536-3732
DOI: 10.1097/SCS.0000000000011959

tions required to achieve the surgical goal. We conducted a survey to assess surgeons' self-reported time to complete these tasks and examined correlations between experience and task completion times. Statistical analysis, augmented by machine learning (ML), was used to evaluate skill levels and performance differences between surgeons.

We believe the findings from this task analysis and timing evaluation will contribute to the development of educational tools and provide a deeper understanding of performance quality for each step of rhinoplasty.

METHODS

Hierarchical Task Analysis

We used the clinical experience of two of the authors who are expert rhinoplasty surgeons to develop the HTA. The primary goal was to identify and break down the steps involved in rhinoplasty into individual tasks, analyze their relationships, and organize them into a systematic step-by-step procedure.

Rhinoplasty techniques are generally classified into 2 main categories: structural rhinoplasty and preservation rhinoplasty.[13–15] Structural rhinoplasty involves reshaping or repositioning nasal structures, often with the addition of grafts to achieve the desired form.[16] Open (external) structural rhinoplasty includes a columellar incision, resulting in an external, visible scar, whereas the closed (endonasal) approach reshapes or augments nasal structures through internal, intranasal incisions.[17,18] To ensure homogeneity in the tasks under analysis, we focused on a single approach: open structural rhinoplasty. Tasks involving preservation techniques, such as maintaining dorsal or keystone structures, were not included in this study.

The steps involved in open structural rhinoplasty have been previously described in the literature[17,19,20] or this study, the authors, who routinely perform rhinoplasty surgery, constructed a task tree that served as the foundation for the HTA. Our HTA was designed with 4 main steps: (1) patient evaluation, (2) preparation, (3) intranasal portion (septoplasty and turbinate remodeling), and (4) dorsal/external (rhinoplasty). This framework provided a clear and organized approach to analyzing and evaluating the tasks involved in open structural rhinoplasty.

Patient Evaluation

The first step involves evaluating and planning the surgery for patients seeking rhinoplasty. As shown in Figure 1, this step includes patient assessment, photodocumentation, simulation, discussion of goals, surgical planning, and workup.

Preparation

The second step in our HTA is preparation as shown in Figure 2. This occurs in the operating room before the start of surgery and involves initiating anesthesia, surgical preparation and incisions, and exposure of the nasal skeleton.

Septoplasty

Septoplasty is the third step in our HTA, and this is shown in Figure 3. This step involves the septal flaps, treating septal deformities, and addressing turbinate hypertrophy if present.

Rhinoplasty

The fourth and final step of our HTA is rhinoplasty, as shown in Figure 4. Figures 5, 6, and 7 further detail the substeps of this phase. This step may include: nasal dorsal work,

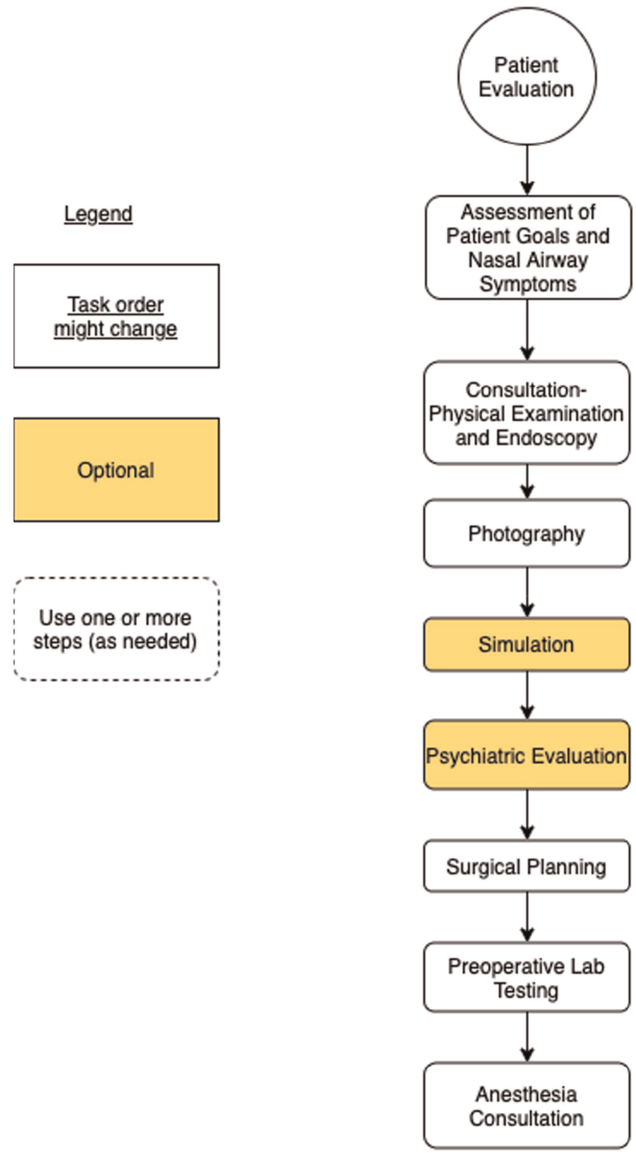

FIGURE 1. Patient evaluation step. The legends on the left side of the figure are applicable to Figures 1–7.

addressing bone and cartilage structures, and nasal tip work, involving grafts and reshaping of tip cartilage.

This stage often requires significant time to ensure proper support for both esthetic and functional outcomes. Techniques will vary depending on the patient's pathology and the surgeon's preferences. During this phase, the septal flaps are approximated, and septal closure is completed. If alar base surgery is part of the surgical plan, it is performed at this stage. The procedure concludes with the closure of all incisions. Finally, splints and dressings are applied to complete the surgery.

Institutional Review Board Approval

The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Demiroglu Bilim University (Protocol no: 44140529 /8743; date: 05.10.2021). All participants consented to participate, and the questionnaire was designed to protect the privacy and anonymity of all responses.

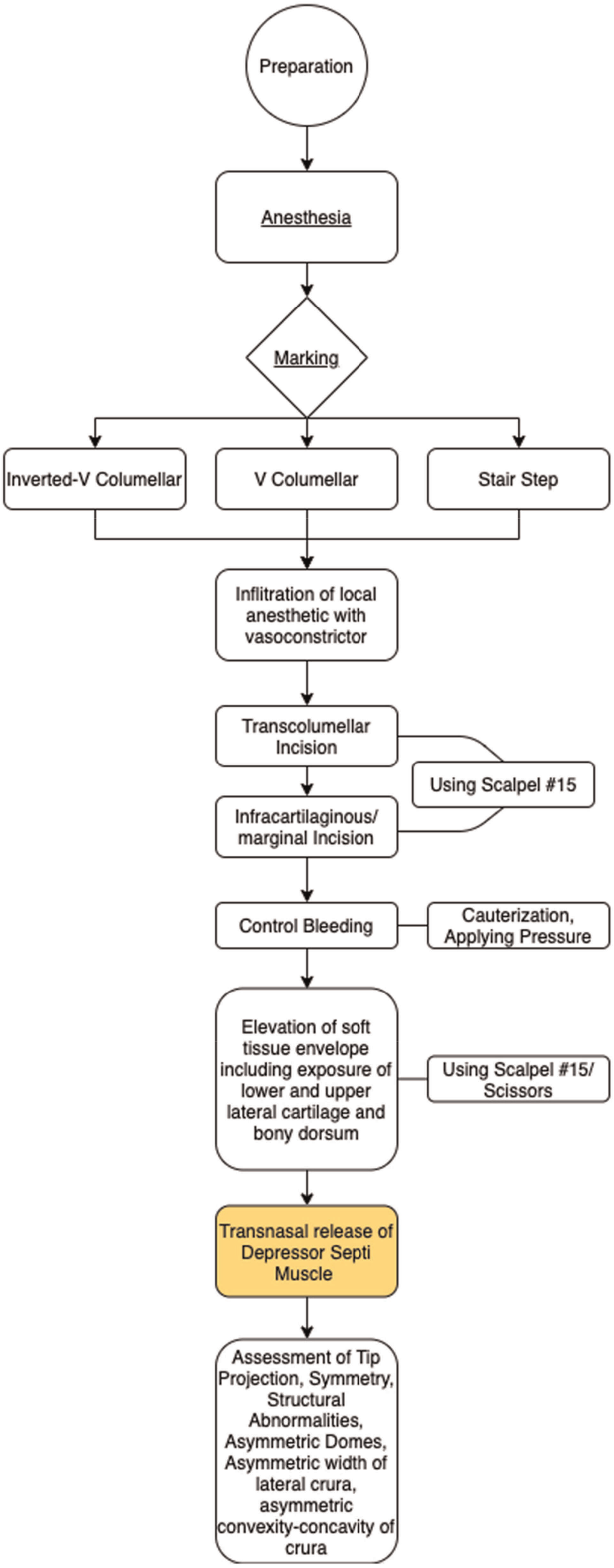

FIGURE 2. Preparation step.

Survey

We surveyed 26 participants, who were recruited through the American Academy of Facial Plastic and Reconstructive Surgery and the Turkish Society of Otolaryngology, Head and Neck Surgery. The survey included questions about the participants' background, experience level, and the amount of time they spent on each task outlined in the HTA (Supplemental Fig. 1, Supplemental Digital Content 1, http://links.lww.com/SCS/I448)

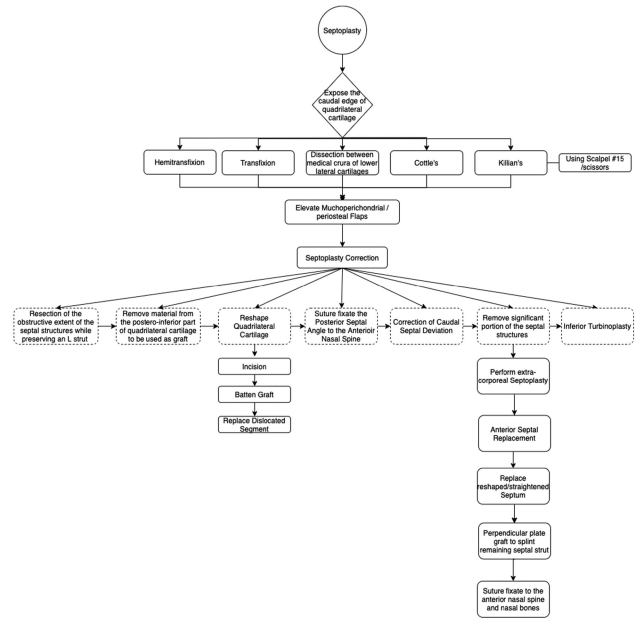

FIGURE 3. Septoplasty step.

Statistical Analysis

For this study, descriptive statistics were calculated using mean, SD, median, minimum, maximum values, frequency, and percentage. The distribution of variables was assessed using the Kolmogorov-Smirnov test. For the comparison of quantitative data, the Mann-Whitney *U* test was employed. Correlations between variables were analyzed using Spearman's correlation analysis. All statistical analyses were performed using SPSS version 28.0.

Machine Learning to Predict Expert Level Based on Time

We hypothesized that the time required to complete specific tasks in rhinoplasty can serve as an indicator of a surgeon's experience level, with a measurable relationship between expertise and task completion times. We further proposed that this relationship could be identified using machine learning algo-

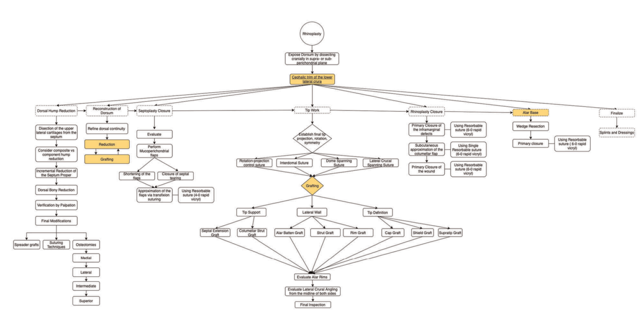

FIGURE 4. Rhinoplasty step.

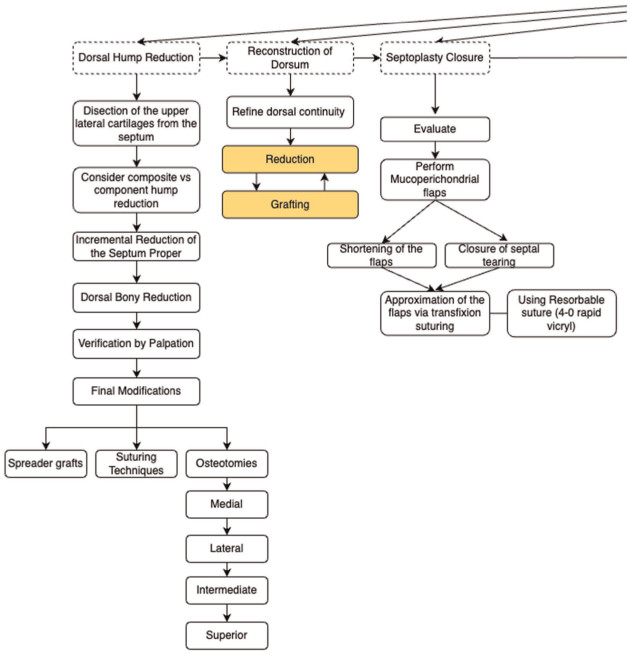

FIGURE 5. Dorsal hump reduction, reconstruction of dorsum, and septoplasty closure subtasks of the rhinoplasty task.

rithms to develop a predictive model, potentially offering greater sensitivity than traditional methods in distinguishing experience levels based on task timing.

To test this hypothesis, we categorized surgeons into 2 groups, “novice” and “experienced,” based on the total number of rhinoplasty cases they had performed. Data preparation involved cleaning the data set and imputing missing values. Spearman correlation analysis was conducted to identify the

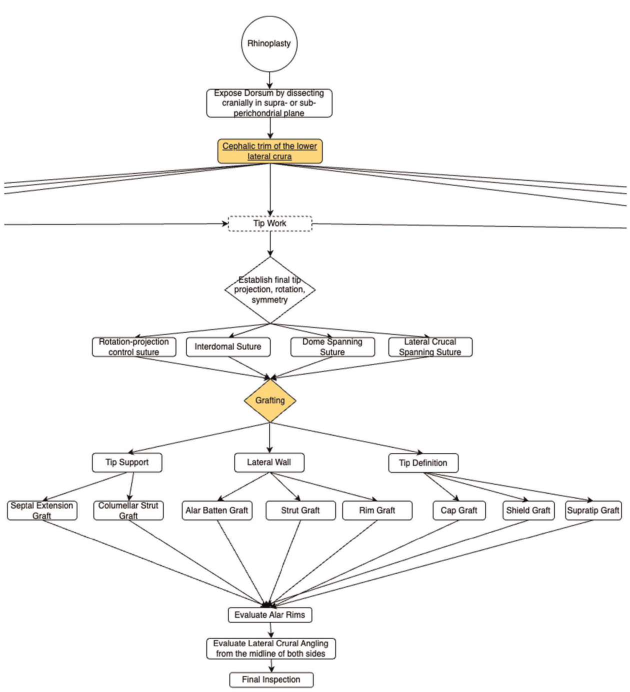

FIGURE 6. Tip work subtask of the rhinoplasty task.

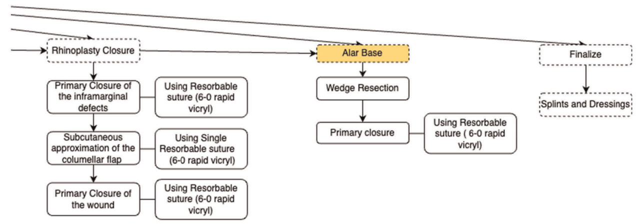

FIGURE 7. Rhinoplasty closure, alar base and finalization of the procedure subtasks of the rhinoplasty task.

strength and direction of associations between task completion times and experience levels, highlighting the tasks most indicative of expertise.

Given the limited size of our data set, we employed statistical techniques to generate synthetic data to enhance the training of the machine learning algorithms. Synthetic data generation is widely used to supplement, augment, or even replace real-world data for training purposes, enabling better model performance and reliability despite constraints in data availability. We generated 200 new sample data to supplement our data using Gaussian Copula.[21,22]

For model evaluation, we used accuracy as the performance metric. The models were trained using Random Forest,[23] Neural Network,[24] LightGBM,[25] and Extremely Randomized Trees.[26]

For Random Forest implementation, we selected hyperparameters to optimize performance and computational efficiency. The model was configured with 300 decision trees to ensure robust aggregation while maintaining manageable training times. The maximum number of leaf nodes per tree was limited to 15,000 to control tree complexity and prevent overfitting. We set a fixed random seed to ensure reproducibility. The model was trained with bootstrapped samples and used the Gini impurity measure to evaluate splits. Extremely Randomized Trees classifier was configured with 300 trees. To manage tree complexity and prevent overfitting, we limited the maximum number of leaf nodes to 15,000. Bootstrapping was enabled to introduce variance in the training samples, while the

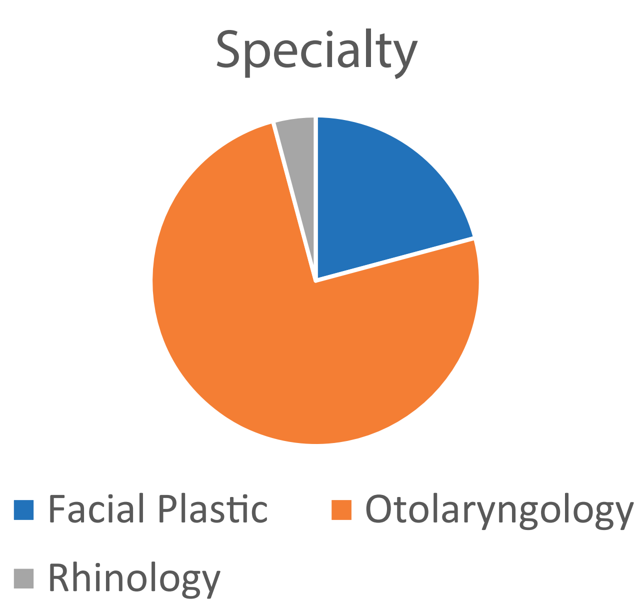

FIGURE 8. Specialty distribution.

criterion for split evaluation was entropy to optimize information gain. Same as Random Forest implementation, a fixed random seed was set to ensure reproducibility.

For LightGBM, the learning rate was set to 0.03, for a balance between convergence speed and predictive accuracy. The number of leaves in each tree was limited to 128, to provide a controlled level of tree complexity. To increase generalization and reduce overfitting, we used a feature subsampling ratio of 0.9 and set the minimum number of samples per leaf to 3. The training consisted of 83 boosting iterations.

Finally, the Neural Network. Our architecture consisted of 4 hidden layers with 128 neurons per layer, using the ReLU activation function for nonlinearity. The training process spanned 17 epochs with a batch size of 32. To prevent overfitting, we used a dropout regularization with a probability of 0.1 and used a weight decay of $1x10^{-6}$. For gradient optimization, we used the Adam optimizer with a learning rate of 0.0003.

RESULTS

The survey results are presented in detail in Supplemental Table 1 (Supplemental Digital Content 2, http://links.lww.com/SCS/I449). Of the 26 respondents who participated in the study, one did not provide any answers, and thus their response set was excluded. Another surgeon only partially completed the questionnaire. The remaining responses were complete and suitable for analysis.

Most of the surgeons who participated in the study were from the otolaryngology specialty (Fig. 8). We categorized the surgeons' experience levels based on the number of rhinoplasty cases they had performed: 0 to 10 cases were classified as novice, 11 to 100 cases as medium-level experience, 101 to 500 cases as experienced, and over 500 cases as very experienced. The majority of participants were classified as either experienced or very experienced; however, a notable proportion (n = 11) were classified as novice (Fig. 9). For statistical analysis, we combined the novice and medium-level groups into a single category, termed "non-expert," while those with experience of 101 cases or more were grouped as "expert." A composite figure has been constructed (Fig. 10).

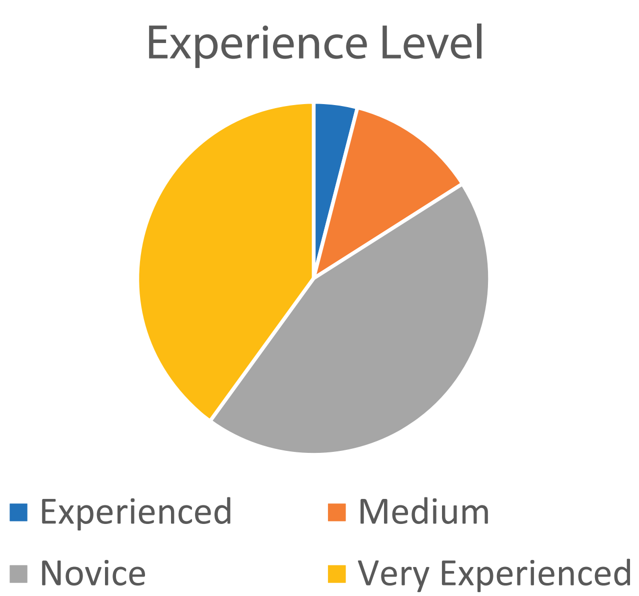

FIGURE 9. Experience level (based on total surgeries) novice: 0–10, medium: 10–100, experienced: 100–1000, very experienced: > 1000.

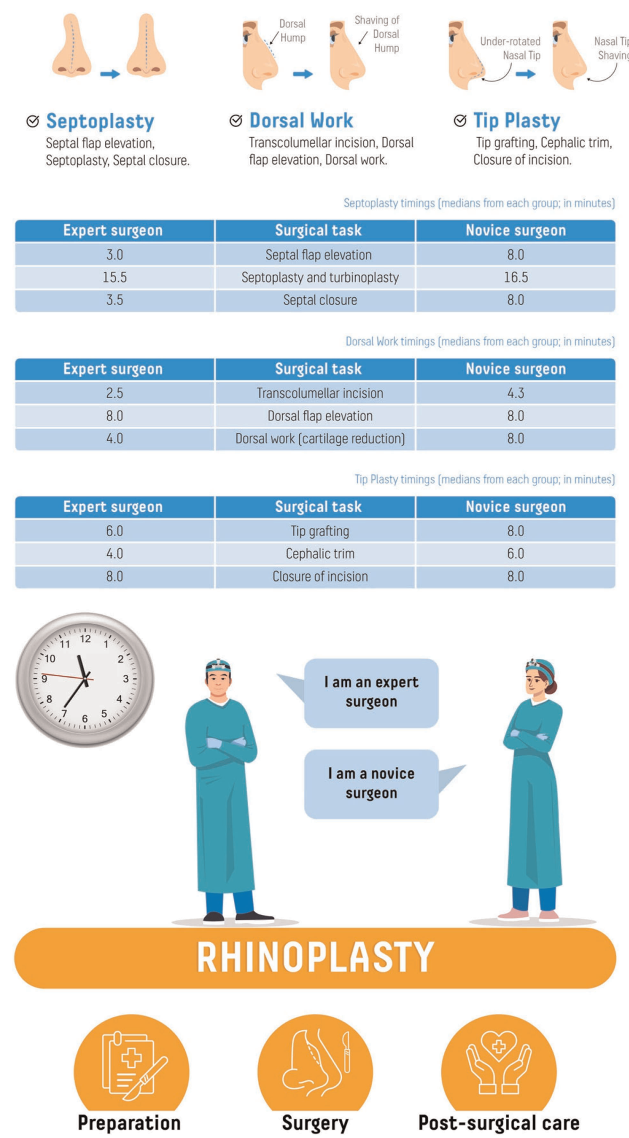

FIGURE 10. A composite figure demonstrating the concept of comparing the durations in minutes spent on different tasks by expert and non-expert surgeons.

Statistical Analysis

The comparison of responses between the expert and non-expert groups is shown in Supplemental Table 2 (Supplemental Digital Content 2, http://links.lww.com/SCS/I449), with significant differences highlighted in bold. Notable differences were observed in responses related to experience, specifically years of practice and the number of surgeries performed or observed. Surgical steps that exhibited significant differences between the two groups included the elevation of (septal) mucoperichondrial-mucoperiosteal flaps, cephalic trim, septoplasty closure, and rhinoplasty closure, with expert surgeons requiring significantly less time for these tasks.

Supplemental Table 3 (Supplemental Digital Content 2, http://links.lww.com/SCS/I449) presents the Spearman correlation analysis, comparing surgical steps with surgeon experience and the number of procedures performed. A negative correlation was found between years of experience and surgical volume, and steps such as cephalic trim, marking, transcolumellar incisions, bleeding control, elevation of the soft tissue envelope, elevation of (septal) mucoperichondrial-mucoperiosteal flaps, inferior turbinoplasty, and septoplasty closure. Four determinants of experience are listed in this table, with various steps showing significant correlations with these determinants. Taken together, the data suggest that more experienced surgeons tend to spend less time on these steps. However, no significant correlations were found between experience/volume and time for the remaining steps, indicating that the number of cases a surgeon has seen did not significantly affect the time spent on many of the tasks.

Machine Learning

Before training the machine learning algorithms, we assessed the correlations between the target variable (expert level) and the input features (timing of each task). The strongest correlations (>0.5) were found for the tasks of elevating (septal) mucoperichondrial-mucoperiosteal flaps, cephalic trim, and septoplasty closure. It is important to note that these correlations were all negative, indicating that as the experience level increased from novice to expert, the time required for these tasks decreased. The test and validation scores of the ML algorithms (Random Forest, Neural Network, LightGBM, and Extra Trees) can be seen in Supplemental Table 4 (Supplemental Digital Content 2, http://links.lww.com/SCS/I449). Also, we evaluated the performance of our machine learning algorithms using precision, recall, and F1-score as metrics. Random Forest achieved the highest precision (0.929), with an F1-score of 0.788 and a recall of 0.684. The Neural Network demonstrated a precision of 0.8, an F1-score of 0.706, and a recall of 0.632. LightGBM exhibited a balanced performance with a precision of 0.778, the highest F1-score of 0.757, and a recall of 0.737. Extremely Randomized Trees showed a strong precision of 0.917 but had a lower F1-score (0.71) and recall (0.579). These results highlight the trade-offs between precision and recall across the algorithms, with LightGBM offering the best overall balance.

We then assessed the contribution of each input feature (timing of tasks) influenced the model's predictions. Based on this analysis, we found that the tasks listed in Supplemental Table 5 (Supplemental Digital Content 2, http://links.lww.com/SCS/I449) had the most significant impact on the model's outcome. These tasks include additional surgical steps such as rhinoplasty closure, injection, transcolumellar incisions, dorsal hump reduction, lateral osteotomies, assessment of lower lateral cartilage, and dorsal hump bone reduction.

DISCUSSION

Hierarchical task analysis is a tool that may be utilized to identify and analyze how surgical tasks are accomplished.[9,12] Previous studies have shown that the time spent on each task may correlate with the skill level and/or experience level of the surgeon.[9] Video recordings of surgery may be utilized for surgical skill assessment.[1,9] However, video recording rhinoplasty is not a common practice,[27] thus we decided to develop our own task tree to be a guide for this HTA study.

Hierarchical task analysis results are generally given according to various scoring metrics, which are generic or subjective.[9,28] In one study, a 4-point ordinal scale described the quality of technical performance for each domain within each task area, with objective descriptors developed from the interviews, error analysis, and steering group refinement.[29]

Open structural rhinoplasty is a surgery of numerous steps, which may vary according to the surgeon or the surgical plan.[16,30] However, there appears to be a general sequence of steps agreed upon by the surgeons in this study, which was included in the analysis. From this, a hierarchical task analysis (HTA) was developed, which, with validation, could serve as a key milestone in the digitalization of rhinoplasty. This HTA can assist in reliably evaluating, executing, and teaching rhinoplasty surgery. Examples of digitalization in rhinoplasty include establishing time frames for each task, as seen in the HTA, standardizing surface anthropometrics before and after surgery,[31] and incorporating intraoperative measurements.[32] In addition, rhinoplasty quality performance measures, developed by the multidisciplinary Rhinoplasty Performance Measure Development Work Group and approved by several relevant associations, have been introduced. These measures include one outcome measure and 3 process measures, and their use in quality initiatives is recommended.[33,34]

Our study, which analyzed surgeons' self-reported times to complete tasks, reveals significant differences between two groups based on years and volume of experience. The steps that showed differences include the elevation of (septal) mucoperichondrial-mucoperiosteal flaps, elevation of the soft tissue envelope, cephalic trim, septoplasty closure, rhinoplasty closure, transcolumellar incisions, control of bleeding, inferior turbinoplasty, and marking. Among these, elevation of (septal) mucoperichondrial-mucoperiosteal flaps and cephalic trim were consistently performed more quickly by expert surgeons, while non-experts (those with fewer than 100 cases) spent more time on these tasks.

The time spent on the "elevation of soft tissue envelope" differed significantly between experts and non-experts based on both total surgeries performed and surgeries in the past 6 months. This may reflect the benefit of both volume and recent experience. Similarly, "control of bleeding" and "transcolumellar incisions" showed significant differences between the two groups based on years of experience and surgeries performed in the last 6 months, suggesting that general experience and recent case load improve surgical dexterity and efficiency. "Inferior turbinoplasty" and "marking" took longer for non-experts, with the time differences correlated with years of experience and the number of surgeries performed in the last 6 months. More experience seems to have shortened the time spent on inferior turbinoplasty, while marking may be more influenced by recent case load, possibly indicating habit development. The time spent on dorsal work, including osteotomies, resections, and tip sutures/grafts (except for cephalic trim), did not differ significantly between the groups, suggesting that these tasks are time-consuming for all levels of expertise. Interestingly, the number of cases performed by the surgeon did not have a statistically significant effect on the time spent on any of the tasks.

In addition to statistical analysis, we applied machine learning (ML) to the survey results to assess whether ML could be more sensitive in determining whether time spent on tasks correlates with experience. After applying ML, tasks such as "injection, transcolumellar incision, dorsal hump reduction, dorsal surgery-lateral osteotomies, assessment of lower lateral cartilages, and dorsal hump bone reduction" were identified as important indicators of expertise, with different variables highlighted compared with the statistical analysis. This difference may be due to ML uncovering higher-order mathematical relationships between the target and input features, whereas correlation analysis primarily detects linear relationships.

We believe that as rhinoplasty becomes more digitized and task-related data is further analyzed, more insights into the relationships and interactions of these tasks will emerge. The combination of HTA and ML will become increasingly beneficial, as the subtasks from HTA will provide more specific data points for detailed analysis through ML. These performance metrics, derived from such studies, will guide trainers and trainees, helping to identify key tasks and subtasks to focus on in training.

Several limitations are present in this study. First, despite our best efforts, we were unable to reach a large number of surgeons, and a larger sample size would have strengthened the study. Another limitation is the subjectivity of the time estimates for each task, as they were based on the surgeons' recall. A more accurate method would involve collecting data through direct surgical observation by dedicated personnel; however, this would be challenging to implement for a large number of surgeons. In future studies, one could record the tasks over a number of cases and also consider evaluating primary versus revision cases, as that may also reveal significant differences between novice and expert surgeons. The task tree could also have been constructed from video recordings of each case. While this is a viable option, it would require obtaining video recordings, addressing patient privacy concerns, and managing additional costs and time constraints. In addition, the incongruence between the results from statistical correlations and machine learning algorithms remains unexplained, and this discrepancy will be explored in future studies.

CONCLUSION

Based on our study, we successfully developed a HTA for open structural rhinoplasty, which highlights key tasks that vary in time according to the surgeon's level of experience. Our findings show that experienced surgeons perform certain tasks, such as the elevation of septal mucoperichondrial-mucoperiosteal flaps, cephalic trim, and septoplasty closure, more efficiently, with significantly less time required compared with non-expert surgeons. These results suggest that specific tasks may be more sensitive to the benefits of increased surgical volume and experience.

In addition to statistical analysis, we applied machine learning techniques, which identified additional variables, such as rhinoplasty closure, injection, and dorsal hump reduction, as important indicators of surgical expertise. This demonstrates the potential of machine learning to detect complex relationships beyond what traditional statistical methods can reveal. Ultimately, our work contributes to the development of a more standardized and evidence-based approach to rhinoplasty training and assessment.

REFERENCES

1. Ghaferi AA, Dimick JB. Considering surgeon skill in randomized clinical trials. *JAMA Surg* 2020;155:599–599
2. Dobratz EJ, Tran V, Hilger PA. Comparison of techniques used to support the nasal tip and their long-term effects on tip position. *Arch Facial Plast Surg* 2010;12:172–179
3. De Ruiter BJ, Lesko RP, Davidson EH. Standardized schematics for facial trauma planning: a clinical education tool. *J Craniofac Surg* 2020;31:1297
4. Ali YH, Alandejani T. Rhinoplasty assisted by endoscopic septoplasty: precise job and an educational tool. *J Craniofac Surg* 2020;31:847–850
5. Guo J, Zhang X, Xu Y, et al. Learning curve analysis of full-length costal cartilage harvesting by plastic surgery residents: a retrospective study. *J Craniofac Surg* 2023;34:e594
6. Mauch JT, Shubinets V, Chang B, et al. Searching for an alternative method to evaluate surgical knowledge: does a drawing and labeling quiz predict in-service exam performance and enhance learning? *J Craniofac Surg* 2020;31:904–907
7. Zammit D, Ponnudurai N, Safran T, et al. Reevaluating the current model of rhinoplasty training and future directions: a role for focused, maneuver-specific simulation. *Plast Reconstr Surg* 2019; 144:597e–605e
8. Stone DH, Upchurch GR, Scali ST. Surgeon credentialing should reflect real-world practice outcomes rather than arbitrary minimum-volume benchmarks. *JAMA Surg* 2021;156:597–598
9. Demirel D, Yu A, Cooper-Baer S, et al. A hierarchical task analysis of shoulder arthroscopy for a virtual arthroscopic tear diagnosis and evaluation platform (VATDEP). *Int J Med Robot* 2017;13:e1799
10. Dials J, Demirel D, Sanchez-Arias R, et al. Skill-level classification and performance evaluation for endoscopic sleeve gastroplasty. *Surg Endosc* 2023;37:4754–4765
11. Lee JA, Close MF, Liu YF, et al. Using intraoperative recordings to evaluate surgical technique and performance in mastoidectomy. *JAMA Otolaryngol Neck Surg* 2020;146:893–899
12. Demirel D, Butler KL, Halic T, et al. A hierarchical task analysis of cricothyroidotomy procedure for a virtual airway skills trainer simulator. *Am J Surg* 2016;212:475–484
13. Daniel RK. The preservation rhinoplasty: a new rhinoplasty revolution. *Aesthet Surg J* 2018;38:228–229. Accessed December 24, 2024. https://academic.oup.com/asj/article-abstract/38/2/228/4793278
14. Patel PN, Abdelwahab M, Most SP. A review and modification of dorsal preservation rhinoplasty techniques. *Facial Plast Surg Aesthetic Med* 2020;22:71–79
15. Toriumi DM. Nasal tip contouring: anatomic basis for management. *Facial Plast Surg Aesthetic Med* 2020;22:10–24
16. Rohrich RJ, Ahmad J. A practical approach to rhinoplasty. *Plast Reconstr Surg* 2016;137:725e–746e
17. Rohrich RJ, Afrooz PN. Primary open rhinoplasty. *Plast Reconstr Surg* 2019;144:102e–117e
18. Çakır B, Öreroğlu AR, Doğan T, et al. A complete subperichondrial dissection technique for rhinoplasty with management of the nasal ligaments. *Aesthet Surg J* 2012;32: 564–574
19. Ponsky D, Eshraghi Y, Guyuron B. The frequency of surgical maneuvers during open rhinoplasty. *Plast Reconstr Surg* 2010;126: 240–244
20. Tasman AJ. Rhinoplasty–indications and techniques. *GMS Curr Top Otorhinolaryngol Head Neck Surg* 2008;6:Doc09
21. Masarotto G, Varin C Gaussian copula marginal regression. Published online 2012. Accessed January 1, 2025. https://projecteuclid.org/journals/electronic-journal-of-statistics/volume-6/issue-none/Gaussian-copula-marginal-regression/10.1214/12-EJS721.short
22. Tasman AJ. Rhinoplasty – indications and techniques. *GMS Curr Top Otorhinolaryngol Head Neck Surg*. 2008;6. Accessed May 22, 2021. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3199847/
23. Rigatti SJ. Random forest. *J Insur Med* 2017;47:31–39
24. Dongare AD, Kharde RR, Kachare AD. Introduction to artificial neural network. *Int J Eng Innov Technol IJEIT* 2012;2:189–194
25. Ke G, Meng Q, Finley T, et al. Lightgbm: a highly efficient gradient boosting decision tree. *Adv Neural Inf Process Syst* 2017;30: 3146–3154
26. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Mach Learn* 2006;63:3–42
27. Celikoyar MM, Aslan CG, Cinar F. An effective method for videorecording the nasal–dorsal part of a rhinoplasty – a multiple case study. *J Vis Commun Med* 2016;39(3-4):112–119
28. Cetinsaya B, Gromski MA, Lee S, et al. A task and performance analysis of endoscopic submucosal dissection (ESD) surgery. *Surg Endosc* 2019;33:592–606
29. Curtis NJ, Foster JD, Miskovic D, et al. Association of surgical skill assessment with clinical outcomes in cancer surgery. *JAMA Surg* 2020;155:590–598
30. Celikoyar MM, Nickas B, Dobratz E, et al. Surgical algorithms in rhinoplasty: a scoping review of the current status. *Aesthetic Plast Surg* 2021;45:2869–2877
31. Celikoyar MM, Pérez MF, Akbaş MI, et al. Facial surface anthropometric features and measurements with an emphasis on rhinoplasty. *Aesthet Surg J* 2022;42:133–148
32. Toriumi DM. New concepts in nasal tip contouring. *Arch Facial Plast Surg* 2006;8:156–185

33. Ishii L, Fedok F, Marcus B, et al. Executive summary: "evidence-based performance measures for rhinoplasty: a multi-disciplinary performance measure set. *Otolaryngol Neck Surg* 2021;165:383–384
34. Manahan MA, Fedok F, Davidson C, et al. Evidence-based performance measures for rhinoplasty: a multidisciplinary performance measure set. *Plast Reconstr Surg* 2021;147:222e–230e

Lateral Support Maneuver: Validation of a Diagnostic Tool Specific to Dynamic Nasal Valve Collapse

Kaitlynne Y. Pak, MD, Alberto Nunez, BSc, and Matthew K. Lee, MD

Background: The modified Cottle (MC) maneuver has limited clinical value due to a lack of specificity. The authors propose a variation utilizing only passive support of the lateral wall in dynamic lateral wall insufficiency (LWI), termed the "lateral support (LS) maneuver".
Objective: Evaluate whether the LS maneuver reduces false positives in LWI diagnosis.
Methods: This cross-sectional study involved 106 volunteers who were evaluated using the NOSE scale, Likert scale, and LWI grade to compare nasal breathing with no intervention (N) versus the MC and LS maneuvers.
Results: In LWI grade 0, the LS showed no significant improvement compared with N. In LWI grade 1 to 3, the LS showed significant improvement in nasal breathing compared with N (paired t test, $P < 0.05$), demonstrating specificity to the presence of LWI. MC had significant improvement in all LWI grades, including grade 0 (false positive). Both maneuvers demonstrated significant improvement across the NOSE severity categories.
Conclusions: Our study demonstrates both the MC and LS maneuvers have high positivity rates across the NOSE severity categories. However, a positive response to the LS maneuver was seen only in those with true, objective LWI (grades 1–3), demonstrating greater specificity than the MC maneuver in assessing dynamic nasal valve collapse.

Key Words: Cottle maneuver, lateral support, modified cottle, nasal valve collapse, nasal valve insufficiency

From the Division of Otolaryngology, Head and Neck Surgery, Cedars Sinai Medical Center, Los Angeles, CA.
Received August 6, 2025.
Accepted for publication August 17, 2025.
Address correspondence and reprint requests to Matthew K. Lee, MD, Division of Otolaryngology, Head and Neck Surgery, Cedars Sinai Medical Center, 8635 W 3rd St #870W, Los Angeles, CA 90048; E-mail: matthew.lee@cshs.org
The authors report no conflicts of interest.
Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.jcraniofacialsurgery.com.

ISSN: 1536-3732
DOI: 10.1097/SCS.0000000000011960

Subjective nasal obstruction can have a substantial impact on quality of life, and can be caused by a variety of factors including a deviated septum, turbinate hypertrophy, chronic rhinosinusitis, allergic rhinitis, or nasal valve collapse (also known as lateral wall insufficiency, or LWI). When LWI is identified, corrective nasal valve surgery can significantly improve disease-specific quality of life.[1,2]

Nasal valve obstruction can be assessed by clinical history and physical examination using the Cottle or modified Cottle maneuver. The Cottle maneuver entails an active, lateral pull of the cheek adjacent to the nose to gently widen or lift the internal nasal valve. The modified Cottle maneuver, on the other hand, involves a cerumen curette or cotton-tip swab being placed within the nares to gently, albeit actively, lateralize the ala and stent the internal nasal valve open. Subjective improvement in nasal airflow and airway resistance is recorded.

Although the Cottle and modified Cottle maneuvers are the current standard diagnostic methods for diagnosing nasal valve obstruction per the American Academy of Otolaryngology-Head and Neck Surgery's (AAO-HNS) Position Statement in 2023, there is a well-established concern that these maneuvers have a low specificity.[3–5] The mechanism of actively lateralizing the internal nasal valve leads to significant subjective improvement in breathing even in those patients without identified nasal valve collapse or reported nasal obstruction. Although multiple studies have assessed the effectiveness of the modified Cottle maneuver in predicting surgical outcomes, a more reliable diagnostic method has not been identified.[3,4,6]

In this study, we propose a revision of the modified Cottle maneuver, termed the "lateral support (LS) maneuver," as a more specific assessment for clinically significant LWI. We used the validated subjective Nasal Obstruction and Septoplasty Effectiveness (NOSE) scale and validated objective grading system for LWI (grade I–III) to quantify subjective nasal obstruction and objective LWI severity. We further delineated the degree of subjective improved nasal breathing with a Likert scale (0–10), completed by the subject without intervention (N), with the modified Cottle (MC), and with the lateral support (LS) maneuver.

METHODS

A cross-sectional, observational survey study was conducted at Cedars-Sinai Medical Center. Institutional review board (IRB) approval was obtained before starting this study.

Subjects included physicians, medical students, surgical technicians, and administrative staff members at Cedars-Sinai Medical Center. A brief survey was provided to collect demographic information and prior nasal surgical history. Baseline NOSE scores were obtained. Participants were asked to grade their subjective nasal breathing using a Likert scale, with 0 indicating complete obstruction and no nasal airflow, and 10 indicating complete nasal patency and superb nasal airflow. The subject was first asked to grade their nasal breathing at baseline without intervention. The examiner then performed a bilateral MC and LS maneuver, and the subject asked to report their nasal breathing using the same Likert scale with each subsequent maneuver The distinction between the 2 maneuvers involved the active, lateral displacement of the lateral nasal wall in the MC maneuver versus the passive support of the lateral wall against dynamic collapse in the LS maneuver (Fig. 1). Paired t testing was used to evaluate the significance between the 2 maneuvers compared with the no intervention score (N): modified cottle to no intervention (MC-N) and lateral support to no intervention (LS-N).