

Development of the “LAsEr Surgical skills Evaluation for Residents” (LASER) Scale Through a Modified Delphi Method

E. Sédillot-Daniel, MD, MS^{1,2} , A. Melkane, MD³, C. Chiesa, MD⁴,
N. Fakhry, MD⁵, V. N. Young, MD⁶, L. J. Smith, MD⁷,
J. R. Lechien, MD, PhD, MS^{8,9,10} , L. Guertin, MD, FRCSC¹,
M. J. Olivier, MD, FRCSC², A. Maniakas, MD, PhD, FRCSC¹¹ ,
R. Jun Lin, MD, FRCSC, MS¹², E. Bissada, MD, FRCSC¹,
A. Christopoulos, MD, FRCSC, MS¹, T. Ayad, MD, FRCSC^{1,2},
and Andrée-Anne Leclerc, MD, MS, FRCSC^{1,2} 

Abstract

Background. The CO₂ laser is frequently used during microlaryngeal surgery (MLS) for a variety of pathology including laryngeal malignancy and stenosis. Learning how to use the laser safely is part of the curriculum for every otolaryngology resident. However, assessment of laryngoscopy technical skills can be challenging for supervisors, making it difficult to adequately provide feedback to trainees.

Objectives. “LAsEr Surgical skills Evaluation for Residents” (LASER) Scale aims to facilitate the evaluation of residents’ performance and promote constructive feedback.

Methods. The initial evaluation grid was based on a literature review of CO₂ laser laryngoscopy (with an emphasis on indications, technique, safety, and efficacy) using Covidence systematic review software (Veritas Health Innovation). The final version was produced after 4 rounds of Delphi surveys.

Results. This study was an international collaboration including 15 otolaryngologists with either laryngology or head and neck surgery subspecialties. Panelists were based in Canada (8), the United States (3), France (1), Spain (1), Belgium (1), and Lebanon (1). The process involved 4 rounds of Delphi surveys. Assessment categories included: anesthesia considerations, pre- and perioperative laser safety measures, and surgical technique. Consensus was reached on final survey completion.

Conclusions. Through a modified Delphi method, a novel scale was developed through an international collaborative effort that evaluates resident skillset in CO₂ laser MLS. Future studies are warranted to validate this assessment tool.

Keywords

laryngology, medical education, resident training, CO₂ laser, microlaryngeal surgery, Delphi

Received February 7, 2024; accepted June 2, 2024.

¹Centre Hospitalier de l’Université de Montréal, Montreal, QC, Canada

²CIUSSS de l’Est de Montréal, Montreal, QC, Canada

³Hotel Dieu de France Hospital, Saint Joseph University, Beirut, Lebanon

⁴Hospital Universitario Donostia, San Sebastian, Spain

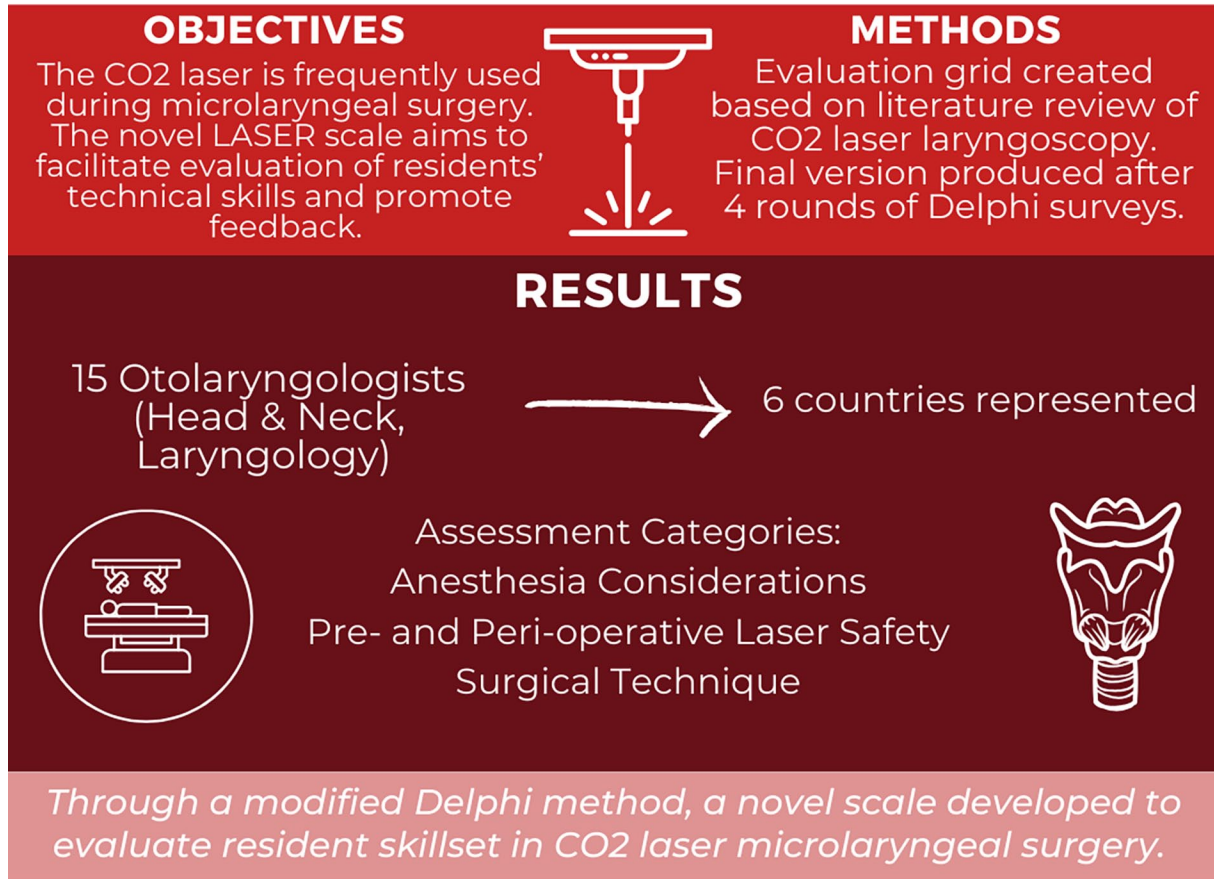
⁵Centre Hospitalier Universitaire La Conception, Aix-Marseille University, Marseille, France



Graphical Abstract

DEVELOPMENT OF THE “LASER SURGICAL SKILLS EVALUATION FOR RESIDENTS” (LASER) SCALE THROUGH A MODIFIED DELPHI METHOD

Sedillot-Daniel E, Melkane A, Chiesa C, Fakhry N, Young VN et al.



CANADIAN SOCIETY OF
OTOLARYNGOLOGY -
HEAD & NECK SURGERY

THE OFFICIAL JOURNAL OF THE CANADIAN SOCIETY OF OTO-HNS

Université 
de Montréal

⁶UCSF Voice and Swallowing Center, Department of Otolaryngology—Head and Neck Surgery, University of California San Francisco, San Francisco, CA, USA

⁷UPMC Voice Center, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

⁸University of Mons, Mons, Belgium

⁹Polyclinic of Poitiers, Elsan Hospital, Poitiers, France

¹⁰EpiCURA Hospital, Baudour, Belgium

¹¹The University of Texas MD Anderson Cancer Center, Houston, TX, USA

¹²Unity Health Toronto, St. Micheal's Hospital, Toronto, ON, Canada

Corresponding Author:

Andrée-Anne Leclerc, MD, MSc, FRCSC, Division of Otolaryngology—Head and Neck Surgery, Hôpital Maisonneuve-Rosemont, CIUSSS de l'Est de Montréal, 5415 Assomption Boulevard, Montreal, QC HIT 2M4, Canada.

Email: andree-anne.leclerc.1@umontreal.ca

Background

Performance assessment tools can provide excellent constructive feedback for trainees. They also tend to reduce biases by allowing improved objectivity. In fact, studies show that trainees are more accepting of lower evaluation scores when being assessed with standardized evaluation grid.¹ Rating scales are an example of detailed evaluation ensuring accurate feedback and encouraging repeated assessments.²

Few operative assessment tools exist for otolaryngology—head and neck surgery residents and the faculty who assess them, such as in parotidectomy, neck dissection, endoscopic sinus surgery, and nasal septoplasty.³⁻⁶ Unfortunately, no assessments exist for microlaryngeal surgery (MLS) or laser procedures despite the frequency of these surgeries within the otolaryngology residency curriculum. This particular evaluation grid specifically assesses CO₂ laser procedures, but it could also be applied to all types of lasers. CO₂ laser MLS can be challenging to teach and learn due to limited visualization of the surgical field and multiple laser safety measures. A task-specific evaluation grid ensuring a detailed assessment of residents' skills set and constructive feedback could improve both trainee performance and patient safety.⁷ This is especially relevant in the current Competency-Based Medical Education (CBME) era where emphasis is made on multiple tasks assessments.

The Delphi methodology is a systematic group communication technique developed with a series of questionnaires or interviews. This process aims to find a consensus in a subject requiring decision-making.⁸ It has previously been used successfully to develop assessment tools such as a neck dissection rating scale.³

The purpose of this study is to develop an evaluation grid related to CO₂ laser safety and technical abilities during MLS.

Methods

An international consensus-building process was used to develop this new teaching tool: the “LASER Surgical skills Evaluation for Residents” (LASER) Scale.

The initial version of the evaluation grid was based on a review of the literature (to be published separately) and the senior author's (A.A.L.) experience. The systematic review focused on general CO₂ laser technique during microlaryngeal laser surgery as well as its safety measures. Covidence software (Veritas Health Innovation) was used for data collection in the systematic review. A total of 141 studies were included into this systematic review. The systematic review will be published independently from this article.

Four rounds of Delphi surveys were carried out to develop the evaluation form with the participation of an international panel of experts. The grid was developed in English to enhance its global applicability. Otolaryngologists were solicited worldwide by email. Recruitment was on a voluntary basis only. Experts were either fellowship-trained laryngologists or head and neck surgeons within the senior authors' (A.A.L. and T.A.) professional network.

The Delphi process took place between February and June 2022. The REDCap platform (Research Electronic Data Capture tools hosted at CIUSSS de l'Est de l'Ile de Montréal) was used for data collection and management for each Delphi survey. At the beginning of each round, a Delphi survey link was sent by email to participants. Due to firewall concerns, this platform was only available to Canadian-based surgeons. International participants instead received an email with an attached PDF taken from the same Delphi survey on REDCap. Email reminders were also sent if necessary. Participants had an average delay of 2 weeks for completing each survey. All surveys were written in English. A Student's *t* test was used to calculate the *P* value between the level of consensus after the second and the fourth survey.

Round 1

The first Delphi round introduced the first version of the evaluation grid, comprising 5 main sections: anesthesia considerations, preoperative safety measures, surgical technique, perioperative safety measures, and global appreciation. Each section contained specific assessment objectives totaling 39 candidate statements. Participants were asked to provide suggestions for new statements and free-text modifications of the existing statements if deemed necessary. During this round, no statements were removed from the grid. Participants were also asked to describe their practice and professional background (age, gender, fellowship fields, years of practice, and site of practice).

Round 2

In the second Delphi survey, experts were asked their opinion on new and modified statements suggested in the previous round. Fifty-five statements were divided in the same 5 main categories. Experts had to rate each statement according to a 5-point Likert scale. Relevance scores (respective Likert points) were graded between strongly disagree (1), disagree (2), neutral (3), agree (4), or strongly agree (5). An ordinal scale was chosen to ensure clear interpretations of statements and acknowledgment of the ratings' impact. At the end of each section, participants had the opportunity to write additional recommendations or modifications. Participants were also specifically asked to regroup or rephrase statements if needed.

Round 3

After removing the statements with an average Likert score of 3 or lower during round 2, there were 45 remaining statements. Modifications were applied to some statements, following previous suggestions. In the third round, each participant received an individualized survey where they had to reconsider their previous answers in light of the group's collective responses, detailed in percentage of agreement and disagreement, for each statement (Figure 1). Again, participants were also specifically asked to regroup or rephrase statements if needed.

Group responses (% in favor, % against) Round 2 answers ● If you want to change your previous answer, write it down (by hand) or in a different color	Ensures warning sign and eye protection goggles are installed at OR entrance (92% in favor, 8% against)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
	Ensures water is readily available in case of fire (88% in favor, 15% against)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
	Ensures fire extinguisher is available in the OR (54% in favor, 46% against)	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

Figure 1. Example of the round 3 survey.

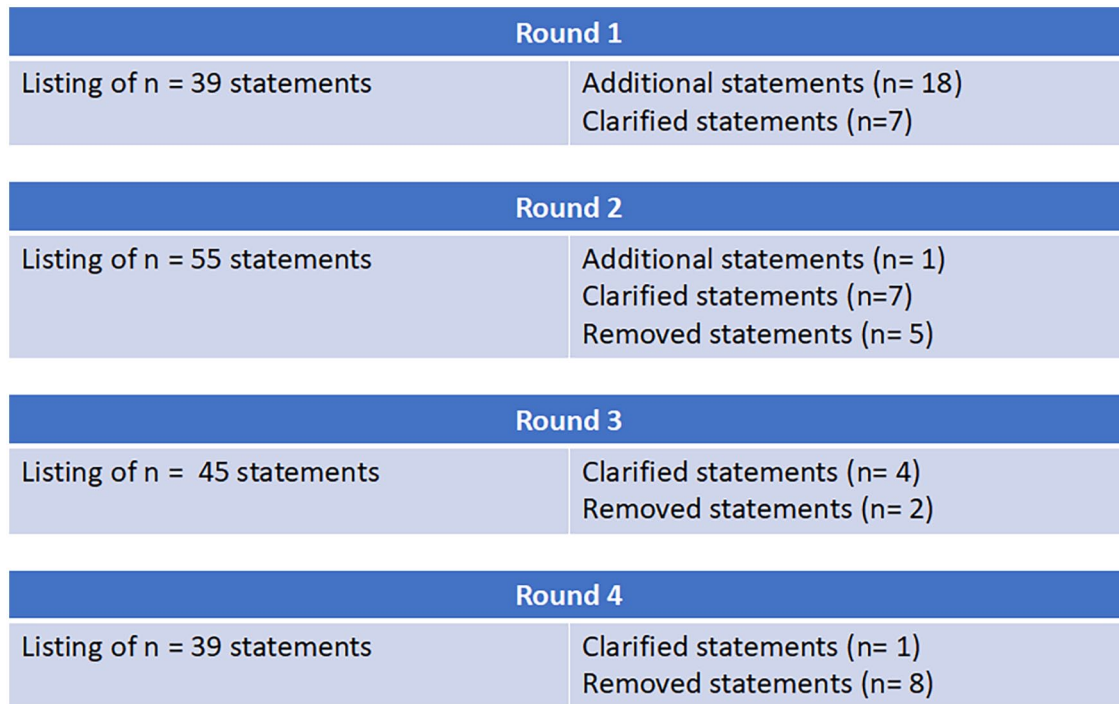


Figure 2. Flow chart of evaluation grid evolution.

Round 4

The fourth survey was very similar to the previous one, in the form of the same individualized questionnaires updated with the results from the third round. Statements with an average Likert score of 3 or lower were removed and previous suggestions were applied, totaling 39 statements (Figure 2). At the end of the survey, participants were asked if they would agree if this version was the final evaluation grid. An optional section was always available at the end of each section for additional recommendations or modifications.

Results

Participants' Demographics Characteristics

Of the 18 otolaryngologists who were contacted, 15 (83%) responded affirmatively. The majority were men (67%). Geographic distribution consisted of 8 surgeons based in Canada, 3 in the United States of America, 1 in France, 1 in Belgium, 1 in Spain, and 1 in Lebanon. Mean age was 45 years old. Nine had completed a head and neck surgery fellowship (60%), five (33%) had a laryngology fellowship, and one was

trained in both. The average practice period after completing a fellowship was 13 years. The mean annual CO₂ laser MLS procedure was 31 cases/participant with an average of 6 residents per expert annually supervised during these procedures. Unfortunately, 2 (13%) panelists were unable to complete the entire process.

Final Evaluation Grid

The final evaluation grid comprised 30 statements divided into 4 categories: anesthesia considerations, preoperative CO₂ laser safety measures, surgical technique, and perioperative CO₂ laser safety measures (Figure 3). To reduce completion time and redundancy, the initially adopted "Global Appreciation" section was removed as suggested by 6 out of 13 experts. For a similar reason, the statement "Demonstrates knowledge of safety measures in case of fire" was also removed. As suggested, the statements "Awareness of laser beam direction" and "Awareness and adjustment of laser power output, spot size and duration of exposure on the same area" were regrouped as: "Precise manipulation and awareness of spot size, duration and direction."

	Unable to perform step	Performs step with much help	Performs step with some help	Performs step with minimal help	Performs step easily and with fluidity	
Anesthesia Considerations						
Discussion of ventilation strategies (preferred and alternatives)	1	2	3	4	5	N/A
Professional and effective communication with anesthesia team (ie, regarding anesthetics administration or techniques)	1	2	3	4	5	N/A
Ensures laser resistant endotracheal tube is used	1	2	3	4	5	N/A
Ensures endotracheal tube balloon is filled with saline/methylene blue/ fluorescein-dyed saline if applicable	1	2	3	4	5	N/A
Ensures O ₂ concentration of 30–35% or lower if applicable	1	2	3	4	5	N/A
Preoperative CO₂ Laser Safety Measures						
Environment Related Considerations						
Ensures team members and staff are wearing appropriate eye protection	1	2	3	4	5	N/A
Ensures appropriate masks (N95, high filtration) are available if applicable	1	2	3	4	5	N/A
Ensures warning sign and eye protection goggles are installed at OR entrance	1	2	3	4	5	N/A
Ensures water is readily available in case of fire	1	2	3	4	5	N/A
Ensures laser pedal is on a dry surface and that only one pedal is within reach for surgeon	1	2	3	4	5	N/A
Patient Related Considerations						
Ensures patient's eyes, skin and hair are protected	1	2	3	4	5	N/A
Installation of subglottic protection (ex. moist cottonoids) when applicable	1	2	3	4	5	N/A
Laser Related Considerations						
Performs adequate laser safety time-out	1	2	3	4	5	N/A
Sets adequate laser parameters and ensures laser on "safe" mode at the beginning of the procedure	1	2	3	4	5	N/A
Ensures laser is tested and works adequately before use	1	2	3	4	5	N/A

Figure 3. (continued)

Surgical Technique						
Laryngoscopy and Exposure						
Ensures proper patient positioning and placing of mouth guard	1	2	3	4	5	N/A
Execution of atraumatic laryngoscopy technique	1	2	3	4	5	N/A
Ability to obtain good initial exposure and to adapt if necessary (securing suspension, counter-pressure, laryngoscope types and sizes, etc.)	1	2	3	4	5	N/A
Inspection of disease location and extension	1	2	3	4	5	N/A
Procedure Performance						
Manipulation of tissues with appropriate instruments	1	2	3	4	5	N/A
Dissection in appropriate plane	1	2	3	4	5	N/A
Adequate traction and additional manipulations if needed	1	2	3	4	5	N/A
Regular removal of char debris	1	2	3	4	5	N/A
Maintains good hemostasis	1	2	3	4	5	N/A
Orientation of surgical specimen if applicable	1	2	3	4	5	N/A
Execution of technique in an ergonomic position	1	2	3	4	5	N/A
Perioperative CO ₂ Laser Safety Measures						
Precise manipulation and awareness of laser spot size, direction and duration of exposure	1	2	3	4	5	N/A
Ensures smoke suctioning throughout the procedure	1	2	3	4	5	N/A
Asks for the laser to be on safe mode when it is not used	1	2	3	4	5	N/A
Ensures tonsillar pillars, oral mucosa, teeth, gingiva, and lips are adequately inspected and that pledgets, mouth props or other instruments are removed at the end of the procedure	1	2	3	4	5	N/A

Figure 3. LASER rating scale in its final version. LASER, LAsER Surgical skills Evaluation for Residents.

Consensus After Delphi Process

A 100% global consensus was reached at the end of the Delphi process regarding the content of the evaluation grid. It was evaluated with the answers to the question: "Do you agree if this latest version is the final evaluation grid?" at the end of the last survey.

For each section, general agreement was also obtained (see Table 1). To illustrate the consensus-building process, the evolution from the second to the last round is highlighted ($P=.209$). Individual statements and sections were modified or removed, as suggested by panelists. After the fourth round, all statements received an average Likert score of 4.25 or higher (according to following system:

Table 1. Consensus on the LASER Evaluation Grid.

Sections	Consensus Delphi 2 (%)	Consensus Delphi 4 (%)
Anesthesia considerations	89.2	98.7
Preoperative CO ₂ laser safety measures	91.3	87.9
Surgical technique	95.8	97.0
Preoperative CO ₂ laser safety measure	87.2	90.8

Abbreviation: LASER, LAser Surgical skills Evaluation for Residents.

Regarding the Delphi process, what are your thoughts on:			
	Not enough	Adequate	Too much
Number of surveys (4 rounds in total)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Response delays for survey completion (approx. 2 weeks)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
What are your thoughts on the REDCap platform (Canadian participants only) and the PDF exchanges?			
<input type="radio"/> Adequate <input type="radio"/> Inadequate			
After its validation, do you plan to use the LASER Rating Scale ?			
<input type="radio"/> Yes <input type="radio"/> No			
Do you have suggestions to improve the Delphi process?			
Do you have any other comments regarding the Delphi process?			

Figure 4. Delphi satisfaction survey.

strongly disagree = 1, disagree = 2, neutral = 3, agree = 4, or strongly agree = 5).

After the final round, the most divided assessment objectives remaining were: “Performs adequate laser safety time-out” (average Likert score of 4.25), “Ensures laser pedal is on a dry surface and that only one pedal is within reach for surgeon” (4.27), “Ensures appropriate masks (N95, high filtration) are available if applicable” (4.33), and “Ensures tonsillar pillars, oral mucosa, teeth, gingiva, and lips are adequately inspected and that pledgets, mouth props or other instruments are removed at the end of the procedure” (4.33).

Post Delphi Satisfaction Survey

According to the post Delphi satisfaction survey (Figure 4), all participants found the number of rounds (4) to be adequate. Ten experts (80%) agreed that an average of 2 weeks was reasonable for survey completion. The use of the REDCap platform was appreciated by all Canadian panelists. Most important, 92% plan to use the LASER rating

scale after its validation. One participant was uncertain because the final rating scale would have to be presented to his colleagues. In a continuous improvement approach, the only suggestion mentioned was the ability to complete online surveys for everyone.

Discussion

The Delphi method aims to recreate a structured group communication, ensuring anonymity, controlled feedback, and statistical data.⁸ It has been used in multiple disciplines such as finance, social studies, and medicine to create a consensus-building process.⁹ For instance, it was used recently to establish an expert practice statement by the American Rhinology Society for skull base reconstruction.¹⁰ More similar to this project, it was also utilized in developing a checklist for pediatric esophagoscopy. Such tools enable trainees to concentrate on crucial steps of the procedure and assist training programs in standardizing the evaluation of trainees.¹¹

The Delphi process has various advantages and disadvantages. First, it is a flexible method that can be utilized across different sectors (eg, healthcare, finance, psychology, etc), and the number of rounds can be adjusted, if necessary, to reach a consensus. The questionnaire format also allows for gathering the opinions of experts who might not have interacted otherwise due to various reasons (geographical location, network connections, scheduling constraints, etc). These surveys also ensure freedom of expression, honesty, and creativity while maintaining anonymity. Particularly during this study's timeframe, the Delphi method facilitated the formation of an international panel. Furthermore, it enables distinguishing topics where there is consensus from those that are more controversial. Ultimately, it is a cost-effective means of collecting multiple opinions on a specific issue.¹²

It is important to note that Delphi literature does not state an ideal number of rounds.¹³ A 4-round Delphi process was, in the authors' opinion, the best way to develop an evaluation grid with an international collaborative effort. This is also reflected in the post Delphi satisfaction survey. Consensus was improved throughout the course of the Delphi, but the *P* value was not statistically significant between the second and fourth rounds (both rates being high). Overall, the development of the LASER evaluation grid was made possible after a successful Delphi process, especially when considering that an acceptable consensus rate is 75% in the literature.¹⁴

This is the first evaluation grid on the technique and safety precautions regarding CO₂ laser application in MLS. The LASER evaluation grid has the potential to offer a detailed evaluation of residents' performance and can be repeated to portray the evolution of their surgical abilities. Similar to the Ottawa Surgical Competency Operating Room Evaluation (OSCORE)¹⁵ and Objective Structured Assessment of Technical Skills (OSATS),¹⁶ the LASER evaluation grid is based on a 5-point scoring system, ranging from "Unable to perform step" to "Performs step easily and with fluidity" (Figure 3). The 5-point Likert scale allows for a comprehensive assessment of residents' skills and understanding, enabling us to capture a range of situations. It allows for granularity and depth in assessing competency, especially in situations where the answer may not be immediately apparent. The 30 evaluated steps aim to offer improved assessments and detailed feedback to residents to optimize their surgical training. Since residents' perspectives on CBME are favorable and CBME has been proven efficient for improving students' skill sets and patient safety, the authors anticipate that the LASER rating scale will also be well received by trainees.⁸ However, the evaluation grid's acceptability will further be assessed during the validity phase.

Interestingly the sections "Surgical Technique" and "Anesthesia Considerations" gained the highest consensus scores (97.0% and 98.7%, respectively). This raises the idea that pre- and perioperative CO₂ laser safety measures are less uniformized internationally. Technical skills set seems to be more homogeneous, whereas safety precautions may vary more according to surgeons' background.

There are a few limitations intrinsic to the Delphi methodology. First, a 4-round process can be time-consuming for panelists. However, since this is the first evaluation grid on CO₂ laser MLS, the authors anticipated the need for 4 surveys: the first, to present the initial assessment objectives based on literature and to collect suggestions and modifications; the second, to grade the statements according to a relevance score; the third, to match the individual answers to the group's collective response; and the last, to confirm if a consensus has been attained. Regardless of being a time-consuming process in general, the participation rate was high (83 %).

The selection of the panelists could have been slightly biased because the sample was based on the senior authors' professional network. However, to limit biases, worldwide otolaryngologists were invited to contribute to the Delphi process. Regardless of their relationships, they are all experts in laryngology or head and neck surgery practicing according to standards of care in their discipline.

The number of assessed objectives was raised as a potential issue to limit the acceptability of the evaluation grid. A few panelists indicated that while many statements seemed relevant, the evaluation grid tended to be redundant when taken as a whole. Therefore the "Global Appreciation" section was removed. After the last round, the grid was downsized to 30 assessment objectives, and it now takes approximately 2 minutes to complete. Challenges exist in trying to find the right balance between crafting an evaluation grid to provide detailed feedback to trainees and making a tool easily used during nonsimulated operative time. Nevertheless, the length of the final rating scale seems to be an adequate compromise.

Conclusion

The LASER is the first evaluation tool assessing CO₂ laser MLS technical skills and safety precautions, created following an international Delphi process with experts in laryngology and head and neck surgery. With this new tool in otolaryngology, the authors aspire to facilitate the assessment of residents' performance and promote constructive feedback exchanges. Given that standardized evaluation grids are increasingly employed in medical training since the introduction of CBME, the authors anticipate a potential positive impact on residents' performance with CO₂ laser cases in MLS. Future work is warranted to validate this evaluation tool.

Acknowledgments

The authors would like to thank Dr Jean-Claude Tabet for his contribution in the first Delphi survey and Miss Vivianne Landry for her participation in the systematic review.

Author Contributions

E.S.D., A.A.L., and T.A. submitted the protocol to the ethic committees and designed the Delphi process. E.S.D. managed the Delphi surveys under the supervision of A.A.L. E.S.D. collected and analyzed the data and wrote the initial version of the manuscript. All authors read and revised the final manuscript.

Consent for Publication

Not applicable.

Data Availability

The datasets used and analyzed in this study are available from the corresponding author on reasonable request.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethics Approval and Consent to Participate

This study was approved by the ethics committees of the CIUSSS de l'Est de l'Île de Montréal and the Centre Hospitalier de l'Université de Montréal (CHUM). MP-12-2022-2729.

ORCID iDs

E. Sédillot-Daniel  <https://orcid.org/0000-0003-2916-903X>

J. R. Lechien  <https://orcid.org/0000-0002-0845-0845>

A. Maniakas  <https://orcid.org/0000-0002-4354-2841>

Andrée-Anne Leclerc  <https://orcid.org/0000-0003-2398-8018>

References

1. Robinson T, Wagner N, Szulewski A, Dudek N, Cheung WJ, Hall AK. Exploring the use of rating scales with entrustment anchors in workplace-based assessment. *Med Educ.* 2021;55(9):1047-1055. doi:10.1111/medu.14573
2. Fahim C, Wagner N, Nousiainen MT, Sonnadara R. Assessment of technical skills competence in the operating room: a systematic and scoping review. *Acad Med.* 2018;93(5):794-808. doi:10.1097/ACM.0000000000001902
3. Mercier É, Guertin L, Bissada E, et al. Assessment of surgical competence for neck dissection: a pilot study. *Can J Surg.* 2022;65(2):E178-E187. doi:10.1503/cjs.018020
4. Obeid AA, Al-Qahtani KH, Ashraf M, Alghamdi FR, Marglani O, Alherabi A. Development and testing for an operative competency assessment tool for nasal septoplasty surgery. *Am J Rhinol Allergy.* 2014;28(4):e163-e167. doi:10.2500/ajra.2014.28.4051
5. Marglani O, Alherabi A, Al-Andejani T, Javer A, Al-Zalabani A, Chalmers A. Development of a tool for Global Rating of Endoscopic Surgical Skills (GRESS) for assessment of otolaryngology residents. *B-ENT.* 2012;8(3):191-195.
6. Yang N, Mercier É, Guertin L, et al. Development, translation, and preliminary validation of the neck dissection assessment tool. *Head Neck.* 2021;43(8):2307-2315. doi:10.1002/hed.26696
7. Martin M, Vashisht B, Frezza E, et al. Competency-based instruction in critical invasive skills improves both resident performance and patient safety. *Surgery.* 1998;124(2):313-317.
8. Rowe G, Wright G. Expert opinions in forecasting: the role of the Delphi technique. In: Armstrong JS, ed. *Principles of Forecasting: A Handbook for Researchers and Practitioners.* Kluwer Academic Publishers; 2001:125-144.
9. Schmalz U, Spinler S, Ringbeck J. Lessons learned from a two-round Delphi-based scenario study. *MethodsX.* 2020;8:101179. doi:10.1016/j.mex.2020.101177
10. Douglas JE, Adappa ND, Choby G, et al. American Rhinologic Society expert practice statement part 1: skull base reconstruction following endoscopic skull base surgery. *Int Forum Allergy Rhinol.* Published online March 30, 2024. doi:10.1002/alr.23346
11. Faucett EA, Wolter NE, Balakrishnan K, et al. Competency-based assessment tool for pediatric esophagoscopy: International Modified Delphi Consensus. *Laryngoscope.* 2021;131(5):1168-1174. doi:10.1002/lary.29126
12. Iqbal S, Pilon-Young L. The Delphi method. *Psychologist.* 2009;22(7):598-601.
13. Aghimien DO, Aigbavboa CO, Oke AE. Critical success factors for digital partnering of construction organisations—a Delphi study. *Eng Constr Archit Manag.* 2020;27(10):3171-3188. doi:10.1108/ECAM-11-2019-0602
14. Chichetti JV. A two-round Delphi study examining consensus of recommended clinical practices for patients with ventricular assist devices as destination therapy. *Prog Transplant.* 2011; 21(1):15-26. doi:10.7182/prtr.21.1.v32706724w42k27
15. Saliken D, Dudek N, Wood TJ, MacEwan M, Gofton WT. Comparison of the Ottawa Surgical Competency Operating Room Evaluation (O-SCORE) to a single-item performance score. *Teach Learn Med.* 2019;31(2):146-153. doi:10.1080/10401334.2018.1503961
16. Hatala R, Cook DA, Brydges R, Hawkins R. Constructing a validity argument for the Objective Structured Assessment of Technical Skills (OSATS): a systematic review of validity evidence. *Adv Health Sci Educ Theory Pract.* 2015;20(5):1149-1175. doi:10.1007/s10459-015-9593-1