Integration of a deep learning basal cell carcinoma detection and tumor mapping algorithm into the Mohs micrographic surgery workflow and effects on clinical staffing: A simulated, retrospective study

Rachael Chacko, BA,[a] Matthew J. Davis, MD,[b] Joshua Levy, PhD,[a,b] and Matthew LeBoeuf, MD, PhD[a,b]

Background: Artificial intelligence (AI) enabled tools have been proposed as 1 solution to improve health care delivery. However, research on downstream effects of AI integration into the clinical workflow is lacking.

Objective: We aim to analyze how integration of an automated basal cell carcinoma detection and tumor mapping algorithm in a Mohs micrographic surgery unit impacts the work efficiency of clinical and laboratory staff.

Methods: Slide, staff, and histotechnician waiting times were analyzed over a 20-day period in a Mohs micrographic surgery unit. A simulated AI workflow was created and the time differences between the real and simulated workflows were compared.

Results: Simulated nonautonomous algorithm integration led to savings of 35.6% of slide waiting time, 18.4% of staff waiting time, and 18.6% of histotechnician waiting time per day. Algorithm integration on days with increased reconstruction complexity resulted in the greatest time savings.

Limitations: One Mohs micrographic surgery unit was analyzed and simulated AI integration was performed retrospectively.

Conclusions: AI integration results in reduced staff waiting times, enabling increased productivity and a streamlined clinical workflow. Schedules containing surgical cases with either increased repair complexity or numerous tumor removal stages stand to benefit most. However, significant logistical challenges must be addressed before broad adoption into clinical practice is realistic. (JAAD Int 2024;15:185-91.)

Key words: artificial intelligence; clinical research; deep learning; general dermatology; medical dermatology; oncology; surgery.

INTRODUCTION

The introduction of artificial intelligence (AI) into medicine has long promised to improve health care delivery and the experiences of both patients and providers. The prevalence of AI applications within health care is increasing rapidly, as evidenced by the

From the Geisel School of Medicine at Dartmouth, Hanover, New Hampshire[a]; and Department of Dermatology, Dartmouth Health, Lebanon, New Hampshire.[b]

Drs Levy and LeBoeuf contributed equally as senior authors on the manuscript.

Funding sources: Dr Levy is supported by NIH grants R24GM141194, P20GM104416 and P20GM130454.

Patient consent: Not applicable.

IRB approval status: Not applicable.

Accepted for publication February 8, 2024.

Correspondence to: Matthew LeBoeuf, MD, PhD, Department of Dermatology, Dartmouth Health, 18 Old Etna Rd, Lebanon, NH 03766. E-mail: Matthew.LeBoeuf@hitchcock.org.

2666-3287

https://doi.org/10.1016/j.jdin.2024.02.014

343 Food and Drug Administration-approved AI-enabled medical devices currently in clinical practice.[1] AI-enabled medical devices aim to provide diagnostic expertise where not otherwise able, improve clinicians' diagnostic accuracy, and increase diagnostic efficiency while reducing fatiguability of redundant tasks.[2,3] Despite the numerous studies on the accuracy of AI-based algorithms, research regarding implementation of these algorithms into clinical w orkflows and the downstream effects on improving staff efficiency is lacking.

Evolving over the last decade, nationwide staffing shortages and increased staffing costs have presented a major challenge to health care delivery.[4] An aging population with increased health care needs will continue to drive demand as highlighted by the United States Bureau of Labor Statistics' estimate that over 195,000 additional registered nurses and 112,000 medical assistants will be necessary by 2031.[5,6] The introduction of AI to maximize the utility of available staffing resources is 1 approach to modify staffing demands. Automated generation of actionable data will maximize the ability of a staff member to complete tasks when they are ready to do so. Improved understanding and study of the downstream effects of AI on process and staff efficiency may provide a solution to addressing the increased staffing demands of health care.

Mohs micrographic surgery (MMS) relies on seamless integration of surgeons, nursing staff, and histotechnicians to provide real-time surgical and histologic care for tumor removal and defect reconstruction. This real-time health care delivery approach offers a unique model system to study how implementation of AI-driven algorithms can improve clinical workflow efficiency. Throughout each day, rate-limiting steps including tissue processing, histologic analysis, and defect reconstruction can impact MMS efficiency and result in increased patient and staff waiting times. AI provides an opportunity to further streamline this process by providing earlier diagnostic information for staff to act on (Supplementary Fig 1, available via Mendeley at https://doi.org/10.17632/4yv8zg4k2p.1).

We have developed an AI-driven algorithm that provides intraoperative tissue grossing and inking recommendations, tissue section quality assessments, histologic basal cell carcinoma (BCC) identification, and tumor map generation to inform additional tumor removal, described further in the methods section below.[7] This algorithm has been shown to rapidly and accurately localize tumor with an area under the curve (AUC) of 0.97. This algorithm was developed with the purpose of reducing the time for the histotechnician to process tissue and the Mohs surgeon to perform histologic examination of the tissue margins and subsequent tumor mapping, which can provide the nursing staff with earlier insight into potential next steps in the process. However, the down-stream effects on staff efficiency from integrating such an algorithm into the clinical workflow have not yet been assessed.

In this study, the BCC detection and tumor mapping algorithm is used in the context of a MMS unit to simulate how automating an important step through AI incorporation may affect clinical and laboratory staff efficiency. Additionally, this paper highlights important concepts in AI implementation, including selection of processes that most benefit from automation and the logistical challenges that must be addressed prior to seamless integration.

CAPSULE SUMMARY

- This study addresses the impact and logistical requirements of implementation of an artificial intelligence algorithm into a real-world clinical workflow.
- Results indicate the potential for increased efficiency and productivity with use of artificial intelligence in Mohs micrographic surgery, particularly in settings where timely reconstruction is performed concurrently with tumor removal.

MATERIALS AND METHODS

The study included 104 consecutive MMS BCC cases performed over a 20-day period. To create a simulated MMS workflow that integrates the BCC detection algorithm, 3 measures were defined to evaluate algorithm integration: slide waiting time, staff waiting time, and histotechnician waiting time, depicted in Fig 1. Slide waiting time is defined as the time between when the histotechnician places the prepared slides next to the microscope and when the surgeon performs histologic analysis of the slides and subsequent tumor mapping. Staff waiting time is the time between when the histotechnician places the prepared slides next to the microscope and when the nursing staff member begins preparation for the next step, which is either an additional stage or defect repair. Histotechnician waiting time is the time between when the histotechnician places the prepared slides next to the microscope and when the surgeon delivers tissue

Abbreviations used:

AI: artificial intelligence
BCC: basal cell carcinoma
MMS: Mohs micrographic surgery

from the next stage in cases in which additional stages of tumor removal are required. On the days of surgery, slide, staff, and histotechnician waiting times were measured for each of the 104 cases. To generate simulated waiting times, the BCC detection algorithm was implemented retrospectively at the immediate time the slides were placed next to the microscope.

The BCC detection algorithm uses a combination of convolutional neural networks and graph neural networks to identify tumor. In recently published work, this algorithm has demonstrated an AUC of 0.97 in the identification of positive tumor margins, comparable to or exceeding findings from previous, similar studies.[7] The algorithm has the capability to provide (1) three-dimensional gross specimen reconstruction with neural radiance fields, (2) tissue completeness assessment with convolutional neural networks to ensure high-quality sections without holes or tears prior to analysis, (3) identification of tumor confounders such as hair follicles with R101-FPN neural network models, and (4) tumor mapping that depicts histological findings on the image of the surgical site to generate surgical recommendations. These results are available for histotechnicians, pathologists, and surgeons through a dynamic web application that provides an interactive and exportable pathology report. The BCC detection algorithm was trained and validated using whole slide images of specimens from patients undergoing MMS for BCC removal that were manually annotated for tumor, benign structures (eg, areas of inflammation and follicles), and layers of skin (eg, epidermis, dermis, etc.) by 4 pathologists. The algorithm is available on GitHub for public use.[7]

To simulate nonautonomous algorithm integration, slide scanning time, algorithm processing time, algorithm runtime, and output review time were measured (Supplementary Fig 2, available via Mendeley at https://doi.org/10.17632/4yv8zg4k2p.1). The sum of these 4 components represented the simulated slide waiting time. Sixty-five cases reflecting the normal distribution of BCC histologic subtypes in MMS clinical practice were selected to undergo slide scanning at 20X using an Aperio AT2 to generate whole slide images in the SVS format with 8 bit color channels. The scanning time per slide was recorded and summed per case to determine total slide scanning time. Whole slide images were assessed using the BCC detection and tumor mapping algorithm, and algorithm processing time and runtime were measured. All cases were deidentified for the comparison of algorithm-versus surgeon-generated tumor maps. Concordance between the tumor maps was established based on a retrospective review and subjective interpretation of visual agreement (yes/no). The proportion of times the 2 maps agreed was recorded as a measure of concordance and 95% confidence intervals were obtained using a normal approximation of the binomial probabilities. No clinical decisions were made using algorithm output.

To generate simulated staff and histotechnician waiting times for each case, the difference between the actual and simulated slide waiting times was calculated, which represented the time saved or lost with algorithm integration. This amount of time was then subtracted or added from the downstream staff and histotechnician waiting times. For analysis of repair approach on waiting times, complex repairs were defined as local flaps, full thickness skin grafts, or interpolation flaps, and simple repairs were defined as linear closures or healing by second intent. For analysis of number of tumor removal stages on waiting times, data from 5 days with the most number of tumor removal stages were compared with data from 5 days with the least number of tumor removal stages.

RESULTS

Removal of skin cancer with MMS relies on a clinical workflow that involves nursing staff, the surgeon, and histotechnicians in the on-site pathology laboratory. Staffing numbers and operating room space can affect the length of time required to remove the tumor with iterative real-time histologic margin analysis and subsequent repair of the resulting surgical defect. Over the 20-day study period, the numbers of nursing staff (3), histotechnicians (2), and operating rooms (5) were held constant. Additional variables hypothesized to affect staff and histotechnician waiting times included number of tumor removal stages per day and repair complexity.

Critical to the decision of whether to implement AI in an autonomous versus nonautonomous manner is the determination as to whether the algorithm performs at or above the level of the human expert. To test this, a cohort of 65 BCC cases reflecting the normal distribution of BCC histologic subtypes in MMS clinical practice were scanned and analyzed by the BCC detection algorithm (Table I). The algorithm-generated tumor maps were compared

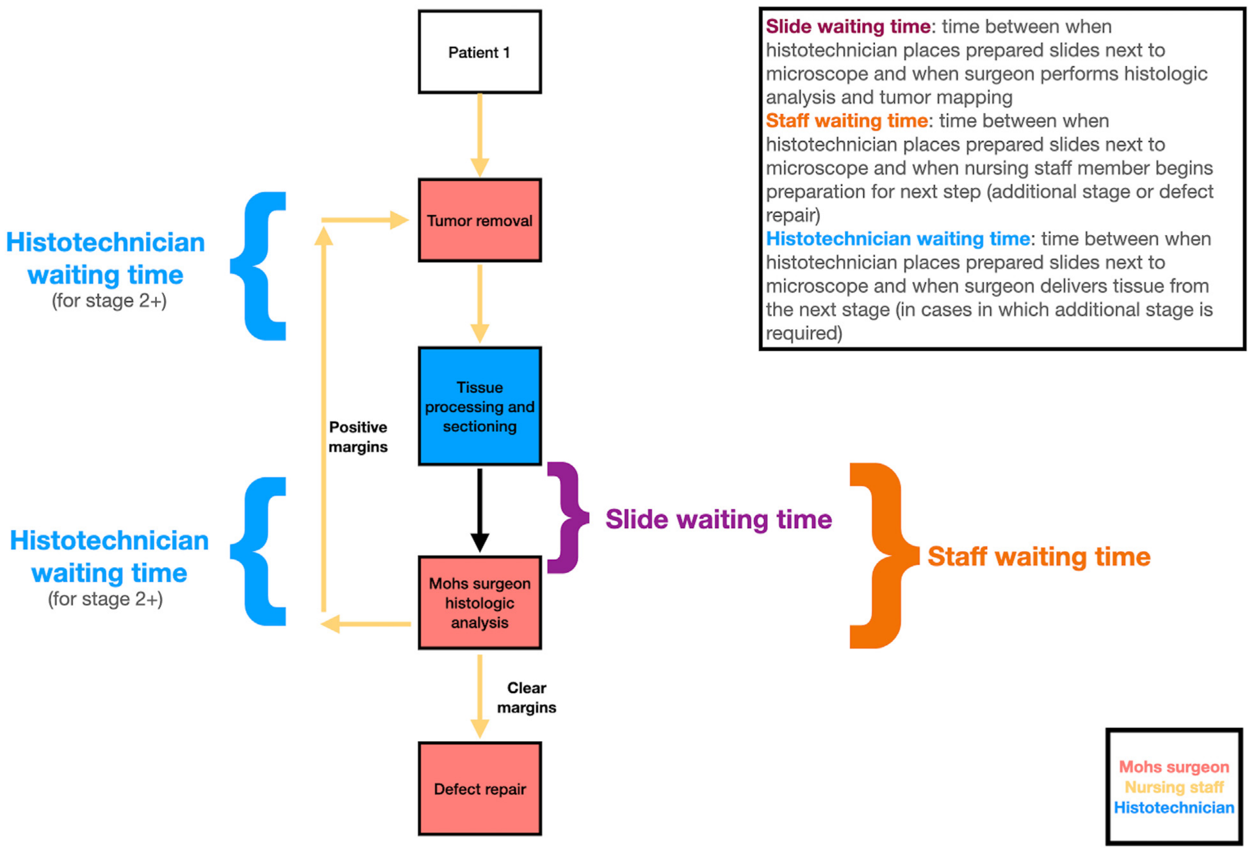

Fig 1. Definitions and depiction of slide, staff, and histotechnician waiting times.

to the hand-drawn tumor maps created by the Mohs surgeon at the time of surgery (Fig 2). The BCC detection algorithm identified the tumor and appropriate location of tumor in 94% (95% CI: 88% to 99%) of surgical cases. The 94% accuracy fell below the presumed human diagnostic accuracy of 98% to 99%, which was based on long term recurrence rates of 1% to 4%, as highlighted in expert treatment guidelines.[8] Therefore, the simulated workflow in this paper used the algorithm in a nonautonomous manner, integrating review of the algorithm output by the surgeon as the final step of the histologic analysis.

To determine how the BCC detection and tumor mapping algorithm affects the MMS workflow, a simulated nonautonomous workflow was created in which histologic assessment was performed by the BCC detection algorithm with subsequent verification of algorithm output by the Mohs surgeon prior to performing the next step of the process. The amount of time required to scan the selected 65 BCC cases and generate output from the algorithm was measured for each case (Supplementary Table I, available via Mendeley at https://doi.org/10.17632/4yv8zg4k2p.1). The average slide scanning time, algorithm computation time, and output review time per case were 10:25 minutes, 2:00 minutes, and 30 seconds, respectively. Scan time was proportional to the number of slides per case, the number of tissue sections per slide, and the size of the tissue sections. Scan times per case ranged from 4:00 to 43:00 minutes. Simulated waiting times were adjusted by adding or subtracting the slide

Table I. Selected basal cell carcinoma case histologic subtypes and algorithm accuracy

Number of cases	65
BCC histologic subtype	
Nodular	37
Superficial	14
Infiltrative	15
Micronodular	5
No tumor	16
Accuracy (CI)	0.94 (0.88-0.99)
Average slide scanning time per case (min)	10:25
Average algorithm processing and execution time per case (min)	2:00
Average output review time per case (s)	30

BCC, Basal cell carcinoma.

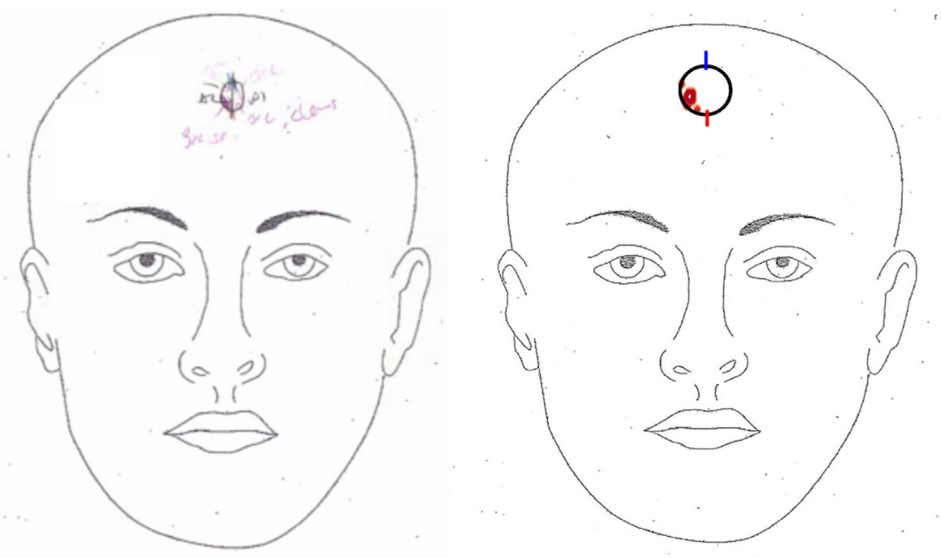

Fig 2. Surgeon-generated (*left*) versus algorithm-generated (*right*) tumor maps.

scanning/algorithm output time from the actual slide waiting time for each of the 104 BCC cases. Nonautonomous algorithm integration resulted in saving 35.69% of slide waiting time (37:28 minutes per day). As a result, 18.2% of staff waiting time and 18.6% of histotechnician time was saved (Table II). Further analysis revealed that 55/104 (52.9%) of cases had increased time savings when the AI algorithm was implemented. Comparison of only cases (55) that benefited from the implementation of the AI algorithm revealed savings of 55%, 28%, and 25% for slide, staff, and histotechnician waiting time respectively over the 20-day period (Table II).

Additional analysis was performed to determine characteristics of MMS cases and days in which the algorithm provided the most benefit to the clinical workflow. Repair complexity and the number of tumor removal stages performed per day were identified as variables that affect the MMS clinical workflow. Comparison of waiting times between the 5 days with the highest and lowest proportion of complex repairs (see Methods) and most and least tumor removal stages revealed more time saved on days with increased complexity and/or increased tumor removal stages (Table III).

DISCUSSION

Implementation of technology into clinical workflows offers 1 approach to increase the efficiency of health care delivery. Here we demonstrate that simulated use of a nonautonomous BCC detection and tumor mapping algorithm within the MMS workflow can reduce slide, staff, and histotechnician waiting times. By aiding the surgeon in performing the real-time histologic examination and tumor mapping, subsequent steps in the MMS process can be carried out in a more efficient manner. Notably, the nursing staff can act at an earlier timepoint to prepare both the operating room and patients for either additional tumor removal or defect reconstruction. The surgeon can move between operating rooms without the normal waiting time required for patient and staff transition. This results in more efficient tumor removal and subsequent defect repair. Histotechnicians receive tissue in the lab at an earlier time point than they would otherwise, decreasing the overall amount of histotechnician time required. Together, algorithm implementation in the simulated scenario increased the parallelization of an otherwise serial process.

Consideration of design of autonomous versus nonautonomous algorithms relies on a number of factors, most importantly task complexity and algorithm accuracy. In the setting of an algorithm accuracy of 94%, the BCC detection algorithm would need to be implemented, at least initially, in a nonautonomous manner. In this setting, this study has identified that not all cases benefit from implementation of the algorithm with number of tumor removal stages and repair complexity being important variables that affect time savings with algorithm implementation. Scoring systems exist in MMS to predict these factors and could be combined with a nonautonomous algorithm-driven system to generate increased operational efficiency while ensuring the gold standard outcome measure

Table II. Nonautonomous basal cell carcinoma algorithm implementation effects on slide, staff, and histotechnician waiting time per day for all cases (104) and those with net positive time savings (55)

	All cases	Only cases with net positive time savings
Mean slide waiting time saved per day (h:mm:ss)	0:37:28*	0:57:47*
Proportion of slide waiting time saved over period	35.69%*	55.03%*
Proportion of staff waiting time saved over period	18.24%*	28.13%*
Proportion of histotechnician waiting time saved over period	18.62%*	25.53%*

Paired t tests were used to assess whether the difference between actual and simulated waiting times were significant, *$P < .05$.

Table III. Waiting time analysis stratified by comparing 5 days with highest proportion of complex repairs to 5 days with lowest proportion of complex repairs and 5 days with most tumor removal stages to 5 days with least number of tumor removal stages

	Simple repair days	Complex repair days	Fewer stage days	More stage days
Slide	0:01:13	0:11:09	0:07:34	0:17:23
Staff	0:01:13	0:11:09	0:07:34	0:17:23
Histotechnician	−0:00:06	0:05:22	0:02:14	0:09:36

Time saved displayed in (h:mm:ss).

of highly accurate tumor removal with low risks for tumor recurrence.[9] While not simulated here, if the algorithm was to be implemented in an autonomous manner, this would likely provide further improvements in staffing time saved per case assuming there would not be a need for the surgeon to perform pathologic analysis and tumor mapping. Additional training of the algorithm to approach the accuracy of a human expert may allow for this in future studies.

This simulated workflow thus far has only been run retrospectively and at a single site. Real-time clinical implementation may affect outcomes or even more likely identify additional clinically relevant workflow variables that have not been accounted for. This study highlights the number of additional considerations that must be taken into account, as highlighted by the average slide scanning time of over 10 minutes resulting in only approximately 50% of the cases benefitting from implementation of the algorithm. Furthermore, implementation of complex AI algorithms requires a high level of computing power that is unlikely to currently be available to the average user. The cost of implementation and complexity of such a robust automated system is significant. AI integration in the MMS clinical workflow requires a slide scanner, elevated levels of computing power, and reliable access to high-speed internet. Not only are these resources expensive, but as with any technology-dependent system, inferior performance of any of these components may significantly impact real-life results. With the introduction of more efficient and automated slide scanners combined with ready access to powerful computing resources, AI technology will provide a more realistic solution to a broader number of end users.

Future studies should involve analyzing the broader implementation of the algorithm into external MMS units with varying numbers of Mohs micrographic surgeons, nursing staff, histotechnicians, operating rooms, cases per day, and cases with complex repairs. Additionally, schedules could be created containing cases with predicted increased number of tumor removal stages and high degrees of repair complexity. Based on the current results, days maximizing these variables may benefit most from implementation of the BCC detection algorithm. As additional tumor detection algorithms are developed, similar workflow analyses may be done to assess the generalizability of this study to other tumor types, including squamous cell carcinoma and melanoma. An additional potential implementation of the BCC algorithm is in a clinical setting where real-time analysis of tumor margins is not available but a patient would benefit from precise and real-time tumor removal. While beyond the scope of this study, this scenario would present a distinct set of logistical challenges and workflow barriers that would need to be solved prior to implementation. We would propose use of a simulated scenario as performed in this study as an initial approach to identify and develop solutions to workflow barriers.

Processes that are repetitive or highly iterative and where multiple or numerous individuals rely on decision making from a single individual are ideal for AI integration. This study in particular helps to provide proof of concept for potential uses of AI in MMS while also highlighting the current limitations to real-world implementation of the technology in this setting.

Regardless of the specific algorithm being implemented, perspectives and feedback from all members of the clinical care team that might be affected by AI integration should be taken into account. Design and testing validation of AI solutions by end users, including providers and medical staff, offers an opportunity to correctly identify important clinical variables that need to be controlled for and the most relevant bottlenecks in the health care delivery process. Early analysis of both positive and negative neighborhood effects, including cost and resource requirements, will help identify clinical settings most likely to benefit from early AI implementation.

Conflicts of interest

None disclosed.

REFERENCES

1. Center for Devices and Radiological Health, U.S. Food and Drug Administration. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. 2022:1.
2. Feng H, Berk-Krauss J, Feng PW, Stein JA. Comparison of dermatologist density between urban and rural counties in the United States. *JAMA Dermatol*. 2018;154(11):1265-1271.
3. Puri P, Comfere N, Drage LA, et al. Deep learning for dermatologists: part II. Current applications. *J Am Acad Dermatol*. 2022;87(6):1352-1360.
4. Chervoni-Knapp T. The staffing shortage pandemic. *J Radiol Nurs*. 2022;41(2):74-75.
5. Bureau of Labor Statistics. U.S. Department of Labor. Occupational Outlook Handbook. Medical Assistants; 2022:1.
6. Bureau of Labor Statistics, U.S. Department of Labor. Occupational Outlook Handbook. Registered Nurses; 2022:1.
7. Levy JJ, Davis MJ, Chacko RS, et al. Intraoperative margin assessment for basal cell carcinoma with deep learning and histologic tumor mapping to surgical site. *NPJ Precis Oncol*. 2024;8(1):2.
8. Kim JYS, Kozlow JH, Mittal B, Moyer J, Olencki T, Rodgers P, Work Group, Invited Reviewers. Guidelines of care for the management of basal cell carcinoma. *J Am Acad Dermatol*. 2018;78(3):540-559.
9. Rivera AE, Webb JM, Cleaver LJ. The Webb and Rivera (WAR) score: a preoperative Mohs surgery assessment tool. *Arch Dermatol*. 2012;148(2):206-210.