


High Interobserver Agreement in the Subjective Classification of 5-Aminolevulinic Acid Fluorescence Levels in Newly Diagnosed Glioblastomas

Mario Mischkulnig, MD ^{1,2} Barbara Kiesel, MD,^{1,2} Martin Borkovec, BSc, BA,^{1,3} Lisa I. Wadiura, MD,^{1,2} Dimitri Benner, MD,¹ Arthur Hosmann, MD,^{1,2} Shawn Hervey-Jumper, MD,⁴ Engelbert Knosp, MD,^{1,2} Karl Roessler, MD,^{1,2} Mitchel S. Berger, MD,⁴ and Georg Widhalm, MD, PhD^{1,2*}

¹Department of Neurosurgery, Medical University Vienna, Waehringer Guertel 18-20, Vienna 1090, Austria

²Central Nervous System Tumours Unit, Comprehensive Cancer Center, Medical University Vienna, Waehringer Guertel 18-20, Vienna 1090, Austria

³Department of Statistics, Ludwig-Maximilians-University, Ludwigstraße 33, Munich 80539, Germany

⁴Department of Neurological Surgery, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, California 94143

Background and Objectives: Fluorescence-guided resection of glioblastomas (GBM) using 5-aminolevulinic acid (5-ALA) improves intraoperative tumor visualization and is thus widely used nowadays. During resection, different fluorescence levels can usually be distinguished within the same tumor. Recently, we demonstrated that strong, vague, and no fluorescence correspond to distinct histopathological characteristics in newly diagnosed GBM. However, the qualitative fluorescence classification by the neurosurgeon is subjective and currently no comprehensive data on interobserver variability is available. The aim of this study was thus to investigate the interobserver variability in the classification of 5-ALA fluorescence levels in newly diagnosed GBM.

Study Design/Materials and Methods: A questionnaire investigating the interobserver variability in 5-ALA fluorescence quantification was performed at a nation-wide neurosurgical oncology meeting. The participants involved in the neurosurgical/neurooncological field were asked to categorize 30 cases of 5-ALA fluorescence images derived from GBM resection on a lecture hall screen according to the widely used three-tier fluorescence classification scheme (negative, vague, or strong fluorescence). Additionally, participants were asked for information on their medical background such as specialty, level of training, and experience with 5-ALA fluorescence-guided procedures. Interobserver agreement was defined as the calculated mean κ values for each observer.

Results: A total of 36 questionnaires were included in the final analysis. The mean average κ value in fluorescence classification within the entire cohort was 0.71 ± 0.12 and 29 (81%) participants had a substantial or almost perfect interobserver agreement (κ values 0.6–1.0). Interobserver agreement was significantly higher in neurosurgeons (mean κ : 0.83) as compared with non-neurosurgeons involved in the neurooncological field (mean κ : 0.52; $P < 0.001$). Furthermore, interobserver agreement was significantly higher in participants who had experience

with at least 25 5-ALA fluorescence-guided surgeries (mean κ : 0.87) compared with less experienced colleagues (mean κ : 0.82; $P = 0.039$).

Conclusion: Our study found a high interobserver agreement in the qualitative classification of different 5-ALA fluorescence levels in newly diagnosed GBM. Interobserver agreement increases significantly in more experienced participants and therefore a high level of experience is crucial for reliable intraoperative fluorescence classification. *Lasers Surg. Med.* © 2020 The Authors. *Lasers in Surgery and Medicine* published by Wiley Periodicals, Inc.

Key words: 5-ALA; fluorescence levels; newly diagnosed glioblastoma; interobserver variability; high interobserver agreement

INTRODUCTION

Surgical tumor resection represents the initial treatment of choice in newly diagnosed glioblastomas (GBM) [1]. In such tumors, extent of resection strongly correlates with patient prognosis and thus maximum safe resection of GBM is the neurosurgical goal [2,3]. However, incomplete resection of the contrast-enhancing GBM tissue is commonly found in up to 80% of cases due to their highly infiltrative growth and

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

*Correspondence to: Georg Widhalm, MD, PhD, Department of Neurosurgery, Medical University Vienna, Waehringer Guertel 18-20, Vienna 1090, Austria. E-mail: georg.widhalm@meduniwien.ac.at

Accepted 23 February 2020

Published online 8 March 2020 in Wiley Online Library (wileyonlinelibrary.com).

DOI 10.1002/lsm.23228

poor intraoperative visualization of the tumor margin [4-6]. In the last two decades, fluorescence-guided surgery using 5-aminolevulinic acid (5-ALA) has been established as powerful technique for improved intraoperative detection of GBM tissue [7]. In particular, Stummer et al. demonstrated a significantly higher rate of complete resections of malignant gliomas and a prolonged progression-free survival in patients with 5-ALA fluorescence-guided surgery compared with conventional white-light procedures [6,8].

Generally, different 5-ALA fluorescence levels can be distinguished during resection of GBM by the neurosurgeon [6,9,10]. While most prior studies applied dichotomous (negative/positive) fluorescence categorization, the second most commonly applied quantification system distinguishing a larger number of fluorescence levels is the three-tier classification in strong, vague, and no fluorescence (see Fig. 1) [6,11]. In a recent study, we demonstrated that these three fluorescence levels are characterized by distinct differences in specific histopathological parameters in newly diagnosed GBM [12]. These different fluorescence levels are thus capable to support the neurosurgeon visualizing specific intratumoral regions (compact vs. infiltrative tumor) to achieve maximal safe resection of GBM [12]. Nevertheless, this currently applied qualitative fluorescence classification system is subjective and in consequence observer-dependent [13-16]. Although this three-tier fluorescence classification system is widely used, no systematical analysis of its interobserver variability is currently available.

The aim of this study is, therefore, to systematically analyze the interobserver variability in the three-tier classification of different 5-ALA fluorescence levels in newly diagnosed GBM by means of a questionnaire-based analysis.

MATERIAL AND METHODS

We performed a questionnaire-based analysis examining the classification of intraoperative images with varying amounts of 5-ALA fluorescence in newly diagnosed GBM. This study was conducted in the course of the 25th Annual Meeting of the Section for Neurosurgical Oncology of the Austrian Society of Neurosurgery in May 2018 and included participants with variable experience

in 5-ALA fluorescence procedures. The present study was approved by the ethics committee of the Medical University of Vienna (EK 419/2008, Amendment).

Acquisition of Fluorescence Images

Video archives were screened for high-quality intraoperative digital images of 5-ALA fluorescence-guided resections in newly diagnosed GBM. The selected images of varying amounts of fluorescence were obtained during different stages of fluorescence-guided resection of the contrast-enhancing compact tumor, infiltration zone, and the resection cavity. Images were only included if postoperative histopathological assessment confirmed the diagnosis of GBM. All patients received a standard dose of 5-ALA (20 mg/kg bodyweight) approximately 3 hours prior to induction of anesthesia. For intraoperative fluorescence visualization, specifically equipped neurosurgical microscopes (Pentero or NC4; Carl Zeiss Surgical GmbH, Oberkochen, Germany) were applied. To guarantee comparable imaging quality, both microscopes used PpIX fluorescence detection hardware with identical wavelength settings and were subject to factory-adjustment of fluorescence visualization.

Questionnaire Design

For the assessment of the variability in the subjective classification of different amounts of 5-ALA fluorescence, we designed a specific questionnaire. The first section of this questionnaire asked for general information about the participants including the home medical center, use of 5-ALA at their center (yes or no), annual case number of 5-ALA surgeries at their center (<20 cases, 20–50 cases, >50–100 cases, >100 cases), the specialty (neurosurgery or other), level of training (resident or attending), personal clinical focus on neurooncology (yes or no), and the approximate total number of 5-ALA surgeries performed by the participants as primary surgeon and assistant, respectively. In the second part of the questionnaire, participants were requested to rate a total of 30 intraoperative images of varying 5-ALA fluorescence according to the



Fig. 1. Illustration of typical qualitative 5-aminolevulinic acid fluorescence levels observed during fluorescence-guided surgery of glioblastomas. (A) Intraoperative image showing a specimen that was classified to show no fluorescence according to all 36 (100%) observers. (B) Image of a specimen categorized as vague fluorescence by 33 of 36 (92%) observers. (C) Photograph of a specimen classified as strong fluorescence by all 36 (100%) observers.

commonly used three-tier classification system (negative, vague, or strong fluorescence; see Fig. 1).

Data Collection

This questionnaire was distributed to the attendees of the Annual Meeting of the Section for Neurosurgical Oncology. Initially, the principles of the three-tiered fluorescence classification system were explained to all participants in a short presentation by the first author (M.M.). The audience was then asked for their anonymous and voluntary participation in our study. Attendees that chose to participate were asked to fill out the questionnaire's general information part first. Thereafter, the auditorium was darkened, and the 30 intraoperative images were presented to the participants in a random order on the lecture hall screen. Each image was presented for 10 seconds and during this time period the participants classified and documented the fluorescence status of each image as negative, vague, or strong fluorescence.

Data Analysis

Statistical analyses were performed using "R" (Version 3.5.3; R Foundation for Statistical Computing, Vienna, Austria). Figures were created using ggplot 2 software (Version 3.2.0; Springer-Verlag; New York, NY). General interobserver variability was assessed calculating Cohen's κ values for the agreement of each pair of observers. Each observer was then assigned the mean κ value of his assessment measured against all other observers and further analyses of the agreement in the complete cohort were based on these values. To detect possible incorrectly completed questionnaires, median absolute deviation (MAD) for the entire cohort as well as the analyzed subgroups was calculated and outliers (distance from median >3 MAD) were excluded. To further evaluate agreement in subgroups according to home medical center, use of 5-ALA and annual number of 5-ALA procedures at the center, specialty, level of training, personal clinical focus, and experience in 5-ALA guided surgery (<25 vs. ≥ 25 cases) as surgeon and assistant we calculated the mean Cohen's κ within the respective subgroups. The distribution of mean κ values between the subgroups was subsequently analyzed for statistically significant differences using the non-parametric Mann-Whitney U test for dichotomous variables and the non-parametric Kruskal-Wallis test for number of annual 5-ALA procedures, where four groups were analyzed. κ values were interpreted according to earlier studies on interobserver agreement in diagnostic tests [17]. The level of significance for inferential statistical analyses was set at $P = 0.05$.

RESULTS

A total of 40 questionnaires were collected after the presentation of 30 intraoperative images of varying amounts of 5-ALA fluorescence in newly diagnosed GBM derived from altogether 22 patients/surgeries. Altogether, four of these questionnaires had to be excluded from final analysis as a result of complete absence of general information on the participant ($n = 2$) and assumed

incorrect completion of the form ($n = 2$) due to distance from median >3 MAD and individual implausible ratings (e.g., rating of no fluorescence in an image rated strongly fluorescent by all other observers). Therefore, 36 questionnaires remained in the final analysis of this study.

General Information on Participants

Of the 36 participants, 26 (72%) were affiliated with academic medical centers and 10 (28%) with non-academic medical centers. A total of 24 (67%) questionnaires were obtained from neurosurgeons, while the remaining 12 (33%) questionnaires were provided by non-neurosurgical participants. According to the level of training in the group of neurosurgeons ($n = 24$), 10 (42%) were attendings and 11 (46%) were in residency. The remaining 3 (12%) neurosurgeons provided no information in that regard. Of 22 neurosurgeons with available data on the personal clinical focus, 17 (77%) participants stated a specialization on neurooncology. A total of 33 (92%) raters provided information on the use of 5-ALA at their center and according to these data 5-ALA was implemented in the majority of cases (30 of 33 participants; 91%). With regard to the annual number of 5-ALA procedures at their center, 4 (15%) participants estimated less than 20 cases, 7 (26%) indicated 20–50 cases, 4 (15%) stated 50–100 cases, and 12 (44%) reported >100 cases. According to the total number of 5-ALA procedures performed by each participating neurosurgeon with available data, 13 (65%) indicated <25 procedures and 7 (35%) ≥ 25 operations as primary surgeon. As assistant, 11 (50%) neurosurgeons stated <25 5-ALA operations and 11 (50%) declared ≥ 25 operations. Further details are provided in Table 1.

Overall Interobserver Variability

According to the classification of images with varying amounts of 5-ALA fluorescence, the mean average κ value within the entire cohort of 36 participants was 0.71 ± 0.12 with values ranging from 0.38 to 0.81. In detail, 29 (81%) of 36 participants had a substantial interobserver agreement (κ value 0.6–0.8; $n = 25$; 69%) or almost perfect interobserver agreement (κ value 0.8–1.0; $n = 4$; 11%). In contrast, the remaining 7 (19%) participants had a moderate interobserver agreement (κ value 0.4–0.6; $n = 5$; 14%) or slight interobserver agreement (κ value 0.2–0.4; $n = 2$, 5%). We never observed participants with a poor interobserver agreement (κ value <0.2) in our study. For details see Figure 2.

Subgroup Analysis by Specifics of the Home Medical Center

We did not find a statistically significant difference in the interobserver variability in the fluorescence classification between participants from academic (mean κ : 0.69 ± 0.14 ; range: 0.37–0.80) and non-academic centers (mean κ : 0.76 ± 0.06 ; range: 0.62–0.83; $p = 0.069$). Furthermore, the interobserver agreement was significantly higher in participants from centers using 5-ALA

TABLE 1. Rater Characteristics

	<i>n</i>	%
Number of raters	36	100
Medical center		
Academic	26	72
Non-academic	10	28
Use of 5-ALA at center		
Yes	30	91
No	3	9
Specialty		
Neurosurgeons	24	67
Non-neurosurgeons	12	33
Level of training ^a		
Resident	11	52
Consultant	10	48
Clinical focus on neurooncology ^a		
Yes	17	77
No	5	23

5-ALA, 5-aminolevulinic acid.

^aWithin neurosurgeons.

(mean κ : 0.78 ± 0.09 ; range: 0.52–0.86) compared with colleagues from centers not using this technique (mean κ : 0.62 ± 0.04 ; range: 0.57–0.65; $p = 0.022$). Moreover, the interobserver agreement between participants from centers with less than 20 annual 5-ALA operations was lower (mean κ : 0.58 ± 0.09 range: 0.48–0.68) compared with centers with 20–50 (mean κ : 0.84 ± 0.03 , range: 0.80–0.87), >50–100 (mean κ : 0.84 ± 0.05 , range: 0.79–0.89) and >100 (mean κ : 0.86 ± 0.05 , range: 0.77–0.90) annual 5-ALA surgeries ($p = 0.007$). Details are given in Figure 2.

Subgroup Analysis by Specialty, Level of Training, and Personal Clinical Focus

According to the specialty, the interobserver agreement in fluorescence classification was significantly higher in the subgroup of neurosurgeons (mean κ : 0.83 ± 0.05 ; range: 0.74–0.88) compared with non-neurosurgeons (mean κ : 0.52 ± 0.12 ; range: 0.31–0.66; $P < 0.001$). Further, the mean κ values of neurosurgical attendings (mean κ : 0.86 ± 0.06 ; range: 0.73–0.90) were slightly higher compared with neurosurgical residents (mean κ : 0.83 ± 0.04 ; range: 0.76–0.87). However, this difference did not reach statistical significance ($P = 0.112$). With regard to the personal clinical focus, the interobserver agreement was significantly higher in neurosurgeons specialized in neurooncology (mean κ : 0.85 ± 0.05 ; range: 0.74–0.90) compared with neurosurgeons with a different clinical focus (mean κ : 0.76 ± 0.04 ; range: 0.72–0.81; $P = 0.006$). See also Figure 3.

Subgroup Analysis by Practical Experience With 5-ALA Procedures

With regard to the total number of 5-ALA fluorescence-guided surgeries performed as primary surgeon, the interobserver agreement was significantly higher among participants with at least 25 procedures (mean κ : 0.87 ± 0.04 ;

range: 0.80–0.89) compared with less experienced neurosurgeons (mean κ : 0.82 ± 0.05 ; range: 0.73–0.87; $P = 0.039$). Finally, a significantly higher interobserver agreement was also present in neurosurgeons that assisted in at least 25 5-ALA operations (mean κ : 0.84 ± 0.05 ; range: 0.74–0.89) compared with neurosurgeons that assisted in less than 25 5-ALA procedures (mean κ : 0.80 ± 0.04 ; range: 0.74–0.84; $P = 0.048$). Details are given in Figure 3.

DISCUSSION

Nowadays, 5-ALA fluorescence-guided surgery is widely applied in the surgical treatment of GBM. During these procedures, distinct fluorescence levels can usually be observed within the same GBM. Frequently, these fluorescence levels are qualitatively classified by the treating neurosurgeon according to a three-tier scheme in strong, vague, and no fluorescence. In a recent study, we found that strong fluorescence was consistent with compact tumor tissue in the majority of cases, whereas vague fluorescence corresponded mainly to infiltrative tumor tissue [12]. Furthermore, we demonstrated a significant correlation of the different fluorescence levels with histopathological criteria of malignancy, proliferation, and microvessel density [12]. Consequently, these different fluorescence levels are capable to support the neurosurgeon visualizing specific intratumoral regions such as compact contrast-enhancing tumor or infiltrative non-enhancing tumor in order to achieve maximal safe resection of GBM. However, no systematical analysis of the interobserver variability in the classification of 5-ALA fluorescence in newly diagnosed GBM between different observers is currently available in the literature.

Preliminary Data

In our previous study on the histopathological analysis of different 5-ALA fluorescence levels, we also investigated the interobserver variability between different neurosurgeons [12]. In this sense, participating neurosurgeons separately classified the same photographs of varying degrees of fluorescence derived from 5-ALA fluorescence-guided surgery of newly diagnosed GBM [12]. According to our earlier data, we found an almost perfect interobserver agreement of all observers [12]. However, this study was limited by the small number of participating neurosurgeons ($n = 6$) [12]. Additionally, the cohort of observers was highly homogenous, comprising only neurosurgeons from the same center and with extensive experience in 5-ALA fluorescence-guided procedures [12].

Present Study

Based on these preliminary data [12], we thus designed a systematical analysis of interobserver variability of different 5-ALA fluorescence levels in a cohort of observers involved in the neurosurgical/neurooncological field. In the present study, we thus performed a questionnaire-based analysis investigating the classification of intraoperative images with varying amounts of fluorescence in newly diagnosed GBM at a specialized nation-wide meeting for neurosurgical oncology.

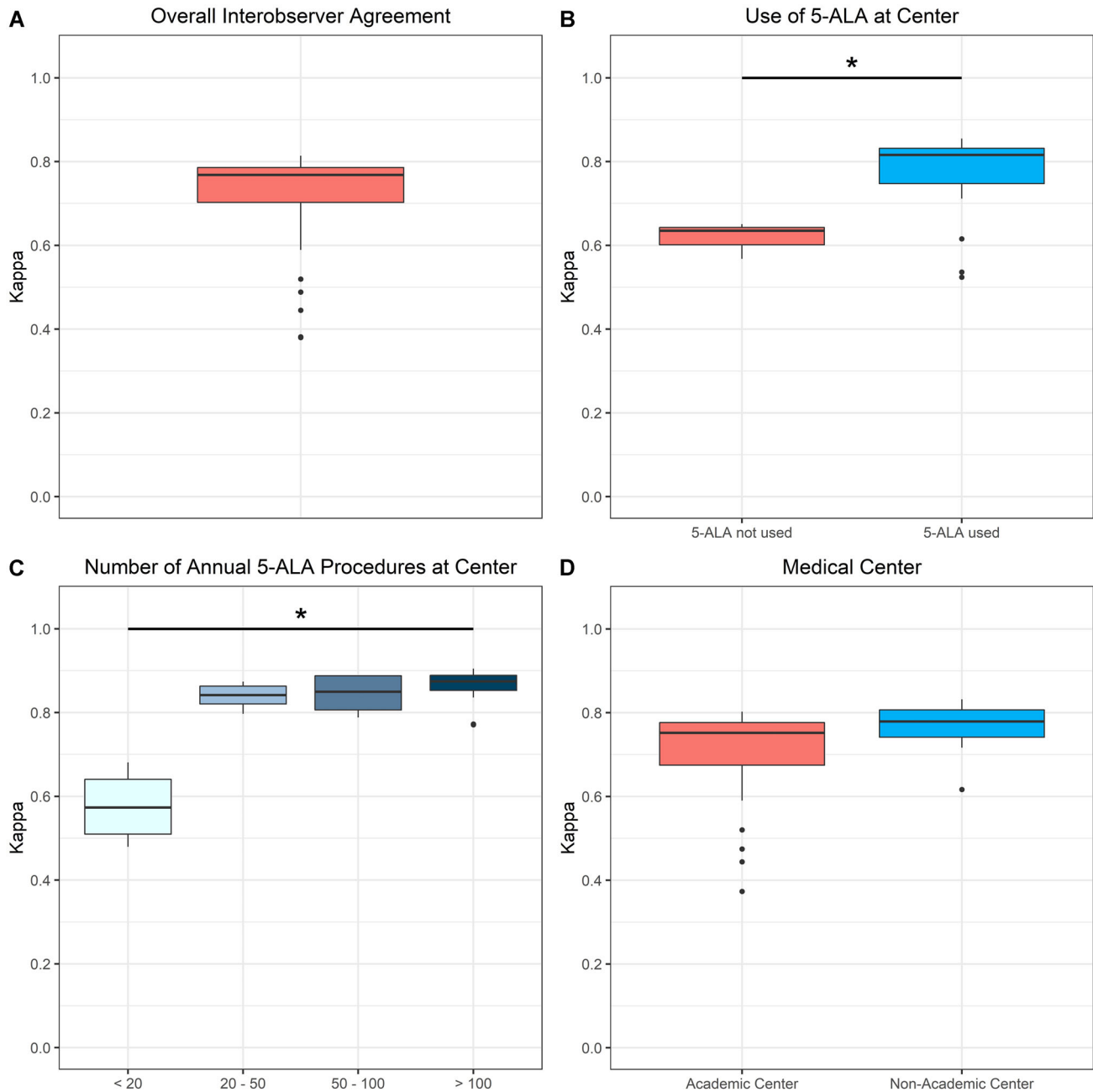


Fig. 2. Boxplot diagrams of mean interobserver agreement in the fluorescence classification in the whole study cohort as well as in distinct subgroups. (A) A substantial mean overall interobserver agreement was observed in the entire study cohort. (B) Interobserver agreement was significantly higher ($P = 0.022$) in observers from centers using 5-aminolevulinic acid (5-ALA) fluorescence-guided surgery as compared with departments without this technique. (C) A significantly higher interobserver agreement ($P = 0.007$) was found in observers from centers performing a higher number as compared with lower number of annual 5-ALA procedures. (D) No significant difference ($P = 0.06$) in interobserver agreement was demonstrated between observers from academic and non-academic centers.

Altogether, the data of 36 observers from different specialties, levels of training and experience with 5-ALA fluorescence-guided surgery were analyzed in this study—representing the most comprehensive analysis of interobserver variability in 5-ALA fluorescence to date.

Overall interobserver variability. According to our data, a substantial or almost perfect interobserver agreement in fluorescence classification was found in the majority (81%) of the 36 observers. It is of note that we never found a poor interobserver agreement in the current

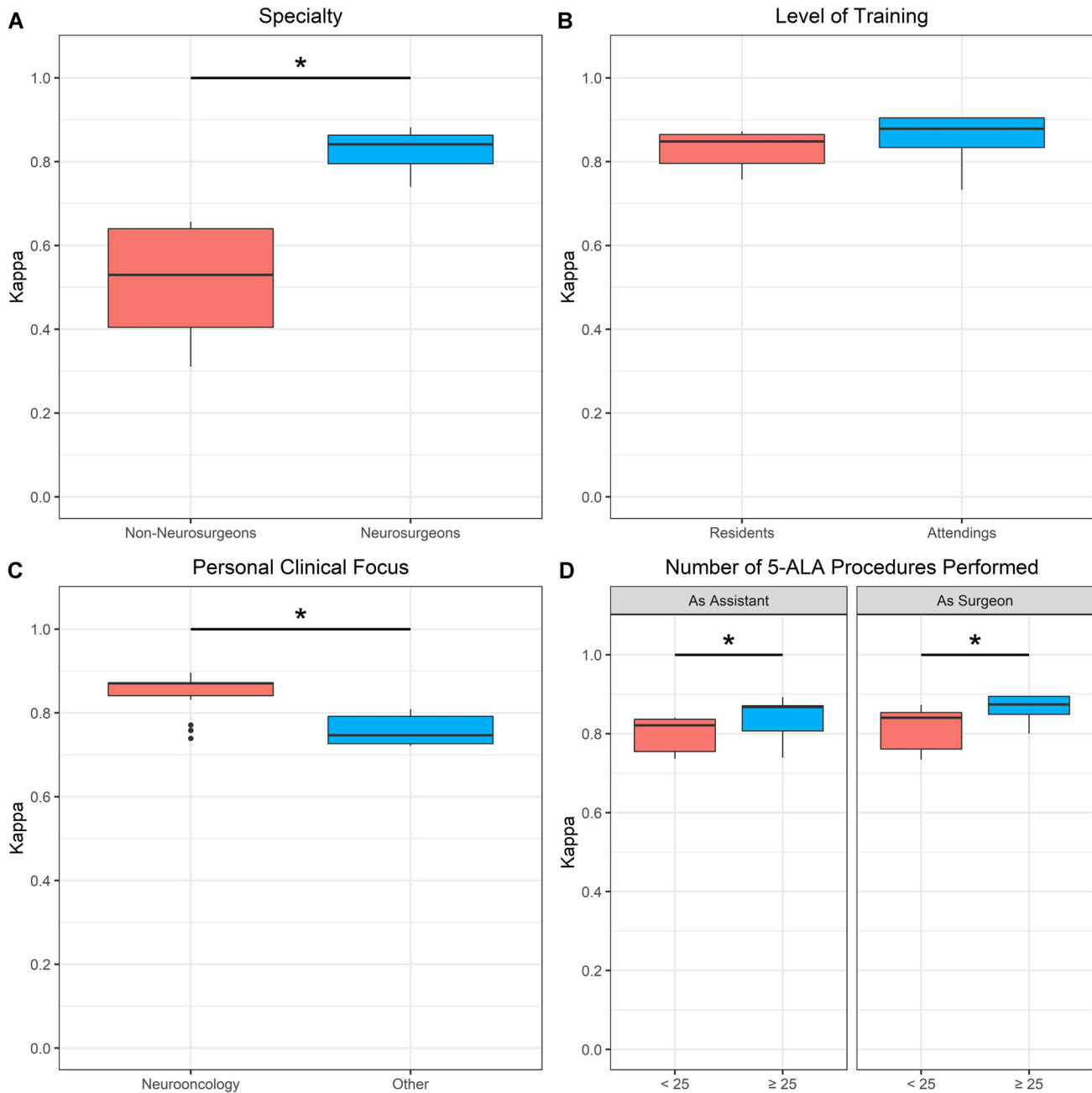


Fig. 3. Boxplot diagrams of interobserver agreement in fluorescence classification in the subgroup of neurosurgeons. (A) Interobserver agreement was significantly higher ($P < 0.001$) in neurosurgeons as compared with non-neurosurgeons. (B) No significant difference ($P = 0.112$) in the interobserver agreement was observed between neurosurgical residents and attendings. (C) Interobserver agreement was significantly higher ($P = 0.006$) in neurosurgeons with personal clinical focus on neurooncology. (D) Interobserver agreement was significantly higher in neurosurgeons who had performed at least 25 5-ALA fluorescence-guided procedures as assistant ($P = 0.048$) and primary surgeon ($P = 0.039$) compared with less experienced observers.

study. In our initial study, we found an almost perfect interobserver agreement for all six participating observers with a relatively small range of κ values between 0.83 and 1.00 [12]. Interobserver agreement in our present study was lower with κ values between 0.38 to 0.81 as compared with

our initial study. These findings are not surprising considering that our current cohort of observers was more heterogeneous with regard to inclusion of different specialties and degree of experience with 5-ALA fluorescence-guided procedures as compared with our

primary study [12]. However, we found a high overall interobserver agreement with a mean average κ value of 0.71 ± 0.12 in our entire study cohort. Taken together, our current data demonstrate a high objectivity of the observer-based categorization of 5-ALA fluorescence levels in a cohort of 36 observers, also including observers with less experience in 5-ALA fluorescence-guided surgery.

Subgroups of neurosurgeons and non-neurosurgeons. The subgroup of neurosurgeons included attendings and residents who were mostly familiar with the clinical application of 5-ALA fluorescence, while the subgroup of non-neurosurgeons consisted of colleagues from other medical specialties and medical students with an interest in neurosurgical oncology. The observed significantly higher agreement within neurosurgeons as compared with non-neurosurgeons suggests that there is a learning curve in observer-based fluorescence categorization and a reliable fluorescence rating can, to some extent, be trained.

Subgroups by formal level of training and experience with fluorescence surgery. In contrast to the difference between neurosurgeons and non-neurosurgeons, there was only a slight, but non-significant difference in the mean κ value between neurosurgical residents and neurosurgical attendings. It is of note that the interobserver variability in fluorescence rating among neurosurgical attendings analyzed in our current study was very similar to the one observed in the raters of our previous examination at our neurosurgical department at the Medical University of Vienna [12]. Furthermore, we found a significantly higher agreement between observers who had performed ≥ 25 5-ALA procedures as primary surgeon as well as assistant in our study. These findings suggest that actual practical experience with 5-ALA fluorescence-guided surgery is a more accurate predictor for reliable fluorescence rating than the formal level of training.

Innovations in objective fluorescence quantification. While the routinely applied subjective classification scheme of 5-ALA-induced fluorescence constitutes a reliable method to detect tumor tissue in newly diagnosed GBM, objective quantitative assessment of fluorescence by spectroscopic probes measuring the precise PpIX accumulation are increasingly examined in first clinical studies to further enhance the diagnostic performance of 5-ALA [12,18-20]. Aside from the completely objective fluorescence quantification, spectroscopic probes also have the potential to detect very low concentrations of PpIX that do not cause any visible fluorescence [18-20]. Therefore, these spectroscopic probes are capable to increase the relatively low sensitivity of visible 5-ALA-induced fluorescence [18-20].

Study limitations. The following limitations of the present study have to be mentioned: (I) first of all, no definitive statement regarding the interobserver variability in recurrent GBM can be made in this study since only newly diagnosed GBM were included. However, the scope of this study was to analyze the interobserver

variability in newly diagnosed GBM with lack of treatment associated changes. (II) Furthermore, no correlation of the subjectively categorized fluorescence levels with specific histopathological parameters was performed in this study. However, a comprehensive examination of the validity of qualitatively analyzed 5-ALA fluorescence levels as compared with histopathology (hematoxylin-eosin staining, CD34, proliferation index) is already provided in a recent study performed by our group [12]. (III) In the current study, the interobserver variability was only analyzed in the frequently applied three-tiered classification system, but not in other less commonly used classification systems with four or even five fluorescence levels [11,14,16,20-28]. Therefore, our study provides an insight only in the interobserver variability of the commonly applied three-tiered fluorescence classification, whereas the calculated values do not apply to other subjective classification systems and discrimination between a larger number of fluorescence levels is likely to be associated with an increased interobserver variability. (IV) Finally, the fluorescence rating of the majority of observers in our study does not necessarily imply correctness of classification. In our opinion, the gold standard to objectively measure fluorescence serving as a reference would be the spectroscopic analysis of fluorescence. Indeed, spectroscopic probes that are capable to objectively measure the quantitative 5-ALA-induced protoporphyrin IX accumulation (PpIX concentration) during tumor resection were recently introduced to the neurosurgical field [19,20]. Therefore, the subjective observer-based fluorescence classification should thus be compared with objective 5-ALA induced protoporphyrin IX accumulation (PpIX concentration) measured by spectroscopic probes serving as a reference in a future study.

CONCLUSIONS

In the present study, we systematically investigated for the first time the interobserver variability of different 5-ALA fluorescence levels in newly diagnosed GBM in cohort of 36 observers involved in the neurosurgical/neurooncological field. According to our data, we found a substantial/almost perfect interobserver agreement in fluorescence classification in the majority of observers in our study. Within the subgroup of neurosurgeons, a further decrease in interobserver variability was observed in raters who had greater practical experience (≥ 25 procedures) with 5-ALA fluorescence-guided procedures, which was a more accurate predictor for reliable fluorescence rating than the formal level of training. To further improve intraoperative 5-ALA fluorescence quantification, spectroscopic probes represent a promising technique that should be further investigated in future clinical studies.

ACKNOWLEDGMENT

We give thanks to Walter Schuetzenauer, who by carefully archiving the intraoperative videos recorded at our department made the selection of fluorescence images possible.

REFERENCES

- Alifireris C, Trafalis DT. Glioblastoma multiforme: Pathogenesis and treatment. *Pharmacol Ther* 2015;152:63–82. <https://doi.org/10.1016/j.pharmthera.2015.05.005>
- Kreth F-W, Thon N, Simon M, et al. Gross total but not incomplete resection of glioblastoma prolongs survival in the era of radiochemotherapy. *Ann Oncol Off J Eur Soc Med Oncol* 2013;24(12):3117–3123. <https://doi.org/10.1093/annonc/mdt388>
- Almeida JP, Chaichana KL, Rincon-Torroella J, Quinones-Hinojosa A. The value of extent of resection of glioblastomas: Clinical evidence and current approach. *Curr Neurol Neurosci Rep* 2015;15(2):517. <https://doi.org/10.1007/s11910-014-0517-x>
- Kowalczyk A, Macdonald RL, Amidei C, et al. Quantitative imaging study of extent of surgical resection and prognosis of malignant astrocytomas. *Neurosurgery* 1997;41(5):1028–1036. Discussion 1036–1038.
- Albert FK, Forsting M, Sartor K, Adams HP, Kunze S. Early postoperative magnetic resonance imaging after resection of malignant glioma: objective evaluation of residual tumor and its influence on regrowth and prognosis. *Neurosurgery* 1994;34(1):45–60. Discussion 60–61.
- Stummer W, Novotny A, Stepp H, Goetz C, Bise K, Reulen HJ. Fluorescence-guided resection of glioblastoma multiforme by using 5-aminolevulinic acid-induced porphyrins: a prospective study in 52 consecutive patients. *J Neurosurg* 2000;93(6):1003–1013. <https://doi.org/10.3171/jns.2000.93.6.1003>
- Osorio JA, Aghi MK. Optimizing glioblastoma resection: Intraoperative mapping and beyond. *CNS Oncol* 2014;3(5):359–366. <https://doi.org/10.2217/cns.14.36>
- Stummer W, Pichlmeier U, Meinel T, Wiestler OD, Zanella F, Reulen HJ. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: A randomised controlled multicentre phase III trial. *Lancet Oncol* 2006;7(5):392–401. [https://doi.org/10.1016/S1470-2045\(06\)70665-9](https://doi.org/10.1016/S1470-2045(06)70665-9)
- Widhalm G, Minchev G, Woehrer A, et al. Strong 5-aminolevulinic acid-induced fluorescence is a novel intraoperative marker for representative tissue samples in stereotactic brain tumor biopsies. *Neurosurg Rev* 2012;35(3):381–391. Discussion 391. <https://doi.org/10.1007/s10143-012-0374-5>
- Chohan MO, Berger MS. 5-Aminolevulinic acid fluorescence guided surgery for recurrent high-grade gliomas. *J Neurooncol* 2018;141:517–522. <https://doi.org/10.1007/s11060-018-2956-8>
- Kamp MA, Krause Molle Z, Munoz-Bendix C, et al. Various shades of red—A systematic analysis of qualitative estimation of ALA-derived fluorescence in neurosurgery. *Neurosurg Rev* 2018;41(1):3–18. <https://doi.org/10.1007/s10143-016-0745-4>
- Kiesel B, Mischkulnig M, Woehrer A, et al. Systematic histopathological analysis of different 5-aminolevulinic acid-induced fluorescence levels in newly diagnosed glioblastomas. *J Neurosurg*. 2018;129(2):341–353. <https://doi.org/10.3171/2017.4.JNS162991>
- Teixidor P, Arráez MÁ, Villalba G, et al. Safety and efficacy of 5-aminolevulinic acid for high grade glioma in usual clinical practice: A prospective cohort study. *PLoS One* 2016;11(2):e0149244. <https://doi.org/10.1371/journal.pone.0149244>
- Valdés PA, Leblond F, Kim A, et al. Quantitative fluorescence in intracranial tumor: implications for ALA-induced PpIX as an intraoperative biomarker. *J Neurosurg*. 2011;115(1):11–17. <https://doi.org/10.3171/2011.2.JNS101451>
- Haj-Hosseini N, Richter J, Andersson-Engels S, Wårdell K. Optical touch pointer for fluorescence guided glioblastoma resection using 5-aminolevulinic acid. *Lasers Surg Med* 2010;42(1):9–14. <https://doi.org/10.1002/lsm.20868>
- Piccirillo SGM, Dietz S, Madhu B, et al. Fluorescence-guided surgical sampling of glioblastoma identifies phenotypically distinct tumour-initiating cell populations in the tumour mass and margin. *Br J Cancer* 2012;107(3):462–468. <https://doi.org/10.1038/bjc.2012.271>
- Preusser M, Charles Janzer R, Felsberg J, et al. Anti-O6-methylguanine-methyltransferase (MGMT) immunohistochemistry in glioblastoma multiforme: observer variability and lack of association with patient survival impede its use as clinical biomarker. *Brain Pathol Zurich Switz* 2008;18(4):520–532. <https://doi.org/10.1111/j.1750-3639.2008.00153.x>
- Cornelius JF, Placke JM, Knipps J, Fischer I, Kamp M, Steiger HJ. Minispectrometer with handheld probe for 5-ALA based fluorescence-guided surgery of brain tumors: Preliminary study for clinical applications. *Photodiagnosis Photodyn Ther* 2017;17:147–153. <https://doi.org/10.1016/j.pdpdt.2016.12.007>
- Valdés PA, Jacobs V, Harris BT, et al. Quantitative fluorescence using 5-aminolevulinic acid-induced protoporphyrin IX biomarker as a surgical adjunct in low-grade glioma surgery. *J Neurosurg* 2015;123(3):771–780. <https://doi.org/10.3171/2014.12.JNS14391>
- Widhalm G, Olson J, Weller J, et al. The value of visible 5-ALA fluorescence and quantitative protoporphyrin IX analysis for improved surgery of suspected low-grade gliomas. *J Neurosurg* 2019. 1–10. <https://doi.org/10.3171/2019.1.JNS182614>
- Belloch JP, Rovira V, Llácer JL, Riesgo PA, Cremades A. Fluorescence-guided surgery in high grade gliomas using an exoscope system. *Acta Neurochir (Wien)* 2014;156(4):653–660. <https://doi.org/10.1007/s00701-013-1976-6>
- Lau D, Hervey-Jumper SL, Chang S, et al. A prospective Phase II clinical trial of 5-aminolevulinic acid to assess the correlation of intraoperative fluorescence intensity and degree of histologic cellularity during resection of high-grade gliomas. *J Neurosurg* 2015;124:1–10. <https://doi.org/10.3171/2015.5.JNS1577>
- Roberts DW, Valdés PA, Harris BT, et al. Coregistered fluorescence-enhanced tumor resection of malignant glioma: Relationships between δ -aminolevulinic acid-induced protoporphyrin IX fluorescence, magnetic resonance imaging enhancement, and neuropathological parameters. Clinical article. *J Neurosurg* 2011;114(3):595–603. <https://doi.org/10.3171/2010.2.JNS091322>
- Bekelis K, Valdés PA, Erkmen K, et al. Quantitative and qualitative 5-aminolevulinic acid-induced protoporphyrin IX fluorescence in skull base meningiomas. *Neurosurg Focus* 2011;30(5):E8. <https://doi.org/10.3171/2011.2.FOCUS1112>
- Valdes PA, Kim A, Brantsch M, et al. δ -aminolevulinic acid-induced protoporphyrin IX concentration correlates with histopathologic markers of malignancy in human gliomas: The need for quantitative fluorescence-guided resection to identify regions of increasing malignancy. *Neuro-Oncol* 2011;13(8):846–856. <https://doi.org/10.1093/neuonc/nor086>
- Valdés PA, Moses ZB, Kim A, et al. Gadolinium- and 5-aminolevulinic acid-induced protoporphyrin IX levels in human gliomas: An ex vivo quantitative study to correlate protoporphyrin IX levels and blood-brain barrier breakdown. *J Neuropathol Exp Neurol* 2012;71(9):806–813. <https://doi.org/10.1097/NEN.0b013e31826775a1>
- Piquer J, Llácer JL, Rovira V, Riesgo P, Rodriguez R, Cremades A. Fluorescence-guided surgery and biopsy in gliomas with an exoscope system. *BioMed Res Int* 2014;2014:207974–207976. <https://doi.org/10.1155/2014/207974>
- Valdés PA, Kim A, Leblond F, et al. Combined fluorescence and reflectance spectroscopy for in vivo quantification of cancer biomarkers in low- and high-grade glioma surgery. *J Biomed Opt* 2011;16(11):116007. <https://doi.org/10.1117/1.3646916>