A Pilot Study Investigating the Minimum Requirements Necessary for Grading Astrocytomas Remotely


OBJECTIVE: To investigate the minimum requirements necessary for grading remote astrocytomas in terms of selected static images and descriptive histologic characteristics.

STUDY DESIGN: A histopathologist examined 106 formalin-fixed, paraffin-embedded tissue samples of low- and high-grade astrocytomas. Interobserver-checked cases were reviewed under a microscope to estimate the accuracy of the conventional glass slide diagnoses. Then cases based on 5 static-digitized images from each case were diagnosed. Next, the grade of each tumor was assessed based on the set of 5 images and the World Health Organization (WHO) description of 8 histologic characteristics defined as crucial in grading astrocytomas. Finally, an evaluation was made using a custom-designed decision support system.

RESULTS: Conventional glass slide diagnosis was 93.9%. Diagnosis based only on the set of 5 images dropped to 81.6%. Diagnosis based on the set of 5 images and the WHO characteristics boosted accuracy to 88.8%. Accuracy improved to 91.8% with the addition of the decision support system.

CONCLUSION: Our findings suggest that a telepathology system might be valuable for accurate grade diagnosis of astrocytomas—providing a means for avoiding diagnostic errors—without blocks or slides having to leave the department. This could significantly reduce the overall time and cost of diagnosis.
Astrocytomas are considered to be among the most lethal and difficult-to-treat forms of cancer. The most significant step in diagnosing astrocytomas is grading—the determination of the degree of tumor abnormality. Grading is performed by expert histopathologists who visually inspect microscopic sections of biopsy material under the microscope. According to the World Health Organization (WHO) grading system, astrocytomas are classified by 3 grades (grades II, III and IV) on the basis of histologic criteria. Although the WHO system is the most popular for grading astrocytomas, the WHO guidelines are limited by the ambiguity of the descriptions used to define each grade. Some of these definitions become clear only by experience. This results in interobserver and intraobserver variability, which has been shown to significantly influence the quality of diagnosis, potentially leading to diagnostic errors. Errors in diagnosing astrocytomas arise from multiple sources, most important being (1) malignancy formation along a biologic continuum, where there is no specific biologic criterion to establish clear and understandable boundaries between low-grade (grade II) and high-grade (grade III or IV) tumors, (2) the tumor’s internal heterogeneity, which means that tumors are formed from various cell populations of different grades and (3) biopsy size and sample, because the most representative tissue region is not always selected for examination.

Recent developments in telepathology and decision support systems have introduced solutions toward more objective histologic grading. In telepathology, which is the electronic transmission of pathologic images from 1 location to another, the main purposes are case referral for expert opinion, consultation, emergency service in the absence of a resident pathologist, case conferences and education. Case referral and consultation are fundamentally important in achieving accurate decisions, especially for cases that are difficult to diagnose. Decision support systems have been introduced as second opinion tools for improving diagnostic accuracy. Such systems attempt to correlate in a systematic-quantitative manner the texture and morphology of the nuclei with the tumors’ malignancy status.

Though the grading of astrocytomas suffers from poor interobserver and intraobserver variability, we have found no studies investigating if telepathology systems could be used for teleconsultation of astrocytomas. Such systems—provided that they offer portability, sufficient image quality and compatibility with routine clinical standards—would promote the exchange of histologic data among different laboratories. In this way, standardization and reproducibility in the grading of astrocytomas could improve. Moreover, the integration of expert-driven decision support systems would contribute to more accurate decisions as second opinion tools.

However, to accept telepathology systems as valuable medical tools for diagnosing astrocytomas, it is essential to investigate their accuracy. Due to the substantial storage requirements, digitization of the entire specimen at different magnifications is impractical for static telepathology systems. Also, in daily clinical practice, physicians’ diagnoses are based on representative tumor regions. With both of these constraints in mind, we designed a pilot study to investigate if it is possible to perform a diagnosis of astrocytomas based on a limited set of images obtained from the tumors’ most representative regions, together with assessments of histologic tumor characteristics. Thus we attempted to evaluate whether or not reviewing high-quality static images of astrocytomas from remote sites is feasible to an acceptable degree of accuracy, enabling offline teleconsultation, case referral for expert opinion, emergency service in the absence of a resident physician, and case conferences for more accurate grade diagnosis of astrocytomas. Additionally, this pilot study was designed in accordance with the most widely accepted clinical protocols for astrocytoma grading—namely, the WHO scheme and the hematoxylin-eosin (H-E) staining—and it incorporated a physician-driven, custom-developed decision support system as a second opinion tool. The accuracy in the grading of astrocytomas was evaluated against conventional microscope glass slide diagnosis.

**Material and Methods**

Archive material from formalin-fixed, paraffin-embedded tissue samples of astrocytomas was obtained from 106 patients who had undergone surgery at the University Hospital of Patras between 1993 and 2002. Patient ages ranged from 11 to 70 years for grade II tumors, from 14 to 77 for grade III tumors’ malignancy status.
tumors and from 13 to 83 for grade IV tumors. All patients were treated with partial or total tumor resection. Most patients with high-grade tumors were postoperatively treated with radiation or chemotherapy or both. On average, 5 H-E–stained sections were generated from the same block for each case (patient). Sections were placed on slides for microscopic examination. Slides were examined to ensure that crucial histologic characteristics (i.e., presence or absence of necrosis) were similar for all sections for each case. Tumor grade was defined as low or high grade according the WHO grading system. Based on the archives, of the 106 biopsies, 33 were classified as low grade (grade II) and 77 as high grade (32 grade III and 40 grade IV).

An experienced histopathologist (P.R.) digitized 5 images (768×576) from the most representative regions of each tumor. The digitization was performed using a light microscopy imaging system consisted of a Zeiss Axiostar-Plus microscope (Zeiss, Germany) connected to a Leica DC 300 F color video camera (Leica, Germany). Two different objectives were used at magnifications of ×40 and ×16. The sampling rate in the object plane was about 0.3 μm per pixel. Three images were captured at a magnification of ×40 for facilitating evaluation of nuclear texture related histologic features, and 2 images were acquired at a magnification of ×16 for validation of cellularity, vascular proliferation and necrosis.

The study comprised the several stages, with each stage performed after a period of 1 month. During stage I (first month, first evaluation), the 106 cases were checked for interobserver variability between the official hospital’s archive diagnosis and a new diagnosis, performed by the collaborating histopathologist (P.R.). In stage II (second month, second evaluation), cases with interobserver concordance were resubmitted to the histopathologist in order to assess intraobserver variability. During stage III (third month, third evaluation), the histopathologist performed grade diagnosis based only on the 5 images obtained from each case using a familiar to her image viewer. At stage IV (fourth month, fourth evaluation), each case’s histologic evaluation data were made available (examples of high- and low-grade images with corresponding supplementary material are shown at Figures 1 and 2, respectively). Then the histopathologist defined the grade of each tumor based on the set of 5 images and the corresponding glass slide histologic data.

For stage V (fifth month, fifth evaluation), the histopathologist evaluated the same data used during stage IV, but this time also had a decision support system available. The decision support system was built as a decision-classification tree. The tree’s structure was designed by choosing the most important features—those features that minimized diversity (i.e., the “goodness” of a split of data into low- and high-grade groups). In this study the Gini’s diversity index was used. The input to the decision tree comprised the scores for the 8 histologic features assigned for each case. Its performance was evaluated using a 10-fold cross-validation process against the “gold standard,” which consisted of those cases checked for interobserver concordance.

**Results**

Concerning interobserver reproducibility, unanimous diagnosis was noted for 31 low-grade cases (88.6% agreement), 30 grade III cases (90.9% agreement) and 37 grade IV cases (97.4% agreement). For the purposes of this study, only the 98 cases for which consensus existed between the histopathologist’s diagnosis and the archive diagnosis were selected for further analysis. For these 98 cases the overall glass slide diagnosis was 93.9% (intraob-

![Figure 1](An example of a high-grade sample image along with corresponding supplementary material (physician’s assessment of the tumor’s histologic characteristics).
server variability). The accuracy of the histopathologist’s assessment for each case when reviewing only the 5 images is illustrated in Table II. Table III shows the accuracy of the assessment when reviewing the 5 images accompanied with the corresponding glass slide histologic evaluation data and Table IV when reviewing the 5 images and the corresponding glass slide histologic evaluation data and having the decision support system available.

Figure 3 illustrates the decision tree structure. The most important feature was recognized as the existence of necrosis, followed by cellularity, pleomorphism, apoptosis and mitoses. Figure 4 depicts the performance of the tree under the 10-fold cross-validation test with respect to the number of nodes. The best performance was obtained using 6 branch nodes, with 88.8% accuracy. Trees with fewer branch nodes resulted to reduced classification accuracies.

Discussion

There have been many attempts to improve telepathology systems, which operate using either “live” images (the local and the expert physicians review images at the same time) or “still,” static images (the expert physician reviews images offline). In live telepathology, promising studies combined decision support systems and live transmission of images through remotely controlled robotic microscopic systems in the case of malignant lymphomas and chronic lymphocytic leukemia. Though efficient, this form of telepathology requires expensive robotic instrumentation, fast networking speeds and real time teleconsultation from participating experts. On the other hand, static telepathology applications are more flexible in the sense that the review of images may be performed offline, simple and routine measuring equipment may be used, the speed of transmission is unimportant and specialized physicians may provide expert opinions even from home (on-call service) without having to travel to the laboratory facilities. Additionally, static telepathology systems resemble to a certain extent the diagnostic procedure followed by histopathologists in daily clinical practice. Indeed, grade diagnosis is performed on representative tumor regions. The latter complies to the WHO guidelines, which suggest that even if a small region of the specimen has, for example, characteristics of higher grade tumors, then the whole specimen should be characterized as higher grade 3.

In 27, the accuracy of diagnosis using static images from a static telepathology system has been shown to be inferior compared with conventional

<table>
<thead>
<tr>
<th>Histologic feature</th>
<th>Assessment</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellularity</td>
<td>Light, mild, marked</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Mitoses</td>
<td>Absent, present</td>
<td>1, 2</td>
</tr>
<tr>
<td>Apoptosis</td>
<td>Absent, present, marked</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Multinucleated cells</td>
<td>Absent, present, numerous</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Giant cells</td>
<td>Absent, present, numerous</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Vascular proliferation</td>
<td>Absent, present, marked</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Necrosis</td>
<td>Absent, present, marked</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Nuclear pleomorphism</td>
<td>Mild, moderate, marked</td>
<td>1, 2, 3</td>
</tr>
</tbody>
</table>
glass slides diagnosis (88.5% vs. 95.5%, respectively). Diagnostic misinterpretations were found to be image selection and quality. Our pilot study was designed to compensate such diagnostic misinterpretations when using static images by taking into account a scoring report of crucial histologic characteristics, such as necrosis and mitoses (Figures 1 and 2), that cannot be accurately assessed on representative images only. In our study, accuracy in conventional glass slide diagnosis was found to be 93.9%. When the histopathologist performed diagnosis based only on the 5 representative images from each case, accuracy dropped to 81.6% (Table II). This reduction was related mainly to image selection and the difficulty in accurate evaluation of necrosis, cellularity and vascular proliferation. When the 5 representative images were combined with an additional description of the corresponding histologic characteristics (Table III), accuracy rose to 88.8%. This additional information proved to be valuable, especially in cases for which the validation of the extent of necrosis, mitoses, apoptosis and cellularity was problematic. Moreover, when the decision support system was made available to the histopathologist, accuracy was further improved to 91.8% (Table IV), which is comparable to conventional glass slide diagnosis accuracy (93.9%). Thus, accurate diagnosis of astrocytomas is feasible using static images and corresponding descriptions of histologic characteristics, without significant loss in diagnostic accuracy.

### Table II
**Truth Table Illustrating the Accuracy of the Histopathologist’s Assessment Based Only on the 5 Images**

<table>
<thead>
<tr>
<th>Low</th>
<th>High</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>7</td>
<td>77.4</td>
</tr>
<tr>
<td>11</td>
<td>56</td>
<td>83.6</td>
</tr>
</tbody>
</table>

### Table III
**Truth Table Illustrating the Accuracy of the Histopathologist’s Assessment Based on the 5 Images Accompanied With the Corresponding Glass Slide Histologic Evaluation Data**

<table>
<thead>
<tr>
<th>Low</th>
<th>High</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>5</td>
<td>83.9</td>
</tr>
<tr>
<td>6</td>
<td>61</td>
<td>91.0</td>
</tr>
</tbody>
</table>

### Table IV
**Truth Table Illustrating the Accuracy of the Histopathologist’s Assessment Based on the 5 Images Accompanied With the Corresponding Glass Slide Histologic Evaluation Data and the Decision Tree**

<table>
<thead>
<tr>
<th>Low</th>
<th>High</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>4</td>
<td>87.1</td>
</tr>
<tr>
<td>4</td>
<td>63</td>
<td>94.0</td>
</tr>
</tbody>
</table>

Because the incidence rate of astrocytomas is approximately 5 cases per 100,000,1 histopathologists residing at remote medical centers are more likely to lack sufficient experience for performing accurate diagnosis. Under these conditions, the general practice is to either direct patients to more specialized centers or to post biopsy material to a central hospital. These procedures are time consuming and increase the cost of diagnosis. Our findings suggest that static telepathology systems might offer a means to overcome the latter limitations, because such systems could enable case referral for expert opinion, emergency service in the absence of a resident physician, case conferences and consultation. In this way, the overall time and cost of diagnosis might decrease significantly, without blocks or slides having to leave the department. More importantly, though, potential diagnostic errors may be avoided, because any physician (expert or nonexpert) may request a fast second opinion from another physician. In this way, standardization, interobserver variability and communication of experts

![Figure 3](image-url)  
**Figure 3** Decision tree structure for discriminating low- from high-grade astrocytic tumors.
residing in different medical centers might improve, to the benefit of patients.

Regarding the decision tree, 4 important points are worth making. First, the decision tree’s accuracy (88.8%) is comparable with similar methods presented in the literature for astrocytomas grading using the WHO grading and H-E–staining protocols (ranging from 88%\(^1\) to 89.7%\(^2\) and 92.1%\(^3\) accuracies). Most previous studies have been based on the aggregate properties of nuclei texture within a region of interest, without taking into account necrosis and vascularity, because their quantification still remains difficult to implement. However, necrosis, primarily, is a crucial criterion in grade assignment according to the WHO system.\(^3\) The proposed decision tree structure enables the inclusion of necrosis in the form of qualitative assessment.

Second, the proposed decision tree can be used interactively by the histopathologist, because a simple assignment of scores for the 8 examined histologic features is only needed for automated grade assessment.

Third, we have shown that the utilization of the decision tree improved diagnostic accuracy. When the histopathologist was provided with only the 5 representative images from each case and the corresponding description of the 8 histologic features, 2 high- and 1 low-grade case were assigned as between low and high grade. When the decision tree was made available, these 3 cases were classified to the proper grade category.

Fourth, the decision tree can be used to induce valuable conclusions regarding the importance of certain histologic features in grade assignment. The existence of necrosis, cellularity and pleomorphism proved to be the most significant factors for discriminating lower from higher grade tumors. For tumors with an absence of necrosis and moderate pleomorphism, the discrimination of low- from high-grade tumors depends on the presence or absence of mitoses. On the other hand, cellularity additionally affects the assignment of a tumor’s grade. For marked cellularity, the tumor is more likely to belong to higher grades, whereas for moderate cellularity and limited necrosis, higher grade tumors are those exhibiting apoptosis and mitoses.

A future application of this study would be the development of a Web-based telepathology system using Web services—that is, automated information services that enable the exchange and integration of large amounts of data over the Internet. The remote physician, by utilizing an easy-to-use Web-based interface, would be able to submit to the system static high-quality images, as well as descriptions of the histologic criteria describing tumors’ malignancy status and supplementary material. Every remote physician would have to register to the system in order to obtain a unique profile, enabling her or him to view all submitted queries and the progress or result of each case’s evaluation process. When a submission is completed by the remote physician (a completed submission will require a set of at least 5 static images from each case and corresponding score assignments of histologic criteria), the system would post an alert to an expert’s physician on-call device (which could be either a GSM cellular phone or a personal digital assistant [PDA] device), informing him or her that a new case has been submitted to the system. The alert would contain information regarding the urgency of the case. After reviewing of each case, the expert physician could submit a response to the system via the Web interface. The response would be made available to the remote physician, who would be notified by an automated e-mail alert that a review has been completed.

The significant advantage of such telepathology system would be that the expert physician could access the system through a common World Wide Web browser, rendering the system platform and location independent. The unique profile management system would enable secure access of the remote physician to only his or her cases. Additionally, no patient data would be submitted to the system, safeguarding patient anonymity. Finally,
Web service platforms ensure that no data would be lost during transmission, thus maintaining the quality of the images.

References