Quantifying the Value of In-house Consultation in Surgical Pathology

Andrew A. Renshaw, MD,¹ Nat E. Pinmar, MD,¹ Michael R. Jiroutek, MS,² and Mary L. Young, MS²

Key Words: Anatomic pathology; Surgical pathology; Cytopathology; Accuracy; Diagnosis; Second opinion; Sensitivity; Quality assurance; Continuous quality improvement

Abstract

In-house consultation is a well-known method to improve diagnostic accuracy and agreement, but the technique has not been well studied. We reviewed the results of in-house consultation in a large private hospital practice setting for a 1-month period and determined its effect on diagnostic accuracy using the final sign-out as the “gold standard.”

During this 1-month period, 352 cases were reviewed as in-house consultations. Initial complete agreement was found in 315 (89.5%) cases. Using the initial diagnosis as the test case and the final sign-out as the gold standard, of the 37 discrepant cases, 4 (1.1%) were thought to represent false-negative results, 1 (0.3%) a false-positive result, 3 (0.9%) differences in type, and 29 (8.2%) differences in diagnostic threshold. Disagreements in 10 cases were thought to be potentially clinically significant. Internal consultation was obtained on approximately 20% of all cases seen in the laboratory, and disagreements were found in 2% of all cases.

Internal consultation has a significant and measurable impact on the practice of surgical pathology.

Diagnostic accuracy is crucial in anatomic pathology, including surgical pathology and cytopathology. A very common method to increase diagnostic agreement and accuracy is to “show cases around,” that is, to obtain internal consultation before officially signing out the case. While such a practice is common and generally agreed to be of value,¹ despite numerous studies on agreement or accuracy in anatomic pathology,²¹ to date no study has examined the value of internal consultation. To do this, we reviewed the results of in-house consultation at our hospital for a 1-month period.

Materials and Methods

The types and number of disagreements obtained from internal consultation at the Baptist Hospital of Miami, Miami, FL, were compared as outlined previously¹³ during a 1-month period (July 2001). Material was limited to surgical pathology and nongynecologic cytology. For the purposes of this report, the final sign-out, which was the consensus reached after review by all observers, was interpreted as the “gold standard” for truth. Such a practice is commonplace, and, indeed, the standard for similar studies of diagnostic accuracy in the cytology literature¹⁴‐¹⁶ has been accepted by the Food and Drug Administration as an acceptable measure of accuracy and is not thought to represent simply a measure of interobserver agreement. Exact details accounting for the lack of accuracy in the review methods have been described elsewhere¹⁷,¹⁸; they are not included in this report since all reviews seemed to be very accurate.

A threshold disagreement is a disagreement in which a lesion is identified by both pathologists and both agree
on its general nature, but disagree as to its degree. Common examples of this would include the difference between atypical ductal hyperplasia and ductal carcinoma in situ and the difference between a diagnosis of “suspicious” and “positive” in cytology. Disagreements related to process type are called type disagreements. For example, a uterine tumor could be either endometrioid or papillary serous. Sensitivity is related to screening and measures whether a lesion is correctly identified. Specificity reflects the ability to recognize the absence of a lesion when it is not there, ie, the benign nature of a lesion, when consensus holds that a malignant process is not present. Clinically significant disagreements were defined as those that might potentially affect treatment, prognosis, or both.

During the study period, there were no formal guidelines in place regarding in-house consultation. Essentially anything that the pathologist thought required a second review was reviewed. The practice situation consists of 5 board-certified pathologists, 3 of whom are subspecialty board-certified in cytopathology, 1 in hematopathology, and 1 in dermatopathology. All 5 pathologists perform both surgical pathology and cytopathology, and their work is essentially restricted to material received at this laboratory. The vast majority (>90%) of all in-house consultations are performed the same day, so that the turnaround time is not significantly altered. While it is true that in-house consultation occurs for 20% of all cases, this is divided among all 5 pathologists, and since many consultations are about specific foci, the amount of time and effort needed to perform the consultation in many cases is less than that for a case being seen for the first time.

**Results**

The distribution of cases by anatomic site for each method is detailed in **Table 1**.

During this 1-month period, 352 cases were reviewed as in-house consultations. Initial complete agreement was found in 315 (89.5%) cases, leaving 37 discrepant cases. Of these discrepant cases, and using the initial diagnosis as the test and the final sign-out as the gold standard, 4 (1.1%) were thought to represent false-negative results, 1 (0.3%) false-positive results, 3 (0.9%) differences in type, and 29 (8.2%) differences in diagnostic threshold.

There were a total of 10 cases (2.8%) in which clinically significant disagreements were present. The details of these cases are shown in **Table 2**.

Internal consultation was obtained on approximately 20% of all cases seen in the laboratory, and disagreements were present in 2% of all cases.

**Discussion**

It is obvious that diagnostic accuracy is important in anatomic pathology, and it is well known that showing cases around can improve the diagnostic accuracy. However, to our knowledge, no previous study has attempted to quantitate the value of such efforts.

Our data strongly suggest that internal consultation has a significant impact on the final sign-out. Disagreements were identified in 10.5% of cases reviewed, and this represented 2% of the total volume of the laboratory. As detailed in Table 2, many of these had a significant impact on the clinical management of patients. In accord with previous reports, differences in diagnostic thresholds were the most common type of disagreement detected. Nevertheless, all types of errors were detected by this method, including false-negative cases.

While we were aware that we showed many cases, we were surprised to find that we were showing nearly 20% of our total volume. Apparently other laboratories also have the same high volume of intradepartemental consultation. Fortunately, in this study, there was agreement concerning the diagnosis in the vast majority of cases. This suggests that this level of consultation may be higher than is needed, but the most appropriate level of intradepartemental consultation is not known. Nevertheless, since there were disagreements in 2% of cases, it might be reasonable to conclude that, assuming great accuracy in case selection, in order to identify all cases with potential disagreement, at least this many cases should be reviewed.

The results of this study, like those of all previous studies that reviewed cases in which the discrepancy can be measured, clearly show that screening error (ie, false-negative
cases) remains an important and serious issue in anatomic pathology. In the year 2001, this issue was taken for granted in gynecologic cytology, but it was only a decade ago that experts stated that their laboratories did not make screening errors. A reasonable legal standard for gynecologic cytology is still being defined. While studies suggest that the rate of screening error is lower in surgical pathology than in gynecologic cytology, it is interesting to note that there is not a single study that has a sensitivity of 100%. The adage “there are two kinds of pathologists, those who have missed gastric cancer and those who will” is a surprisingly honest assessment of this issue from a bygone era, an era in which pathologists had more time, reviewed fewer cases, were not expected to identify such small lesions, and did not face litigation for every “mistake.” With more specimens to be reviewed, larger specimens (11-gauge breast cores instead of 14-gauge), an increased number of levels (3 for prostate, 5 for breast), and smaller clinically significant foci (atypical small acinar proliferation, atypical ductal hyperplasia, micrometastases in sentinel nodes), the sensitivity of routine surgical pathology can only decrease. In this environment, efforts must be made, just as they have been in gynecologic cytology, to define what is an acceptable error. In this context, the data are clear and consistent: some lesions, such as small foci of gastric cancer or prostate cancer, cannot always be identified, and missing such a lesion should not by itself define substandard care.

We have shown that intradepartmental consultation has a significant impact on the final diagnosis. During the period of this study, there were disagreements for 2% of all cases in the laboratory. While differences in diagnostic threshold remain the most common areas of disagreement, false-negative and false-positive cases also were identified. By identifying areas of disagreement and allowing a consensus to be reached, intradepartmental consultation is a valid way to improve diagnostic accuracy in a laboratory.

From the Departments of 1 Pathology, Baptist Hospital of Miami, Miami, FL; and 2 Biostatistics, University of North Carolina, Chapel Hill.

Address reprint requests to Dr Renshaw: Dept of Pathology, Baptist Hospital of Miami, 8900 N Kendall Dr, Miami, FL 33176.

References


Table 21

“Potentially Clinically Significant” Disagreements Detected by Intradepartmental Consultation

<table>
<thead>
<tr>
<th>Specimen</th>
<th>Original Diagnosis</th>
<th>Second Diagnosis</th>
<th>Type of Disagreement</th>
<th>Clinical Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast core</td>
<td>No calcifications</td>
<td>Calcifications</td>
<td>False-negative</td>
<td>No biopsy</td>
</tr>
<tr>
<td>Skin biopsy</td>
<td>No viral changes</td>
<td>Herpes</td>
<td>False-negative</td>
<td>Antiviral therapy</td>
</tr>
<tr>
<td>Colon biopsy</td>
<td>Tubular adenoma</td>
<td>And severe atypia/carcinoma in situ</td>
<td>False-negative</td>
<td>Surgery</td>
</tr>
<tr>
<td>Prostate resection</td>
<td>No carcinoma (hormone treatment)</td>
<td>Carcinoma</td>
<td>False-negative</td>
<td>—</td>
</tr>
<tr>
<td>Breast core</td>
<td>Atypical ductal hyperplasia</td>
<td>Focal atypia</td>
<td>Threshold</td>
<td>No surgery</td>
</tr>
<tr>
<td>Colon biopsy</td>
<td>Ischemia</td>
<td>Nonspecific changes</td>
<td>Threshold</td>
<td>Continues birth control pills</td>
</tr>
<tr>
<td>Colon biopsy</td>
<td>Ulcer with reactive changes</td>
<td>Ulcer, carcinoma cannot be ruled out</td>
<td>Threshold</td>
<td>Rebiopsy</td>
</tr>
<tr>
<td>Cervix biopsy</td>
<td>Low-grade squamous intraepithelial lesion</td>
<td>Reactive changes</td>
<td>Threshold</td>
<td>Rebiopsy</td>
</tr>
<tr>
<td>Breast fine-needle aspirate</td>
<td>“Suspicious” for malignancy</td>
<td>Positive for malignancy</td>
<td>Threshold</td>
<td>Mastectomy</td>
</tr>
<tr>
<td>Ascites</td>
<td>Suspicious for malignancy</td>
<td>Positive for malignancy</td>
<td>Threshold</td>
<td>Chemotherapy</td>
</tr>
</tbody>
</table>


