Histological Classification of Gastric Adenocarcinoma for Epidemiological Research: Concordance between Pathologists

Atsuko Shibata, Teri A. Longacre, Balaram Puligandla, Julie Parsonnet, and Laurel A. Habel

Abstract

Epidemiology of gastric adenocarcinoma suggests that intestinal-type and diffuse-type cancers develop through distinct causal pathways. To examine the differences in risk factors and molecular changes between the histological types, reliable data on pathological typing are essential. We evaluated the concordance between two pathologists in assessment of 95 gastric adenocarcinomas for Laurén classification and tumor grade. Two pathologists, each blinded to the other’s assessment, reviewed H&E-stained slides of gastric tumor. The responses of the two pathologists for pathological type were considered as concordant if they fell on one of the three categories (intestinal type, diffuse type, or other). Tumor grade was classified into three categories (well, moderately, or poorly differentiated). The pathologists agreed on the classification of histological type for 71 of 92 (77%) tumors. r coefficient was 0.59 (95% confidence interval, 0.44–0.73). Concordance for tumor grade was 87%, with a r coefficient of 0.72 (95% confidence interval, 0.57–0.87). Both observed concordance and r coefficient for histological type and tumor grade were similar across three calendar periods of study. Interobserver agreement was virtually identical between tumors with biopsy specimens only and those with surgical specimens. Although the level of disagreement for histological type observed in this study is comparable with that in other studies, the resulting misclassification would lead to the reduction in observed differences in prevalence and odds ratio estimates between two histological types.

Introduction

The most common form of stomach cancer is adenocarcinoma. Laurén (2) classified gastric adenocarcinoma into two histological types, intestinal and diffuse, according to morphological features of tumor. The decrease in intestinal-type tumors largely accounts for the declining trends of stomach cancer incidence (3), suggesting larger influences of environmental factors on intestinal-type cancer than diffuse-type cancer. Preneoplastic lesions, such as intestinal metaplasia, are often associated with intestinal-type cancer but are infrequently associated with diffuse-type cancer. Both observations suggest that the two histological types develop through distinct pathways (4).

Materials and Methods

Study Subjects. Study subjects were identified for a study of Helicobacter pylori serology and molecular changes in gastric cancer from the members of the KPMCP of Northern California who had participated in the MPHC program between 1964 and 1969. Diagnoses of gastric cancer between the date of participation in the MPHC program and June 1996 were ascertained by linkage of the MPHC database with the KPMCP’s local tumor registry and computer-stored hospitalization records. The study protocol has been approved by the human subjects panels of Stanford University and the Kaiser Foundation Research Institute.

We first screened pathology reports and H&E-stained slides to confirm the eligibility of subjects. The following tumors were excluded from this study: adenocarcinoma of the gastroesophageal junction; gastric neoplasia other than adenocarcinoma; metastatic tumor to the stomach from another organ; and metastatic cancer of gastric origin without concomitant histological material from primary tumor in the stomach. The present analysis includes 95 cases of gastric adenocarcinoma, and tumor specimens of these cases were reviewed by two study pathologists before August 1999.

The abbreviations used are: CI, confidence interval; IHC, immunohistochemistry; KPMCP, Kaiser Permanente Medical Care Program; MPHC, multiphasic health checkup; OR, odds ratio.

Received 6/6/00; revised 11/2/00; accepted 11/4/00. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 Supported in part by American Cancer Society Institutional Research Grant 32-36 and NIH Grants R01CA73011 and R35CA49761.

2 To whom requests for reprints should be addressed, at Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California 94305, and Department of Pathology [B. P.] and Division of Research [L. A. H.], Kaiser Foundation Research Institute, Oakland, California 94611.
Short Communication: Concordance of Gastric Cancer Typing

Concordance of Histological Type

The initial responses on histological type were available for 92 tumors. Intestinal-type tumors were the most common, accounting for over 50% of the tumors according to either pathologist. Overall, the pathologists agreed on the major classification for 71 of 92 (77%) cases (Table 1). The \( \kappa \) coefficient was 0.59 (95% CI, 0.44–0.73). The observed concordance for the three calendar periods was 79%, 79%, and 69%, with \( \kappa \) coefficients of 0.62 (95% CI, 0.57–0.87), 0.62 (95% CI, 0.43–0.82), and 0.40 (95% CI, 0.01 to 0.81). The level of agreement was almost identical, regardless of whether only biopsy specimens were available (concordance, 77%; \( \kappa \), 0.58) or surgical specimens were available (concordance, 79%; \( \kappa \), 0.61).

Concordance of Tumor Grade

Responses on tumor grade were available for 91 tumors. Poorly differentiated tumors were the most common (67% and 66% by two pathologists, respectively). Overall agreement for tumor grade was 87%, with a \( \kappa \) coefficient of 0.72 (95% CI, 0.57–0.87; Table 2). The observed concordance for the three calendar periods was 93%, 83%, and 88%, with \( \kappa \) coefficients of 0.85 (95% CI, 0.64–1.00), 0.64 (95% CI, 0.42–0.86), and 0.73 (95% CI, 0.37–1.00). Interobserver agreement was virtually identical between cases with biopsy specimens only (concordance, 88%; \( \kappa \), 0.70) and those with surgical specimens (concordance, 86%; \( \kappa \), 0.72).

Effects of Misclassification on Prevalence and OR Estimates

Example 1: Prevalence of p53 IHC-positive Tumors. If 10% of 100 intestinal-type tumors are misclassified as diffuse type, that a certain percentage (10% or 15%) of intestinal-type tumors was misclassified as diffuse type and that the same percentage of diffuse-type tumors was misclassified as intestinal type, independent of p53 IHC and \( H. pylori \) status (i.e., non-differential misclassification). The two percentages used in the examples were chosen to represent the level of disagreement between pathologists observed in this study. For the purpose of these illustrations, we assumed, although it is unrealistic, that there was no measurement error in p53 IHC and \( H. pylori \) status.

For the example of p53 IHC, we assumed that the true prevalence of p53-positive tumors was 60% in intestinal-type tumors and 30% in diffuse-type tumors (reasonable estimates based on work by A. S.) \(^5\). We estimated the difference between the two percentages observed in a study including 200 gastric tumors (100 intestinal-type and 100 diffuse-type tumors) when a specified level of misclassification occurred.

Histological Review of Tumor Specimens. H&E-stained slides from endoscopic biopsies and gastrectomies were obtained from pathology departments and a storage facility of the KPMCP. Biopsy specimens alone were available for 42 of 95 (44%) subjects. For each subject, one pathologist reviewed all of the available slides (mean = 4.7 slides/subject, including slides that did not contain tumor) and selected slides containing most representative tumor sections. The selected slides (mean = 1.8 slides/subject) were forwarded to the second pathologist. The two pathologists, each of whom was blinded to the other’s assessment, recorded their assessment for histological type and grade of tumor on an evaluation form. For each tumor, pathologists were instructed to identify a type (intestinal type, diffuse type, or other) and indicate the degree of differentiation (well, moderately, or poorly differentiated; Ref. 5).

Initial review of the 95 tumors was conducted between November 1997 and July 1999. The pathologists held joint review sessions in July 1998 and April 1999 to resolve disagreements between their assessments. Data analyzed in this study came from the initial assessments made by each pathologist before these joint reviews.

Assessment of Interobserver Agreement. Responses of the two pathologists regarding histological type were considered as concordant if their responses both fell on one of the three categories (i.e., intestinal type, diffuse type, or other). Observed concordance was calculated as a percentage of tumors with concordant responses (concordance was calculated as a percentage of tumors with concordant responses, \( 50 \) of 92). The observed concordance for the three calendar periods was 93%, 83%, and 88%, with \( \kappa \) coefficients of 0.85 (95% CI, 0.64–1.00), 0.64 (95% CI, 0.42–0.86), and 0.73 (95% CI, 0.37–1.00). Interobserver agreement was virtually identical between cases with biopsy specimens only (concordance, 88%; \( \kappa \), 0.70) and those with surgical specimens (concordance, 86%; \( \kappa \), 0.72).

Example 1: Prevalence of p53 IHC-positive Tumors. If 10% of 100 intestinal-type tumors are misclassified as diffuse type,
and 10% of 100 diffuse-type tumors are misclassified as intestinal type. 100 tumors classified as intestinal type include 90 “true” intestinal-type and 10 “true” diffuse-type tumors. If 60% of “true” intestinal-type and 30% of “true” diffuse-type tumors are p53 IHC positive (Table 3A), the number of p53 IHC-positive tumors in 100 “nominal” intestinal-type tumors is 90 × 0.6 + 10 × 0.3 = 57. Similarly, 33 of 100 tumors classified as diffuse type are p53 IHC positive. Consequently, the difference in prevalence of p53 IHC tumors between the two types is 57% – 33% = 24%, which is smaller than the true difference of 30%. Similarly, if 15% of tumors are misclassified, the difference in p53 IHC positivity between intestinal-type and diffuse-type tumors is 21%.

**Example 2: OR for Gastric Cancer Associated with *H. pylori* Infection.** When 10% of case subjects are misclassified, the observed OR for intestinal-type tumors increases from 5.0 to 5.4 (Table 3B), whereas the observed OR for diffuse-type tumors decreases from 10.0 to 8.9, resulting in a smaller ratio of the two ORs (1.6 rather than 2). If the misclassification occurs in 15% of tumors, the ORs for intestinal-type and diffuse-type tumors are even closer to each other.

**Discussion**

Frequently used histological classifications for gastric adenocarcinoma (intestinal-type and diffuse-type carcinoma) are attributed to Laurén (2). Laurén himself observed a stronger male preponderance and older age at diagnosis for intestinal-type cancer than for diffuse-type cancer (2), which was corroborated by Hanai et al. (7).

Histological classification of cancers, as is the case with all measurements, is prone to error. Histological typing and grading are inherently subjective because they depend on a pathologist’s judgement in the absence of objective “gold standard” measures. Pathologists would be more likely to disagree for tumors with overlapping histological patterns, which consist of as high as 15% of all gastric cancers (1, 8). Given the absence of a “gold standard,” examining the level of agreement between two or more pathologists is a feasible alternative to estimate the reliability of histological data.

Our study showed an overall agreement for Laurén classification between two pathologists in 77% of the 92 gastric adenocarcinomas; a k coefficient of 0.59 is considered as “fair to good” agreement (6). Both observed concordance and k coefficient for histological type were consistent across the three calendar periods and regardless of whether surgical specimens of tumor were available in addition to biopsy specimens. The level of agreement was higher for tumor grade, which was somewhat surprising to us because we had expected that decrease in prevalence of p53 IHC tumors between the two types is 57% – 33% = 24%, which is smaller than the true difference of 30%. Similarly, if 15% of tumors are misclassified, the difference in p53 IHC positivity between intestinal-type and diffuse-type tumors is 21%.

**Table 2** Overall concordance in tumor grade assessment between pathologists

<table>
<thead>
<tr>
<th>Pathologist 2</th>
<th>Well differentiated</th>
<th>Moderately differentiated</th>
<th>Poorly differentiated</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathologist 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well differentiated</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Moderately differentiated</td>
<td>0</td>
<td>21</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>Poorly differentiated</td>
<td>0</td>
<td>6</td>
<td>55</td>
<td>61</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>28</td>
<td>60</td>
<td>91</td>
</tr>
</tbody>
</table>

*Observed concordance = (3 + 21 + 55)/91 = 87%. k coefficient = 0.72 (95% CI 0.57–0.87).*

**Table 3** Effects of misclassification on prevalence and OR estimates in two hypothetical examples

<table>
<thead>
<tr>
<th>Level of histological misclassification</th>
<th>A. Observed prevalences of p53 immunohistochemistry-positive tumors by histological type</th>
<th>B. Observed ORs for <em>H. pylori</em> infection and gastric cancer risk by histological type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intestinal-type cancer</td>
<td>60%</td>
</tr>
<tr>
<td>Difference</td>
<td>30%</td>
<td>24%</td>
</tr>
<tr>
<td>Ratio of the two ORs</td>
<td>2.0</td>
<td>1.6</td>
</tr>
</tbody>
</table>

and OR estimates between histological types, such that those differences might be discounted. It would be difficult to assess the degree and direction of bias for actual studies, in which additional measurement error occurs in other variables, and histological misclassification might not be independent of other marker measurements.

Extrapolation of our findings to other settings may be limited due to some aspects of study design. First, a second pathologist reviewed only a subset of slides that a first pathologist considered as most representative, which may have resulted in overestimated k coefficients. Second, the joint review by two pathologists may have influenced subsequent typing of tumors, although we observed no increase in concordance over the calendar periods.

In conclusion, we observed reasonably good agreement.
between pathologists for Laurén classification and grade of gastric cancers. A central review of pathology material and resolution of disagreements between pathologists will be practical alternatives to improve the reliability of histological data in epidemiological studies.

Acknowledgments

We thank Maria Isabel Garcia, Lucy Bonds, Bonnie Bell, and Bill Frank for technical assistance; the pathology departments of Kaiser Permanente Medical Centers in Northern California for tumor specimens; and Dr. Gary Friedman for access to the KPMCP databases.

References
