EDITORIAL

Counterpoint

Histologic Grade as a Prognostic Factor in Breast Carcinoma

Harry B. Burke, M.D., Ph.D.[1]
Donald Earl Henson, M.D.[2]

[1] Bioinformatics and Health Services Research, Department of Medicine, New York Medical College, Valhalla, New York.

[2] Division of Cancer Prevention, National Cancer Institute, National Institutes of Health, Bethesda, Maryland.

In this issue of *Cancer,* Dr. Roberti reviews the role of histologic grade in the prognosis of breast carcinoma and wonders why, because it is available, it has not been widely used in predicting outcome.[1] The position of this editorial is that there must be some fundamental reason, after 100 years of progress on histologic grade, that confusion persists regarding its prognostic value.

The systematic use of morphologic variation at the cellular level of analysis as a prognostic factor in cancer has been fraught with controversy. Currently, there is no universally agreed on set of necessary and sufficient conditions for the definition of histologic grade in breast carcinoma. There has been uncertainty regarding the identification of what variation was important, how the variation should be organized, and whether it should be integrated into a staging or index system.

An additional issue is that grading system criteria have been selected based on their ability to create subgroups of patients using histologic distinctions to produce significant differences in outcome. There are two problems with this approach. First, there are many possible criteria that can create significant differences between subgroups and there is no analytic method for finding the best criteria.[2] Second, statistical significance is not necessarily accuracy. Significance is the chance that two or more distributions of variables, as represented by their parameter estimates, for example, means and variances, are really the same. Accuracy assesses the strength of association between two or more variables.[3,4] In general, accuracy quantifies how good a variable is at predicting another variable. Specifically, we are interested in the strength of association between grade and survival, i.e., how good is grade at predicting survival.

Fundamentally, grade remains controversial because it confounds two types of time. One type is how long the tumor has been growing and the other is how rapidly it has been growing. A "high grade" tumor could be an indolent tumor that grew for a long time prior to discovery and will continue to be slow growing; alternatively, it could be an aggressive tumor of recent origin that will continue to be rapidly growing. Because one can never know when a tumor originated, it may not be possible on histologic grounds to separate

Supported in part by a research grant from the U.S. Army Medical Research and Development Command Breast Cancer Research Program (DAMD 17-94-J-4383).

See reply to counterpoint on pages 1706–7 and referenced original article on pages 1708–16, this issue.

Address for reprints: Donald Earl Henson, M.D., Division of Cancer Prevention, National Cancer Institute, Bethesda, MD 20892.

Received May 9, 1997; accepted May 15, 1997.

a slowly growing tumor from a rapidly growing tumor. In other words, one cannot always distinguish how long the tumor has been growing from how fast it has been growing. The extent to which time ambiguity exists in grade is the extent to which grade's prediction variance will increase and consequently the extent to which its prediction accuracy will decrease. This limits grade's independent prognostic value and its ability to add significant prognostic value when placed in a system that includes other time-related factors such as tumor size.

The mechanical theory of cancer, a view espoused by Halsted,[5] assumes that cancer spreads from the primary tumor to the regional lymph nodes and then to distant sites of the body. This view is the basis of the TNM staging system. For the mechanical theory, the primary purpose of a prognostic system is to capture the spread of the cancer because cancer spread is believed to be the best indicator of outcome. The three elements of the TNM staging system (local tumor, regional lymph node, and distant metastasis[6]) are believed to reflect directly the spread of cancer, i.e., the extent of disease. Grade is not one of the TNM variables because it does not fit into this mechanical epistemology; it does not directly reflect the spread of the cancer. However, even if grade could have been subsumed within the mechanical theory of cancer, it would not have replaced tumor size in breast carcinoma. Using the Surveillance, Epidemiology, and End Results data of the National Cancer Institute for 1983–1987 and the area under the receiver operating characteristic curve as the measure of accuracy (Az), we found the Az for grade alone to be .634 and the Az for tumor size alone to be .737 ($P < .05$) for 5-year survival. Furthermore, grade does not add prognostic accuracy to tumor size; the Az for tumor size and grade combined was .749, which was not significant when compared with tumor size alone. In addition, grade could not have been added to the TNM staging system because the system is a bin model comprised of five levels of tumor characteristics (T), four levels of regional lymph node involvement (N), and two levels of distant metastasis (M).[7] Adding the 4 levels of grade to the 40 bins of the TNM ($5T \times 4N \times 2M$) would have created 160 bins and made it too complex to be useful.[7]

What is the future of grade as a prognostic factor in breast carcinoma? If we no longer accept the mechanical theory of cancer spread, grade becomes a possible prognostic factor. In addition, because the TNM staging system is not very accurate[8] new computer-based prognostic systems are being developed.[7] Computer-based prognostic systems are more accurate in predicting outcome and they do not have a limitation on the number of variables that can be used. Can grade be an independent prognostic factor in a computer-based system or can grade substitute for another more difficult to assess factor such as lymph node status?

We evaluated the ability of grade to predict 5-year breast carcinoma survival using data from the National Cancer Institute's SEER program.[9] The data were collected between 1983–1987 and the patients were followed for at least 5 years. The variables were tumor size, local extent of disease, lymph node status, and histologic grade. The criteria used to determine grade were neither standardized nor explicitly reported. The data set did not include cases with metastatic disease because grade is infrequently reported in these patients. Only 14,704 of the 48,643 cases were graded (30%). All analyses without grade were performed on the full data set of 48,643 cases. An analysis using the subset of graded cases favors grade because it is almost certainly the case that the variance of grade would increase if all the cases were graded. The area under the receiver operating characteristic curve was the measure of prediction accuracy. We used the logistic regression statistical method to create our models (SAS Institute, Cary, NC) and all results were performed on the test data set.

The predictive accuracy of tumor size, local tumor extent, and lymph status was .794. Adding histologic grade slightly increased the Az to .797, but this was not significant. Therefore, in a statistical model with traditional prognostic factors, grade does not add prognostic accuracy.

Can histologic grade substitute for a factor that is becoming difficult to evaluate (e.g., lymph node status). To answer this question, we created a logistic regression model in which grade was the predictor and lymph node metastasis (detected vs. not detected) was the outcome. This addressed the issue of how well grade can take the place of lymph node status as a prognostic factor (in other words, to what extent does their prognostic information overlap?). If their predictions completely overlap, then the observed Az would be 1.0; if there was no overlap, then the observed Az would be .5. Again using the SEER data set, we found an Az of .589, which indicated that there was very little predictive overlap. Therefore, grade is not an effective surrogate for nodal status.

If grade is to be a useful prognostic factor in the future it must improve predictive accuracy for women with small tumors and few involved lymph nodes when used in predictive models that include the new molecular genetic prognostic factors. The data set from Duke University, kindly provided by Dr. Jeffrey Marks, includes patients with early stage breast carcinoma. These data were described in a previous arti-

cle.[10] Briefly, all patients were pathologic TNM Stage I or early Stage II. Early Stage II included all TNM Stage II patients except those with five or more positive lymph nodes. The variables were age, race, tumor size, positive lymph nodes, TNM lymph node status, nuclear grade, histologic grade, p53, c-*erb* B-2 (HER-2/*neu*), estrogen receptor status (ER) and progesterone receptor status (PR), vascular invasion, adjuvant therapy (tamoxifen, chemotherapy), and radiation therapy. Patients who underwent a lumpectomy received radiation therapy. Patients who underwent a modified radical mastectomy did not receive radiation therapy. There were 229 cases, 226 of which had complete data for all variables except ER and PR status. Because many individual patient ER and PR values were missing, both variables were removed from the data set. The 5-year survival rate was 70%. The logistic regression statistical method was used to create the models and a prediction endpoint of 5-year overall survival.

Neither histologic grade nor nuclear grade added any predictive power to the new molecular genetic prognostic factors in the logistic regression model. The predictive accuracy for all factors excluding histologic and nuclear grade was .733; when histologic grade was added the Az was .738 (not significant), when nuclear grade was added the Az was .736 (not significant), and when both were added the Az was .740 (not significant).

Overall, the accuracy of the Duke University logistic regression models was lower than the SEER logistic regression models because outcome prediction for early stage breast carcinoma was more difficult than outcome prediction for early and late stage breast carcinoma. In the Duke data set, the TNM staging system performed at chance level when predicting the outcome of women with early stage breast carcinoma, the Az was .567.[11]

Histologic grade alone has modest prognostic value. However, grade does not significantly increase the predictive accuracy of computer-based prognostic systems, either in data sets that represent all stages of breast carcinoma and contain traditional predictive factors or in data sets that represent early stage breast carcinoma and contain the new molecular genetic prognostic factors.

REFERENCES

1. Roberti NE. The role of histologic grading in the prognosis of patients with carcinoma of the breast: is this a neglected opportunity? *Cancer* 1997;80:1708–66.
2. Burke HB. Integrating multiple clinical tests to increase predictive accuracy. In: Hanausek M, Walaszek, Z, editors. Methods in molecular biology: tumor marker protocols. Mahwah, NJ: Lawrence Erlbaum Associates, 1997. In press.
3. Hays WL. Statistics. 5th edition. Fort Worth, TX: Harcourt Brace, 1991:335–8.
4. Burke HB. Evaluating artificial neural networks for medical applications. Proceedings of the 1997 International Congress of Neural Networks, 1997.
5. Halsted WS. The results of radical operations for the cure of carcinoma of the breast. *Ann Surg* 1907;46:1–19.
6. Beahrs OH, Henson DE, Hutter RVP, Kennedy BJ, editors. Manual for staging of cancer. 4th edition. Philadelphia: J.B. Lippincott, 1992.
7. Burke HB, Henson DE. Criteria for prognostic factors and for an enhanced prognostic system. *Cancer* 1993;72:3131–5.
8. Burke HB, Goodman PH, Rosen DB, Henson DE, Weinstein JN, Harrell FE Jr., et al. Artificial neural networks improve the accuracy of cancer survival prediction. *Cancer* 1997;79:868–73.
9. Shambaugh EM, Ries LG, Young JL Jr. Extent of disease: new 4-digit schemes, codes and coding instructions. Bethesda, MD: National Institutes of Health, National Cancer Institute, Biometry Branch, 1984.
10. Marks JR, Wu K, Berry D, Bandarenko N, Kerns BJ, Iglehart JD. Overexpression of p53 and HER-2/neu proteins as prognostic markers in early stage breast cancer. *Ann Surg* 1994;219:332–41.
11. Burke HB, Hoang A, Iglehart JD, Marks JR. Predicting response to adjuvant and radiation therapy in early stage breast cancer. *Cancer*. In press.