Histologic Grading of Breast Cancer: Linkage of Patient Outcome with Level of Pathologist Agreement

Leslie W. Dalton, M.D., Sarah E. Pinder, M.R.C.Path., Christopher E. Elston, F.R.C.Path., Ian O. Ellis, F.R.C.Path., David L. Page, M.D., William D. Dupont, Ph.D., Roger W. Blamey, F.R.C.S.

Austin Pathology Associates (LWD), Austin, Texas; The City Hospital NHS Trust (SEP, CEE, IOE, RWB), Nottingham, United Kingdom; and Vanderbilt University (DLP, WDD), Nashville, Tennessee

In the histologic grading of invasive breast cancer with the Nottingham modification of the Scarff-Bloom-Richardson grading scheme (NSBR), it has been found that when pathologists disagree, they tend not to disagree by much. However, if tumor grade is to be used as an important parameter in making treatment decisions, then even this generally small degree of pathologist variability in assessing grade needs to be correlated with patient outcome.

Findings from the Nottingham/Tenovus Primary Breast Cancer Study were used for patient outcome data. Kaplan-Meier survival curves were constructed for NSBR scores grouped according to the level at which pathologists tend to agree in assessing grade, from a reproducibility perspective. For example, if a given tumor were assessed by several pathologists as having either an NSBR score of 5 or 6, then what is the correct score—the intermediate-grade Score 6 assessments or the low-grade Score 5 assessments? By "regrouping" the Nottingham outcome data such that data from patients with Score 5 tumors are grouped with patients having Score 6 tumors (a 5–6 group), then the level in which the pathologists agreed with each other (that the tumor was either score 5 or 6) is better matched with patient outcome.

In response to the above example, it was not surprising to find that patients with Score 5–6 tumors had a probability of survival between the established low and intermediate NSBR final combined grades. However, it is the discussion of this approach that highlights that optimal use of grading requires awareness of the level of pathologist agreement and understanding the value of pathologists' reaching consensus in assessments. Also, knowledge of possible clinical decision thresholds can help in providing relevant interpretations of grading results.

KEY WORDS: Agreement, Breast cancer, Breast carcinoma, Grade, Prognosis, Reproducibility.

That the histologic grading of invasive breast carcinoma, in particular the Nottingham modification of the Scarff-Bloom-Richardson (NSBR) grading scheme, provides significant prognostic information has been clearly shown (1). One of the reasons that the application of this valuable prognostic indicator has been limited is concerns about reproducibility. Several studies have addressed the reproducibility issue and have found that pathologists grading of breast cancer is to a degree reproducible when the Nottingham or similar method is used (2–9). Ideally, use of the NSBR method would show that pathologists grades were to a high degree reproducible within a given Nottingham score. This is not the case and, admittedly, the level of pathologist reproducibility could conceivably be problematic if grade were being used to strictly classify patients into a treatment regimen whereby patients with a given NSBR score were treated in one way and patients with a score of just one away were treated in an entirely different manner. For optimal application of the NSBR grading method, this concern needs to be addressed. This need to study pathologist reproducibility in grading with respect to the end point of patient outcome is in response to concerns previously discussed (10).

In using the NSBR method and to arrive at a final combined NSBR grade, first an NSBR score is assessed. Among the seven possible NSBR scores, Scores 3, 4, and 5 tumors are grouped as low-grade tumors; Scores 6 and 7 are intermediate grade; and Scores 8 and 9 tumors are consolidated as high-grade neoplasms. In previous work (1), the optimal grouping of scores into grade was determined by patient outcome, and the pertinent patient out-

come is shown in the patient survival curve depicted in Figure 1; note that there is excellent and significant stratification of patients into a low-grade, favorable-prognosis group from those with poorer prognosis.

From a reproducibility perspective, the need to group NSBR scores in a certain way also becomes evident. One study (4) showed that in a given case, pathologists tend to cluster in their opinions around two adjacent NSBR scores. This study involved 25 pathologists who graded a representative slide from each of 10 breast cancer cases. In each case, the majority of opinions were concentrated around a particular NSBR score and a score that was either one less or one greater than the single most favored score. When an individual pathologist scores a tumor, it is not known whether the adjacent score, or "clustering," would be in a direction of +1 or −1 from the initial evaluation, unless, of course, the initial score assessed is at the NSBR limits of Score 3 or Score 9. It then follows that strictly from an individual reproducibility perspective, the NSBR scheme is best divided into score 3–4, 3-4-5, 4-5-6, 5-6-7, 6-7-8, 7-8-9, and 8–9 groups (*e.g.,* an initial assessment of Score 4, with ±1 clustering, results in a 3–4-5 individual reproducibility group). It requires the assessment by other pathologists, or an assessment by consensus, to detect whether the adjacent score is either +1 or −1 (or, less common, even a greater difference) from the initial individual opinion. By exclusion of a +1 or −1 possibility (or again, less common, even greater discrepancies), consensus grading results in the "adjacent score groups" of 3–4, 4–5, 5–6, 6–7, 7–8, 8–9, or termed the consensus reproducibility groups. Thus, from a reproducibility perspective, a consensus is considered reached if the possible score assessments in a given case can be narrowed

to two neighboring scores as this seems to be the natural tendency of what pathologists can easily achieve.

An example is the situation in which several pathologists in assessing the grade of a given tumor are evenly split as to whether the tumor is a Score 5 or a Score 6 tumor (a 5–6 tumor). In the prior reproducibility study (4), Case 9 was such an example with almost an even split among 25 pathologists. With such a situation, we wanted to visualize the survival probability if patient outcome data from both Score 5 and Score 6 tumors were pooled (as well as the other reproducibility groups) and in this way we might better match the variability in score assignment (what pathologists can do) with patient outcome.

The question then becomes how best to reconcile the prognostically determined low-, intermediate-, and high-grade consolidations with the overlapping ranges inherent with reproducibility groups, such as the 5–6 example. The goal was to determine whether attempting to answer this question could better define the practice of NSBR grading.

PROCEDURE

The Nottingham/Tenovus Primary Breast Cancer Study was established in 1973 to provide a means for the evaluation of possible prognostic factors in invasive breast cancer (1). At the time of this study, 1949 of the patients enrolled in the Nottingham study had an NSBR score assessed. Of these patients, 454 had died and the median time until death was 38 months. Of the 1394 who remained alive, the median follow-up was 54 months; 707 patients of those alive had been followed for more than 5 years, and 230 had been followed for more than 10 years. The Nottingham/Tenovus project has resulted in a valuable resource for correlating patient outcome with tumor data.

In this investigation the Nottingham/Tenovus follow-up data were studied to view the effect on patient outcome after grouping of scores from a reproducibility perspective. Aside from including patients from more recent years, this was accomplished by nothing more than using the same Nottingham data as reported previously but regrouping the data with the reproducibility question in mind. Therefore, the population of patients and subsequent evaluations of the population are essentially the same as previously reported.

From this regrouping of data, patient survival curves were constructed to visualize patient follow-up according to reproducibility-determined NSBR groups as compared with the prognostically determined categorization. These Kaplan-Meier survival curves (11) of the reproducibility groups
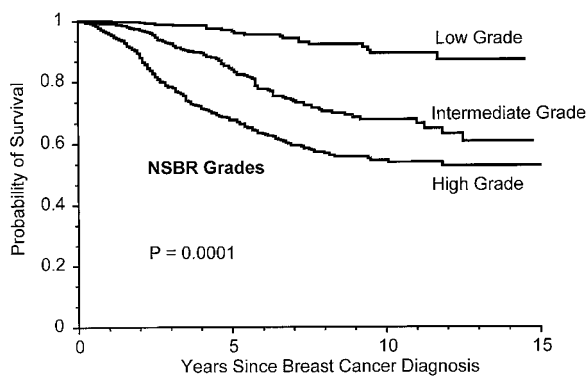
| Grade	Patients at Risk		
Low	381	185	78
Intermediate	661	244	68
High	907	278	84

FIGURE 1. Comparison of probability of patient survival with grouping of tumors into Nottingham Scores 3, 4, 5 (low grade), Scores 6, 7 (intermediate grade), and Scores 8, 9 (high grade).

are shown in Figures 2 and 3, and the prognostically determined categorization is shown in Figure 1.

RESULTS

The results of this rather simple rearrangement of the Nottingham data are not surprising, and it is the relevant discussion drawn from this exercise that warrants the most attention. However, as for the results, it is readily seen by inspection of the Kaplan-Meier curves that the grouping of NSBR scores according to reproducibility groups achieves essentially the same degree of patient stratification as achieved with grouping according to the 3–4-5 low-, 6–7 intermediate-, and 8–9 high-grade classifications. This is not surprising because both ways of grouping have no more than three NSBR scores consolidated into any group. Also, both the individual reproducibility categorization and the prognostically determined groups include a 3–4-5 consolidation. In the consensus reproducibility categorization, there is redundancy of 6–7 and 8–9 consolidations with the prognostically determined groups.

It should be emphasized that despite some redundancy, the reproducibility-determined groups and the prognostically determined groups have differences in implied usage. For example, a population of patients with near 90% 10-year survival probability is stratified in the low-grade (Score 3–4-5) prognostically determined classification. The 3–4-5 reproducibility determined category, of course, has the identical patient follow-up. In the prognostically determined categorization, a diagnosis of a low-grade tumor would be rendered if an individual pathologist assessed a given tumor as having any of the NSBR Scores 3, 4, or 5. However, a 3–4-5 tumor in the reproducibility-determined categorization is for tumors for which an individual pathologist assesses a score of 4, with Scores 3 and
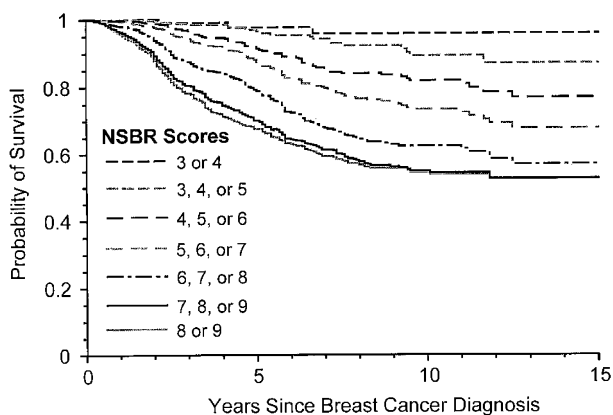
FIGURE 2. Comparison of probability of patient survival with patient stratification according to the reproducibility groupings as relevant to grading by an individual pathologist.
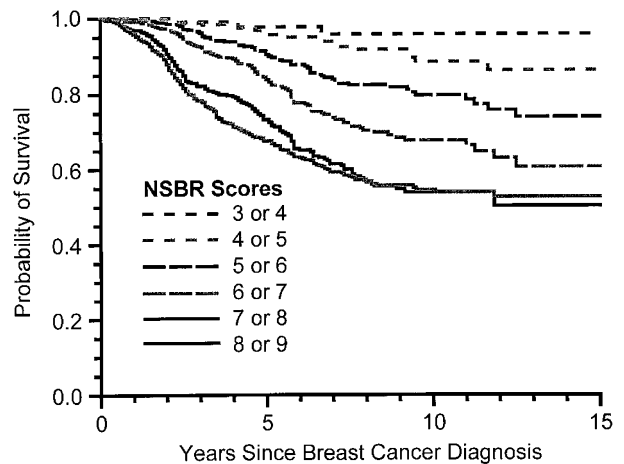
FIGURE 3. Comparison of probability of patient survival with patient stratification according to the reproducibility groupings as relevant to grading by consensus.

5 becoming inclusive because of the tendency for a ±1 level of agreement among pathologists.

The result is that entirely from a reproducibility perspective, a tumor that is given a score of 5 by an individual pathologist cannot be automatically placed into the low-grade 3–4-5 prognostically determined group but would instead represent a 4–5-6 tumor. It should come as no surprise that inclusion of Score 6 patient survival data shows that a 4–5-6 tumor has a prognosis that is worse (Fig. 2) than the low-grade categorization shown in Figure 1 and intermediate between the prognostically determined low and intermediate grades (again, Figs. 1 and 2). For more precise classification of a 4–5-6 tumor, obtaining a consensus assessment might help. A patient with a 4–5-6 tumor could be placed in the favorable low-grade prognostic category if additional assessment(s) by other pathologists all are Score 5 (a "solid 5") or the tumor is a 4–5 tumor. A 4–5 tumor (Fig. 3) has a prognosis very near the low-grade (3–4-5) consolidation. If by consensus assessment the "adjacent" score is indeed a 6, then the 4–5-6 tumor becomes a 5–6 tumor, in which the survival probability is less than 90% at 10 years (Fig. 3).

Assuming a consensus assessment, the other reproducibility group that crosses a boundary of NSBR final combined grade are 7–8 tumors (Fig. 3). The 7–8 group has a prognosis very similar to high-grade tumors.

DISCUSSION

Finding that the probability of survival of a 5–6 reproducibility group is between the survival probability of the established low and intermediate NSBR grades was virtually self-evident, so why this exercise in rearranging data? The reasoning should become clear when also considered is a suggested

clinical "rule of thumb." It has been suggested that breast cancer patients with greater than 90% probability of survival likely are not in need of adjuvant cytotoxic chemotherapy (12). The thought is that when predicted mortality reaches 10% or less, only a very few patients will benefit from adjuvant therapy given that adjuvant chemotherapy reduces mortality by only 20 to 30%, and the few who might benefit from therapy are balanced by the negative aspects of therapy. This is an especially compelling proposition in that tumors of different grade seem to respond differently to chemotherapy (13). Therefore, a patient outcome threshold of 90% survival serves as an illustrative landmark. Histologic grade as a prognostic parameter yields near 90% 10-year survival in tumors that are assessed as being low grade. However, what is seen by this study is that a 5–6 consolidation, which includes a low-grade score, yields a probability of survival below 90% (Fig. 3).

A way to address the 5–6 concern is first to look at the extremes. A lesser concern about reproducibility in grading is warranted the more contrasting a given NSBR score is from the scores that border a decision threshold. There should be little, if any, trepidation on the part of either clinician or pathologist that a tumor assessed as being a high-grade tumor is in truth a low-grade tumor. This is supported by several reproducibility studies in which high- *versus* low-grade discrepancies are rare (Table 1). It follows that the appropriate level of concern about reproducibility is best visualized by the depiction of a gray zone that represents those scores near a given decision threshold. Score assessments that are distant from a decision threshold are more black and white and more certain (Fig. 4). The existence and handling of "gray areas" or "fuzziness" is becoming more accepted and refined under such concepts as Bayesian belief networks (14).

For example, if a treatment decision threshold (*e.g.*, use of the "90% rule of thumb") calls for withholding adjuvant treatment to patients with low-grade tumors, then the decision to treat all patients with high-grade tumors should be readily made, as high-grade tumors are quite in contrast ("in the black") from the 5–6 low- to intermediate-grade boundary of NSBR scores. So, in a given case, there may be disagreement ("lack of reproducibility") among pathologists as to whether a tumor is intermediate or high grade (with commensurate degradation of a κ score), but this disagreement is moot if the clinical decision requires a low-grade designation. In such cases, the knowledge of what a tumor "is not" can be more important than what a tumor "is."

The delineation of possible reproducibility gray zones becomes even more meaningful when the more widely accepted prognostic parameters of lymph node status and tumor size are also consid-

TABLE 1. Number of High- *Versus* Low-Grade Discrepancies and Percentage Intermediate-Grade Assessments in Eight Reproducibility Studies in Which All Used a Semiquantitative Method that Was the Nottingham or Similar Method

Reference	Number of Evaluations	High/Low Discrepancies	Percentage Intermediate Grade
Frierson *et al.*, 1995 (2)[a]	75 cases	4 cases (5.3%)	32
Frierson *et al.*, 1995 (2)[b]	450	4 or 5 (<2%)	NA
Robbins *et al.*, 1995 (3)[c]	100	0	32
Robbins *et al.*, 1995 (3)[c]	100	0	27
Dalton *et al.*, 1994 (4)[d]	250	1 or 2 (<1%)	41
Dalton *et al.*, 1994 (4)[e]	60	0	38
Harvey *et al.*, 1992 (5)[f]	152	0	43
Theissig *et al.*, 1990 (6)[g]	332	1 (<1%)	46
Hopton *et al.*, 1989 (7)[h]	1748	21 (<2%)	45
Davis *et al.*, 1986 (8)[i]	2640	45 (<2%)	48
Delides *et al.*, 1982 (9)[j]	158 cases	44 cases (28%)	NA

[a] Of six pathologists grading 75 tumors, only four cases had any high/low discrepancies (actual data not in reference; H.F. Frierson, personal communication, 1999).

[b] See [a]: $6 \times 75 = 450$ individual evaluations in which the high/low discrepancy is four or five depending on whether the four who differed or the five who differed represented the most inappropriate discrepancy (actual data not in reference, personal communication with Dr. Frierson).

[c] At the consensus level (two pathology groups grading 50 cases equals 100 evaluations), there were no high/low discrepancies in training set.

[d] Twenty-five pathologists graded 10 cases for 25×10) 250 evaluations. Only one or two of 250 individual opinions were of a third grade when the other opinions were spread over the other two grades in each of the 10 cases. One or two is used in that it is not known whether the one who differed or the two who differed represent the most inappropriate discrepancy.

[e] At the consensus level (six pathology groups grading 10 cases equals 60 evaluations), there were no high/low discrepancies.

[f] Two pathologists graded 76 cases for 152 evaluations.

[g] The grading of a consensus of two pathologists compared with an expert senior pathologist's in 166 cases for 332 evaluations.

[h] One of 10 pathologists graded 874 cases and comparison was made to a "control" pathologist who graded all cases ($874 \times 2 = 1748$ evaluations).

[i] Local pathologists submitted 1320 cases for central review. Thus, 1320 cases were evaluated twice ($1320 \times 2 = 2640$ evaluations).

[j] Six pathologists from the same group using the World Health Organization's method graded 158 tumors. Numeric data were not given to the level of grading by individual pathologist. By far the most disappointing performance, and this was possibly due to level of training as five of the pathologists were still pathologists in training (fourth year), and the sixth was their "tutor."

ered. Axillary lymph node status, tumor size, and tumor grade compose the three variables of the Nottingham Prognostic Index (NPI) (15). With respect to the NPI, if an invasive breast cancer is less than 2 cm and is lymph node negative, then a low-grade tumor is well within the confines of the favorable prognostic category and is considered to be within an "excellent prognostic" subgroup. Furthermore, intermediate-grade tumors are also associated with favorable prognosis, provided, again, that lymph nodes are negative and tumors are of small size (tumors having an NPI of less than 3.4). Therefore, with the intermediate-grade category serving as a "buffer," it is extremely unlikely, especially after consensus assessment, that a node-negative patient with a small, low-grade tumor
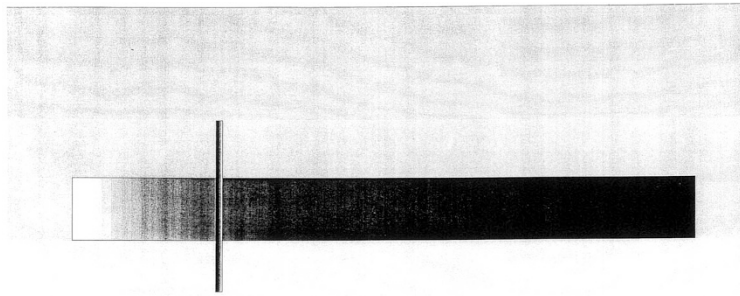
FIGURE 4. Reproducibility studies of breast cancer tumor grading are most consistent, with level of pathologist agreement yielding a gray area adjacent to a given decision threshold and with evaluations more disparate from the decision threshold becoming more certain, or "black and white."

would receive an incorrect treatment recommendation as a result of a lack of reproducibility in grading. Such an inappropriate classification would require a high-/low-grade discrepancy, which, especially if taken to the level of consensus assessment, should be extremely rare (Table 1). So, if tumor grading can assist in finding a "safe harbor" for identifying patients who meet the "90% rule of thumb," it would be for patients who have tumors that are small, low grade, and lymph-node negative. In this mammographic era, these are becoming more frequent.

The use of consensus assessment is emphasized here. Obtaining consensus does require some extra effort. However, with respect to the ease that glass slides can be passed around—or even mailed— consensus assessments should be far from onerous, and obtaining consensus is—or should be—a common practice in surgical pathology (16, 17), especially intradepartmental consensus. Although not emphasized here, disparate discrepancies also can occur (±more than 1 NSBR score), and consensus assessment is especially effective in identifying larger discrepancies. Also, there is a lesser chance of identifying those in need of additional training if consensus grading does not become a part of everyday practice, and of the quality improvement/ assurance program of any pathology group.

In the evaluation of a given case, the question then becomes, "How many opinions are required to feel comfortable that a reasonable consensus has been reached?" From prior work (4), more than 90% of NSBR score assessments were either one of two consecutive scores. This indicates that if two pathologists assess a tumor as having two adjacent scores, then the chance that a third opinion will yield a third NSBR score (a spread of three scores) is less than 10%. This is supported by Table 1, which shows that a spread across several scores is indeed

rare as shown by the rarity of low- *versus* high-grade discrepancies (granted, a high- *versus* low-grade discrepancy requires a spread across four scores). Also from prior work (4), of the six pathology groups that participated, four groups were composed of three individuals, and of the 40 separate consensus assessments by these groups (10 cases × four groups = 40 evaluations), none of the assessments was of a high–low discrepancy and only 4 (10%) resulted in either a low- *versus* intermediate-grade or an intermediate- *versus* high-grade discrepancy (actual data not in reference but obtained by additional review of data). Therefore, if two or three pathologists can achieve consensus, then the chances that additional opinions will differ significantly are not great. Although generally a consensus reached among two or three pathologists should suffice, a greater number of assessments or referral to an expert may be warranted in situations of disparate opinions or when a pathologist lacks experience in grading.

Despite efforts to obtain consensus in grading, it will remain that because of the inherent tendency for pathologist assessments in grading to cluster around adjacent NSBR scores, a subset of tumors will have score assessments that span whatever decision threshold might be set. There can be no dogmatic recommendation of how to handle such gray area cases, as variables are many with respect to the practice of medicine in various locations. However, referral to an expert in breast pathology is a possible solution as an expert may be willing to commit confidently to a single score. Or a pathology group or department may want to identify a specific member(s) in their group to become especially practiced in grading, and this person could serve as the source for the most heavily weighted assessment. This person might also be responsible

for monitoring how the grading of his or her group might compare with pathologists in other groups, and this may require occasional interlaboratory comparisons. In the course of this study, it was certainly informative for the first author to find how his grading compared with the Nottingham group. Such an exercise, as formally described (3), can be helpful.

Another approach in dealing with the overlap problem is simply to use the survival probabilities generated from the reproducibility-oriented consolidations. This approach is not to settle on a single NSBR score but to allow neighboring scores to be factored into prognosis, which would entail using Figures 2 and 3 instead of Figure 1 for survival probability information. This has some appeal in that it directly matches the capability of what pathologists can easily do with patient outcome. It is a strength of the quantitative and robust nature of the NSBR method that allows such flexibility.

The exact approach used in dealing with the reproducibility concerns addressed here probably should be decided at the local level based on what can best be communicated to and understood by the relevant clinicians who will treat patients. This can only be accomplished if pathologists are fully aware of where the level of pathologist agreement can lead to a degree of uncertainty and, perhaps more important, where certainty is achieved. Although the actual technical process of grading is independent of any given clinical decision threshold, a pathologist's awareness of possible clinical decision thresholds should increase his or her ability to give a reasonable interpretation of grading results.

That histologic grading of invasive breast cancer is a valuable prognostic feature remains the cornerstone for the need to grade, along with the ease and cost-effectiveness of grading. Grading is already being used prospectively in the clinical treatment of breast cancer patients (18). To be of more widespread value, grading must be rigorously and consistently applied, or doubts as to value will understandably persist, especially as exemplified by opinions quoting studies in which uniformity and consistency in grading were not apparent (19). The NSBR provides a platform for uniformity, and it is hoped that the discussion here will serve as a guide toward the optimal practice of NSBR grading.

The final conclusion of this exercise is not to change the NSBR method in any way but to emphasize that proper grading not only entails time at the microscope with subsequent summing of scores but also requires awareness of the level of pathologist agreement, knowledge of relevant clinical thresholds, and the value of consensus assessment. With this additional information in mind, we should feel even more confident of how to interpret our grade assessments. So, when it comes to grading invasive breast cancer, let's just do it (20, 21)!

REFERENCES

1. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. The value of histologic grade in breast cancer: experience from a large study with long term follow-up. Histopathology 1991;19:403–10.
2. Frierson HF, Wolber RA, Berean KW, Franquemont DW, Gaffey MJ, Boyd JC, et al. Interobserver reproducibility of the Nottingham modification of the Bloom and Richardson histologic grading scheme for infiltrating ductal carcinoma. Am J Clin Pathol 1995;103:195–8.
3. Robbins P, Pinder S, de Klerk, Dawkins H, Harvey J, Sterrett G, et al. Histologic grading of breast carcinomas: a study of interobserver agreement. Hum Pathol 1995;26:873–9.
4. Dalton LW, Page DL, DuPont WD. Histologic grading of breast carcinoma: a reproducibility study. Cancer 1994;73:2765–70.
5. Harvey JM, De Klerk NH, Sterrett GF. Histological grading in breast cancer: interobserver agreement, and relation to other prognostic variables including ploidy. Pathology 1992;24:63–8.
6. Theissig F, Kunze KD, Haroske G, Meyer W. Histologic grading of breast cancer: interobserver reproducibility and prognostic significance. Pathol Res Pract 1990;186:732–6.
7. Hopton DS, Thorogood J, Clayden AD, MacKinnon D. Observer variation in histologic grading of breast cancer. Eur J Surg Oncol 1988;15:21–3.
8. Davis BW, Gelber RD, Goldhirsch A, Hartmann WH, Locher GW, Reed R, et al. Prognostic significance of tumor grade in clinical trials of adjuvant therapy for breast cancer with axillary lymph node metastasis. Cancer 1986;58:2662–70.
9. Delides GS, Garas G, Georgouli G, Jiortziotis D, Lecca J, Liva T, et al. Intralaboratory variations in the grading of breast carcinoma. Arch Pathol Lab Med 1982;106:126–8.
10. Henson DE. End points and significance of reproducibility in pathology. Arch Pathol Lab Med 1989;113:830–1.
11. Lawless JF. Statistical models and methods for lifetime data. Wiley series in probability and statistics. New York: Wiley; 1982.
12. Allred DC, Harvey JM, Berardo M. Prognostic and predictive factors in breast cancer assessed by immunohistochemical analysis. Mod Pathol 1998;11:155–68.
13. Pinder SE, Murray S, Ellis IO, Trihia H, Elston CW, Gelber RD, et al. The importance of the histologic grade of invasive breast carcinoma and response to chemotherapy. Cancer 1998;83:1529–39.
14. Kronqvist P, Montironi R, Yrjo UIC, Bartels PH, Thompson D. Management of uncertainty in breast cancer grading with bayesian belief networks. Analyt Quant Cytol Histol 1995;17:300–8.
15. Todd JH, Dowle C, Williams MR, Elston CW, Ellis IO, Hinton CP, et al. Confirmation of a prognostic index in primary breast cancer. Br J Cancer 1987;56:489–92.
16. Elston CW, Ellis IO. Assessment of histologic grade. In: Elston CW, Ellis IO, editors. Systemic pathology: the breast. 3rd ed., Vol. 13. Edinburgh: Churchill Livingston; 1998. p. 365–84.
17. Leslie KO, Fechner RE, Kempson RL. Second opinions in surgical pathology. Am J Clin Pathol 1996;106(Suppl 1):S58–64.
18. Blamey RW. Clinical aspects of malignant breast lesions. In: Elston CW, Ellis IO, editors. Systemic pathology: the breast. 3rd ed., Vol. 13. Edinburgh: Churchill Livingston; 1998. p. 501–14.
19. Burke HB, Henson DE. Histologic grade as a prognostic factor in breast carcinoma. Cancer 1997;80:1703–4.
20. Page DL, Ellis IO, Elston CW. Histologic grading of breast cancer: let's do it. Am J Clin Pathol 1995;103:123–4.
21. Roberti NE. The role of histologic grading in the prognosis of patients with carcinoma of the breast: is this a neglected opportunity? Cancer 1997; 80:1708–16.