C-erbB-2 immunostaining: Problems with interpretation

E W Kay, C J Barry Walsh, M Cassidy, B Curran, M Leader

Abstract

Aims—To assess the consistency and reproducibility of assessment of c-erbB-2 immunostaining, and to examine some of the problems relating to inter- and intraobserver variability in the documentation of positive staining; to profile the spectrum of cytoplasmic and membranous staining in a wide range of tumour types.

Methods—A total of 283 neoplasms were examined for immunohistochemical expression of the c-erbB-2 oncoprotein. Three independent observers were required to assess intensity both of membrane and cytoplasmic staining on a three point and then a four point scale. Extent of positive staining was also assessed on a two point scale. A minimum of two weeks elapsed between assessments using the differing scales.

Results—Positive membrane staining was documented by one or more observers in 16-6% of tumours examined. This positivity was largely restricted to bladder, renal, and breast tumours. The overall level of disagreement as to the presence or absence of membranous staining was 11-3%. Cytoplasmic staining was identified in 55-5% of tumours studied. The level of disagreement as to the presence or absence of cytoplasmic staining was 26-5%.

Conclusions—Intraobserver variability was minimal, indicating that each pathologist was adhering to internal reproducible standards. Interobserver variability was greater, indicating that the interpretation of c-erbB-2 immunostaining may require set guidelines. It is suggested that assessment should be referenced to a standard positive control, that a three tier system for grading of intensity and a two tier system for grading of extent should be adopted, and that the evaluation should be agreed by at least two pathologists. The presence of cytoplasmic staining should continue to be routinely recorded until its biological role and clinical implications are fully understood.

(J Clin Pathol 1994;47:816–822)

The c-erbB-2 proto-oncogene is thought to have a role in normal cell proliferation and development. The gene is located on chromosome 17 (q21), encodes a 4-6 kilobase mRNA, and produces a 185 kilodalton glycoprotein. This protein product is a transmembrane receptor with tyrosine kinase activity and shows 78% homology with the intracytoplasmic domain of the epidermal growth factor receptor.1-11

Amplification and overexpression of the proto-oncogene c-erbB-2 (HER2/neu) has been reported in a wide variety of tumour types, predominantly of epithelial origin, including female breast,1-14 ovary,15 transitional cell carcinoma (TCC),16-20 pancreas,21 stomach,22-23 kidney,24-25 and salivary gland.16-27 In contrast, it is rarely described in other epithelial tissues26 or mesenchymal neoplasms, such as cartilage.28

Despite this impressive list of human tumours studied, the clinical relevance of some of the results in terms of prognosis is uncertain30-32 and institutional results, in many instances, vary. McCann et al found positive membranous staining in 2% of the bladder carcinomas studied,33 while Coombs et al found a 14% incidence of positive staining.34 McCann et al showed positive staining in 17% of breast carcinomas, while a 30% positivity rate was reported by Slamon et al in their study of breast carcinomas. In this department a small proportion of c-erbB-2 staining in breast carcinoma showed some positivity in 26% of cases (unpublished observations).

How should these variations in results be interpreted? In this department three large multivariate studies are in train and in each study the prognostic value of c-erbB-2 immunoreactivity is being assessed. However, in the course of these studies each investigator encountered problems with the interpretation of the staining reaction.

This prompted the current study in which we sought to examine some of the problems relating to inter- and intraobserver variability in the documentation of positive staining. We also profiled the spectrum of cytoplasmic and membranous staining in a wide range of tumours, some of which had never been examined before, for the presence of c-erbB-2 oncogene expression. From our results we can draw some conclusions which may help to explain the occurrence and relevance of positive results published by different institutions.

Methods

A total of 283 neoplasms were examined for immunohistochemical expression of the
c-erbB-2 oncoprotein. The monoclonal antibody NCL-CB11 (Novocastra) was used except for the breast carcinomas, where a polyclonal antibody pAb1 (Triton Biosciences) was used instead. Specimens were drawn from six separate institutions and all tissue had been routinely formalin fixed before embedding in paraffin wax.

The study included a total of 223 carcinomas or malignant epithelial lesions. These comprised 37 breast carcinomas, 40 colorectal adenocarcinomas, 37 transitional cell carcinomas (TCC) of bladder, 39 renal cell carcinomas, 20 squamous carcinomas of head and neck, 40 cervical intraepithelial neoplasias (grade 3) and 10 basal cell carcinomas. In addition, 40 non-epithelial malignancies were examined. These comprised 10 lymphomas (five Hodgkin’s and five non-Hodgkin’s), 10 gliomas, 20 soft tissue sarcomas (five each of leiomyosarcomas, liposarcomas, synovial sarcomas and clear cell sarcomas). Ten cystosarcoma phyllodes of the breast were also examined.

Sections were dewaxed and endogenous peroxidase blocked by 3% hydrogen peroxide in methanol for 30 minutes at room temperature. The sections were then incubated overnight with the primary antibody NCL-CB11 at a concentration of 1 in 20, followed by exposure to biotinylated rabbit anti-mouse immunoglobulin (Dakopatts UK) at a concentration of 1 in 200 for 40 minutes at room temperature. After rinsing with TBS (TRIS-buffered saline) avidin-biotin complex/horseradish peroxidase (Dakopatts UK) was applied for 30 minutes and colour developed using a standard dianamobenzidine technique.

A known positive breast carcinoma was included as a positive control. As a negative control the same procedure was followed but with omission of the primary antibody. Both staining intensity and proportion of positive cells were graded (table 1). A similar procedure was followed for the breast tumours, but the antibody applied was pAb1 (Triton Biosciences). Secondary antibody swine antirabbit immunoglobulin was applied.

In each case both membrane and cytoplasmic staining were assessed by three experienced histopathologists with an average of nine years’ experience each in histopathology. In keeping with other published methods of assessment, each observer recorded staining intensity for both patterns of staining on a three point (0 = no staining; + = weak staining; ++ = strong staining) and then a four point (0 = no staining; + = weak staining; ++ = moderate staining; +++ = strong staining) scale. Extent of staining was also assessed on a two point scale (1 = < 50% positive tumour cells; 2 = > 50% positive tumour cells). For each tumour type a minimum of two weeks elapse between assessments of the differing scales of intensity in an attempt to render the results of a single observer as independent as possible. All assessments were then compared and contrasted to establish both intraobserver and interobserver reproducibility.

The $\chi^2$ statistic was used to measure levels of agreement between observers. In addition, the $\chi^2$ test for symmetry was applied to evaluate the level of systematic error between observers, as described by Lee et al.34

### Results

The findings are tabulated in table 1. Agreed positive membrane staining was identified in 15 (5.3%) of the tumours examined. With the exception of positive staining in the epithelial component of a single phyllodes tumour, these were all carcinomas of breast, bladder, or kidney. In a further 32 (11.3%) tumours at least one observer disagreed as to the presence of membrane staining. Most of these tumours (28/32) were also carcinomas of breast, bladder, and kidney. This group also included three examples of CIN (grade 3) and a single glioma.

Agreement as to the presence of cytoplasmic staining was also assessed for each tumour type.

### Table 1 Membranous and cytoplasmic staining with c-erbB-2 in all tumours studied

<table>
<thead>
<tr>
<th>Tumour type</th>
<th>Membrane positivity</th>
<th>Membrane intensity</th>
<th>Membrane extent</th>
<th>Cytoplasmic positivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Disagree</td>
<td>Agree</td>
</tr>
<tr>
<td>Synovial sarcoma (n=5)</td>
<td>0</td>
<td>5 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Leiomyosarcoma (n=5)</td>
<td>0</td>
<td>5 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Liposarcoma (n=5)</td>
<td>0</td>
<td>5 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Clear cell sarcoma (n=5)</td>
<td>0</td>
<td>5 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Glioma (n=10)</td>
<td>0</td>
<td>9 (90%)</td>
<td>1 (10%)</td>
<td>3 (30%)</td>
</tr>
<tr>
<td>Lymphoma (n=10)</td>
<td>0</td>
<td>10 (100%)</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Basal cell carcinoma (n=10)</td>
<td>0</td>
<td>10 (100%)</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Squamous cell carcinoma (n=20)</td>
<td>0</td>
<td>20 (100%)</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Colonic adenocarcinoma (n=40)</td>
<td>0</td>
<td>40 (100%)</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Cervical intraepithelial neoplasia (n=40)</td>
<td>0</td>
<td>37 (92.5%)</td>
<td>3 (7.5%)</td>
<td>2 (5%)</td>
</tr>
<tr>
<td>Renal cell carcinoma (n=39)</td>
<td>3 (7.7%)</td>
<td>26 (66.7%)</td>
<td>10 (25.6%)</td>
<td>2 (67%)</td>
</tr>
<tr>
<td>Transitional cell carcinoma (n=37)</td>
<td>8 (21.6%)</td>
<td>15 (40.6%)</td>
<td>14 (37.8%)</td>
<td>4 (50%)</td>
</tr>
<tr>
<td>Breast carcinoma (n=37)</td>
<td>3 (8%)</td>
<td>30 (81%)</td>
<td>4 (11%)</td>
<td>2 (67%)</td>
</tr>
<tr>
<td>Phyllodes (epithelial) (n=10)</td>
<td>1 (10%)</td>
<td>9 (90%)</td>
<td>0</td>
<td>1 (100%)</td>
</tr>
<tr>
<td>Phyllodes (stromal) (n=10)</td>
<td>0</td>
<td>10 (100%)</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Total (n=283)</td>
<td>15 (5.3%)</td>
<td>236 (83.4%)</td>
<td>32 (11.3%)</td>
<td>9 (60%)</td>
</tr>
</tbody>
</table>

J Clin Pathol. First published as 10.1136/jcp.47.9.816 on 1 September 1994.
staining for c-erbB-2 was reached in 82 (29%)
tumours. Positively staining tumours were
identified in all categories of neoplasm,
excluded lymphoma and three of the four
categories of sarcoma. In
the clear cell sarcoma group all three
observers identified cytoplasmic positivity
in two of the five cases examined. In a further 75
(26.5%) cases at least one observer identified
cytoplasmic positivity. The only tumour
where complete agreement as to the
absence of cytoplasmic c-erbB-2 positivity was
reached were the leiomyosarcomas. Some of
the positively staining tumour types and the
problems encountered are illustrated in the
figure.

In those tumours where all three observers
agreed as to the presence of membranous
staining there was considerable disagreement
both as to its intensity and extent. Using a
three point scale, agreement on staining intensity
was reached in nine of 15 cases. Use of a
four point scale reduced the level of agreement
to seven of 15 cases. In only a minority of
cases (6/15) was complete agreement reached
as to the extent of positive staining. When
intensity and extent of staining were evaluated
together, complete agreement was reached in
only four of the 15 cases.

In a total of 32 cases membranous positivity
was identified by two (10 cases) observers
or by one (22 cases) observer. In the former
group there was a better level of agreement on
intensity (8/10) and extent (8/10).

Analysis of the results for cytoplasmic posi-
tivity gave similar findings, but with higher
levels of disagreement as to the extent and
intensity of staining.

The detailed analysis of these results is
given in table 2, showing the \( \kappa \) statistics for all
pairs of raters with 95% confidence limits.
The \( \chi^2 \) test for frequency symmetry is also
shown. Given that a \( \kappa \) of 0.7 to 0.75 is
considered to indicate adequate agreement,
this level was achieved in only three groups,
all of which involved assessment of membrane
intensity and all of which involved observers
designated B and C. The 95% confidence
intervals for these \( \kappa \) statistics are also shown.

The \( \chi^2 \) test for symmetry reached signifi-
cance (\( p \leq 0.05 \)) in most cases, suggesting
that there was individual rater bias rather than
random disagreement.

Given that the measured value for many of
the tumours was zero (a fact that may distort
the total figures), and working on the premise
that it is probably more important to discrimi-
nate consistently and reliably between a posi-
tive value and a negative value, further
statistical analysis was applied to the results
recorded as simply positive or negative.
Agreement between observers was assessed
using Cohen’s \( \kappa \) modified for use with ordi-
nal rating scales. Again, a \( \kappa \) of \( \geq 0.7 \) is taken as
indicating satisfactory agreement and \( \kappa \) values
whose 95% confidence intervals do not
include 0.7 were deemed to indicate poor
agreement. The results (table 3) showed simi-
lar levels of disagreement and again suggest
individual rater bias.
C-erbB-2 immunostaining: problems with interpretation

(D) Agreed strong positive membranous staining for c-erbB-2 in transitional cell carcinoma.
(E) Agreed weak membranous staining with c-erbB-2 in renal cell carcinomas. In comparison with (D) the presence of membranous positivity is questionable.
(F) Agreed strong cytoplasmic positivity for c-erbB-2 in renal cell carcinoma.
(G) This illustrates the difficulty of interpreting the precise staining pattern against a background of strong cytoplasmic reactivity with c-erbB-2. All three observers recorded strong cytoplasmic positivity present in this case, but two of three observers also scored this case as showing strong membranous staining. The third observer discounted the membranous pattern interpreting it instead as cytoplasmic condensation at the membrane (renal cell carcinoma).
(H) Two of three observers recorded strong membranous positivity with c-erbB-2 in this case. The third observer gave this a negative score. This was the only positive area and represented one high power field of the tumour (renal cell carcinoma).

Discussion

The primary aim of this paper was to address the universal problem of histochemical markers—that is inter- and intraobserver variation and the requirement for quality standards. The need for standardisation is heightened by the potential use of markers such as c-erbB-2 as prognostic variables in clinical practice.

The results of this study show that intraobserver variability was minimal, indicating that each pathologist was adhering to internal standards which were consistently reproducible. However, there was a high rate of interobserver variability. With respect to
Table 2  Statistical analysis of results shown for all tumours and for carcinomas in which most positive staining was found

<table>
<thead>
<tr>
<th>Observers</th>
<th>Attribute rated</th>
<th>PObs</th>
<th>PExp</th>
<th>Kappa</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All tumours</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A&amp;B Membrane intensity (4)</td>
<td>0.950</td>
<td>0.873</td>
<td>0.609</td>
<td>0.513 to 0.705</td>
<td>16</td>
</tr>
<tr>
<td>A&amp;C Membrane intensity (3)</td>
<td>0.944</td>
<td>0.847</td>
<td>0.617</td>
<td>0.510 to 0.715</td>
<td>10</td>
</tr>
<tr>
<td>B&amp;C Membrane intensity (3)</td>
<td>0.953</td>
<td>0.860</td>
<td>0.623</td>
<td>0.521 to 0.721</td>
<td>21</td>
</tr>
<tr>
<td>A&amp;B Membrane intensity (3)</td>
<td>0.962</td>
<td>0.877</td>
<td>0.717</td>
<td>0.615 to 0.818</td>
<td>6.6</td>
</tr>
<tr>
<td>A&amp;C Cytoplasm intensity (4)</td>
<td>0.978</td>
<td>0.844</td>
<td>0.660</td>
<td>0.572 to 0.746</td>
<td>49</td>
</tr>
<tr>
<td>B&amp;C Cytoplasm intensity (4)</td>
<td>0.947</td>
<td>0.866</td>
<td>0.541</td>
<td>0.450 to 0.634</td>
<td>12.4</td>
</tr>
<tr>
<td>A&amp;B Cytoplasm intensity (3)</td>
<td>0.958</td>
<td>0.841</td>
<td>0.637</td>
<td>0.544 to 0.726</td>
<td>17.7</td>
</tr>
<tr>
<td>B&amp;C Cytoplasm intensity (3)</td>
<td>0.960</td>
<td>0.847</td>
<td>0.630</td>
<td>0.541 to 0.717</td>
<td>20.5</td>
</tr>
<tr>
<td>A&amp;B Membrane extent (3)</td>
<td>0.938</td>
<td>0.860</td>
<td>0.562</td>
<td>0.464 to 0.661</td>
<td>8.8</td>
</tr>
<tr>
<td>B&amp;C Membrane extent (3)</td>
<td>0.907</td>
<td>0.855</td>
<td>0.563</td>
<td>0.471 to 0.645</td>
<td>15.7</td>
</tr>
<tr>
<td>A&amp;B Cytoplasm extent (3)</td>
<td>0.959</td>
<td>0.894</td>
<td>0.614</td>
<td>0.512 to 0.717</td>
<td>27</td>
</tr>
<tr>
<td>B&amp;C Cytoplasm extent (3)</td>
<td>0.813</td>
<td>0.676</td>
<td>0.424</td>
<td>0.327 to 0.521</td>
<td>9.6</td>
</tr>
</tbody>
</table>

Table 3  Kappa coefficients for each pair of observers recording staining as simple positive or negative

<table>
<thead>
<tr>
<th>Observers</th>
<th>Attribute rated</th>
<th>PObs</th>
<th>PExp</th>
<th>Kappa</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All tumours</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A&amp;B Membrane intensity (4)</td>
<td>0.926</td>
<td>0.797</td>
<td>0.635</td>
<td>0.582 to 0.784</td>
<td></td>
</tr>
<tr>
<td>A&amp;C Membrane intensity (4)</td>
<td>0.900</td>
<td>0.811</td>
<td>0.472</td>
<td>0.360 to 0.583</td>
<td></td>
</tr>
<tr>
<td>B&amp;C Membrane intensity (4)</td>
<td>0.957</td>
<td>0.874</td>
<td>0.754</td>
<td>0.637 to 0.871</td>
<td></td>
</tr>
<tr>
<td>A&amp;B Membrane intensity (3)</td>
<td>0.921</td>
<td>0.796</td>
<td>0.613</td>
<td>0.498 to 0.727</td>
<td></td>
</tr>
<tr>
<td>A&amp;C Membrane intensity (3)</td>
<td>0.903</td>
<td>0.808</td>
<td>0.496</td>
<td>0.384 to 0.609</td>
<td></td>
</tr>
<tr>
<td>B&amp;C Membrane intensity (3)</td>
<td>0.955</td>
<td>0.834</td>
<td>0.731</td>
<td>0.617 to 0.844</td>
<td></td>
</tr>
<tr>
<td>A&amp;B Cytoplasm intensity (4)</td>
<td>0.829</td>
<td>0.493</td>
<td>0.662</td>
<td>0.550 to 0.773</td>
<td></td>
</tr>
<tr>
<td>A&amp;C Cytoplasm intensity (4)</td>
<td>0.788</td>
<td>0.518</td>
<td>0.560</td>
<td>0.445 to 0.676</td>
<td></td>
</tr>
<tr>
<td>B&amp;C Cytoplasm intensity (4)</td>
<td>0.824</td>
<td>0.497</td>
<td>0.650</td>
<td>0.536 to 0.764</td>
<td></td>
</tr>
<tr>
<td>A&amp;B Cytoplasm intensity (3)</td>
<td>0.837</td>
<td>0.498</td>
<td>0.675</td>
<td>0.564 to 0.786</td>
<td></td>
</tr>
<tr>
<td>A&amp;C Cytoplasm intensity (3)</td>
<td>0.836</td>
<td>0.497</td>
<td>0.674</td>
<td>0.563 to 0.779</td>
<td></td>
</tr>
<tr>
<td>B&amp;C Cytoplasm intensity (3)</td>
<td>0.872</td>
<td>0.600</td>
<td>0.680</td>
<td>0.587 to 0.772</td>
<td></td>
</tr>
</tbody>
</table>

Comparison of our findings with those of several other published studies illustrates the problem of such variability. A result of 21-6% positive staining for transitional cell carcinoma is broadly in line with the range of 14-30% quoted.23-47 In general, the tumour groups in which agreed positive membrane staining was identified were in keeping with other published results on c-erbB-2. However, our figure of 8% positivity in breast carcinomas is lower than that reported by other authors.5 Reproducibility between our findings and those reported by others may have been related to interobserver variation. Factors such as tissue fixation and antibodies may also have an important role.13 19-22 27 33 40 43-57 Our ability accurately to compare results is hindered by the lack of detail recorded in terms of staining type, intensity, extent and location. In those tumour types where our results are at variance with those reported, the possibility that the findings may be accentuated in our own laboratory is raised. The way to explore this would be to repeat the study using tumours stained in other laboratories. However, as many of the other tumour groups demonstrated a level of positivity similar to those of other published studies, the likeliness of our methodology being solely at fault is not very high.

The issue of the presence and importance of cytoplasmic immunoreactivity with c-erbB-2 as opposed to membranous staining, is, as yet, undetermined. Many investigators have chosen to ignore the cytoplasmic component present28 48 58 but a few have tentatively explored its possible association with membrane staining29 44 51 56 and clinical outcome.49 There are several putative explanations for the

membrane staining—for example, overall agreement as to its presence was 90%, but there was a disagreement rate of 40% with regard to intensity in those cases with agreed positive membrane staining. Increasing the complexity of the grading system did not improve the level of interobserver agreement. There was also a high level of interobserver variability in the assessment of whether cytoplasmic staining was positive or negative. The K2 statistics and y2 tests for symmetry suggest that a significant proportion of the variability is due to responses set by individual observers rather than that the picture is vague. The observers were demonstrating a difference in approach, despite the fact that before embarking on this study there was an assumed uniformity of approach.

The findings may provide some explanation for the variability in published results regarding the presence and clinical relevance of c-erbB-2 membrane staining in different tumour groups, a subject of extensive study in a comprehensive range of malignancies.19 21-23 26 27 29 30 33 36-46
C-erbB-2 immunostaining: problems with interpretation

cytoplasmic location of staining, including suggestions that it represents an intermediate or alternative protein product. This uncertain background leads us to suggest that its presence warrants continued documentation. However, the level of interobserver variation revealed in this study introduces a note of caution in drawing meaningful conclusions.

Although our work has highlighted some limitations in the interpretation of c-erbB-2, its future value lies in its usefulness as a prognostic indicator, a feature which we have not addressed in this study. However, differences in interpretation of staining by different observers will lead to variations in emphasis on its prognostic value. The prognostic value of c-erbB-2 IHC in breast carcinoma is influenced by interpretation of staining. Such major differences in interobserver interpretation as outlined in this study suggest that the role of c-erbB-2 oncoprotein as a prognostic indicator should be more carefully evaluated.

In conclusion, the findings of this study suggest that interpretation of c-erbB-2 immunostaining requires set guidelines. The following are suggested:

1. Assessment of intensity of positive membrane staining should be referenced to a standard positive control, such as breast carcinoma stained at the time and under the same conditions as the test case.

2. Use of three tier system for grading of intensity: 0 = negative, 1+ = weak, 2+ = strong.

In addition, the evaluation of c-erbB-2 might be improved by assessment by more than one pathologist.

The presence of cytoplasmic staining is, as yet, of uncertain prognostic, morphological, or biological importance, but we feel that this staining pattern should be routinely recorded in all tumour types. However, six criteria to those suggested for evaluation of membrane staining may perhaps need to be instituted.

Much remains to be elucidated regarding the role of c-erbB-2 in tumour biology and its value as a prognostic indicator seems to be restricted to a few tumour types. Although speculative, perhaps its greatest application will be as one of a panel of oncogene markers as is the situation with other immunohistochemical agents.

We thank Ms Ronán Conroy, Department of Statistics, RCSI, and Ms Saundra Dalton for typing the manuscript.


822


35 Cohen J. Weighted kappa: nominal scale agreement with provision for scale disagreement or partial credit. *Psychol Bull* 1968;70:213-20.


