

# Immunohistochemical demonstration of oestrogen and progesterone receptors: correlation of standards achieved on in house tumours with that achieved on external quality assessment material in over 150 laboratories from 26 countries

A Rhodes, B Jasani, A J Balaton, K D Miller

## Abstract

**Aims**—To investigate the sensitivity of immunohistochemical (IHC) assays for oestrogen receptors (ER) and progesterone receptors (PR) achieved by laboratories on breast tumours fixed and processed in their own department, and to compare this with the degree of sensitivity they achieve on tumours circulated as part of an external quality assessment (EQA) programme.

**Methods**—On 10 occasions between April 1994 and June 1998, histological sections from breast cancers showing various degrees of expression of ER and PR were circulated for IHC staining to laboratories participating in the UK national external quality assessment scheme for immunocytochemistry (UK NEQAS-ICC). The staining of these tumours, in addition to that of tumours fixed and processed in the participants own laboratories (in house tumours), was assessed by a panel of four assessors, using the established UK NEQAS-ICC scoring system. For a selected assessment run, the degree of expression of participants in house tumours was evaluated by means of the semiquantitative quick score method.

**Results**—Although the scores awarded for the staining of in house tumours were generally higher than those awarded for the staining of UK NEQAS tumours, there was also a significant positive correlation between the two sets of scores. Using the quick score method of evaluation for one of the assessment runs, 47% of in house tumours were classified as having a high degree of ER expression. Of the remaining cases, a significant proportion initially classified as having only low or medium expression of ER were found to have higher expression when stained by the organising laboratory. The UK NEQAS-ICC centre's routine assay for hormonal receptors was found to be 90–100% efficient in achieving optimal demonstration of breast tumours from over 150 different laboratories.

**Conclusions**—The significant positive correlation between the results obtained on the UK NEQAS tumours and the in house tumours provides evidence for the view that results achieved on EQA mate-

rial are accurate indicators of in house laboratory performance. Although most laboratories adequately detected tumours with high receptor expression, a large proportion of in house tumours classified initially by participants' staining as being of low or medium ER expression had a higher degree of expression when stained by the UK NEQAS-ICC centre. The efficiency of the organising centre's routine IHC method for ER and PR in optimally demonstrating participants in house breast tumours shows that variations in fixation and tissue preparation are not limiting factors preventing a different laboratory achieving optimal demonstration.

(J Clin Pathol 2000;53:292-301)

Keywords: immunohistochemistry; oestrogen receptors; progesterone receptors; external quality assessment

Recent leading articles have emphasised the importance of establishing the oestrogen receptor (ER) status of women with breast cancer.<sup>1,2</sup> Other articles have reported on the degree of variability that exists between laboratories when demonstrating ER by immunohistochemistry (IHC) on the same cases.<sup>3,4</sup> The largest of these studies looked at the results obtained by 200 different participants of an external quality assessment (EQA) programme on slides circulated to these laboratories by the EQA scheme and containing tumours with differing degrees of ER expression.<sup>3</sup>

Although many view the results of EQA as a useful gauge of a laboratory's ability to perform adequate staining for ER and progesterone receptors (PR) on paraffin wax embedded sections,<sup>5,6</sup> it could be argued that the system suffers from a number of drawbacks. Distributed material is limited and consists of tissue that has been fixed, processed, and prepared under different conditions, no matter how slight, to those used by the participating laboratory. This is thought to be important because there is a view that the IHC assay optimised for use on in house material cannot be expected to produce results of the same quality on tissues fixed and processed in a different laboratory. Consequently, it is thought by some that the quality of IHC achieved on material distributed by an EQA scheme does not reflect

Department of  
Histopathology, UCL  
Medical School,  
University Street,  
London WC1E 6JJ, UK  
A Rhodes

Department of  
Histopathology, UCL  
Medical School  
K D Miller

Department of  
Pathology, University  
of Wales College of  
Medicine, Cardiff  
CF4 4XN, UK  
B Jasani

Centre de Pathologie,  
20 Avenue de la Gare,  
91570 Bievres, France  
A J Balaton

Correspondence to:  
Mr Rhodes  
email: rmkdher@ucl.ac.uk

Accepted for publication  
20 October 1999

Table 1 Details of tumours circulated by UK NEQAS-CC for assessments between April 1994 and June 1998

Run no.	Date	Number of labs	Tumour type(s)	Receptor	Degree of expression	
					IHC	Cytosol
26	4/94	37	IDC	ER	100%, +++	Not performed
29	2/95	66	IDC (mets)	ER	100%, +++	Not performed
32*	12/95	98	Inf./DCIS	ER	100%, +++	Not performed
			ILC		80-90% ++	Not performed
34*	8/96	105	IDC	ER	90-100%, +++	ER = 248
			ILC		75-80% , ++	ER = 29
			IDC		0%	ER = 0
36	1/97	118	IDC	PR/ER†	90%, +++	ER = 66, PR = 307
38*	4/97	175	IDC	ER	90-100%, +++	ER = 122
			IDC		90%, ++	ER = 41
			IDC		0%	ER = 7
39*	8/97	176	IDC	PR/ER†	99-100%, +++	ER = 309, PR = 1582
			ILC		75%, ++	ER = 11, PR = 21
			IDC		1%, +	ER = 7, PR = 6
40*	12/97	192	IDC	ER	99%, +++	ER=15
			IDC		99%, +++	ER=11
			IDC		90%, ++	ER = 12
41	4/98	178	IDC	ER and PR‡	ER = 75-80%, ++	ER = 10
					PR = 0%	PR = 2
42*	6/98	205	IDC	ER	100%, +++	ER = 70
			IDC		90%, ++	ER = 41
			IDC		15%, +	ER = 65

All tumours were fixed for 24 hours in 10% neutral buffered formalin.

ER/PR expression of tumours by IHC is described in terms of the proportion of invasive nuclei staining (%) and the average staining intensity (+, ++, +++).

Cytosol values of expression are in fmol/mg protein.

\*Composite blocks containing more than one tumour were used for these assessments.

†Participants not stocking an antibody to PR were asked to demonstrate ER instead. On these runs tumours were chosen that have very similar expression of ER and PR, as determined by IHC at the organising laboratory.

‡On this assessment, participants were asked to demonstrate both ER and PR.

ER, oestrogen receptor; IDC, infiltrating ductal carcinoma; IHC, immunohistochemistry; Inf., infiltrating; mets, metastatic; PR, progesterone receptor.

the standards achieved by a laboratory on tumours fixed and processed in house, which might account for a large percentage of its work load. Problems of this type are encountered with most EQA programmes that circulate material for analysis, and have yet to be resolved completely.<sup>7-11</sup>

To investigate these limitations, our study evaluates the immunostaining of the UK NEQAS-ICC organising laboratory and tests its validity as a reference standard for ER and PR. It then investigates and compares the performance achieved by laboratories on in house tumours to that achieved by the same laboratories on the tumours circulated by the scheme.

## Materials and methods

### TUMOURS CIRCULATED BY THE EQA SCHEME

Laboratories participating in the UK NEQAS-ICC programme for steroid hormonal receptors were sent, at each assessment, two unstained slides containing histological tissue sections of formalin fixed and paraffin wax processed breast tumours showing different degrees of hormonal expression. Each participant was asked to demonstrate ER and/or PR and to return the stained slide(s) to the UK NEQAS coordinating centre for assessment of staining quality. Table 1 shows details of the tumours circulated at each assessment from April 1994 to June 1998. Although tested in the organisers laboratory, most of these tumours were fixed and processed in the laboratories of participants, from where they were kindly donated. Whenever possible, tumours that also had their receptor status determined

Table 2 Details, as given on the participants returning questionnaire, of the 152 in house breast carcinomas stained and submitted for assessment (run 41)

Type of tumour	n	%
Adenocarcinoma	5	3.3%
Breast carcinoma (unspecified)	66	43.4%
Ductal carcinoma (grade unspecified)	47	30.9%
Infiltrating ductal carcinoma (grade I)	4	2.6%
Infiltrating ductal carcinoma (grade II)	13	8.6%
Infiltrating ductal carcinoma (grade III)	5	3.3%
Colloid breast carcinoma	1	0.7%
Cribiform	1	0.7%
Infiltrating lobular carcinoma	4	2.6%
Intracystic papillary	1	0.7%
Metastatic disease	4	2.6%
Mucinous	1	0.7%
Tubular carcinoma	2	1.3%
Total	152	100%

biochemically by the ligand binding assay (LBA) were used.

### IN HOUSE TUMOURS

In parallel with the immunostaining of slides circulated by the scheme, participants were asked to demonstrate the same receptor on their own in house tumour, and to submit this stained slide for assessment of staining quality, along with two additional unstained tissue sections from the same tissue block. Tables 2 and 3 give examples of the types of tumours submitted by participants and the various fixatives used. These unstained slides were then stained by the UK NEQAS organising laboratory along with its routine workload, utilising its routine methodology and reagents, for the demonstration of hormonal receptors (table 4). They were then coded and filed alongside the participants own immunostaining before routine assessment. The UK NEQAS-ICC organising centre also stained the participants' in house tumours so that a comparison could be

Table 3 Details of fixation for the 152 in house breast carcinomas submitted for run 41

Type of fixative	Duration	Temperature	n	%
10% Formal saline	18–48 hours	Ambient	80	52.6%
10% Formal saline	72 hours	Ambient	2	1.3%
10% Formal saline	24 hours	37°C	2	1.3%
10% Neutral buffered formalin	24–48 hours	Ambient	44	28.9%
10% Neutral buffered formalin	72 hours	Ambient	1	0.7%
10% Neutral buffered formalin	5 hours	Ambient	1	0.7%
Acidic formal alcohol	24–48 hours	Ambient	3	2.0%
Acidic formal alcohol	10 hours	25°C	1	0.7%
Acidic formal alcohol	48 hours	Ambient	1	0.7%
Bouin's fluid	12–24 hours	Ambient	2	1.3%
Bouin's Hollande	Not given	Ambient	1	0.7%
Bouin Dubosq-Brasil	6–24 hours	Ambient	3	2.0%
Carson's fluid	24 hours	Ambient	1	0.7%
Formal calcium	24 hours	Ambient	1	0.7%
7.5% Formalin	24 hours	Ambient	1	0.7%
Information not given			8	5.3%
Total			152	100%

Unless stated otherwise, it was assumed fixation was performed at room temperature (ambient).

made between this staining and that produced by the participant on the same case.

#### ASSESSMENT OF SLIDES

An expert panel of four, comprising consultant pathologists and biomedical and clinical scientists, assessed the quality of the IHC independently on a single blind basis, with each assessor awarding marks out of 5 for each of the coded slides. The four individual marks were then added together to give a total mark out of 20. Marks were awarded by comparing the proportion and intensity of tumour nuclei staining in the participant's slide, to that achieved on duplicate sections of the same cases by the UK NEQAS-ICC organising centre. A total mark > 12 out of 20 indicates acceptable immunostaining and a pass at assessment, a mark of 10–12 out of 20 is considered to be suboptimal and borderline, whereas a total mark < 10 out of 20 is given for staining that is of unacceptable quality and represents a failure at assessment. One of the main criteria by which staining is deemed unacceptable is when < 10% of receptor positive tumour nuclei are clearly demonstrated in a tumour that has been shown by the UK NEQAS-ICC organising centre to express > 10% ER or PR positive nuclei. To ensure assessor concordance, within  $\pm 1$  mark, the slides were marked in batches of 20. On the marking of the 20th slide, all scores were read out, when there was a difference of greater than 1 mark between any of the assessors' individual marks, the respective slide

was reviewed until a consensus was reached—that is, all the assessors gave the same mark, within  $\pm 1$  mark.

For a selected assessment (run 41), the participants in house staining and that achieved by the organising centre on sections from the same tumours was assessed using the semi-quantitative quick score method of evaluation.<sup>12–14</sup> With this method, the intensity of the immunohistochemical reaction as viewed under the light microscope was recorded as either: 0, negative (no staining of any nuclei even at high magnification); 1, weak (only visible at high magnification); 2, moderate (readily visible at low magnification); or 3, strong (strikingly positive even at low power magnification). The proportion of tumour nuclei showing positive staining was also recorded as either: zero (0), approximately 1–25% (1), 26–50% (2), 51–75% (3), or 76–100% (4). The score for intensity was then added to the score for proportion, giving the quick score with a range of 0–7. In the case of composite blocks (n = 17), the tumour showing the lowest amount of expression was assessed using this method, and where participants had submitted normal breast tissue, these were not submitted for quick score evaluation. Before the assessors' (AR, BJ) evaluation of these slides, all were randomised using numbers generated by a Microsoft Excel program. Determination of the degree of concordance between the quick scores of the participants' immunostaining of their own in house tumour and the quick scores of the UK NEQAS laboratory's immunostaining of the same tumour was established using Cohen's  $\kappa$  coefficient. The proportion of cases that showed the same or a higher degree of expression when stained by the organising laboratory was analysed by means of the  $\chi^2$  test, as was the proportion of cases that showed the same or a higher degree of expression when stained by the participant's laboratory. The degree of assessor concordance when using the quick score method to evaluate stained slides was measured using Goodman and Kruskal's  $\gamma$  statistic.

#### VALIDATION OF ASSESSMENT THRESHOLDS AND OPTIMAL SENSITIVITY

Of the participating laboratories, six were identified as having published studies clinically

Table 4 The main technical parameters used by participants and the organising laboratory of UK NEQAS-ICC in the immunohistochemical demonstration of oestrogen receptors (ER) and progesterone receptors (PR)

	Oestrogen receptor: run 41			Progesterone receptor: run 39		
	Ag retrieval	1° Ab clone	Detection	Ag retrieval	1° Ab clone	Detection
Clinically validated labs (6)	MW (4) PC (2)	1D5 (5) 6F11 (1)	LSAB (1) St ABC (3) ABC (2)	MW (3) PC (2) None (1)	1A6 (3) KD68 (3)	LSAB (1) St ABC (4) ABC (1)
All UK NEQAS-ICC participating labs	MW (96) PC (60) PC in MW (18) Other (4)	1D5 (131) 6F11 (41) Other (6)	St ABC (61) ABC (50) LSAB (23) Other (44)	MW (106) PC (55) PC in MW (5) Other (10)	1A6 (76) 1D5 (52) 6F11 (10) Other (38)	St ABC (64) ABC (43) LSAB (41) Other (28)
UK NEQAS organising lab	PC	1D5	St ABC	PC	1A6	St ABC

Numbers shown in parenthesis relate to the number of laboratories using each antibody or system for run 41 (ER – 178 returns) and run 39 (PR and ER, – 176 returns)

In the runs for which PR was requested, laboratories not stocking an antibody to PR were requested to demonstrate ER. 1°, primary; ABC, avidin biotin complex; Ab, antibody; Ag, antigen; MW, microwave oven pretreatment; LSAB, labelled St ABC; "Other", antibodies or systems each used by less than four participants, or incomplete data; PC, pressure cooker pretreatment; St ABC, streptomyces ABC.

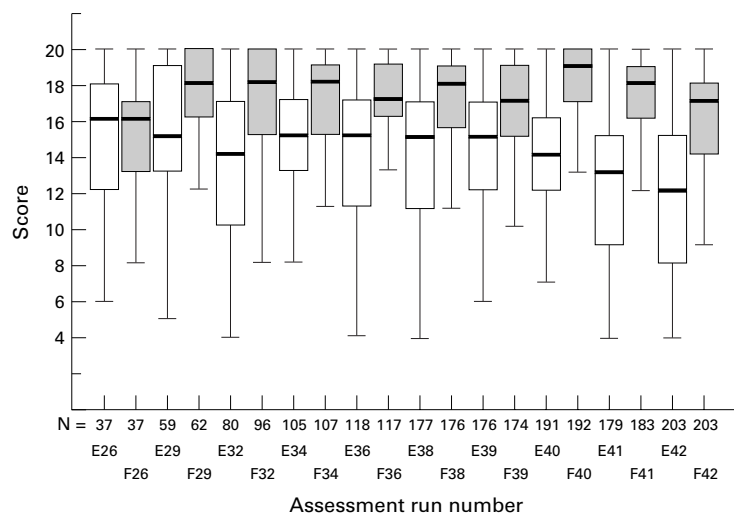


Figure 1 Box plot to show the relation between the UK NEQAS scores achieved by participants on the slides circulated by UK NEQAS (unshaded boxes labelled with run number and the letter "E") and the scores achieved by the same participants on their own in house control slides (shaded boxes labelled with the run number and the letter "F") between April 1994 and June 1998. The maximum score attainable was 20 and the minimum score attainable was 4. The bold line across each box indicates the median score. N, number of laboratories participating in each assessment run.

validating their technique. These studies are not referred to here because this would identify the laboratories concerned, and in so doing transgress the UK NEQAS code of practice, which confers anonymity to all participants.<sup>15</sup> The proportion of these laboratories achieving staining comparable with that of the UK NEQAS organising centre on the tumours circulated at assessment by the scheme and on their own in house tumours was determined for all the assessments between April 1994 and June 1998. For the in house tumours, the Wilcoxon two sample matched pairs signed ranks test was also used to test for differences in distributions between the marks awarded to the expert centres and those awarded to the UK NEQAS organising laboratory for staining of the same tumours.

#### COMPARISON OF THE SCORES AWARDED TO PARTICIPANTS FOR STAINING OF IN HOUSE TUMOURS AND THE STAINING OF UK NEQAS-ICC TUMOURS

The UK NEQAS scores for all the assessments conducted for ER and/or PR between April 1994 and June 1998 were compared by the use of box plots and by establishing the proportion

of participants achieving acceptable staining at each assessment on the two sets of slides. The Wilcoxon two sample matched pairs test was used to test for differences in distributions. Spearman's rank coefficient was used to test for correlations between the results achieved by participants on the UK NEQAS tumours circulated at assessment and on their own in house tumours.

## Results

### VALIDATION OF ASSESSMENT THRESHOLDS AND OPTIMAL SENSITIVITY

#### Comparison of the scores of expert centres and those of the UK NEQAS organising centre on UK NEQAS tumours

For all the assessments (but one) conducted between April 1994–June 1998, all of the six expert centres demonstrated > 10% nuclei in tumours deemed by the organising centre to be hormonal receptor positive. The exception was assessment run 42, where two of the expert laboratories stained < 10% of invasive nuclei in a low expressing infiltrating ductal carcinoma, considered by the UK NEQAS organising centre to be ER positive.

#### Correlation with biochemical values of UK NEQAS tumours circulated at assessment

Of the tumours used for assessments by the scheme between April 1994 and June 1998, 17 of 21 had been initially tested before assessment using both the LBA and IHC. Of these, all were similarly receptor positive or negative with either assay, using an arbitrary threshold value of 10% or greater of invasive tumour nuclei stained by IHC or 10 fmol/mg protein or greater with the LBA, as designating receptor positive status.

#### Comparison of the scores of expert centres and those of the UK NEQAS organising centre on the same in house tumours

The Wilcoxon two sample matched pairs signed ranks test was used to test for differences in distributions between the scores awarded to the expert centres for the quality of immunostaining of their own in house tumour and the marks awarded to the UK NEQAS organising laboratory for staining of the same tumours. This test revealed no significant difference, either in the routine scores

Table 5 Assessment runs April 1994 to June 1998; differences in the distribution of marks for UK NEQAS and in house slides, measures of correlation between the two sets of scores, and details of the pass rates at each assessment

Run no.	Distribution differences (E) and (F)		Spearman's correlation (r)			Pass rate	
	Wilcoxon (Z)	p Value	r	SE	p Value	NEQ (E)	In house (F)
26	-0.189	0.850	0.436	0.152	0.007	73%	81%
29	-3.612	< 0.0001	0.314	0.126	0.018	74%	94%
32	-5.705	< 0.0001	0.462	0.099	< 0.0001	60%	89%
34	-4.962	< 0.0001	0.485	0.092	< 0.0001	75%	90%
36	-7.333	< 0.0001	0.336	0.090	< 0.0001	70%	97%
38	-7.990	< 0.0001	0.280	0.070	< 0.0001	68%	91%
39	-6.464	< 0.0001	0.356	0.069	< 0.0001	73%	94%
40	-10.705	< 0.0001	0.307	0.068	< 0.0001	64%	97%
41	-10.345	< 0.0001	0.304	0.068	< 0.0001	56%	98%
42	-10.769	< 0.0001	0.327	0.064	< 0.0001	43%	87%

Pass rate refers to the proportion of participants achieving a score > 12 out of 20 (acceptable staining). NEQ (E) refers to UK NEQAS slides circulated to each participating laboratory. In house (F) refers to the participants own in house tumours.



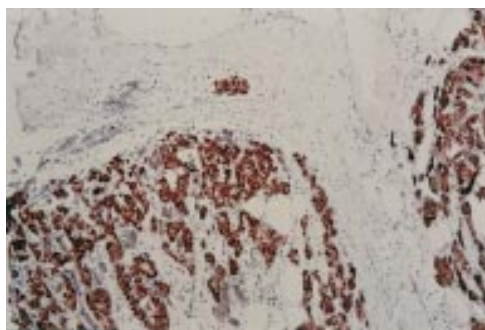


Figure 2 Results of immunohistochemistry for oestrogen receptors (ER) performed by the UK NEQAS organising laboratory on the low expressing (cytosol assay ER, 10 fmol/mg protein), ER positive, infiltrating ductal carcinoma circulated by UK NEQAS-ICC for assessment run 41.

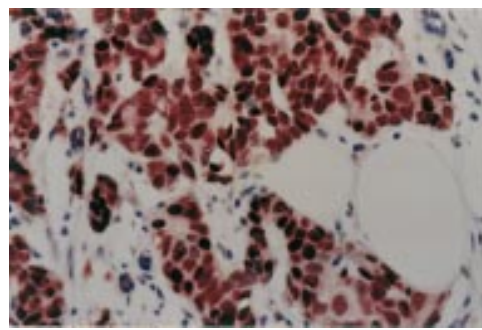


Figure 3 High power detail of the same section shown in fig 2.

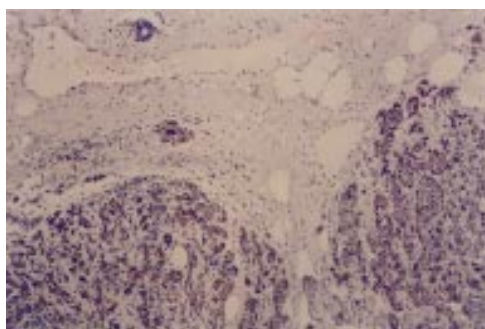


Figure 4 Results of immunohistochemistry for oestrogen receptors (ER) performed by laboratory "X" on the low expressing infiltrating ductal carcinoma shown in figs 2 and 3. The UK NEQAS score awarded to laboratory "X" for this staining was 8 out of 20. Laboratory "X" considered this tumour to be ER negative.

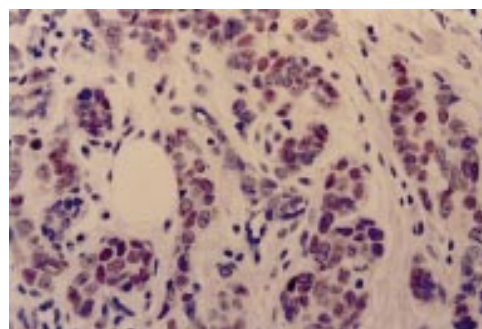


Figure 5 High power detail of the same section shown in fig 4.



Figure 6 Results of immunohistochemistry for oestrogen receptors (ER) performed by laboratory "X" on the high expressing in house tumour submitted by laboratory "X" for run 41 (UK NEQAS score, 16 out of 20, quick score, 4).

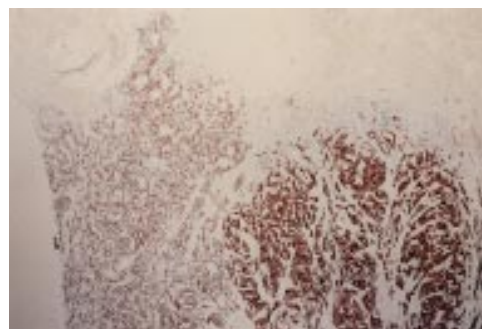


Figure 7 Results of immunohistochemistry for oestrogen receptors (ER) performed by the UK NEQAS organising laboratory on the tumour of laboratory "X" shown in fig 6 (UK NEQAS score, 20 out of 20, quick score, 7).

( $Z = -0.170$ ;  $p = 0.865$ ) or the quick scores generated for run 41 ( $Z = -0.647$ ;  $p = 0.518$ ).

COMPARISON OF THE UK NEQAS SCORES AWARDED AT ASSESSMENT FOR THE QUALITY OF IMMUNOSTAINING ON THE SLIDES CIRCULATED AND THOSE ACHIEVED BY THE SAME PARTICIPANTS ON IN HOUSE SLIDES FROM APRIL 1994 TO JUNE 1998

For nine out of the 10 assessments analysed, the median for the scores that participants achieved on in house tumours was higher than the median for the scores achieved on UK NEQAS tumours (fig 1). The Wilcoxon signed ranks test showed a highly significant difference in the distribution of marks for these nine

runs ( $p < 0.0001$ ; two tailed; table 5). The interquartile range was also frequently smaller for the scores achieved on in house sections, indicating less spread in the results. The proportion of participants achieving acceptable staining was always greater on the in house tissues than on the tumours circulated at assessment. These differences ranged from just 8% for run 26 to 44% for run 42 (table 5). However, Spearman's test showed a significant positive correlation between the routine scores awarded for the staining of UK NEQAS tumours and the staining of in house tumours. This relation is seen for all the assessments for ER/PR conducted between April 1994 and June 1998 (table 5; figs 2-9).

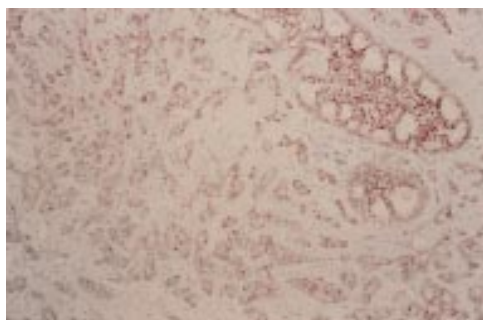


Figure 8 Results of immunohistochemistry for oestrogen receptors (ER) performed by laboratory "Y" on the high expressing in house tumour submitted by laboratory "Y" (UK NEQAS score, 13 out of 20, quick score, 3). Laboratory "Y" scored 6 out of 20 on the low expressing infiltrating ductal carcinoma shown in figs 2–5 and considered this tumour to be ER negative.

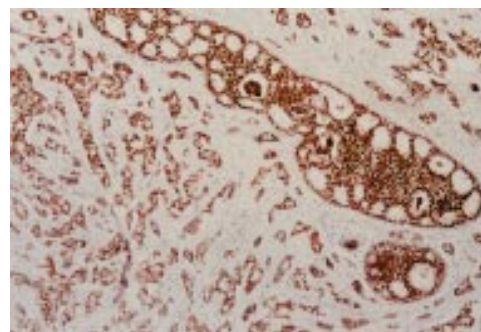


Figure 9 Results of immunohistochemistry (IHC) for oestrogen receptors (ER) performed by the UK NEQAS organising laboratory on the tumour of laboratory "Y" shown in fig 8 (UK NEQAS score, 20 out of 20, quick score, 7).

#### COMPARISON OF THE QUICK SCORES (RUN 41) ON PARTICIPANTS' STAINING OF THEIR OWN IN HOUSE TUMOURS AND THE QUICK SCORES OF THE SAME TUMOURS WHEN STAINED BY THE UK NEQAS-ICC ORGANISING CENTRE

The number of participants who submitted two unstained slides containing sections of breast tumour, along with their own laboratories immunostaining of that tumour, for run 41 was 152 (85% of the total returns). The remaining 26 participants (15%) did not provide unstained slides, or only ones of normal breast tissue. Table 2 details the types of tumours submitted, as described in the returned questionnaires. The initial Wilcoxon test indicated a highly significant difference between the quick scores for the in house tumours when stained by the participant and when stained by the organising laboratory ( $Z = -6.814$ ;  $p < 0.0001$ ; two tailed). Table 6 shows the degree of expression of the 152 tumours as evaluated using the quick score method on both the slides stained by the participants and duplicate slides stained by the organising laboratory. The proportion of cases designated as high expressers was 51.3% ( $n = 78$ ) by the participants' staining and 80.9% ( $n = 123$ ) by the UK NEQAS organising laboratory, with concordance on 72 cases (55.8%;  $\kappa$  coefficient,  $-0.091$ ;  $p = 0.043$ ). The proportion of cases designated as medium expressers was 27.6% ( $n = 42$ ) by the participants' staining and 12.5% ( $n = 19$ ) by the organising laboratory, with concordance on seven cases (13.0%;  $\kappa$  coefficient,  $-0.495$ ;  $p < 0.0001$ ). The proportion of cases designated as low expressers was 15.8% ( $n = 24$ ) by the participants' staining and 4.6% ( $n = 7$ ) by the organising laboratory, with concordance on

just three cases (10.7%;  $\kappa$  coefficient,  $-0.316$ ;  $p < 0.0001$ ). Lastly, the proportion of cases designated as negative was 5.3% ( $n = 8$ ) by the participants' staining and 2.0% ( $n = 3$ ) by the organising laboratory, with concordance on three cases (37.5%;  $\kappa$  coefficient, not applicable). Overall, there was agreement on the degree of expression, as defined by the participants' staining and that of the organising laboratory, in 96 of the 152 cases (63.2%;  $\kappa$  coefficient,  $-0.026$ ;  $p = 0.291$ ).

Table 7 details the analysis of the tumours showing less than high ER expression by the participants' IHC. Of the 42 cases classified as medium ER expressers and obtaining quick scores of 4 and 5, 69% ( $p = 0.014$ ) were shown to have higher expression when stained by the UK NEQAS organising laboratory and achieved quick scores that were higher by 2 marks or more. Of the 24 cases initially classified as low expressers and obtaining quick scores of 3 or 2, 83% ( $p = 0.001$ ) were shown to have higher expression when stained by the UK NEQAS organising laboratory, with quick scores that were higher by 2 marks or more (table 7). Lastly, five of the eight cases classified by the participants' IHC as being ER negative and having quick scores of zero were shown to be ER positive when stained by the UK NEQAS reference laboratory, with one having a quick score of 2, two quick scores of 3, and two quick scores of 6.

#### EVALUATION OF THE EFFICIENCY OF THE UK NEQAS ORGANISING LABORATORY'S ROUTINE METHOD IN STAINING TUMOURS SUBMITTED FROM PARTICIPATING LABORATORIES

Using the standard UK NEQAS scoring system, for six of the seven assessment runs at

Table 6 The degree of oestrogen receptor (ER) expression of 152 tumours from 152 laboratories participating in assessment run 41, as defined by the participants' IHC assays and the UK NEQAS organising laboratory's IHC assay

Degree of ER expression	Participant laboratory		Organising laboratory		Level of concordance			$\kappa$ Value	p Value
	n	%	n	%	n	%	% of total		
High expressing tumours (quick scores of 7 and 6)	78	51.3	123	80.9	72	55.8	47.4	-0.091	0.043
Medium expressing tumours (quick scores of 5 and 4)	42	27.6	19	12.5	7	13.0	4.6	-0.495	< 0.0001
Low expressing tumours (quick scores of 3 and 2)	24	15.8	7	4.6	3	10.7	2.0	-0.316	< 0.0001
Negative tumours (quick scores of 0)	8	5.3	3	2.0	3	37.5	2.0	NA	NA
All tumours (quick scores the same, within $\pm 1$ mark)					96	63.2	63.2	-0.026	0.291

ICH, immunohistochemistry; NA, not applicable.

Table 7 The proportion of the 74 participating laboratories in house tumours submitted for run 41 that showed a higher degree of oestrogen receptor (ER) expression\* when tested by the UK NEQAS organising laboratory

ER expression (participant)	n	ER expression (organising lab)	n	Quick score (participant)	Quick score (organising lab)	%	$\chi^2$ Value	p Value
Medium	42	High	29	5	7	69%	6.095	0.014
				4	6 or 7			
Low	24	High/Medium	20	3	> 4	83%	10.667	0.001
				2	> 3			
Negative	8	High	5	0	6, 6	63%	0.500	0.480
		Low			3, 3			
					2			

\*Achieved a quick score that was greater by 2 marks or more.

which the UK NEQAS organising laboratory stained participants' in house slides, the median of the scores awarded to the organising laboratory was either equal or greater to the median of the scores awarded to the participants on these same in house slides (fig 10). The interquartile range for the scores achieved by the organising laboratory was smaller than the interquartile range of the participants scores on all seven occasions, indicating less spread in the results. The Wilcoxon signed rank test showed that the distribution of scores awarded to the UK NEQAS organising laboratory was significantly higher overall ( $Z = -6.190$ ;  $p < 0.0001$ ; two tailed), and individually in four of the seven assessments. For the remaining three runs, there was no significant difference between the two sets of scores (table 8).

Using the quick score method of evaluation, the technique used by the UK NEQAS organising laboratory was 99% efficient ( $p < 0.0001$ ) in demonstrating the 152 different tumours submitted by participants for run 41, at either the first or second attempt. The overall efficiency achieved by participants using various different methods was 65% ( $p < 0.0001$ ; table 9).

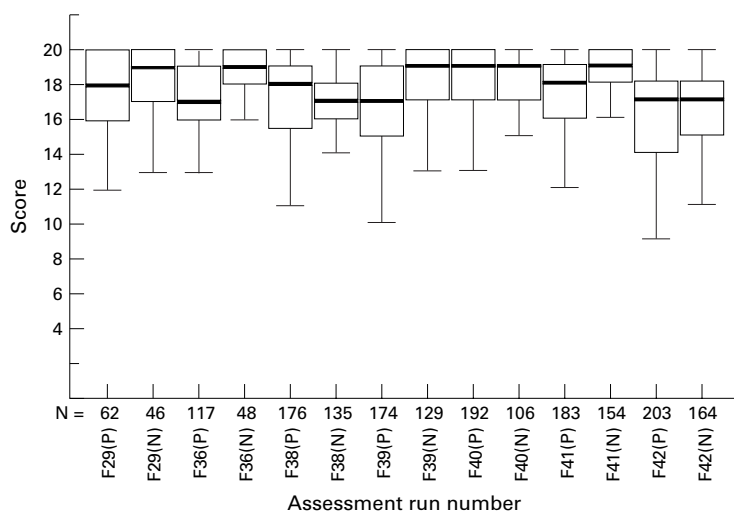


Figure 10 Box plot to compare the scores achieved by participants on their own in house breast tumours and the scores awarded for the UK NEQAS organising laboratory's immunostaining of duplicate sections of the same tumours. The plot shows seven assessment runs between February 1995 and April 1998 for which the UK NEQAS-ICC organising laboratory stained participants' in house slides. Run numbers are labelled F29–F42. The boxes labelled "P" refer to the participants' scores, whereas the boxes labelled "N" refer to the scores awarded to the UK NEQAS organising laboratory.

#### MEASURES OF ASSESSOR CONCORDANCE WHEN EVALUATING IHC SENSITIVITY BY THE QUICK SCORE METHOD FOR RUN 41

Goodman and Kruskal's  $\gamma$  statistic showed highly significant observer concordance between the assessors (AR, BJ) when using the quick score method to evaluate slides, with values of 0.949 ( $p < 0.0001$ ) and 0.960 ( $p < 0.0001$ ) for the staining of in house tumours by the participants and by the UK NEQAS organising laboratory, respectively.

#### Discussion

For the accurate assessment of the results achieved by different laboratories participating in EQA it is essential to validate the standards against which optimal sensitivity is defined. In our study, we have sought to validate these standards in various ways. Comparison of the results deemed to be optimal by the organising centre with those achieved by participants of the scheme who are known to have clinically validated their results has revealed many similarities, both in the proportion of the UK NEQAS tumours confirmed to be receptor positive and in the quick scores generated on the in house tumours for run 41. Of the tumours used for assessment by the scheme between April 1994 and June 1998, 81% had been initially tested using both the LBA and IHC. Of these, all were similarly receptor positive or negative with either assay, using a threshold value of 10% or greater of invasive tumour nuclei stained by IHC and 10 fmol/mg protein or greater with the LBA, as designating receptor positive status. Although the use of any threshold value is arbitrary, we have used this cut off point because of its use in several studies that correlate IHC receptor assay results with clinical and biochemical values.<sup>16–20</sup> We have also shown previously that this is the threshold most commonly used by the laboratories participating in UK NEQAS-ICC.<sup>3</sup> Also imperative to our study is the reproducibility of the methods of evaluation used to assess the quality of IHC. The reproducibility of the routine UK NEQAS scoring system was ensured at assessment by the checking of assessor concordance after every 20 slides. A highly significant degree of concordance with the quick score evaluations was confirmed by Goodman and Kruskal's  $\gamma$  statistic.

Between April 1994 and June 1998, UK NEQAS-ICC conducted 10 assessment runs for ER or PR. During this period, the pass rate on in house tumours remained high (81–97%), whereas that on the distributed UK NEQAS



Table 8 The differences in distributions of the routine marks awarded to participants for the staining of their own in house tumour and those awarded to the UK NEQAS-ICC organising laboratory for staining of the same tumour (February 1995 to June 1998)

Run no.	Distribution differences		Pass rate (scores > 12/20)	
	Wilcoxon (Z)	p Value	Participant	NEQAS organising lab*
29	-2.079	0.038	93%	100%
36	-2.752	0.006	97%	98%
38	-0.105	0.916	93%	98%
39	-4.416	< 0.0001	94%	97%
40	-0.606	0.544	99%	97%
41	-5.982	< 0.0001	97%	99%
42	-0.026	0.979	87%	90%

\*The scores awarded to the UK NEQAS-ICC organising laboratory for immunochemical staining of participants in house tumours was performed for research purposes only.

slides fell, particularly for the later runs (runs 40–42). The reasons for these differences are twofold. First, many in house slides submitted for assessment contain just one tumour with high ER/PR expression, and thus are easier to stain than the UK NEQAS tumours. For example, for the one assessment (run 41) subjected to quick score evaluation, at least 47% of the in house tumours submitted were judged to be high expressing tumours by both the participant and the UK NEQAS organising laboratory (table 6). In contrast, the scheme has circulated slides from composite blocks comprising ER positive/PR positive tumours with progressively lower amounts of expression, particularly for runs 40, 41, and 42 (table 1).

The second reason for the difference does not appear to be possible biological differences in ER expression of the tumours used, but more probably differences in the way the tissues have been prepared. In cellular pathology, a multitude of variables affect a specimen the moment it is removed from the patient—for example, delay in fixation, type of fixation, duration of fixation, fixation temperature, paraffin wax processing schedule, and so on.<sup>21–23</sup> Individually, or in combination, these variables might have an effect on the efficiency and reliability of the immunohistochemical assay to demonstrate various antigens.<sup>23</sup> To minimise their effect, the technologist will have adjusted the methodology (over a period of time) to achieve consistently optimum results, according to his or her laboratory's expectations of the desired standard. When presented with tissue subjected to a different set of fixation and processing variables, as is the case with EQA and referred material, the efficiency of the in house method might fail to achieve the optimal result to varying degrees. This probably is the most likely reason for the differences seen between the scores that participants achieved for in house tumours and for the UK NEQAS

slides, over the range of antigen expression included in the tumours examined.

Although the results achieved by participants on the UK NEQAS tumours were significantly different to those obtained on the in house tumours, Spearman's coefficient revealed that there was a significant positive correlation between the two sets of scores (table 5). To understand why there is a significant difference, and yet still a significant correlation, it is necessary to consider the relation between IHC assay sensitivity and the degree of receptor expression by the tumours under investigation. Suboptimal IHC assay sensitivity when used to stain ER or PR in a low expressing receptor positive tumour (for example, a UK NEQAS tumour) usually results in < 10% of invasive nuclei being stained and a failure at assessment. When applied to a high expressing in house tumour, this same degree of IHC sensitivity is also suboptimal because some invasive receptor positive nuclei that should be demonstrated are not. However, this is unlikely to result in < 10% of the tumour nuclei being stained, purely on the basis of the large number of receptor epitopes available. Consequently, participants who fail on the low expressing receptor positive UK NEQAS tumour tend to achieve lower scores than they should on their high expressing in house tumour. However, they do not usually fail (a score < 10 out of 20) if the proportion of nuclei demonstrated is equal to, or greater than, the designated 10% threshold. Conversely participants with high IHC assay sensitivity, who achieve a relatively high score on the low expressing receptor positive UK NEQAS tumour, tend to score very high marks on their in house tumour. This relation between assay sensitivity and the proportion of nuclei demonstrated in low expressing ER positive tumours and high expressing ER positive tumours is illustrated in figs 2–9.

The main implication of this correlation is that the IHC sensitivity achieved by laboratories on tumours circulated by UK NEQAS-ICC at assessment is a reflection of the sensitivity that the same laboratories achieve on tumours fixed and processed in their own laboratory (in house tumours). This is the first time evidence has been obtained in support of the view that the IHC results achieved on EQA material are accurate indicators of in house laboratory performance.

It has been shown previously that there is a significant positive correlation between the sensitivity achieved by the same laboratories on tumours of differing expression when these tumours are stained as a composite block.<sup>3</sup> In

Table 9 Relative efficiency of the immunohistochemical (IHC) assay of the UK NEQAS organising laboratory in achieving optimal demonstration of oestrogen receptors in 152 breast carcinomas, fixed and processed in 152 different laboratories

Efficiency of IHC assays	n	%	$\chi^2$ Value	p Value
Proportion of 152 cases showing the same* or a greater level of expression** (UK NEQAS organising laboratory's assay)	150	99%	144.105	> 0.0001
Proportion of 152 cases showing the same† or a greater level of expression‡ (participating laboratories' assays)	99	65%	13.921	> 0.0001

\*IHC assay results awarded the same quick score as participant ( $\pm 1$  mark)

\*\*IHC assay results awarded a quick score that was greater by 2 marks or more than that awarded for participant's IHC.

†IHC assay results awarded the same quick score as UK NEQAS organising laboratory ( $\pm 1$  mark).

‡IHC assay results awarded a quick score that was greater by 2 marks or more than that awarded for UK NEQAS organising laboratory's IHC.



our study, we show that there is a similar correlation between suboptimal demonstration of in house tumours and suboptimal demonstration (< 10% nuclei staining) of relatively low expressing ER positive tumours circulated by an EQA scheme.

At present, there is a tendency to overlook the implications of suboptimal staining of tumours with relatively high amounts of ER or PR expression. Reiner and colleagues<sup>12</sup> and MacGrogan and colleagues<sup>24</sup> showed that patients whose carcinomas contained high numbers of hormone receptor positive cells (> 30%, > 50%, > 70%) had a better overall survival than those patients whose tumours had fewer receptor expressing cells. This was in general agreement with the results of Barnes and colleagues<sup>13 14</sup> and Walker *et al*, who showed a high rate of recurrence occurring in patients whose tumours contained high proportions of ER negative cells.<sup>25</sup> Hawkins has suggested that it is possible for ER IHC results to be divided into a minimum of four categories (negative, low, medium, and high) and still provide prognostic/predictive information similar to that provided, as a continuum, by a sensitive and quantitative biochemical assay.<sup>26</sup> Our present study found that 69% ( $p = 0.014$ ) of tumours initially classified as medium expressers by participants' staining were subsequently shown to be high expressers when stained by the organising centre, and that 83% ( $p = 0.001$ ) of those classified as low expressers were shown to be medium or high expressers (table 7). In addition, five of eight completely ER negative tumours were found to be ER positive. Two of these tumours were subsequently classified as high expressers, with quick scores of 6, and three low expressers, with quick scores of 3 or 2, although all with 10% or greater of the tumour nuclei staining. Obviously, the clinical importance of producing false negative staining is greater than that of suboptimally staining relatively high expressing ER positive tumours.<sup>1</sup> However, all these tumours were from patients who, according to the criteria of Reiner *et al*,<sup>12</sup> Barnes *et al*,<sup>13 14</sup> Walker *et al*,<sup>25</sup> and Hawkins,<sup>26</sup> would have a better overall survival than that predicted by the initial in house IHC assays.

This raises the important question as to whether anything short of optimal immunostaining is acceptable for ER testing, the results of which are likely to influence overall clinical management. A number of other immunocytochemical markers also appear to fall into this category, c-erb-2 being one of the best examples.<sup>27-30</sup>

Evaluation of the efficiency of the UK NEQAS reference laboratory's routine technique in analysing the in house tumours submitted between February 1995 and June 1998 using the standard UK NEQAS scoring system gives an efficiency ranging from 90% to 100% (table 8). The in depth analysis of run 41, using the quick score method to evaluate staining of 152 in house tumours, shows the organising laboratory's routine method to be 99% efficient in achieving an equivalent or greater sensitivity to that of the participating

laboratory from where the tumours were submitted (table 9). This degree of sensitivity was achieved on the first, or second, attempt.

These results clearly indicate that the variations in fixation and paraffin wax processing that have been used by participating laboratories on the tumours submitted for assessment to date are not limiting factors preventing a different laboratory achieving a similar or greater degree of sensitivity for hormonal receptors. This is supported by a detailed analysis conducted by Williams *et al*, which showed that variations in immunostaining as a result of variations in fixation and processing regimens could be overcome by heat mediated antigen retrieval.<sup>23</sup>

The organising laboratory's standard technique uses routine commercial antibodies and reagents—that is, the same antibodies or reagents used by numerous participants at assessment and by some of the laboratories that have clinically validated their results (table 4). Although some participants use different clones to the ones used by the UK NEQAS organising laboratory, the technical comparisons performed after recent assessments do not show one clone to ER or PR to be significantly superior to another.<sup>31</sup> The same applies for the differing secondary detection systems. This leaves the efficiency of the heat mediated antigen retrieval step as the most likely factor preventing some participants from achieving optimal demonstration of hormonal receptors. Because all the clones to ER currently used at assessment, and most of those used to PR, necessitate heat mediated antigen retrieval for use on routinely processed tissues,<sup>32-36</sup> the degree of sensitivity ultimately achieved with these clones is directly dependant on how well the heat mediated antigen retrieval step has been performed. A multicentre study, involving 15 French laboratories, found the duration of the antigen retrieval step to be the crucial factor preventing some of the participating laboratories producing adequate results for ER on tissues fixed in a different laboratory.<sup>37</sup> Subsequent extension of the heat mediated antigen retrieval time allowed these laboratories to achieve optimal results. This supports the findings of our study, which suggest inefficiencies in the heat mediated antigen retrieval step might be the most important factor responsible for poor IHC demonstration of hormonal receptors. It is beyond the remit of our present investigation to provide an in depth technical analysis of the different variables that affect the efficiency of heat mediated antigen retrieval. To date, the few publications on this subject have mainly restricted their investigations to the efficiencies of the buffers used.<sup>38-41</sup> Other papers have compared the efficiency of the various heating methods available—for example, microwave ovens versus pressure cookers.<sup>42 43</sup> However, a comprehensive study is clearly required to compare the relative merits of all the different systems, particularly with respect to their efficiencies in the demonstration of low ER/PR positive breast tumours that have been fixed and processed under differing conditions. This would provide valuable information in the

formulation of recommended technical guidelines for optimal IHC demonstration of ER and PR. In turn, this should help in the standardisation of the technique, increase the sensitivity of detection for some laboratories, and ultimately help ensure that hormonal receptor positive cases are not erroneously reported as hormonal receptor negative.

We thank E Anderson, D Barnes, R Baumann, L Bobrow, V LeDoussal, and R Golouh for providing us with invaluable assistance, and all the participants of UK NEQAS-ICC, without whom this study would not have been possible.

- Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 1998;351:1451-67.
- Elledge RM, Osborne CK. Oestrogen receptors and breast cancer: it is time for individualised treatment based on oestrogen receptor status. *BMJ* 1997;314:1843-44.
- Rhodes A, Jasani B, Barnes D, et al. The reliability of immunohistochemical demonstration of oestrogen receptors in routine practice: interlaboratory variance in the sensitivity of detection and evaluation of scoring systems. *J Clin Pathol* 2000;53:125-30.
- Lambkin HA, Dunne P, McCarthy PM. Standardisation of estrogen-receptor analysis by immunohistochemistry: an assessment of interlaboratory performance in Ireland. *Applied Immunohistochemistry* 1998;6:103-7.
- Barnes DM, Millis RR, Beex LVA, et al. Increased use of immunohistochemistry for oestrogen receptor measurement in mammary carcinoma: the need for quality assurance. *Eur J Cancer* 1998;34:1677-82.
- Groupe d'Etude des Récepteurs Hormonaux par Immunohistochemie. FNCLCC/AFAQP. Recommandations pour l'évaluation immunohistochimique des récepteurs hormonaux sur coupes en paraffine dans les carcinomes mammaires. *Ann Pathol* 1996;16:144-8.
- Shahangian S. Proficiency testing in laboratory medicine: uses and limitations. *Arch Pathol Lab Med* 1998;122:15-30.
- Naito HK, Kwak Y-S, Hartfield JL, et al. Matrix effects on proficiency testing materials: impact on accuracy of cholesterol measurements in laboratories in the nation's largest hospital system. *Arch Pathol Lab Med* 1993;117:345-51.
- Ross JW, Myers GL, Gilmore BF, et al. Matrix effects and the accuracy of cholesterol analysis. *Arch Pathol Lab Med* 1993;117:393-400.
- Lasky FD. Achieving accuracy for routine clinical chemistry methods by using patient specimen correlations to assign calibrator values: a means of managing matrix effects. *Arch Pathol Lab Med* 1993;117:412-19.
- Long T. Statistical power in the detection of matrix effects. *Arch Pathol Lab Med* 1993;117:387-92.
- Reiner A, Nuemeister B, Spona J, et al. Immunocytochemical localisation of oestrogen and progesterone receptors and prognosis in human primary breast cancer. *Cancer Res* 1990;50:7057-61.
- Barnes DM, Harris WH, Smith P, et al. Immunohistochemical determination of oestrogen receptors: comparison of different methods of assessment of staining and correlation with clinical outcome of breast cancer patients. *Br J Cancer* 1996;74:1445-51.
- Barnes DM, Millis RR. Oestrogen receptors: the history, the relevance and the methods of evaluation. In: Kirkham N, Lemoine NR, eds. *Progress in pathology* 2. Edinburgh: Churchill Livingstone, 1995:89-114.
- Participation in UK NEQAS. In: *UK NEQAS report and directory*, 3rd ed. Sheffield: UK NEQAS Office, 1998:15.
- Pellicer EM, Sundblad A. Evaluation of antibodies to oestrogen receptors. *Applied Immunohistochemistry* 1994;2:141.
- De Mascarel I, Soubeyran I, MacGrogan G, et al. Immunohistochemical analysis of estrogen receptors in 938 breast carcinomas. Concordance with biochemical assay and prognostic significance. *Applied Immunohistochemistry* 1995;3:222-31.
- Pertschuk LP, Feldman JG, Kim Y-D, et al. Estrogen receptor immunocytochemistry in paraffin embedded tissues with ER1D5 predicts breast cancer endocrine response more accurately than H222Spg in frozen sections or cytosol-based ligand-binding assays. *Cancer* 1996;77:2514-19.
- Ferno M, Andersson C, Giitta F, et al. Oestrogen receptor analysis of paraffin sections and cytosol samples of primary breast cancer in relation to outcome to adjuvant tamoxifen therapy. *Acta Oncol* 1996;35:17-22.
- Soubeyran I, Quenel N, Coindre J-M, et al. PS2: a marker improving prediction of response to neoadjuvant tamoxifen in post-menopausal breast cancer patients. *Br J Cancer* 1996;74:1120-5.
- Larsson LI. Tissue preparation methods for light microscopic immunohistochemistry. *Applied Immunohistochemistry* 1993;1:2-16.
- van Diest PJ, van Dam P, Henzen-Longmans SC, et al. A scoring system for immunohistochemical staining: consensus report of the task force for basic research EORTC-GCCG. *J Clin Pathol* 1997;50:801-4.
- Williams JH, Mepham BL, Wright DH. Tissue preparation for immunocytochemistry. *J Clin Pathol* 1997;50:422-8.
- MacGrogan G, Soubeyran I, De Mascarel I, et al. Immunohistochemical detection of progesterone receptors in breast invasive ductal carcinomas: a correlative study of 942 cases. *Applied Immunohistochemistry* 1996;4:219-27.
- Walker KJ, Bouzubar N, Robertson J, et al. Immunocytochemical localization of estrogen receptor in human breast cancer. *Cancer Res* 1998;48:6517-22.
- Hawkins RA. How best to express oestrogen receptor activity. *Br J Cancer* 1996;74:1329-30.
- Allred DC, Harvey JM, Berado M, et al. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod Pathol* 1998;11:155-68.
- Press MF, Hung G, Godolphin W, et al. Sensitivity of HER-2/neu antibodies in archival tissue samples: potential source of error in immunohistochemical studies of oncogene expression. *Cancer Res* 1994;54:2771-7.
- Pegram MD, Lipton A, Hayes DF, et al. Phase II study of receptor-enhanced chemosensitivity using recombinant humanized anti-p185<sup>HER2/neu</sup> monoclonal antibody plus cisplatin in patients with HER2/neu-overexpressing metastatic breast cancer refractory to chemotherapy treatment. *J Clin Oncol* 1998;16:2659-71.
- Seidman AD, Beselga J, Yao T-J, et al. HER2/neu-overexpression and clinical taxane sensitivity: a multivariate analysis in patients with metastatic breast cancer (MBC). *Proceedings of the American Society of Clinical Oncology* 1996;15:104.
- Rhodes A. UK NEQAS for immunocytochemistry: review of run 44. *J Cell Pathol* 1999;4:114-18.
- Saati TAL, Clamens S, Cohen-Knafo E, et al. Production of monoclonal antibodies to human oestrogen receptor (ER) protein using recombinant ER (RER). *Int J Cancer* 1993;55:651-4.
- Hendricks JB, Wilkinson EJ. Comparison of two antibodies for evaluation of estrogen receptors in paraffin-embedded tumours. *Mod Pathol* 1993;6:765-70.
- Sannino P, Shousha S. Demonstration of oestrogen receptors in paraffin sections of breast carcinoma using microwave oven processing. *J Pathol* 1993;170(suppl):201.
- Kell DL, Kamel OW, Rouse RV. Immunohistochemical analysis of breast carcinoma estrogen and progesterone receptors in paraffin-embedded tissue, correlation of clones ER1D5 and 1A6 with a cytosol-based hormone receptor assay. *Applied Immunohistochemistry* 1993;1:275-81.
- Leong ASY, Milios J. Comparison of antibodies to oestrogen and progesterone receptors and the influence of microwave antigen retrieval. *Applied Immunohistochemistry* 1993;1:282-8.
- Balaton AJ, Mathieu MC, Le Doussal V (on behalf of the Groupe d'Etude des Récepteurs Hormonaux par Immunohistochemie/FNCLCC). Optimization of heat-induced epitope retrieval for estrogen receptor determination by immunohistochemistry on paraffin sections: results of a multi-centric comparative study. *Applied Immunohistochemistry* 1996;4:259-63.
- Morgan JM, Navabi H, Schmid KW, et al. Possible role of tissue bound calcium ions in citrate-mediated high temperature antigen retrieval. *J Pathol* 1994;174:301-7.
- Pileri SA, Roncador G, Ceccarelli C, et al. Antigen retrieval techniques in immunohistochemistry: comparison of different methods. *J Pathol* 1997;183:116-23.
- Cattoretti G, Suurmeijer AJH. Antigen unmasking on formalin-fixed paraffin embedded tissues using microwaves: a review. *Adv Anat Pathol* 1995;2:2-9.
- Leong ASY, Milios J, Leong FJ. Epitope retrieval with microwaves: a comparison of citrate buffer and EDTA with three commercial solutions. *Applied Immunohistochemistry* 1996;4:201-7.
- Norton AJ, Jordan S, Yeomans P. Brief, high temperature heat denaturation (pressure cooking): a simple and effective method of antigen retrieval for routinely processed tissues. *J Pathol* 1994;173:371-9.
- Miller K, Auld J, Jessup E, et al. Antigen unmasking in formalin fixed routinely processed paraffin wax-embedded sections by pressure cooking: a comparison with microwave oven heating and traditional methods. *Adv Anat Pathol* 1995;2:60-4.