J Clin Pathol 1999;52:184–192

Fully automated microvessel counting and hot spot selection by image processing of whole tumour sections in invasive breast cancer

Jeroen A M Beliën, Semir Somi, Johannes S de Jong, Paul J van Diest, Jan P A Baak

Abstract

***Background*—Manual counting of microvessels is subjective and may lead to unacceptable interobserver variability, which may explain conflicting results.**

***Aims*—To develop and test an automated method for microvessel counting and objective selection of the hot spot, based on image processing of whole sections, and to compare this with manual selection of a hot spot and counting of microvessels.**

***Methods*—Microvessels were stained by CD31 immunohistochemistry in 10 cases of invasive breast cancer. The number of microvessels was counted manually in a subjectively selected hot spot, and also in the same complete tumour sections by interactive and automated image processing methods. An algorithm identified the hot spots from microvessel maps of the whole tumour section.**

***Results*—No significant difference in manual microvessel counts was found between two observers within the same hot spot, and counts were significantly correlated. However, when the hot spot was reselected, significantly different results were found between repeated counts by the same observer. Counting all microvessels manually within the entire tumour section resulted in significantly different hot spots than manual counts in selected hot spots by the same observer. Within the entire tumour section no significant differences were found between the hot spots of the manual and automated methods using an automated microscope. The hot spot was found using an eight connective path search algorithm, was located at or near the border of the tumour, and (depending on the size of the hot spot) did not always contain the field with the largest number of microvessels.**

***Conclusions*—The automated counting of microvessels is preferable to the manual method because of the reduction in measurement time when the complete tumour is scanned, the greater accuracy and objectivity of hot spot selection, and the possibility of visual inspection and relocation of each measurement field afterwards.**

(*J Clin Pathol* 1999;52:184–192)

Keywords: breast cancer; image analysis; angiogenesis; automation

Department of Pathology, Academic Hospital Vrije Universiteit, PO Box 7057, 1007 MB Amsterdam, The Netherlands
J A M Beliën
S Somi
J S de Jong
P J van Diest
J P A Baak

Correspondence to:
Prof Dr Baak.
email: jam.belien@azvu.nl

Accepted for publication
19 October 1998

In 1971 Folkman proposed that tumour growth requires the development of new microvessels (angiogenesis).[1 2] Since then, several studies have shown that these new vessels are important for nutrient supply and also for tumour growth by local production and delivery of tumour growth factors.[3] Moreover, new blood vessels provide a route for tumour cells to reach the circulation, allowing metastatic spread.[4]

For different human cancers, microvessel density in the tumour has generally been found to be of independent prognostic significance and to be a predictor of metastatic disease.[2 5] Breast cancer in particular has been studied with regard to the clinical impact of microvessel density (for an overview see Fox[6]), and the initial studies provided promising results.[7–17] However, these results could not be confirmed in several well conducted studies, and no prognostic value was found.[18–23] This can probably be explained mainly by differences in methodology.[5 24]

The correlations between angiogenesis and tumour growth, metastatic spread, and prognosis are based on measurements of tumour microvessel densities highlighted by immunohistochemistry, using antibodies directed against endothelium (for an overview of the specificity and sensitivity of commonly used antibodies, see Weidner[5]). At present, the microvessel density in tumours is mainly assessed by a manual count of the number of microvessels in what appears to be the most vascular area of the tumour—called the hot spot—using a protocol described by Weidner *et al.*[16 17]

With standardised conditions, such as section preparation, staining, careful and reproducible selection of the hot spot, and a strict protocol for defining microvessels, adequate reproducibility can be achieved.[25] Despite these precautions, manual counting of microvessels, and especially the manual selection of the hot spot, is still subjective and therefore not always fully reproducible.[26 27] Therefore, it is desirable to automate the counting of microvessels as well as the selection of the hot spot.

Methods for counting microvessels based on interactive image analysis or interactive image morphometry have been proposed in several studies.[8 14 19 27–37] Although the subjectivity introduced by interobserver variation in recognition and counting of microvessels was greatly reduced by automation, the hot spot was still manually selected and so remained subjective. Recently, van der Laak *et al* presented an

Table 1 Results of experiments 1 and 2: manually counting microvessels in 10 connected fields in a subjectively selected hot spot by two observers in 10 cases of invasive breast cancer

		Experiment 1		*Experiment 2*	
Tumour	*Largest diameter (cm)*	*Observer*	*Number of microvessels and (microvessels/mm²)*	*Observer*	*Number of microvessels and (microvessels/mm²)*
T1	1.5	1A	102 (169.9)	1B	75 (125.0)
		2	93 (154.9)	3	75 (125.0)
T2	2.8	1A	308 (513.1)	1B	223 (371.5)
		2	311 (518.1)	3	223 (371.5)
T3	5	1A	55 (91.6)	1B	62 (103.3)
		2	56 (93.3)	3	60 (100.0)
T4	2	1A	135 (224.9)	1B	145 (241.6)
		2	147 (244.9)	3	141 (234.9)
T5	3	1A	91 (151.6)	1B	73 (121.6)
		2	97 (161.6)	3	68 (113.3)
T6	2.5	1A	93 (154.9)	1B	79 (131.6)
		2	95 (158.3)	3	74 (123.3)
T7	1.3	1A	103 (171.6)	1B	96 (159.9)
		2	92 (153.3)	3	97 (161.6)
T8	2.3	1A	123 (204.9)	1B	111 (184.9)
		2	114 (189.9)	3	109 (181.6)
T9	2.8	1A	139 (231.6)	1B	95 (158.3)
		2	150 (249.9)	3	97 (161.6)
T10	2.1	1A	77 (128.3)	1B	57 (95.0)
		2	71 (118.3)	3	59 (98.3)

Observer 1A and 1B are the same person.

improved semiautomated procedure to quantify tumour vascularity in a human melanoma xenograft model by true colour image analysis.[38] Not only was interobserver variation reduced, but a further step was made towards objectifying hot spot selection. Instead of manually selecting the hot spot first and analysing it afterwards, they recorded the entire tumour section and stored the images. These were then analysed by semiautomated image analysis, allowing interactive correction when necessary. The average value of the three images with the highest vessel density was taken as the "automated hot spot." This procedure indeed is more objective, but does not reflect the true nature of a hot spot, since the three fields might be spread out across the entire tumour, not forming a real hot spot.

In this paper we present for the first time an automated method to identify microvessels in whole tumour sections by image processing, which allows fully objective selection of hot spots and counting of microvessels. The automated system will also be validated by careful comparison of manual counts of microvessels in whole tumour sections.

Methods

Fresh operation specimens from 10 patients were cut into slices of approximately 0.5 cm and tumour size was measured (table 1). The material was fixed in neutral 4% buffered formaldehyde. Representative tumour samples were embedded in paraffin, with special care that the periphery of the tumour was sampled. Sections (4 μm thick) were cut and mounted on 3-aminopropyl-triethoxysilane (Sigma) coated slides for enhanced CD31 immunohistochemistry and for routine staining with haematoxylin and eosin to assess the diagnosis (for choice of CD31, see Parums *et al*[39] and Vermeulen *et al*[40]).

A microscope TV system was used, consisting of an Axioplan microscope (Zeiss) equipped with a 20× plan apochromat objective with a numerical aperture of 0.60. Images were recorded by a XC-77-CE CCD black and white camera (Sony) with a CCD cell of size 11×11 μm. The section was moved with an automatic scanning stage (Märzhäuser) with a step size of 0.25 μm, and an autofocus device (Zeiss) operating on the TV signal took care of the focusing (step size 0.025 μm). The specimen was illuminated with a stabilised halogen light source and filtered with a monochrome filter at the wavelength of the maximum absorption of the CD31 stain (λ = 420 nm, $\Delta\lambda$ = 10 nm). Images of 512×512 picture elements were digitised with an S2200 video board (Data Cell and Active Silicon) in eight bit grey value resolution. The pixel to pixel distance at the specimen level was 0.54 μm. Image processing was performed on a Sparc 10 model 30 workstation (Sun Microsystems) running under the UNIX operating system (SunOS, Sun Microsystems) with a colour monitor, at a spatial resolution of 1152×900 in eight bit resolution.

The automatic microvessel counter was developed and evaluated within the Pathology Image Processing Environment software, which is based on the multilevel interactive image processing environment SCIL_Image, version 1.3 (Dutch Vision Systems); this allows storage of image data as well as numerical data resulting from measurements in combined databases.

IMMUNOHISTOCHEMISTRY

For immunohistochemical staining of the CD31 antigen an avidin–biotinyl peroxidase technique was used.[41] In brief, after the 4 μm thick slides underwent dewaxing and rehydration, endogenous peroxidase activity was blocked by incubation for 30 minutes in 0.3% (vol/vol) hydrogen peroxide in methanol. The slides were heated at 100°C in a 0.01 M citrate buffer (pH 6.0) for 15 minutes for antigen retrieval. Thereafter, the slides were preincubated with normal rabbit serum in a 1:50 dilution in phosphate buffered saline with 1% bovine serum albumin (PBS/BSA) for 10 min-

utes. Subsequently, the slides were incubated with the mouse monoclonal antibody JC70 (Dako; see also Parums *et al*[39]), in a 1:40 dilution of PBS/BSA for 16 hours at 4°C. Thereafter the slides were incubated with biotinylated rabbit antimouse antibody diluted 1:500 in PBS/BSA for 30 minutes at room temperature. The slides were subsequently incubated with avidin–biotinyl peroxidase complex diluted 1:1000 in PBS/BSA for 30 minutes, and then with biotinylated tyramine in a 1:1000 dilution in PBS for 10 minutes (BT enhancement),[41] and then again with avidin–biotinyl peroxidase complex diluted 1:200 in PBS/BSA for 30 minutes. 3,3'-Diaminobenzindine tetrahydrochloride was used as a chromogen. Between steps, the slides were rinsed for 5–10 minutes in PBS three times. All sections were lightly counterstained (20 seconds) with haematoxylin, dehydrated, and mounted.

MICROVESSEL COUNTING

The following six counting experiments were performed on all 10 specimens.

Experiment 1: Interobserver agreement of traditional microvessel counting

The hot spot was identified by two observers (JAMB (observer 1A) and SS (observer 2)) according to the criteria defined by Weidner *et al*,[16 17] with some minor modifications. Each observer individually identified a hot spot by scanning at low power (×25 and ×100). If different hot spots were identified, the observers then had to agree on which one was used for assessing the microvessel count in 10 consecutive fields at ×400 magnification (grid area at ×400 corresponds to 0.06 mm^2). The number of microvessels within the hot spot was assessed using an ocular grid and the forbidden lines method[42] to facilitate the counting. The microvessels within a grid area of one field were counted individually by both observers consecutively, using a single headed microscope, before the next field was selected by the first observer. The location of the hot spot was noted, so that we could compare the results with those obtained in experiment 2.

Experiment 2: Intraobserver and interobserver agreement of traditional microvessel counting

Experiment 1 was repeated "blind" four months later by two observers (JAMB (observer 1B) and PJvD (observer 3)) to assess intraobserver and interobserver agreement when using the traditional manual assessment of hot spot and microvessel counting.

Experiment 3: Manual microvessel map development

Manual counting of all microvessels in the entire tumour area was done by one observer (SS) at a total magnification of ×400, using an ocular grid and the forbidden lines method of Gundersen.[42] Counts per field were recorded to create a microvessel map of the whole tumour section, which was used for hot spot calculation by a custom designed hot spot search program. This program reads the map and produces—depending on parameter settings—hot spots of different shapes and sizes. Variables which can be set are: (1) the shape of the hot spot; a square or a random path; (2) the size of the hot spot; and (3) the type of random path—either four connective (where only fields north, south, west, and east of a given field are examined) or eight connective (the same as four but also including fields northwest, northeast, southwest, and southeast). When the user has selected a square shaped hot spot, the program locates the square hot spot by moving a window of corresponding dimensions over the map. The hot spot is the window with the highest number of microvessels. When a random path has been selected, the program uses an algorithm based on recursion. The output consists of a path or square starting with or containing the field having the highest number of microvessels, and the square or path with the overall highest number of microvessels. If several different squares or paths are found having the same total number of microvessels, the program will display them all. An example of a microvessel map and results from the hot spot search program are shown in fig 1. As shown in this figure, the location of a hot spot could change depending on the initial parameter settings. The output of the hot spot search program was used to compare hot spots of different shape with the same size, and to compare hot spots of different sizes with the same shape.

Experiment 4: Identifying microvessels in the entire tumour area by a fully automated image processing method

Since the coordinates of each individual microvessel were recorded, a microvessel map was also created here to assess the hot spot (fig 1).

Experiment 5: System performance

The performance of automated system was assessed by also counting microvessels manually in exactly the same fields as in experiment 4. In the entire tumour area, consecutive fields from start to end were automatically scanned using the automated scanning stage, and the microvessels were manually counted in each individual field by one observer (JAMB). A microvessel map was also created (fig 1) and processed with the hot spot search program. The results of this experiment were then compared with those of experiment 4 to evaluate the system performance.

Experiment 6: Reproducibility of the automated system

To test the reproducibility of the automated microvessel counting program, one observer (JAMB) first defined a measurement area on one specimen (T6) and then stored this definition. The automated counting method was then started, using this measurement area definition. The measurement results were stored and the program was restarted with the same measurement area definition, and the corresponding measurement results were also stored. The two datasets were then compared for agreement in the number of microvessels.

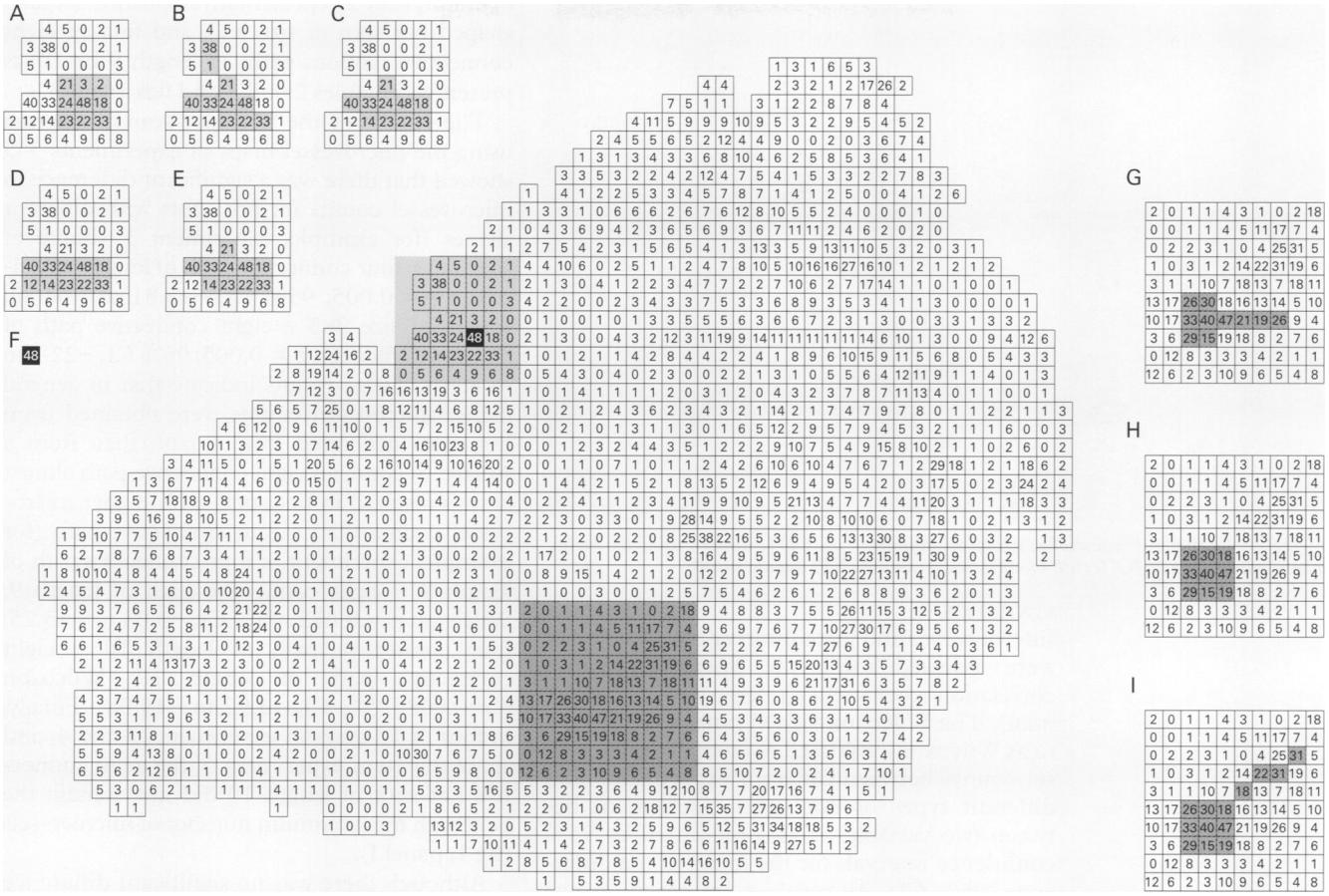

Figure 1 Microvessel map of tumour T2 (central oval area of figure) as generated by the fully automated image processing method which was used as input for the hot spot search program. Panels A to E correspond to the light grey area within the microvessel map and parts G-I with the dark grey area. The map of T2 is shown together with the results (hot spots are light or dark grey) from the automated hot spot search program: (A) Sq3-WM; (B) 8-CRP10-WM; (C) 8-CRP10-SM; (D) 4-CRP10-SM; (E) 4-CRP10-WM; (F) field containing the maximum number of microvessels; (G) 4-CRP10-OPT; (H) Sq3-OPT; (I) 8-CRP10-OPT. Abbreviations: Sq3, hot spot as square of 3×3 fields; 4-CRP10, hot spot as a four connective random path of 10 fields; 8-CRP10, hot spot as an eight connective random path of 10 fields; OPT, square or path containing the greatest number of microvessels possible (ie, with or without field with maximum number of microvessels); SM, path starting with field with maximum number of microvessels; WM, square or path containing field with maximum number of microvessels.

IMAGE PROCESSING

The image processing methodology was tuned to favour the detection of microvessels. The image processing procedure was as follows:

(1) Shading correction—Before every image analysis session, the microscope was carefully adjusted (using Köhler illumination among others) to ensure standardised conditions. The images were linearly corrected for shading with two empty images, one illuminated and one dark current image.[43] The corrected grey values thus provided a measure for the local optical density.

(2) Segmentation—An initial rigorous low grey value threshold of 60 was set to select CD31 immunostained objects. This resulted in a segmentation of the darkest pixels in the image, belonging to the microvessels. These seed points were then propagated to complete vessels using a mask image obtained by setting a relatively high threshold. This last threshold was related to the mode of the illuminated image used for shading correction (threshold = mode of illuminated image − 60).

The described segmentation procedure eliminated 95–98% of non-microvessels, while only a few microvessels did not survive this procedure owing to sectioning or low optical density.

(3) Object features—We measured several features (such as area and perimeter[43 44]) of the objects remaining after segmentation. If an object had an area smaller than 30 μm^2 it was excluded from further analysis on the assumption that objects of these dimensions are not microvessels. Measurement data and images concerning all other potential microvessel objects were stored in databases. As well as measurement information and numerical features of each object, the measurement database also contained the position of the object on the slide to allow relocation and visual inspection or verification of the computer classification through the microscope. All objects measured could be displayed in a composite image. Each object could then be selected and fully automatically relocated to the nearest 0.25 μm for inspection if required.

The complete tumour area in the specimen was measured and all objects were placed in an image and measurement database. Each object was selected and relocated automatically for inspection, after which it was placed in one of the two possible classes: microvessels or non-microvessels (such as tissue folding, plasma cells, and darkly stained nuclei). All the objects used in this study were inspected twice in this way to ensure that each object had been classified correctly. When in doubt, the object was not regarded as a microvessel.

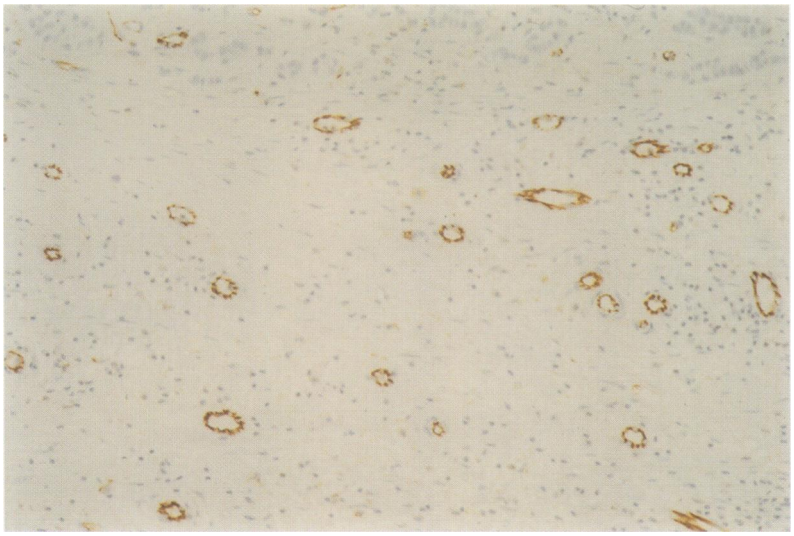

Figure 2 Example of immunohistochemical staining of the CD31 antigen using the mouse monoclonal antibody JC70 with biotinylated tyramine (BT) enhancement.

STATISTICS

Interobserver and intraobserver reproducibility were assessed using the Spearman (rank) correlation coefficient (referred to as Spearman). The Wilcoxon signed rank test (referred to as Wilcoxon) was used to compare microvessel counts between different experiments and different types of hot spots. Agreement between two variables is also presented by 95% confidence intervals for the mean of the difference (95% CI). All p values are based on two tailed testing.

Results

IMMUNOHISTOCHEMISTRY

Immunohistochemical staining of the CD31 antigen using the mouse monoclonal antibody JC70 with BT enhancement resulted in homogeneous, generally crisp and easy to interpret staining, which was confined to endothelial and plasma cells, with only minor background staining in some cases (fig 2). Plasma cells, when present, were easily recognised morphologically and were not included (manual counting) or were removed (automated counting).

HOT SPOT SEARCH PROGRAM

Since the hot spot search program generates an enormous amount of information depending on the shape and size of the hot spot defined by the operator, only a representative selection (square shaped hot spot of size 3×3, and four or eight connective random paths of length 9 or 10) is presented in tables 2 and 3 and figs 1 and 3.

The results of the hot spot search program using the microvessel maps of experiments 3–5 showed that there was a significant difference in microvessel counts for hot spots with different shapes (for example, experiment 3: square of size 3×3 *v* four connective path of length 9: Wilcoxon p = 0.005; 95% CI, −15.81 to −64.89; square of size 3×3 *v* eight connective path of length 9: Wilcoxon p = 0.005; 95% CI, −22.3 to −74.70). These results indicate that in general higher microvessel counts were obtained from hot spots generated from a path than from a square. Second, an eight connective path almost always resulted in hot spots with higher microvessel counts than a four connective path (for example, experiment 3: four connective path of length 9 *v* eight connective path of length 9: Wilcoxon p = 0.012; 95% CI, −3.03 to −13.25; four connective path of length 10 *v* eight connective path of length 10: Wilcoxon p = 0.008; 95% CI, −3.92 to −19.08). Finally, in four of 10 cases in experiments 3 and 4, and six of the 10 in experiment 5, the eight connective hot spot of length 10 did not contain the field with the maximum number of microvessels (fig 1, panel I).

Although there was no significant difference between the absolute number of microvessels in a hot spot between experiments 4 and 5, in two of 10 cases the hot spot search algorithm assigned completely different locations to the hot spot. The differences in microvessel counts in those different hot spots were, however, small. Interestingly, the location of all hot spots was always at or near the periphery of the tumour (see also fig 1).

Besides comparing hot spots with the same size but different shape, fig 3 also shows that, independent of shape, different hot spot sizes may result in different rankings according to the number of microvessels in that hot spot. This especially holds for tumours T4 and T6 which have different ranks at hot spot sizes of 0.08, 0.15, 0.23, and 0.30 mm^2. Another interesting phenomenon could be seen in fig 3C. Tumours having large numbers of microvessels

Table 2 Results of experiment 3: manually counting microvessels (/mm^2) in the entire tumour area in 10 cases of invasive breast cancer, and calculating hot spot microvessel counts in different ways

	Method													
	Sq3		*4-CRP9*			*8-CRP9*			*4-CRP10*			*8-CRP10*		
Tumour	*WM*	*OPT*	*SM*	*WM*	*OPT*	*SM*	*WM*	*OPT*	*SM*	*WM*	*OPT*	*SM*	*WM*	*OPT*
T1	101.8	161.0	151.8	151.8	187.0	162.9	162.9	196.2	151.6	151.6	184.9	159.9	159.9	193.3
T2	440.6	440.6	514.6	538.7	538.7	538.7	538.7	538.7	513.1	519.8	519.8	519.8	519.8	519.8
T3	96.3	96.3	122.2	125.9	125.9	133.3	133.3	133.3	120.0	125.0	125.0	133.3	133.3	133.3
T4	209.2	209.2	312.8	312.8	312.8	331.3	333.2	333.2	289.9	291.6	291.6	316.5	328.2	328.2
T5	185.1	192.5	220.3	220.3	249.9	240.6	240.6	257.3	228.2	228.2	246.6	246.6	246.6	258.2
T6	155.5	155.5	175.9	179.6	179.6	196.2	196.2	196.2	171.6	176.6	176.6	196.6	196.6	196.6
T7	161.0	161.0	172.2	177.7	177.7	177.7	177.7	177.7	168.3	169.9	169.9	169.9	171.6	171.6
T8	157.3	194.4	190.7	190.7	216.6	224.0	224.0	231.4	194.9	194.9	213.2	223.2	223.2	228.2
T9	268.4	268.4	279.5	281.4	281.4	281.4	283.2	283.2	273.2	274.9	274.9	279.9	281.6	281.6
T10	124.0	129.6	135.1	137.0	142.5	140.7	144.4	146.2	131.6	134.9	141.6	138.3	141.6	148.3

Sq3, hot spot as square of 3×3 fields; 4-CRP9/4-CRP10, hot spot as a four connective random path of 9/10 fields; 8-CRP9/8-CRP10, hot spot as an eight connective random path of 9/10 fields; OPT, square or path containing the highest number of microvessels possible (that is, with or without field with maximum number of microvessels); SM, path starts with field with maximum number of microvessels; WM, square or path contains field with maximum number of microvessels.

Table 3 Results of experiment 4 (fully automated counting of microvessels (per mm²) in the entire tumour area) and experiment 5 (manually counting microvessels in exactly the same relocated fields using an automated scanning stage)

	Method, experiment 4						*Method, experiment 5*					
	4-CRP10			*8-CRP10*			*4-CRP10*			*8-CRP10*		
Tumour	*SM*	*WM*	*OPT*	*SM*	*WM*	*OPT*	*SM*	*WM*	*OPT*	*SM*	*WM*	*OPT*
T1	155.0	164.2	178.7	168.2	172.1	197.1	160.3	172.1	174.7	176.0	181.3	191.8
T2	341.6	353.4	371.8	353.4	360.0	383.6	350.8	362.6	375.7	362.6	371.8	392.8
T3	147.1	151.1	151.1	160.3	160.3	160.3	122.2	122.2	122.2	127.4	127.4	127.4
T4	197.1	201.0	261.4	237.8	237.8	275.9	194.4	195.7	260.1	236.5	236.5	275.9
T5	260.1	260.1	260.1	268.0	268.0	268.0	266.7	266.7	266.7	277.2	277.2	277.2
T6	143.2	148.5	212.8	176.0	176.0	216.8	140.6	145.8	202.3	172.1	172.1	206.3
T7	152.4	157.7	157.7	173.4	177.4	177.4	155.0	164.2	164.2	173.4	185.2	185.2
T8	202.3	204.9	204.9	202.3	207.6	207.6	155.0	155.0	174.7	155.0	155.0	182.6
T9	312.7	323.2	323.2	327.1	327.1	327.1	314.0	325.8	325.8	328.4	329.7	329.7
T10	115.6	123.5	123.5	124.8	127.4	127.4	80.1	80.1	118.2	95.9	95.9	124.8

4-CRP10, hot spot as a four connective random path of 10 fields; 8-CRP10, hot spot as an eight connective random path of 10 fields; OPT, square or path containing the highest number of microvessels possible (that is, with or without field with maximum number of microvessels); SM, path starts with field with maximum number of microvessels; WM, square or path contains field with maximum number of microvessels.

in small hot spots may have smaller average numbers of microvessels in larger hot spots.

MICROVESSEL COUNTING

Given the results from the hot spot search program presented above, data of experiments 3–5 are presented as hot spots of eight connective paths of 10 fields only.

The average times needed for each of the first five experiments were 12 minutes (range 7–20), 13 minutes (range 7–18), eight hours (range 4–15), four hours (range 1.5–6), and six hours (range 4–12), respectively. The results of the first five experiments are shown in tables 1 to 4 and figs 1 and 3.

The results of experiments 1 and 2 (table 1) showed that assessment of the number of microvessels within the same (preselected) hot spot did not differ significantly between two different observers (observer 1A *v* 2: Spearman $r = 0.88$, $p = 0.001$; Wilcoxon $p = 0.88$; 95% CI, −9.99 to 9.99; observer 1B *v* 3: Spearman $r = 0.99$, $p < 0.001$; Wilcoxon $p = 0.16$; 95% CI, −5.36 to 1.06). Although the intraobserver agreement (observer 1A *v* 1B) was high (Spearman $r = 0.90$, $p < 0.001$), there was a significant difference between these two counts (Wilcoxon $p = 0.02$; 95% CI, 2.4 to 67.58). These results suggested that in general two observers arrived at a comparable hot spot microvessel count when the same fields were screened; however, even though the observers had to come to an agreement on the selection of a hot spot, their re-assessed counts differed from the initial counts because they did not select exactly the same fields or hot spots.

Comparing the manually assessed number of microvessels/mm² in a subjectively selected hot spot (experiment 1; SS) with the manually assessed number of microvessels/mm² in a hot spot selected by the hot spot search program from the whole tumour section (experiment 3; SS) showed that in all tumours, the hot spot of experiment 3 contained more microvessels (Wilcoxon $p = 0.005$; 95% CI, −10.73 to −30.91).

The results of fully automated counting of microvessels (experiment 4) and counting microvessels manually but using an automated stage (experiment 5) in whole tumour sections (tables 3 and 4) showed that fully automated counting without correction for missed microvessels or erroneously included non-microvessels did not result in significantly different results when exactly the same tumour area was scanned manually using the automated scanning stage (experiment 4 *v* experiment 5: Wilcoxon $p = 0.441$; 95% CI, −7.50 to 2.78). Furthermore, although staining quality of all specimens was good, the false negative percentage of automatically counted microvessels fell to optimal levels (below 5%) when staining intensity increased (table 4).

Although the size of one single field in experiments 1, 2, and 3 did not exactly correspond to the field size of experiments 4 and 5 (0.06003 mm² *v* 0.07612 mm², respectively) there was no significant difference between hot spots of experiments 3 and 4 and experiments 3 and 5 (experiment 3 *v* 4: Wilcoxon $p = 0.799$; 95% CI, −9.31 to 25.83; experiment 3 *v* 5: Wilcoxon $p = 0.445$; 95% CI, −12.59 to 24.39). Furthermore, there was no significant difference between experiment 1 (average score of two observers) and experi-

Table 4 Combined results of experiment 4 (counting microvessels fully automated in the entire tumour area) and experiment 5 (manually counting microvessels in exactly the same relocated fields using an automated scanning stage)

Tumour	*Subjective scoring of staining intensity*	*Total number of vessels counted by*		*Vessels missed by computer (% false negatives)*	*Non-vessels counted as vessels by computer*
		Computer	*Observer*		
T1	+	5391	5280	641 (12.1)	752
T2	+++	9150	8961	255 (2.8)	444
T3	+	1899	1586	173 (10.9)	486
T4	+++	8922	8804	398 (4.5)	516
T5	+++	6573	6154	214 (3.5)	633
T6	+++	3276	3082	164 (5.3)	358
T7	+	3342	3518	559 (15.9)	383
T8	+	8320	8738	1363 (15.6)	945
T9	++	13008	13131	841 (6.4)	718
T10	++	5468	5179	333 (6.4)	622

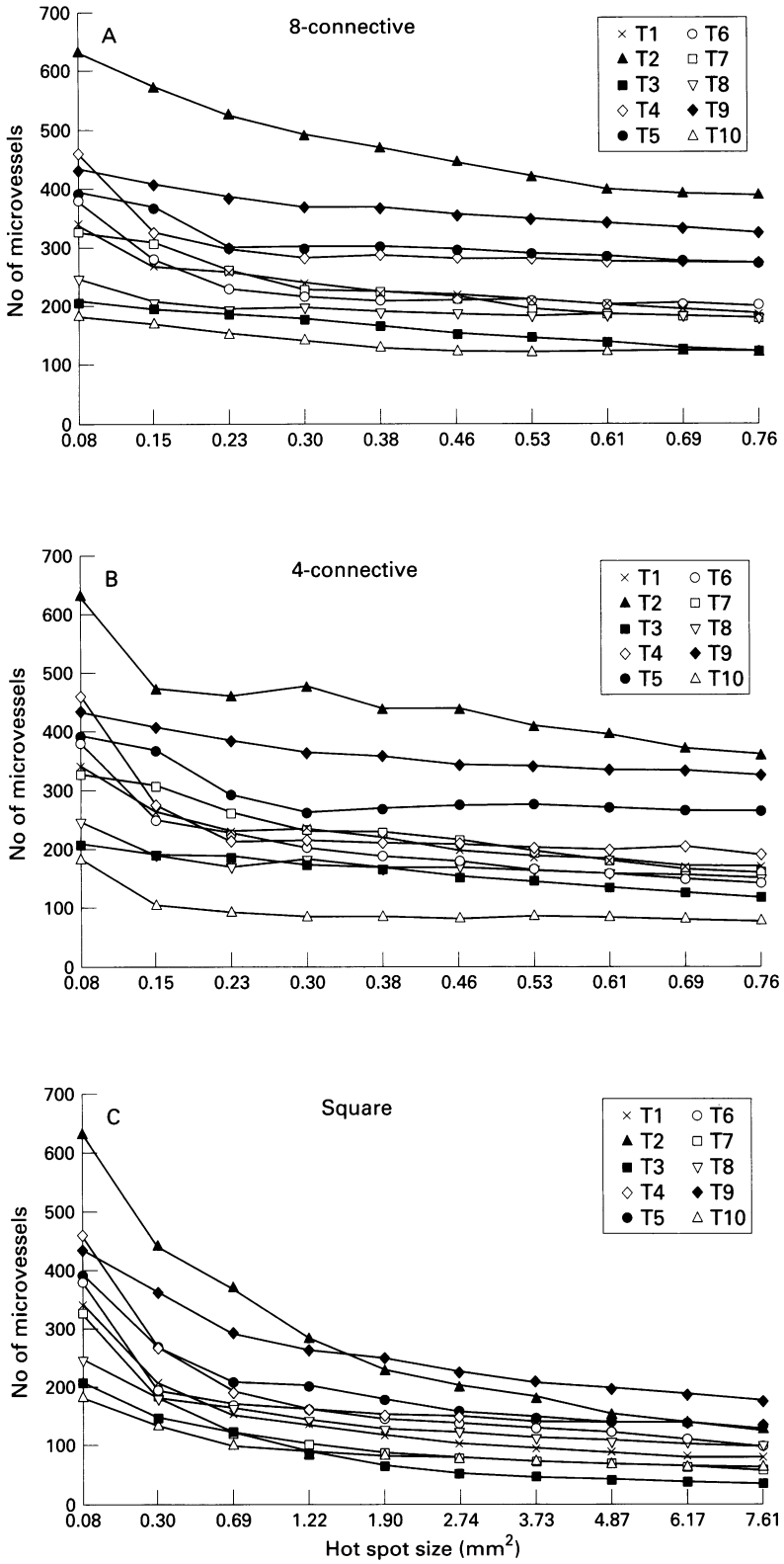

Figure 3 The hot spots of different sizes and shapes (A, eight connective; B, four connective; C, square) found after analysing the microvessel map of tumours T1–T10 using the hot spot search program. A single field size equals 0.07612 mm². (A) and (B) show the hot spots of size 1 to 10; (C) shows the hot spots of size 1×1 to 10×10.

ment 4, and between experiment 1 and experiment 5 (experiment 1 *v* 4: Wilcoxon p = 0.074; experiment 1 *v* 5: Wilcoxon p = 0.114), although the microvessel counts of experiments 4 and 5 were in general higher (experiment 1 *v* 4: 95% CI, −71.70 to 21.44; experiment 1 *v* 5: 95% CI, −77.40 to 17.68). A significant difference was found between the results of experiment 2 (average score of two observers) and experiments 4 and 5 (experiment 2 *v* 4: Wilcoxon p = 0.005; 95% CI, −103.48 to −18.90; experiment 2 *v* 5: Wilcoxon p = 0.007; 95% CI, −105.17 to −26.67), probably owing to selection of different hot spots.

Finally, the reproducibility test of the automated method (experiment 6) on tumour T6 resulted in excellent agreement (95%), in that repeating the measurement resulted in approximately the same microvessel count. The small differences can be explained by different focusing and by microvessels on the borders of the fields of vision.

Discussion

Several well conducted retrospective studies have provided contradictory results on the prognostic value of tumour microvessel density in patients with invasive breast cancer.[7-23] These contradictory results are most probably caused by differences in methodology.[5 24]

Although the counting of microvessels has been made more objective by introducing morphometry[8 45] and image processing,[8 14 19 27-37] the most observer dependent and therefore most subjective step remained—the selection of the hot spot.

In the consensus report by Vermeulen and colleagues[40] and also in a paper by Barbareschi *et al*,[27] it was stated that automated hot spot selection would solve the problem of subjectivity caused by manual assessment of intratumour microvessel density. To our knowledge, there have been only two attempts to perform completely automated hot spot selection and counting of microvessels.[8 38] Fox *et al*,[8] who were unfortunately unsuccessful, stated that "their lack of success was probably due to the low tolerance required for their image analysis package to recognise solely the immunostaining." Van Der Laak *et al* were more successful.[38] They were able to select the three fields with the highest vessel density after semiautomated image analysis. Whether this selection was done manually or automatically is not clear. Although this last method is more objective, it does not, however, reflect the true nature of a hot spot (one or several fields connected to each other), in that the three fields might be spread out across the entire tumour. Although their results were promising, they also concluded that more reproducible immunohistochemical staining procedures were needed to reduce the somewhat interactive nature of their segmentation algorithm.

Thus, in order to develop a fully automated system for reproducible and objective selection of the hot spot together with microvessel counting, it is essential to use a microvessel staining technique that is not only reliable but is also optimally sensitive and specific for microvessels. Since routine formalin fixed, paraffin embedded specimens are mostly used, we employed JC70, an anti-CD31 antibody: on the basis of its sensitivity and specificity this is the optimal choice available.[39 40] Using additional incubation with biotinylated tyramine, the quality of the CD31 staining improves

markedly, in that the staining becomes more intense and crisp (fig 2), which facilitates automated analysis using image processing.

Based on this improvement in the staining method, we developed an automated method to identify microvessels in whole tumour sections by image processing. Using the coordinates of each microvessel, a geographical microvessel map is created which is used as the input for a dedicated hot spot search program (central part of fig 1). This program searches the map for vascular hot spots of user defined shape and size (fig 1, A–I). This allows fully automated, and therefore more objective, selection of hot spots and counting of microvessels, overcoming almost all the apparent disadvantages of (interactive) computer aided microvessel measurements so far as discussed by Barbareschi *et al* and Vermeulen *et al*.[27 40]

To determine whether the results obtained using the automated counting and hot spot selection method are more objective, we performed six experiments. First of all we assessed the number of microvessels manually in a selected hot spot to determine whether or not the counting procedure, based on Weidner's protocol, is itself reproducible. From the results of other studies and from those of experiments 1 and 2, we can state that when exactly the same fields are examined, trained observers in general arrive at comparable microvessel counts. If the selection of the hot spot is taken into account, different observers in general still count the same microvessels in exactly the same fields, but results by the same observer between two experiments may differ significantly.

Thus the first step in reducing subjectivity would be to count microvessels in the entire tumour area and select the hot spot based on those numbers. Experiment 3 shows that the number of microvessels within the hot spot selected by the hot spot program is in general higher than the number found using Weidner's manual method. This experiment also showed that the shape of a hot spot (a square or random path) may provide different results. Whether or not the shape of the hot spot has any impact on prognostic value is unclear and should be studied in a large group of patients. Another interesting result is that when a hot spot consisted of more than one field (in this case 10 fields), in four of the 10 cases analysed the hot spot search program found a hot spot which did not include the field with the maximum number of microvessels, indicating that manual selection of the hot spot, as done using the traditional method of Weidner, may be biased by looking for the field with the largest number of microvessels.

Although the results are encouraging and result in a more objective hot spot selection, counting the whole tumour area manually is not an attractive approach since it is time consuming, labour intensive, and requires constant attention of the observer. We therefore explored the possibilities of automating this task using image processing.

Again, when comparing hot spots of the same size but with different shapes, an eight connective random path resulted in the hot spot with highest microvessel counts per mm^2 (see fig 3), and in many cases the hot spot did not contain the field with the maximum number of microvessels. Although the staining method has been improved, the automated method scores relatively best (low false negative rate) on specimens that have been graded subjectively as the most intense, indicating that standardisation of immunohistochemical staining procedures remains necessary. The automated counting method without any corrections by an observer performs excellently compared with manual assessment using an automated scanning stage. However, a comparison of the absolute numbers of microvessels with the exact location of the hot spot within the tumour between experiments 4 and 5 showed that in two of 10 cases the hot spot search program assigned completely different locations to the hot spot. The differences found, however, were caused by only small differences in the numbers of microvessels per field, owing to imperfect segmentation of microvessels. Most of those imperfections result in the splitting of a single relatively large microvessel into several single unconnected fragments of that microvessel. Those are counted separately by the computer, where the human eye would count them as a single microvessel. Some of these errors may be solved in future releases of the program by providing a priori knowledge of microvessel structure and shape at the expense of throughput.

Weidner *et al* showed that the predictive value of microvessel density decreases when the field size is smaller than 0.19 mm^2.[17] Fox *et al* stated that "microvessel density is also determined by the tumour area measured, currently ranging from 0.12 to 0.74 mm^2. Too small an area will always be highly vascular and too large an area will dilute out the hot spot".[8] These latter phenomena are shown in fig 3. It is clear that, independently of shape, the use of different hot spot sizes may result in different rankings of the tumours according to the number of microvessels in that hot spot. This especially holds for tumours T4 and T6 which had different ranks at hot spot sizes of 0.08, 0.15, 0.23, and 0.30 mm^2 (fig 3).

However, given the reproducibility, the increased accuracy and objective selection of the hot spot, the possibility of visual inspection and relocation of each measurement field afterwards, and the reduction in measurement time compared with manual counting of the complete tumour, we conclude that the automated method of counting microvessels is preferable over the traditional manual method.

The throughput of the system was on average 3.5 h/cm^2, but preliminary tests using an optimisation of the current microvessel counting program have already shown that measurement time can be reduced to 1.4 h/cm^2 without loss of precision and accuracy. This optimisation will be implemented in the program, making the measurement time acceptable for future investigations which will have to focus on clinical evaluation of this

automated method. We are planning a large prospective study in breast cancer using the fully automated method, which will allow us to find the optimum size and shape of the hot spot. Of course, the method can and will be applied in a similar way to various other tumours.

In conclusion, we have developed a fully automated image processing method to identify microvessel hot spots and assess microvessel counts by scanning whole tumour sections. This method will allow large numbers of patients to be analysed, and will lead to more reproducible and more accurate microvessel counting and optimal prognostic classification for clinical trials.

In part supported by grant No 95-930 of the Dutch Cancer Society and by grant No 28-2015 of the Praeventiefonds.

1 Folkman J. Tumour angiogenesis: therapeutic implications. *N Engl J Med* 1971;**285**:1182–6.
2 Folkman J. Clinical applications of research on angiogenesis. *N Engl J Med* 1995;**333**:1757–63.
3 Rak J, Filmus J, Kerbel RS. Reciprocal paracrine interactions between tumor cells and endothelial cells: the "angiogenesis progression" hypothesis. *Eur J Cancer* 1996;**32A**: 2438–50.
4 Fidler IJ, Ellis LM. The implications of angiogenesis for the biology and therapy of cancer metastasis. *Cell* 1994;**79**: 185–8.
5 Weidner N. Intratumor microvessel density as a prognostic factor in cancer. *Am J Pathol* 1995;**147**:9–19.
6 Fox SB. Tumour angiogenesis and prognosis. *Histopathology* 1997;**20**:294–301.
7 Bosari S, Lee AKC, DeLellis RA, *et al.* Microvessel quantitation and prognosis in invasive breast carcinoma. *Hum Pathol* 1992;**23**:755–61.
8 Fox SB, Leek RD, Weekes MP, *et al.* Quantitation and prognostic value of breast cancer angiogenesis: comparison of microvessel density, Chalkley count, and computer image analysis. *J Pathol* 1995;**177**:275–83.
9 Gasparini G, Weidner N, Bevilacqua P, *et al.* Tumor microvessel density, p53 expression, tumor size, and peritumoral lymphatic vessel invasion are relevant prognostic markers in node-negative breast carcinoma. *J Clin Oncol* 1994;**12**: 454–66.
10 Horak ER, Leek R, Klenk N, *et al.* Angiogenesis, assessed by platelet/endothelial cell adhesion molecule antibodies, as indicator of node metastases and survival in breast cancer. *Lancet* 1992;**340**:1120–4.
11 Obermair A, Kurz C, Czerwenka K, *et al.* Microvessel density and vessel invasion in lymph-node-negative breast cancer: effect on recurrence-free survival. *Int J Cancer* 1995;**62**:126–31.
12 Ogawa Y, Chung Y-S, Nakata B, *et al.* Microvessel quantitation in invasive breast cancer by staining for factor VIII-related antigen. *Br J Cancer* 1995;**71**:1297–301.
13 Ravazoula P, Hatjikondi O, Kardamakis D, *et al.* Angiogenesis and metastatic potential in breast carcinoma. *Breast* 1996;**5**:418–21.
14 Simpson JF, Ahn C, Battifora H, *et al.* Endothelial area as a prognostic indicator for invasive breast carcinoma. *Cancer* 1996;**77**:2077–85.
15 Toi M, Kashitani J, Tominaga T. Tumor angiogenesis is a independent prognostic indicator in primary breast carcinoma. *Int J Cancer* 1993;**55**:371–4.
16 Weidner N, Folkman J, Pozza F, *et al.* Tumor angiogenesis: a new significant and independent prognostic indicator in early-stage breast carcinoma. *J Natl Cancer Inst* 1992;**84**: 1875–87.
17 Weidner N, Semple JP, Welch WR, *et al.* Tumor angiogenesis and metastasis—correlation in invasive breast carcinoma. *N Engl J Med* 1991;**324**:1–8.
18 Costello P, McCann A, Carney DN, *et al.* Prognostic significance of microvessel density in lymph node negative breast carcinoma. *Hum Pathol* 1995;**26**:1181–4.
19 Goulding H, Nik Abdul Rashid NF, Robertson JF, *et al.* Assessment of angiogenesis in breast carcinoma: an important factor in prognosis. *Hum Pathol* 1995;**26**:1196–200.
20 Hall NR, Fish DE, Hunt N, *et al.* Is the relationship between angiogenesis and metastasis in breast cancer real? *Surg Oncol* 1992;**1**:223–9.
21 Van Hoef MEHM, Knox WF, Dhesi SS, *et al.* Assessment of tumour vascularity as a prognostic factor in lymph node negative invasive breast cancer. *Eur J Cancer* 1993;**29A**: 1141–5.
22 Morphopoulos G, Pearson, Ryder WD, *et al.* Tumour angiogenesis as a prognostic marker in infiltrating lobular carcinoma of the breast. *J Pathol* 1996;**180**:44–9.
23 Siitonen SM, Haapasalo HK, Rantala IS, *et al.* Comparison of different immunohistochemical methods in the assessment of angiogenesis: lack of prognostic value in a group of 77 selected node-negative breast carcinomas. *Mod Pathol* 1995;**8**:745–52.
24 Page DL, Jensen RA. Angiogenesis in human breast carcinoma: what is the question? *Hum Pathol* 1995;**26**: 1173–4.
25 De Jong JS, Van Diest PJ, Baak JPA. Heterogeneity and reproducibility of microvessel counts in breast cancer. *Lab Invest* 1995;**73**:922–6.
26 Axelsson K, Ljung BME, Moore DH, *et al.* Tumor angiogenesis as a prognostic assay for invasive ductal breast carcinoma. *J Natl Cancer Inst* 1995;**87**:997–1008.
27 Barbareschi M, Weidner N, Gasparini G, *et al.* Microvessel density quantification in breast carcinomas: assessment by light microscopy vs. a computer-aided image analysis system. *Appl Immunohistochem* 1995;**3**:75–84.
28 Bigler SA, Deering RE, Brawer MK. Comparison of microscopic vascularity in benign and malignant prostate tissue. *Hum Pathol* 1993;**24**:220–6.
29 Brawer MK, Deering RE, Brown M, *et al.* Predictors of pathologic stage in prostatic carcinoma: the role of neovascularity. *Cancer* 1994;**73**:678–87.
30 Charpin C, Devictor B, Bergeret D, *et al.* CD31 quantitative immunocytochemical assays in breast carcinomas: correlation with current prognostic factors. *Am J Clin Pathol* 1995;**103**:443–8.
31 Kohlberger PD, Obermair A, Sliutz G, *et al.* Quantitative immunohistochemistry of factor VIII-related antigen in breast carcinoma: a comparison of computer-assisted image analysis with established counting methods. *Am J Clin Pathol* 1996;**105**:705–10.
32 Rucklidge GJ, Travis AJ. An automatic objective estimation of vascularization of normal and tumor-invaded brain tissue using image analysis. *Anal Quant Cytol Histol* 1989;**11**:286–90.
33 Smolle J, Soyer HP, Hofmann-Wellenhof R, *et al.* Vascular architecture of melanocytic skin tumors: a quantitative immunohistochemical study using automated image analysis. *Path Res Pract* 1989;**185**:740–5.
34 Srivastava A, Laidler P, Davies RP, *et al.* The prognostic significance of tumor vascularity in intermediate-thickness (0.76–4.0 mm thick) skin melanoma: a quantitative histologic study. *Am J Pathol* 1988;**133**:419–23.
35 Visscher DW, Smilanetz S, Drozdowicz S, *et al.* Prognostic significance of image morphometric microvessel enumeration in breast carcinoma. *Anal Quant Cytol Histol* 1993;**15**: 88–92.
36 Wakui S, Furusato M, Itoh T, *et al.* Tumour angiogenesis in prostatic carcinoma with and without bone marrow metastasis: a morphometric study. *J Pathol* 1992;**168**:257–62.
37 Wesseling P, Van Der Laak JAWM, De Leeuw H, *et al.* Quantitative immunohistological analysis of the microvasculature in untreated human glioblastoma multiforme: computer-assisted image analysis of whole-tumor sections. *J Neurosurg* 1994;**81**:902–9.
38 Van Der Laak JAWM, Westphal JR, Schalkwijk LJM, *et al.* An improved procedure to quantify tumour vascularity using true colour image analysis. Comparison with the manual hot-spot procedure in a human melanoma xenograft model. *J Pathol* 1998;**184**:136–43.
39 Parums DV, Cordell JL, Micklem K, *et al.* JC70: a new monoclonal antibody that detects vascular endothelium associated antigen on routinely processed tissue sections. *J Clin Pathol* 1990;**43**:752–7.
40 Vermeulen PB, Gasparini G, Fox SB, *et al.* Quantification of angiogenesis in solid human tumours: an international consensus on the methodology and criteria of evaluation. *Eur J Cancer* 1996;**32**:2474–84.
41 Adams JC. Biotin amplification of biotin and horseradish peroxidase signals in histochemical stains. *J Histochem Cytochem* 1992;**40**:1457–63.
42 Gundersen HJG. Note on the estimation of the numerical density of arbitrary profiles: the edge effect. *J Microsc* 1997; **111**:219–23.
43 Ten Kate TK, Beliën JAM, Smeulders AWM, *et al.* Method for counting mitoses by image processing in Feulgen stained breast cancer sections. *Cytometry* 1993;**14**:241–50.
44 Beliën JAM, Baak JPA, Van Diest PJ, *et al.* Counting mitoses by image processing in Feulgen stained breast cancer sections: the influence of resolution. *Cytometry* 1997;**28**: 135–40; 1997;**29**:189.
45 Chalkley H. Method for the quantitative morphological analysis of tissues. *J Natl Cancer Inst* 1943;**4**:47–53.