Journal of Microscopy, Vol. 235, Pt 1 2009, pp. 106–118
Received 27 March 2008; accepted 6 March 2009

Statistical analysis of labelling patterns of mammary carcinoma cell nuclei on histological sections

T. MATTFELDT*, S. ECKEL†, F. FLEISCHER‡ & V. SCHMIDT†
**Institute of Pathology, Ulm University, Ulm, Germany*
†Institute of Stochastics, Ulm University, Ulm, Germany
‡Medical Data Services/Biostatistics, Boehringer-Ingelheim Pharma GmbH & Co. KG, Ingelheim, Germany

Key words. Computer-intensive methods, diversity, K-function, mammary carcinoma, pair correlation function, pathology, point process, Simpson index, spatial statistics, stereology, stochastic geometry.

Summary

It is of central interest for tumour biology to explore the mechanisms of tumour cell proliferation. In this study, methods of spatial statistics were used to study the spatial distribution of proliferating cells within tumour tissue quantitatively and objectively. Mammary cancer tissue was studied as an example. It was attempted to clarify whether cell division occurs entirely at random (random labelling), i.e. the process of division occurs at random, independently from the state of the neighbouring nuclei, or whether the spatial distribution of proliferation is more complex, e.g. in the form of actively proliferating clusters alternating with relatively silent zones. In the case of random labelling, the reduced second moment functions $K(r)$ of the labelled and the unlabelled nuclei would be identical. The same would hold for the pair correlation functions $g(r)$. The alternative hypothesis is that the second-order properties of the processes of the labelled and the unlabelled nuclei are different. Twenty cases of invasive ductal mammary carcinomas were studied. The nuclei of proliferating cells were stained immunohistochemically with the monoclonal antibody MIB-1, which detects specifically the proliferation-associated nuclear antigen Ki 67. The planar coordinates of the tumor cell nucleus profiles from two rectangular visual fields per case were recorded. For each visual field, the following investigations were performed: estimation of the explorative summary characteristics $K(r)$ and $g(r)$, fitting of the parameters of a stationary Strauss hard-core model to the observed point patterns, estimation of two distance-dependent Simpson indices and Monte Carlo tests of all individual patterns on the null hypothesis of random labelling. Significant differences between the mean K-functions and the mean g-functions of the labelled and the unlabelled nuclei were found. Moreover, the mean interaction parameter γ of the stationary Strauss hard-core model was significantly higher for the labelled nuclei than for the unlabelled nuclei. The estimates of the two distance-dependent Simpson indices showed a tendency of points with the same label towards a positive spatial correlation. In the Monte Carlo tests, the null hypothesis of random labelling was rejected for the majority of the visual fields. These four lines of investigation led to the concordant conclusion that the labelling of mammary carcinoma nuclei by MIB-1 is not simply random. The data suggest that the second-order properties of the point process of the labelled nuclei are significantly different from those of the unlabelled nuclei. In particular, the process of the labelled nuclei shows a higher degree of clustering (increased strength of interaction) than the process of the unlabelled points.

Introduction

In tumour biology, the proliferative behaviour of the neoplastic cells in malignant tumours is of central interest. High proliferative activity of the tumour cells is often associated with rapid tumour progression and poor prognosis. A major target of tumour therapies, such as anti-neoplastic agents and irradiation, is a significant decrease or total stop of tumour cell proliferation. Hence, it is of fundamental scientific importance to understand the principles that are governing tumour cell proliferation. This process can be studied on the sub-cellular level using methods of molecular biology and biochemistry. Another basic approach is the morphological (microscopical) study of cell division. Here, one wants to find out how

Correspondence to: Prof. Dr. T. Mattfeldt. Tel: +49 731 500 56322; fax: +49-731-500-56384; e-mail: torsten.mattfeldt@uni-ulm.de

large the fraction of dividing cells is and in which manner the proliferative activity is spatially distributed within the tissue.

In diagnostic histopathology, it is a routine application to use immunohistochemical stains to study the labelling pattern of the tumour cell nuclei. Using such methods, it is possible to classify observed tumour cell nucleus profiles on histological sections in a binary manner into two categories, i.e. stained versus unstained nuclei (labelled vs. unlabelled and positive vs. negative nuclei). In the domain of tumour biology, a popular immunohistochemical marker is the MIB-1 stain (Gerdes *et al.*, 1992). In this method, a monoclonal antibody versus the antigen Ki 67 is used, which is specifically expressed in the nuclei of proliferating (dividing) cells, whereas the nuclei of non-dividing cells are negative (Figs 1a and b). In diagnostic histopathology, it is usually sufficient to perform a semiquantitative estimation of the percentage of positively labelled nucleus profiles. For a closer evaluation of such images, spatial statistics may help in two ways. First, the percentage of positively labelled nucleus profiles may be objectively counted (quantitation) after interactive or automatic segmentation of all nuclei. Second, spatial statistics may help to clarify whether the proliferation is governed by certain deterministic or stochastic principles, or, alternatively, whether the process of cell division occurs purely at random in the tumour cells. For example, it would be plausible to assume that proliferation starts from clusters (nests) of actively proliferating cells, whereas other regions scarcely divide. In some cases, e.g. in the follicles of the lymph nodes, this clustering is so obvious that it can be discerned by microscopical inspection without sophisticated statistical analyses.

In malignant epithelial tumours arising from glands (adenocarcinomas), the centres of the tumour cell nucleus profiles on sections may usually be considered as stationary and isotropic planar point processes. Random labelling means here that the binary marks (label 0 or 1) are ascribed to the points independently at random (Diggle, 2003, pp. 48–49). Mammary carcinoma (breast cancer) is the most frequent malignant tumour in women. It was the aim of this study to find out whether the null hypothesis of random labelling applies to the proliferating subset of the tumour cells in the most frequent type of mammary carcinoma, i.e. invasive ductal mammary cancer. Usually, the fraction of MIB-1-labelled nuclei is rather low in such tumours (5–10%), and by visual inspection alone, one cannot safely determine whether they occur randomly in the tissue or in groups (Figs 1a and b). Hence, further progress necessitates the use of methods of spatial statistics here. In the case of random labelling, the reduced second moment functions $K(r)$ of the labelled and the unlabelled points would be identical (Diggle, 2003, pp. 48–49). The same would hold for the pair correlation functions $g(r)$. From these reasons, it was decided to focus the attention on these two second-order summary statistics. In the related

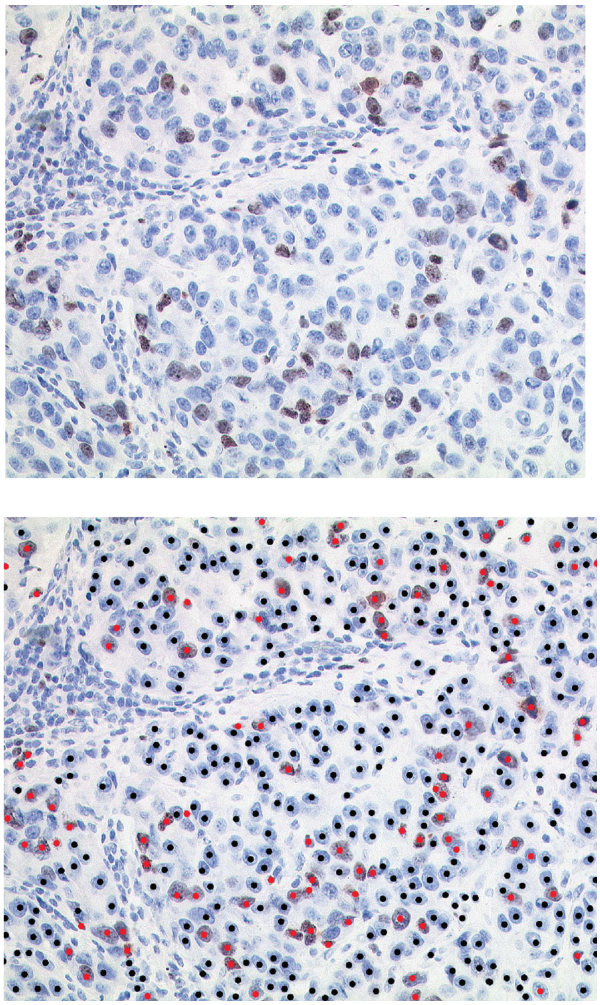

Fig. 1. (a) Histological section of a case of an invasive ductal mammary carcinoma. By means of the immunohistochemical MIB-1 stain, the nuclei expressing the proliferation-associated antigen Ki-67 are specifically stained. As a result, the dividing nuclei are stained brown (positive). The non-dividing nuclei remain in the background colour, blue (negative). (b) Same visual field as depicted in Fig. 1(a) but here the nuclei have already been detected (marked). The unlabelled nuclei (blue) are marked with black dots; the labelled nuclei (brown) are marked with red dots.

context of the independence of two point processes, several studies on replicated point patterns arising in microscopy have been published recently (Diggle *et al.*, 2006; Eglen *et al.*, 2006; Webster *et al.*, 2006).

Materials and methods

Materials

Twenty cases of invasive mammary ductal carcinomas were studied by light microscopy using paraffin sections. The best section in each case containing tumour tissue was selected for

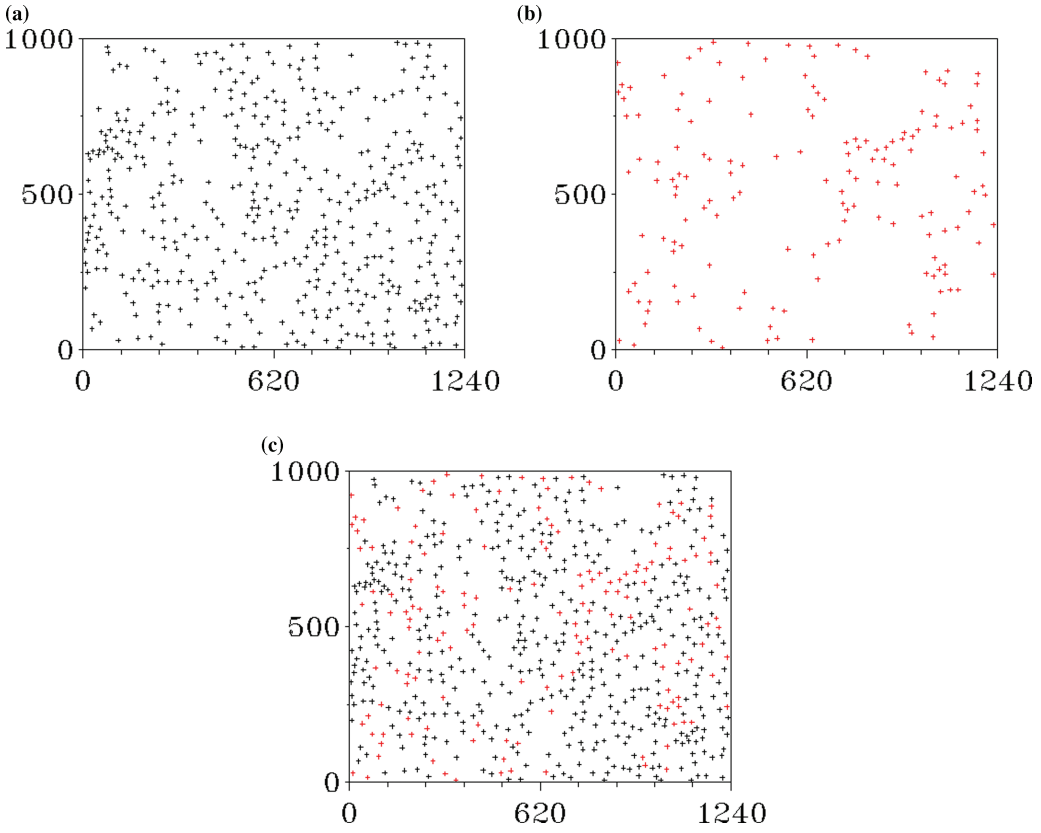

Fig. 2. (a) Point pattern of the unlabelled (negative) nuclei in a window of the size 1240 × 1000 pixels. (b) Same as Fig. 2(a) but for the labelled (positive) nuclei. (c) The complete marked point pattern is shown, with the coordinates of the unlabelled nuclei as black crosses and the coordinates of the labelled nuclei as red crosses.

the MIB-1 stain according to technical quality criteria. Two rectangular visual fields of the size 1240 × 1000 pixels per case were recorded and directly stored in computer memory after image acquisition with a CCD camera. This size amounts to 440 μm × 354 μm at the level of the tissue. The planar coordinates of the centres of tumor cell nucleus profiles were recorded (384–1387 points per field, of which 3–27% were labelled) together with the marks (Figs 2a–c).

For each visual field, the investigations were performed as described in sections 'Explorative Analysis of Point Patterns', 'Parametric Modelling', 'Distance-Dependent Characteristics of Diversity' and 'A Monte Carlo Rank Test on Random Labelling' below. First, the patterns of the unlabelled and the labelled points were considered separately. In section 'Explorative Analysis of Point Patterns', we consider an explorative analysis of the two kinds of point patterns, i.e. they are explored without assuming any particular point process model. In section 'Parametric Modelling', parametric modelling of both types of patterns was performed on the basis of Strauss hard-core processes (Mattfeldt *et al.*, 2007). Thereafter, computations are described in which the unlabelled and the labelled points were considered simultaneously for each field. For this purpose, we suggest using two distance-dependent Simpson indices (section 'Distance-Dependent Characteristics of Diversity') and a Monte Carlo rank test, in which the integral deviation between the K-functions of the unlabelled and the labelled points is used as test statistics (section 'A Monte Carlo Rank Test on Random Labelling').

Explorative analysis of point patterns

Usually, exploratory methods of data analysis are the first step to characterize a planar point process quantitatively. The most basic information is an estimate of the intensity λ of the point process, i.e. the mean number of points per area. If $X = \{X_n\}$ is a stationary and isotropic point process and W is a sampling window, then $\lambda = \frac{E(X(W))}{|W|}$, where $X(W)$ is the number of points of X in W, and $|W|$ denotes the area of W. A natural estimator for the intensity λ is given by $\widehat{\lambda} = \frac{X(W)}{|W|}$. Although the intensity is a single quantity, first-order functions and second-order functions (summary statistics) provide a series of values in which the latter ones are functions of the inter-point distance r. As examples of first-order functions, let us mention

the nearest-neighbour distribution function, the empty-space function and the *J*-function that can be used to study the spatial structure of a point pattern. Moreover, the *K*-function and the pair correlation function are examples of second-order functions (see, e.g. Stoyan *et al.*, 1995; Diggle, 2003; Illian *et al.*, 2008). In particular, Ripley's *K*-function $K(r)$ (reduced second moment function) is one of the most popular functions of explorative spatial point pattern analysis (Ripley, 1988; Stoyan *et al.*, 1995; Illian *et al.*, 2008). Intuitively, $K(r)$ is the mean number of other points of the process lying within a circle of radius r, centred about a typical point of the process, divided by the intensity of the process:

$$K(r) = \frac{E(\text{number of other points of } X \text{ within distance } r \text{ from } (x, y) \mid X \text{ has point at } (x, y))}{\lambda}, \quad (1)$$

where the symbol '|' denotes 'conditional to'. In analogy to a probability density function, which is the derivative of a cumulative distribution function, there is a counterpart to the *K*-function, namely the 'pair correlation function' $g(r)$, which may be obtained after differentiation and normalization of $K(r)$:

$$g(r) = \frac{1}{2\pi r}\frac{dK(r)}{dr}. \quad (2)$$

In the case of a planar Poisson point process, we obtain $g(r) \equiv 1$ for all r. Values above 1 indicate clustering, whereas values below 1 are related to regularity in the pattern. The pair correlation function may also be defined as the product density of second order of the point process divided by the square of the intensity for the purpose of normalization (Stoyan & Stoyan, 1994, p. 249; Stoyan *et al.*, 1995, p. 129; Illian *et al.*, 2008, p. 219). Hence, another estimator for $g(r)$ is given by

$$\widehat{g}(r) = \frac{\widehat{\varrho^{(2)}}(r)}{\widehat{\lambda}^2}, \quad (3)$$

where

$$\widehat{\varrho^{(2)}}(r) = \frac{1}{2\pi r}\sum_{X_i, X_j \in W_{i \neq j}} \frac{k_h(r - ||X_i - X_j||)}{|W_{X_i} \cap W_{X_j}|} \quad (4)$$

is an estimator for $\varrho^{(2)}(r)$, the product density of second order. In the estimator $\widehat{\varrho^{(2)}}(r)$, the denominator $|W_{X_i} \cap W_{X_j}|$ is used for edge correction, where $W_{X_i} = W + X_i$ denotes the sampling window shifted by X_i (see Stoyan & Stoyan, 1994, p. 284). The term $k_h(x)$ denotes a kernel function, which is used for smoothing. We used the Epanechnikov kernel

$$k_h(x) = \frac{3}{4h}\left(1 - \frac{x^2}{h^2}\right)1_{(-h,h)}(x), \quad (5)$$

with a bandwidth $h = 0.1/\sqrt{\widehat{\lambda}}$ according to Krasnoperov & Stoyan (2004). The estimation of the *K*-function was performed using the translation-corrected estimator according to eq. (15.11) in Stoyan & Stoyan (1994) (see also Illian *et al.*, 2008, p. 228, eq. (4.3.27)). Our estimator of $g(r)$ in Eqs (3) and (4) is the same as in eq. (15.15) in Stoyan & Stoyan (1994) (see also Illian *et al.*, 2008, p. 230, eq. (4.3.29)). For the explorative point pattern analysis, the software package SPATSTAT was used (Baddeley & Turner, 2005, 2006; Mattfeldt *et al.*, 2006, 2007).

In addition to these computations, each estimated *g*-function per visual field was evaluated with a method published by Stoyan & Schnabel (1990) (see also Stoyan & Stoyan, 1994, pp. 250–258; Mattfeldt *et al.*, 2006; Illian *et al.*, 2008, p. 241). This procedure includes identification of the first maximum $g_{\max}$ and the next following minimum $g_{\min}$, with the corresponding r values $r_{\max}$ and $r_{\min}$ for each *g*-function, where $r_{\min} > r_{\max}$. From these data, the statistic

$$M = \frac{g_{\max} - g_{\min}}{r_{\min} - r_{\max}} \quad (6)$$

is computed (Stoyan & Stoyan, 1994, p. 251; Illian *et al.*, 2008, p. 241). The statistic M is related to the global degree of order in the spatial point pattern. Large values indicate a high degree of order and may be expected, e.g. in the case of point patterns with an element of long-range order. The statistic may be used as a non-parametric index to summarize the course of the *g*-function by a single quantity. Even simpler is the difference value $\Delta_g = g_{\max} - g_{\min}$.

Ultimately, the methods of this section lead to the estimation of four *g*-functions per case (two patterns with *g*-functions related to the unlabelled and two to the labelled points, respectively). From these four estimates, the two mean *g*-functions were computed for each case, with equal weights ascribed to the estimates from both patterns. After averaging in this manner, 20 mean *g*-functions for the unlabelled and the labelled points were obtained. Another possibility would have been to compute a weighted mean and have the number of points as weights. It was attempted to test the mean *g*-functions of the unlabelled points and the labelled points *r*-wise for significant differences. As the *g*-functions are estimated from the same visual fields, one has to deal with paired observations. As a normal distribution of the differences between the values of the *g*-functions of the two types for a given r could not be safely surmised, the non-parametric Wilcoxon matched-pairs signed-rank test (Sachs, 2003, pp. 411–413) was used instead of a paired *t*-test. As these rank tests were performed for 22 selected values of r (see Table 1), the significance levels were adjusted by applying the Bonferroni correction. A result was considered as significant at the level of 0.05 when the P value was less than $0.05/22 = 0.0023$.

Parametric modelling

Gibbs processes (Markov point processes) are flexible models for point processes with interaction (Stoyan & Stoyan, 1994; Stoyan *et al.*, 1995; Baddeley & Turner, 2000; Van Lieshout,

Table 1. Local comparisons of *g*-functions. For the *r* values in the left column, the mean values of the *g*-functions are given for the unlabelled and the labelled nuclei.

r	Unlabelled $\hat{\bar{g}}(r)$	Labelled $\hat{\bar{g}}(r)$	Difference	*P* value
5	0.00000	0.29258	0.29258	<0.001
10	0.00127	0.70074	0.69947	>0.05
15	0.24768	1.33178	1.08410	<0.001
20	1.13593	1.94351	0.80758	<0.001
25	1.46384	2.33754	0.87370	<0.001
30	1.35353	2.29150	0.93797	<0.001
35	1.22940	2.00732	0.77792	<0.001
40	1.17549	1.77813	0.60265	<0.001
45	1.16746	1.61141	0.44395	>0.05
50	1.15192	1.50813	0.35621	>0.05
55	1.13997	1.42883	0.28886	>0.05
60	1.13603	1.35491	0.21888	>0.05
65	1.12911	1.30302	0.17391	>0.05
70	1.11672	1.28056	0.16385	>0.05
75	1.10675	1.27387	0.16712	>0.05
80	1.09886	1.28003	0.18117	>0.05
85	1.10246	1.26078	0.15832	>0.05
90	1.10575	1.22666	0.12091	>0.05
95	1.08833	1.17965	0.09132	>0.05
100	1.10102	1.13606	0.03504	>0.05
150	1.05846	1.13965	0.08120	>0.05
200	1.04896	1.07587	0.02691	>0.05

2002; Diggle, 2003; Møller & Waagepetersen, 2004; Illian *et al.*, 2008). In the present context, we decided to use the 'stationary Strauss hard-core point process' as a candidate model because it takes into account a hard-core property and allows repulsion as well as clustering of the points at different domains of *r* values, depending on the model parameters (see Takacs & Fiksel, 1986; Diggle *et al.*, 1994; Goulard *et al.*, 1996; Baddeley & Turner, 2000; Mattfeldt *et al.*, 2007; Illian *et al.*, 2008, and references therein). The classical Strauss model and the Strauss hard-core model are both examples of the Gibbs processes. These processes may be defined in terms of their pair potential or of their probability density (Takacs & Fiksel, 1986; Goulard *et al.*, 1996; Baddeley & Turner, 2000; Illian *et al.*, 2008). The probability density of the stationary Strauss hard-core process in a sampling window W (e.g. of unit area) is given by

$$f(\mathbf{x}|(r_0, R, \beta, \gamma)) = \alpha\beta^{n(\mathbf{x})}\gamma^{s_R(\mathbf{x})}\mathbf{1}_{\{|x_i-x_j|>r_0:\forall\{x_i,x_j\}\subseteq\mathbf{x},\ x_i\neq x_j\}}, \quad (7)$$

where $n(\mathbf{x})$ is the number of points in the pattern $\mathbf{x} = \{x_i\}$, and $s_R(\mathbf{x})$ is the number of distinct unordered pairs of points that have a distance to each other that is less or equal to R. Note that α is a (usually intractable) normalizing constant and that the Strauss hard-core process can be completely defined by its four model parameters r_0, R, β and γ, where r_0 is the hard-core distance, R is the interaction range (interaction radius), β is a constant factor contributed by each point to the probability density and related to the intensity and γ is the strength of interaction (see Baddeley & Turner, 2005; Mattfeldt *et al.*, 2007). The meaning of the indicator function $\mathbf{1}$ in Eq. (7) is that it becomes 0 if the point configuration $\mathbf{x}$ contains at least one point pair $\{x_i, x_j\} \subseteq \mathbf{x}$ with a distance less or equal to r_0, otherwise the indicator function equals 1. For point pairs of a distance between r_0 and R, values of $\gamma > 1$ indicate clustering, whereas $\gamma < 1$ is related to regularity. If $\gamma = 1$, a pure hard-core process is obtained. For point pair distances larger than R, there is no more pairwise interaction. Note that contrary to the case of a classical Strauss process, for Strauss hard-core processes, the interaction parameter γ can assume any non-negative value, in particular, a value larger than 1.

For practical modelling, the software package SPATSTAT (Baddeley & Turner, 2005, 2006) was used with R 2.2.0 under Linux. The tumour cell nucleus profile's mid-point coordinates of all 40 images were read into a computer and the fitting procedure described below was performed for each individual image. The hard-core distance r_0 was estimated for each visual field as the minimum value of the observed inter-point distances, which is a maximum likelihood estimator. For simulations of the hard-core Strauss process, the largest integer number below this value was used as model parameter for r_0. The interaction radius R was estimated according to the profile pseudo-likelihood method (Baddeley & Turner, 2005, 2006). This procedure was used to find the value of R between 20 and 100 pixels in steps of 1 pixel with the maximum pseudo-likelihood for a given image. Edge correction was performed by translation correction (for details, see Ohser, 1983). The estimated value of R was then used together with that of r_0 for the subsequent model fitting, which ultimately yielded an estimate of γ (Baddeley & Turner, 2005, 2006; Mattfeldt *et al.*, 2007). For comparison between the model parameters estimated pairwise per case for the two classes of point patterns, the non-parametric Wilcoxon matched-pairs signed-rank test was used (Sachs, 2003).

Distance-dependent characteristics of diversity

A well-established statistical measure for the degree of diversity of a marked point process is the global Simpson index D. Let us consider a marked random point process with m types of points. The Simpson index is generally defined as follows:

$$D = 1 - \sum_{i=1}^{m} \frac{\lambda_i^2}{\lambda^2}. \quad (8a)$$

For the special case of two types of marks, which we consider here, we have $m = 2$, hence:

$$D = 1 - \frac{\lambda_1^2 + \lambda_2^2}{\lambda^2}, \quad (8b)$$

where λ_1 and λ_2 are the intensities of the unlabelled and the labelled points, respectively. For $m = 2$, the index reaches its maximum value if $\lambda_1 = \lambda_2 = \lambda/2$, and in this case, we obtain $D = 0.5$. The larger the difference between the two intensities, the lower becomes D. A high D value indicates a state near the equilibrium of the intensities of the two kinds of points. If one of the two point types dominates strongly, D is lowered as compared with the maximum value 0.5. The diversity is then said to be low, as only few point pairs will have points of different kinds in this situation, just as 'diversity' in everyday language indicates the amount of mixing of different types of things. In the form of Eqs (8a) and (8b), D is a distance-independent (global) characteristics of diversity. It is easy to see that D is the distance-independent probability to select a point pair at random belonging to different components because we have $D = 1 - \lambda_1^2/\lambda^2 - \lambda_2^2/\lambda^2$. In our material, we found, on average, approximately 10.7% of the tumor cell nucleus profiles labelled and 89.3% of the profiles unlabelled, leading to $D = 1 - 0.893^2 - 0.107^2 = 0.191$.

The concept of a global Simpson index has been generalized to the distance-dependent Simpson indices (Shimatani, 2001; Eckel *et al.*, 2007). Let us first consider the summary statistics $\alpha(r)$ for marked point processes with m types of points:

$$\alpha(r) = 1 - \sum_{i=1}^{m} \frac{\lambda_i^2 K_{ii}(r)}{\lambda^2 K(r)}, \tag{9a}$$

which reduces for the case $m = 2$ to

$$\alpha(r) = 1 - \frac{\lambda_1^2 K_{11}(r)}{\lambda^2 K(r)} - \frac{\lambda_2^2 K_{22}(r)}{\lambda^2 K(r)}, \tag{9b}$$

where $K_{11}(r)$ is the K-function of the unlabelled points, $K_{22}(r)$ is the K-function of the labelled points and $K(r)$ is the K-function of all points. This function is related to the probability to select a point pair at random belonging to different components conditional to the event that it has distance less than r. Under random labelling, we have $K_{11}(r) = K_{22}(r) = K(r)$ (Diggle & Chetwynd, 1991; Diggle, 2003); hence, $\alpha(r) \equiv D$ for all values of r. If $\alpha(r) < D$, the point pattern has a smaller diversity at distances r than in the case of random labelling. In other words, the points are less mixed than one would expect if the labelling were random. If $\alpha(r) < D$, this indicates that the two classes of point processes are more clustered together with their own species than in the case of pure random labelling. This behaviour was observed for our data and, in general, in cell biology, it seems more relevant than $\alpha(r) > D$. Theoretically, however, the case $\alpha(r) > D$ is also conceivable. It would mean that the point pattern has a larger diversity at distances r than in the case of random labelling. This would imply that cell types with different marks are more often associated than in the case of random labelling, which seems unlikely for biological cells.

In the same spirit, another distance-dependent indicator of diversity has been defined (Eckel *et al.*, 2007):

$$\beta(r) = 1 - \sum_{i=1}^{m} \frac{\lambda_i^2 g_{ii}(r)}{\lambda^2 g(r)} \text{ and} \tag{10a}$$

$$\beta(r) = 1 - \frac{\lambda_1^2 g_{11}(r)}{\lambda^2 g(r)} - \frac{\lambda_2^2 g_{22}(r)}{\lambda^2 g(r)}, \tag{10b}$$

where $g_{11}(r)$ is the g-function of the unlabelled points, $g_{22}(r)$ is the g-function of the labelled points and $g(r)$ is the g-function of all points. The quantity $\beta(r)$ is related to the probability to select a point pair at random belonging to different components conditional to the event that it has distance r. Under random labelling, we obtain $\beta(r) \equiv D$ for all r. If $\beta(r) < D$, the point pattern has a smaller diversity at distance r than in the case of random labelling. If $\beta(r) > D$, the point pattern has a larger diversity at distance r than in the case of random labelling. For the estimation of the distance-dependent Simpson indices, the software package GeoStoch, a Java-based open-library system, was used (Mayer *et al.*, 2004). To compute 90% confidence intervals for $\alpha(r)$ and $\beta(r)$, 1000 simulations of random labelling were performed for each real pattern.

A Monte Carlo rank test on random labelling

In this section, we suggest a simulation-based test on random labelling for all the 40 individual marked point patterns. We proceed again from the well-known fact that in the case of random labelling, we have $K_{11}(r) = K_{22}(r) = K(r)$ (Diggle & Chetwynd, 1991; Diggle, 2003, pp. 48–49). Hence, a natural test statistic to test for random labelling should be the integral deviation of the estimated K-functions for the unlabelled and the labelled points of a binary point pattern. It may be estimated as follows:

$$T_K = \sum_{l=1}^{s} |\widehat{K_{11}}(r_l) - \widehat{K_{22}}(r_l)|\, \Delta r, \tag{11}$$

where Δr is the increment of r values at which the K-functions are computed; here, $\Delta r = 1$. The larger is T_K, the stronger is the deviation of the marked point pattern from random labelling. To implement this idea in practice, 999 realizations of the empirical point pattern based on independent labelling were simulated. Given the observed locations (the coordinates) and the observed values of the labels, the new labels were assigned at random to the locations, which were kept unchanged. The number of labelled and unlabelled points was held constant, i.e. random labelling was simulated by permutation (Fig. 3). For each simulation, the functions $K_{11}(r)$ and $K_{22}(r)$ were estimated at the distances $r_1, \ldots, r_s$. In practice, we estimated $K_{11}(r)$ and $K_{22}(r)$ from $r_1 = 1$ to $r_s = 50$ pixels in steps of $\Delta r = 1$ pixel (see Fig. 4). Then, the values of the test statistic T_K for the 999 simulations were computed according to Eq. (11). In the same manner, T_K was computed for the single real pattern. The resulting $999 + 1 = 1000$ values of T_K

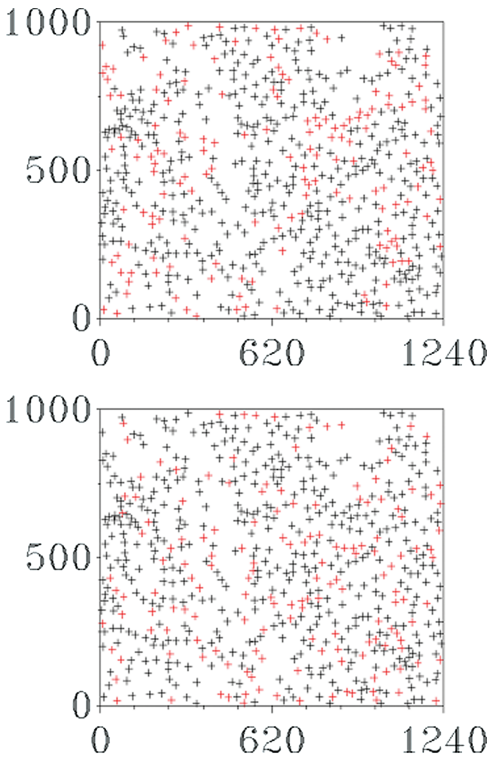

Fig. 3. Simulation of random labelling. Upper panel: a real pattern with unlabelled points in black and labelled points in red. Lower panel: simulation of random labelling applied to the same field.

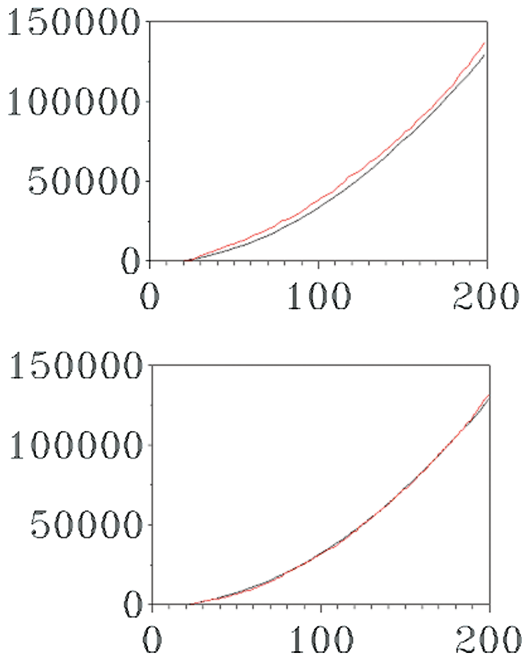

Fig. 4. Further illustration of simulation of random labelling. Upper panel: the two *K*-functions for the real pattern (black: unlabelled nuclei; red: labelled nuclei). Lower panel: the two *K*-functions obtained from a simulation of random labelling of the same pattern.

were sorted by size in ascending order. The rank of the value of T_K originating from the real pattern in this sequence was determined. The hypothesis of random labelling was rejected for the pattern if the rank of this T_K value was $\in [951, 1000]$. For the simulations needed for the Monte Carlo rank tests, the software package SPATSTAT was used (Baddeley & Turner, 2005, 2006).

Results

Explorative point pattern analysis

For practically all 40 patterns, it was visually obvious that the *g*-functions of the unlabelled and the labelled nuclei were different. Usually the *g*-function of the labelled points rose steeper and attained a higher first maximum than the *g*-function of the unlabelled points of the same pattern (Fig. 5a). This difference was significant for the mean *g*-functions per class, estimated by *r*-wise averaging between the cases (Fig. 5b). Statistical comparison using the Wilcoxon matched-pairs signed-rank test disclosed significant differences between the mean g values for $r = 5$ and $r = 15$ to $r = 40$ pixels (Table 1). Note distinctly higher mean values of $g(r)$ for the labelled nuclei at small distances r. For higher r values, significant differences could not be found any more. Analogous changes were found with respect to the mean *K*-functions of both processes (Fig. 6). Moreover, significant differences were found between the *g*-functions of the unlabelled and the labelled nuclei in terms of various non-parametric summary characteristics, i.e. $g_{\max}$, $r_{\min}$ and Δ_g (see Table 2). No significant differences were found with respect to $r_{\max}$, $g_{\min}$ and M. The intensity of the labelled points was much lower than the intensity of the unlabelled points (see Table 3). This difference was accompanied by a significantly higher hard-core distance for the process of the labelled points.

Parametric modelling

The model parameters of the two groups are shown in Table 3. The interaction parameter γ of the Strauss hard-core model was strongly increased for the labelled points as compared with the unlabelled points. By contrast, there was no significant difference with respect to the interaction distance R between the two kinds of point processes.

Distance-dependent Simpson indices

The estimates of the two mean distance-dependent Simpson indices are shown in Figs 7(a) and (b). Both show a concordant behaviour. The curves lie distinctly below the horizontal line at D, which corresponds to random labelling. Hence, for small

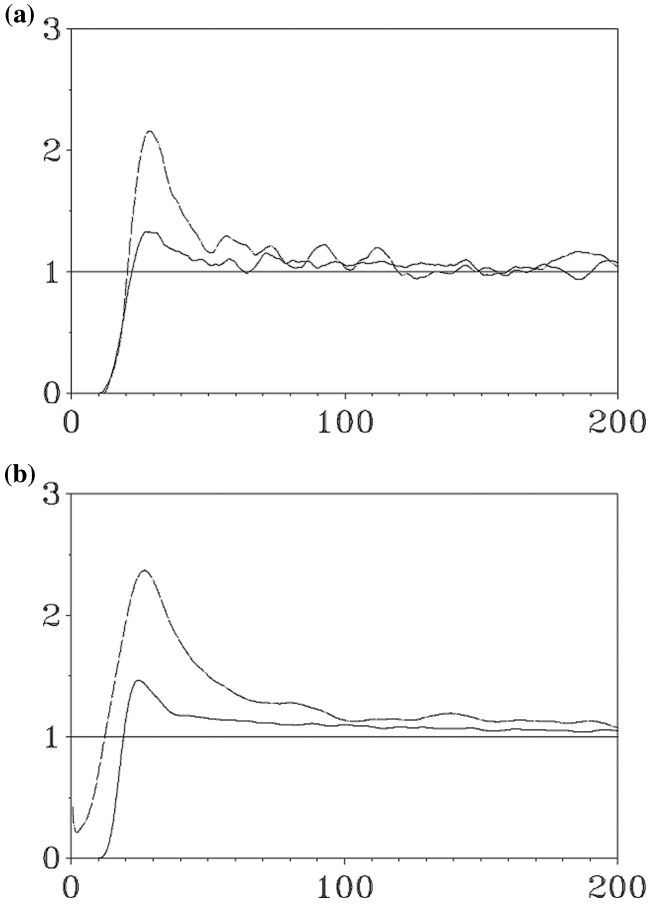

Fig. 5. (a) Pair correlation functions obtained from a selected visual field. Pair correlation function for the unlabelled nuclei: solid; for the labelled nuclei: dashed. (b) Mean pair correlation functions of all 20 cases for the unlabelled nuclei: solid; for the labelled nuclei: dashed.

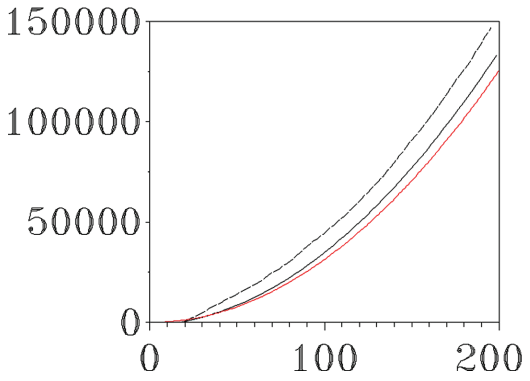

Fig. 6. Mean K-functions for the two types of nuclei. Mean K-function of the unlabelled nuclei: continuous; mean K-function of the labelled nuclei: dashed. The Poisson case is shown in red for comparison.

Table 2. Comparison of labelled and unlabelled nuclei in terms of explorative statistics. Summary characteristics of g-functions are given as mean values for the unlabelled and the labelled nuclei. The symbols r_{max}, g_{max}, r_{min}, g_{min}, M and Δ_g are explained in section 'Explorative Analysis of Point Patterns'. The difference between the mean values of these quantities was tested for significance by the Wilcoxon matched-pairs signed-rank test (right column). $\bar{x}$: mean value, SD: standard deviation.

	Unlabelled nuclei		Labelled nuclei		
Estimate	$\bar{x}$	SD	$\bar{x}$	SD	P value
r_{max} (pixel)	25.70	3.55	27.13	5.66	>0.05
g_{max}	1.582	0.295	2.559	0.655	<0.001
r_{min} (pixel)	40.46	6.53	53.18	12.69	<0.001
g_{min}	1.113	0.162	1.244	0.377	>0.05
M	0.035	0.017	0.051	0.030	>0.05
Δ_g	0.469	0.191	1.314	0.699	<0.001

distances r, the mixed point pattern shows less diversity than one would expect for random labelling. This difference is significant, as the mean curves lie outside a 90% pointwise confidence interval in which $\alpha(r)$ and $\beta(r)$ should lie in the case of random labelling. Note that the difference from D is significant but not very large for both indices.

The Monte Carlo rank tests

The Monte Carlo rank tests yielded a significant result for the majority of the 40 individual patterns. The null hypothesis of random labelling was rejected for 26 patterns and accepted for 14 patterns.

Discussion

Labelling pattern of the proliferating tumour cell nuclei

All four lines of investigation led to the concordant conclusion that the labelling of mammary carcinoma nuclei by MIB-1 does not simply result from a random labelling of the nuclei. A battery of methods was used for our study, which gives more reliability than a test based on a single criterium. The data suggest that the second-order properties of the point process of the labelled nuclei are significantly different from those of the unlabelled nuclei. This finding is not compatible with random labelling. In particular, the process of the labelled nuclei shows a higher degree of clustering for low r values than the process of the unlabelled points. This can be concluded from the higher values of g_{max} and Δ_g for the labelled points. From the viewpoint of parametric modelling, the stronger clustering tendency of the labelled points is corroborated by a highly significant increase of the interaction parameter γ of the Strauss hard-core model, which indicates the strength of interaction between the points. Also, in previous investigations on another biomedical point process, we found that changes of g_{max} and γ went into the same direction (Mattfeldt *et al.*, 2007).

Table 3. Comparison of unlabelled and labelled nuclei in terms of parametric modelling. The estimated model parameters of the Strauss hard-core point processes are given as the mean values for the unlabelled and the labelled nuclei. The symbols λ, r_0, R and γ are explained in section 'Parametric Modelling'. The difference between the mean values of these quantities was tested for significance by the Wilcoxon matched-pairs signed-rank test (right column). N(nucl/field): number of nucleus profiles per visual field, $\bar{x}$: mean value, SD: standard deviation.

	Unlabelled nuclei		Labelled nuclei		
	$\bar{x}$	SD	$\bar{x}$	SD	P value
Intensity					
N (nucl/field)	741	217	89	47	<0.001
λ (points/pixel^2)	0.0005957	0.000175	0.00007177	0.0000379	<0.001
Strauss hard-core model					
r_0 (pixel)	14.75	1.18	18.25	1.81	<0.001
R (pixel)	39.12	18.94	44.52	14.28	>0.05
γ	0.874	0.334	3.164	2.010	<0.001

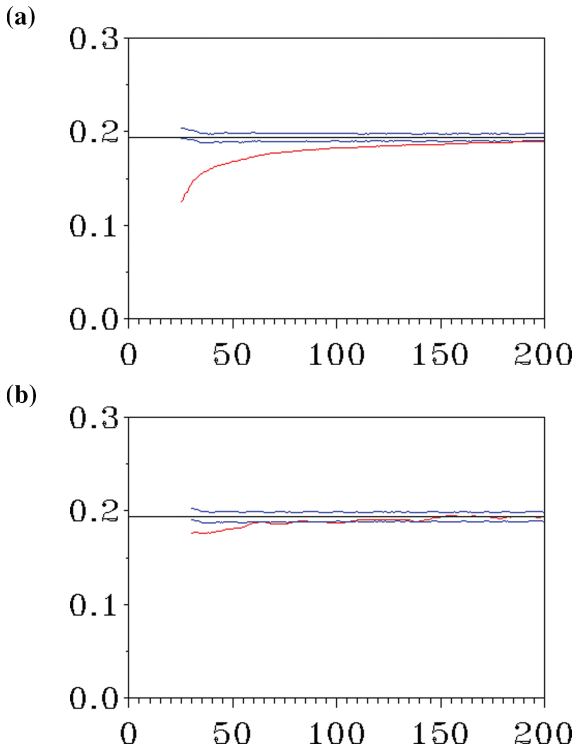

Fig. 7. (a) Mean estimated distance-dependent Simpson index $\hat{\alpha}(r)$ for all 20 cases (red curve). The classical distance-independent Simpson index $\hat{D}$ is plotted as a horizontal black line. The blue curves above and below this line indicate 90% confidence intervals of $\alpha(r)$ for the case of random labelling for the value $\hat{D} = 0.191$, i.e. the mean estimated value of D for our 20 cases. (b) The analogous plot for the mean estimated distance-dependent diversity index $\hat{\beta}(r)$.

The distance-dependent Simpson indices show in which manner the labelling pattern of our points differs from random labelling. Both indices are lowered as compared with D, which means that for small distances r, the mixed point pattern shows less diversity than one would expect for random labelling. This statement indicates that both unlabelled and labelled points aggregate more strongly together with points of their own kind, than one would expect at random labelling. Biologically, it favours the hypothesis that in mammary carcinomas, cell division occurs in clusters of proliferative activity. Actively proliferating zones seem to alternate with relatively silent zones. This behaviour is known from many normal tissues such as the normal lymph node in which proliferation is practically restricted to the lymph follicles. It is also seen in normal epithelial tissues such as the gut mucosa or the skin in which the proliferation is largely concentrated in the basal layers.

On the whole, the results of the Monte Carlo rank test also support the conclusion that the labelling of proliferating tumour cell nuclei is not random. For the majority of the individual patterns (26/40), the hypothesis of random labelling was rejected. One may however wonder why the null hypothesis was not rejected for a still higher number, or even for all patterns. It can be seen from the results on the distance-dependent Simpson indices that the deviation of the labelling from randomness is statistically significant but the difference is only small. In this constellation, the power of hypothesis tests on individual patterns will be relatively low. We conclude that the labelling of the points is presumably structured in all patterns but our Monte Carlo test does not detect this structure in all patterns. Probably, it fails for those patterns in which the deviation from random labelling is the lowest. Moreover, it is typical for biological data that variation between individuals is large, and this may make it difficult to find differences between groups.

The methods presented in this paper are not new from the viewpoint of spatial statistics but we consider them as innovative in the context of histopathology. A microscopical application that has been studied rather thoroughly consists

of testing patterns of cells for independence of two point processes (see Diggle *et al.*, 2006; Eglen *et al.*, 2006, and references therein). However, this concept must be sharply distinguished from random labelling (see Diggle, 2003, pp. 48–49). If the hypothesis of independence of two point processes is considered, we have two point processes, each of which has (or is labelled with) a different binary mark. In this case, the null hypothesis is: are the two processes independent, i.e. are the two types of events generated by a pair of independent univariate processes? (Diggle, 2003, p. 48). The classical example from microscopy are the amacrine cells of the retina in which two populations can be distinguished, the 'on-cells' and the 'off-cells' (Diggle *et al.*, 2006; Eglen *et al.*, 2006). None of the two may be considered as a subset of the other. By contrast, if we consider random labelling, we have only one point process, of which a subset is positively labelled, whereas the remainder is negative. Notably, from the viewpoint of histopathology, the latter case is much more relevant. In routine histological work, immunohistochemical methods are used that stain a specific subset of the cells (or of the nuclei), whereas the others remain unstained (see Fig. 1a). Only one antibody (one stain) is used for a section. Hence, one obtains only unlabelled or labelled cells but not two different labels. Thus, the natural approach for a statistical analysis of cells stained by routine immunohistochemical methods is to test them on random labelling, and one would not test on independence. Using immunohistochemistry, a test on the independence of two labelled point processes would surmise the application of two stains to the same section, which is unusual in routine procedures. A test on random labelling in the context of immunohistochemistry was performed here for the first time, as far as we could determine.

Furthermore, the distance-dependent indices of diversity have scarcely been used in applications to real data before at all (Shimatani, 2001; Eckel *et al.*, 2007). As far as we know, our contribution is the first paper in which the indices $\alpha(r)$ and $\beta(r)$ have been applied to histopathological data. The battery of methods presented in this paper is by no means restricted to the study of proliferation. It can be used whenever particles such as cell nuclei are stained in such a manner that all nucleus profiles can be classified in a binary manner as negative or positive. This holds not only for the MIB-1 stain but also for many other immunohistochemical stains used in normal histology and histopathology, such as, stains for hormone receptors in mammary carcinomas. Only the model assumptions of isotropy and stationarity must be surmised. Note that for other cases, there are anisotropic versions of the K- and g-functions as well as a non-stationary version of the K-function. This opens a wide field of applications in practice. Hence, although our results are related to a specific case series (labelling for MIB-1 in mammary cancer), the methodology appears to be widely applicable.

Methodological aspects

Simulation of random labelling. Simulation of random labelling by permutation of the marks is not the only solution for this problem. In addition to random permutation of the marks (i.e. sampling without replacement), it is also possible to sample with replacement as follows. Given the observed locations (the coordinates), each point is labelled as 0 or 1, uniformly and independently at random. The probability of labelling is selected according to the estimated labelling fraction of the individual pattern for which the simulations were performed. This method leads to slightly varying numbers of labelled points in the simulations, whereas the number of labelled and unlabelled points is constant if the permutation method is used. On the whole, the permutation method is more usual. Sampling with replacement was performed in this manner in the present context too. It led to exactly the same results in the Monte Carlo test on random labelling (rejection of the hypothesis of random labelling in 26 of 40 patterns).

Choice of the test statistics. In this paper, the attention was focused on the functions $K_{11}(r)$ and $K_{22}(r)$ when testing for random labelling. It is also possible to estimate the cross K-function $K_{12}(r)$. Under random labelling, all three functions should be identical, i.e. $K_{11}(r) = K_{22}(r) = K_{12}(r)$ (see Diggle & Chetwynd, 1991, eq. (3); Diggle, 2003, eq. (4.12)). Hence, one is free to decide whether one bases a test on random labelling on a comparison of $K_{11}(r)$ and $K_{22}(r)$, on a comparison of $K_{11}(r)$ and $K_{12}(r)$ or on a comparison of $K_{12}(r)$ and $K_{22}(r)$ (Diggle & Chetwynd, 1991). For this study, the function $K_{12}(r)$ was estimated with SPATSTAT using the function K_{cross}, and plotted for all the 40 patterns. In the majority of the patterns, the plot of $K_{12}(r)$ lay quite near to the plot of $K_{11}(r)$, whereas $K_{22}(r)$ usually lay distinctly more separate (Fig. 8). On the other hand, $K_{11}(r)$ and $K_{22}(r)$ were needed for the Monte Carlo rank tests and for the estimation of $\alpha(r)$ anyway, and their meaning is easier to grasp intuitively than the meaning of $K_{12}(r)$. Hence, we preferred to work with $K_{11}(r)$ and $K_{22}(r)$ (see also Diggle & Chetwynd, 1991). Analogous considerations led us to favour the estimation of $g_{11}(r)$ and $g_{22}(r)$ instead of $g_{12}(r)$.

Choice of the kernel parameters. For the estimation of the second-order function $g(r)$, kernel smoothing is strongly recommended. The choice of the kernel type and the bandwidth strongly influences the estimate of $g(r)$. It was decided to use the popular Epanechnikov kernel. It is the default kernel implemented in SPATSTAT, in which the recommendations given in Stoyan & Stoyan (1994, pp. 284–285) are exactly followed. Our bandwidth was selected according to the formula $h\hat{\lambda}^{-\frac{1}{2}}$, with $h = 0.1$, which is a default value (see, e.g. Krasnoperov & Stoyan, 2004). However, it should be pointed out that in some cases, other kernels, such as the box kernel, may lead to slightly more favourable results

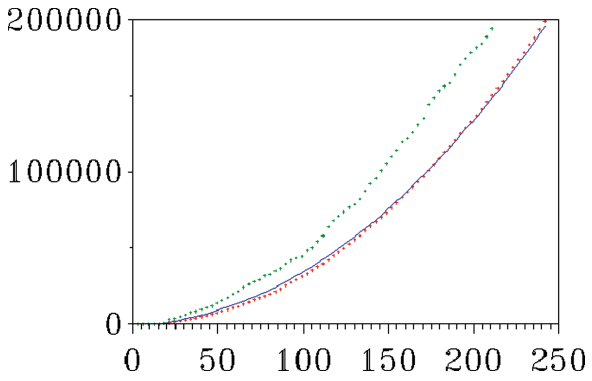

Fig. 8. Plots of estimates of $K_{11}(r)$ (blue, continuous), $K_{22}(r)$ (green) and $K_{12}(r)$ (red) from a selected pattern.

(Stoyan & Stoyan, 2000, pp. 649–650). The latter kernel is also implemented in SPATSTAT and can be used with the syntax 'kernel = rectangular'. We are currently exploring the benefits and drawbacks of the Epanechnikov kernel, the box kernel and other kernels for our data but this methodological comparison is beyond the scope of this paper.

Validation of the point process model assumptions. The patterns of the unlabelled and the labelled nucleus profiles were parametrically modelled as stationary Strauss hard-core processes. To check whether this model was realistic for our point patterns, 99 simulations were performed per pattern using SPATSTAT on the basis of the four estimated model parameters for the pattern ($\hat{\lambda}$, $\hat{r_0}$, $\hat{\gamma}$ and $\hat{R}$). The goodness of fit was judged graphically from the plots of the K-function estimated from the real patterns and the plots of the envelopes of the K-functions obtained from the simulations, in a range up to $r = 250$ pixels. A perfect fit was found in 47 of the 80 patterns in which the K-function of the real pattern lay everywhere between the envelopes of the K-functions from the simulated patterns (Figs 9a and b). In other patterns, slight deviations of the empirical K-function below the lower or above the upper envelopes were noted (Fig. 9c). Drastic discrepancies were found for some patterns but they were the exception. It must be admitted that not all patterns are entirely explained by the model. The intensity, the hard-core distance and the general course of the K-function were, however, generally well preserved. From our point of view, the Strauss hard-core process does not yields an absolutely perfect fit but may be considered as a realistic approximation for practical purposes, which may be superimposed by unknown other processes. Nevertheless, a word of caution seems appropriate. For the Strauss processes, it may happen that simulated realizations of the model tend to be more clustered than the observed pattern (see, e.g. Møller, 1999). However, as the main aim of the present study was merely to compare the degree of regularity/clustering between the different patterns, this special methodological problem was not attacked here.

Merits of the K-function and the pair correlation function in studies on random labelling. In section 'Explorative Analysis of Point Patterns', a major emphasis was put on the estimation of the pair correlation function $g(r)$. While $K(r)$ is monotonically increasing, $g(r)$ has a non-monotonic course in which characteristic points can be identified, such as the first maximum and the first minimum. This yields two uniquely defined points by which a particular estimate of $g(r)$ can be numerically characterized. In contrast to the model parameters γ and R mentioned in section 'Parametric Modelling', the coordinate values of these two points are entirely non-parametric indicators and do not surmise any particular point process model. Such characteristic points, however, can only be identified from $g(r)$ and not from $K(r)$ (see Figs 5 and 6). They can be used for group comparisons. On the other hand, $K(r)$ is indispensable for the Monte Carlo rank tests of the individual marked point patterns on random labelling. The distance-dependent indices of diversity can be computed on the basis of both $K(r)$ and $g(r)$. On the whole, these considerations lead to the recommendation to estimate both summary statistics if the question of random labelling is considered.

Computational aspects

Some of the investigations described in this paper belong to the category of computer-intensive methods. Typical for this domain is that rather simple calculations are performed with a very high number of replications (Mattfeldt & Fleischer, 2006). In particular, this holds for simulation and re-sampling techniques that can be performed in an acceptable time limit only with computers of high performance (Møller & Waagepetersen, 2004; Mattfeldt & Fleischer, 2006). For the present study, this statement applied, in particular, to the computations described in sections 'Distance-Dependent Characteristics of Diversity' and 'A Monte Carlo Rank Test on Random Labelling': computation of the distance-dependent Simpson indices with confidence intervals by simulations and the Monte Carlo rank tests on random labelling for individual marked point patterns. It should be noted that this data structure (a limited number of equally structured low-dimensional samples) is rather typical for biomedical studies. If the analysis of each sample needs computer-intensive methods, such data are ideally suited for high-performance clusters. Such a cluster was also used in this study (a Linux cluster with 32 nodes, each node equipped with 2 AMD Opteron single-core or dual-core processors (AMD, Markham, ON)). In such an architecture, parallel processing of the data sets (here our 40 point patterns) can be simply performed on the cluster by processing the computations in batch mode. The parallelization can be achieved here by ascribing separate batch jobs to the statistical analysis of individual patterns. Proceeding like this, the increase of speed due to the multiple nodes of the cluster is fully exploited, without any modification

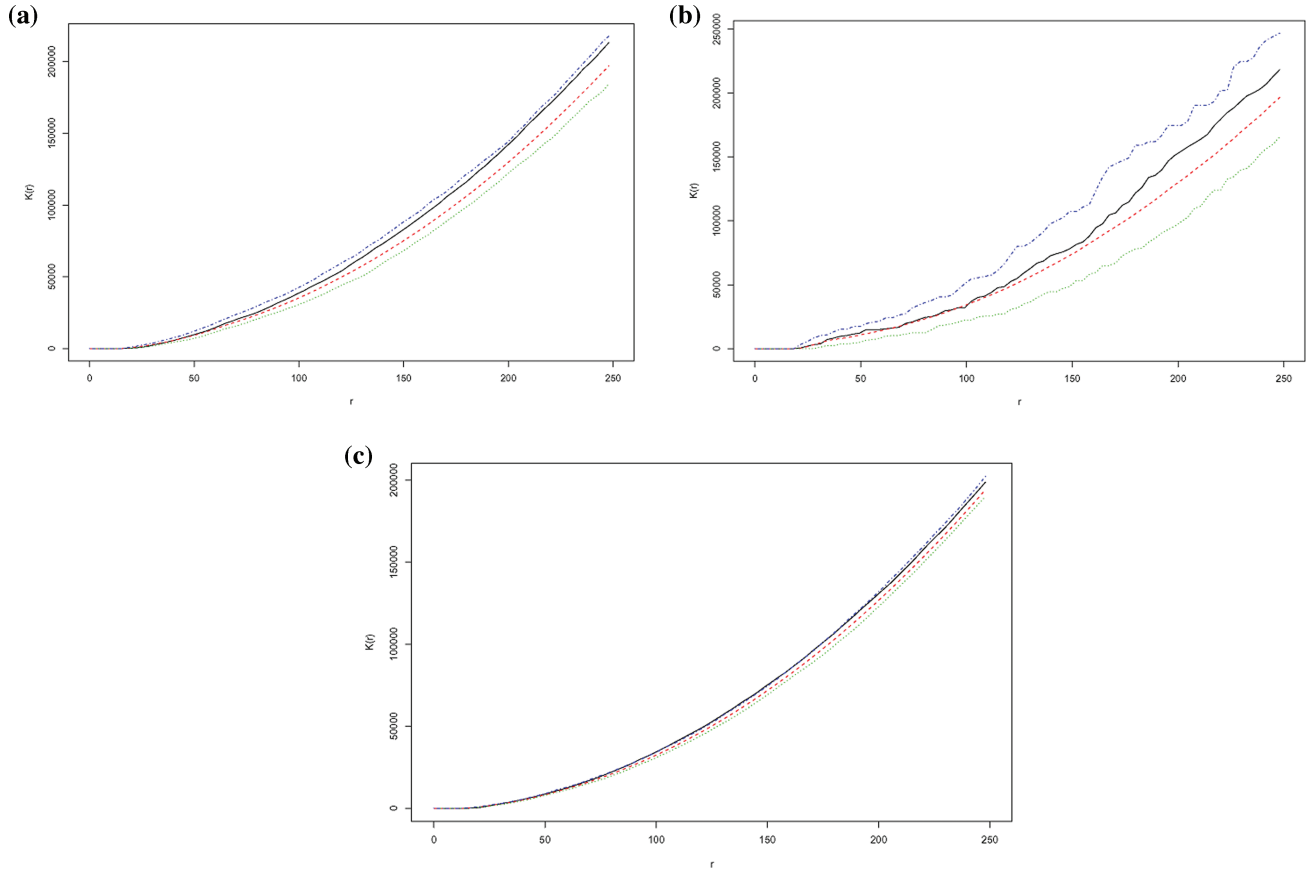

Fig. 9. Graphical test of selected patterns of tumour cell nuclei on compatibility with the model of a stationary Strauss hard-core process. The lower envelope, mean and upper envelope of the *K*-functions obtained from 99 simulated patterns of the model are shown in green, red and blue, respectively. The *K*-function estimated from the real pattern is shown in black. (a) From a selected pattern of unlabelled nuclei: perfect fit. (b) From a selected pattern of labelled nuclei: perfect fit. (c) From a selected pattern of unlabelled nuclei: slight deviation of the *K*-function of the real pattern from the *K*-function of the simulated model.

of the program code. In this setting, shell scripts using, e.g. R and SPATSTAT, which run on a stand-alone Linux workstation, can be executed without any changes in the cluster, as there is no necessity to make adaptations of the source code to parallel programming libraries, such as Message Passing Interface (MPI) or other libraries (Sloan, 2005).

Stereological considerations

In this study, the labelling pattern of tumor cell nucleus profiles on sections was studied by methods of spatial statistics. It must be borne in mind that these points are not the nuclei but traces of the nuclei on sections. It would be naive to surmise that our functions $\hat{g}(r)$ and $\hat{K}(r)$ are unbiased estimators of the functions $g_3(r)$ and $K_3(r)$ of the true nuclei. These are the second-order summary statistics of the mid-points of the particles living in 3D space, i.e. the 3D versions of the functions investigated here. It is easy to see that direct inference is not possible in this way for general particles if no assumptions on their shape and size distribution can be made (Mattfeldt, 2005). However, if some model assumptions are valid, we have $g_3(r) \approx g(r)$. This relation holds in good approximation if we deal with sections of spherical particles with constant diameters or spherical particles with limited size variation (Stoyan *et al.*, 1995, pp. 377ff.) This shape model is not too unrealistic for mammary carcinoma cell nuclei, which look mostly nearly circular on sections (see Fig. 1a). Also, in our cases, giant nucleus profiles, strongly exceeding the size of the average population in terms of their diameter, did not emerge (see Fig. 1a). Such nuclei are well known to pathologists from other tumour types, but in our cases, (ductal adenocarcinomas of the mammary gland), they play no role. These qualitative observations indicate that our analysis on planar sections is probably valid also for the true nuclei in the first approximation. The true labelling pattern would be the marked 3D point pattern of the proliferating and non-proliferating nuclei in the 3D tissue. Such patterns are, however, difficult to visualize directly, which would require

special 3D microscopic methods. For the time being, one has to live in routine applications with planar sections. The aforementioned conclusions should hold, bearing in mind that they are based on the model of roughly spherical nuclei with limited size variation.

Acknowledgements

Thanks are due to Rolf Kunft for skilful technical assistance. Stefanie Eckel is supported by a grant of the graduate college 1100 at the University of Ulm. The authors thank both referees for their constructive comments, which helped to improve an earlier version of the paper.

References

Baddeley, A. & Turner, R. (2000) Practical maximum pseudolikelihood for spatial point patterns (with discussion). *Austral. N. Z. J. Stat.* **42**, 283–322.

Baddeley, A. & Turner, R. (2005) Spatstat: An R package for analyzing spatial point patterns. *J. Statist. Software* **12**(6), 1–42.

Baddeley, A. & Turner, R. (2006) Modelling spatial point patterns in R. *Case Studies in Spatial Point Pattern Modelling. Lecture Notes in Statistics 185* (ed. by A. Baddeley, P. Gregori, J. Mateu, R. Stoica & D. Stoyan), pp. 23–74. Springer, New York.

Diggle, P.J. (2003) *Statistical Analysis of Spatial Point Patterns*, 2nd edn. Arnold, London.

Diggle, P.J. & Chetwynd, A.G. (1991) Second-order analysis of spatial clustering for inhomogeneous populations. *Biometrics* **47**, 1155–1163.

Diggle, P.J., Fiksel, T., Grabarnik, P., Ogata, Y., Stoyan, D. & Tanemura, M. (1994) On parameter estimation for pairwise interaction point processes. *Int. Statist. Rev.* **62**, 99–117.

Diggle, P.J., Eglen, S.J. & Troy, J.B. (2006) Modelling the bivariate spatial distribution of amacrine cells. *Case Studies in Spatial Point Pattern Modelling. Lecture Notes in Statistics 185* (ed. by A. Baddeley, P. Gregori, J. Mateu, R. Stoica & D. Stoyan) pp. 215–233. Springer, New York.

Eckel, S., Fleischer, F. Grabarnik, P. & Schmidt, V. (2007) An investigation on the spatial correlations for relative purchasing power in Baden-Württemberg. *Adv. Statist. Anal.* **92**, 135–152.

Eglen, S.J., Diggle, P.J. & Troy, J.B. (2006) Functional independence of on- and off-centre beta retinal ganglion cell mosaics. *Vis. Neurosci.* **22**, 859–871.

Gerdes, J., Becker, M.H., Key, G. & Cattoretti, G. (1992) Immunohistological detection of tumour growth fraction (Ki-67 antigen) in formalin-fixed and routinely processed tissues. *J. Pathol.* **168**, 85–86.

Goulard, M., Särkkä, A. & Grabarnik, P. (1996) Parameter estimation for marked Gibbs point processes through the maximum pseudo-likelihood method. *Scand. J. Stat.* **23**, 365–379.

Illian, J., Penttinen, A., Stoyan, H. & Stoyan, D. (2008) *Statistical Analysis and Modelling of Spatial Point Patterns*. Wiley, Chichester.

Krasnoperov, R.A. & Stoyan, D. (2004) Second-order stereology of spatial fibre systems. *J. Microsc.* **216**, 156–164.

Mattfeldt, T. (2005) Explorative statistical analysis of planar point processes in microscopy. *J. Microsc.* **220**, 131–139.

Mattfeldt, T. & Fleischer, F. (2006) Computer-intensive methods for statistical inference from stereological data. In *Proceedings of the International Conference on Spatial Statistiscs and Stochastic Geometry (S4G)*, Prague (ed. by R. Lechnerova, I. Saxl & V. Beneš), pp. 349–360.

Mattfeldt, T. Eckel, S., Fleischer, F. & Schmidt, V. (2006) Statistical analysis of reduced pair correlation functions of capillaries in the prostate gland. *J. Microsc.* **223**, 107–119.

Mattfeldt, T., Eckel, S., Fleischer, F. & Schmidt, V. (2007) Statistical modelling of the geometry of planar sections of prostatic capillaries on the basis of stationary Strauss hard-core processes. *J. Microsc.* **228**, 272–281.

Mayer, J., Schmidt, V. & Schweiggert, F. (2004) A unified simulation framework for spatial stochastic models. *Simulat. Model. Pract. Theory* **12**, 307–326.

Møller, J. (1999) Markov chain Monte Carlo and spatial point processes. *Stochastic Geometry: Likelihood and Computation* (ed. by O. Barndorff-Nielsen, W. Kendall & M.N.M. Lieshout), pp. 141–172, Chapman and Hall, Boca Raton.

Møller, J. & Waagepetersen, R.P. (2004) *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall, Boca Raton.

Ohser, J. (1983) On estimators for the reduced second moment measure of point processes. *Mathemat. Operationsforsch. Stat.* **14**, 63–71.

Ripley, B.D. (1988) *Statistical Inference for Spatial Processes*. Cambridge University Press, Cambridge.

Sachs, L. (2003) *Angewandte Statistik*, 11th edn. Springer, Berlin.

Shimatani, K. (2001) Multivariate point processes and spatial variation of species diversity. *Forest Ecol. Manage.* **142**, 215–229.

Sloan, J.D. (2005) *High Performance Linux Clusters with OSCAR, Rocks, openMosix, and MPI*. O'Reilly, Sebastopol.

Stoyan, D. & Schnabel, H.-D. (1990) Description of relations between spatial variability of microstructure and mechanical strength of alumina ceramics. *Ceram. Int.* **16**, 11–18.

Stoyan, D. & Stoyan, H. (1994) *Fractals, random shapes and point fields. Methods of Geometrical Statistics*. Wiley, Chichester.

Stoyan, D. & Stoyan, H. (2000) Improving ratio estimators of second order point process characteristics. *Scand. J. Stat.* **27**, 641–656.

Stoyan, D., Kendall, W.S. & Mecke, J. (1995) *Stochastic Geometry and Its Applications*, 2nd ed. Wiley, Chichester.

Takacs, R. & Fiksel, T. (1986) Interaction pair potentials for a system of ants' nests. *Biometr. J.* **28**, 1007–1013.

Van Lieshout, M.N.M. (2002) *Markov Point Processes and Their Applications*. World Scientific, Singapore.

Webster, A., Diggle, P.J., Clough, H.E., Green, R.B. & French, N.P. (2006) Strain-typing transmissible spongiform encephalopathies using replicated spatial data. *Case Studies in Spatial Point Pattern Modelling. Lecture Notes in Statistics 185* (ed. by A. Baddeley, P. Gregori, J. Mateu, R. Stoica & D. Stoyan), pp. 197–214. Springer, New York.