INTERFACE

rsif.royalsocietypublishing.org

Research

Cite this article: Dini S, Binder BJ, Fischer SC, Mattheyer C, Schmitz A, Stelzer EHK, Bean NG, Green JEF. 2016 Identifying the necrotic zone boundary in tumour spheroids with pair-correlation functions. *J. R. Soc. Interface* **13**: 20160649.
http://dx.doi.org/10.1098/rsif.2016.0649

Received: 16 August 2016
Accepted: 15 September 2016

Subject Category:
Life Sciences–Mathematics interface

Subject Areas:
biomathematics, computational biology, biometrics

Keywords:
tumour spheroid, necrotic zone, pair-correlation function

Author for correspondence:
S. Dini
e-mail: saber.dini@adelaide.edu.au

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.3491550.

THE ROYAL SOCIETY PUBLISHING

Identifying the necrotic zone boundary in tumour spheroids with pair-correlation functions

S. Dini[1], B. J. Binder[1], S. C. Fischer[2], C. Mattheyer[2], A. Schmitz[2], E. H. K. Stelzer[2], N. G. Bean[1,3] and J. E. F. Green[1]

[1]School of Mathematical Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia
[2]Department of Biological Sciences (IZN, FB 15), Buchmann Institute for Molecular Life Sciences (BMLS), Goethe Universität Frankfurt am Main, 60438 Frankfurt am Main, Germany
[3]ARC Centre of Excellence for Mathematical and Statistical Frontiers, Parkville, Victoria 3010, Australia

SD, 0000-0002-4613-3672; JEFG, 0000-0001-5061-9563

Automatic identification of the necrotic zone boundary is important in the assessment of treatments on *in vitro* tumour spheroids. This has been difficult especially when the difference in cell density between the necrotic and viable zones of a tumour spheroid is small. To help overcome this problem, we developed novel one-dimensional pair-correlation functions (PCFs) to provide quantitative estimates of the radial distance of the necrotic zone boundary from the centre of a tumour spheroid. We validate our approach on synthetic tumour spheroids in which the position of the necrotic zone boundary is known *a priori*. It is then applied to nine real tumour spheroids imaged with light sheet-based fluorescence microscopy. PCF estimates of the necrotic zone boundary are compared with those of a human expert and an existing standard computational method.

1. Introduction

Tumour spheroids are *in vitro* cell aggregates grown from a smaller number of cells initially placed in a non-adhesive environment [1–3]. They provide a way to study cancer cell behaviours and interactions in a well-controlled environment, while mimicking the *in vivo* arrangement of cells more closely than monolayer cultures. Importantly, tumour spheroids are used in drug testing assays [1,4], and assessments of the effectiveness of anti-tumour treatments often rely on quantitative measures of the cell distribution within the spheroid [5,6].

In a sufficiently large spheroid, only the cells in the outermost cell layers can receive enough oxygen and other nutrients to proliferate. This region is termed the proliferative zone, while deeper inside the spheroid, quiescent and necrotic zones are formed [3]. Cells in the quiescent zone remain viable but do not proliferate, while the innermost cells die due to an accumulation of toxic waste products and a lack of oxygen and nutrient supply, forming the necrotic zone [7]. The schematic diagram of figure 1 illustrates the necrotic, quiescent and proliferative zones within a tumour spheroid. Identifying and quantifying these three regions is important in the analysis of comparative assays on tumour spheroids [8–10] and mathematical models of the tumour growth process [11–14].

In this paper, we analyse the *in vitro* distribution of cells in nine (homotypic) tumour spheroids. Using light sheet-based fluorescence microscopy in combination with optical clearing, high-quality three-dimensional images are generated [15]. Subsequently, the images are processed with a three-dimensional segmentation method to obtain a point cloud representing the cell distribution. The cell distribution gives a point pattern that is subsequently analysed, with the aim of being able to provide an estimate of the position of the necrotic zone boundary, i.e. the distance from the spheroid centre to where the necrotic zone transitions into the quiescent zone.

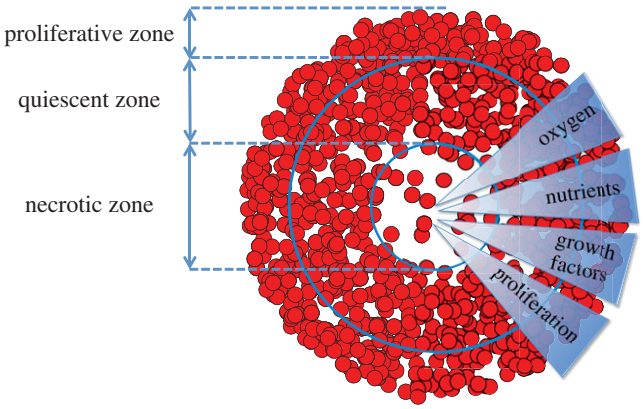

Figure 1. Schematic of the necrotic, quiescent and proliferative zones within a tumour spheroid. In the necrotic zone, cell death occurs due to an accumulation of toxic waste products and a lack of oxygen and nutrient supply [7]. The quiescent zone is a region of viable and non-proliferative cells. In the proliferative zone, cells receive enough oxygen and nutrients to proliferate [3]. (Online version in colour.)

Standard density-based spatial clustering and data clustering methods (DBSCAN and k-means) are implemented to identify the boundary of the necrotic zone [16–18]. However, we find that using such existing methods can fail, or produce unreliable results when the difference in the cell density between the quiescent zone and necrotic zone is small. Therefore, we offer an alternative, statistically based, approach by developing a one-dimensional pair-correlation function (PCF) to identify the necrotic zone boundary in tumour spheroids.

The PCF is a second-order summary statistic commonly used for analysing point patterns in cell biology [19–24]. Typically, PCFs describe the relative frequency of Euclidean distances between pairs of data points, indicating the extent of deviations from complete spatial randomness (CSR) [25–29]. The PCF for a stationary and isotropic point process is defined as

$$g(r) = \frac{\varrho(r)}{\lambda^2}, \tag{1.1}$$

where r is usually the Euclidean distance between points, $\varrho(r)$ is the second-order product density (frequency of points separated by a distance r) and λ is the intensity of the point process [27]. When points are distributed uniformly at random (i.e. CSR) $\varrho(r) = \lambda^2$ and thus $g(r) = 1$ at all distances. Consequently, aggregation and segregation length-scales correspond to $g(r) > 1$ and $g(r) < 1$, respectively. Therefore, we can quantify spatial features by estimating the PCF for point patterns [30–33].

However, in some situations the Euclidean distance between points is not the most appropriate distance to study. For example, in a scratch assay used to assess wound healing *in vitro*, the cells move into the wounded region as a front, which is approximately a straight line. Binder & Simpson [34] used a one-dimensional PCF to quantify the spatial patterning of the cells in the Cartesian direction perpendicular to the front in both experimental images and simulations. They also showed that there was no spatial structure in the Cartesian direction parallel to the front. In other situations, the Euclidean distance between points may not be the quantity of interest. For example, Binder *et al*. [35] analysed the angular separation of filaments emanating from two-dimensional images of yeast colonies with a one-dimensional angular PCF. Similarly, in attempting to identify the necrotic zone boundary in tumour spheroids, we are concerned primarily with variations of cell density in one particular direction (the radial direction).

This then motivates us to formulate a projected one-dimensional PCF to analyse three-dimensional spatial point patterns with respect to the directions (radial, polar and azimuthal) of the spherical coordinates. In the derivation of this PCF, the (usual) conditions of stationarity and isotropy are relaxed because the projected point processes are in general non-stationary and anisotropic. We examine the accuracy of our method by estimating the PCF for simulated CSR and regular spatial patterns in the spherical coordinate system. In addition, we generate synthetic datasets of cell distributions in tumour spheroids, and demonstrate that the PCF can accurately identify the necrotic zone boundary. This helps with the interpretation and analysis of the PCF results for nine experimental datasets, which provide a 'proof of concept' for the usefulness of our approach.

2. Mathematical methods

We derive a projected one-dimensional (non-periodic) PCF to analyse three-dimensional spatial point patterns described in spherical coordinates, (r, θ, ϕ). The aim is to use a sample of N data points to estimate the PCF of the underlying point process. The sample, or point pattern, is a finite subset of three-dimensional space that we can define by

$$S = \{\underset{\sim}{b}_i = (r_i,\theta_i,\phi_i) | 0 \le r_i \le R,\ 0 \le \theta_i \le \pi,\ 0 \le \phi_i < 2\pi,$$
$$i = 1,2,\ldots,N\},$$

where $\underset{\sim}{b}_i$ is the position vector of the ith sample point, and $R = \max(r_i)$ for $i = 1,2,\ldots,N$. Without loss of generality, we consider the projection, Q, of S onto the interval $[0, L]$, to obtain the projected point pattern

$$S_a = \{a_i = Q(\underset{\sim}{b}_i) | a_i \in [0, L],\ i = 1,2,\ldots,N\},$$

where a_i is the projected position of the ith sample point. Note that for the radial, polar and azimuthal projections, we have $a_i = r_i$ and $L = R$, $a_i = \theta_i$ and $L = \pi$, and $a_i = \phi_i$ and $L = 2\pi$, respectively. The approach now taken in deriving the PCF is based on normalizing the proportion of pairs of projected data points, $G(\Delta a)$, that are separated by a distance Δa in S_a. The normalization is with respect to the probability of observing such pairs in the projection of the Poisson process, $\bar{G}(\Delta a)$. Owing to the discrete nature of points, $G(\Delta a)$ has to be estimated using the average over an interval (numerator of $G(\Delta a)$ in equation (2.1)). The quantities $G(\Delta a)$ and $\bar{G}(\Delta a)$ are analogous to the numerator and denominator in equation (1.1).

To evaluate $G(\Delta a)$, we introduce the bandwidth h, and obtain the expression

$$G(\Delta a) = \frac{(1/h)\sum_{i=1}^{N}\sum_{j>i}^{N} 1_{(0,h]}(|a_j - a_i| - \Delta a)}{\binom{N}{2}}, \tag{2.1}$$

where

$$1_{(0,h]}(x) = \begin{cases} 1 & \text{if } x \in (0,h] \\ 0 & \text{otherwise.} \end{cases}$$

Note that the denominator in equation (2.1) accounts for the total number of all possible combinations of pairs of data points.

To evaluate $\bar{G}(\Delta a)$, we first consider the homogeneous Poisson process, **P**, which is synonymous with CSR. Depending on the form of the projection operator, Q, the projection of the Poisson process, P_a, can be non-stationary and anisotropic (e.g. radial projection). However, as **P** is the Poisson process, the probability density function f_a for the projected points of $\mathbf{P}_a$ is known. The projected points, $\mathbf{P}_a$, can be considered as samples drawn from a random variable A with probability density function f_a.

In order to find the probability of having a pair of points that are separated by a distance Δa in $\mathbf{P}_a$, we can use the density at a location $a \in [0, L]$, and the cumulative distribution of the points in the intervals $(a + \Delta a, a + \Delta a + h]$ and $[a - \Delta a - h, a - \Delta a)$. Integrating over the interval [0, L] then gives

$$\begin{aligned}\bar{G}(\Delta a) = &\frac{1}{h}\int_0^{L-\Delta a} f_a(a) P(a + \Delta a < A < a + \Delta a + h)\,\mathrm{d}a \\ &+ \frac{1}{h}\int_{\Delta a}^{L} f_a(a) P(a - \Delta a - h < A < a - \Delta a)\,\mathrm{d}a.\end{aligned} \tag{2.2}$$

As the probability of finding a pair of points a distance Δa apart does not depend on the order of counting the possible pairs (left-wise or right-wise), it can be shown that the two integrals in equation (2.2) are equal. When written in terms of the probability density function, equation (2.2) becomes

$$\bar{G}(\Delta a) = \frac{2}{h}\int_0^{L-\Delta a} f_a(a) \int_{a+\Delta a}^{a+\Delta a+h} f_a(t)\,\mathrm{d}t\,\mathrm{d}a. \tag{2.3}$$

Using equations (2.1) and (2.3), the estimate of the non-periodic PCF is then given by

$$g(\Delta a) = \frac{G(\Delta a)}{\bar{G}(\Delta a)}, \tag{2.4}$$

where the probability density function, f_a, for each of the three projections is

$$\left.\begin{aligned} f_r(r) &= \begin{cases} \dfrac{3r^2}{R^3} & \text{for } r \in [0,\ R] \\ 0 & \text{otherwise,} \end{cases} \\ f_\theta(\theta) &= \begin{cases} \dfrac{\sin(\theta)}{2} & \text{for } \theta \in [0,\ \pi] \\ 0 & \text{otherwise,} \end{cases} \\ \text{and} \quad f_\phi(\phi) &= \begin{cases} \dfrac{1}{2\pi} & \text{for } \phi \in [0,\ 2\pi) \\ 0 & \text{otherwise.} \end{cases} \end{aligned}\right\} \tag{2.5}$$

2.1. Complete spatial randomness and regular spatial patterns

The method is validated by evaluating the non-periodic PCF, equation (2.4), for simulated CSR and regular spatial patterns within a sphere of radius R. The point patterns are shown in the panels of the top row of figure 2. From left to right they are: (i) CSR pattern, (ii) segregated clusters of points in spherical shells, (iii) segregated clusters of points that are locally aggregated around prescribed angles of ϕ, and (iv) segregated clusters of points in conical shells. The three panels directly below each test pattern in figure 2 correspond to the PCF evaluation of the radial, azimuthal and polar projections from 1000 simulations.

At short and intermediate length-scales the results for the simulated CSR pattern, in the first column of figure 2, indicate that there is no spatial structure (in any direction) as $g \approx 1$. However, we see that there is significant deviation from unity in the PCF signals at large distances. This deviation at large distances appears to be inconsistent with our formulation of the PCF, as we might expect the signal to be close to unity at all distances. The explanation for this disparity is the division of small numbers in equation (2.4), where the observed frequency and expected frequency of pairs of points at large distances are both small. The results show that the non-periodic PCF is a reliable predictor of CSR at short and intermediate length-scales.

We now consider the regular spatial patterns. They were chosen such that for each spatial pattern there is only spatial structure expected in one of the three corresponding projected patterns. For example, we discuss the results for the points distributed in spherical shells, in the second column of figure 2. As expected, the PCF indicates that there is no spatial structure for the azimuthal and polar projections (figure 2*g*,*h*). For the radial projection (figure 2*f*), we see a series of five decreasing peaks for increasing values of Δr in the PCF. The highest peak at $\Delta r = 0$ corresponds to pairs of points within each of the five spherical shells (e.g. pairs of red points, pairs of green points, etc.). The smallest peak at $\Delta r = 200$ corresponds to a pair of points with one point belonging to the innermost spherical shell (red) and the other to the outermost spherical shell (black). Further information about the spatial patterning can be easily obtained from the signal (e.g. the distance between the peaks in the signal is a measure of distance between the spherical shells). A similar discussion holds for the two sets of results in the last two columns of figure 2. This demonstrates the application of the PCF to quantify spatial structures in point patterns.

2.2. Periodic pair-correlation function

In our analysis of the simulated CSR point pattern of figure 2, we observed that the non-periodic PCF, equation (2.4), had significant deviations from unity at large distances—due to the low frequency of occurrence of large separations. This can lead to the incorrect assessment that a point pattern has spatial structure at large distances. This problem of the division of small numbers at large distances can be overcome by defining a periodic PCF.

Following the work of Agnew *et al*. [19] (and others), we re-define the distance between two points in equation (2.1) as

$$\begin{aligned}||a_j - a_i|| = \min(&|a_j - a_i|, L - |a_j - a_i|) \\ &\text{for } i \neq j \in \{1, \ldots, N\},\end{aligned} \tag{2.6}$$

with the consequence that

$$0 \leq ||a_j - a_i|| \leq \frac{L}{2}.$$

Using equations (2.1) and (2.6), the proportion of pairs of data points, $G_p(\Delta a)$, separated by a distance Δa is given by

$$G_p(\Delta a) = \frac{(1/h)\sum_{i=1}^{N}\sum_{j>i}^{N} \mathbf{1}_{(0,h]}(||a_j - a_i|| - \Delta a)}{\binom{N}{2}}. \tag{2.7}$$

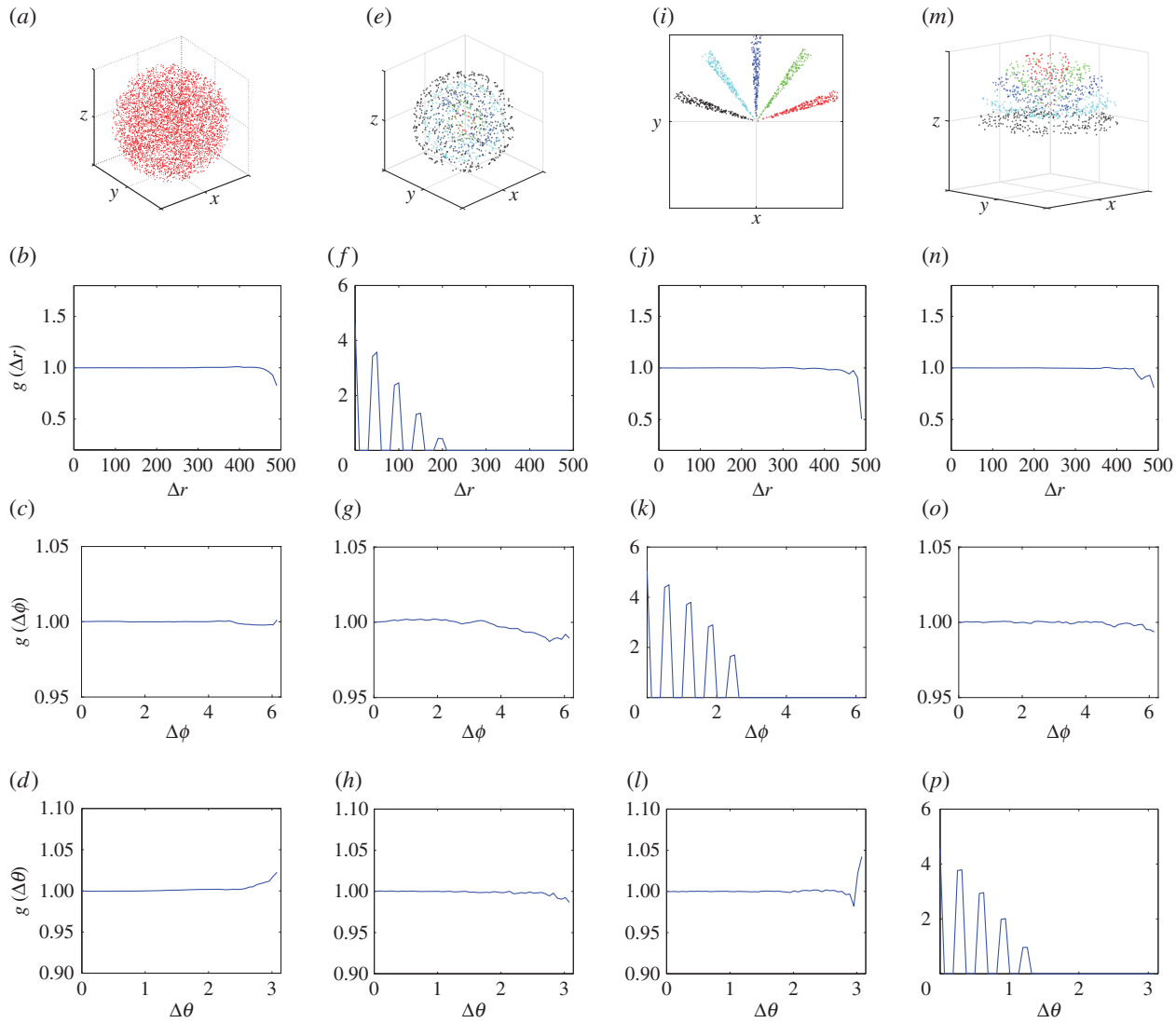

Figure 2. CSR and regular point patterns, $h = L/50$. (*a*) CSR pattern, (*e*) segregated clusters of points in spherical shells, (*i*) segregated clusters of points that are locally aggregated around prescribed angles of ϕ, and (*m*) segregated clusters of points in conical shells. In the CSR pattern, $N = 5000$ and $R = 500$. In all of the regular patterns, $N = 1000$. The centres of clusters are equally spaced with distance (*e*) $R/10$, (*i*) $\pi/5$, (*m*) $\pi/10$, and the points are distributed uniformly in each cluster in an interval of size (*e*) $R/100$, (*i*) $\pi/50$, (*m*) $\pi/50$. The remaining rows are for the averaged non-periodic PCF from 1000 simulated patterns. (*b,f,j,n*) Radial projection. (*c,g,k,o*) Azimuthal projection. (*d,h,l,p*) Polar projection. (Online version in colour.)

This implies that the number of pairs of points separated by a distance $L - \Delta a$ will be added to the number of pairs of points separated by a distance Δa (for $\Delta a \leq L/2$). Therefore, we also need to re-formulate the normalization term for the periodic PCF, which yields

$$\bar{G}_p(\Delta a) = \frac{2}{h}\int_0^{L-\Delta a} f_a(a) \int_{a+\Delta a}^{a+\Delta a+h} f_a(t)\,\mathrm{d}t\,\mathrm{d}a + \frac{2}{h}\int_0^{\Delta a} f_a(a) \int_{a+L-\Delta a-h}^{a+L-\Delta a} f_a(t)\,\mathrm{d}t\,\mathrm{d}a. \tag{2.8}$$

Equations (2.7) and (2.8) then give the periodic PCF

$$g_p(\Delta a) = \frac{G_p(\Delta a)}{\bar{G}_p(\Delta a)}. \tag{2.9}$$

When evaluating the periodic PCF, equations (2.9), for the simulated CSR pattern in figure 2, we observed a reduction in the deviations from unity at large distances (results not shown).

For the azimuthal projection, the periodic distance that separates two points has a clear physical interpretation. It is simply the acute angle that separates the two points a_i and a_j ($i \neq j$). However, the physical interpretation of the distance at which departures from unity occur in the periodic PCF for the radial and polar projections are, in isolation, unclear. This is because the periodic PCF cannot distinguish the distances $L - \Delta a$ and Δa. Therefore, we evaluate both the non-periodic PCF and periodic PCF for synthetic (§3) and experimental datasets of cell distributions within a tumour spheroid (§4).

3. Synthetic tumour spheroids

We first analyse synthetic tumour spheroid datasets, to help with interpreting the PCFs for the nine experimental datasets. A spatial model is used to generate point patterns with two zones of uniform density. For each synthetic dataset, a total of N points are distributed within a sphere of radius R. The points are distributed under the assumption that there are two zones of uniform cell density, partitioned by a radial distance, $r = B$, representing the necrotic zone boundary. The inner zone, $r < B$, is the necrotic zone of the synthetic tumour. The outer zone, $B < r < R$, corresponds to a viable zone (i.e. the quiescent and proliferative zones together).

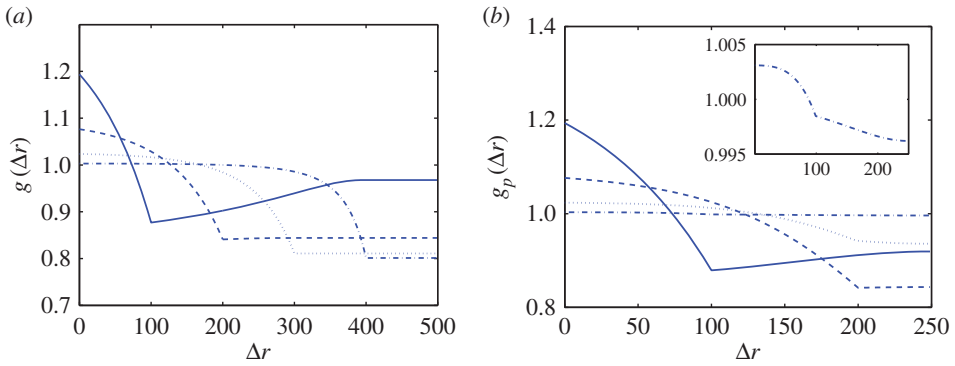

Figure 3. Analytic PCFs for synthetic tumour spheroids, $N = 5000$, $\Delta N = 1000$, $R = 500$ and $h = R/50$. Solid curves, $B = 400$. Dashed curves, $B = 300$. Dotted curves, $B = 200$. Dashed-dotted curves, $B = 100$. (*a*) Non-period PCF. (*b*) Periodic PCF. The inner graph in (*b*) is for $B = 100$ and it illustrates the small variations of g_p in this case. (Online version in colour.)

To ensure that the viable zone has a larger cell density than the necrotic zone, we distribute uniformly at random $N - \Delta N$ points throughout the entire spatial domain, along with additional ΔN points in the viable zone only. There is no spatial structure in the azimuthal and polar projections of the synthetic datasets. Therefore, we analyse the radial projection of the synthetic datasets with the non-periodic and periodic PCF.

As the point process is known for the synthetic tumour spheroids, we can derive analytic expressions for the non-periodic and periodic PCF, see electronic supplementary material, appendix A. Results for four values of the necrotic boundary are shown in figure 3. We observe that the distance at which a slope discontinuity occurs in the non-periodic PCF (figure 3*a*), $\Delta r = s$, uniquely identifies the width of the viable zone, $W = s$. This means that the necrotic boundary, $B = R - W$, can be identified from the non-periodic PCF. But the slope discontinuity in the periodic PCF (at $\Delta r = s$ in figure 3*b*) identifies either the width of the viable zone, $W = s$, or the necrotic boundary, $B = s$, with the non-uniqueness being due to periodicity, as discussed in §2. Therefore, we cannot uniquely determine the necrotic zone boundary from the distance at which the slope discontinuity occurs in the periodic PCF.

We now examine estimates of the non-periodic and periodic PCF from samples of 1000 synthetic tumour spheroids. The central slice from synthetic tumour spheroids for four values of the necrotic boundary are shown in the first row of figure 4, where the difference in density between the necrotic and viable zones is visually undetectable (i.e. visually indistinguishable from CSR spatial patterning). In the panels of the second and third row, directly below each test pattern, are the corresponding average non-periodic and periodic PCFs (solid curves). The upper and lower broken curves are the 95% confidence intervals, which are the 97.5 and 2.5 percentiles of the 1000 simulations. The arrows identify *critical points* in the estimates of the PCFs, similar to that seen for the points of slope discontinuity in the continuous PCFs (figure 3).

At short to moderate distances, we find a comparable amount of (small) variability in both the non-periodic and periodic PCFs (broken curves, second and third rows, figure 4). However, at moderate to large distances there is a much greater variability in the non-periodic PCF than that of the periodic PCF. This implies that we have greater confidence inferring salient features of the periodic PCF at large distances. It is therefore advantageous to examine both PCFs together when assessing the spatial pattern of tumours.

To assess the PCFs of the synthetic tumour spheroids, it is useful to consider statistical significance of the results, i.e. our confidence in distinguishing the results from the CSR state (null case). This is done by presenting the 2.5 and 97.5 percentiles of the PCFs of simulated CSR point patterns (the grey regions in the second and third rows of figure 4). In each case, 1000 CSR point patterns are generated with the same number of points as the synthetic spheroid. Then, a point pattern's PCF signal found within the grey region could be interpreted as not distinctive from the CSR point process. Therefore, this makes it difficult to estimate the critical points of the PCFs with just one point pattern (or a small sample size) for $B = 100$ in figure 4*b*,*c*, where the broken curves essentially bound the grey CSR regions. By contrast, the two broken curves at the critical points in figure 4*j*,*k*,*n*,*o* both lie below the grey CSR region. This indicates that it is possible to provide an estimate of the necrotic boundary from just one point pattern in the cases when $B = 300$ and $B = 400$. We note that this contrast is mainly due to the difference in density between the necrotic and viable zones, rather than the increase in the necrotic boundary (see electronic supplementary material, appendix B, for a fixed value of $B = 200$ and varying ΔN).

The statistical significance of the PCF results is further examined by comparison to those for the normalized density [35]. This first-order statistic is derived by considering the proportion of points at a distance r from the origin

$$F(r) = \frac{(1/h)\sum_{i=1}^{N} \mathbf{1}_{(0,h]}(r_i - r)}{N}. \tag{3.1}$$

Equation (3.1) is normalized by the radial projection of a Poisson process

$$\bar{F}(r) = \frac{1}{h}\int_r^{r+h} \frac{3t^2}{R^3}\,\mathrm{d}t = \frac{(r+h)^3 - r^3}{hR^3},$$

to obtain the normalized density

$$f(r) = \frac{F(r)}{\bar{F}(r)}. \tag{3.2}$$

In the bottom row of figure 4, three of the four critical points lie within the grey CSR region and the broken curves essentially bound the grey CSR region. This illustrates that it is difficult to provide an estimate of the necrotic zone boundary from just one point pattern (or small sample size) with the normalized density, demonstrating a greater confidence in the statistical significance of the second-order PCF method.

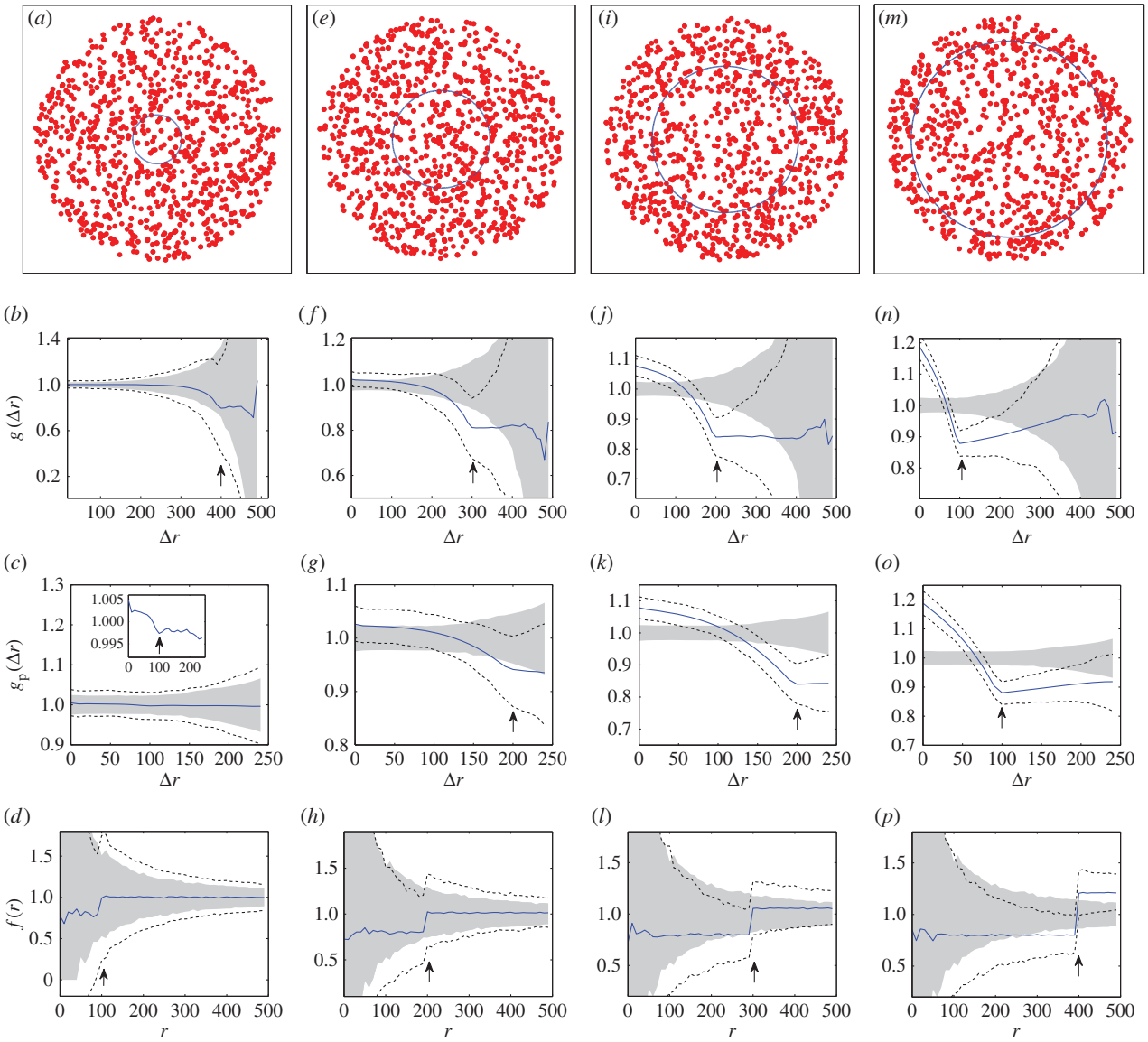

Figure 4. Synthetic tumour spheroid point patterns, $N = 5000$, $\Delta N = 1000$, $R = 500$ and $h = R/50$. Columns: left to right $B = \{100, 200, 300, 400\}$. (*a,e,i,m*) Central slice of synthetic tumour spheroid. The remaining rows are for statistics from 1000 realizations. (*b,f,j,n*) Average non-periodic PCF, solid curves. (*c,g,k,o*) Average periodic PCF, solid curves. (*d,h,l,p*) Average normalized density, solid curves. The upper and lower broken curves are for the 97.5 and 2.5 percentiles, and the arrows are for the known distances of the critical points. The shaded region is for the 97.5 and 2.5 percentiles of 1000 CSR point patterns. (Online version in colour.)

4. Real tumour spheroids

We have shown that the PCF method can provide a reliable estimate of the (known) radial distance of the necrotic zone boundary, B, from the centre of a synthetic tumour spheroid. However, in practice, there are two important distinctions to consider when examining real tumour spheroids.

The first distinction is that real tumour spheroids can be ellipsoidal in shape, for example, see figure 5. Therefore, we consider an ellipsoidal spatial domain centred at the origin of a Cartesian coordinate system and aligned with the axes,

$$\frac{x^2}{X^2} + \frac{y^2}{Y^2} + \frac{z^2}{Z^2} \leq 1,$$

where X, Y, Z are the lengths of the three semi-principle axes. We now project, or map the points to the interval [0, 1] via

$$a = \sqrt{\frac{x^2}{X^2} + \frac{y^2}{Y^2} + \frac{z^2}{Z^2}}, \quad \text{with } 0 \leq a \leq 1. \tag{4.1}$$

The previous PCF analysis for the radial distance of spherical synthetic tumours holds for this non-dimensional problem, with $r = a$ and $R = 1$ in equation (2.5). Therefore, we can estimate the dimensionless necrotic zone boundary, $\tilde{B} \leq 1$, for this projection of the data points. The inverse mapping then provides estimates for the lengths of the three semi-principle axes of the ellipsoidal necrotic zone boundary

$$B_X = \tilde{B}X, \quad B_Y = \tilde{B}Y \quad \text{and} \quad B_Z = \tilde{B}Z. \tag{4.2}$$

The second distinction is that the necrotic zone boundary in a real tumour spheroid is unknown, but it can be estimated by a human expert. Without prior knowledge of the PCF estimates, human expert estimates for the necrotic zone boundary in nine tumour spheroids were obtained (table 2). In addition to this, we automate the data collection process and subsequent evaluation of the PCF estimates of the necrotic zone boundary for each tumour (see §§4.1 and 4.2). Together, this allows for an unbiased comparison between

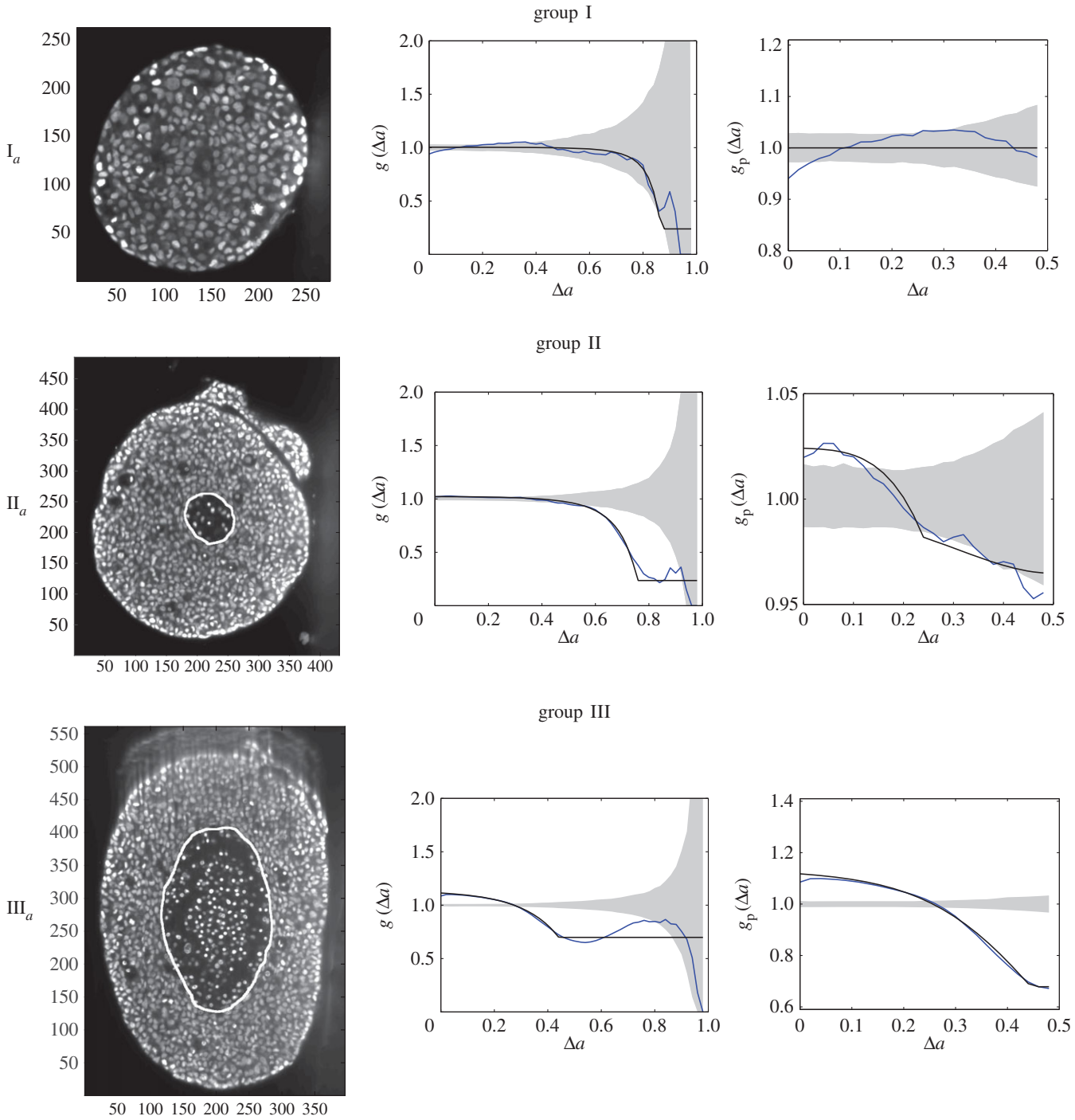

Figure 5. Central image slice and corresponding point pattern analysis for one tumour spheroid from each of the three groups. The rows top to bottom correspond to I_a, II_a and III_a, respectively. The first column is the central slice of the tumour spheroids. The unit of the length in the images is a micrometre. The white curves superimposed on the central image slices outline the necrotic zone identified by a human assessor. The second and third columns are the non-periodic PCF and periodic PCF respectively, $h = 0.02$. The blue curves are for the PCF statistics. The black curves are the analytic PCFs best fit to the statistics (blue curves), using a nonlinear least-squares method. The shaded region is for the 97.5 and 2.5 percentiles of 1000 CSR point patterns. (Online version in colour.)

the human expert and PCF estimates of the necrotic zone boundary for each tumour (tables 2 and 3).

4.1. Experiments and data collection

Nine homotypic tumour spheroids were formed from the breast cancer cell line T47D by the liquid overlay method [36]. An initial number, N_s, of seed cells (table 1) together with medium were placed in a convex well, which is coated with a non-adhesive layer. Mature spheroids were obtained after being cultured for 12 days in the incubator, each with N_e number of cells (table 1). The spheroids were then removed from the wells, fixed, stained with Draq 5 to label the nuclei, and optically cleared. Imaging of the spheroids was performed with a digital scanned laser light-sheet fluorescence microscope [37]. They were categorized, by visual inspection, into three groups: (i) spheroids with no visible necrotic core, (ii) spheroids with a small necrotic core and, (iii) spheroids with a large necrotic core. A central slice of the image stack of one tumour spheroid from each of the three groups is shown in figure 5, with central slices from the remaining six tumour spheroids presented in electronic supplementary material, figure S2 (appendix C).

The point patterns or raw datasets are the positions of the nuclei of the cells. Detecting the cell nuclei in the

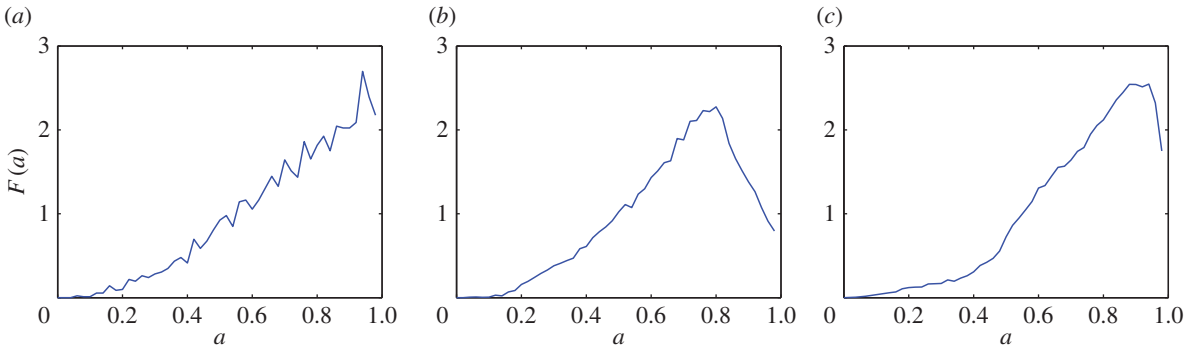

Figure 6. Removal of data points associated with the irregularities of the surface for three tumour spheroids. Points with $a > \tilde{A}$ are removed from the raw datasets. (*a*) Tumour spheroid I_a, $\tilde{A} = 0.94$. (*b*) Tumour spheroid II_a, $\tilde{A} = 0.80$. (*c*) Tumour spheroid III_a, $\tilde{A} = 0.94$. (Online version in colour.)

Table 1. Data for nine tumour spheroids. Group I: no visible necrotic core. Group II: small necrotic core. Group III: large necrotic core. *X*, *Y*, *Z* are the lengths of the three semi-principle axes in micrometres for each ellipsoidal spatial domain. N_s is the number of initially seeded cells for each tumour spheroid. N_e is the total number of cells in the raw dataset for each mature tumour spheroid. *N* is the number of cells in each data subset (i.e. ellipsoidal spatial domain) used in the spatial analysis.

	group I			group II			group III		
	I_a	I_b	I_c	II_a	II_b	II_c	III_a	III_b	III_c
X	93	102	114	173	173	145	169	190	169
Y	115	107	149	190	176	201	250	196	233
Z	110	88	133	145	159	178	218	165	180
N_s	500	1000	2000	5000	5000	5000	10 000	10 000	10 000
N_e	4597	3983	10 334	25 806	23 739	25 943	36 732	32 916	35 200
N	3607	3127	7900	16 356	15 593	17 864	30 555	21 226	24 729

three-dimensional images of the spheroids is achieved by applying a custom multi-scale Laplacian of Gaussian (LoG) detection algorithm [38]. Further details of the detection method can be found in the electronic supplementary material, appendix D.

Subsets of the raw data points are obtained by removing data points associated with the irregularities of the surface of the tumour spheroids. For each data subset, a spatial domain is defined by an ellipsoid centred at the origin with the three semi-principal axes aligned with the Cartesian axes. This is done systematically. (i) Find the smallest convex set of points (i.e. the convex hull) that contains all the raw data points [39]. (ii) Fit an ellipsoid to the convex hull, using a linear least-squares algorithm [40], to obtain initial estimates of the lengths of the three semi-principal axes, X^*, Y^* and Z^*, with $X^* \leq Y^*$. (iii) The origin is chosen as the centre of mass in each dataset. The MATLAB function #pca is used to find three orthogonal directions (principle components) in which each dataset has the largest variances. The point pattern is then rotated so that the three principle axes of each dataset coincide with the Cartesian coordinate system [41]. (iv) The data points are projected onto the interval [0, 1] using equation (4.1), with $X = X^*$, $Y = Y^*$ and $Z = Z^*$. The global maximum of $F(a)$, given by equation (3.1), provides an estimate of the non-dimensional distance, $a = \tilde{A}$, associated with the surface irregularities of each tumour spheroid (figure 6). (v) Data points with $a > \tilde{A}$ are removed from each raw dataset and the inverse map, equation (4.2), provides the lengths of the three semi-principal axes, $X = \tilde{A}X^*$, $Y = \tilde{A}Y^*$ and $Z = \tilde{A}Z^*$, (with $X \leq Y$) for the ellipsoidal spatial domains (table 1).

The subsets of *N* data points for each of the nine tumours (see electronic supplementary material, appendix E) are subsequently analysed with the PCF methods.

4.2. Spatial analysis

The nine subsets of data points, with ellipsoidal spatial domains, are mapped to the interval [0, 1] using equation (4.1). The non-periodic and periodic PCFs (blue curves) are shown in figure 5, and electronic supplementary material figure S2 (appendix C). The PCFs for the real tumour spheroids can be examined in a similar way to that of the synthetic tumour spheroids, as discussed in §3. However, unlike the synthetic tumour spheroids, the non-dimensional necrotic zone boundary, $\tilde{B}$, is unknown in the real tumour spheroids. Additionally, it is difficult to provide an estimate of the necrotic zone boundary by visual examination of the PCFs alone. To overcome this difficulty, and to automate the PCF estimates of necrotic zone boundary, we fit the analytical PCFs (see electronic supplementary material, appendix A) to the statistical PCFs. A nonlinear least-squares method with two parameters ΔN and $\tilde{B}$ is used to find the best fit (e.g. see black curves in figure 5). The point at which there is slope discontinuity in the fitted PCF is taken as the critical point used in determining the estimates for the necrotic zone boundary, $\tilde{B}$. Equation (4.2) gives estimates for the lengths, B_X, B_Y and B_Z, of the three semi-principle axes of the ellipsoidal necrotic zone boundary in each tumour spheroid. The two lengths B_X and B_Y (with $B_X \leq B_Y$) for each tumour spheroid are recorded in the second and third row of table 2.

Table 2. Necrotic zone boundary estimates in micrometres for the nine tumour spheroids. The italicized PCF estimates are for fitted PCFs (and therefore critical points) that reside within the 97.5 and 2.5 percentiles of 1000 CSR point patterns (e.g. see first row of results for spheroid I_a in figure 5).

		group I			group II			group III		
		I_a	I_b	I_c	II_a	II_b	II_c	III_a	III_b	III_c
human	B_X	—	—	—	36	36	34	81	83	93
	B_Y	—	—	—	43	40	53	138	90	128
non-periodic PCF	B_X	*12*	*6*	*23*	42	46	34	94	92	96
	B_Y	*15*	*6*	*30*	46	47	46	140	94	132
periodic PCF	B_X	—	—	*52*	42	49	30	94	101	95
	B_Y	—	—	*69*	46	50	42	139	104	131
DBSCAN	B_X	84	—	—	43	47	33	77	74	80
	B_Y	103	—	—	47	48	45	114	77	110

Table 3. Percentage difference in computed estimates relative to the human estimates, for group II and III tumour spheroids. The smallest (absolute) percentage differences are italicized. The rightmost column shows the average (absolute) percentage difference for each method.

		group II			group III			
		II_a (%)	II_b (%)	II_c (%)	III_a (%)	III_b (%)	III_c (%)	average (%)
non-periodic PCF	B_X	*15.9*	*26.9*	*0.1*	16.5	10.7	3.2	*12.2*
	B_Y	*7.6*	*18.2*	*−12.5*	1.4	*4.7*	3.5	*8*
periodic PCF	B_X	*15.9*	35.2	−11.8	16.5	21.6	*2.1*	17.2
	B_Y	*7.6*	25.8	−20.1	*0.6*	15.9	*2.7*	12.1
DBSCAN	B_X	18.1	30.7	−4.2	*−4.1*	*−10.6*	−13.8	13.6
	B_Y	10.0	21.2	−14.1	−17.1	−14.7	−13.4	15.1

The italicized PCF estimates in table 2 are for fitted PCFs (and therefore critical points) that reside within the CSR shaded regions (e.g. see black curves for spheroid I_a in figure 5). This means that we can reject these estimates as they may be considered as not distinctive from the CSR point process. This is the case for all three group I tumour spheroids with no visible necrotic cores, and because the group I tumour spheroids are the smallest in size it is reasonable to infer that the innermost cells still have sufficient nutrient and oxygen to remain viable [42].

We compare the non-periodic and periodic PCF estimates for B_X and B_Y with those of a human. The human estimates are based on visual examination of a central image slice of each tumour, where white curves are superimposed onto the images to identify the necrotic core boundary in each tumour spheroid (e.g. see central image slices in figure 5). Fitting an ellipse to the white curves in each image then provides human estimates for B_X and B_Y, which are shown in the first row of table 2.

Recorded in table 3 (first and second rows) are the percentage difference in the PCF estimates relative to the human estimates for the group II and III tumour spheroids, with visible necrotic cores. The averaged results (rightmost column) show that the non-periodic PCF estimates have the smallest (absolute) percentage difference, when compared with those of the human. A similar result is found when calculating the overall mean squared error (MSE) for each PCF method; non-periodic PCF MSE = 45 and periodic PCF MSE = 96. We believe that the main difference between the PCF and human estimates can be attributed to the fact that the human estimates are based on a single two-dimensional central slice of each spheroid, whereas the PCF estimates are based on the three-dimensional point pattern data of each spheroid. Other contributing factors in the percentage difference are likely to include the processing method of the raw data and the spatial model used for the estimation of the critical point in the PCFs. We also note that there appears to be a positive bias in the estimates (21 out of 24 in table 3), and further investigation of this is left to future research.

To conclude the analysis, we implement an existing method commonly used to evaluate spatial clustering, using the three-dimensional point pattern data of each spheroid. The density-based spatial clustering of applications with noise (DBSCAN) algorithm classifies points in high-density regions (e.g. viable zone) as a cluster [16]. Points that are in low-density regions (e.g. necrotic zone) are classified as outliers. A subset of outliers for each tumour spheroid is used to calculate the DBSCAN estimates shown in the bottom row of table 2 (see electronic supplementary material, appendix F). The difference in the DBSCAN and human estimates is comparable to the difference in the PCF and human estimates (see bottom row of table 3), with an overall MSE = 138. The results demonstrate that the PCF method is a potentially useful alternative to

existing standard clustering methods in providing estimates for the necrotic zone boundary in tumour spheroids.

5. Discussion

We have derived one-dimensional non-periodic and periodic PCFs for the radial, azimuthal and polar projections of a point pattern. This is different from the usual Euclidean distance based PCF [25–29]. Analysis of spatial structure with one-dimensional PCFs is useful in situations where the Euclidean distance is not the main quantity of interest, as is the case in identifying the necrotic zone boundary in tumour spheroids.

We analysed synthetic tumour spheroids (i.e. simulated datasets) with our PCFs to illustrate the methodology for identifying the necrotic zone boundary. It was found that a critical point (at which there is a slope discontinuity) in the non-periodic PCFs corresponds to the width of the viable zone in synthetic tumour spheroids, and this allows us to identify the necrotic zone boundary. However, for small sample sizes, when the non-periodic PCF is used, the critical point can be obscured by noise (at large distances). Using the periodic PCF, we reduced the noise in the PCFs, but with the drawback that the critical point corresponds to either the width of the viable zone, or the necrotic zone boundary itself. However, by examining both the non-periodic and periodic PCFs, we are able to provide a more statistically significant (i.e. it lies farther from CSR) estimate of the necrotic zone boundary than that obtained using the normalized cell density.

The PCF method was modified to provide estimates for the lengths of the three semi-principle axes of an ellipsoidal shaped necrotic zone boundary, and then applied to three-dimensional point patterns of nine experimental tumour spheroids. The PCF estimates were compared with those of a human and the DBSCAN method [16]. The primary difference in the computed and human estimates was attributed to the human estimates being based only on a two-dimensional slice (image) of each tumour spheroid. The average percentage differences of the PCF and DBSCAN method were comparable (see rightmost column in table 3), and this demonstrates that the PCF method potentially has merit as an alternative to the existing DBSCAN method.

The focus of this study has been to provide estimates for the radial distance(s) of the necrotic core boundary from the centre of a tumour spheroid. These estimates provide a simple metric to classify and categorize tumour spheroids, which has potential application to high-throughput comparative assays [8–10]. For example, the PCF method could be used to investigate population-level variability in the size of the necrotic zone by using a larger sample of mature tumour spheroids from the same cell line, grown from the same number of seeded cells. This would allow one to study differences in nutrient consumption between cell types, or the effects of different cell culture methods, or drug treatments. The automatic evaluation of PCF estimates has clear advantage over manual human estimates in the assessment of such high-throughput comparative assays.

The PCF method also has two main advantages over the DBSCAN method. The first being that the DBSCAN method is semi-automatic, requiring human input for each tumour spheroid analysed. The second advantage is that only the PCF method provides reliable estimates for point pattern data that is visually indistinguishable from the CSR point process (see electronic supplementary material, appendix F).

Although this work is concerned with homotypic spheroids there are approaches that aim at more complex spheroids, including different cell types and heterogeneities in the microenvironment. Our data analysis can be readily applied to datasets from such complex spheroids. Furthermore, adjusting the segementation method would allow for the extraction of cell position information from histological stains and enable the analysis of sections of cancer patient tumours. Therefore, the PCF method could potentially have a role in diagnostic testing and personalized cancer treatment.

More generally, and in addition to the estimates for the necrotic zone boundary, it is important to understand that the PCFs can provide multi-scale spatial information on tumour spheroids (e.g. figure 2). Previous studies have shown that the PCF is close to a sufficient summary statistic, essentially capturing all the spatial information in a given point pattern [34,43]. Therefore, our PCF method could be implemented in combination with inferencing algorithms such as approximate Bayesian computation [43], which require close to sufficient summary statistics, to parametrize tumour growth models for specific cell types and culture conditions [11–14].

Competing interests. We declare we have no competing interests.

Funding. The work of S.D. and J.E.F.G. was supported by an Australian Research Council Discovery Early Career Researcher Award (DE130100031) to J.E.F.G. S.D. also acknowledges a University of Adelaide Full Fees Scholarship. B.J.B. was supported by a National Health and Medical Research Council Project grant no. (APP1069757) and Australian Research Council Discovery Project Grant (DP160102644). N.G.B. was supported by the Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers (CE140100049). S.C.F., C.M., A.S. and E.H.K.S. are supported by the Deutsche Forschungsgemeinschaft (DFG) and the Cluster of Excellence for Macromolecular Complexes (CEF-MC II, EXC115).

Authors' contributions. The experiments were conceived and designed by C.M. and E.H.K.S. The experiments were performed by C.M. and the data analysis performed by A.S. The PCF method was conceived and designed by S.D., B.J.B., J.E.F.G. and N.G.B., with input from S.C.F. The mathematical derivations, numerical computations and interpretation of results were carried out by S.D. under the guidance of B.J.B., J.E.F.G., N.G.B. and S.C.F. The manuscript was written by S.D., B.J.B., J.E.F.G., S.C.F. and N.G.B. All authors edited the manuscript and approved submission.

References

1. Hirschhaeuser F, Menne H, Dittfeld C, West J, Mueller-Klieser W, Kunz-Schughart LA. 2010 Multicellular tumor spheroids: an underestimated tool is catching up again. *J. Biotechnol*. **148**, 3–15. (doi:10.1016/j.jbiotec.2010.01.012)
2. Mueller-Klieser W. 1987 Multicellular spheroids. *J. Cancer Res. Clin. Oncol*. **113**, 101–122. (doi:10.1007/BF00391431)
3. Sutherland RM. 1988 Cell and environment interactions in tumor microregions: the multicell spheroid model. *Science* **240**, 177–184. (doi:10.1126/science.2451290)
4. Loessner D, Stok KS, Lutolf MP, Hutmacher DW, Clements JA, Rizzi SC. 2010 Bioengineered 3D platform to explore cell–ECM interactions and drug

resistance of epithelial ovarian cancer cells. *Biomaterials* **31**, 8494–8506. (doi:10.1016/j.biomaterials.2010.07.064)

5. Blacher S *et al.* 2014 Cell invasion in the spheroid sprouting assay: a spatial organisation analysis adaptable to cell behaviour. *PLoS ONE* **9**, e97019. (doi:10.1371/journal.pone.0097019)
6. Vinci M *et al.* 2012 Advances in establishment and analysis of three-dimensional tumor spheroid-based functional assays for target validation and drug evaluation. *BMC Biol*. **10**, 29. (doi:10.1186/1741-7007-10-29)
7. Dairkee SH, Deng G, Stampfer MR, Waldman FM, Smith HS. 1995 Selective cell culture of primary breast carcinoma. *Cancer Res*. **55**, 2516–2519.
8. Deisboeck T, Berens M, Kansal A, Torquato S, Stemmer-Rachamimov A, Chiocca E. 2001 Pattern of self-organization in tumour systems: complex growth dynamics in a novel brain tumour spheroid model. *Cell Prolif*. **34**, 115–134. (doi:10.1046/j.1365-2184.2001.00202.x)
9. Dufau I *et al.* 2012 Multicellular tumor spheroid model to evaluate spatio-temporal dynamics effect of chemotherapeutics: application to the gemcitabine/CHK1 inhibitor combination in pancreatic cancer. *BMC Cancer* **12**, 15. (doi:10.1186/1471-2407-12-15)
10. Garg S, Fischer S, Schuman E, Stelzer E. 2015 Lateral assembly of N-cadherin drives tissue integrity by stabilizing adherens junctions. *J. R. Soc. Interface* **12**, 20141055. (doi:10.1098/rsif.2014.1055)
11. Araujo R, McElwain D. 2004 A history of the study of solid tumour growth: the contribution of mathematical modelling. *Bull. Math. Biol*. **66**, 1039–1091. (doi:10.1016/j.bulm.2003.11.002)
12. Casciari JJ, Sotirchos SV, Sutherland RM. 1992 Mathematical modelling of microenvironment and growth in EMT6/Ro multicellular tumour spheroids. *Cell Prolif*. **25**, 1–22. (doi:10.1111/j.1365-2184.1992.tb01433.x)
13. Grimes DR, Kelly C, Bloch K, Partridge M. 2014 A method for estimating the oxygen consumption rate in multicellular tumour spheroids. *J. R. Soc. Interface* **11**, 20131124. (doi:10.1098/rsif.2013.1124)
14. Stein AM, Demuth T, Mobley D, Berens M, Sander LM. 2007 A mathematical model of glioblastoma tumor spheroid invasion in a three-dimensional *in vitro* experiment. *Biophys. J*. **92**, 356–365. (doi:10.1529/biophysj.106.093468)
15. Stelzer EHK. 2015 Light-sheet fluorescence microscopy for quantitative biology. *Nat. Methods* **12**, 23–26. (doi:10.1038/nmeth.3219)
16. Ester M, Kriegel HP, Sander J, Xu X. 1996 A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. of the Second International Conf. on Knowledge Discovery and Data Mining* (*KDD*), *Portland*, *Oregon*, vol. 96, pp. 226–231. See http://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf.
17. Hartigan JA, Wong MA. 1979 Algorithm AS 136: A K-means clustering algorithm. *Appl. Stat*. **28**, 100–108. (doi:10.2307/2346830)
18. Jain AK, Dubes RC. 1988 *Algorithms for clustering data*. Upper Saddle River, NJ: Prentice-Hall, Inc.
19. Agnew DJG, Green JEF, Brown TM, Simpson MJ, Binder BJ. 2014 Distinguishing between mechanisms of cell aggregation using pair-correlation functions. *J. Theor. Biol*. **352**, 022705.
20. Binny RN, Plank MJ, James A. 2015 Spatial moment dynamics for collective cell movement incorporating a neighbour-dependent directional bias. *J. R. Soc. Interface* **12**, 20150228. (doi:10.1098/rsif.2015.0228)
21. Fozard JA, Kirkham GR, Buttery LD, King JR, Jensen OE, Byrne HM. 2011 Techniques for analysing pattern formation in populations of stem cells and their progeny. *BMC Bioinform*. **12**, 396. (doi:10.1186/1471-2105-12-396)
22. Mattfeldt T, Gottfried HW, Frey H, Vogel U. 1993 Second-order stereology of prostatic adenocarcinoma and normal prostatic tissue. *Acta Stereol*. **12**, 203–208.
23. Riedel IH, Kruse K, Howard J. 2005 A self-organized vortex array of hydrodynamically entrained sperm cells. *Science* **309**, 300–303. (doi:10.1126/science.1110329)
24. Treloar KK, Simpson MJ, Binder BJ, McElwain DS, Baker RE. 2014 Assessing the role of spatial correlations during collective cell spreading. *Sci. Rep*. **4**, 5713. (doi:10.1038/srep05713)
25. Chandler D, Percus JK. 1988 Introduction to modern statistical mechanics. *Phys. Today* **41**, 114. (doi:10.1063/1.2811680)
26. Dieckmann U, Law R, Metz JA. 2000 *The geometry of ecological interactions: simplifying spatial complexity*, vol. 1. Cambridge, UK: Cambridge University Press.
27. Illian J, Penttinen A, Stoyan H, Stoyan D. 2008 *Statistical analysis and modelling of spatial point patterns*, vol. 70. Hoboken, NJ: John Wiley & Sons.
28. Jones BJ, Martínez VJ, Saar E, Trimble V. 2005 Scaling laws in the distribution of galaxies. *Rev. Mod. Phys*. **76**, 1211. (doi:10.1103/RevModPhys.76.1211)
29. Martínez VJ, Saar E. 2010 *Statistics of the galaxy distribution*. Boca Raton, FL: CRC Press. See http://www.crcnetbase.com/doi/book/10.1201/9781420036169.
30. Ripley BD. 1976 The second-order analysis of stationary point processes. *J. appl. Probability* **13**, 255–266. See https://www.cambridge.org/core/journals/journal-of-applied-probability/article/the-second-order-analysis-of-stationary-point-processes/63ECBD08B25F19C844C8FEC5B88C5266.
31. Ohser J, Mücklich F. 2000 *Statistical analysis of microstructures in materials science*. Chichester, UK: Wiley. See http://au.wiley.com/WileyCDA/WileyTitle/productCd-0471974862.html.
32. Stoyan D, Stoyan H. 1996 Estimating pair correlation functions of planar cluster processes. *Biom. J*. **38**, 259–271. (doi:10.1002/bimj.4710380302)
33. Stoyan D, Stoyan H. 2000 Improving ratio estimators of second order point process characteristics. *Scand. J. Stat*. **27**, 641–656. (doi:10.1111/1467-9469.00213)
34. Binder BJ, Simpson MJ. 2013 Quantifying spatial structure in experimental observations and agent-based simulations using pair-correlation functions. *Phys. Rev. E* **88**, 022705. (doi:10.1103/PhysRevE.88.022705)
35. Binder BJ, Sundstrom JF, Gardner JM, Jiranek V, Oliver SG. 2015 Quantifying two-dimensional filamentous and invasive growth spatial patterns in yeast colonies. *PLoS Comput. Biol*. **11**, e1004070. (doi:10.1371/journal.pcbi.1004070)
36. Carlsson J, Yuhas JM. 1984 Liquid-overlay culture of cellular spheroids. In *Spheroids in cancer research: methods and perspectives* (eds H Acker, J Carlsson, R Durand, RM Sutherland), pp. 1–23. Berlin, Germany: Springer. See http://dx.doi.org/10.1007/978-3-642-82340-4_1.
37. Keller PJ, Schmidt AD, Wittbrodt J, Stelzer EH. 2008 Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science* **322**, 1065–1069. (doi:10.1126/science.1162493)
38. Lowe DG. 2004 Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis*. **60**, 91–110. (doi:10.1023/B:VISI.0000029664.99615.94)
39. De Berg M, Van Kreveld M, Overmars M, Schwarzkopf OC. 2000 *Computational geometry*. Berlin, Germany: Springer.
40. Rosin PL. 1993 A note on the least squares fitting of ellipses. *Pattern Recognit. Lett*. **14**, 799–808. (doi:10.1016/0167-8655(93)90062-I)
41. Jolliffe I. 2014 *Principal component analysis*. Hoboken, NJ: John Wiley & Sons, Ltd.
42. Greenspan HP. 1972 Models for the growth of a solid tumor by diffusion. *Stud. Appl. Math*. **51**, 317–340. (doi:10.1002/sapm1972514317)
43. Johnston ST, Simpson MJ, McElwain DS, Binder BJ, Ross JV. 2014 Interpreting scratch assays using pair density dynamics and approximate Bayesian computation. *Open Biol*. **4**, 140097. (doi:10.1098/rsob.140097)