British Journal of Cancer (2005), 1 – 6

www.bjcancer.com

Full Paper

Spread of human cancer cells occurs with probabilities indicative of a nongenetic mechanism

JS Michaelson*[,1,2,4], JA Cheongsiatmoy[2], F Dewey[2], MJ Silverstein[5], D Sgroi[1,4], B Smith[2,3,5] and KK Tanabe[2,3]

[1]Department of Pathology, Massachusetts General Hospital, Boston, MA, USA; [2]Department of Surgery, Massachusetts General Hospital, Boston, MA, USA; [3]Department of Surgery, Harvard Medical School, Boston, MA, USA; [4]Department of Pathology, Harvard Medical School, Boston, MA, USA; [5]Department of Surgery, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

There has been much uncertainty as to whether metastasis requires mutation at the time of spread. Here, we use clinical data to calculate the probability of the spread of melanoma and breast cancer cells. These calculations reveal that the probability of the spread of cancer cells is relatively high for small tumours (~1 event of spread for every 500 cells for melanomas of 0.1 mm) and declines as tumours increase in size (~1 event of spread for every 10^8 cells for melanomas of 12 mm). The probability of spread of breast cancer cells from the lymph nodes to the periphery is ~1 event of spread for every 10^8 cells in the nodal masses, which have a mean diameter of 5 mm, while the probability of spread of cancer cells from the breast to the periphery when the primary masses are 5 mm is also ~1 event of spread for every 10^8 cells. Thus, the occurrence of an event of spread from the breast to the lymph nodes appears not to increase the propensity of the progeny of those cells to spread from the lymph nodes to the periphery. These values indicate that the spread of human breast cancer and melanoma cells is unlikely to occur by a mechanism requiring mutation at the time of spread.
British Journal of Cancer advance online publication, 8 November 2005; doi:10.1038/sj.bjc.6602848 www.bjcancer.com

Keywords: metastasis; probability; mutation

There has been much uncertainty as to whether metastasis requires mutation at the time of spread (Cifone and Fidler, 1981; Fidler, 1983; Sobel, 1990; Welch *et al*, 2000; Yokota, 2000; Bernards and Weinberg, 2002; Couzin, 2003; Bernards, 2003; Van't Veer and Weigelt, 2003; Yang *et al*, 2004). Mutations have a number of characteristic features, in terms of the rates of their occurrence and other qualities, which are diagnostic: mutations are rare, a phenotype conferred on a cell by mutation is inherited by the progeny of the cell, and the rates of the appearance of phenotypes caused by mutations either remains constant over time for those phenotypes requiring only a single mutation, or increase in frequency for those phenotypes requiring the accumulation of multiple mutations. We have recently shown that from clinical data it is possible to measure the rates of metastatic spread, expressed in terms of the probability of spread per cell (Michaelson, 1999; Michaelson *et al*, 2002, 2003). Here we use this methology to measure the probability of spread per cell for human breast cancer and melanoma. The values of these probabilities are inconsistent with metastasis occurring by a process of mutation.

*Correspondence: Dr JS Michaelson, Division of Surgical Oncology, Cox Building Room 626, Massachusetts General Hospital, 100 Blossom Street, Boston, MA 02114, USA;
E-mail: michaelj@helix.mgh.harvard.edu
Revised 15 June 2005; accepted 7 September 2005

METHODS

Data

Data on the relationship between tumour size and breast cancer survival is from the USC/Van Nuys population (Silverstein, 2000; Michaelson *et al*, 2002, 2003), from Tabar *et al* (2000) and Tubiana and colleagues (Koscielny *et al*, 1984; Tubiana and Koscielny, 1990, 1991). For details and the general equivalence of these survival values (based on the 15-year Kaplan–Meier survival rate, based upon Karrison *et al* (1999) finding that it is not until this point in time that the survival rate become clear), see Michaelson *et al* (2002, 2003). Data on the relationship between tumour thickness and melanoma survival are 10-year Kaplan–Meier disease survival values from Balch *et al*, 2001).

Tumour diameters (breast cancer) and thickness (melanoma) were taken at pathological analysis. Since we shall be interested in tumour size in terms of the number of cells that they contain, N, we can generate rough estimates of the value of N that are quite satisfactory for our purposes here by converting values of tumour diameter or thickness, D, into values of cell number, N, assuming spherical geometry and a density of s (here we shall use 10^8 cells/cm^3 as a biologically plausible estimate of, s, as outlined in Boon *et al*, 1982; Pesce and Colacino, 1986; Van der Linden *et al*, 1986; and Michaelson *et al*, 2002). While this estimate of the value of s is biologically reasonable, for the purposes of the calculations made here, it need not be precise, as any error in the estimation of s by an order of magnitude or more will not change the general lessons

drawn here on the nature of cancer spread, although it will affect the precise values of the probability of spread.

Node-positive patients are defined as those patients with one or more lymph nodes found to have cancer upon pathological analysis. The range of nodes examined among women in the USC/Van Nuys populations varied from 1 to 47; however, only 10% of women had fewer than 10 nodes examined, while only 2% of women had fewer than five nodes examined. The mean and median number of nodes examined was 16.8 and 17, with a s.d. of 7.2.

Information on the size of the cancer metastases in the lymph nodes was collected from microscope slides from 16 node-positive patients, chosen at random, among MGH patients with invasive breast cancer diagnosed in 1993, among which there were 49 positive lymph nodes. Microscopic imagines of each node were

Table 1 Data and calculations of the probability of the spread of cancer cells

Cancer	Population	Size range	Nominal tumour size (mm)	Cell number (*N*)	Manifestation of metasatasis	Fraction of patients with a manifestation of metastasis (*L*) (%)	Probability of spread ($p = -\ln(1-L)/N$)
Breast cancer	Tabar *et al*	10–14	12	9.05E+07	Cancer death	13	$1/6.5 \times 10^{8}$
Breast cancer	Tabar *et al*	15–19	17	2.57E+08	Cancer death	20	$1/1.2 \times 10^{9}$
Breast cancer	Tabar *et al*	20–29	25	8.18E+08	Cancer death	45	$1/1.4 \times 10^{9}$
Breast cancer	Tabar *et al*	30–49	39	3.11E+09	Cancer death	56	$1/3.9 \times 10^{9}$
Breast cancer	Tubiana *et al*	10–25	17	2.57E+08	Cancer death	27	$1/8.2 \times 10^{8}$
Breast cancer	Tubiana *et al*	26–35	30	1.41E+09	Cancer death	42	$1/2.6 \times 10^{9}$
Breast cancer	Tubiana *et al*	36–45	40	3.35E+09	Cancer death	55	$1/4.1 \times 10^{9}$
Breast cancer	Tubiana *et al*	46–55	50	6.54E+09	Cancer death	66	$1/6.1 \times 10^{9}$
Breast cancer	Tubiana *et al*	56–65	60	1.13E+10	Cancer death	78	$1/7.4 \times 10^{9}$
Breast cancer	Tubiana *et al*	66–75	70	1.80E+10	Cancer death	83	$1/1.0 \times 10^{10}$
Breast cancer	Tubiana *et al*	76–85	80	2.68E+10	Cancer death	81	$1/1.6 \times 10^{10}$
Breast cancer	Tubiana *et al*	86–95	90	3.82E+10	Cancer death	92	$1/1.5 \times 10^{10}$
Breast cancer	Van nuys	10–14	12	9.05E+07	Cancer death	14	$1/6.0 \times 10^{8}$
Breast cancer	Van nuys	15–19	17	2.57E+08	Cancer death	28	$1/7.8 \times 10^{8}$
Breast cancer	Van nuys	20–29	25	8.18E+08	Cancer death	33	$1/2.0 \times 10^{9}$
Breast cancer	Van nuys	30–49	39	3.11E+09	Cancer death	46	$1/4.0 \times 10^{9}$
Breast cancer	Van nuys	15–19	17	2.57E+08	Cancer in nodes	32	$1/6.7 \times 10^{8}$
Breast cancer	Van nuys	20–24	22	5.58E+08	Cancer in nodes	39	$1/1.1 \times 10^{9}$
Breast cancer	Van nuys	25–29	27	1.03E+09	Cancer in nodes	39	$1/2.1 \times 10^{9}$
Breast cancer	Van nuys	30–34	32	1.72E+09	Cancer in nodes	50	$1/2.5 \times 10^{9}$
Breast cancer	Van nuys	35–39	37	2.65E+09	Cancer in nodes	54	$1/3.4 \times 10^{9}$
Breast cancer	Van nuys	40–44	42	3.88E+09	Cancer in nodes	63	$1/3.9 \times 10^{9}$
Breast cancer	Van nuys	45–50	47	5.44E+09	Cancer in nodes	54	$1/7.1 \times 10^{9}$
Breast cancer	Van nuys	50–54	52	7.36E+09	Cancer in nodes	59	$1/8.3 \times 10^{9}$
Melanoma	Balch *et al*	—	0.10	5.3E+01	Cancer death	11	1/480
Melanoma	Balch *et al*	—	0.23	6.8E+02	Cancer death	9	1/7100
Melanoma	Balch *et al*	—	0.37	2.6E+03	Cancer death	10	$1/2.4 \times 10^{4}$
Melanoma	Balch *et al*	—	0.44	4.3E+03	Cancer death	14	$1/2.9 \times 10^{4}$
Melanoma	Balch *et al*	—	0.50	6.7E+03	Cancer death	18	$1/3.4 \times 10^{4}$
Melanoma	Balch *et al*	—	0.60	1.2E+04	Cancer death	22	$1/4.5 \times 10^{4}$
Melanoma	Balch *et al*	—	0.70	1.8E+04	Cancer death	18	$1/9.1 \times 10^{4}$
Melanoma	Balch *et al*	—	0.84	3.1E+04	Cancer death	20	$1/1.4 \times 10^{5}$
Melanoma	Balch *et al*	—	0.91	3.9E+04	Cancer death	23	$1/1.5 \times 10^{5}$
Melanoma	Balch *et al*	—	0.97	4.8E+04	Cancer death	21	$1/2.0 \times 10^{5}$
Melanoma	Balch *et al*	—	1.11	7.1E+04	Cancer death	20	$1/3.1 \times 10^{5}$
Melanoma	Balch *et al*	—	1.17	8.5E+04	Cancer death	27	$1/2.7 \times 10^{5}$
Melanoma	Balch *et al*	—	1.31	1.2E+05	Cancer death	30	$1/3.3 \times 10^{5}$
Melanoma	Balch *et al*	—	1.38	1.4E+05	Cancer death	25	$1/4.8 \times 10^{5}$
Melanoma	Balch *et al*	—	1.51	1.8E+05	Cancer death	30	$1/5.0 \times 10^{5}$
Melanoma	Balch *et al*	—	1.58	2.1E+05	Cancer death	25	$1/7.1 \times 10^{5}$
Melanoma	Balch *et al*	—	1.71	2.6E+05	Cancer death	30	$1/7.1 \times 10^{5}$
Melanoma	Balch *et al*	—	1.85	3.3E+05	Cancer death	32	$1/8.3 \times 10^{5}$
Melanoma	Balch *et al*	—	1.88	3.5E+05	Cancer death	38	$1/7.1 \times 10^{5}$
Melanoma	Balch *et al*	—	1.98	4.1E+05	Cancer death	44	$1/7.1 \times 10^{5}$
Melanoma	Balch *et al*	—	2.11	4.9E+05	Cancer death	40	$1/1.0 \times 10^{6}$
Melanoma	Balch *et al*	—	2.25	6.0E+05	Cancer death	45	$1/1.0 \times 10^{6}$
Melanoma	Balch *et al*	—	2.52	8.3E+05	Cancer death	36	$1/1.9 \times 10^{6}$
Melanoma	Balch *et al*	—	2.72	1.1E+06	Cancer death	45	$1/1.8 \times 10^{6}$
Melanoma	Balch *et al*	—	3.02	1.4E+06	Cancer death	47	$1/2.3 \times 10^{6}$
Melanoma	Balch *et al*	—	3.39	2.0E+06	Cancer death	50	$1/2.9 \times 10^{6}$
Melanoma	Balch *et al*	—	3.72	2.7E+06	Cancer death	54	$1/3.4 \times 10^{6}$
Melanoma	Balch *et al*	—	4.26	4.1E+06	Cancer death	54	$1/5.3 \times 10^{6}$
Melanoma	Balch *et al*	—	4.73	5.5E+06	Cancer death	55	$1/7.1 \times 10^{6}$
Melanoma	Balch *et al*	—	5.27	7.7E+06	Cancer death	59	$1/8.3 \times 10^{6}$
Melanoma	Balch *et al*	—	5.77	1.0E+07	Cancer death	57	$1/1.2 \times 10^{7}$
Melanoma	Balch *et al*	—	6.74	1.6E+07	Cancer death	63	$1/1.6 \times 10^{7}$
Melanoma	Balch *et al*	—	7.75	2.4E+07	Cancer death	65	$1/2.3 \times 10^{7}$
Melanoma	Balch *et al*	—	12.32	9.8E+07	Cancer death	76	$1/6.7 \times 10^{7}$

captured with a Nikon Eclipse E400 microscope equipped with an Insight digital camera (Diagnostic Instruments Inc., Sterling Heights, MI, USA), used to capture high-resolution noninterpolated image with a $\times 10$ objective. The image measurements were calibrated by comparison to an image of the 1 mm grating on a haemocytometer. The longest dimension of the metastasis was measured from prints made of the images. All slides were reviewed by a qualified breast pathologist (DS). In some nodal metastases, noncancerous cells were apparent within the mass of cancer, and the sizes of these cancerous and noncancerous areas were measured. Thus, the values labelled 'Diameter (corrected)' (Table 3) were corrected with respect to the cancerous component of each metastasis, and thus were representative of the size that a mass of cancer would have had, had it shown the same number of cells but no noncancerous component.

Mathematical methods

Following the line of thinking outlined previously (Michaelson *et al*, 2002), let us define p as the probability of a single successful event of metastatic spread prior to surgery per unit of tumour volume, s. When the value of s is chosen so as to be to be equivalent to the volume of a cell, then p, from a practical standpoint, is also the probability of spread per cell, N. Note that by defining p on a per volume or per-cell basis, we are not assuming that every cell in the tumour has the potential to spread. For example, if we find that in a specific context, $p = 1$-in-ten-billion, then we shall not mean that every cell in the tumour mass will have such a chance of spread. Rather, this simply means that for every ten billion cells in a tumour, there will be about one event of metastatic spread. Note also that we have defined p as the probability of an *event* of spread, which can be either the spread of a single cell or a cluster of cells. Additionally, since we are defining p in terms of *successful* events of spread, that is events of spread that go on to give rise to evident cancer in the local nodes, or to give rise to distant metastatic disease, we are not concerned with those events of spread that do not result in such manifestations of metastasis. Let us define L as the fraction of patients displaying the occurrence of such an event of spread. If we are interested in examining the probability of the lethal spread of cells to the periphery, resulting in metastatic disease, then L will be the fraction of patients dying of the cancers; while if we are interested in measuring the nonlethal spread of cancer cells to the lymph nodes, then L will be the fraction of patients with cancer found in the nodes upon pathological analysis. It follows that (1-L) will be the fraction of the fraction of patients *not* displaying the occurrence of spread Similarly, as p is the per-cell probability of an event of spread, the probability that there will *not* be an event of spread will be (1-p), and the overall probability that a tumour of N cells has not given rise to one or more such metastases will be $(1\text{-}p)^N$. It has long been appreciated that for small values of p, $(1\text{-}p)^N$ can very well be approximated by e^{-Np}, and thus:

$$1 - L = e^{-Np} \tag{1}$$

Rearranging provides a way to estimate the probability of spread (p) per cell (N):

$$p = -\ln(1 - L)/N \tag{2}$$

RESULTS

From clinical data, we are able to observe the consequences of several examples of the spread of cancer cells, and their probabilities: the lethal spread of breast cancer and melanoma cells from the primary site to the periphery ($p_{\text{BC-overall}}$ and $p_{\text{MEL-overall}}$); and the nonlethal spread of breast cancer cells from the primary site to the local nodes ($p_{\text{BC-to-nodes}}$), and the lethal spread of breast cancer cells from the lymph nodes to the periphery ($p_{\text{BC-from-nodes}}$) (Tables 1 and 2). To see the general approach for estimating the values of these probabilities, consider the simple example of a group of patients with tumours containing a billion cells ($N = 10^9$ cells, ~ 3 cm), of whom 10% have died of metastatic disease ($L = 0.1$). If we assume, for explanatory purposes, that each death was the result of the spread of a single cell from the primary site to the periphery (an assumption not made in the math outlined in equations (1) and (2) above) then it follows that the probability (p) of lethal spread is approximately 1 event of spread for every ten billion cells in the primary mass ($p \approx L/N = 0.1/10^9$). Similarly, if about 1% patients with a different type of tumour, but of the same size ($N = 10^9$ cells), have died of metastatic disease ($L = 0.01$), then it follows that the probability of lethal spread (p) is about 1 event of spread for every hundred billion cells ($p \approx L/N = 0.01/10^9$).

Equation (2) provides the technique for accurately quantifying the probabilities of these types of cancer spread. In three instances ($p_{\text{BC-overall}}$, $p_{\text{MEL-overall}}$ and $p_{\text{BC-to-nodes}}$, Tables 1 and 2), data are available on spread from tumours of various sizes, revealing that the probability values do not remain constant as tumours grow, but declines as tumours increase in size (Figures 1 and 2). For example, the probability of lethal spread of melanoma cells from the primary site in the skin to the periphery ($p_{\text{MEL-overall}}$) is ~ 1 event of spread for every 500 cells for melanomas of 0.1 mm, but

Table 2 The values of the probabilities of various events of cancer spread

Metastatic event	Probability of spread per cell $p = \text{-ln(1-}L\text{)/}N$	Source of information for the value of L	Value of p the probability of spread per cell for tumour masses of ~ 5 mm	Nature of relationship between the value of p and tumour size
Lethal spread of breast cancer from the primary site in the breast to periphery, pathway unknown	$p_{\text{BC-overall}}$	$L_{\text{BC-overall}}$ = the fraction of breast cancer deaths among all patients	8.27×10^{-9}	$p = aN^b$ $a_{\text{BC-overall}} \approx 0.000056$ $b_{\text{BC-overall}} \approx -0.56203$
Lethal spread of melanoma from the primary site in the skin to periphery, pathway unknown	$p_{\text{MEL-overall}}$	$L_{\text{MEL-overall}}$ = the fraction of melanoma deaths among all patients	1.23×10^{-7}	$p = aN^b$ $a_{\text{MEL-overall}} \approx 0.027$ $b_{\text{MEL-overall}} \approx -0.7836$
Nonlethal spread of breast cancer from the primary site in the breast to the lymph nodes	$p_{\text{BC-to-nodes}}$	$L_{\text{BC-to-nodes}}$ = the fraction of node positive patients among all patients	1.75×10^{-8}	$p = aN^b$ $a_{\text{BC-to-nodes}} \approx 0.000092$ $b_{\text{BC-to-nodes}} \approx -0.69251$
Lethal spread of breast cancer from the lymph nodes to the periphery	$p_{\text{BC-from-nodes}}$	$L_{\text{BC-from-nodes}} = 6.08\%$, the lethal contribution per positive lymph node	7.96×10^{-9}[a]	Undefined

[a]The size of nodal metastases was found to have a mean value of 5.3 mm (Table 3). The value shown here for $p_{\text{BC-overall}}$ is for a mass of 5 mm; for 5.3 mm, $p_{\text{BC-overall}} = 7.49 \times 10^{-09}$.

100,000-fold lower ($\sim$1 event of spread for every 10^8 cells) for 12 mm tumours (Figure 1). As we have reported previously (Michaelson *et al*, 2002), a similar decline in the probability of spread per cell is also seen as tumours become larger for the overall probability of lethal spread of breast cancer cells from the primary site in the breast to the periphery ($p_{\text{BC-overall}}$, Figure 1). As can also be seen in Figure 1, a similar decline in the probability of spread per cell occurs for the nonlethal spread of breast cancer cells from the primary site in the breast to the lymph nodes ($p_{\text{BC-to-nodes}}$, Figure 2) (Tables 1 and 2). Furthermore, in each of the three contexts this decline occurs in a highly predictable fashion with N, such that it is well fit (Figures 1 and 2) to a power function of the form:

$$p = aN^b \tag{3}$$

Values for a and b for each of these three types of metastatic spread are shown in Table 2. b has a negative value of approximately -0.5 to -0.8, reflecting the fact that the value of p declines as tumours increase in size. a can be thought of as the probability of spread for the very first cell in the tumour, because $p = a$ when $N = 1$. Note that the parameter a is approximately 500-fold higher for melanoma than for breast cancer, reflecting the long-appreciated greater propensity of melanoma to give rise to metastases.

There are a number of possible explanations for why the probability of spread per cell, p, declines as tumours become larger. It could be that only a subpopulation of tumour cells are capable of metastasising, and that the relative abundance of these cells decline as tumours grow. Another possibility is that the decline in the per-cell probability of spread is the result of the simple geometrical constraints posed to the escape of cells from the primary mass (Padera *et al*, 2002), which become more formidable as tumours increase in size. As shown in the Supplementary material, such possibilities are mathematically possible, and are testable in experimental systems (Figure 3).

The spread of cancer cells can occur in single steps, such as the spread of a cell directly from the primary site to the periphery, or in multiple steps, such as the initial spread of a cell from the primary site to a local lymph nodes followed by the subsequent spread of one of the progeny of that cell away from the node to the periphery. By measuring the breast cancer death rate among subpopulations of patients sorted by both the size of the primary mass and the number of such positive lymph nodes, we have recently found that the presence of each positive node is associated with an extra 6.08% chance of death (Michaelson *et al*, 2003). It is possible to use this information with equation (2) to measure the probability of the spread of breast cancer cells from the nodes to the periphery ($p_{\text{BC-from-nodes}}$). To carry out this calculation, we set

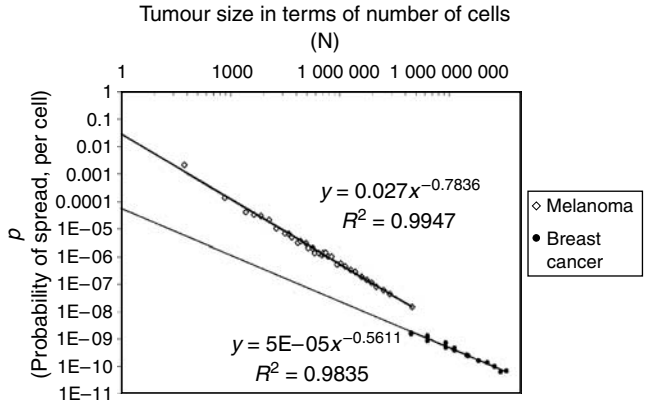

Figure 1 Calculations of the probability of lethal spread of breast cancer and melanoma cells, as a function of tumour size, and the close fit of the data to equation (2). ($R^2 = 0.98$ for breast cancer, ($R^2 = 0.9$ for melanoma). Shown here are the overall values for the probability of lethal spread of cancer cells from the primary site to the periphery for breast cancer ($p_{\text{BC-overall}}$) and melanoma ($p_{\text{MEL-overall}}$) using tumour size/survival data for all patients (Table 1). Note the close fit to the power function, equation (3).

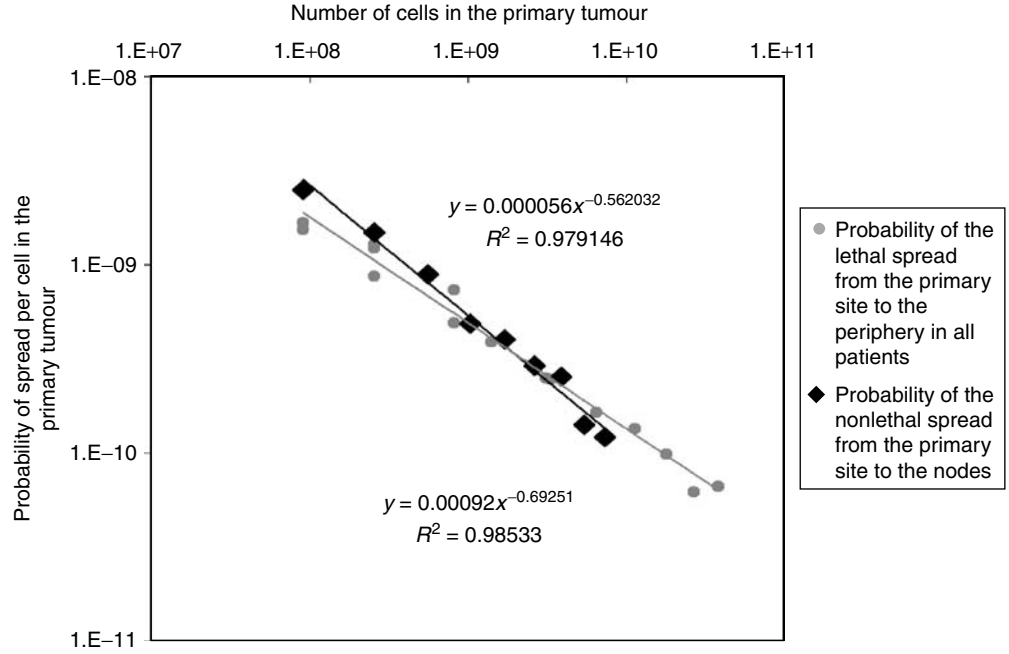

Figure 2 Calculations of the probability of lethal spread of breast cancer from the primary site to the periphery ($p_{\text{BC-overall}}$) by equation (2) and using tumour size/survival data for all patients (Table 1), and the probability of nonlethal spread of breast cancer from the primary site to the lymph nodes ($p_{\text{BC-to-nodes}}$) by equation (2) and using tumour size/nodal status data (Table 1). Note that in both cases the relationship between the probability of spread and tumour size is well fit by a power function, equation (3).

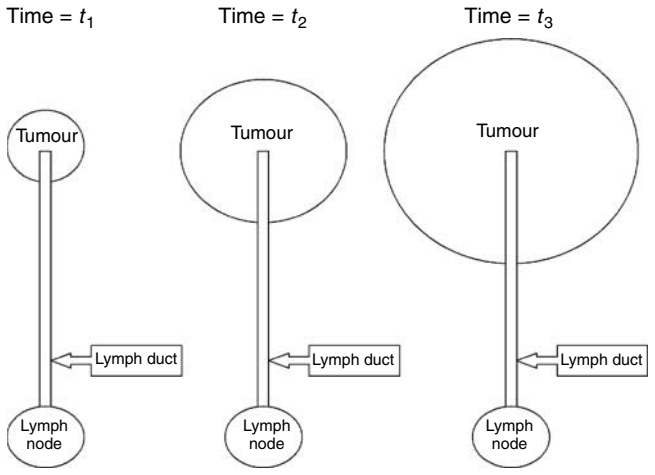

Figure 3 Schematic of Geometrical Model #1. Shown is a highly idealised image of a tumour mass and a lymph duct leading to a local lymph node.

the value of $L = 0.0608$, and to determine the value of N, we collected data on the sizes of metastases in lymph nodes, revealing a mean size of 5.3 mm, which is equivalent to $N = 7.84 \times 10^6$ cells (Table 3). It follows with equation (2) that the probability of the spread of cancer cells from lymph nodes $p_{\text{BC-from-nodes}} = 7.96 \times 10^{-09} \approx 1$ event of spread for every 10^8 cells. This value is remarkably close to the value for the probability of the lethal spread of breast cancer cells from the primary mass in the breast when the primary mass is also 5.3 mm ($p_{\text{BC-overall}} = 7.49 \times 10^{-09}$), as calculated by extrapolation of equation (3). This reveals that the occurrence of an event of spread of cancer cells from the primary site in the breast to the local lymph nodes does not appreciably change the tendency of the progeny of those cancer cells to make yet a second event of spread from the lymph nodes to the periphery.

DISCUSSION

It has often been wondered whether mutation at the time of spread is a requirement for metastasis (Cifone and Fidler, 1981; Fidler, 1983; Sobel, 1990; Welch *et al*, 2000; Yokota, 2000; Bernards and Weinberg, 2002; Couzin, 2003; Bernards, 2003), but the values of the probabilities of metastatic spread of breast cancer and melanoma cells revealed by equation (2) are difficult to reconcile with such genetic changes due to several reasons: *First*, the value of the probability of spread for the smallest melanomas (0.1 mm), at ~ 1 event of spread for every 500 cells, is many orders of magnitude greater than that expected for a genetic change. *Second*, the occurrence of one event of spread (the spread of breast cancer cells from the breast to the local lymph nodes) does not appear to increase the chance of a second event of spread (the spread of breast cancer cells from the local lymph nodes to the periphery). In other words, the occurrence of the initial event of spread does not lead to a cell-heritable change in the tendency of the progeny of that cell to spread. *Third*, the data shown here reveal that the probability of metastatic spread per cell declines as tumours increase in size. While this decline is consistent with a number of explanations that are mechanical, (using this term in the sense in which it is used in physics: 'pertaining to the relations of force and matter'), such as the effect of tumour geometry on the ease of the escape of cells from the primary mass (see Supplementary material), it is not what would be expected for genetic events. Indeed, the probability of genetic events over time should be expected either to remain constant (if only a single genetic event is required) or to increase with time (if the accumulation of multiple genetic events is required). Taken together, these findings would appear to be in agreement with the viewpoint put forward by Bernards and Weinberg 'that the tendency to metastasise is largely determined by the identities of mutant alleles that are acquired relatively early during multistep tumorigenesis', and that 'genes and genetic changes specifically and exclusively involved in orchestrating the process of metastasis do not exist' (Bernards and Weinberg, 2002).

Table 3 Sizes of the invasive breast cancer metastases seen in the lymph nodes

Node number	Patient number	Diameter (mm)	Fraction of the metastatic area containing cancer (%)	Diameter (corrected)
49	16	0.3	100	0.3
32	9	1	100	1
20	3	1	100	1
31	8	1.1	100	1.1
43	14	1.4	100	1.4
41	14	1.9	65	1.5
17	2	2	100	2
21	4	2	95	2
39	12	2.1	100	2.1
13	2	2.5	93	2.4
29	7	2.7	95	2.6
15	2	3	95	2.9
30	7	3	100	3
36	12	3	95	3
2	1	3.2	10	3.2
22	5	3.4	95	3.3
5	1	3.4	93	3.3
34	10	3.7	95	3.7
37	12	3.9	100	3.9
4	1	3.9	86	3.6
47	15	4.3	95	4.2
35	11	4.5	95	4.4
40	13	5.4	95	5.2
6	1	5.4	86	5
16	2	5.4	100	5.4
23	5	5.5	76	4.8
33	10	5.5	78	4.9
24	5	6	80	5.4
48	15	6	90	5.7
42	14	6.4	10	6.4
45	14	7.1	60	5.5
10	2	7.1	100	7.1
11	2	7.1	100	7.1
12	2	7.7	95	7.5
18	2	8	43	5.2
19	2	9	72	7.6
46	14	9	90	8.5
28	6	9	80	8
1	1	9.8	84	9
44	14	10	49	7
9	2	10	65	8.1
3	1	10	86	9.3
14	2	10	95	9.7
7	1	10.7	55	7.9
26	6	11	35	6.5
8	2	11	95	10.7
25	6	12.5	45	8.4
27	6	14.5	85	13.4
38	12	15	100	15
Average		5.95	86	5.31

Supplementary Information accompanies the paper on British Journal of Cancer website (http://www.nature.com/bjc)

REFERENCES

Balch CM, Soong SJ, Gershenwald JE, Thompson JF, Reintgen DS, Cascinelli N, Urist M, McMasters KM, Ross MI, Kirkwood JM, Atkins MB, Thompson JA, Coit DG, Byrd D, Desmond R, Zhang Y, Liu PY, Lyman GH, Morabito A (2001) Prognostic factors analysis of 17,600 melanoma patients: validation of the American Joint Committee on Cancer melanoma staging system. *J Clin Oncol* **19:** 3622 – 3634

Bernards R (2003) Clues for migration. *Nature* **425:** 247 – 248

Bernards R, Weinberg RA (2002) A progression puzzle. *Nature* **418:** 823

Boon ME, Trott PA, van Kaam H, Kurver PJ, Leach A, Baak JP (1982) Morphometry and cytodiagnosis of breast lesions. *Virchows Arch [Pathol Anat]* **396:** 9 – 18

Cifone MA, Fidler IJ (1981) Increasing metastatic potential is associated with increasing genetic instability of clones isolated from murine neoplasms. *Proc Natl Acad Sci USA* **78:** 6949 – 6952

Couzin J (2003) A clash over genes that foretell metastasis. *Science* **299:** 1005

Fidler IJ (1983) The Ernst W. Bertner Memorial Award lecture: the evolution of biological heterogeneity in metastatic neoplasms. *Symp Fundam Cancer Res* **36:** 5 – 26

Karrison TG, Ferguson DJ, Meier P (1999) Dormancy of mammary carcinoma after mastectomy. *J Natl Cancer Inst* **91:** 80 – 85

Koscielny S, Tubiana M, Le MG, Valleron AJ, Mouriesse H, Contesso G, Sarrazin D (1984) Breast cancer: relationship between the size of the primary tumor and the probability of metastatic dissemination. *Br J Cancer* **49:** 709 – 715

Michaelson J (1999) The role of molecular discreteness in normal and cancerous growth. *Anticancer Res* **19:** 4853 – 4867

Michaelson JS, Silverstein M, Sgroi D, Cheongsiatmoy JA, Taghian A, Powell S, Hughes K, Comegno A, Tanabe KK, Smith B (2003) The effect of tumor size and nodal status on the lethality of breast cancer. *Cancer* **98:** 2133 – 2143

Michaelson JS, Silverstein M, Wyatt J, Weber G, Moore R, Halpern E, Kopans DB, Hughes K (2002) The prediction of breast cancer survival from tumor size. *Cancer* **95:** 713 – 723

Padera TP, Kadambi A, di Tomaso E, Carreira CM, Brown EB, Boucher Y, Choi NC, Mathisen D, Wain J, Mark EJ, Munn LL, Jain RK (2002) Lymphatic metastasis in the absence of functional intratumor lymphatics. *Science* **296:** 1883 – 1886

Pesce C, Colacino R (1986) Morphometry of the breast fibroadenoma. *Pathol Res Pract* **181:** 718 – 720

Silverstein MJ (2000) The Van Nuys Breast Center: the first free-standing multidisciplinary breast center. *Surg Oncol Clin North Am* **9:** 159 – 175

Sobel ME (1990) Metastasis suppressor genes. *J Natl Cancer Inst* **82:** 267 – 276

Tabar L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF, Krusemo UB, Tot T, Smith RA (2000) The Swedish two-county trial twenty years later. Updated mortality results and new insights from long-term follow-up. *Radiol Clin North Am* **38:** 625 – 651

Tubiana M, Koscielny S (1990) The natural history of human breast cancer: implications for a screening stratergy. *Int J Radiation Oncol Biol Phys* **19:** 1117 – 1120

Tubiana M, Koscielny S (1991) Natural history of human breast cancer: recent data and clinical implications. *Breast Cancer Res Treatment* **18:** 125 – 140

Van der Linden HC, Baak JP, Smeulders AW, Lindeman J, Meyer CJ (1986) Morphometry of breast cancer. I. Comparison of the primary tumours and the axillary lymph node metastases. *Pathol Res Pract* **181:** 236 – 242

Van't Veer LJ, Weigelt B (2003) Road map to metastasis. *Nat Med* **8:** 999 – 1000

Welch DR, Steeg PS, Rinker-Schaeffer CW (2000) Molecular biology of breast cancer metastasis. Genetic regulation of human breast carcinoma metastasis. *Breast Cancer Res* **2:** 408 – 416

Yang J, Mani SA, Donaher JL, Ramaswamy S, Itzykson RA, Come C, Savagner P, Gitelman I, Richardson A, Weinberg RA (2004) Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117:** 927 – 939

Yokota J (2000) Tumor progression and metastasis. *Carcinogenesis* **21:** 487 – 503

SUPPLEMENTARY MATERIAL

Some possible mechanisms for why the per-cell probably of spread declines with tumor size

One of the striking findings of the calculations described above is that when data were analyzed for the tumors of various sizes, the value of the probability of spread per-cell appeared to decline in a very characteristic fashion, such that this decline is closely fit by a power function of the form of eq. (3). There are several possible explanations for why this decline might occur.

Explanations arising from cellular heterogeneity

As we noted at the outset, while we have defined p in terms of the probability of spread per-cell, we do not assume that every cell in a tumor has the capacity to spread. Indeed, such heterogeneity could be one explanation for why the probability per-cell in the tumor as a whole declines as tumors increase in size. Consider the case in which there are two kinds of cells in a tumor, those that have the capacity to spread to the periphery, with a constant probability p_s., and those that do have the capacity to spread to the periphery, with the probability p_u=0. Let us first imagine that cells start out having this capacity to spread, but can lose it in a cell heritable fashion. Let us further imagine that in each cell division there is a chance (p_l) that the dividing cell and its progeny irreversibly loses the capacity to spread to the periphery. There is, then, a subpopulation of cells that have the ability to form distant lethal metastases described by:

$$N_s = N(1-p_l)^n \tag{5}$$

where N_s is the number of cells that retain the ability to spread in each cell division, n is the number of cell divisions, and N is the total number of tumor cells. Assuming simple exponential growth of the tumor from one cell, we have:

$$n = \ln(N)/\ln(2) \tag{6}$$

Again, for reasons well known to us, if L is the fraction of patients dying from breast cancer:

$$\ln(1-L) = -N_s p_s \tag{7}$$

and, because this subpopulation of cells comprises the only set of cells in the tumor that have the capacity to form distant lethal metastases, we also have:

$$\ln(1-L) = -Np \tag{8}$$

where p is the probability per-cell of distant lethal metastatic spread for the tumor as a whole. Combining equations (5)-(8) and solving for p yields:

$$p = p_s(1-p_l)^{\ln N/\ln 2} \tag{9}$$

A simple rearrangement of (9) gives:

$$p = p_s N^{\ln(1-p_l)/\ln 2} \tag{10}$$

or:

$$p = aN^b \tag{3}$$

for $a = p_s$ and $b = ln(1-p_l)/ln2$. For example, let us take $p_l = 1/3$. p_l would take on such a value if in each cell division there was a 1-in-3 chance that the dividing cell and its progeny irreversibly lost their capacity to spread to the periphery. It follows that the value of b will be –0.58, which is nearly exactly that given by empirical data. Both the form of the expression derived from this model, then, and the N-dependency of the expression, which is given by the parameter b, closely approximate the expression for the probability per-cell of distant lethal metastatic spread as derived from empirical data; the empirically derived relationship is given by a similar power function, with a value of b=*-0.5611*.

Explanations arising from tumor geometry

Another category of explanations arises from the simple geometrical barriers to the escape of cells from tumor masses, which become more formidable as size increases (Figure 3). Recall that as a sphere increases in size, neither its surface area, nor its diameter, will increase as much as its volume. If these are the places from which cancer cells leave the mass, then their escape will become progressively more difficult as tumor size increases, and the probability of escape per-cell in the mass as a whole will decline. Let us examine the consequences of such geometrical constraints.

Geometrical model #1

Consider N to be the number of cells in the tumor of radius r where s is the density of cells per cc. Thus:

$$N = s(4/3)\pi r^3 \tag{11}$$

Let N_s be the number of cells in a subpopulation of cells that are capable of spreading. Let us further consider the case shown in Figure C above, where the only cells with this capacity are those cells that are next to the lymph duct. If the cells each have a diameter of d, and c is the number of cells that form a ring around the lymph duct, which is of constant diameter, it follows that the number of cells next to the lymph duct is:

$$N_s = cr/d \tag{12}$$

Combination with (11) yields:

$$N_s = N^{1/3} q \tag{13}$$

where:

$$q = c/[d(4/3s\pi)^{1/3}] \tag{14}$$

Let p_s, a constant, be the probability of an event of spread per-cell for N_s cells in the subpopulation of cells. Let p, a variable, be the probability of an event of spread, per-cell in the tumor as a whole, comprised of N cells. Let L be the fraction of patients dying of breast cancer. For reasons well known to us:

$$1 - L = e^{-p_s N_s} \tag{15}$$

and, because we are assuming that only the cells in the subpopulation immediately adjacent to the lymph duct can give rise to lethal metastatic spread:

$$1 - L = pN \tag{16}$$

thus

$$-\ln(1-L) = p_s N_s \tag{17}$$

and

$$-\ln(1-L) = pN \tag{18}$$

setting (17) and (18) equal to one another and rearranging yields:

$$p = p_s N_s / N \tag{19}$$

combining (13) and (19) yields:

$$p = qN^{0.333}/N \tag{20}$$

or

$$p = aN^b \tag{3}$$

where $a=q^* p_s$ and $b=-0.666$. Thus, simple geometrical constraints to the spread of cancer cells can give rise to a probability of spread per-cell that declines as tumors increase in size in such a way that this decline conforms to a power function of the form of eq. (3), the N-dependence of which is close to the value found empirically.

Geometrical model #2

We have also been able to show that a simple geometrical model for lethal metastatic spread through the lymphatic system based on a random walk closely approximates the value of *b*, the empirically derived parameter describing the N-dependency of the probability of spread per-cell. In this model, we assume that each cell is equally likely to take a step in any one of the six possible spatial directions, and that if a series of steps results in the cell arriving at a lymph duct, the cell then has a constant intrinsic probability of lethal spread through the lymphatic system. Thus, the probability per-cell of lethal metastatic spread is defined by two factors: the simple geometrical problem of arriving at the lymph duct, and a constant probability that it will spread once it has reached the lymph duct. Using an unbounded random walk model for tumor cells that yields a three-dimensional Gaussian distribution of the probable locations of each cell in the tumor, and assuming that the primary route of spread is through a lymphatic duct modeled as a single point on the surface of the tumor, we were able to arrive at a probability per-cell of lethal metastatic spread through the lymphatic system given, as a function of time and position, by:

$$P_t = (1/4\pi t)^{3/2} e^{r^2/4} r \sin(\theta) dr d\theta d\phi \tag{21}$$

where r, θ, and ϕ are canonical spherical coordinates. Integrating over all cells in the tumor and until the time of spread, and dividing by the number of cells in the tumor yields an average probability per-cell of lethal spread for the tumor as a whole:

$$\langle P \rangle = 1/N \iiint (1/4\pi t)^{3/2} e^{[D(\theta,\phi,r)]^2/4t} r \sin(\theta) dr d\theta d\phi \tag{22}$$

where $\phi \in [0, 2\pi]$, $\theta \in [0, \pi]$, $r \in [0, (N/C)^{1/3}]$ and D (θ, ϕ, r) is the distance from any point (θ, ϕ, r) to the lymph duct, as given in spherical coordinates and assuming a constant cell density of 10^8 cells/cc:

$$D(\theta,\phi,r) = [(r\sin\theta\cos\phi)^2 + (r\sin\theta\sin\phi)^2 + (r\cos\theta - 0.0062N^{1/3})^2]^{1/2} \tag{23}$$

We were unable to find a closed-form solution to integral (22), but numerical integration and graphical analysis reveals that the points generated are very well fit ($R^2 > 0.99$) by an equation of the form (data not shown):

$$p = aN^b \tag{3}$$

where the value of b is -0.5562, which is almost exactly that given by empirical data: b = -0.5611.

Testing these possibilities

While the modeling described above reveals that the explanations of cell heterogeneity and geometry are possible explanations, they tell us nothing about whether they are the actual explanations. However, these possibilities should be testable in experimental systems with transplantable tumors that give rise to metastases if the relationship between tumor size and the fraction of animals with signs of metastasis are such that the probability of spread per-cell conforms to eq. (3), with the parameter *b* lying in the range ~ -0.5 to -0.8. The hypothesis of cell heterogeneity would be born out by disaggregating tumors of various sizes, injecting the cells intravenously, and determining if the number of metastatic colonies created agreed with that found for the intact tumors of various sizes. The hypothesis of geometry could be tested in a variety of ways in which the geometry of the tumor mass was manipulated. For example, eq. (3) leads to the prediction that the number of metastatic colonies found in two animals containing the same mass of the same tumor will not be same if the tumor is present in one animal as a single transplanted mass (for example in a single limb) while present in the second animal as four smaller masses (for example, by transplanting into each of four limbs), each mass of which is one-fourth of the mass in the animal with a single transplanted tumor. A variety of other manipulations, such as growing the transplanted tumor between glass plates, would alter the geometry of the tumor, and thus might be expected to affect the relationship between tumor size and the probability of spread per-cell.