The Laboratory of Quantitative Medicine
Technical Report #1
August 31, 2009

The Binary-Biological Model of Cancer Metastasis

A mathematical framework for predicting the risk of cancer death and other manifestations of the spread of cancer cells, for isolating the impact of primary tumor size, nodal status, and other prognostic factors on the risk of death, and for measuring the probabilities of the spread of cancer cells

James S Michaelson PhD[1,2,3]

Departments of Pathology[1]and Surgery[2], Massachusetts General Hospital and the Department of Pathology, Harvard Medical School, Boston, Massachusetts

Correspondence to James S. Michaelson Ph.D., Division of Surgical Oncology, Cox Building Room 626, Massachusetts General Hospital, 100 Blossom Street, Boston, Massachusetts, 02114
TEL 617 501 0590 FAX 617 724 3895
Email: michaelj@helix.mgh.harvard.edu

Here, I shall outline a general theory of the spread of cancer. This mathematics can be used to understand both the lethal and non-lethal manifestations of such spread. In particular, it can be used to answer a range of practical problems concerned with the estimation of the risk of cancer death for individual patients. It can also be used in the estimation of population-wide cancer death rates, especially in the context of screening. This mathematics has also been used to drive a series of web-based calculators (http://www.CancerMath.net), which provide medical professionals with estimates of the risk of cancer death for individual patients, together with estimates of the impact of various treatment choices on that risk of death.

Parts of this theory, in a somewhat piecemeal fashion, have been used in a number of published studies[1,2,3,4,5,6,7,8,9,10,11,12]. Here I provide the theory in an integrated fashion.

What we shall attempt to do

One of the central tasks of this theory is to provide a mathematical framework that can capture the relationships between primary tumors size, nodal status, prognostic factors, and cancer lethality. This involves the development of mathematical techniques for teasing out the independent impact of tumors size, nodal status, prognostic factors, and cancer lethality (the *SizeAssessment* and *PrognosticMeasurement* methods), and for taking information on a patient's tumor size, nodal status, and other prognostic factors and estimating a patient's risk of death (the *SizeOnly*, *Size+Nodes*, and *Size+Nodes+PrognosticFactors* (*SNP*) methods). This approach also results in the development of an expression, the *NodalSizeOnly* equation, which can relate tumor size to the chance of cancer in the nodes, and another expression, the *PrimarySizeOnly* equation, which can relate tumor size to the chance of death in node negative patients. This approach has proven to be empirically sound, as the accuracy of these methods has been confirmed with actual outcome data from large populations of patients with breast carcinoma, melanoma, and renal cell carcinoma.

Methods of the Binary-Biological Model of Cancer Metastasis

Method	Purpose
ProbabilityEstimation Equation (#7)	Estimates the probability of the spread of cancer cells, per cell, from data on the size of the mass from which the spread occurs and the lethal or non-lethal consequences of the spread.
SizeOnly Equation (#1)	Relates tumor size to the chance of cancer death*
TimeOnly Equation (#14)	Relates time to the chance of cancer death
PrimarySizeOnly Equation (#1c)	Relates tumor size to the chance of cancer death for node negative patients
NodalSizeOnly Equation (#1n)	Relates tumor size to the chance of cancer in the lymph nodes
Size+Nodes method	Integrates information on tumor and number of positive nodes into an estimate of the chance of cancer death*
Nodal Lethality Equation (#17)	Calculates the lethal contribution, per positive node, of cancer in the lymph nodes
SizeAssessment method	Determines whether a prognostic factor makes an independent contribution to the risk of cancer death, or is merely correlated with tumor size
PrognosticMeasurement method	Provides a quantitative measure of each prognostic factor's contribution to cancer lethality, through the introduction of a parameter, *g*, inserted into the *SizeOnly* equation
SNAP (*Size+Nodes+PrognosticFactor*) method	Integrates information on primary tumor size, nodal status, and other prognostic factors into an estimate of the chance of cancer death*

* When only tumor size and nodal status are known, the *SNAP* method reduces to the *Size+Nodes* method for estimating the risk of cancer death from information on tumor size and number of positive nodes, while when only

size is known, the *Size+Nodes* and *SNAP* methods reduce to the *SizeOnly* method for estimating the risk of cancer death from information on tumor size.

Motivation

The development of these equations was made possible by capturing, in mathematical terms, a quality of cancer that is deceptively obvious, but exceedingly useful; cancer is a disease of cells, and cells are irreducibly discrete entities[13]. Cells come in integers only. There can be 1, 3, or a million and three cancer cells in a lymph node, but never 1.3 cells. Thus, when cells move from one location to another, giving rise to a new cancer phenotypes, such events of spread must inevitably be discrete, either/or, events. Either a cancer cell has spread from the primary site to the periphery, causing death, or it hasn't. Either a cancer cell has spread from the primary site to a local node, causing cancer in that node, or it hasn't. Either a cancer cell has spread from a node to the periphery, causing death, or it hasn't. This either/or quality of spread of cancer cells from one location to another allowed us to assign probability values for such events of spread[14,15], and from such a basis, we have been able to derive the equations used for relating the various features of tumor size, nodal status, prognostic factors, and cancer lethality. We have called this approach the *binary-biological model of cancer metastasis*. We have borrowed both the name, and the inspiration, for our model from statistical mechanics in physics, where the large-scale physical properties of matter are understood as the macroscopic consequences of the underlying microscopic events of molecules.

The discrete approach of the *binary-biological model of cancer metastasis* is but a specific example of a general project, which we call *binary-biology*, that we have undertaken for understanding the macroscopic features of multicellular systems as the aggregate consequences of the many *either/or* events going on among the discrete components of which we are comprised: molecules, atoms, electrons, photons, genes. As we have reported elsewhere, by considering multicellular organization in this way, it has been possible to see how the normal cellular populations of the body can grow to predictable sizes, at predictable times, and to predictable shapes[13]. Such modeling has also provided a way to see how normal cellular populations become cancerous cellular populations[13].

Definitions

For consistency with previous publications, Equations #1-#4 are numbered in agreement with references 28, and thus will appear out of order when presented below.

Here we shall examine macroscopic features of cancer - tumor size, nodal status, and cancer survival - in terms of the underlying microscopic spread of cancer cells, occurring with a definable probability of spread per cell, p_x. We call this approach the binary-biological model of cancer metastasis.

Let us define L_x as the fraction of patients who display a manifestation of the spread of cancer cells: $L_{to\text{-}nodes}$ will be the fraction of patients with cancer in the lymph nodes, $L_{primary}$ will be the fraction of node negative patients dying of cancer metastasis and L will be the fraction of all patients, both node positive and node negative, who die of cancer. For reasons outlined above, $L_{primary}$ and L are defined in terms of the 15-year cancer-specific Kaplan Meier death rate.

Let us define p_x as the probability of a single successful event of spread for each of the N cells from which the event of spread originated. Then, $p_{to\text{-}nodes}$ is the probability of a single successful event of spread to the nodes for each of the N cells in a primary tumor (thus related to $L_{to\text{-}nodes}$), while $p_{primary}$ is the probability of a single successful lethal event of spread to the periphery for each of the N cells in a primary tumor in node negative patients (related to $L_{primary}$), and p is the probability of a single successful lethal event of spread to the periphery for each of the N cells in a primary tumor in all patients in the population, regardless of nodal status (related to L). Let us also define $p_{from\text{-}nodes}$ as the probability of a lethal event of spread from a lymph node to the periphery, and $p_{node\text{-}to\text{-}node}$ as the probability of a non-lethal event of spread from lymph node to lymph node.

Note that we have framed the definition of p_x so that it does not consider events of spread that do not lead to a macroscopic manifestation, nor does this definition assume p_x to be constant[28,30], nor does it require that every cell in the mass of cancer have the potential for spread, although this may be the case. For further discussion, see the next paper in this series and reference #30.

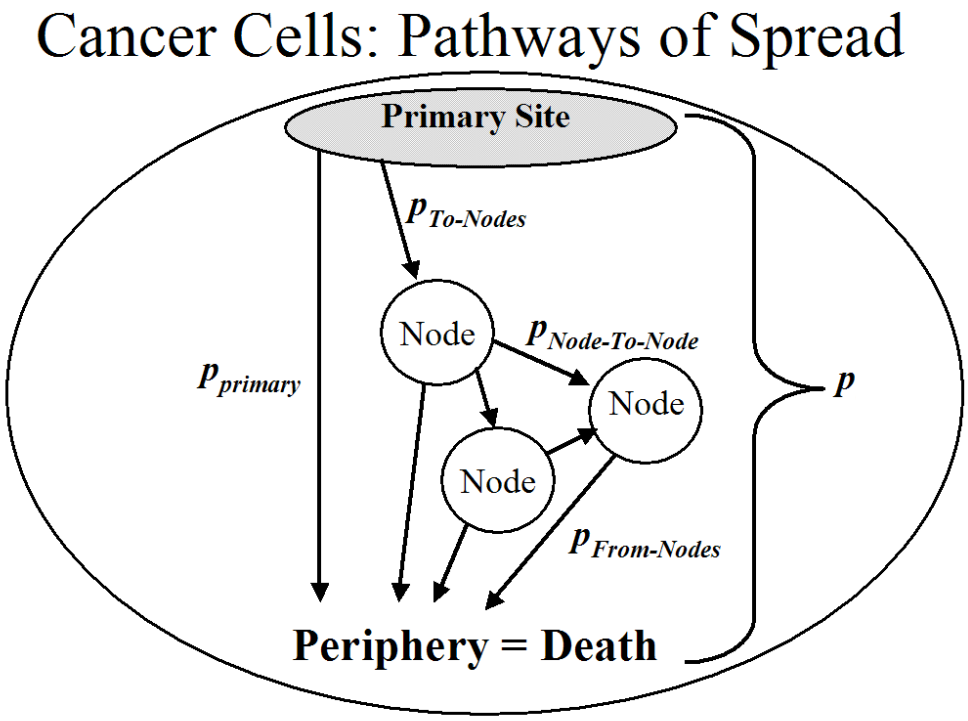

FIGURE 1
Cancer cells: Pathways of Spread

TABLE I

Pathways of the Spread of Cancer Cells and their Probabilities

Pathway of Spread	Seen In	Fraction of patients with a manifestation of spread	Probability of spread, per cell
Lethal spread from the primary site to the periphery (directly only)	Fraction of patients dying (=15-year Kaplan-Meier cancer death rate) among node-negative patients	$L_{primary}$	$p_{primary}$
Non-lethal spread from the primary site to the local lymph nodes	The fraction of patients with positive nodes	$L_{to\text{-}nodes}$	$p_{to\text{-}nodes}$
Lethal spread from the lymph nodes to the periphery	Lethal contribution per positive lymph node	-	$p_{from\text{-}nodes}$
Non-lethal spread from lymph node to lymph node	The numbers of positive nodes (M)	-	$p_{node\text{-}to\text{-}node}$
Lethal spread from the primary site to the periphery (The aggregate consequence of the pathways of spread characterized by	Fraction of patients dying (=15-year Kaplan-Meier cancer death rate) among all patients	L	p

$p_{primary}$, $p_{to\text{-}nodes}$, $p_{from\text{-}nodes}$ and $p_{node\text{-}to\text{-}node}$)			

Units

Tumors grow over time. Let us define C as *the number of cells present for the number of days that a tumor has accumulated from any point in time*, that is, the number of "cell-days" a tumor has accumulated. Since there is abundant evidence that over the range of sizes for which most tumors are seen (approximately 1 mm to 2 cm) invasive breast cancer growth is exponential[16,17], the relationship between the number of cells in a tumor (N), and time (t), can be expressed as:

$$N = N_0 e^{rt} \qquad (*)$$

where N_o is the number of cells at time t=0, and e is the exponential constant, t is time, and r is a constant, such that:

$$r = \ln(2) / t_D \qquad (*)$$

where t_D is the tumor doubling time. It follows that C will equal the integral of Eq. #1, evaluated from t=0 (where $N=N_0$) to t=(ln N)/r:

$$C = \int e^{rt} = (N - N_0) / r \approx N / r \qquad \text{if} \qquad No << N \qquad (*)$$

It has often been noted that while breast cancer growth is approximately exponential over the range where it is evident clinically (approximately 1 mm to 2 cm), it may display faster growth when the tumors are smaller, a growth pattern often called Gompertzian, in reflection of one type of equation which captures such growth[17]. However, as can be seen from Equation #14, as long as exponential growth occurs between values N_0 and N, $N_0 << N$ and $F \approx 1$ at $N=N_0$, all three conditions of which are fulfilled for breast cancer for $N_0 \approx 5x10^4$, which roughly corresponds to a tumor of 1 mm, then such Gompertzian growth will not materially affect the values of C.

We may wish to estimate the probability of spread in terms of C. In analogy to the treatment used above, let us define p' as the *probability of a single event of lethal spread per C cell-days*, such that:

$$F = (1 - p')^C \qquad (*)$$

together with Equation #14, and with a treatment analogous to that used for Equations #4 to #6, it follows that:

$$p' = rp \qquad (*)$$

Note that it is simply a matter of preference whether we wish to estimate the probability of spread in terms of p per cell (N), or in terms of p' per cell-day (C). However, these two values are probably relevant to two different types of biological mechanism of cancer spread. Biological mechanisms that occur once per cell, such as mutational events that occur at cell division[18], should be expected to occur with a probability such as p, which is calculated in terms of N, the number of cells in the tumor at the time of surgery. On the other hand, biological mechanisms of a more mechanical nature, such as cell detachment, re-engraftment, evasion of immune attack, or successful angiogenesis[19,20,21,22] might occur at any time, and should be expected to occur with a probability such as p', which is calculated in terms of the number of "cell-days", C, that a tumor of N cells has accumulated. Note, however, that whether the mechanism of spread occurs by events that can occur once per cell (p) or continuously (i.e. per cell per day) (p'), both lead to Equation #1, (with values of a and b as shown in Equations # 10 and #11), which as we shall show below, provides a remarkably good capture of the relationship between tumor size and survival. This reveals that the actual survival data are equally compatible with biological mechanisms that would occur once per cell (p) or continuously in time (p').

How we measure cancer lethality

Karrison and colleagues[23] had found that little lethality occurs 15-years after diagnosis, and we have found a similar hazard function for melanoma (see *http://www.cancer-math.net/*). Thus, we have relied upon the 15-year cancer-specific Kaplan Meier death rate as our measurements of the cancer death rate (*L*).

The *ProbabilityEstimation* Equation

Since p_x is the probability of a single successful event of lethal spread, the probability per cell that there will not be an event of spread is $(1-p_x)$ and the overall probability that a tumor of *N* cells will not give rise to one or more lethal metastases will be $(1-p_x)^N$. Thus, for a population of patients all with tumors of identical size, the fraction of patients who have not had an event of spread, $1-L_x$, will be:

$$1 - L_x = (1 - p_x)^N \tag{5}$$

For small values of p_x:

$$1 - L_x = e^{-Np_x} \tag{6}$$

or:

$$p_x = -\ln(1 - L_x) / N \tag{7}$$

Equation #7, which we call the *ProbabilityEstimation* Equation, provides a way to estimate the value of p_x from information on the fraction of patients with a manifestation of spread, L_x, for a group of patients with tumors of size *N*.

The *SizeOnly* Equation

The number of cells, *N*, can be estimated by assuming a density of *s* (in cells/cm^3) and spherical geometry, i.e.:

$$N = s\pi\frac{4}{3}\left(\frac{D}{2}\right)^3 \tag{8}$$

with a plausible value for *s* set at 10^8 cells/cc[24-26].

Using Equation #7, we have found[28,30] for both lethal and non-lethal spread of breast carcinoma and melanoma that the value of *p* declines gradually as tumors increase in size, *N*, and indeed can be closely fit by a power function:

$$p_x = aN^b \tag{9}$$

where $b \approx -2/3$.

Let us define:

$$b = (Z/3) - 1 \tag{10}$$

and

$$a = Q/[(\pi/6)s]^{Z/3} \tag{11}$$

Combining Equation #6-11 leads to an expression for relating the fraction of patients with any manifestation of the spread of cancer cells, L_x (such as cancer lethality, *L*, or cancer in the nodes, $L_{to\text{-}nodes}$), and tumor size (diameter or thickness), *D*:

$$L_X = 1 - e^{-QD^Z} \tag{12}$$

This, when applied to the relationship between the fraction of all patients dying of cancer, *L*, and tumor diameter or thickness, *D*, leads to an expression that we have called the *SizeOnly* Equation:

$$L = 1 - e^{-QD^Z} \tag{1}$$

The *sSizeOnly* Equation

As noted above,, we have found from information on the fraction of patients dying of cancer, *L*, for a group of patients with tumors of size *N* that the value of *p* does not remain constant as tumors increase in size, *N*, but declines gradually such that the relationship between *p* and *N* can be closely fit by a power function, Equation #9 with *b*=-2/3, or:

$$p = aN^{-2/3} \tag{9s}$$

The likely explanation for why the value of *p* declines as tumors increase in size is geometry. After all, the larger the tumor, the more cells must be "pushed aside" before any single cell can get out. For a mathematical examination of this point, see reference 12.

Combining Equation #6-10 leads to an expression for relating the fraction of patients dying of cancer, and tumor size, *D*, that we have called the *SimplifredSizeOnly* Equation or *sSizeOnly* Equation:

$$L = 1 - e^{-QD} \tag{1s}$$

We have found for melanoma, renal cell carcinoma, and breast carcinoma, the relationship between tumor size and the risk of cancer death is well capture but only captured by the *SizeOnly* Equation, but also by the sSizeOnly Equation. This makes it possible to carry out a comparison of the lethalities of different cancers and different subtypes of the same cancer by a comparison of the values of their *Q* parameters (the *SizeAssessment* and *PrognosticMeasurement* method), as well as providing a unified method for integrating tumor size, nodal status, and other prognostic factors into a estimate of the risk of cancer death, the *Size+Nodes+PrognosticFactors* (*SNAP*) method.

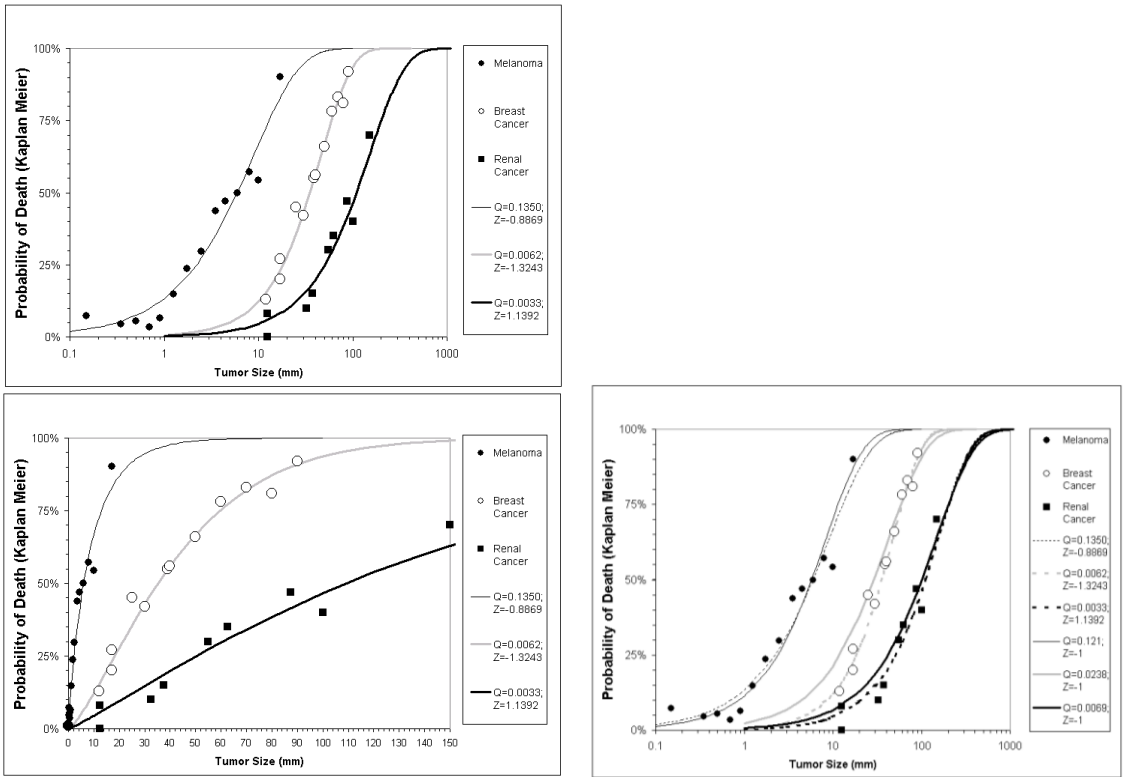

FIGURE * **Fit of Size/Lethality data for melanoma, breast, and renal cell carcinoma to the *SizeOnly* Equation (#1).** TOP: Size on log scale. BOTTOM: Size on conventional scale.
C:\X_MelanomaAnalysisByExcell 1 1 04___Survival Analysis of Basic Dataset by Size\groupsOf300\Renal\[FIGS.xls]Sheet1!$C$1

FIGURE *. Fit of Size/Lethality data for melanoma, breast, and renal cell carcinoma to the *SizeOnly* Equation (#1) with and without *Z*=1.

The *PrimarySizeOnly* Equation

Similarly, the relationship between the risk of death for node negative patients, $L_{primary}$, and tumor size, D, may be captured by a variant of the *SizeOnly* Equation, the *PrimarySizeOnly* Equation:

$$L_{primary} = 1 - e^{-(Q * j_{primary}) D^Z} \quad (1c)$$

$j_{primary}$ can be thought of as representing that fraction of events of spread occurring directly from the primary site to the periphery, causing death, while 1- $j_{primary}$ can be thought of that fraction of events of spread that begin at the primary site, spread to the nodes, and give rise to one or more progeny cells that spread to the periphery, causing death.

The *NodalSizeOnly* Equation

Similarly, the relationship between the fraction of patients with cancer in the nodes, $L_{to\text{-}nodes}$, and tumor size, D, may be captured by another variant of the *SizeOnly* Equation, the *NodalSizeOnly* Equation:

$$L_{To-Nodes} = 1 - e^{-Q_n D^Z} \quad (1n)$$

A first approximation regression method for estimating the value of the parameters of the *SizeOnly* Equation

The fit of data to these equations of the form of the *SizeOnly* Equation can be easily determined by regression by transforming them:

$$-\ln(1 - L_X) = Q_X D^Z \quad (13)$$

A *pseudo-Monte Carlo* method for refining the values of equation parameters

While these variants of the *SizeOnly* Equation accurately estimate the risk of cancer spread (L_x) for a group of patients all of whom have tumors of the same size (D), a group of W patients with tumors of various sizes will have an average chance of spread, L_{xW}, such that:

$$L_{xW} = 1 - (1/W) \sum_{i=1}^{w} e^{-Q_x D i^Z} \quad (1b)$$

where D_i is the size of the tumor in patient i. Reversing Equation #1b makes it possible to refine the estimate of the value of Q by a *pseudo-Monte Carlo* method, readily accomplished by spreadsheet software, in which we vary the value of Q until the average of the values of L_{xW} generated by Equation #1b for the W patients agrees with the actual value of L_{xW} for the population of W patients, as determined by Kaplan-Meier analysis.

The *Size+Nodes* method

The combined impact of tumor size and nodal status can be embraced by considering that the overall risk of cancer death, L, is the consequence of the lethal spread of cancer cells from the nodes, L_{nodes}, and the lethal spread of cancer cells from the primary site, $L_{primary}$, such that:

$$L = L_{primary} + L_{nodes} - (L_{primary} * L_{nodes}) \quad (4)$$

We have called this approach the *Size+Nodes* method. Let us define $L_{per\text{-}node}$ as the lethal contribution per positive node.. The value of $L_{per\text{-}node}$ can be derived from first principles, as we did for the *SizeOnly* Equation, as resulting from the spread of cancer cells from the nodes to the periphery (see Equation #9n below), or $L_{per\text{-}node}$ can be characterized operationally. Thus, if M is the number of nodes found to be positive in a group of patients with the same number of positive nodes, then the overall lethal contribution from those M nodes can be expected to be approximately:

$$L_{nodes} \approx M * L_{per-node} \quad (14)$$

Equation #14 becomes inaccurate for high values of M or $L_{per\text{-}node}$, but a more accurate estimate of L_{nodes} can be made by:

$$L_{nodes} = 1 - e^{-(M * L_{per-node})} \quad (2)$$

TABLE II

The *SizeOnly* and *Size+Nodes* Methods for Estimating the Risk of Cancer Death

(values are approximations: see other Technical Reports for more up to date values.)

When both tumor size and nodal status are known, the *Size+Nodes* estimates the risk of cancer death from information on tumor size and nodal status

$$L = L_{primary} + L_{nodes} - (L_{primary} * L_{nodes})$$

(eq. (4))

Source of Lethality	***Method of Estimation***	***Independent Variable***	***Parameters****	***Interpretation***
The Lethal Contribution from Cancer at the Primary Site Directly to the Periphery	$L_{primary} = 1 - e^{-(Q * j_{primary}) D^Z}$ eq. (1c)	*D* = Tumor Size: *For Breast Carcinoma:* Diameter (mm) *For Melanoma:* Thickness (mm)	*For Breast Carcinoma:* *Q*= 0.0062 *Z*=1.34 $j_{primary}$=0.661 if nodal status is known $j_{primary}$ =1 if nodal status is unknown *For Melanoma:* *Q*= 0.1428 *Z*= 0.89 $j_{primary}$=0.801 if nodal status is known $j_{primary}$=1 if nodal status is unknown	*The lethal contribution of the primary mass increases gradually with tumor size*
The Lethal Contribution from Cancer in the Lymph Nodes	$L_{nodes} = 1 - e^{-(M * R)}$ eq. (2)	*M* = The Number of Positive Nodes	*For Breast carcinoma:* *R*=0.0608 *For Melanoma* *R*=0.22527	*The presence of each positive lymph node contributes approximately "R" extra chance of death*

When only size is known, the *Size+Nodes* method reduces to the *SizeOnly* method for estimating the risk of cancer death from information on tumor size

$$L = 1 - e^{-QD^Z}$$

eq. (1)

TABLE III

A simple mnemonic for the *Size+Nodes* method			
	Lethal contribution for each mm of primary tumor size		Lethal contribution for each positive lymph node
Risk of Breast Carcinoma Death=	~1% per mm	+	~6% per node

Risk of Melanoma Death=	~8% per mm	+	~23% per node

The *Nodal Lethality* Equation

Re-arranging Equation #2 provides a way to estimate the lethal contribution per node, $L_{per\text{-}node}$:

$$L_{per\text{-}node} = (ln\ (1- L_{nodes})\ ) / M \quad (15)$$

While the value of L_{nodes} can be calculated by re-arranging Equation #4:

$$L_{nodes} = (L - L_{primary}) / (1- L_{primary}) \quad (16)$$

It follows that the lethal contribution per positive node, $L_{per\text{-}node}$, for a group of W patients with M positive nodes, tumors of size D, and a 15-year Kaplan-Meier death rate of L_W will be:

$$L_{per\text{-}node} = (ln\ (1- ((\ L_W - (1-e^{-(Q*jprimary)D\wedge z})) / ((-e^{-(Q*\ jprimary)D\wedge z}))))) / M \quad (17)$$

We shall call Equation #17 the *Nodal Lethality* Equation. While Equation #17 defines the value of $L_{per\text{-}node}$ for subpopulations of patients with specific numbers of positive nodes (M), the value of $L_{per\text{-}node}$ can be estimated for groups of patients with various numbers of nodes by a *pseudo-Monte Carlo* method, in a fashion analogous to that captured by Equation #1b.

Using the same line of reasoning outlined above, we are now able to estimate the probability of lethal spread per cell from the lymph nodes to the periphery:

$$p_{From-Nodes} = -\ln(1 - L_{per-node}) / N_{nodes} \quad (7n)$$

The value of N_{nodes} can be estimated with Equation #8 if we have data on the size of nodal deposits, which may be made by digital microscopy.

It follows that the probability of the spread of cancer cells from the primary site to the periphery, p, for a primary mass that is the same size as the mass of cancer seen in a node can also fits a power function:

$$p = aN_{nodes}^{\ b} \quad (9n)$$

The *SizeAssessment* method.

We have shown that for breast cancer and melanoma, the relationship between the risk of cancer death, L (defined operationally as the 15-year Kaplan-Meier cancer-specific death rate), and tumor size, D, is well captured by a simple expression, the *SizeOnly* equation:

$$L = 1 - e^{-QD^Z} \quad (1)$$

The *SizeOnly* equation makes it possible to determine whether a prognostic factor independently contributes to lethality, or is simply correlated with tumor size. In this test, which we call the *SizeAssessment* method, the actual 15-year cancer specific Kaplan-Meier death rate for a group of patients with a prognostic factor, $L_{empirical}$, is compared the predicted death rate, $L_{predicted}$, that would be expected by the *SizeOnly* equation for patients with tumors of these sizes. The statistical significance of the independent lethal contribution of this prognostic factor is assessed by comparing the difference of $L_{predicted}$ minus $L_{empirical}$ by an independent, two-sample Student's t-test with a threshold of 0.05.

The *SizeAssessment* method can be used to not only assess a prognostic factor's impact on the spread of cancer to the periphery, leading to death, but also to assess a prognostic factor's impact on the spread of cancer to the nodes, since both manifestations of the spread of cancer cells are well captured by equations of the form of the *SizeOnly* equation. The expression that relates tumor size to the chance of cancer in the nodes is called the *NodalSizeOnly* equation:

$$L_{to-nodes} = 1 - e^{-Q_{nodes} D^Z} \quad (2)$$

where $L_{to\text{-}nodes}$ is the fraction of patients found to have cancer in their lymph nodes.

The *PrognosticMeasurement* method measures the magnitude of a prognostic factor's independent contribution to lethality:

The magnitude of a prognostic factor's impact on lethality can be incorporated into the *SizeOnly* and *NodalSizeOnly* equations by adding multipliers for each prognostic factor, which we call *g* parameters:

$$L_X = 1 - e^{-Q(g_1 * g_2 * g_3 * g_4 * ..) D^Z} \quad (3)$$

The value of the *g* parameter for the lethal contribution of each prognostic factor in the *SizeOnly* equation can be determined by a *pseudo-Monte Carlo* method. We call this technique for quantifying a prognostic factors impact on lethality the *PrognosticMeasurement* method.

It follows that the independent contribution of a prognostic factor to the chance of spread of cancer to the nodes can also be considered in terms of g_n parameters inserted into the *NodalSizeOnly* equation.

While the value of a *g* parameter is an abstraction, there are two practical ways to comprehend the nature of its magnitude. *First*, since the *g* parameter sits next to the tumor size, *D*, in the *SizeOnly* equation, patients with tumors with a prognostic factor for which *g*=2 can be expected to have the same death rate as patients with tumors without the factor but with tumors of twice the size. Likewise, patients with tumors with a prognostic factor for which *g*=0.5 can be expected to have the same death rate as patients with tumors without the factor but with tumors of half the size. *Second*, as we have noted in the first paper in this series, the *SizeOnly* equation is roughly linear over most of its range. Thus, patients with tumors with a prognostic factor for which *g*=2 can be expected to have the roughly twice the death rate as patients with tumors without the factor, while patients with tumors with a prognostic factor for which *g*=0.5 can be expected to have the roughly half the death rate as patients with tumors without the factor.

It may well be expected that some prognostic factors will be found by the *SizeAssessment* method to make independent contributions to lethality, but whose contributions to lethality, as measured by the *PrognosticMeasurement* method, may be found to be trivial in magnitude. For example, while the lobular and ductal carcinoma phenotypes make statistically significant contributions to lethality, as determined by by the *SizeAssessment* method, as measured by the *PrognosticMeasurement* method, the value of the *g* parameter for lobular was 0.9032 and the value of its *g* parameter for ductal was 1.057. This means that patients with lobular carcinomas have roughly 90% of the death rate of patients with ductal carcinoma (Table 2). For this reason, we have chosen the somewhat arbitrary term "*marked*" to identify those prognostic factors found to have independent contribution to lethality with a *p* value greater than 0.05 by the *SizeAssessment* method, and whose *g* parameter is either <0.74 or >1.33.

The *Size+Nodes+PrognosticFactors* (*SNAP*) method combines tumor size, nodal status, and other prognostic factors into estimates of the risk of death.

Once the value of each prognostic factor's *g* parameter is known, we are able to combine information on tumor size, nodal status, and other prognostic factors with three linked equations to estimate of the risk of death, *L* for each patient:

$$L = L_{primary} + L_{nodes} - \left(L_{primary} * L_{nodes}\right) \quad (4)$$

where

$$L_{primary} = 1 - e^{-(Q * j_{primary})(g_1 * g_2 * g_3 * g_4 * ..) D^Z} \quad (5)$$

and

$$L_{nodes} = 1 - e^{-(M * L_{per-node})} \quad (6)$$

where *M* is the number of local lymph nodes found to be positive for cancer, and $L_{per\text{-}node}$ is the lethal contribution for each positive node. We call this technique the *Size+Nodes+PrognosticFactors* (*SNAP*) method (Table 1).

This provides a way to integrate information on primary tumor size, nodal status, and other prognostic factors into an estimate of the risk of cancer death for each patient, an approach we have called the *SNAP* (*Size+Nodes+Prognostic Factors*) method.

TABLE IV

THE *SizeAssessment* METHOD FOR DETERMINING THE LETHAL IMPACT OF PROGNOSTIC FACTORS SUSPECTED OF CONTRIBUTING TO CANCER LETHALITY	
Step 1	**Identify the sub-population of *W* patients with a candidate prognostic factor**
Step 2	**Use the *SizeOnly* Equation** $L = 1 - e^{-QD^Z}$ **(1)** **to calculate L_W, the average of all of the individual risks of cancer lethality (*L*) expected for the sizes of the tumors seen in the sub-population of *W* patients carrying the candidate prognostic factor**
Step 3	**Calculate L_{ACTUAL}, the actual 15-year Kaplan-Meier melanoma death rate for the sub-population of *W* patients carrying the candidate prognostic factor**
Step 4	**If $L_W \neq L_{ACTUAL}$, then it can be concluded that the candidate prognostic factor makes an independent contribution to cancer lethality (the statistical basis of this inequality can be assessed by the standard errors of the Kaplan-Meier and *SizeOnly* calculations)**
A similar calculation can be made with respect to the propensity of cancer to spread to the nodes, based on the observed and expected fraction of patients with cancer in the nodes, $L_{to\text{-}nodes}$, as determined by the *NodalSizeOnly* Equation (#1n): $L_{To-Nodes} = 1 - e^{-Q_{Nodes} D^Z}$	

TABLE V

THE *PrognosticMeasurement* METHOD FOR QUANTIFYING THE LETHAL IMPACT OF PROGNOSTIC FACTORS THAT CONTRIBUTE TO CANCER LETHALITY
For prognostic factors found the *SizeAssessment* method to act by increasing the propensity for lethal spread, use a *pseudo-Monte Carlo* method to capture the magnitude of the impact of the prognostic factor with this variant of the *SizeOnly* Equation: $L = 1 - e^{-(Q*g_1*g_2\ldots)D^Z}$ **(1b)** **where each *g* parameter is provides a measure of the impact of a prognostic factor**

TABLE VI

The *SNAP* (*Size+Nodes+PrognosticMarkers*) Method for Estimating the Risk of Cancer Death from Information on Tumor Size, Nodal Status, and Other Prognostic Factors

$$L = L_{primary} + L_{nodes} - (L_{primary} * L_{nodes})$$

(eq. (4))

Source of Lethality	*Method of Estimation*	*Independent Variable*	*Parameters*	*Interpretation*
The lethal contribution from cancer at the primary site	$L_{primary} = 1 - e^{-(Q * j_{primary}) * (g_1 * g_2 * g_3 * g_4 * ..) D^Z}$ (1c)	*D* = Tumor Size: *For Breast Carcinoma:* Diameter (mm) *For Melanoma:* Thickness (mm)	*For Breast Carcinoma:* *Q* = 0.0118395 *Z* = 1 $j_{primary}$=0.661 if nodal status is known $j_{primary}$ =1 if nodal status is unknown See Table 2 for *g* parameter values *For Melanoma:* *Q*= 0.1428 *Z*= 0.89 $j_{primary}$ =0.801 if nodal status is known $j_{primary}$ =1 if nodal status is unknown See Table 5 for *g* parameter values	*The lethal contribution of the primary mass increases gradually with tumor size, and the amount of that lethal contribution is influenced by prognostic factors, as captured by the g parameters in Equation 1d*
The lethal contribution from cancer in the lymph nodes	$L_{nodes} = 1 - e^{-(M * L_{per-node})}$ eq. (2)	*M* = The Number of Positive Nodes	*For Breast carcinoma:* $L_{per\text{-}node}$ = 0.0608 *For Melanoma:* $L_{per\text{-}node}$ = 0.2253	*The presence of each positive lymph node contributes approximately "$L_{per\text{-}node}$" extra chance of death*

The *SNAP* (*Size+Nodes+PrognosticMarkers*) method reduces to:

- **the *Size+Nodes* method, when only size and nodal status are known.**
- **the *SizeOnly* method, when only size is known.**

Numeric Computation of SNAP constants: an iterative method

Although the general method of calculating the SNAP constants has been explained, there has yet to be defined a method of generating these parameters quickly and efficiently. We have written a script in Matlab which can calculate the Q parameter for the entire population, and subsequently calculate $L_{per\text{-}node}$ and $J_{primary}$. We have also generated a script that can quickly calculate g-parameters for any desired number of groups. The method seeks iterated convergence of the mean calculated values for a cohort, to some given err tolerance ε. Typical values for ε are 0.0001, but this method has worked often with values as low as 10^{-7}.

The method follows these steps:

1. Calculate L_{actual} for cohort using the 15-Year Kaplan-Meier Lethality
2. Set Q=1;
3. For up to 1000 repetitions:
    a. Calculate $L_{estimate}=1-\exp(-Q*D)$ for each member of the cohort;
    b. Calculate the mean these $L_{estimate}$ values;
    c. If: $|L_{actual}-L_{estimate}| > \varepsilon * L_{actual}$
        Set $Q_{new}= (1+L_{actual}-L_{estimate})*Q_{old}$;
        Repeat Step a;
    d. Else:
        Break out of continuous loop;
        Return Q value;

Using this pseudo code, it is possible to calculate the parameter Q. The calculation of g-parameters can be thought of as calculating a Q parameter for a subset of the population.

1. Using the 15-Year Kaplan-Meier Lethality, Calculate L_{actual} for the subset of the cohort that fits the desired g-parameter criterion
2. Set $Q_{subset}=1$;
3. For up to 1000 repetitions:
    a. Calculate $L_{estimate}=1-\exp(-Q_{subset}*D)$ for each member of the cohort;
    b. Calculate the mean these $L_{estimate}$ values;
    c. If: $|L_{actual}-L_{estimate}| > \varepsilon * L_{actual}$
        Set $Q_{new}= (1+L_{actual}-L_{estimate})*Q_{old}$;
        Repeat Step a;
    d. Else:
        Break out of continuous loop;
        Return Q_{subset} value;
    e. g-parameter = Q_{subset}/Q_{all}

Where Q_{all} is the Q parameter calculated for all patients. Similarly, $J_{primary}$ can be thought of as a g-parameter that describes node-negativity, and therefore could be calculated among the node-negative patients.

1. Using the 15-Year Kaplan-Meier Lethality, Calculate L_{actual} for the node negative patients
2. Set $Q_{negative}=1$;
3. For up to 1000 repetitions:
    a. Calculate $L_{estimate}=1-\exp(-Q_{negative}*D)$ for each member of the cohort;
    b. Calculate the mean these $L_{estimate}$ values;
    c. If: $|L_{actual}-L_{estimate}| > \varepsilon * L_{actual}$
        Set $Q_{new}= (1+L_{actual}-L_{estimate})*Q_{old}$;
        Repeat Step a;
    d. Else:
        Break out of continuous loop;
        Return Q_{subset} value;
    e. $J_{primary} = Q_{negative}/Q_{all}$

Calculating the $L_{per\text{-}node}$ parameter uses a similar iterative method, but adapts the method to use the total SNAP lethality equation rather than the SizeOnly equation.

1. Using the 15-Year Kaplan-Meier Lethality, Calculate L_{actual} for the node negative patients
2. Set $L_{per\text{-}node}=1$;
3. For up to 10,000 repetitions:
    a. Calculate $L_{estimate}=L_{primary}+L_{nodes}-(L_{primary}*L_{nodes})$ where: $L_{primary}=1-exp(-Q*D)$ and $L_{nodes}=1-exp(-L_{per\text{-}node}*M)$
    b. Calculate the mean these $L_{estimate}$ values;
    c. If: $|L_{actual}-L_{estimate}| > \varepsilon * L_{actual}$
        Set $L_{pernode\text{-}new}= (1+L_{actual}-L_{estimate})*L_{pernode\text{-}old}$;
        Repeat Step a;
    d. Else:
        Break out of continuous loop;
        Return $L_{per\text{-}node}$ value;

Using these methods, we are able to calculate all of the constants necessary for the mathematical methods used in our lab. However, the calculation of g-parameters represents a particularly time consuming process, because of the necessary selection of each subgroup. We have written a script that generates the desired g-parameters for each permutation of subgroup. The script, *getgparameters.m* takes as inputs tumor size, survival information, censoring, and matrix of categorical data, *D*. Each column in *D* contains a vector of categorical information for each patient in the overall group.[1] The code generates a Q parameter for the entire group, and then for each column in D, generates a list of unique entries. It then selects the tumorsize, survival information, and censoring data for each entry in that unique list, and calculates the g-parameter, p-value, and 95% confidence interval. The pseudo code for this is:

1 Note that Matlab does not allow null values or strings in numerical arrays, and it is therefore required that every entry in D contains a numerical value.

1. Calculate L_{actual} for cohort using the 15-Year Kaplan-Meier Lethality
2. Set Q=1;
3. For up to 1000 repetitions:
4. Calculate $L_{estimate}$=1-exp(-Q*D) for each member of the cohort;
5. Calculate the mean these $L_{estimate}$ values;
6. If: $|L_{actual}-L_{estimate}| > \varepsilon * L_{actual}$
7. Set $Q_{new}= (1+L_{actual}-L_{estimate})*Q_{old}$;
8. Repeat Step a;
9. Else:
10. Break out of continuous loop;
11. Return Q value;
12. For each column in D
    1. The script generates a list of unique categorical variables for that column.
    2. For each unique entry
        1. Select tumorsize, survival information, and censor data for patients who have that specific value in this column
        2. Calculate L_{actual} for cohort using the 15-Year Kaplan-Meier Lethality
        3. Set Qi=1;
        4. For up to 10000 repetitions:
        5. Calculate $L_{estimate}$=1-exp(-Qi*D) for each member of the cohort;
        6. Calculate the mean these $L_{estimate}$ values;
        7. If: $|L_{actual}-L_{estimate}| > \varepsilon * L_{actual}$
        8. Set $Q_{new}= (1+L_{actual}-L_{estimate})*Q_{old}$;
        9. Repeat Step a;
        10. Else:
        11. Break out of continuous loop;
            1. Return Q value
            2. g-paramter = Q*i*/Q
        2. Repeat for next column in D;
        3. End

The *TimeOnly* Equation

There is abundant evidence that over the size ranges where most tumors are seen, growth appears to be exponential:

$$D = D_0 e^{rt} \tag{18}$$

where the growth rate constant, r is related to the tumor doubling time, D_T:

$$r = \ln(2) / D_T \tag{19}$$

The density-dependant change in growth rate that occurs above and below the tumor sizes seen clinically[27] has relatively little impact on the appearance of exponential growth that is evident over the rather narrow range of sizes seen clinically.

We can simplify Equation #11 by considering the $D_0 = 1mm$:

$$D = e^{rt} \tag{20}$$

By combining the *sSizeOnly* Equation (#1s) with the exponential growth equation (#13), we can derive an expression for relating the risk of cancer death and time, which we call the *TimeOnly* Equation:

$$L = 1 - e^{-Qe^{rt}} \tag{21}$$

Equation #14 allows us to see directly the rate of the increase in death rate with time, and, as can be seen in FIGURE S2, over the size ranges where most breast cancers are seen clinically (~10-~150 mm), Equation #14 is remarkably linear, such that R^2 =0.99. This demonstrates why the increase in the risk of cancer death appears to be roughly constant over time. This linear relationship between time and the risk of death shown in FIGURE S2 is the direct, if again unobvious, consequence of the most generally accepted mechanism of breast cancer death, which is by the spread of cancer cells occurring with definable probabilities per cell (Equation #7).

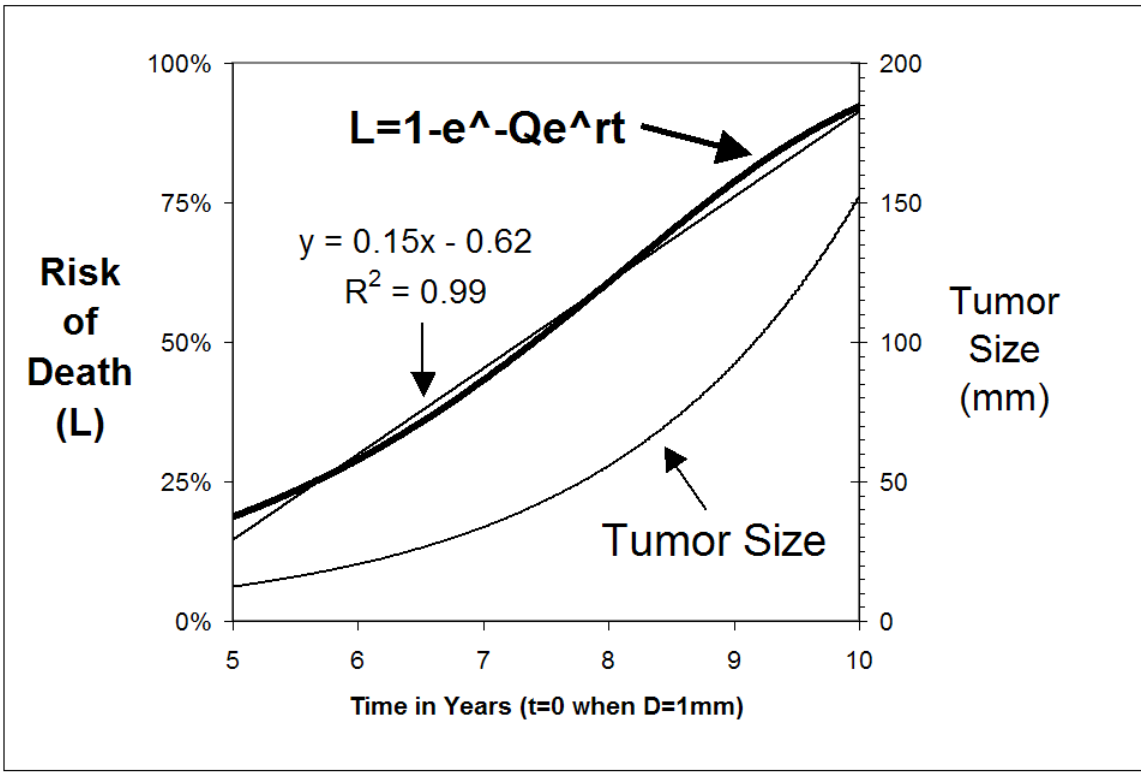

FIGURE S2

C:\My Documents (Death & Delay 9 2002)_Excell All\Time\[Time.xls]Time (5)'!$I$4

Computer Simulation Model of Cancer Screening

The *binary biological model of cancer metastasis* has also provided the basis of a biologically-based computer simulation model of cancer growth, spread, and detection[28,29-31], which can calculate such values as the relationship between the screening interval and the cancer survival rate. To execute this model, we have collected data[32] on the rate of breast cancer growth[33], the sizes at which breast cancers become detectable on mammographic and clinical grounds[34], and the relationship between tumor size and survival[35], which can make the simulation a realistic tool for determining the consequences of various usages of screening. The simulation calculates the simultaneous change in tumor size, lethality, and detectability that occurs over time for breast cancer, and then estimates that likely survival rate for various women, particularly as regards age, under various usages of screening. The simulation does not consider the detection of DCIS, now ~20% of breast cancer cancers[36], and thus probably provides a conservative estimate of the impact of screening on the breast cancer death rate[37]. The simulation was written in Visual Basic 6, and is available from Dr. Michaelson.

The core of the simulation is based upon a day-to-day increase in tumor cell number (*N*), expected for exponential growth (Equation #1 below) and the relationship between tumor size (*N*) and the fraction (*L*) of women with lethal metastatic disease (Equation #2 below):

(a) $$N_{today} = N_{yesterday} + (g * N_{yesterday})$$

where *g*, is the fraction of live cells today will be replaced by two live cells tomorrow (see Equation #4 below).

(6) $$L = 1 - e^{-Np}$$

The value of *p*, the probability, per cell, of an event of spread of breast cancer from the primary site in the breast to the periphery leading to death, which we have found[35] is calculated with the expression:

(c) $$p = 0.00005017 * N^{0.5575}$$

The value of *g* in Equation #1 is related to the tumor doubling time by the expression:

(d) $$g = (2 \wedge (1 / \textit{doubling time})) - 1$$

Unless otherwise stated, the *doubling time* used is the empirically based value of 130 days, as outlined in reference 33.

There is ample empirical support for the exponential growth of breast carcinomas[33,38-43], as captured by Equation #1. There is also ample empirical support for the relationship between tumor size and risk of death that is captured by Equations #2[29,35]. Equation #2 is biologically plausible, as it has been derived from a consideration of the most generally accepted mechanism of breast cancer death, which is by the spread of cancer cells[44]. Equation #2 has also been found to accurately predict the relationship between tumor size and the risk of death in five separate populations of breast carcinoma patients, as well as in sub-populations of patients whose tumors were detected at screening.[35]

At each daily iteration, the simulation determines whether the tumor would be detectable on clinical grounds, if it has reached size *Sp*, while mammographic detection is set to occur at the time of screening, if tumors have reached size *Sm*. Values of *N* are converted into values of tumor diameter, *D*, by assuming spherical geometry and cellular density *s* (in cells/mm^3), such that $s=10^8$ cells/cc, based on data on the size of cancer cells[45-47]. Values for *Sm* and *Sp* are expressed in terms of distributions captured by 5-part step functions, each value of which is applicable to 20% of the population. For *Sp*, the values for five part step function are 9 mm, 14 mm, 20 mm, 26 mm, and 69 mm, based on the sizes of the cancers in the control (no screening) population in the Swedish Two-country trial of mammography[1-3,34], while the distribution of values of *Sp* is captured by the five part step function with values of 3.5 mm, 5.5 mm, 7 mm, 9.2 mm, and 11.7 mm estimated from data on the sizes of the cancers seen at screening by the method in reference 34. The effect of age on this distribution is captured by multiplying each value with a *modifier*, such that the *modifier=2.16025 - (Age * 0.01785)*, again based upon empirical measures described in reference 34. This impact on the change in cancer detectablity is believed to be due to the age-associated decrease that occurs in the radiographic density of the breast.[48]

The simulation uses age-specific data on cancer incidence[49], and life expectancy[50] to calculate the benefit and marginal benefit of screening from *L* in Equation #2. Values are derived in terms of both

"years of life lost" and "cancer free years of life lost". The latter combines both the years of life a woman with metastatic disease looses to cancer plus the years of life she will live with metastatic disease (on average, 5.74 years, for the women at this institution). Population-wide benefits expected in a population with the age structure in the USA were estimated using USA census data[50] ("USA reduction in death"), as well as for an imaginary population with a uniform distribution of ages ("un-age-structure-adjusted reduction in death"); the latter allows a comparison of populations (such as US and UK) that have different age structures.

The simulation derived costs estimates, in terms of dollars/cancer free year of live saved, as determined by the direct cost of the mammogram and screening's indirect costs incurred by false positives[51]. Sickles and colleagues[52,53] found that screening mammograms can be carried out for $50, while the year 2007 Medicare reimbursement is $81.86, and insurance reimbursements above and below this value are common[54]. The cost ascribable to false positives have been found to be $25.32/mammogram.[55]

The simulation calculates the consequences of screening at 3651 intervals (1 day to 10 years, plus no screening) for women ages 20 to 85 yielding 241,296 age/interval permutations. Permuting these into all possible life-long screening strategies leads to an enormous number of possibilities, almost all of which are inefficient, i.e., for most permutations, there is another permutation that can achieve a greater reduction in death with fewer mammograms. However, the small number of efficient strategies can be identified as those in which all individuals (i.e. all ages) receive the same marginal benefit, a technique known as the equimarginal method[56,57]. The simulation isolated 2200 such efficient strategies.

[1] Michaelson J, Reducing Delay in the Detection and Treatment of Breast Cancer. Adv Imag Onc In Press 2007

[2] Michaelson J, Mammographic Screening: Impact on Survival in CANCER IMAGING Ed: M.A. Hayat in press 2007

[3] Blanchard K, Weissman J, MoyB, PuriD, Kopans D,, , Kaine E, MooreR, Halpern E, Hughes K, Tanabe K, Smith B Michaelson J, Mammographic screening: Patterns of use and estimated impact on breast carcinoma survival Cancer 101, 495-507 2004

[4] Michaelson JS, Silverstein M, Sgroi D, Cheongsiatmoy JA, Taghian A, Powell S, Hughes K, Comegno A, Tanabe KK, Smith B The effect of tumor size and lymph node status on breast carcinoma lethality. CANCER 98:2133-43 2003

[5] Michaelson JS, Satija S, Kopans DB, Moore RA, Silverstein, M, Comegno A, Hughes K, Taghian A, Powell S, Smith, B Gauging the Impact of Breast Cancer Screening, in Terms of Tumor Size and Death Rate Cancer 98:2114-24 2003

[6] Michaelson JS Silverstein M, Wyatt J Weber G Moore R Kopans DB, Hughes, K. Predicting the survival of patients with breast carcinoma using tumor size CANCER 95: 713-723 2002

[7] Michaelson JS Satija S, Moore R Weber G Garland G Phuri, D. Kopans DB The Pattern of Breast Cancer Screening Utilization and its Consequences CANCER 94:37-43 2002

[8] Michaelson JS Using Information on Breast Cancer Growth, Spread, and Detectability to Find the Best Ways To Use Screening to Reduce Breast Cancer Death Woman's Imaging 3:54-57 2001

[9] Michaelson, JS, Kopans, DB, Cady, B. The Breast Cancer Screening Interval is Important. Cancer 2000 88:1282-1284

[10] Michaelson, J. The Role Of Molecular Discreteness In Normal And Cancerous Growth. Anticancer Research.

[11] Michaelson, J, Halpern, E, Kopans, D. A Computer Simulation Method For Estimating The Optimal Intervals For Breast Cancer Screening. Radiology

[12] Michaelson JS, Cheongsiatmoy JA Dewey F, Silverstein M, Sgroi D Smith B. Tanabe KK, The Spread of Human Cancer Cells Occurs with Probabilities Indicative of A Non Genetic Mechanism British Journal of Cancer 93:1244-1249 2005

[13] Michaelson, J. The Role Of Molecular Discreteness In Normal And Cancerous Growth. *Anticancer Research*. 19: 4853-4867 1999

[14] Michaelson JS, Cheongsiatmoy JA Dewey F, Silverstein M, Sgroi D Smith B. Tanabe KK, The Spread of Human Cancer Cells Occurs with Probabilities Indicative of A Non Genetic Mechanism *British Journal of Cancer* 93:1244-1249 2005

[15] Michaelson JS Silverstein M, Wyatt J Weber G Moore R Kopans DB, Hughes, K. Predicting the survival of patients with breast carcinoma using tumor size *CANCER* 95: 713-723 2002

[16] Surbone A, and Norton L. Kinetic Concepts in the Treatment of Breast Cancer. Annals of the New York Academy of Science. 698:48-62 1993

[17] Norton L. A Gompertzian Model of Human Breast Cancer Growth Cancer Research. 1988; 48:7067-7071.

[18] Kuukasjarvi T Karhu R Tanner M Kahkonen M Schaffer A Nupponen N Pennanen S Kallioniemi A Kallioniemi OP Isola J Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. *Cancer Res* 57:1597-604 1997

[19] Folkman J. Angiogenesis in cancer, vascular, rheumatoid and other disease. Nat Med:27-31 1995

[20] Folkman J, D'Amore PA . Blood vessel formation: what is its molecular basis? *Cell* 87:1153-5 1996

[21] Pluda JM Tumor-associated angiogenesis: mechanisms, clinical implications, and therapeutic strategies. *Semin Oncol.* 24:203-18 1997

[22] Hanahan D, Folkman J. Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell* 86:353-64 1996
[23] Karrison TG Ferguson DJ Meier P Dormancy of mammary carcinoma after mastectomy. *J Natl Cancer Inst* 91:80-5 1991
24. Van der Linden HC, Baak JP, Smeulders AW, Lindeman J, Meyer CJ. Morphometry of breast cancer. I. Comparison of the primary tumours and the axillary lymph node metastases. *Pathol Res Pract.* 1986;181:236-242.
25. Boon ME, Trott PA, van Kaam H, Kurver PJ, Leach A, Baak JP. Morphometry and cytodiagnosis of breast lesions. *Virchows Arch A Pathol Anat Histol.* 1982;396:9-18.
26. Pesce C, Colacino R. Morphometry of the breast fibroadenoma. *Pathol Res Pract.* 1986;181:718-720.
[27] Norton L. A Gompertzian model of human breast cancer growth. *Cancer Res*. 1988 Dec 15;48(24 Pt 1):7067-71.
[28] Michaelson J, Mammographic Screening: Impact on Survival in *CANCER IMAGING* Ed: M.A. Hayat in press 2007
[29] Michaelson, J. The Role Of Molecular Discreteness In Normal And Cancerous Growth. *Anticancer Research*. 19: 4853-4867 1999
[30] Michaelson J Halpern E Kopans D. A Computer Simulation Method For Estimating The Optimal Intervals For Breast Cancer Screening. *Radiology*. 212:551-560 1999
[31] Michaelson JS Using Information on Breast Cancer Growth, Spread, and Detectability to Find the Best Ways To Use Screening to Reduce Breast Cancer Death Program *Journal of Woman's Imaging* 3:54-57 2001
[32] Michaelson JS Satija S Moore R Weber G Garland G Kopans DB, Observations on Invasive Breast Cancers Diagnosed in a Service Screening and Diagnostic Breast Imaging Program *Journal of Woman's Imaging* 3:99-104 2001
[33] Michaelson JS Satija S, Moore R Weber G Garland G Kopans DB Estimates of the Breast Cancer Growth Rate and Sojourn Time from Screening Database Information *Journal of Women's Imaging* 5:3-10 2003
[34] Michaelson JS Satija S, Moore R Weber G Garland G Kopans DB, Hughes, K. Estimates of the Sizes at which Breast Cancers Become Detectable on Mammographic and on Clinical Grounds *Journal of Women's Imaging* 5:10-19 2003
[35]Michaelson JS Silverstein M Wyatt J Weber G Moore R Kopans DB HughesK. The Prediction of Breast Cancer Survival from Tumor Size *Cancer* 95:713-723 2002
[36] Silverstein MJ Ductal carcinoma in situ of the breast. In: Annu Rev Med (2000) 51:17-32 ISSN: 0066-4219
[37] Mokbel K Towards optimal management of ductal carcinoma in situ of the breast. In: Eur J Surg Oncol (2003 Mar) 29(2):191-7
[38] Fournier DV Weber E Hoefken W Bauerm M Kubli F Barther V. Growth Rate of 147 Mammary Carcinoma. Cancer. 1980; 45:2198-2207.
[39] Lundgren B. Observations on the growth Rate of Breast Carcinomas and it's Possible Implications for Lead Time Cancer. 1977; 40:1722-1725.
[40] Spratt JA Von Fournier D Spratt JS Weber EE Mammographic Assessment of Human Breast Cancer Growth and Duration Cancer. 1993; 71:2020-2026.
[41] Spratt JS, Spratt T. Rates of Growth of Pulmonary Metasases and Survival Annals of Surgery. 1964; 159:161-171.
[42] Peer PGM van DickJ AAM Hendricks JHCL Holland R Verbeek ALM. Age-Dependent Growth of Primary Breast Cancer. Cancer. 1993; 71:3547-3551.
[43] Kuroishi T Tominaga S Morimoto T Tashiro H et al Tumor Growth Rate and Prognosis of Breast Cancer Mainly Detected by Mass Screening Japanese Journal of Cancer Research. 1990; 81:454-462

[44] Michaelson JS, Cheongsiatmoy JA Dewey F, Silverstein M, Sgroi D Smith B. Tanabe KK. The Spread of Human Cancer Cells Occurs with Probabilities Indicative of A Non Genetic Mechanism British Journal of Cancer 93:1244-1249 2005

[45] Van der Linden HC Baak JP Smeulders AW Lindeman J Meyer CJ Morphometry of breast cancer. I. Comparison of the primary tumours and the axillary lymph node metastases. *Pathol Res Pract* 181:236-42 1986

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

[47] Pesce C Colacino R Morphometry of the breast fibroadenoma. *Pathol Res Pract* 181:718-20 1986

[48] Michaelson JS, Satija S, Kopans DB, Moore RA, Silverstein, M, Comegno A, Hughes K, Taghian A, Powell S, Smith, B Gauging the Impact of Breast Cancer Screening, in Terms of Tumor Size and Death Rate *Cancer* 98:2114-24 2003

[49] Ries LAG, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg L, Edwards BK (eds). SEER Cancer Statistics Review, 1973-1997, National Cancer Institute. Bethesda, MD, 2000.

[50] Anonymous. Statistical Abstracts of the United States 1998 US Government Printing Office

[51] Tammy O. Tengs; Miriam E. Adams; Joseph S. Pliskin; Dana Gelb Safran; Joanna E. Siegel; Milton C. Weinstein; and John D. Graham. Five-Hundred Life-Saving Interventions and Their Cost-Effectiveness *Risk Analysis* 15:369-390 1995

[52] Sickles EA Weber WN Galvin HB Ominsky SH Sollitto RA Mammographic screening: how to operate successfully at low cost. Radiology (1986 Jul) 160(1):95-7

[53] Sickles EA Low-cost mass screening for breast cancer with mammography AJR Am J Roentgenol 1992 158:55-7

[54] Brice, J. Small change for big medicine: the crisis in breast imaging services. *Diagnostic Imaging* 2000 22:42-49

[55] Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. *N Engl J Med* 1998;338:1089-96.

[56] Samuelson PA, Nordhaus WD, ECONOMICS, 16th edition, Irwin McGraw-Hill, 1998

[57] Friedman, David D. Price Theory: An Intermediate Text South-Western Publishing Co. 1990