Original Article

How Cancer at the Primary Site and in the Lymph Nodes Contributes to the Risk of Cancer Death

James S. Michaelson, PhD[1,2,3]; L. Leon Chen, BS[2]; Melvin J. Silverstein, MD[4]; Martin C. Mihm, Jr, MD[1,3]; Arthur J. Sober, MD[5,6]; Kenneth K. Tanabe, MD[2,7]; Barbara L. Smith, MD, PhD[2,7]; and Jerry Younger, MD[8,9]

BACKGROUND: It has long been appreciated that tumor size, lymph node status, and patient survival are related qualities, although how to isolate their interactions has not been obvious, nor has it been obvious how to integrate tumor size and lymph node status into predictions of the risk of death for individual patients. **METHODS:** The authors describe a mathematical method, the binary-biological model of cancer metastasis, based on the spread of cancer cells, in which the equations capture the relations between tumor size, lymph node status, and cancer lethality. **RESULTS:** For melanoma, renal cell carcinoma, and breast carcinoma, the relation between tumor size and the risk of cancer death was captured by the *SizeOnly* equation. For melanoma and breast carcinoma, the relation between tumor size and the presence of cancer in the lymph nodes was captured by using the *NodalSizeOnly* equation. For lymph node-negative melanoma and breast carcinoma, the relation between tumor size and risk of death was captured by the *PrimarySizeOnly* equation. For breast carcinoma, the model indicated that each positive lymph node contributed ~6% extra risk of death, whereas each millimeter of greatest primary tumor dimension contributed ~1% risk of death. For melanoma, each positive lymph node contributed ~23% risk of death, whereas each millimeter of primary melanoma thickness contributed ~8% risk of death. This information was captured by a pair of linked equations, the *Size+Nodes* method. **CONCLUSIONS:** Both tumor size and the number of positive lymph nodes made independent contributions to the risk of cancer death, as estimated by using the *Size+Nodes* method. **Cancer 2009;000:000–000. © 2009 American Cancer Society.**

KEY WORDS: cancer death, risk, prediction, tumor size, lymph node status, breast cancer, melanoma.

It has long been appreciated that primary tumor size, lymph node status, and survival are related qualities for many cancers, although it has not been obvious how to isolate the interactions between each component nor how to integrate information on tumor size and lymph node status into predictions of the risk of death.[1,2] We have observed that, for breast carcinoma, a simple expression, the *SizeOnly* equation, accurately captures the relation between greatest primary tumor dimension and lethality.[3,4] By double-sorting

Corresponding author: James S. Michaelson, PhD, Division of Surgical Oncology, Yawkey Building 7, Massachusetts General Hospital, 32 Fruit Street, Boston, MA 02114; Fax: (617) 724-3895; michaelj@helix.mgh.harvard.edu

[1]Department of Pathology, Massachusetts General Hospital, Boston, Massachusetts; [2]Department of Surgery, Massachusetts General Hospital, Boston, Massachusetts; [3]Department of Pathology, Harvard Medical School, Boston, Massachusetts; [4]Breast Service, Hoag Memorial Hospital Presbyterian, Newport Beach, California; [5]Department of Dermatology, Massachusetts General Hospital, Boston, Massachusetts; [6]Department of Dermatology, Harvard Medical School, Boston, Massachusetts; [7]Department of Surgery, Harvard Medical School, Boston, Massachusetts; [8]Department of Medicine, Harvard Medical School, Boston, Massachusetts; [9]Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts

Received: December 8, 2008; **Revised:** April 9, 2009; **Accepted:** April 20, 2009

Published online Month 00, 2009 in Wiley InterScience (www.interscience.wiley.com)

DOI: 10.1002/cncr.24592, www.interscience.wiley.com

patients with breast carcinoma by tumor size and lymph node status, we also have observed that each positive lymph node is associated with ~6% risk of death, whereas the lethal contribution from cancer at the primary site can be estimated from the greatest tumor dimension with a variant of the *SizeOnly* equation such that each millimeter of greatest primary tumor dimension is associated with ~1% risk of death. It follows that the overall risk of death is the sum of the risks of death from the lymph nodes and the primary site, as captured by a pair of linked equations: the *Size+Nodes* method.[5] In the absence of information regarding lymph node status, the *Size+Nodes* method reduces to the *SizeOnly* equation.

The *SizeOnly* and *Size+Nodes* methods not only provide highly practical and accurate tools for estimating breast carcinoma outcome, they also are biologically plausible, because their equations have been derived by a consideration of the most generally accepted mechanism of cancer death, ie, by the spread of cancer cells occurring with definable probabilities per cell.[6-8] We call this approach the binary-biological model of cancer metastasis. We have described various aspects of this mathematical framework; herein, we present this mathematical framework in full.

Although it has been established that the *SizeOnly* and *Size+Nodes* methods accurately capture the relation between tumor size, lymph node status, and the risk of death for breast carcinoma, their applicability to other cancers has not been examined. Herein, we outline a more general and efficient mathematical technique (the *Nodal Lethality* and *PrimarySizeOnly* equations) for isolating the impact of tumor size and lymph node status on the risk of cancer death. We examine the applicability of the *SizeOnly* method for relating tumor size to the risk of death for breast carcinoma, melanoma, and renal cell carcinoma and the applicability of the *Size+Nodes* method for relating tumor size and lymph node status to the risk of death from breast carcinoma and melanoma. We also examine the relation between primary tumor size and the risk of cancer in the lymph nodes for breast carcinoma and melanoma. In the accompanying articles, we extend these finding to examine the reason for the lethal contributions of cancer at the primary site and in the lymph nodes,[8] the impact of prognostic factors other than primary tumor size and lymph node status on the risk of cancer death,[9] and the application of these finding to the development of web-based calculators that physicians can use to estimate the risk of cancer death (available at: http://www.CancerMath.net).[9]

MATERIALS AND METHODS

Data

Information regarding tumor thickness and survival was obtained for 2770 patients with melanoma who were seen at Massachusetts General Hospital from January 1, 1970 to May 1, 2002; lymph node status was known for 664 patients. Information concerning tumor size, lymph node status, and breast carcinoma survival was derived from 1352 patients who had invasive breast carcinoma diagnosed between 1966 and 1990 at the University of Southern California/Van Nuys Breast Center, and their survival was estimated as of December 31, 2000.[3] Information regarding renal cell carcinoma survival was taken from the reports by Delahunt et al[10] and Hafez et al.[11]

Karrison et al[12] reported that little lethality occurs 15 years after diagnosis, and we observed a similar hazard function for melanoma (available at: http://www.CancerMath.net). Thus, we have relied on the 15-year cancer-specific Kaplan-Meier death rate as our measurements of the cancer death rate (L).

In addition, to confirm our findings, we used breast carcinoma data from 362,491 patients and melanoma data from 90,771 patients in the Surveillance, Epidemiology, and End Results dataset and breast carcinoma data from 11,271 patients who were seen at Massachusetts General Hospital and Brigham and Women's Hospital between 1960 and 2003.[9] Further details can be found within the technical reports available at www.CancerMath.net.

Mathematical Essentials: The Binary-Biological Model of Cancer Metastasis

For consistency with previous publications, Equations 1 through 4 are numbered in agreement with our previous reports[3,5] and, thus, appear out of order when presented below.

In this report and in 1 of the accompanying articles,[8] we examine macroscopic features of cancer—tumor size, lymph node 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 with 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 lymph node-negative patients dying of cancer metastasis; and L will be the fraction of all patients, both lymph node positive and lymph 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, $p_{primary\text{-}to\text{-}nodes}$ is the probability of a single successful event of spread to the lymph nodes for each of the N cells in a primary tumor (thus related to $L_{to\text{-}nodes}$), whereas $p_{primary\text{-}to\text{-}periphery}$ 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 lymph 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 lymph node status (related to L). Let us also define $p_{node\text{-}to\text{-}periphery}$ 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 nonlethal 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 that p_x is constant,[3,7,8] 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 accompanying article this series and the report by Michaelson et al.[7,8]

Because 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 1 or more lethal metastases will be $(1 - p_x)^N$. Thus, in a population of patients with tumors of identical size, the fraction that has 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 number of cells, N, can be estimated by assuming a density of s (in cells/cm^3) and spherical geometry, ie,

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

with a plausible value for s set at 10^8 cells/mL.[13-15]

By using Equation 7, we have observed that,[3,7] for both lethal and nonlethal spread of breast carcinoma and melanoma, the value of p declines gradually as tumors increase in size, N, and indeed can be closely fit by a power function,

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

in which $b \approx -2/3$. Let us define

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

and

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

Combining Equations 6 through 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 lymph nodes [$L_{to\text{-}nodes}$]), and tumor size (greatest dimension or thickness [D]:

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

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

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

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

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

The term $j_{primary}$ can be thought of as representing that fraction of events of spread occurring directly from the primary site to the periphery, causing death, whereas $1 - j_{primary}$ can be thought of as that fraction of events of spread that begin at the primary site, spread to the lymph nodes, and give rise to 1 or more progeny cells that spread to the periphery, causing death. Similarly, the relation between the fraction of patients with cancer in the lymph nodes, $L_{to\text{-}nodes}$, and tumor size, D, may be captured by another variant of the *SizeOnly* equation, the *NodalSizeOnly* equation:

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

The fit of data to these equations can be determined easily in a regression by transforming them:

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

Although these variants of the *SizeOnly* equation accurately estimate the risk of cancer spread (L_x) for a group of patients with 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

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

in which 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 using a pseudo-Monte Carlo method (readily accomplished by using spreadsheet software) in which we vary the value of Q until the average of the values of L_{xW} generated by Equation 1b for W patients agrees with the actual value of L_{xW} for the population of W patients, as determined by Kaplan-Meier analysis.

The combined impact of tumor size and lymph node 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 lymph 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.[5] We define $L_{per\text{-}node}$ as the lethal contribution per positive lymph node.[5] 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 lymph nodes to the periphery (see Equation 9n, below), or $L_{per\text{-}node}$ can be characterized operationally. Thus, if M is the number of lymph nodes identified as positive in a group of patients with the same number of positive lymph nodes, then the overall lethal contribution from those M lymph 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 $L_{per\text{-}node}$. However, a more accurate estimate of L_{nodes} can be calculated by

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

Rearranging Equation 2 provides a way to estimate the lethal contribution per lymph node, $L_{per\text{-}node}$,

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

whereas the value of L_{nodes} can be calculated by rearranging Equation 4:

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

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

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

We call Equation 17 the *Nodal Lethality* equation. Although Equation 17 defines the value of $L_{per\text{-}node}$ for subpopulations of patients with specific numbers of positive lymph nodes (M), the value of $L_{per\text{-}node}$ can be estimated for groups of patients with various numbers of

J_ID: Z7B Customer A_ID: C2696-08 Cadmus Art: CNCR24592 Date: 25-July-09 Stage: Page: 5

lymph nodes by using a pseudo-Monte Carlo method in a fashion analogous to that captured by Equation 1b.

By 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:

$$p_{node\text{-}to\text{-}periphery} = -\ln(1 - L_{per\text{-}node})/N_{nodes}. \quad (7n)$$

The value of N_{nodes} can be estimated with Equation 8 if we have data on the size of lymph node deposits, which may be obtained by using 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 observed in a lymph node also fits a power function:

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

Finally, it is widely believed that there are other factors, including those defined by gene expression arrays, that may be expected to have an impact on the risk of death.[16-19] Therapeutic innovations, such as the use of adjuvant chemotherapy, hormone therapy, and radiation therapy, also may be expected to have such an impact on the risk of death.[20] Comparing the actual 15-year Kaplan-Meier cancer death rate of the patients with such variables with the 15-year Kaplan-Meier cancer death rate expected by the *SizeOnly* equation (Equation 1) using information regarding tumor size provides a way to determine whether a prognostic factor actually makes an independent contribution to cancer lethality, which we call the *SizeAssessment* method, as outlined in the third article in this series.[9] Similarly, comparing the fraction of actual lymph node-positive patients with the fraction of lymph node-positive patients expected by the *NodalSizeOnly* equation (Equation 1n) provides another way to determine whether a prognostic factor makes a contribution to a cancer's propensity for spread. The actual magnitude of the impact of such a factor may be incorporated into the *SizeOnly*, *NodalSizeOnly*, and *Size+Nodes* methods by adding multipliers we call g parameters, which are specific for each prognostic factor:

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

The value of the g parameter for a prognostic factor can be determined by using a pseudo-Monte Carlo method in a fashion analogous to that captured by Equation 1b. We call this approach to quantifying the lethal impact of prognostic factors the *PrognosticMeasurement* method. From an operational standpoint, we identify prognostic factors that make a "marked independent contribution to lethality" if, by using the *SizeAssessment* method, the difference between the observed and expected values exceeds the sum of the 95% confidence intervals and if the value of the g parameter calculated by the *PrognosticMeasurement* method is <0.75 or >1.25.

It follows that Equation 1g can be combined with Equation 1c to generate an expression that makes a more accurate estimate of the lethal contribution from the primary site:

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

When combined with Equations 2 and 4, this provides a way to integrate information concerning primary tumor size, lymph node status, and other prognostic factors into an estimate of the risk of cancer death for each patient, an approach that we call the *Size+Nodes+Prognostic Factors* (*SNAP*) method.

RESULTS

The Relation Between Tumor Size and Cancer Death Is Captured Well by the SizeOnly *Equation*

The fit of tumor size/lethality data to the *SizeOnly* equation (Equation 1),

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

can be tested in a regression by transforming it into Equation 13. Figure 1 illustrates that, for melanoma, such a regression reveals that $Z = 0.89$ and $Q = 0.134$ ($R^2 = 0.86$); whereas, for renal cell carcinoma, $Z = 1.14$ and $Q = 0.0033$ ($R^2 = 0.90$); and, for breast carcinoma, $Z = 1.33$ and $Q = 0.0061$ ($R^2 = 0.97$). Refining the value of Q by using a pseudo-Monte Carlo method (Equation 1b) yields a slightly more accurate value of $Q = 0.1428$ for melanoma and $Q = 0.0062$ for breast carcinoma F1

ID: selvarajn | Black Lining: [ON] | Time: 14:57 | Path: N:/3b2/CNCR/Vol00000/090511/APPFile/C2CNCR090511

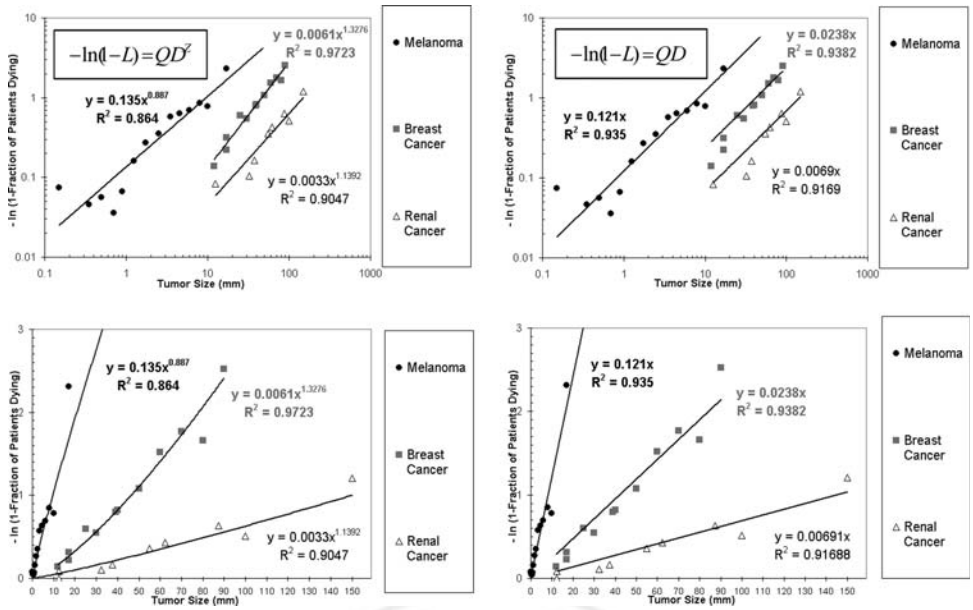

FIGURE 1. The fit of melanoma, breast cancer, and renal cell carcinoma size (*D*)/lethality (*L*) data to the *SizeOnly* equation (Equation 1) was assessed by using Equation 13. On the right, the fit of these data to the *SizeOnly* equation (Equation. 1) is examined with $Z = 1$. The conversion of tumor sizes into values of cell number (*N*) was calculated using Equation 8.

Table 1. Methods of the Binary-Biological Model of Cancer Metastasis

Method	Purpose
ProbabilityEstimation equation (Eq. 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 nonlethal consequences of the spread
SizeOnly equation (Eq. 1)	Relates tumor size to the chance of cancer death
PrimarySizeOnly equation (Eq. 1c)	Relates tumor size to the chance of cancer death for lymph node-negative patients
NodalSizeOnly equation (Eq. 1n)	Relates tumor size to the chance of cancer in the lymph nodes
Size+Nodes method	Integrates information on tumor and number of positive lymph nodes into an estimate of the chance of cancer death
NodalLethality equation (Eq. 17)	Calculates the lethal contribution, per positive lymph 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
Size+Nodes+Prognostic Factor (*SNAP*) method	Integrates information on primary tumor size, lymph node status, and other prognostic factors into an estimate of the chance of cancer death

*When only tumor size and lymph node 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 the number of positive lymph nodes; whereas, 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.

(Table 1). The close fit of these data to the *SizeOnly* equation (Equation. 1) is illustrated in Figure 2.

The fit of these data to the *SizeOnly* equation (Equation 1) also is close when $Z = 1$ (Fig. 1, right). This simplifies the *SizeOnly* equation (Equation 1) and makes it possible to perform 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, as we discuss in the third article in this series.[9]

J_ID: Z7B Customer A_ID: C2696-08 Cadmus Art: CNCR24592 Date: 25-July-09 Stage: Page: 7

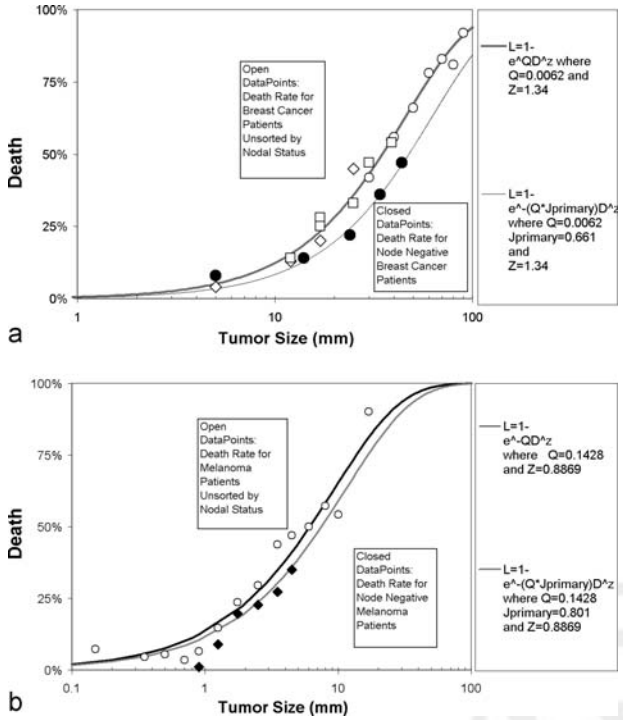

FIGURE 2. (a) This chart illustrates the relation between tumor size (*D*) and cancer death (*L*) for all patients with breast carcinoma (the *NodalSizeOnly* equation [Equation 1]) and for patients with lymph node-negative breast carcinoma (the *PrimarySizeOnly* equation [Equation 1c]). (b) This chart illustrates the relation between tumor size and cancer death for all patients with melanoma (the *NodalSizeOnly* equation [Equation 1]) and for patients with lymph node-negative melanoma (the *PrimarySizeOnly* equation [Equation 1c]). $J_{primary}$ indicates the fraction of events of spread occurring directly from the primary site to the periphery.

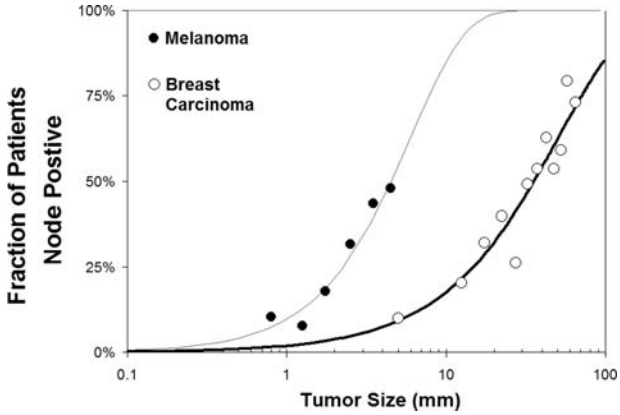

FIGURE 3. This chart illustrates the relation between primary tumor size and the risk of lymph node positivity, as captured by the *NodalSizeOnly* equation (Equation 1n). Values of Q_n and Z were determined by a regression of Equation 13 (for breast carcinoma, $Q_n = 0.019$, $Z = 1.0041$, and $R^2 = 0.902$; for melanoma, $Q_n = 0.1018$, $Z = 1.207$, and $R^2 = 0.8851$)

The Relation Between Tumor Size and the Fraction of Lymph Node-Positive Patients Is Captured Well by the NodalSizeOnly *Equation*

The fit of tumor size/lymph node positivity data to the *NodalSizeOnly* equation (Equation 1n),

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

can be tested in a regression by transforming it into Equation 13. Such a regression reveals that for melanoma, $Q_n = 0.1018$ and $Z = 1.207$ ($R^2 = 0.885$) whereas, for breast carcinoma, $Q_n = 0.019$ and $Z = 1.0041$ ($R^2 = 0.90$). The capacity of the *NodalSizeOnly* equation (Equation. 1n) using these parameters to capture the relation between tumor size and the fraction of patients that are lymph node positive can be observed in Figure 3.

The Relation Between Tumor Size and Cancer Death for Lymph Node-Negative Patients Is Captured Well by the PrimarySizeOnly *Equation*

We previously observed that, for breast carcinoma, the relation between greatest tumor dimension (*D*) and the risk of cancer death for lymph node-negative patients ($L_{primary}$) is well fit to the *PrimarySizeOnly* equation (Equation 1c),

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

for which $j_{primary} = 0.661$, as determined by a pseudo-Monte Carlo method (Equation 1c) (Fig. 2a). Figure 2b illustrates that this also is the case for melanoma, such that $j_{primary} = 0.801$.

Each Positive Lymph Node Is Associated With an Extra Risk of Lethality: ~23% per Positive Lymph Node for Melanoma and ~6% per Positive Lymph Node for Breast Carcinoma

The impact of lymph node status on breast carcinoma and melanoma lethality cannot be observed directly from the survival of patients who have various numbers of positive lymph nodes, because tumor size and lymph node status are conflated: as the size increases, both the fraction of

ID: selvarajn | Black Lining: [ON] | Time: 14:57 | Path: N:/3b2/CNCR/Vol00000/090511/APPFile/C2CNCR090511

AQ9 **Table 2.** Cancer Death Rates for Groups of Melanoma and Breast Carcinoma Patients Sorted by Lymph Node Status

Patient Group	Average No. of Positive Lymph Nodes (*M*)	No. of Patients	Mean Tumor Thickness or Dimension, mm (*D*)	Actual 15-Year Disease-Specific Death: $L_{overall}$, %	The Level of Death Expected for Lymph Node-Negative Patients With Tumors of These Sizes: $L_{primary} = 1-e^{-(Q^*jprimary)D^z}$ (Eq. 1c), %	Extra Death Ascribable to Positive Lymph Nodes, per Lymph Node ($L_{per\text{-}node}$) (Eq. 17), %
Melanoma						
Lymph node negative	0	487	2.35	21	21	—
No. of positive lymph nodes						
1	1	92	3.24	40	25	21
2	2	36	3.43	55	27	24
3	3	21	4.14	67	32	24
4-5	4.5	12	5.11	77	34	23
6-7	6.5	10	3.64	86	30	24
Breast carcinoma						
Lymph node negative	0	790	16.0	20	20	—
No. of positive lymph nodes						
1	1	130	25.7	26	26	0.16
2	2	71	29.0	34	29	3.96
3	3	46	30.7	37	32	2.66
4	4	47	35.4	57	36	9.97
≥1	5.28	443	35.3	49	35	4.52
≥2	7.06	313	39.3	58	38	5.35
≥3	8.76	233	42.9	64	42	5.58
≥4	9.85	196	45.0	69	42	6.21
≥5	11.70	149	48.1	73	46	5.84
≥6	13.11	123	52.3	75	49	5.48

patients with positive lymph nodes and the average number of positive lymph nodes increase whereas, as the number of lymph nodes increases, so does the size (Table 2). Equations 1c and 17 allow us to disentangle the impact of lymph node status and tumor size and to calculate the value of the extra lethality, $L_{per\text{-}node}$, associated with each positive lymph node for patients who have various numbers of positive lymph nodes, *M*.

We can observe the general approach by considering the 92 patients who had melanoma with only 1 positive lymph node (Table 2). This group of patients had a 15-year Kaplan-Meier melanoma death rate of 40%. Equation 1c (see above) tells us that another 92 patients who had tumors of the same thicknesses but who were lymph node negative would have had a 25% death rate (Table 2). Thus, for patients with 1 positive lymph node, the lethal contribution ascribable to cancer in that lymph node, $L_{per\text{-}node}$, is approximately 20% ([40%-25%]/[100%-25%]).

Estimating the value of $L_{per\text{-}node}$ for patients with ≥2 positive lymph nodes requires a more complicated calculation, which can be accomplished with the *Nodal Lethality* equation (Equation 17):

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

For example, for the 36 patients who had melanoma with 2 positive lymph nodes, the lethal contribution per lymph node calculated with Equation 17 was 24% ($L_{per\text{-}node} \approx 0.24$) whereas, for the patients who had 3 positive lymph nodes, $L_{per\text{-}node} \approx 24\%$; for patients with 4 or 5 positive lymph nodes, $L_{per\text{-}node} \approx 23\%$; and, for patients with 6 or 7 positive lymph nodes, $L_{per\text{-}node} \approx 24\%$ (Table 2). Thus, we can conclude that no matter how many positive lymph nodes are identified in a patient

Table 3. The *Size+Nodes* Method for Estimating the Risk of Cancer Death†

Source of Lethality	Method of Estimation	Independent Variable	Parameters	Interpretation
The lethal contribution from cancer at the primary site directly to the periphery	$L_{primary}=1-e^{-(Q^*jprimary)D^Z}$ (Eq. 1c)	D = tumor size: for *breast carcinoma*, greatest dimension in mm; for *melanoma*, thickness in mm	For *breast carcinoma*: $Q = 0.0062$, $Z = 1.34$, and $j_{primary}=0.661$ (if lymph node status is known) or ($j_{primary}=1$ (if lymph node status is unknown); for *melanoma*: $Q = 0.1428$, $Z = 0.89$, and $j_{primary}=0.801$ (if lymph node status is known) or $j_{primary}=1$ (if lymph node status is unknown)	The lethal contribution of the primary mass increases gradually with tumor size
The lethal contribution from cancer in the lymph nodes	$L_{nodes}=1-e^{-(M^*Lper\text{-}node)}$ (Eq. 2)	M = the number of positive lymph nodes	For *breast carcinoma*, $L_{per\text{-}node}=0.0608$; for *melanoma*, $L_{per\text{-}node}=0.22527$	The presence of each positive lymph node contributes approximately "$L_{per\text{-}node}$" extra chance of death

A Simple Mnemonic for the *Size+Nodes* method	Lethal Contribution for Each Millimeter of Primary Tumor Size	+	Lethal Contribution for Each Positive Lymph Node
Risk of breast carcinoma death≈	~1% per mm	+	~6% per lymph node
Risk of melanoma death≈	~8% per mm	+	~23% per lymph node

† When both tumor size and lymph node status are known, the *Size+Nodes* estimates the risk of cancer death from information on tumor size and lymph node status: $L = L_{primary}+L_{nodes}-(L_{primary}{}^*L_{nodes})$ (Eq. 4).

with melanoma, each positive lymph node is associated with an approximately 20% to 25% extra chance of death.

Similarly, when we examine groups of patients with breast carcinoma with various numbers of positive lymph nodes, the *Nodal Lethality* equation (Equation 17) tells us that each positive lymph node is associated with an approximately 6% extra chance of death ($L_{per\text{-}node} \approx 0.06$) (Table 2). This finding agrees with our previously reported observations,[5] which were made by sorting patients according to both tumor size and the number of positive lymph nodes, that each positive lymph node is associated with an approximately 6% extra chance of death.

The values for $L_{per\text{-}node}$ shown in Table 2 reflect subgroups of patients with various numbers of positive lymph nodes. However, we are able to make yet more accurate estimates of the value of $L_{per\text{-}node}$ by using all lymph node-positive patients and applying a pseudo-Monte Carlo method to the *Nodal Lethality* equation (Equation 17), revealing that $L_{per\text{-}node} = 0.0608$ for breast carcinoma and $L_{per\text{-}node} = 0.22,527$ for melanoma.

The Size+Nodes *Method Accurately Predicts the Risk of Cancer Death*

Incorporating the information outlined above for estimating the independent contribution of primary tumor size (Equation 1c) and the number of positive lymph nodes ($L_{per\text{-}node} \approx 6\%$ per positive lymph node for breast carcinoma, $L_{per\text{-}node} \approx 23\%$ for melanoma; Equation 2), provides a technique, the *Size+Nodes* method (Tables 1 and 3), for integrating tumor size and lymph node status into an estimate of the risk of death (L) for each patient:

$$L = L_{primary} + L_{nodes} - (L^*_{primary} L_{nodes}), \qquad (4)$$

in which

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

and

$$L_{nodes} = 1 - e^{-(M^* Lper\text{-}node)}. \qquad (2)$$

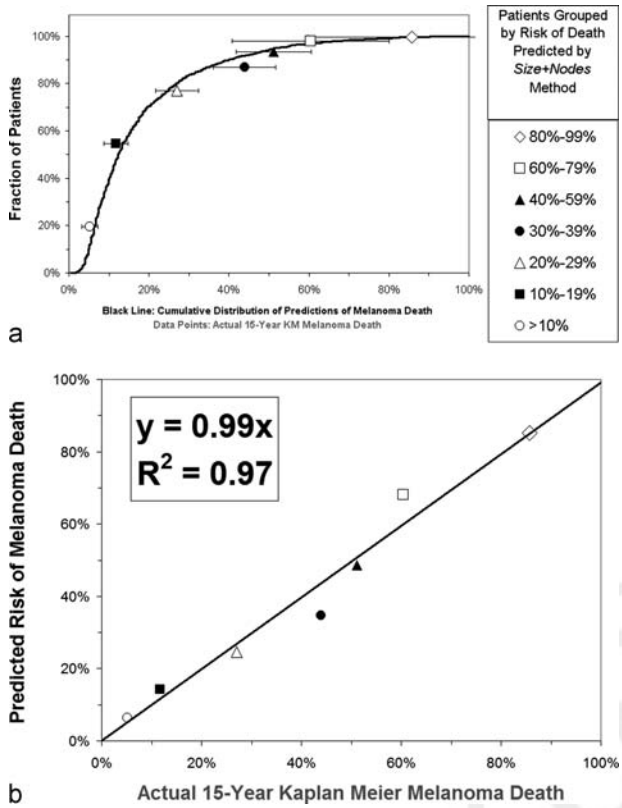

FIGURE 4. These charts illustrate the capacity of the *Size+Nodes* method (Equation 4) to stratify patients with melanoma according to the risk of death. (a) Cumulative distributions of the estimates values of the risk of death were estimated from tumor size and lymph node status information by using the *Size+Nodes* method (Table 1). (b) This scatter plot compares the risk of death estimated from tumor size and lymph node status information using the *Size+Nodes* method with the actual 15-year Kaplan-Meier death rates for 9 groups according to the value of their risk of death estimated from tumor size and the number of positive lymph nodes according to the *Size+Nodes* method. See Table 5 for values. For comparable data in breast carcinoma, see Michaelson et al.[4]

The *Size+Nodes* method can be used to accurately estimate the risk of death for patients with breast carcinoma and can accurately stratify these patients into groups of incrementally increasing lethality, as we reported previously.[5] In contrast, classifying patients with breast carcinoma by lymph node positivity, by tumor classification, or by stage creates groups of women with wide and overlapping levels of lethality.[7] In this study, we demonstrate that this also is true for patients with melanoma (Figs. 4 and 5) (Tables 4 and 5). For example, when we subdivided patients with melanoma into 6 separate groups based on the risk of death predicted by the *Size+Nodes* method (0%-10% estimated risk of death, 10%-19%,

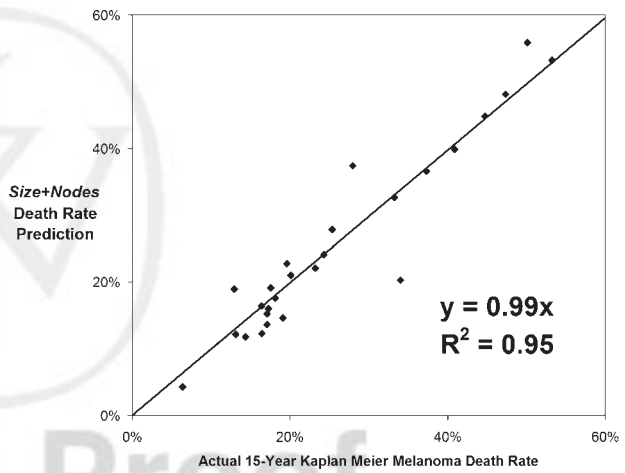

FIGURE 5. This chart illustrates a comparison of the risk of death estimated by using the *Size+Nodes* method and actual death rates among patients sorted in various ways (see Tables 4 and 5).

Table 4. Stratification of Melanoma Patients by Risk of Death Estimated by the *SizeOnly* and *Size+Nodes* Methods

	***SizeOnly* Method**			***Size+Nodes* Method**		
Risk Group, %	Risk of Melanoma Death, %	Actual 15-Year Melanoma Death, %	No. of Patients	Risk of Melanoma Death, %	Actual 15-Year Melanoma Death. %	No. of Patients
0-10	6.42	5.17	1061	6.45	4.65	1084
10-19	14.31	11.71	814	14.19	12.62	859
20-29	24.56	27.02	441	24.77	28.32	366
30-39	34.63	43.88	238	34.89	38.41	179
40-59	48.60	51.20	168	48.61	50.36	191
60-79	68.12	60.42	35	68.16	62.79	59
80-99	85.26	85.71	7	86.72	91.38	26

J_ID: Z7B Customer A_ID: C2696-08 Cadmus Art: CNCR24592 Date: 25-July-09 Stage: Page: 11

Table 5. Comparison of Risk Estimates Made by the *SizeOnly* and *Size+Nodes* Methods to Empirical Death Rates Among Various Groups of Melanoma Patients

		95% CI, %	
		$L_{Predicted}$	
Factor (*N*)	$L_{Empirical}$	*SizeOnly*	*Size+Nodes*
Lymph node status (664)			
Negative (487)	21.02-5.42	24.77-1.06	20.12-0.91
Positive (177)	53.21-9.64	33.29-2.10	53.21-2.67
No. of positive lymph nodes			
1 (92)	39.87-14.71	30.27-2.81	40.91-2.07
1-2 (127)	44.85-12.08	31.04-2.35	44.71-1.99
1-3 (148)	48.14-10.82	31.96-2.15	47.36-2.08
≥2 (85)	66.24-11.92	36.56-3.00	66.81-3.12
≥3 (50)	74.46-13.69	39.03-4.05	75.67-3.23
Clark level (2492)			
2 (773)	4.25-2.20	6.42-0.22	6.41-0.22
3 (655)	11.78-3.30	14.16-0.67	14.34-0.81
4 (964)	27.83-3.94	24.53-0.78	25.32-0.98
5 (100)	55.87-13.74	46.73-3.13	50.14-3.65
Site (2747)			
Trunk (1017)	19.12-3.21	16.36-0.81	17.57-1.01
Face (238)	16.39-7.23	16.72-1.71	16.40-1.90
External ear (71)	14.61-10.52	17.54-2.75	19.09-3.15
Upper limb and shoulder (594)	15.26-3.87	17.59-1.15	17.08-1.28
Lower limb and hip (647)	16.03-3.38	18.36-1.04	17.24-1.27
Scalp and neck (180)	24.12-7.36	22.31-2.45	24.30-3.05
Histology (2742)			
Superficial spreading (1610)	12.17-2.16	12.84-0.47	13.12-0.56
Lentigo malignant (221)	18.92-8.99	13.08-1.56	12.93-1.63
Malignant (453)	22.08-4.36	22.46-1.37	23.23-1.57
Acral lentiginous (68)	37.42-14.30	25.92-3.85	27.96-4.89
Nodular (351)	32.63-6.28	31.79-1.60	33.26-1.91
Desmoplastic (39)	20.27-14.79	35.78-5.31	34.04-5.32
Ulceration (1040)			
Absent (856)	13.63-4.93	17.03-0.85	17.09-0.83
Present (184)	36.53-8.41	34.28-1.17	37.27-2.97
Sex (2762)			
Women (1299)	12.27-2.25	16.33-0.70	16.42-0.77
Men (1463)	22.71-2.89	18.76-0.75	19.61-0.88
All patients (2770)	17.59-1.84	17.59-0.51	18.11-0.59

95% CI indicates 95% confidence interval.

20%-29%, etc) (Table 4), the actual Kaplan-Meier cancer death rate for each group agreed remarkably closely with the risk of death estimated by the *Size+Nodes* method ($R^2 = 0.97$) (Fig. 4). Furthermore, when it was tested on groups of patients with melanoma sorted into 28 categories according to a variety of patient characteristics, including Clark level, sex, body location, histologic subtype, and ulceration, the estimate of the risk of death made by the *Size+Nodes* method agreed within the 95% confidence interval with the actual death rate measured by the Kaplan-Meier method in every case (Table 5). There was a remarkably linear correlation between the estimations of the risk of death made by the *Size+Nodes* method and the actual melanoma death rates for these 28 groups ($R^2 = 0.95$) (Fig. 5).

The Comparative Outcomes of the Size+Nodes *and* SizeOnly *Methods*

When we used it to stratify patients according to the risk of death, the *Size+Nodes* method appeared to be slightly better than the *SizeOnly* method. For example, for melanoma, the *Size+Nodes* method placed 26 patients in the highest risk group and 1084 patients in the lowest risk groups (the group with an 80%-99% estimated risk of death vs the group with a 0%-10% estimated risk of death), whereas the *SizeOnly* method assigned 7 patients and 1061 patients to these groups, respectively (Table 4). However, the most striking advantage of the *Size+Nodes* method can be observed in its superior ability to estimate the risk of death for groups of patients sorted by lymph node status (Table 5). Indeed, compared with the *Size+Nodes* method, the *SizeOnly* method overestimates the chance of death of lymph node-negative patients and underestimates the chance of death for lymph node-positive patients (Table 5).

DISCUSSION

The data presented in the current study provide an integrated method for calculating the impact of the size of the primary tumor and the number of positive lymph nodes on the risk of cancer death. For melanoma, renal cell carcinoma, and breast carcinoma, we have established that the relation between tumor size and the risk of cancer death is captured well by using a simple expression, the *SizeOnly* equation. For melanoma and breast carcinoma, the relation between tumor size and the presence of cancer in the lymph nodes also is captured by a variant of the *SizeOnly* equation, the *NodalSizeOnly* equation. For lymph node-

ID: selvarajn | Black Lining: [ON] | Time: 14:58 | Path: N:/3b2/CNCR/Vol00000/090511/APPFile/C2CNCR090511

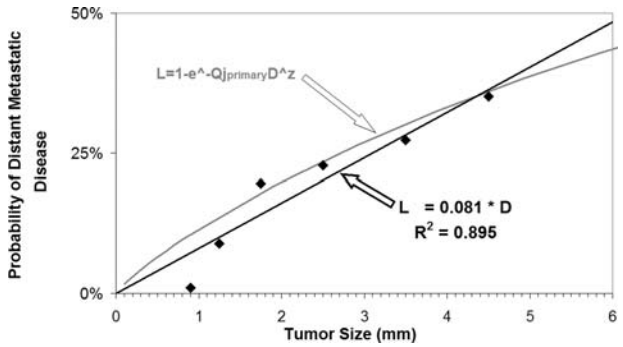

FIGURE 6. This chart illustrates the approximate linearity of the relation between the size of the primary mass (*D*) and the melanoma death rate (*L*) in patients with lymph node-negative disease. Solid diamonds indicate 15-year Kaplan-Meier melanoma death rates for groups of lymph node-negative patients sorted by primary tumor size. Shown is a linear regression of data on cancer death among lymph node-negative melanoma patients sorted by primary tumor size (black line) as well as the fit of Equation 1c to these data (gray line). AQ7 $J_{primary}$ indicates the fraction of events of spread occurring directly from the primary site to the periphery. For comparable data in patients with breast carcinoma, see Michaelson et al.[4]

negative melanoma and breast carcinoma, the relation between tumor size and the risk of death is captured by the *PrimarySizeOnly* equation. For breast carcinoma, we observed that each positive lymph node contributed an approximately 6% risk of death, whereas each millimeter of greatest primary tumor dimension contributed an approximately 1% risk of death. For melanoma, each positive lymph node contributed an approximately 23% risk of death, whereas each millimeter of primary melanoma thickness contributed an approximately 8% F6 risk of death (Fig. 6). This information is captured by a series of linked equations, the *Size+Nodes* method (Table 3).

These observations raise 3 questions: Why do tumor size and lymph node status contribute to the risk of cancer death in this way? How do factors others than tumor size and lymph node status contribute to the risk of cancer death? How can this information be used to provide physicians with practical tools for estimating the risk of cancer death for each patient? We have answered these questions in the 2 accompanying articles in this series.[8,9]

Conflict of Interest Disclosures

Supported by institutional funding of the Massachusetts General Hospital.

References

1. Balch CM, Soong SJ, Gershenwald JE, et al. Prognostic factors analysis of 17,600 melanoma patients: validation of the American Joint Committee on Cancer melanoma staging system. *J Clin Oncol*. 2001;19:3622-3634.
2. White RR, Stanley WE, Johnson JL, Tyler DS, Seigler HF. Long-term survival in 2505 patients with melanoma with regional lymph node metastasis. *Ann Surg*. 2002;235:879-887.
3. Michaelson JS, Silverstein M, Wyatt J, et al. Predicting the survival of patients with breast carcinoma using tumor size. *Cancer*. 2002;95:713-723.
4. Michaelson JS, Satija S, Kopans D, et al. Gauging the impact of breast carcinoma screening in terms of tumor size and death rate. *Cancer*. 2003;98:2114-2124.
5. Michaelson JS, Silverstein M, Sgroi D, et al. The effect of tumor size and lymph node status on breast carcinoma lethality. *Cancer*. 2003;98:2133-2143.
6. Michaelson J. The role of molecular discreteness in normal and cancerous growth. *Anticancer Res*. 1999;19:4853-4867.
7. Michaelson JS, Cheongsiatmoy JA, Dewey F, et al. Spread of human cancer cells occurs with probabilities indicative of a nongenetic mechanism. *Br J Cancer*. 2005;93:1244-1249.
8. Michaelson JS, Chen LL, Silverstein MJ, et al. Why cancer at the primary site and in the lymph nodes contributes to the risk of cancer death. *Cancer*. 2009;115:000-000.
9. Chen LL, Nolan ME, Silverstein MJ, et al. The impact of primary tumor size, lymph node status, and other prognostic factors on the risk of cancer death. *Cancer*. 2009;115:000-000. AQ4
10. Delahunt B, Kittelson JM, McCredie MR, Reeve AE, Stewart JH, Bilous AM. Prognostic importance of tumor size for localized conventional (clear cell) renal cell carcinoma: assessment of TNM T1 and T2 tumor categories and comparison with other prognostic parameters. *Cancer*. 2002;94:658-664.
11. Hafez KS, Fergany AF, Novick AC. Nephron sparing surgery for localized renal cell carcinoma: impact of tumor size on patient survival, tumor recurrence and TNM staging. *J Urol*. 1999;162:1930-1933.
12. Karrison TG, Ferguson DJ, Meier P. Dormancy of mammary carcinoma after mastectomy. *J Natl Cancer Inst*. 1999;91:80-85.
13. 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.
14. 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.

15. Pesce C, Colacino R. Morphometry of the breast fibroadenoma. *Pathol Res Pract*. 1986;181:718-720.
16. Twombly R. Breast cancer gene microarrays pass muster. *J Natl Cancer Inst*. 2006;98:1438-1440.
17. Buyse M, Loi S, van't Veer L, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst*. 2006;98:1183-1192.
18. Simon R. Development and evaluation of therapeutically relevant predictive classifiers using gene expression profiling. *J Natl Cancer Inst*. 2006;98:1169-1171.
19. Fan C, Oh DS, Wessels L, et al. Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med*. 2006;355:560-569.
20. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet*. 2005;365:1687-1717.

0000

How Cancer at the Primary Site and in the Lymph Nodes Contributes to the Risk of Cancer Death

James S. Michaelson, L. Leon Chen, Melvin J. Silverstein, Martin C. Mihm Jr, Arthur J. Sober, Kenneth K. Tanabe, Barbara L. Smith, and Jerry Younger

In this report, the authors have described a mathematical method, the binary-biological model of cancer metastasis, based on the spread of cancer cells, in which the equations capture the relations between tumor size, lymph node status, and cancer lethality.

AQ1: Note that Reference 20 was renumbered Reference 9, in the order it was cited in the text, and the remaining citations and listed references were renumbered accordingly.

AQ2: Please indicate the most recent date you viewed the Cancer Math web site.

AQ3: Note that Tables 2 and 3 were renumbered Tables 3 and 2, respectively, in the order they were mentioned in the text.

AQ4: Note that the third article in the series (originally Reference 20) was renumbered Reference 9, in the order it was cited in the text, and the remaining list and corresponding citations were renumbered accordingly: Please check the list and citations carefully.

AQ5: In the Figure 1 legend, please define the terms "QD" and "QD^{Z}".

AQ6: In the Figure 2 legend, please verify that *Jprimary* is defined correctly and also define the terms e^, Q, and ^Z.

AQ7: In the Figure 6 legend, please verify that *Jprimary* is defined correctly and also define the terms e^, Q, and ^Z.

AQ8: Please check all tables carefully, including titles, headings, data, data alignment, abbreviations, and footnotes, and mark your corrections where needed.

AQ9: Note that the original Tables 2 and 3 were renumbered Tables 3 and 2, respectively, in the order they were mentioned in the text.