

Laboratory for Quantitative Medicine
Technical Report #2
April 10, 2009

Equation Parameters

This report summarizes all the equation parameters used by the web-based calculators, and provides information on the methodologies by which these parameters were derived.

L. Leon Chen, Raymond A. Jean, Matthew E. Nolan, James S. Michaelson,
PhD^{1,2,3}

Departments of Pathology¹ and Surgery², Massachusetts General Hospital and the Department of Pathology³, 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

The manuscripts available on our site are provided for your personal use only and may not be re-transmitted or redistributed without written permissions from the paper's publisher. You may quote this report as personal communication, after notifying Dr. James Michaelson PhD (michaelj@helix.mgh.harvard.edu). You may not upload any of this site's material to any public server, on-line service, or bulletin board without prior written permission from the publisher and authors. You may not make copies for any commercial purpose. Reproduction of materials retrieved from this web site is subject to the U.S. Copyright Act of 1976, Title 17 U.S.C.

I. Equations

For discussions on the theoretical framework, derivations, and other details of these equations, please see technical report #1, Mathematical Methods.

L_t is the lethality at time t . Herein $t = 15$ years, and all parameters are derived based upon this. D , the diameter of the tumor, may have an exponent Z (D^Z). For breast cancer, $Z = 1$, and Q becomes Q_s , where s stands for simplified. All tumor sizes are considered in millimeters (mm).

SizeOnly (S)	$L_t = 1 - e^{-Q_s \cdot D}$
	$L_t = L_{primary} + L_{nodes} - (L_{primary} \cdot L_{nodes})$
Size+Nodes (SN)	where $L_{primary} = 1 - e^{-(Q_s \cdot J_{primary}) \cdot D}$ $L_{nodes} = 1 - e^{-R \cdot M}$
Size+PrognosticFactors (SP)	$L_t = 1 - e^{-(Q_s \cdot (g1 \cdot g2 \cdot g3 \cdot \dots \cdot g_x)) \cdot D}$
	$L_t = L_{primary} + L_{nodes} - (L_{primary} \cdot L_{nodes})$
Size+Nodes+PrognosticFactors (SNP)	where $L_{primary} = 1 - e^{-(Q_s \cdot J_{primary} \cdot (g1 \cdot g2 \cdot g3 \cdot \dots \cdot g_x)) \cdot D}$ $L_{nodes} = 1 - e^{-R \cdot M}$

The following *nodal* equations estimate the chance of having positive nodes, given the size of the tumor and possibly certain prognostic factors.

Nodal SizeOnly (Sn)	$P_n = 1 - e^{-Q_{sn} \cdot D}$
Nodal Size+PrognosticFactors (SPn)	$P_n = 1 - e^{-(Q_{sn} \cdot (gn1 \cdot gn2 \cdot gn3 \cdot \dots \cdot gn_x)) \cdot D}$

II. Main Parameters

Breast Cancer

Parameters are derived from the U.S. Surveillance, Epidemiology, and End Results (SEER) breast cancer database (details in appendix A). For breast cancer, the population is restricted to those having invasive tumors 1-5 cm in size, with 0-7 positive lymph nodes. For details on the methodology used to derive these parameters, please see appendix B.

Parameter	Value	Notes
Qs	0.0118395 (1973-2004)	
	0.014751 (1973-1982)	used in treatment calc.
	0.010054 (1988-2004)	used in outcome calc.
Jprimary	0.8057	
R	0.07581	

Melanoma

Parameters are derived from the U.S. Surveillance, Epidemiology, and End Results (SEER) melanoma database (details in appendix A). The subset of patients with known tumor size (90,771) and the subset with known number of nodes (20,205) were used. For details on the methodology used to derive these parameters, please see appendix B.

Parameter	Value
Qs	0.1455
Jprimary	0.801
Z	1
R	0.22527

Renal Cell Carcinoma

Parameters are derived from the following two papers in the literature:

1. 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*. 94:658-64 2002
2. 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. *Journal of Urology*. 162(11):1930-6 1999

Parameter	Value
Q	0.0033
Z	1.1392

III. g Parameters

Breast Cancer

These g-parameters were derived using the main parameters values above, with the exception of ER/PR status, which is derived using $Q = 0.0101485$ for reasons expounded upon in appendix B. The g-parameters of all considered prognostic factors are listed below. Note: a prognostic factor that is unknown will correspond to a g-parameter of 1.

	Prognostic factor	g-parameter
Age	21-30	1.25449
	31-40	1.12668
	41-50	0.86605
	51-60	1.01903
	61-70	1.01715
	71-80	1.02005
	81-90	1.16455
Grade	1 – well differentiated	0.4324
	2 – moderately differentiated	0.857
	3 – poorly differentiated	1.1224
	4 - undifferentiated	1.26495
Estrogen Receptor / Progesterone Receptor Status	ER+ / PR+	0.91545
	ER+ / PR-	1.13885
	ER- / PR+	1.0462
	ER- / PR-	1.1902
HER2 Status*	HER2+	1.515
	HER2-	0.9662
Histological Type	Ductal	1.0573
	Lobular	0.9032
	Ductal and lobular	0.85725
	Mucinous	0.46458
	Medullary	0.59946
	Tubular	0.2752
	Comedo	0.86451
	Scirrhous	1.63135
	Inflammatory	3.313
	Paget's disease	1.4535
	Papillary	0.54138
	Cribriform	0.9636

*the g-parameters for HER2 are derived instead from the Partners Breast Cancer Database, as information on HER2 is nonexistent in the SEER database. We recognize that this approach is not ideal, but it is supported by the fact that all other g-parameters remain remarkably consistent across these two datasets (see appendix C).

Melanoma

These g-parameters were derived using the main melanoma parameters. Note: a prognostic factor that is unknown will correspond to a g-parameter of 1. N.S. = not statistically significant.

Prognostic factor		g-parameter
Gender	Male	1.188
	Female	0.770
Age	1-20	0.414
	21-30	0.630
	31-40	0.720
	41-50	0.885
	51-60	1 (N.S.)
	61-70	1.195
	71-80	1.378
	81-90	1.593
	>90	1 (N.S.)
	Clark Level (with Ulceration status)	I (not ulcerated)
II (not ulcerated)		0.412
II (ulcerated)		1 (N.S.)
III (not ulcerated)		0.787
III (ulcerated)		1.375
IV (not ulcerated)		1 (N.S.)
IV (ulcerated)		1.348
V (not ulcerated)		1.256
V (ulcerated)		1 (N.S.)
Histology		Desmoplastic
	Nodular	1.335
	Acral Lentiginous	1.773
	Superficial Spreading	0.842
	Lentigo Maligna	1 (N.S.)
	Spindle Cell	1 (N.S.)
	Epithelioid Cell	1 (N.S.)
	Epithelioid, Spindle (Mixed)	0.471
	Regressing	1 (N.S.)
	Giant Pigmented Nevus	1 (N.S.)
Location	External Ear	1 (N.S.)
	Face	1 (N.S.)
	Scalp or Neck	1.689
	Trunk	1.088
	Upper Limb or Shoulder	0.784
	Lower Limb or Hip	0.889

Renal Cell Carcinoma

There are currently no g-parameters in the renal cell carcinoma model.

IV. Nodal Parameters

All nodal parameters were derived from the same sources as the main parameters and g-parameters. However, these parameters were derived using only the subset of patients that had nodal status information. Appendix B contains more details on the methodology. Q_n is Q_{sn} for $Z = 1$, where s stands for simplified, and n stands for nodal.

Breast Cancer

Parameter	Value
Q_{sn}	0.017446

Prognostic factor	gn-parameter	
Age	21-30	1.21965
	31-40	1.26315
	41-50	1.1953
	51-60	1.0832
	61-70	0.88798
	71-80	0.7962
	81-90	0.78902
Grade	1 – well differentiated	0.8244
	2 – moderately differentiated	1.1723
	3 – poorly differentiated	1.22467
	4 - undifferentiated	1.0725
Estrogen Receptor / Progesterone Receptor Status	ER+ / PR+	1.03025
	ER+ / PR-	1.00501
	ER- / PR+	1.02325
	ER- / PR-	0.90615
Histological Type	Ductal	1.0639
	Lobular	0.87525
	Ductal and lobular	1.30763
	Mucinous	0.27398
	Medullary	0.64677
	Tubular	0.41965
	Comedo	0.72415
	Scirrhous	0.90837
	Inflammatory	2.61032
	Paget's disease	1.44607
	Papillary	0.41083
	Cribriform	0.9636

	Apocrine	1.1164
	Phyllodes	0.0097

Melanoma

Parameter	Value
Qn	0.1186
Zn	1

	Prognostic factor	gn-parameter
Age	1-20	1.511
	21-30	1.331
	31-40	1.283
	41-50	1.119
	51-60	1 (N.S.)
	61-70	0.902
	71-80	0.761
	81-90	0.816
	>90	1 (N.S.)
Clark Level (with Ulceration status)	I (not ulcerated)	0.017
	II (not ulcerated)	0.663
	II (ulcerated)	1 (N.S.)
	III (not ulcerated)	0.691
	III (ulcerated)	1 (N.S.)
	IV (not ulcerated)	0.943
	IV (ulcerated)	1 (N.S.)
	V (not ulcerated)	1 (N.S.)
	V (ulcerated)	1.632
Histology	Desmoplastic	0.202
	Nodular	1.178
	Acral Lentiginous	1 (N.S.)
	Superficial Spreading	1 (N.S.)
	Lentigo Maligna	0.329
	Spindle Cell	0.549
	Epithelioid Cell	1 (N.S.)
	Epithelioid, Spindle (Mixed)	1 (N.S.)
	Regressing	1 (N.S.)
	Giant Pigmented Nevus	1 (N.S.)
Location	External Ear	0.707
	Face	0.655
	Scalp or Neck	1 (N.S.)
	Trunk	1.243
	Upper Limb or Shoulder	0.775
	Lower Limb or Hip	1.222

Renal Cell Carcinoma

There is currently no model of spread to the lymph nodes for renal cell carcinoma.

V. APPENDIX A – Data Sources

SEER Dataset (Breast Cancer)

Breast cancer and melanoma data was extracted from the US Surveillance, Epidemiology, and End Results (SEER) dataset, which is provided by the National Cancer Institute. It consists of 17 regional registries across the US, with data from 1973 to 2005. An extensive effort was made to clean the data, for example by considering only first malignant primary tumors. The total number of patients for breast cancer is 508,861, and the total number for melanoma is 251,083.

Partners Healthcare Breast Cancer Database

This database consists of 24,771 breast cancer patients diagnosed at the Massachusetts General Hospital or the Brigham and Women's hospital between 1968 and 2007.

MGH Melanoma Database

This database consists of 2,770 melanoma patients seen at the Massachusetts General Hospital between 1/1/1970 and 5/1/2002.

VI. APPENDIX B – Methodology

Q or Qs:

1. Consider the population of cancer patients with tumor size, follow-up time, and cancer-specific death information.
2. Calculate the Kaplan-Meier cancer-specific mortality rate at 15 years, which we call L15 (empirical).
3. Calculate the value of Q such that the error between L15 (empirical) and the mean of L15 (predicted) is minimized. L15 (predicted) is the predicted lethality of each patient in the cohort using the SizeOnly equation.

Jprimary:

1. Limit the population of cancer patients to only those that are known to be node negative.
2. Calculate the Kaplan-Meier cancer-specific mortality rate at 15 years, which we call L15 (empirical).
3. Calculate the value of Jprimary such that the error between L15 (empirical) and the mean of L15 (predicted) is minimized. L15 (predicted) is the predicted lethality of each patient in this cohort using the Size+Nodes equation with $R = 0$.

R:

1. Limit the population of cancer patients to only those whose nodal status is known, and in the case of node positive, the number of positive nodes.
2. Calculate the Kaplan-Meier cancer-specific mortality rate at 15 years, which we call L15 (empirical).
3. Calculate the value of R such that the error between L15 (empirical) and the mean of L15 (predicted) is minimized. L15 (predicted) is the predicted lethality of each patient in this cohort using the Size+Nodes equation with M being the number of positive nodes for each patient.

VII. APPENDIX C – Comparison of g-parameters between two datasets

To demonstrate the flexibility and robustness of this approach, g-parameters derived from the Partners dataset (appendix A) and g-parameters derived from the SEER dataset (appendix A) are compared against one another. The table shows that though there are slight variations in the g-parameters across the two datasets, overall they tend to be similar in terms of overall trends. g-parameters corresponding to ER and PR status are shown separately to illustrate this.

Prognostic factor	g-parameter		
	SEER dataset	Partners dataset	
Age	21-30	1.25449	-
	31-40	1.12668	1.0196
	41-50	0.86605	0.8379
	51-60	1.01903	0.971
	61-70	1.01715	1.0555
	71-80	1.02005	1.1518
	81-90	1.16455	-
Grade	1	0.432	0.2735
	2	0.857	0.8563
	3	1.122	1.0625
	Undifferentiated	1.265	2.25
Estrogen Receptor Status	ER+	0.9482	0.798
	ER-	1.1808	1.149
Progesterone Receptor Status	PR+	0.91805	0.6795
	PR-	1.1654	1.557
Histological Type	Ductal	1.0573	1.0405
	Lobular	0.9032	0.9378
	Ductal and lobular	0.85725	0.6651
	Mucinous	0.46458	0.7615
	Medullary	0.59946	0.893
	Tubular	0.2752	0.12402
	Comedo	0.86451	1.102
	Scirrhus	1.63135	-
Inflammatory	3.313	1.356	

Paget's disease	1.4535	1.0205
Papillary	0.54138	-
Cribriform	0.9636	-
