

Laboratory for Quantitative Medicine
Technical Report #3
June 24, 2008

Validation: SizeOnly Equation

This report presents results demonstrating the validity and accuracy of the SizeOnly equation.

L. Leon Chen, 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. Introduction

This report summarizes all verification and validation studies that have been performed on the SizeOnly equation, for breast cancer. The data is grouped and permuted in multiple ways, demonstrating the result that the SizeOnly equation even by itself is quite accurate.

The parameter of Qs (refer to Technical Report #2: Equation Parameters) was derived using the SEER dataset. Section III verifies that this derived value indeed provides accurate results over a wide range of data stratification strategies. Section IV then validates this in an entirely separate dataset, using the Partners Healthcare Breast Cancer Database.

Accuracy can be determined in a number of ways, but one of the most straightforward is plotting predicted vs. empirical results, and seeing how close it fits the $y = x$ line. This line, with a slope of 1, represents a one-to-one correspondence between what was predicted and what actually happened. Therefore, we can fit the graph to a line, and measure the slope of that line. A slope of greater than 1 tends toward over-prediction, whereas a slope of less than 1 tends toward under-prediction. We seek both a slope of 1 as well as a R^2 value of 1.

The graphs display visually what ranges of values are closest to the $y = x$ line, and thus give the most accurate results. For most cases, the highest accuracy is achieved at lower lethality values, and becomes less accurate towards higher lethality values. However, this is in part due to the natural variations that are present in the data itself.

II. Datasets

SEER Dataset (Breast Cancer)

Breast cancer 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 2004. An extensive effort was made to clean the data, for example by considering only first malignant primary tumors. Also, analysis is restricted to patients with 1-50 mm tumors and 0-7 positive lymph nodes. These patients total 362,491.

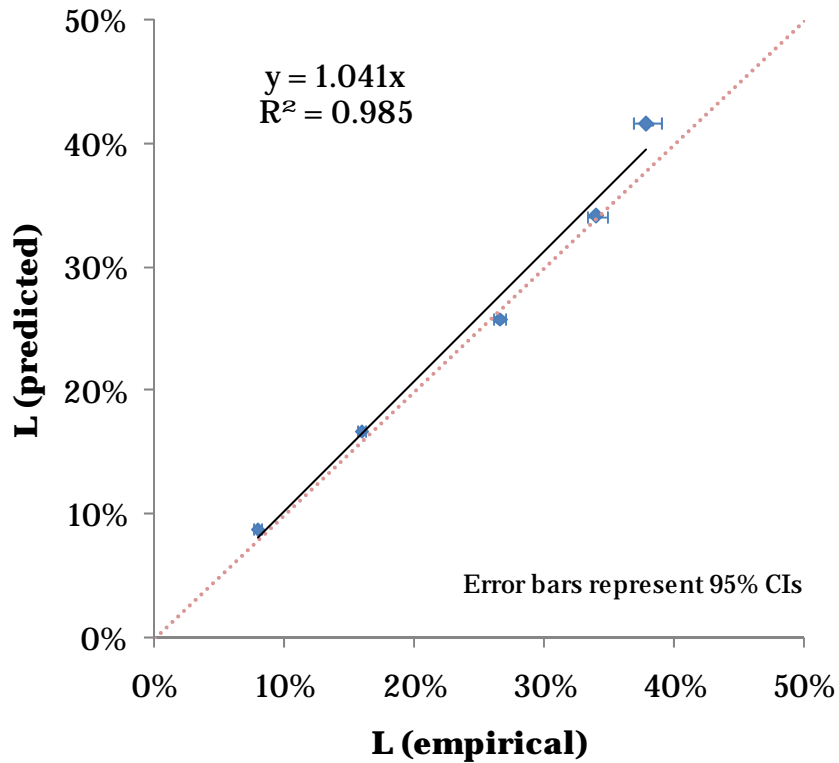
Partners Healthcare Breast Cancer Database

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. Analysis is restricted to patients with 1-50 mm tumors and 0-7 positive lymph nodes, a subset with 7,907 patients.

III. Verification on the SEER dataset

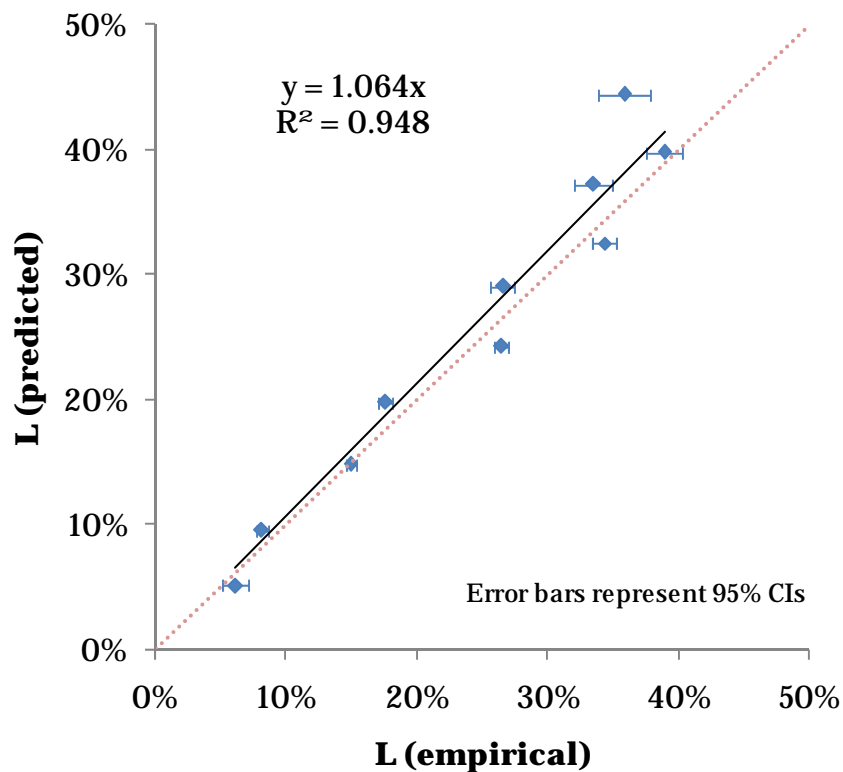
Grouped by 10 mm tumor size bins

Group	N	L empirical (SEM)	L predicted (SEM)	Difference (pred – emp)
1-10 mm	79542	7.89% (0.19%)	8.84% (0.01%)	0.96%
11-20 mm	153328	15.92% (0.16%)	16.80% (0.01%)	0.88%
21-30 mm	81722	26.57% (0.24%)	25.85% (0.01%)	-0.72%
31-40 mm	32763	34.02% (0.39%)	34.17% (0.01%)	0.15%
41-50 mm	15136	37.88% (0.58%)	41.65% (0.02%)	3.76%
<i>Mean (std. dev.)</i>				1.01% (1.68%)
<i>Mean weighted by N (std. dev.)</i>				0.59% (1.01%)
<i>Root Mean Square (std. dev.)</i>				1.81% (2.47%)
<i>Root Mean Square weighted by N (std. dev.)</i>				1.08% (1.16%)



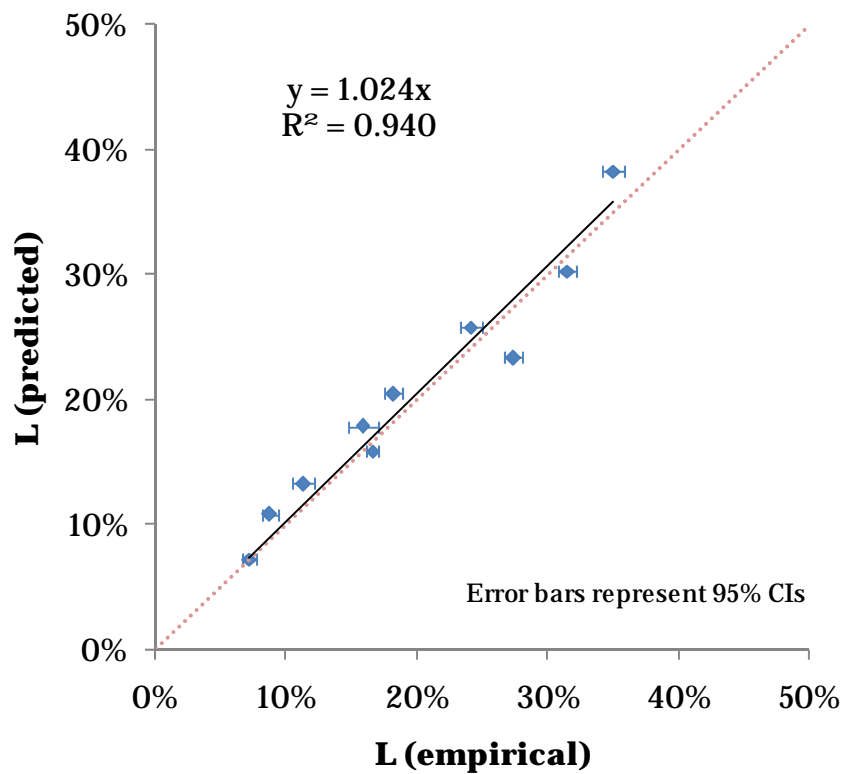
Grouped by 5 mm tumor size bins

Group	N	L empirical (SEM)	L predicted (SEM)	Difference (pred – emp)
1-5 mm	12919	6.08% (0.49%)	5.11% (0.01%)	-0.97%
6-10 mm	66623	8.15% (0.21%)	9.57% (0.01%)	1.42%
11-15 mm	91912	14.89% (0.20%)	14.84% (0.00%)	-0.06%
16-20 mm	61416	17.56% (0.29%)	19.75% (0.01%)	2.19%
21-25 mm	54761	26.43% (0.29%)	24.28% (0.01%)	-2.16%
26-30 mm	26961	26.54% (0.46%)	29.05% (0.01%)	2.51%
31-35 mm	21281	34.31% (0.45%)	32.52% (0.01%)	-1.78%
36-40 mm	11482	33.44% (0.77%)	37.23% (0.01%)	3.79%
41-45 mm	8985	38.90% (0.70%)	39.76% (0.01%)	0.85%
46-50 mm	6151	35.81% (1.04%)	44.41% (0.01%)	8.60%
<i>Mean (std. dev.)</i>				1.44% (3.17%)
<i>Mean weighted by N (std. dev.)</i>				0.63% (1.99%)
<i>Root Mean Square (std. dev.)</i>				3.33% (4.47%)
<i>Root Mean Square weighted by N (std. dev.)</i>				1.99% (2.21%)



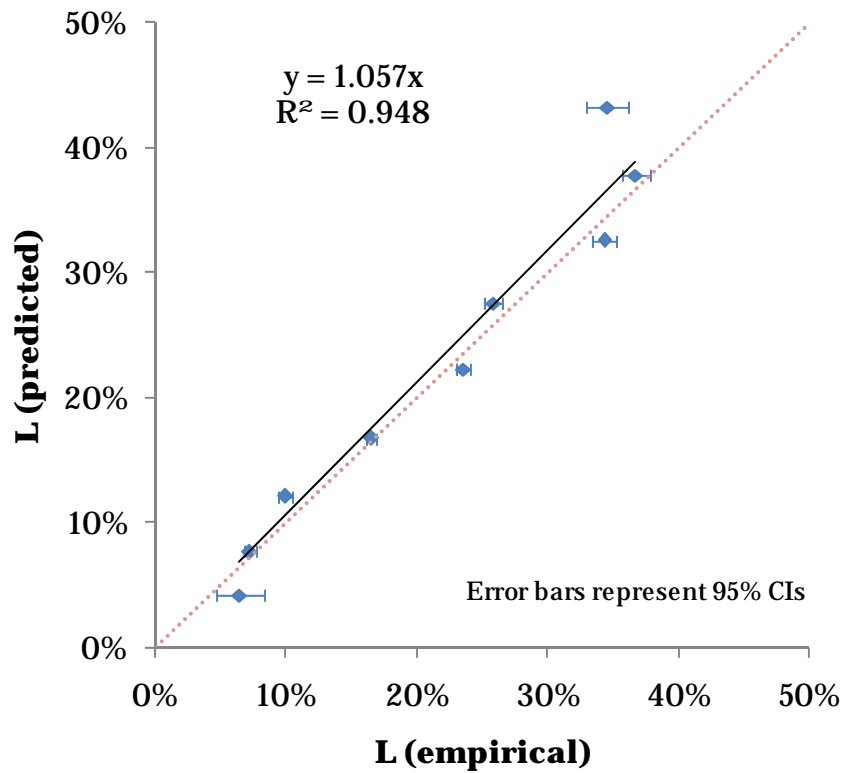
Grouped by 10% tumor size percentiles

Group	N	L empirical (SEM)	L predicted (SEM)	Difference (pred – emp)
1-8 mm	44510	7.16% (0.25%)	7.26% (0.01%)	0.09%
9-10 mm	35032	8.76% (0.30%)	10.86% (0.00%)	2.10%
11-13 mm	36804	11.29% (0.39%)	13.28% (0.00%)	2.00%
14-15 mm	55108	16.59% (0.24%)	15.87% (0.00%)	-0.72%
16-17 mm	16894	15.81% (0.58%)	17.82% (0.00%)	2.01%
18-20 mm	44522	18.15% (0.33%)	20.48% (0.00%)	2.33%
21-23 mm	30854	27.40% (0.35%)	23.34% (0.00%)	-4.06%
24-27 mm	29468	24.07% (0.44%)	25.78% (0.00%)	1.71%
28-33 mm	33673	31.44% (0.37%)	30.29% (0.01%)	-1.15%
34-50 mm	35626	34.96% (0.41%)	38.25% (0.02%)	3.30%
<i>Mean (std. dev.)</i>				0.76% (2.22%)
<i>Mean weighted by N (std. dev.)</i>				0.70% (2.09%)
<i>Root Mean Square (std. dev.)</i>				2.24% (2.25%)
<i>Root Mean Square weighted by N (std. dev.)</i>				2.10% (2.08%)



Grouped by 5% predicted lethality bins using the SizeOnly equation

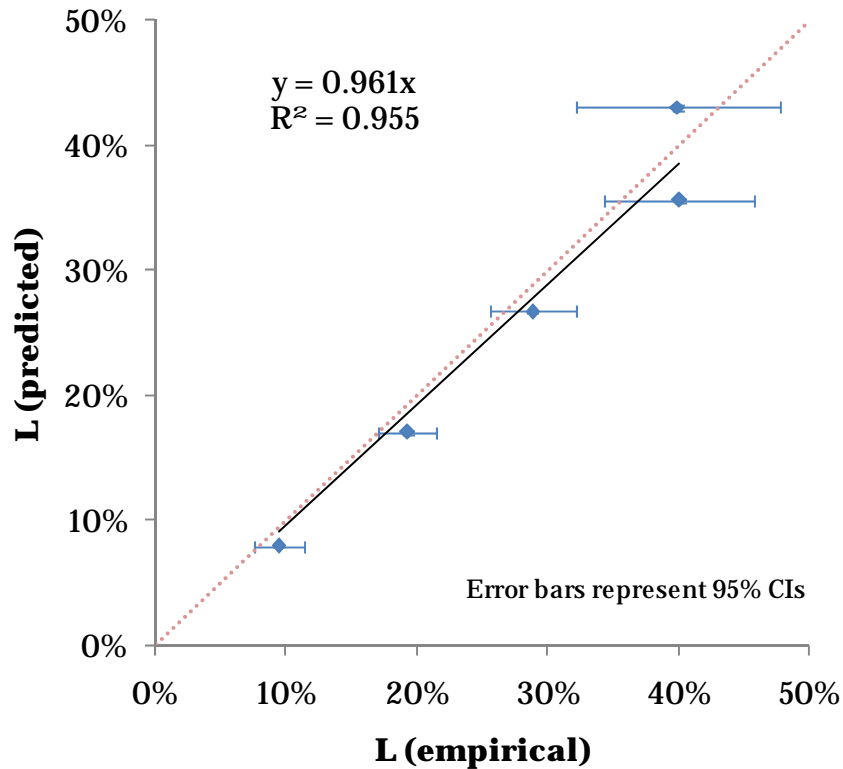
Group	N	L empirical (SEM)	L predicted (SEM)	Difference (pred – emp)
0-5%	5243	6.45% (0.92%)	4.18% (0.01%)	-2.27%
6-10%	39267	7.22% (0.26%)	7.67% (0.01%)	0.44%
11-15%	71836	9.92% (0.24%)	12.10% (0.01%)	2.18%
16-20%	84047	16.49% (0.21%)	16.74% (0.00%)	0.25%
21-25%	66790	23.52% (0.25%)	22.25% (0.01%)	-1.27%
26-30%	47409	25.80% (0.35%)	27.57% (0.01%)	1.77%
31-35%	21849	34.28% (0.45%)	32.58% (0.01%)	-1.69%
36-40%	16110	36.70% (0.56%)	37.78% (0.01%)	1.07%
41-45%	9940	34.52% (0.83%)	43.21% (0.02%)	8.70%
<i>Mean (std. dev.)</i>				1.02% (3.26%)
<i>Mean weighted by N (std. dev.)</i>				0.69% (1.80%)
<i>Root Mean Square (std. dev.)</i>				3.24% (4.95%)
<i>Root Mean Square weighted by N (std. dev.)</i>				1.83% (2.20%)



IV. Validation on the Partners dataset

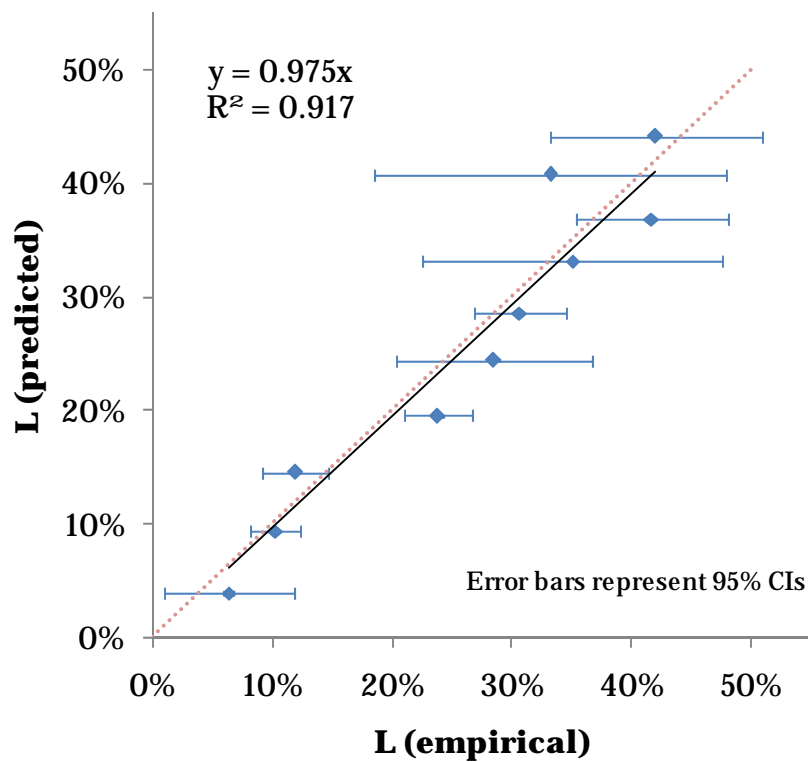
Grouped by 10 mm tumor size bins

Group	N	L empirical (SEM)	L predicted (SEM)	Difference (pred – emp)
1-10 mm	2591	9.41% (0.96%)	7.95% (0.06%)	-1.46%
11-20 mm	3004	19.29% (1.13%)	17.03% (0.05%)	-2.26%
21-30 mm	1499	28.83% (1.68%)	26.72% (0.06%)	-2.11%
31-40 mm	569	40.07% (2.92%)	35.60% (0.08%)	-4.47%
41-50 mm	244	39.91% (3.98%)	43.06% (0.11%)	3.15%
<i>Mean (std. dev.)</i>				-1.43% (2.80%)
<i>Mean weighted by N (std. dev.)</i>				-1.96% (1.71%)
<i>Root Mean Square (std. dev.)</i>				2.88% (2.67%)
<i>Root Mean Square weighted by N (std. dev.)</i>				2.49% (2.67%)



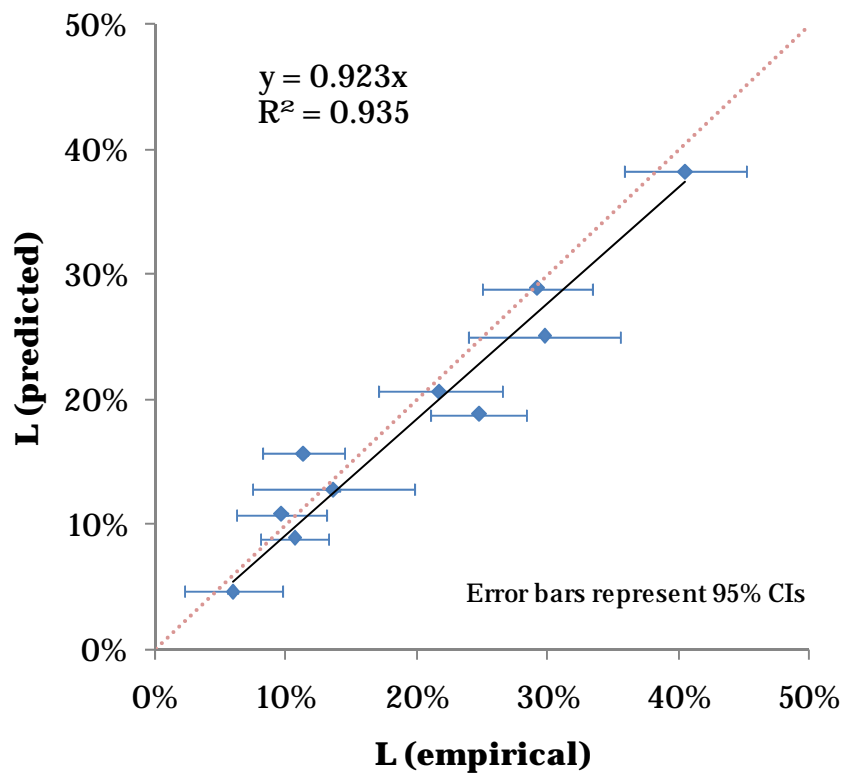
Grouped by 5 mm tumor size bins

Group	N	L empirical (SEM)	L predicted (SEM)	Difference (pred – emp)
1-5 mm	678	6.35% (2.78%)	3.92% (0.06%)	-2.44%
6-10 mm	1913	10.29% (1.06%)	9.38% (0.03%)	-0.92%
11-15 mm	1516	11.88% (1.43%)	14.56% (0.04%)	2.68%
16-20 mm	1488	23.82% (1.48%)	19.53% (0.03%)	-4.28%
21-25 mm	654	28.51% (4.18%)	24.38% (0.05%)	-4.13%
26-30 mm	845	30.66% (1.97%)	28.54% (0.03%)	-2.12%
31-35 mm	189	35.03% (6.37%)	33.13% (0.08%)	-1.90%
36-40 mm	380	41.65% (3.25%)	36.83% (0.04%)	-4.82%
41-45 mm	76	33.20% (7.53%)	40.74% (0.10%)	7.53%
46-50 mm	168	42.04% (4.49%)	44.11% (0.04%)	2.07%
<i>Mean (std. dev.)</i>				-0.83% (3.87%)
<i>Mean weighted by N (std. dev.)</i>				-1.45% (3.38%)
<i>Root Mean Square (std. dev.)</i>				3.76% (4.09%)
<i>Root Mean Square weighted by N (std. dev.)</i>				3.52% (4.48%)



Grouped by 10% tumor size percentiles

Group	N	L empirical (SEM)	L predicted (SEM)	Difference (pred – emp)
1.00-6.00 mm	901	5.88% (1.91%)	4.65% (0.06%)	-1.24%
6.01-8.17 mm	979	10.61% (1.31%)	8.88% (0.02%)	-1.72%
8.18-10.00 mm	711	9.61% (1.74%)	10.85% (0.02%)	1.24%
10.01-12.00 mm	580	13.64% (3.15%)	12.80% (0.02%)	-0.84%
12.01-15.00 mm	936	11.28% (1.57%)	15.66% (0.03%)	4.37%
15.01-18.07 mm	909	24.71% (1.86%)	18.84% (0.02%)	-5.87%
18.08-20.00 mm	579	21.75% (2.39%)	20.62% (0.03%)	-1.13%
20.01-27.78 mm	816	29.74% (2.96%)	25.02% (0.06%)	-4.72%
27.79-32.00 mm	727	29.16% (2.12%)	28.92% (0.04%)	-0.24%
32.01-50.00 mm	769	40.46% (2.40%)	38.21% (0.13%)	-2.25%
<i>Mean (std. dev.)</i>				-1.24% (2.86%)
<i>Mean weighted by N (std. dev.)</i>				-1.27% (3.22%)
<i>Root Mean Square (std. dev.)</i>				2.98% (3.47%)
<i>Root Mean Square weighted by N (std. dev.)</i>				3.31% (3.95%)



Grouped by 10% predicted lethality bins using the SizeOnly equation

Group	N	L empirical (SEM)	L predicted (SEM)	Difference (pred – emp)
0-10%	1880	9.04% (1.07%)	6.85% (0.06%)	-2.19%
11-20%	3244	17.10% (1.07%)	15.11% (0.06%)	-2.00%
21-30%	1970	27.47% (1.49%)	25.33% (0.07%)	-2.13%
31-40%	589	39.61% (2.89%)	35.73% (0.08%)	-3.88%
41-50%	224	41.37% (4.10%)	43.39% (0.09%)	2.02%
<i>Mean (std. dev.)</i>				-1.64% (2.18%)
<i>Mean weighted by N (std. dev.)</i>				-2.10% (1.64%)
<i>Root Mean Square (std. dev.)</i>				2.55% (2.19%)
<i>Root Mean Square weighted by N (std. dev.)</i>				2.56% (2.54%)

