Laboratory for Quantitative Medicine

Technical Report #10

March 22, 2010

Technical report for the paper:

"CancerMath.net: Web-Based Calculators for Breast Carcinoma"

Running title: CancerMath.net Calculators

L. Leon Chen[1], Devon Bush[1], Alan Fong, Barbara Smith MD PhD[1, 4], Kenneth K Tanabe MD[1, 4], Jerry Younger MD[3, 5], James S. Michaelson PhD[1, 2, 6]

Departments of Surgery[1] , Pathology[2], Medicine[3] Massachusetts General Hospital, Boston, Massachusetts, USA
Departments Surgery [4], Medicine[5], Pathology[6], Harvard Medical School, Boston, Massachusetts, USA

Correspondence to James S. Michaelson Ph.D., Division of Surgical Oncology, Massachusetts General Hospital, 2nd Floor, 65 Landsdowne St, Cambridge, Massachusetts, 02139
TEL 617-501-0590 FAX 617-724-3895
Email: michaelj@helix.mgh.harvard.edu

Citation: Technical Report #10 - Technical report for the paper: "CancerMath.net: Wed-based Calculators for Breast Carcinoma" (July 28, 2009, updated November 28, 2009 and March 22, 2010) at: http://cancer.lifemath.net/about/techreports/index.php

ABSTRACT

PURPOSE: Predicting the risk of breast carcinoma death is both a challenge, and an essential requirement for arriving at the best treatment for each patient.

METHODS: CancerMath.net calculators were created with HTML, JavaScript, PHP, using the XML/SWF Charts v5.07 package and Adobe Flash to animate and display the graphs. The risk of cancer death was calculated by the *SNAP* (*Size+Nodes+PrognosticFactor*) method of the *binary biological model of cancer metastasis*, executed in Javascript, from information on tumor size, nodal status, and other prognostic factors. Accuracy was tested against two large breast carcinoma datasets: 7,907 patients seen at two academic hospitals and 362,491 patients from the SEER national dataset.

RESULTS: We describe a series of web-based calculators available at http://www.CancerMath.net: 1) An *outcome calculator*, which calculates expected survival information at the time of diagnosis given the current standard of care treatment; 2) a *conditional survival calculator*, which calculates expected survival information at various times after the time of diagnosis given the current standard of care treatment; 3) a *treatment calculator*, which calculates expected survival information at the time of diagnosis with and without a range of adjuvant therapy options; 4) A *nodal status calculator*, which calculates chance that cancer will be present in the nodes. The CancerMath calculators also provide the patient's classification (T, N, and M) and stage. The CancerMath calculators were found to be highly accurate and specific, as seen by their capacity for stratifying patients into groups differing by as little as a 2% risk of death

CONCLUSIONS: The CancerMath.net breast carcinoma calculators provide accurate and useful estimates of the risk of death, which can aid in a analysis of the various adjuvant therapy options available to each patient.

INTRODUCTION

Predicting the risk of breast carcinoma death is both a challenge, and an essential requirement for arriving at the best treatment for each patient[1-4]. Physicians regularly make such estimates by staging, experience, and intuition. Estimates of breast carcinoma survival may also be made by the online case-matching tool of the FinProg Study[5]. The most widely used web-tool for estimating the risk of breast carcinoma death for each patient is Adjuvant! Online[6,7,8,9], which sorts patients in a number of risk groups, and then uses data from the SEER national dataset, together with data from the metaanalyses of adjuvant therapy[10,11], to make empirically-based estimates of the risk of relapse and mortality at ten years after diagnosis, and the impact which various adjuvant treatment choices might have on relapse and mortality.

We have developed a new mathematical framework, the *binary biological model of cancer metastasis*, for capturing the features of the lethal and non-lethal spread of cancer cells[13,15,1712]. This framework includes a series of linked equations, the *SNAP* (*Size+Nodes+PrognosticMarkers*) method, which can be used to integrate information on tumor size, nodal status, and other prognostic factors into an estimate of the risk of cancer death for each patient.[13,14,15] The *SNAP* method has proven to be remarkably accurate in stratifying patients and in estimating the risk of breast carcinoma death, as confirmed in multiple populations of patients (see reference 15 and the various technical reports available at http://www.lifemath.net/cancer/about/techreports/index.php).

Here we describe the creation of a set of web-based calculators, available at http://www.CancerMath.net, for breast carcinoma patients and their physicians. These calculators use of the *SNAP* method to estimate each patient's risk of death to cancer, as well as the risk of death to causes other than cancer, the overall death rate, and life expectancy, together with information on the impact which various adjuvant treatments can be expected to have on these measures of survival (*breast carcinoma treatment* calculator). The CancerMath.net website also contains calculators which can provide survival estimates at the time of diagnosis that can expected from the current standard of care treatment (the *breast carcinoma outcome* calculator), survival estimates at various times after the time of diagnosis (the *breast carcinoma conditional survival* calculator), and estimates of the chance that cancer will be present in the nodes (the *breast carcinoma nodal status* calculator).

METHODS

Construction of the calculators

The calculators were written in JavaScript, PHP, and HTML, using XML/SWF Charts v5.07 package along with Adobe Flash to animate and display the graphs.

Patient data

Two datasets were used to test both the underlying mathematics behind our calculators and the Adjuvant! Online breast cancer calculator:

SEER Dataset: 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. The total number of patients is 508,861. Analysis was restricted to patients with 1-50 mm tumors and 0-7 positive lymph nodes. The number of these patients was 362,491. Testing of the *breast carcinoma outcome* calculator whose parameters were designed to capture survival of women with breast carcinoma after 1987, when most eligible patients were receiving adjuvant chemotherapy of hormone therapy, is restricted to patients with 1-50 mm tumors and 0-7 positive lymph nodes, diagnosed after 1987, since this was the target population for this calculator. The number of these patients was 293,576.

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 was restricted to patients with 1-50 mm tumors and 0-7 positive lymph nodes. The number of these patients was 7,907. Testing of the *breast carcinoma outcome* calculator whose parameters were designed to capture survival of women with breast carcinoma after 1987, when most eligible patients were receiving adjuvant chemotherapy of hormone therapy, is restricted to patients with 1-50 mm tumors and 0-7 positive lymph nodes, diagnosed after 1987, since this was the target population for this calculator. The number of these patients was 6,415.

Mathematical methods

The essential mathematical framework used by the calculators, the *binary biological model of cancer metastasis*, includes a series of linked equations, the *SNAP* (*Size+Nodes+PrognosticMarkers*) method, which can be used to integrate information on tumor size, nodal status, and other prognostic factors into an estimate of the risk of cancer death for each patient (TABLE I), as well as an expression, *NodalSizeOnly*, Equation, for relating tumor size to the chance that cancer in present in the lymph nodes. Details can be found in reference 15 and the various technical reports available at http://www.lifemath.net/cancer/about/techreports/index.php).

The impact of adjuvant therapy on outcome is identical to Adjuvant! Online, from the reductions in death reported by the metaanalyses[10,11], as summarized by Radvin et al[5,6], and the values were determined by decompiling the java file that drives the Adjuvant! Online calculator.

RESULTS

The CancerMath calculators

Four types of breast carcinoma calculators are available at CancerMath.net website:

1) A *breast carcinoma outcome* calculator (Figure 1b), which calculates survival information at the time of diagnosis that can expected from the current standard of care treatment;

2) A *breast carcinoma conditional survival* calculator (Figure 1c), which calculates survival information at various times after the time of diagnosis that can expected from the current standard of care treatment;

3) A *breast carcinoma treatment* calculator (Figure 1d, which calculates survival information at the time of diagnosis that can expected with and without a range of adjuvant therapy options.

4) A *breast carcinoma nodal status* calculator (Figure 1e), which calculates chance that cancer will be present in the nodes.

The CancerMath.net website contains calculators for a variety of other uses, which have either been described previously[16], or which will be described in detail in the future. These include calculators for estimating the risk of nipple involvement, which may be used when deciding whether to use nipple-sparing mastectomy (as described in Rusby et al[16]), calculators for melanoma survival (Jean et al, in preparation) a calculator for estimating renal cell carcinoma survival[13], and a calculator for estimating the chance that cancer will be present in the lymph nodes for melanoma patients (Jean et al, in preparation).

The math behind the CancerMath calculators

All four calculators accept information on patient age, tumor size, ER/PR/Her2 status, histology, and grade. The *breast carcinoma treatment, outcome calculator* and *conditional survival* calculators also accept information on number of positive nodes, and provide estimates of life expectancy, the 15-year disease-specific Kaplan-Meier death rate (FIGURE 2), as well as the risk of death over the next 15 years to cancer, to causes other than cancer, and to all causes (FIGURE 3). The *treatment calculator* provides these estimates with and without various adjuvant treatment options.

The essential mathematics used by the calculators for calculating cancer lethality is the *SNAP* (*Size+Nodes+PrognosticMarkers*) method, which combines information on tumor size, number of positive nodes, and other prognostic factors, to provide an estimate of the 15-Year Kaplan-Meier cancer-specific death rate, *L* (TABLE I).[13,14,15,17] Of the parameters used to drive the *SNAP* calculation, perhaps the most important is *Q*, which provides the key metric for measuring cancer lethality, and which is derived as a measure of the probability of the lethal spread of cancer cells.[14] *Q* can either measure the difference in lethality seen when comparing cancers - its value is tenfold higher for melanoma than for breast carcinoma and at fifteen-fold greater for melanoma than for renal cell carcinoma[13] – or it can measure more subtle differences in lethality when comparing subtypes of a single cancer, captured by *g* parameters, which sit next to *Q* and modify its value (TABLE I). The CancerMath calculators use these *g* parameters to take into account age, ER/PR/Her2 status, histology, and grade into the estimates of the risk of death.[15] *Q* can also measure differences in lethality that have occurred over time, as the result of changes in treatment. For example, from the data on adjuvant therapy reported by Marioto et al[18], we have found that between 1975 and 1983 only 10% of women received adjuvant therapy, while from 1987 onward, a relatively constant majority of women (~80%) were receiving such therapy.[19] Reflecting these changes, we have found that the value of the *Q* parameter of the patients in SEER dataset before 1983 is 0.014751, while the value of *Q* after 1987 is 0.010054. Thus, using the pre-1983 value of *Q*=0.014751, makes it possible to generate an intrinsic underlying risk of death in the *breast carcinoma treatment calculator* which reflects the risk of underlying risk of death that occurred in error of the absence of adjuvant therapy, while using the post-1987 value of *Q*=0.010054 to drive the *breast carcinoma outcome and conditional survival calculators*, makes it possible to generate the risk of death that can be expected at the present time.

The risk of death for each of the first 15 years after diagnosis is determined by 15 sequential multiplications of the probability of death to cancer at 15 years, *L*, generated, as described above by the *SNAP* method (TABLE I), by a 15-part step function derived from the hazard function, which we

measured from all 362,491 breast carcinoma patients in the SEER (FIGURE 3). Analysis of the hazard function by ourselves[13], and by Karrison et al[20], reveals that while there is a measurable risk of cancer death for at least 15-year after diagnosis, by 15 years, ~90% of this risk has occurred. Furthermore, while there are subtle differences in this hazard function when comparing different groups of patients[21], we have found these differences make only a small practical impact on both the curves displayed and on the estimates of life expectancy provided by the CancerMath calculators (Chen et al, in preparation).

The CancerMath.net calculators calculate the impact of the cancer lethality, as described in the previous paragraph, on life expectancy with US census data on life expectancy by age[23]. These values are calculated in terms of both years, and days, of life that can be expected to be lost by cancer.

The *breast carcinoma treatment calculator* estimates the impact of adjuvant therapy in the same way that is used by Adjuvant! Online; by taking the reduction in death reported by the metaanalyses[10,11], as summarized by Radvin et al[5,6], and multiplying this reduction by the underlying risk of death specific to each patient. As noted above, CancerMath calculators calculate underlying risk of death, *L*, by the *SNAP* method, using by the pre-1983 value of Q=0.010054 (TABLE I), so as to recreate the death rate seen in a population of patients in whom there was essentially no adjuvant therapy. The values used in the *breast carcinoma treatment calculator* to estimating the impact of adjuvant therapy, taken form Radvin et al, act by multiplying this underlying risk of death, as can be seen in TABLE II.

The code behind the CancerMath calculators

The JavaScript code for the calculators, together with documentation, can be viewed in the browser by selecting "View→Source" in the browser menu. Here we outline the code for the *breast carcinoma treatment calculator*, but the *outcome* and *conditional survival* calculators have a similar structure. The code begins by loading several lengthy arrays, such as the life expectancy tables, and proceeds through a series of sequential "Steps', which are numbered below, and are also identified in the source code which is visible in the browser:

1. **STEP 1** The program collects information that the user has entered into the web form:
    1. Tumor size (in centimeters, to 1 decimal point)
    2. Whether nodal status is known, and if so, the number of positive nodes (0, 1, 2, etc.)
    3. Age
    4. Tumor prognostic factors: ER/PR/Her2 status, histology, Bloom-Richardson grade
    5. Adjuvant therapy options
2. **STEP 2** The program calculates yearly and cumulative breast cancer, non-breast cancer, and total death rates for each of the 15 years after diagnosis, based on the entered user information:
    1. The program loads information on the value of the parameters Q, Z, $j_{primaty}$ and $L_{per\text{-}node}$ (TABLE I), which are needed to execute the *SNAP* calculation (**STEP 2.b** below) for the probability of cancer death to cancer at 15 years.
    2. The program loads information on whether nodal status is known
    3. **STEP 2.a** The program loads the *g* parameters determined by the user input, and computes the product of all of them.
    4. **STEP 2.b** The program calculates the 15-year Kaplan-Meier cancer death rate, *L*, using the *SNAP* method (TABLES 1 and 2) from information on tumor size (**STEP 1** above), number of positive lymph nodes (**STEP 1** above), and other prognostic factors, as captured by the product of the *g* parameters (**STEP 2.a** above).
    5. **STEP 2.c** The program calculates 15 values for the breast cancer death rate in each of the 15 years after diagnosis. It accomplishes this for each year by multiplying the 15-year Kaplan-Meier cancer death rate, *L*, (calculated in **STEP 2.b** above) by the fraction of the total lethality which can be expected in each year. The total lethality expected in each year is a pre-computed 15-part step function derived from the breast carcinoma hazard function, which we have derived from data on all 362,491 breast carcinoma

patients in the SEER dataset for whom we have complete tumor size and nodal status information[22].

6. **STEP 2.d** The program calculates 15 values for the non-cancer death rate in each of the 15 years after diagnosis. It accomplishes this for each year by multiplying the fraction of patients not dying of cancer (=1 – (death rate calculated in **STEP 2.b**)) times the yearly risk of death due to non-cancer causes for the given age. The values for the yearly probability of death due to all non-cancer causes for ages 0 to 100 were taken from the National Vital Statistics Report (herein referred to as "NVSR")[23], while the values for ages 101 to 123 were extrapolated using the methodology described in the NVSR. Before creating the array values (nvsr_death_prob_yearly), we corrected them to account for the ~3% of deaths that can be ascribed to breast cancer. These values were loaded at the top of the program, before **STEP 1** as noted above.
7. **STEP 2.e** The program calculates 15 values for the overall death rate in each of the 15 years after diagnosis. It accomplishes this for each year by summing the cancer death rate (**STEP 2.c**) and the non-cancer death rate (**STEP 2.d**).
8. **STEP 2.f** The program calculates 15 values for cumulative breast cancer, non-breast cancer, and total death rates by summing the respective yearly values computed in the steps above.

3. **STEP 3** The program calculates the mean number of years of life left that can be expected for the cancer patient:
   1. **STEP 3.a** The program loads the value at year 0 for the number of people out of a group of 100,000 who survive to the user-specified age, based on yearly probabilities of death given by the NVSR.
   2. **STEP 3.b** For year 1 through year 15, the program multiplies the number of people out of the group of 100,000 who survive to the appropriate age (age+1 at year 1, age+2 at year 2, etc.) by the corresponding cumulative overall death rate (**STEP 2.f**). This applies the additional risk from cancer.
   3. **STEP 3.c** The program calculates the survival difference at year 15 by subtracting the calculated number of individuals surviving to year 15 from the NVSR-given value for the corresponding age (age+15).
   4. **STEP 3.d** The program then calculates 15 values for the total number of years lived by all surviving individuals in the group of 100,000 between each year, by taking the average of the number of individuals surviving to a given year and the number of individuals surviving to the following year.
   5. **STEP 3.e** The program calculates the total number of years lived by surviving individuals past each year, from year 0 to year 15. It begins at year 15, by taking the remaining years of life expected for the corresponding age (age+15), and subtracting away the total number of years that is expected to be lost because of cancer. The life expectancy in years for each age group is calculated as the number of people out of the group of 100,000 who survive to that age (from NVSR) multiplied by the residual life expectancy at that age (also from NVSR data). That expected number is the survival difference calculated in **STEP 3.c** multiplied by the additional number of years beyond the age at year 15 to reach age 101.
   6. **STEP 3.f** Then, working backwards from year 14 to year 0, the program calculates the total number of years lived by surviving individuals past each year by adding this value for the following year to the total number of years lived between that year and the following year (**STEP 3.d**). For example, the total number of years lived by surviving individuals past year 14 is the total number of years lived by surviving individuals past year 15 plus the total number of years lived between year 14 and year 15.

7. **STEP 3.g** The program then calculates the mean life expectancy for the cancer patient by dividing the new total number of years lived by individuals of the specified age (the value at year 0 from **STEP 3.f**) by the number of people out of the group of 100,000 who survive to that age (**STEP 3.f**).
8. **STEP 3.h** The program calculates the expected years of life lost due to cancer, by subtracting the calculated life expectancy (**STEP 3.a**) from the NVSR-given life expectancy for the specified age.

4. **STEP 4** The program calculates the yearly and cumulative breast cancer death rates with therapy:
   1. **STEP 4.a** The program calculates the "risk-reduction" value based on the combination of therapies entered by the user and the information collected in **STEP 1.c** and **STEP 1.d**, consistent with the assumptions of Adjuvant! Online[5,6].
   2. **STEP 4.b** The program calculates 15 values for the breast cancer death rate with therapy in each of the 15 years after diagnosis by multiplying the 15-year Kaplan-Meier cancer death rate, *L*, (calculated in **STEP 2.b**) by the "risk-reduction" value computed above, and by the fraction of the total lethality which can be expected in each year (the 15-part step function described in **STEP 2.c** that captures the breast carcinoma hazard function). This step is analogous to that carried out in **STEP 2.c** described above, for the death rate that would occur in the absence of adjuvant therapy.
   3. **STEP 4.c** The program calculates 15 values for the cumulative breast cancer death rate in each of the 15 years after diagnosis by summing the respective yearly risks of cancer death, with therapy, (**STEP 2**) from the time of diagnosis. This step is analogous to that carried out in **STEP 2.f** described above, for the cumulative cancer death rate that would occur in the absence of adjuvant therapy.
   4. **STEP 4d** The program calculates 15 values for the cumulative overall death rate, with therapy, in each of the 15 years after diagnosis by summing the respective yearly risks of the sum of the cancer death rates with therapy (**STEP 4.c)** and non-cancer death (**STEP 2.c**) from the time of diagnosis.

5. **STEP 5** The program calculates life expectancy gained from therapy:
   1. **STEP 5.a** Using the method outlined in the National Vital Statistics Report[23], the program first substitutes the NVSR-given yearly non-cancer probability of death with the yearly overall risk of death with therapy, calculated in **STEP 4** (which includes death due to cancer as well as causes other than cancer).
   2. **STEP 5.b** Next, the program calculates the new total number of years lived by individuals of the specified age group past that age, by using the values from **STEP 3**
   3. **STEP 5.c** The program then calculates the mean life expectancy with therapy by dividing the new total number of years lived by individuals of the specified age (**STEP 2.d**) by the number of people out of the group of 100,000 who survive to that age (**STEP 1.b**).
   4. **STEP 5.d** The program calculates the life expectancy gained from therapy by subtracting the mean life expectancy with therapy (**STEP 2.e**) from the mean life expectancy for the cancer patient (**STEP 3**).

6. **STEP 6** The program graphs the risk curves for cancer (**STEP 2b**), cancer with therapy (**STEP 1.a**), non-cancer (**STEP 1.a**), overall (**STEP 2.d**), and overall with therapy (**STEP 5.b**) in the user-specified mode, either as mortality curves, survival curves, a bar graph, a pie chart, or a pictogram. For the outcome calculator, the program displays the life expectancy (**STEP 3.a**), the life expectancy lost to cancer (**STEP 3.a**), and the 15-year Kaplan-Meier cancer-specific death rate (**STEP 1**). For the treatment calculator, the program displays these values, as

well as the risk reduction value from therapy (**STEP 1**) and the life expectancy gained from therapy (**STEP 5.d**).

7. **STEP 7** The program computes grade and stage, according to the AJCC criteria[24]

The information provided by the CancerMath calculators

Adjuvant! Online was the pioneer in the field of cancer web-calculators, and its widespread use among practitioners demonstrated the demand for such information[25]. As such, it provided us with the starting point from which we could consider what additional information should be available in a second generation of breast carcinoma calculators:

- While Adjuvant! Online provides its measure of lethality at the single time point (10 years), the CancerMath calculators provide information on the risk of death for each of the first 15-years after diagnosis (FIGURE 3). The CancerMath calculators were built to provide information over such a time span because the analysis of the risk of cancer death over the very long term by Karrison et al[20], using a dataset with a relatively small set of patients that had been followed over a very long period of time, and by ourselves, using the SEER dataset (Chen et in preparation), revealed that there is a measurable risk of death for breast carcinoma lethality for the first 15 years after diagnosis, but that by year 15 at least 90% of that risk has passed. Values provided by the CancerMath calculators for each year include: the risk of death to cancer, the risk of death to causes other than cancer, and the overall risk of death. In addition, the CancerMath calculators display the 15-year Kaplan Meier cancer specific death rate. The CancerMath *breast carcinoma treatment* calculator also provides information on the impact that various adjuvant treatments can be expected to have on these measures of lethality.
- Adjuvant! Online lumps patients into rather large risk groups, which the mathematics of the CancerMath calculators suggest contain individual patients with very different levels of risk. For example, the Adjuvant! Online calculator considers as a single group all patients with tumors 3-5cm and 1-3 positive nodes, and assigns a risk of cancer death of approximately 45% to this group. However, this group contains patients with tumor of 3 cm and 1 positive nodes, which the CancerMath calculators indicate have about a 27% 15-year cancer-specific death rate, while other patients in this Adjuvant! Online group that have tumors of 5 cm and 3 positive nodes, but have a 47% risk of death, as indicated by the CancerMath calculators.
- As noted by Radvin and his colleagues, "The outcomes of ductal and lobular cancers were accurately predicted, but for other histological subtypes, the predicted outcomes by Adjuvant! are too pessimistic".[8] This is not the case for the CancerMath calculators, which use *g*-parameters to make such calculations (see especially the values for tubular carcinoma and Mucinous Carcinoma in Figures 2o and 2s)
- The CancerMath calculators provide information on life expectancy, expressed in terms of both days of life and years of life, together with information on how life expectancy is shortened by the cancer diagnosis. The CancerMath *breast carcinoma treatment* calculator provides information on the benefit, in terms of both days of life and years of life, which various adjuvant treatments can offer.
- The survival information provided by the CancerMath calculators can be viewed in a variety of formats: in terms of death curves, survival curves, pie charts, or in terms of "smiley-face" charts, which have been provided to present the information is a fashion that may be more comprehensible to the lay person.
- The CancerMath calculators also take the information entered and provide the patient's classification (T, N, and M) and stage.
- The Adjuvant! Online calculator only provides information on the underlying 10-year risk of death that is valid at the time of diagnosis. This information becomes invalid as time passes, as the longer a women remains disease free, the greater is her chance of being cured of cancer. The CancerMath *breast carcinoma conditional survival* calculator fills this gap by providing information on the risk values for each of the first 15 years after diagnosis.

- While the Adjuvant! Online calculation is carried out by a compiled java file, which is essentially invisible to the user, the CancerMath Java Script code that drives calculators is completely visible in the browser, together with abundant documentation. In addition, the CancerMath website contains an exhaustive set of Technical Reports, which explains the underlying math, the determination of the values of the parameters, and the various tests of the accuracy of the mathematics.

The accuracy of the CancerMath calculations:

To test the accuracy of the *SNAP* calculations used in the *breast carcinoma outcome* calculator (Figure 1b), individuals in the SEER and Partners datasets were sorted into groups of various types and the predicted survival value calculated by the *SNAP* method were compared with the actual 15-year cancer specific Kaplan-Meier death rates for each group. For example, the *SNAP* method was used to sort the 293,576 such patients from the SEER dataset who were diagnosed after 1987 (the target population for this calculator) into groups of differing by a 2% risk of death (i.e. those patients expected to have 0-2% risk of death, 2%-4% risk of death, 4%-6%, etc). A Kaplan-Meier survival analysis for group revealed that the expected and observed survival values agreed within 2% for the 97% of patients with up to a 48% risk of death, while for the remaining 3% of patients with greater than a 48% chance of death, the expected and observed survival values for each group agreed within 7%. Additionally, when patients in both the SEER and Partners populations were sorted by tumor size, nodal status, grade, age, histology, ER/PR status, sex, and race, the agreement between the expected and observed survival values also proved to be excellent (FIGURE 2).

Accuracy over the whole 15 year time point was tested, for both breast cancer and non- breast cancer lethality, by exporting the code to Matlab, and comparing the predicted survival curves with the actual survival curves for the patients from the SEER dataset who were diagnosed after 1987, again, the target population for the *breast carcinoma outcome* calculator. By this method, the calculator was found to be generally accurate, under-predicts non-breast cancer death rates by only 1-2% points. The greatest departure between observed and expected was seen for patients 70-80 years old, where the observed and expected curves differed by 6%. Overall estimates of lethality share the same biases, but again are largely accurate. The least well captured subgroups of patients were those differing with respect to ER status, with the calculator under-predicting short-term lethality for very lethal and ER- cancers, but is off by 7% at most in those cases. These values can be seen in Technical Report 12b at the CancerMath website[26].

DISCUSSION

Here we have described the creation of a set of web based calculators, which physicians can use when making treatment choices for individual patients. The CancerMath *breast carcinoma outcome* calculator has been found to capable of giving estimates of survival that are accurate to within a few percents for most patients, as seen by examining patients in two large populations, sorted by tumor size, nodal status, grade, age, histology, ER/PR status, sex, and race. The specificity and accuracy of the underlying mathematics of the CancerMath calculators has also been confirmed by its ability to stratify patients into groups differing by as little as 2% risk of death.

The mathematics used by the CancerMath calculators, the *binary-biological model of cancer metastasis*, was built by considering that each cell in a tumor will *either* spread to the periphery, thus leading to death, *or* it will not[13,14,27,30]. From this *either/or* quality, we were able to assign a probability value for the spread of cells, and from this basis, we derived the equations of the CancerMath calculators. In building our math from this intrinsically *either/or* quality, we took advantage a fundamental feature of all of microscopic entities, not only cells, but also molecules, atoms, electrons, photons, and genes (see: http://www.lifemath.net/binbio.html and reference 27). Indeed, we have found this to be a useful starting point for building mathematical tools for understanding a number of features of multicellullular systems, of which cancer lethality is but one example. For example, by examining the considering as discrete, *either/or*, events occur in embryonic signaling, we have found that the growth of the tissues, organs, and anatomical structures to predictable sizes, at predictable times, and to predictable shapes can arise as a natural consequence of such discreteness.[27] Similarly, by examining the molecular mitotic signaling events that go on within cells as discrete, *either/or*, events, we have found that growth of each tissue to a normal size may also be the natural consequence of the discrete, *either/or*, nature of the events that occur among oncogenes and tumor suppressor gene products[27]. The same mathematics has also provided a way to see why mutations in some of these genes will lead to premalignant growth, while other combinations of mutations will lead to outright cancerous growth[27]. Using the same the *binary-biological model of cancer metastasis*, we have been able to create a computer simulation of cancer growth, spread, and detection, which could calculate such things as the impact of various mammographic screening intervals among women of various ages on the breast carcinoma survival rate[30-34]. A similar analysis of the interaction between chemotherapeutic agents and cells has made it possible to create a mathematical method for examining the impact of the timing and dosage of chemotherapeutic agent on the outcome of such therapy.[27] Thus, we would suggest that a mathematical consideration of the discrete, *either/or* events that go on among the cells, molecules, atoms, electrons, photons, and genes provides a versatile toolkit for understanding many aspects of normal and abnormal multicellularity, including cancer lethality, as we have seen at here in the CancerMath calculators.

One the advantages of the *binary biological* framework that drives the CancerMath calculators is that this mathematics can work with as little information as is at hand (the *SNAP* method can generate an estimate of a patient's 15-year death rate with just tumor size), as well as with as much information as is desired. The methods for measuring the lethal impact of prognostic factors and determining the values of the parameters used to calculate outcome (the *SizeAssessment* and *PrognosticMeasurment* methods) are described in reference 13, and an ongoing effort by our group is to collect such data on additional prognostic factors and add this information to the CancerMath calculators. We also welcome other researchers who wish to have other prognostic factors included in subsequent editions of the CancerMath calculators to send along their data. The *SizeAssessment* method also can be used to measure the impact of local recurrence on survival, thus offering a way to include the impact of adjuvant radiation therapy in subsequent versions of the CancerMath calculators.[28,29] Finally, the CancerMath *breast carcinoma treatment* calculator, like Adjuvant! Online, uses the reductions in death reported by the metaanalyses[10,11] as summarized by Radvin et al[6], as a risk reduction. That is, both calculators estimate the impact of adjuvant therapy by multiplying the reduction in death found in the trials by the underlying risk of death specific to each patient. This is a leap of faith by both CancerMath and Adjuvant! Online, but such a calculation arguably captures the general thinking by oncologists when considering the potential benefit of adjuvant therapy. As we have noted previously[30], the *SizeAssessment* method offers a way to quantify

the impact of adjuvant therapy from data outside of trials, and we are currently collecting such data to address the actual impact of adjuvant therapy on outcome for patients with tumors of various characteristics.

It is not unreasonable to ask why the CancerMath calculators are needed. Providing physicians with the most accurate estimates of survival applicable to each patient, rather relying on rough guesstimates would seem to be intuitively valuable, but why? Arguably, the reason for such a goal is for medicine to find the way to use the resources available, which are, by definition, limited, to achieve the maximal possible extension in life[1-]**Error! Bookmark not defined.**. But, to reach such a maximal possible extension in life, physicians and patients need to be able to know what the extension in life particular treatments will offer. It is for this reason that the CancerMath calculator give impact measurements not only in terms of death rates, but most importantly, in terms of life expectancy, the ultimate measure by which one decides whether a certain therapy has value. For example, the CancerMath *breast carcinoma treatment* calculator shows that a 34 year old ER-/PR- patient who has a 1cm mass and 1 positive node can be expected to gain 843 days of life from CMF therapy, while an 82 year old woman with these characteristics will only gain 47 days of life.

Although the CancerMath calculators make calculations for single individuals, the very same code can be exported out of the browser to calculate impact that various treatment choices will have on a population-wide basis. When combined with cost information, this code should also be able to derive cost/benefit values. Additionally, as noted above, the very same *binary-biological model of cancer metastasis* which drives the CancerMath calculators also drives a computer simulation model of breast cancer growth, lethality, and detection[31,32,33,34], which we have created to provide information on the impact on breast carcinoma survival of various usages of mammographic screening, particularly the impact of various screening intervals among women of various ages. We are now in the process of combining the two models so as to make an integrated model of breast cancer screening and treatment. Such an extension of the CancerMath calculators offer the possibility of identifying ways to utilize the financial resources now available for the treatment and detection of breast carcinoma so as to reach the greatest possible level of breast carcinoma survival.

The *binary biological* mathematical framework that drives the CancerMath calculators was created to capture cancer lethality arising from the spread of cancer cells. Indeed, we have found this mathematics to also capture many features of melanoma lethality, as well as being capable of providing highly accurate estimates of the risk of death for melanoma patients. This has lead to the development of a suite of melanoma calculators located at the CancerMath website (http://www.lifemath.net/cancer/melanoma/outcome/index.php, Jean et al, in preparation). There is no reason why the same *binary biological* mathematical framework could not also be used to create calculators for other cancers in which the main cause of death is the spread of cancer cells. Even in the absence of such a biologically-motivated mathematical framework, calculators can be made on a strictly empirical basis. Indeed, we have created just such a calculator for providing individualized information on benefit, in terms of days of life, that may be expected from the various Class-A preventive interventions recommended by the U.S. Preventive Services Task Force (http://www.lifemath.net/preventive/). Thus, we would suggest that the CancerMath *breast carcinoma* calculators provide an example of the sort of web-based tools that can be created to provide physicians and patients with the highly accurate, patient-specific information they need to reach the greatest possible savings in life.

TABLE 1

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

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

L *= 15 year* Cancer-specific Kaplan-Meier Death Rate

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 Diameter	*Q* = 0.014751 for the *treatment calculator* *Q* = 0.010054 for the *outcome and conditional survival calculators* *Z* = 1 $j_{primary}$=0.8057 if nodal status is known $j_{primary}$ =1 if nodal status is unknown *g* parameters:	*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*
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	$L_{per\text{-}node}$ = 0.07581	*The presence of each positive lymph node contributes approximately 8% extra chance of death*

g parameters:

Grade 1	0.4324
Grade 2	0.8570
Grade 3	1.1224
Age 21 – 30	1.2545
Age 31 – 40	1.1267
Age 41 – 50	0.8661
Age 51 – 60	1.0190
Age 61 – 70	1.0172
Age 71 – 80	1.0201
Age 81 – 90	1.1646
Ductal	1.0573
Lobular	0.9032
Intraductal and LCIS	0.8573
Mucinous	0.4646
Medullary	0.5995
Tubular	0.2752
Comedo	0.8645
Scirrhous	1.6314
Inflammatory	3.3130
Paget's disease	1.4535
Papillary	0.5414
Cribriform	0.9636
ER+ / PR+	0.9155
ER+ / PR-	1.1389
ER- / PR+	1.0462
ER- / PR-	1.1902
HER2 POSITIVE,	**1.5150**
HER2 NEGATIVE,	**0.9662**

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.**

Note: the value of several parameters differ slightly from those outlined in reference 13, which were derived from the overall SEER population. See Materials and Methods and Results sections for discussion.

TABLE II
Values Used for the Reduction in Death that may be Expected for Various Adjuvant Therapy Regimens

Adjuvant Therapy	**Age**	**ER Unknown**	**ER+**	**ER-**
Hormonal Therapy (Tamoxifen, Aromatase Inhibitor, Tamoxifen to Aromatase Inhibitor, Ovarian Ablation, Ovarian Ablation + Tamoxifen)	age < 50	20%	32%	0%
	age >= 50 and < 60	21%	32%	0%
	age >= 60	23%	32%	0%
Chemotherapy (CMF-like)	age < 50	30%	30%	30%
	age >= 50 and < 60	18%	16%	22%
	age >= 60	10%	8%	15%
Chemotherapy (Anthracyclines)	age < 50	41%	41%	41%
	age >= 50 and < 60	31%	29%	34%
	age >= 60	24%	23%	29%
Chemotherapy (1st generation regimens: CA*4, CMF, FE(50)C*6)	age < 50	30%	30%	30%
	age >= 50 and < 60	18%	16%	22%
	age >= 60	10%	8%	15%
Chemotherapy (2nd generation regimens: CA*4+T*4, DC*4, CEF*6, CAF*6, FAC*6, FE(100)C*6, E*4+CMF*4)	age < 50	44%	44%	44%
	age >= 50 and < 60	34%	33%	38%
	age >= 60	28%	26%	32%
Chemotherapy (3rd generation regimens: TAC*6, FE(100)C*3+D*3, CA*4+T*4)	age < 50	55%	55%	55%
	age >= 50 and < 60	47%	45%	49%
	age >= 60	42%	40%	45%

See Materials and methods sections for the basis of these values.

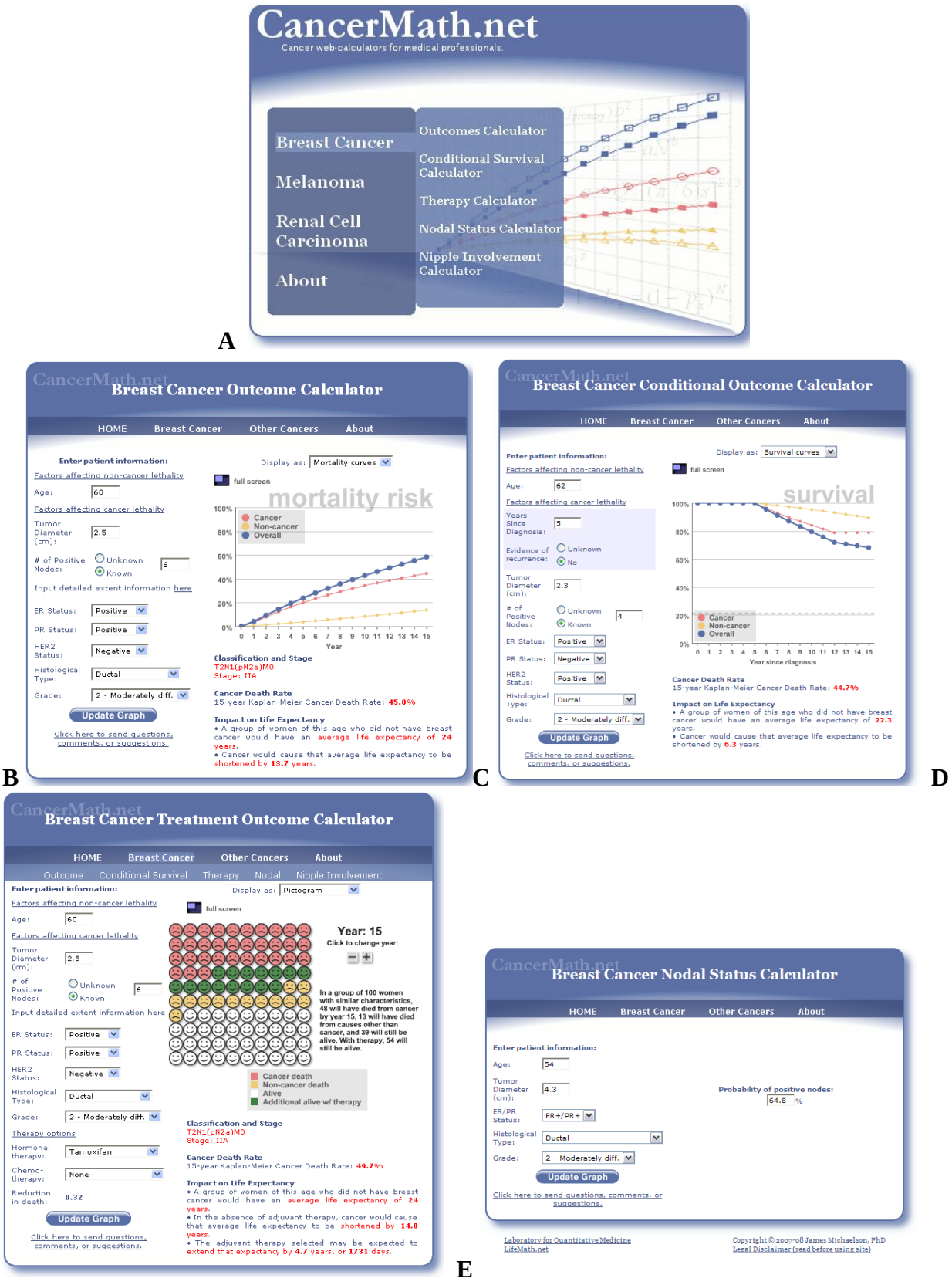


Figure 1. *CancerMath.net* breast carcinoma calculators.

Verification of the SNAP method. LIST OF AL 36 TABLE and Figure 2's shown below:
All patients after 1987, parameters for the outcome calculator, see TABLE I.

TABLE and Figure 2a: Verification on the SEER dataset. Patients grouped by 10% predicted lethality bins
TABLE and Figure 2b: Verification on the SEER dataset. Patients grouped by 5% predicted lethality bins
TABLE and Figure 2c: Verification on the SEER dataset. Patients grouped by 2% predicted lethality bins
TABLE and Figure 2d: Verification on the SEER dataset. Patients grouped by 1% predicted lethality bins
TABLE and Figure 2e: Verification on the SEER dataset. Patients grouped by 10% predicted lethality percentiles
TABLE and Figure 2f: Verification on the SEER dataset. Patients grouped by 5% predicted lethality percentiles
TABLE and Figure 2g: Verification on the SEER dataset. Patients grouped by 10 mm tumor size bins
TABLE and Figure 2h: Verification on the SEER dataset. Patients grouped by 5 mm tumor size bins
TABLE and Figure 2i: Verification on the SEER dataset. Patients grouped by 10% tumor size percentiles
TABLE and Figure 2j: Verification on the SEER dataset. Patients grouped by lymph nodes positivity status
TABLE and Figure 2k: Verification on the SEER dataset. Patients grouped by number of positive lymph nodes
TABLE and Figure 2l: Verification on the SEER dataset. Patients grouped by tumor grade
TABLE and Figure 2m: Verification on the SEER dataset. Patients grouped by estrogen and progesterone receptor
TABLE and Figure 2n: Verification on the SEER dataset Patients grouped by histological type
TABLE and Figure 2o: Verification on the SEER dataset. Permutations of 10 mm tumor size bins and number of positive lymph nodes
TABLE and Figure 2p: Verification on the SEER dataset. Permutations of 10 mm tumor size bins and tumor grade
TABLE and Figure 2q: Verification on the SEER dataset. Permutations of 10 mm tumor size bins and ER/PR receptor
TABLE and Figure 2r: Verification on the SEER dataset. Permutations of 10 mm tumor size bins and histological type
TABLE and Figure 2s: Verification on the SEER dataset. Permutations of number of positive lymph nodes and EER/PR receptor status
TABLE and Figure 2t: Verification on the SEER dataset. Permutations of tumor grade and ER/PR status
TABLE and Figure 2u: Verification on the SEER dataset. Permutations of tumor grade and histological type
TABLE and Figure 2v: Verification of the SNAP method on the SEER dataset. Grouped by race
TABLE and Figure 2w: Verification of the SNAP method on the SEER dataset Grouped by sex

TABLE and Figure 2aa: Verification on the Partners dataset. Patients grouped by 10% predicted lethality bins
TABLE and Figure 2bb: Verification on the Partners dataset. Patients grouped by 5% predicted lethality bins
TABLE and Figure 2cc: Verification on the Partners dataset. Patients grouped by 2% predicted lethality bins
TABLE and Figure 2dd: Verification on the Partners dataset. Patients grouped by 20% predicted lethality %'s
TABLE and Figure 2ee: Verification on the Partners dataset. Patients grouped by 10% predicted lethality %'s
TABLE and Figure 2ff: Verification on the Partners dataset. Patients grouped by 10 mm tumor size bins
TABLE and Figure 2gg: Verification on the Partners dataset. Patients grouped by 5 mm tumor size bins
TABLE and Figure 2hh: Verification on the Partners dataset. Patients grouped by 10% tumor size percentiles
TABLE and Figure 2ii: Verification on the Partners dataset. Patients grouped by number of positive lymph nodes
TABLE and Figure 2ijj: Verification on the Partners dataset. Patients grouped by tumor grade
TABLE and Figure 2ikk: Verification on the Partners dataset. Patients grouped by ER receptor status
TABLE and Figure 2ill: Verification on the Partners dataset. Patients grouped by histological type
TABLE and Figure 2mm: Verification on the Partners dataset. Permutations of 10 mm tumor size bins and tumor grade

I. VALIDATION WITH SEER DATASETS

Table and Figure 2a: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 10% predicted lethality bins using the Size+Nodes+PrognosticFactors equation

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0-10%	113839	6.30%	0.40%	5.57%	0.01%	-0.73%
10.1-20%	90667	13.84%	0.48%	14.50%	0.02%	0.66%
20.1-30%	46747	23.96%	0.87%	24.49%	0.03%	0.53%
30.1-40%	24348	34.95%	1.36%	34.49%	0.04%	-0.46%
40.1-50%	12845	44.60%	1.82%	44.43%	0.05%	-0.17%
50.1-60%	4556	59.27%	3.16%	53.77%	0.08%	-5.50%
					Mean (std. dev.)	-0.95% (2.30%)
					Mean weighted by N (std. dev.)	-0.13% (0.99)
					Root Mean Square (std. dev.)	2.30% (3.50%)
					Root Mean Square weighted by N (std. dev.)	0.91% (1.07)

†60.1-70%, 70.1-80% and 80.1+% groups are not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

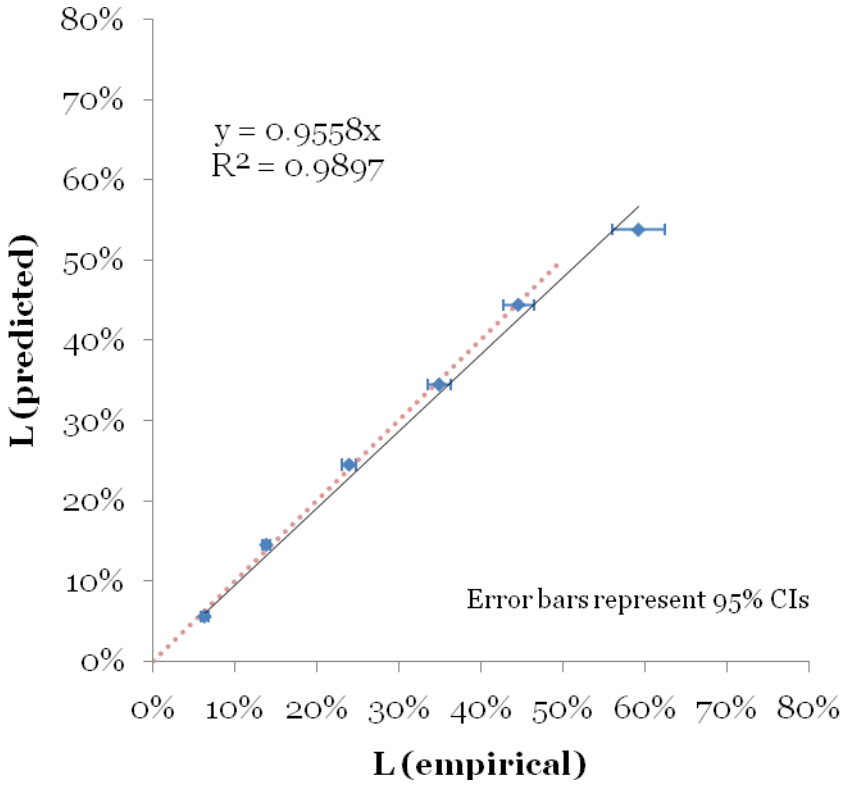


Table and Figure 2b: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 5% predicted lethality bins using the Size+Nodes+PrognosticFactors equation

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0-5%	49022	3.88%	(0.55%)	3.08%	(0.01%)	-0.80%
5.1-10%	64817	7.56%	(0.51%)	7.45%	(0.01%)	-0.11%
10.1-15%	51578	11.73%	(0.59%)	12.34%	(0.01%)	0.61%
15.1-20%	39089	16.67%	(0.79%)	17.35%	(0.01%)	0.68%
20.1-25%	26809	21.92%	(1.14%)	22.38%	(0.02%)	0.45%
25.1-30%	19938	26.66%	(1.33%)	27.33%	(0.02%)	0.67%
30.1-35%	13951	32.91%	(1.81%)	32.35%	(0.02%)	-0.56%
35.1-40%	10397	37.68%	(2.05%)	37.35%	(0.03%)	-0.32%
40.1-45%	7538	41.64%	(2.30%)	42.38%	(0.03%)	0.74%
45.1-50%	5307	48.82%	(3.00%)	47.35%	(0.04%)	-1.47%
50.1-55%	3138	57.24%	(3.91%)	52.25%	(0.05%)	-4.98%
55.1-60%	1418	63.92%	(5.18%)	57.12%	(0.07%)	-6.79%
					Mean (std. dev.)	-0.99% (2.42%)
					Mean weighted by N (std. dev.)	-0.01% (0.79%)
					Root Mean Square (std. dev.)	2.52% (3.79%)
					Root Mean Square weighted by N (std. dev.)	0.76% (0.90%)

†60.1-65.0%, 65.1-70%, 70.1-75%, 80.1+% groups are not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

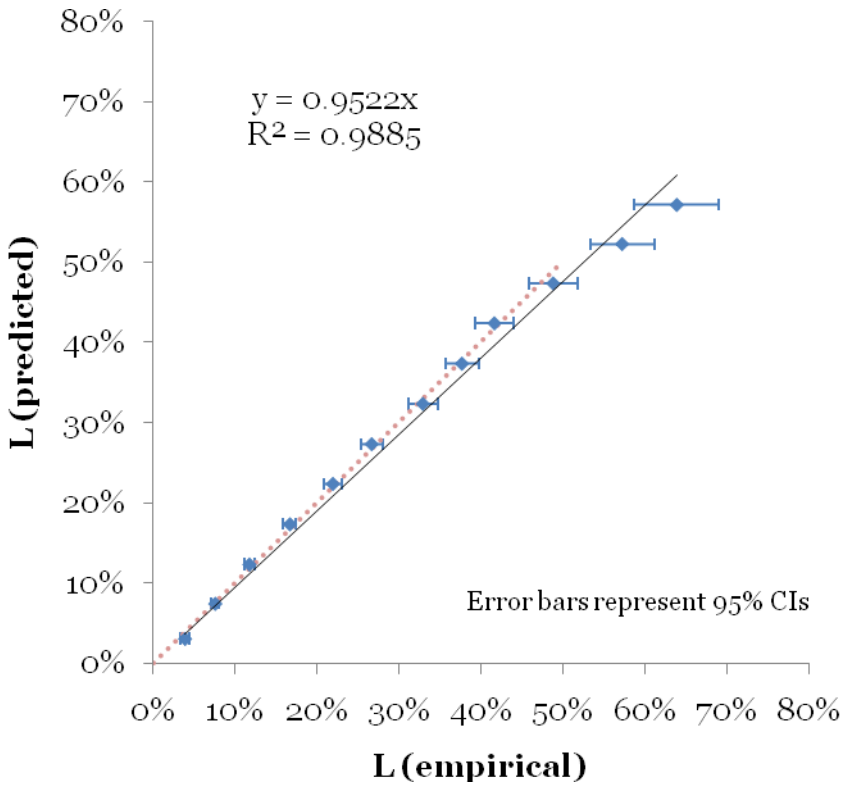


Table and Figure 2c: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 2% predicted lethality bins using the Size+Nodes+PrognosticFactors equation

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0-2%	10517	2.23%	(0.77%)	1.31%	(0.01%)	-0.92%
3-4%	25313	4.10%	(0.87%)	3.07%	(0.01%)	-1.03%
5-6%	26559	4.86%	(0.64%)	4.99%	(0.01%)	0.13%
7-8%	27230	7.41%	(0.83%)	7.01%	(0.01%)	-0.40%
9-10%	24220	8.82%	(0.85%)	9.02%	(0.01%)	0.21%
11-12%	22387	10.13%	(0.88%)	10.97%	(0.01%)	0.85%
13-14%	21147	12.23%	(0.89%)	12.96%	(0.01%)	0.72%
15-16%	17883	14.92%	(1.12%)	15.06%	(0.01%)	0.14%
17-18%	15394	16.75%	(1.38%)	17.01%	(0.01%)	0.27%
19-20%	13856	17.89%	(1.29%)	19.03%	(0.01%)	1.15%
21-22%	11825	20.51%	(1.64%)	21.00%	(0.01%)	0.49%
23-24%	10093	23.07%	(1.87%)	22.95%	(0.01%)	-0.12%
25-26%	9946	23.70%	(1.92%)	24.99%	(0.01%)	1.29%
27-28%	7504	26.19%	(2.19%)	27.00%	(0.01%)	0.81%
29-30%	7379	28.60%	(2.15%)	28.95%	(0.01%)	0.35%
31-32%	6174	30.21%	(2.53%)	30.97%	(0.01%)	0.76%
33-34%	5471	34.40%	(3.06%)	33.01%	(0.02%)	-1.39%
35-36%	4680	35.41%	(2.95%)	34.99%	(0.02%)	-0.42%
37-38%	4203	34.93%	(2.99%)	36.97%	(0.02%)	2.04%
39-40%	3820	42.54%	(3.73%)	38.94%	(0.02%)	-3.59%
41-42%	3327	40.79%	(3.48%)	40.99%	(0.02%)	0.19%
43-44%	2886	41.45%	(3.55%)	43.02%	(0.02%)	1.57%
45-46%	2537	42.98%	(3.93%)	44.97%	(0.02%)	1.99%
47-48%	2214	50.32%	(4.80%)	46.98%	(0.02%)	-3.34%
49-50%	1881	51.08%	(5.00%)	48.98%	(0.03%)	-2.09%
51-52%	1471	53.98%	(5.75%)	50.97%	(0.03%)	-3.01%
53-54%	1175	59.62%	(5.95%)	52.93%	(0.03%)	-6.69%
55-56%	887	60.87%	(7.12%)	54.94%	(0.04%)	-5.93%
57-58%	600	61.45%	(7.99%)	56.94%	(0.05%)	-4.51%
					Mean (std. dev.)	-0.71% (2.27%)
					Mean weighted by N (std. dev.)	-0.05% (0.96%)
					Root Mean Square (std. dev.)	2.34% (3.28%)
					Root Mean Square weighted by N (std. dev.)	0.94% (1.19%)

†59-60%, 61-62% … 87-88%, 89-90% not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

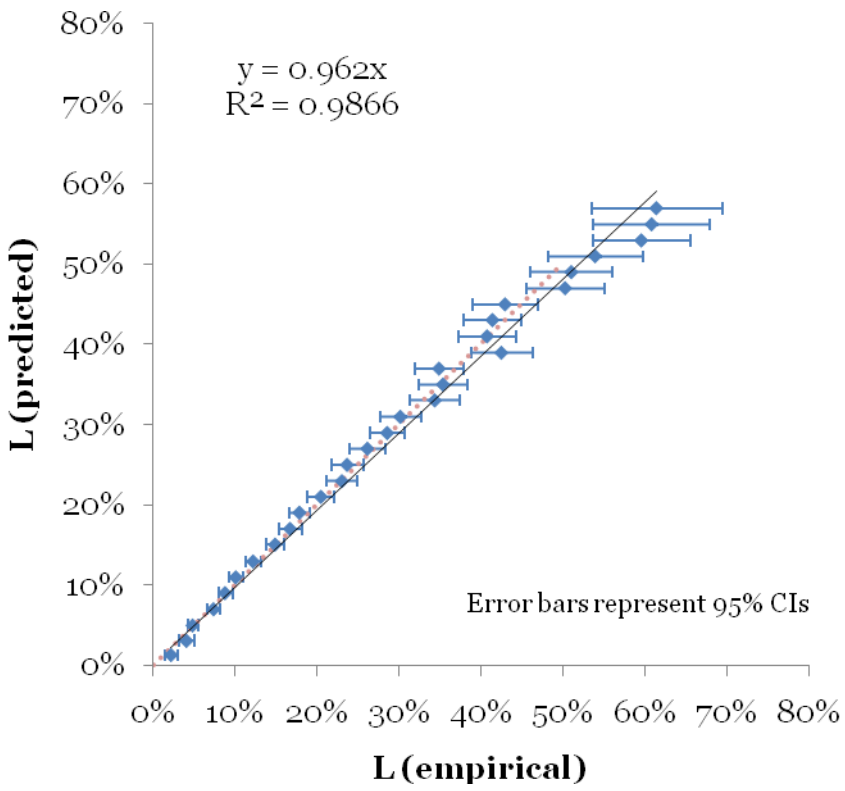

y = 0.962x
R² = 0.9866
Error bars represent 95% CIs
L (predicted)
L (empirical)
0%
10%
20%
30%
40%
50%
60%
70%
80%

Table and Figure 2d: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 1% predicted lethality bins using the Size+Nodes+PrognosticFactors equation

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
1	3153	2.58%	1.65%	0.67%	0.01%	-1.91%
2	7364	2.11%	0.87%	1.58%	0.01%	-0.53%
3	11392	2.69%	0.75%	2.55%	0.01%	-0.14%
4	13921	5.07%	1.36%	3.50%	0.00%	-1.57%
5	13192	4.40%	0.88%	4.49%	0.00%	0.09%
6	13367	5.27%	0.91%	5.48%	0.00%	0.21%
7	14690	7.09%	1.04%	6.53%	0.00%	-0.56%
8	12540	7.89%	1.39%	7.58%	0.00%	-0.32%
9	11605	7.51%	1.02%	8.50%	0.00%	0.99%
10	12615	10.25%	1.43%	9.50%	0.00%	-0.75%
11	11780	10.19%	1.20%	10.50%	0.01%	0.31%
12	10607	10.04%	1.30%	11.50%	0.01%	1.46%
13	11394	11.57%	1.18%	12.48%	0.01%	0.91%
14	9753	13.01%	1.36%	13.51%	0.01%	0.50%
15	8044	14.83%	1.71%	14.51%	0.01%	-0.33%
16	9839	14.98%	1.48%	15.51%	0.01%	0.54%
17	7879	16.06%	1.84%	16.53%	0.01%	0.46%
18	7515	17.49%	2.10%	17.52%	0.01%	0.03%
19	6681	17.63%	1.90%	18.50%	0.01%	0.88%
20	7175	18.15%	1.75%	19.53%	0.01%	1.38%
21	5918	18.84%	2.08%	20.50%	0.01%	1.66%
22	5907	21.96%	2.45%	21.51%	0.01%	-0.46%
23	5322	24.10%	2.50%	22.49%	0.01%	-1.61%
24	4771	21.03%	2.46%	23.47%	0.01%	2.44%
25	4891	22.70%	2.80%	24.51%	0.01%	1.81%
26	5055	24.63%	2.66%	25.46%	0.01%	0.83%
27	3629	25.20%	3.04%	26.49%	0.01%	1.28%
28	3875	27.15%	3.16%	27.48%	0.01%	0.33%
29	3903	28.88%	3.06%	28.50%	0.01%	-0.38%
30	3476	28.31%	3.00%	29.46%	0.01%	1.15%
31	3218	29.27%	3.42%	30.48%	0.01%	1.21%
32	2956	31.21%	3.68%	31.50%	0.01%	0.29%
33	2630	34.86%	4.38%	32.50%	0.01%	-2.35%
34	2841	34.07%	4.28%	33.49%	0.01%	-0.59%
35	2306	36.36%	4.52%	34.49%	0.01%	-1.87%
36	2374	34.41%	3.81%	35.48%	0.01%	1.07%
37	2319	34.18%	4.06%	36.53%	0.01%	2.35%
38	1884	35.80%	4.42%	37.51%	0.01%	1.71%
39	2082	43.13%	5.07%	38.47%	0.01%	-4.66%
40	1738	41.90%	5.56%	39.51%	0.01%	-2.39%
41	1758	42.31%	4.74%	40.51%	0.01%	-1.80%
42	1569	38.96%	5.04%	41.52%	0.01%	2.56%
43	1423	42.25%	5.05%	42.52%	0.01%	0.27%
44	1463	40.67%	4.99%	43.49%	0.01%	2.82%
45	1325	43.74%	5.65%	44.51%	0.01%	0.77%

46	1212	41.34%	4.88%	45.48%	0.02%	4.14%
47	1168	47.81%	5.86%	46.52%	0.02%	-1.28%
48	1046	53.43%	7.96%	47.50%	0.02%	-5.93%
49	958	48.33%	6.25%	48.51%	0.02%	0.18%
50	923	54.37%	8.52%	49.47%	0.02%	-4.90%
51	771	54.51%	8.04%	50.49%	0.02%	-4.02%
52	700	52.60%	7.36%	51.49%	0.02%	-1.11%
53	664	57.07%	8.24%	52.50%	0.02%	-4.56%
54	511	62.69%	8.77%	53.48%	0.02%	-9.21%
					Mean (std. dev.)	-0.34% (2.36%)
					Mean weighted by N (std. dev.)	-0.09% (1.21%)
					Root Mean Square (std. dev.)	2.36% (3.62%)
					Root Mean Square weighted by N (std. dev.)	1.20% (1.61%)

†55%,56%,57% … 87%, 88% groups are not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

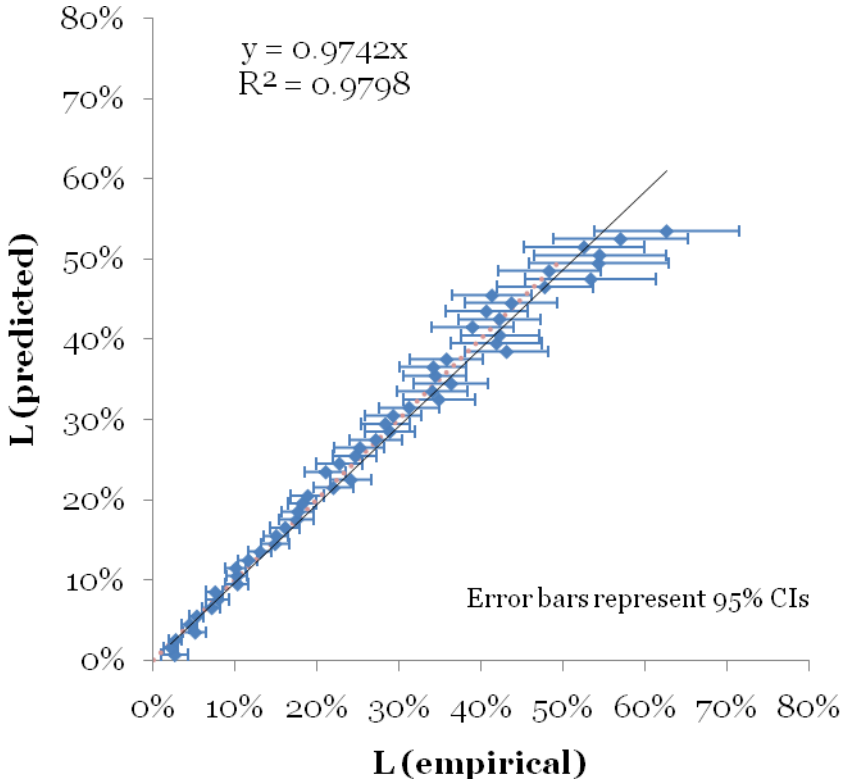


Table and Figure 2e: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 10% predicted lethality percentiles using the Size+Nodes+PrognosticFactors equation

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0-3.5%	28735	2.98%	0.63%	2.26%	0.01%	-0.71%
3.6-5.7%	29067	4.78%	0.66%	4.55%	0.01%	-0.23%
5.8-7.9%	29563	7.11%	0.77%	6.76%	0.01%	-0.36%
7.9-10.3%	30056	9.08%	0.78%	9.07%	0.01%	-0.01%
10.4-12.9%	29497	10.50%	0.75%	11.61%	0.01%	1.11%
13.0-16.2%	29594	14.32%	0.84%	14.54%	0.01%	0.22%
16.3-20.3%	29620	17.28%	0.92%	18.17%	0.01%	0.89%
20.4-25.9%	29779	22.51%	1.09%	22.97%	0.02%	0.45%
26.0-34.5%	28228	29.74%	1.20%	29.85%	0.03%	0.11%
34.6+%	29437	44.49%	1.22%	43.41%	0.08%	-1.08%
					Mean (std. dev.)	0.04% (0.68%)
					Mean weighted by N (std. dev.)	0.04% (0.68%)
					Root Mean Square (std. dev.)	0.64% (0.70%)
					Root Mean Square weighted by N (std. dev.)	0.64% (0.70%)

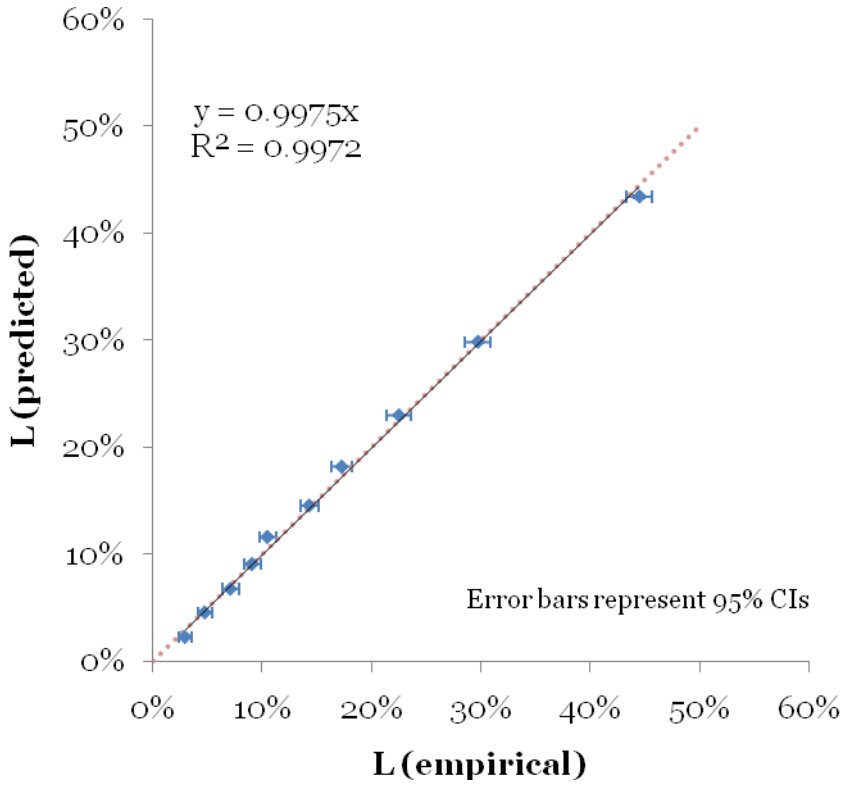


Table and Figure 2f: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 5% predicted lethality percentiles using the Size+Nodes+PrognosticFactors equation

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0-2.4%	14679	2.45%	0.72%	1.57%	0.01%	-0.88%
2.5-3.5%	14697	3.71%	1.00%	3.00%	0.00%	-0.71%
3.6-4.6%	14660	4.57%	1.05%	4.06%	0.01%	-0.51%
4.7-5.7%	14800	4.79%	0.80%	5.17%	0.01%	0.38%
5.8-6.8%	14663	6.76%	1.05%	6.27%	0.00%	-0.49%
6.9-7.9%	15192	7.74%	1.13%	7.40%	0.01%	-0.34%
8.0-9.2%	14171	7.45%	0.94%	8.54%	0.01%	1.09%
9.3-10.2%	14563	10.91%	1.30%	9.70%	0.01%	-1.20%
10.3-11.6%	14656	10.07%	1.07%	10.94%	0.01%	0.87%
11.7-12.9%	14805	10.94%	1.06%	12.26%	0.01%	1.32%
13.0-14.5%	14577	13.07%	1.16%	13.69%	0.01%	0.63%
14.6-16.2%	14696	15.45%	1.23%	15.34%	0.01%	-0.10%
16.3-18.0%	14664	16.78%	1.40%	17.08%	0.01%	0.30%
18.1-20.2%	14680	17.71%	1.24%	19.12%	0.01%	1.41%
20.3-22.7%	14684	22.47%	1.63%	21.42%	0.01%	-1.05%
22.8-25.7%	14786	22.15%	1.42%	24.21%	0.01%	2.06%
25.8-29.4%	14568	27.49%	1.62%	27.56%	0.02%	0.07%
29.5-34.4%	14677	31.94%	1.74%	31.88%	0.02%	-0.05%
34.5-42.0%	14678	38.46%	1.72%	37.99%	0.03%	-0.46%
42.1+%	14679	50.45%	1.73%	48.86%	0.09%	-1.58%
					Mean (std. dev.)	0.04% (0.96%)
					Mean weighted by N (std. dev.)	0.04% (0.96%)
					Root Mean Square (std. dev.)	0.94% (1.04%)
					Root Mean Square weighted by N (std. dev.)	0.94% (1.04%)

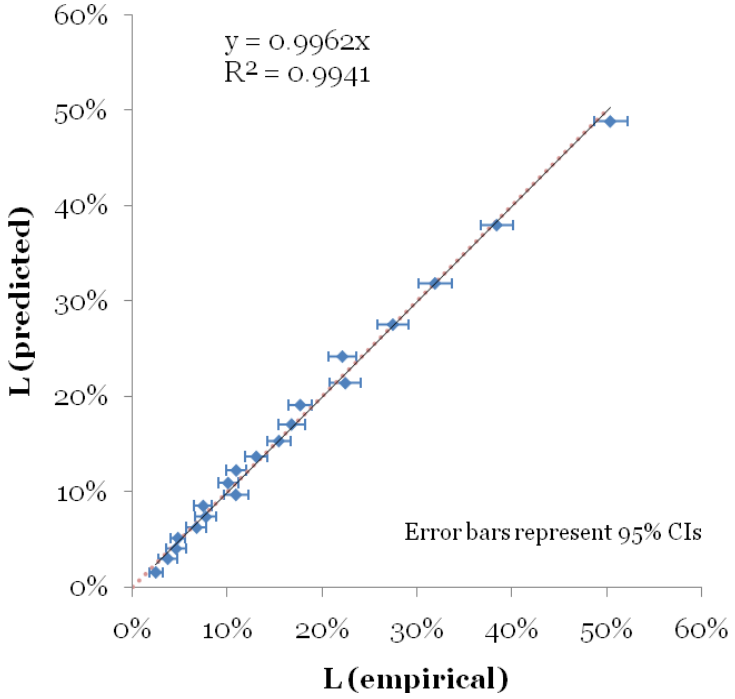


Table and Figure 2g: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 10 mm tumor size bins

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
1-10 mm	72129	6.83%	(0.46%)	6.18%	(0.04%)	-0.64%
11-20 mm	128000	13.95%	(0.44%)	13.86%	(0.05%)	-0.09%
21-30 mm	61307	24.68%	(0.76%)	23.78%	(0.08%)	-0.89%
31-40 mm	22011	32.72%	(1.37%)	32.38%	(0.15%)	-0.33%
41-50 mm	10129	36.67%	(2.07%)	38.90%	(0.23%)	2.23%
					Mean (std. dev.)	-0.06% (1.25%)
					Mean weighted by N (std. dev.)	-0.33% (0.53)
					Root Mean Square (std. dev.)	1.12% (1.45%)
					Root Mean Square weighted by N (std. dev.)	0.58% (0.62)

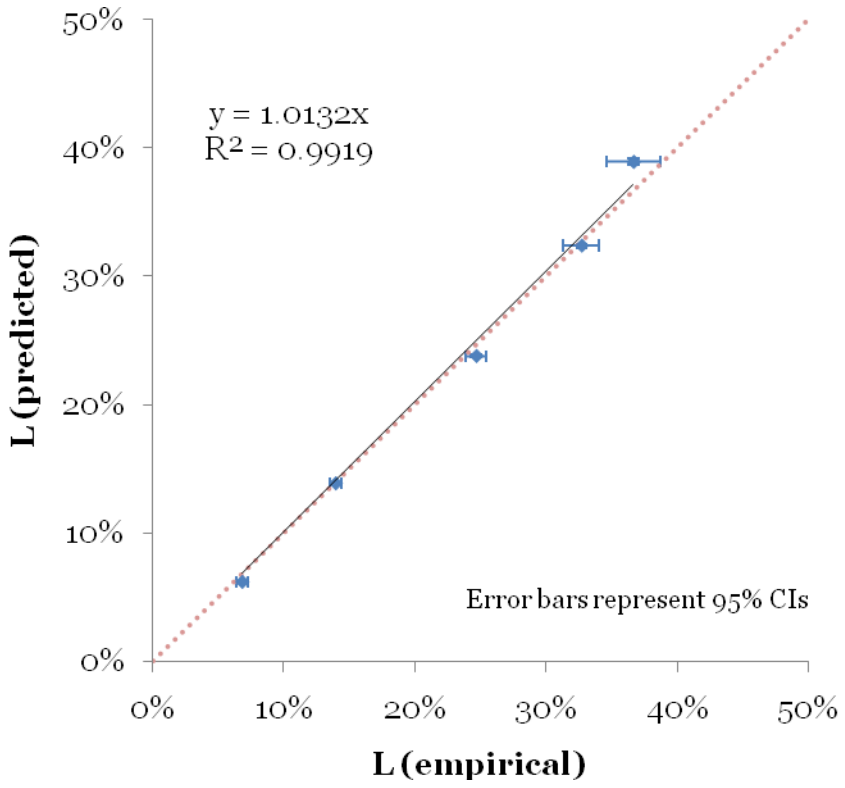


Table and Figure 2i: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 5 mm tumor size bins

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
1-5 mm	12229	5.66%	(1.12%)	3.66%	(0.08%)	-2.00%
6-10 mm	59900	7.02%	(0.50%)	6.70%	(0.04%)	-0.32%
11-15 mm	72892	11.83%	(0.58%)	11.61%	(0.05%)	-0.22%
16-20 mm	55108	16.70%	(0.69%)	16.84%	(0.07%)	0.14%
21-25 mm	37894	22.96%	(0.94%)	22.02%	(0.10%)	-0.94%
26-30 mm	23413	27.35%	(1.24%)	26.63%	(0.14%)	-0.72%
31-35 mm	12049	30.87%	(1.87%)	30.89%	(0.20%)	0.03%
36-40 mm	9962	34.76%	(2.01%)	34.18%	(0.22%)	-0.58%
41-45 mm	4733	34.16%	(3.14%)	37.61%	(0.34%)	3.44%
46-50 mm	5396	38.62%	(2.75%)	40.04%	(0.32%)	1.41%
					Mean (std. dev.)	0.02% (1.48%)
					Mean weighted by N (std. dev.)	-0.29% (0.56)
					Root Mean Square (std. dev.)	1.40% (1.92%)
					Root Mean Square weighted by N (std. dev.)	0.61% (0.67)

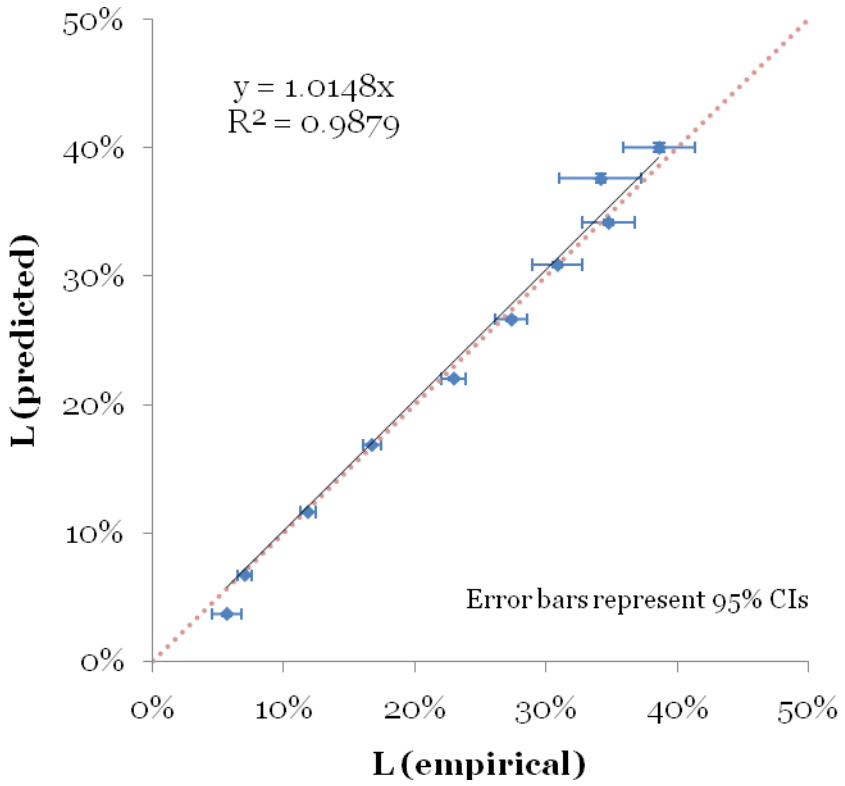


Table and Figure 2i: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 10% tumor size percentiles

Group	N	L empirical (SEM)	L predicted (SEM)	Difference (pred – emp)
1-7 mm	28270	5.28% (0.67%)	4.37% (0.05%)	-1.74%
8-9 mm	21414	6.49% (0.83%)	6.46% (0.07%)	-0.37%
10-12 mm	31275	9.03% (0.82%)	8.40% (0.07%)	-0.59%
13-14 mm	34538	10.86% (0.85%)	10.84% (0.07%)	-2.28%
15-16 mm	36259	13.19% (0.75%)	13.45% (0.08%)	-0.24%
17-19 mm	24023	15.91% (1.16%)	16.03% (0.11%)	0.49%
20-22 mm	36118	18.79% (0.83%)	19.02% (0.10%)	-4.50%
23-25 mm	26126	23.83% (1.11%)	22.81% (0.12%)	0.44%
26-33 mm	27261	27.77% (1.18%)	27.03% (0.13%)	-1.84%
34-50 mm	28292	34.33% (1.20%)	35.11% (0.14%)	2.05%
			Mean (std. dev.)	-0.86% (1.82%)
			Mean weighted by N (std. dev.)	-1.00% (2.06)
			Root Mean Square (std. dev.)	1.92% (2.48%)
			Root Mean Square weighted by N (std. dev.)	2.20% (3.06)

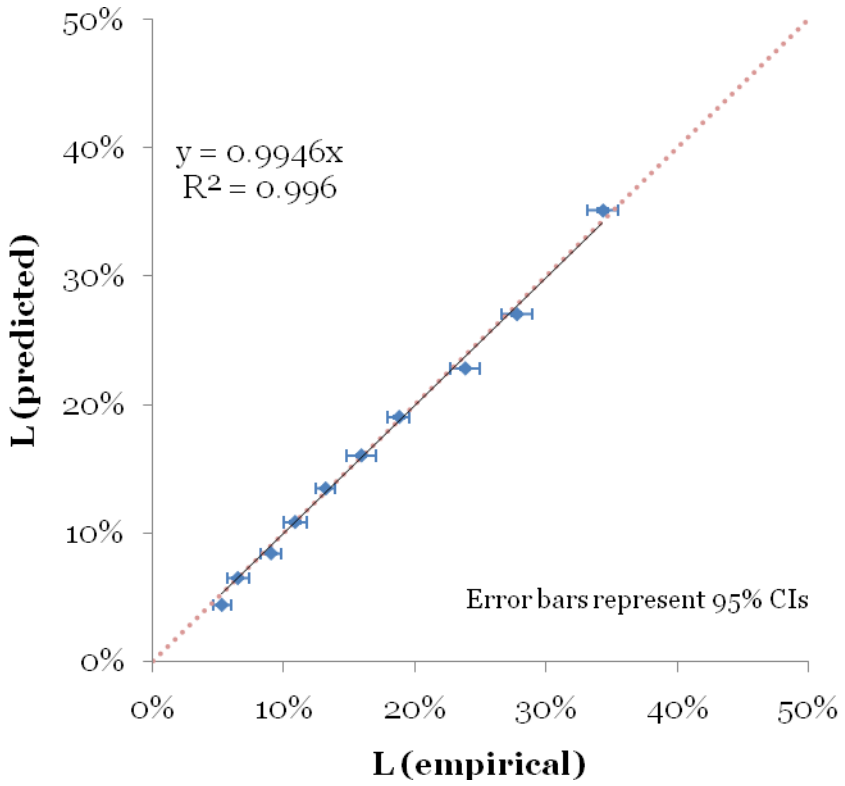


Table and Figure 2j: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by lymph nodes positivity status

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
Negative	206572	11.07%	(0.31%)	11.01%	(0.03%)	-0.06%
Positive	87004	30.56%	(0.72%)	28.86%	(0.08%)	-1.71%
					Mean (std. dev.)	-0.89% (1.17%)
					Mean weighted by N (std. dev.)	-0.55% (0.66%)
					Root Mean Square (std. dev.)	1.21% (1.44%)
					Root Mean Square weighted by N (std. dev.)	0.72% (0.85%)

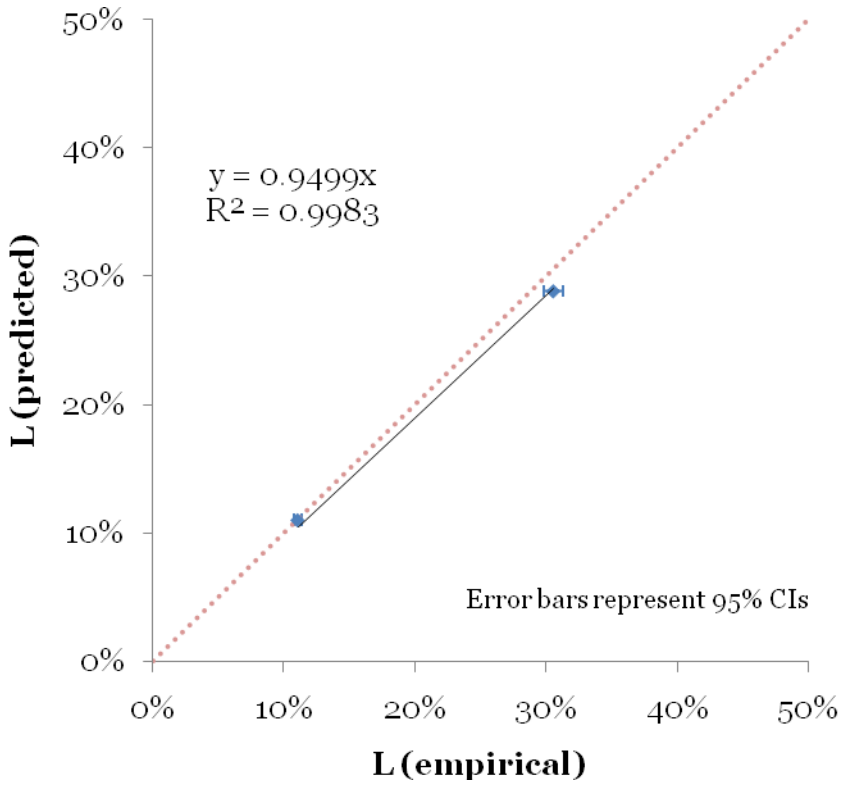


Table and Figure 2k: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by number of positive lymph nodes

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0	206572	11.07%	(0.31%)	11.01%	(0.03%)	-0.06%
1	39610	22.88%	(1.05%)	20.63%	(0.08%)	-2.25%
2	18710	28.29%	(1.47%)	27.74%	(0.11%)	-0.55%
3	10578	35.61%	(2.15%)	34.27%	(0.14%)	-1.33%
4	6804	40.05%	(2.49%)	39.62%	(0.16%)	-0.43%
5	4807	43.50%	(3.03%)	44.40%	(0.17%)	0.90%
6	3656	47.37%	(3.50%)	48.73%	(0.18%)	1.36%
7	2839	51.67%	(3.75%)	53.02%	(0.20%)	1.35%
					Mean (std. dev.)	-0.13% (1.29%)
					Mean weighted by N (std. dev.)	-0.39% (0.85%)
					Root Mean Square (std. dev.)	1.21% (1.28%)
					Root Mean Square weighted by N (std. dev.)	0.89% (1.44%)

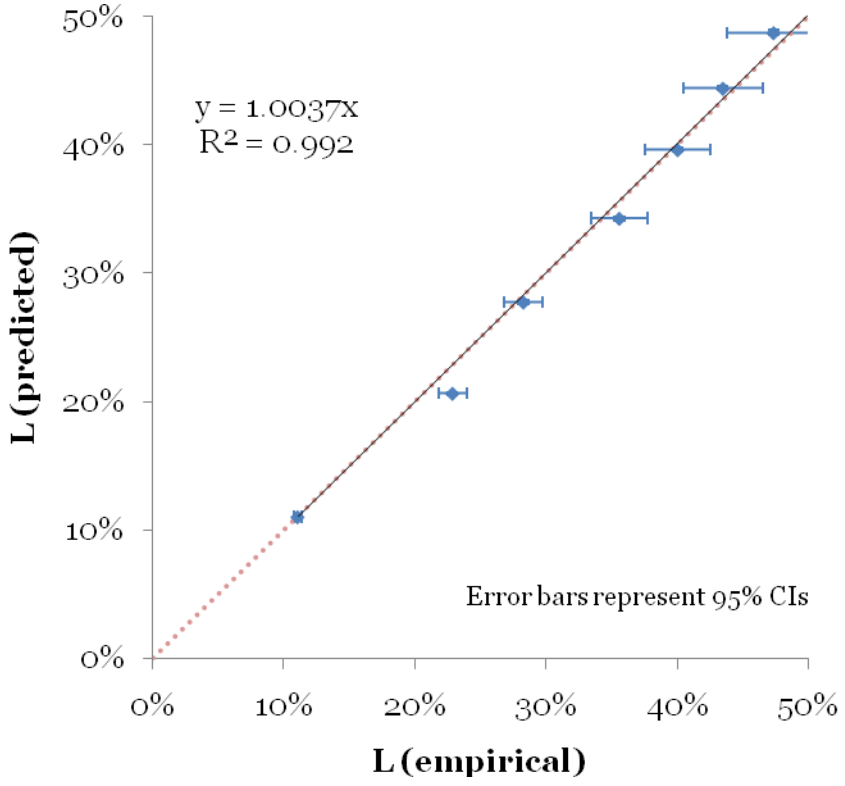


Table and Figure 2l: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by tumor grade

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
grade 1	49344	6.53%	(0.85%)	6.28%	(0.06%)	-0.24%
grade 2	108483	14.57%	(0.64%)	14.67%	(0.06%)	0.11%
grade 3	88369	23.33%	(0.64%)	23.37%	(0.08%)	0.04%
					Mean (std. dev.)	-0.03% (0.19%)
					Mean weighted by N (std. dev.)	-0.01% (0.15%)
					Root Mean Square (std. dev.)	0.05% (0.18%)
					Root Mean Square weighted by N (std. dev.)	0.12% (0.11%)

†grade 4 also exists in the dataset, but is not included in the calculation of the mean or displayed on graph; grade 4 no longer exists

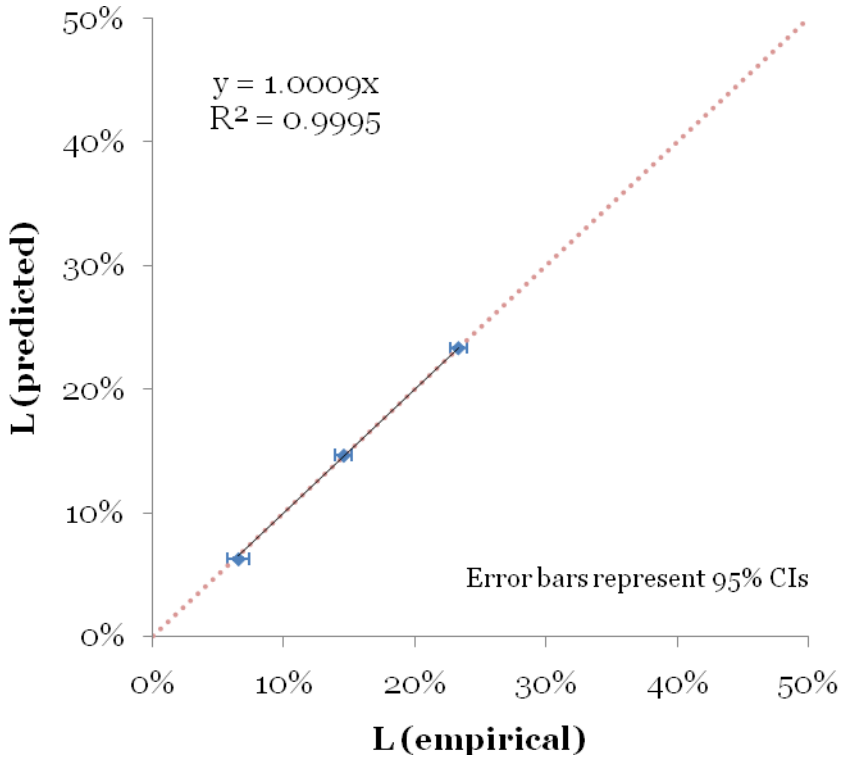


Table and Figure 2m: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by estrogen and progesterone receptor status

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
ER+/PR+	151742	14.68%	(0.84%)	14.01%	(0.06%)	-0.66%
ER+/PR-	28880	18.51%	(1.71%)	17.63%	(0.15%)	-0.88%
ER-/PR+	5519	18.16%	(1.58%)	18.16%	(0.33%)	0.00%
ER-/PR-	44672	22.25%	(0.94%)	22.96%	(0.12%)	0.71%
					Mean (std. dev.)	-0.21% (0.72%)
					Mean weighted by N (std. dev.)	-0.41% (0.97%)
					Root Mean Square (std. dev.)	0.66% (0.57%)
					Root Mean Square weighted by N (std. dev.)	0.94% (1.20%)

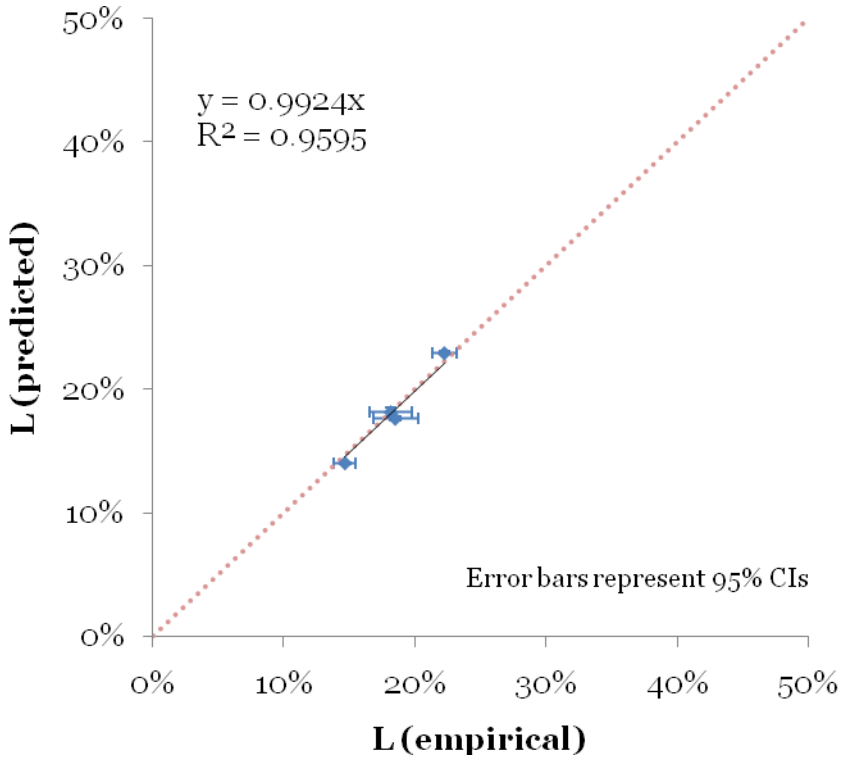


Table and Figure 2n: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by histological type

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
Ductal	215920	17.25%	(0.35%)	17.11%	(0.05%)	-0.14%
Lobular	21169	16.88%	(1.52%)	16.44%	(0.16%)	-0.44%
Intraductal+ LCIS	22282	15.31%	(1.51%)	14.67%	(0.15%)	-0.64%
Mucinous	7258	6.99%	(1.46%)	5.42%	(0.13%)	-1.57%
Comedo	3078	18.00%	(1.95%)	16.23%	(0.40%)	-1.77%
Scirrhous	369	24.91%	(7.83%)	24.41%	(1.44%)	-0.51%
Papillary	536	8.15%	(5.70%)	7.45%	(0.56%)	-0.70%
Cribriform	708	7.34%	(5.14%)	7.84%	(0.53%)	0.50%
					Mean (std. dev.)	-0.66% (0.73%)
					Mean weighted by N (std. dev.)	-0.26% (0.30%)
					Root Mean Square (std. dev.)	0.95% (1.09%)
					Root Mean Square weighted by N (std. dev.)	0.38% (0.52%)

†Medullary, Tubular, inflammatory, Paget's disease, Apocrine, and Phyllodes histologies are not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

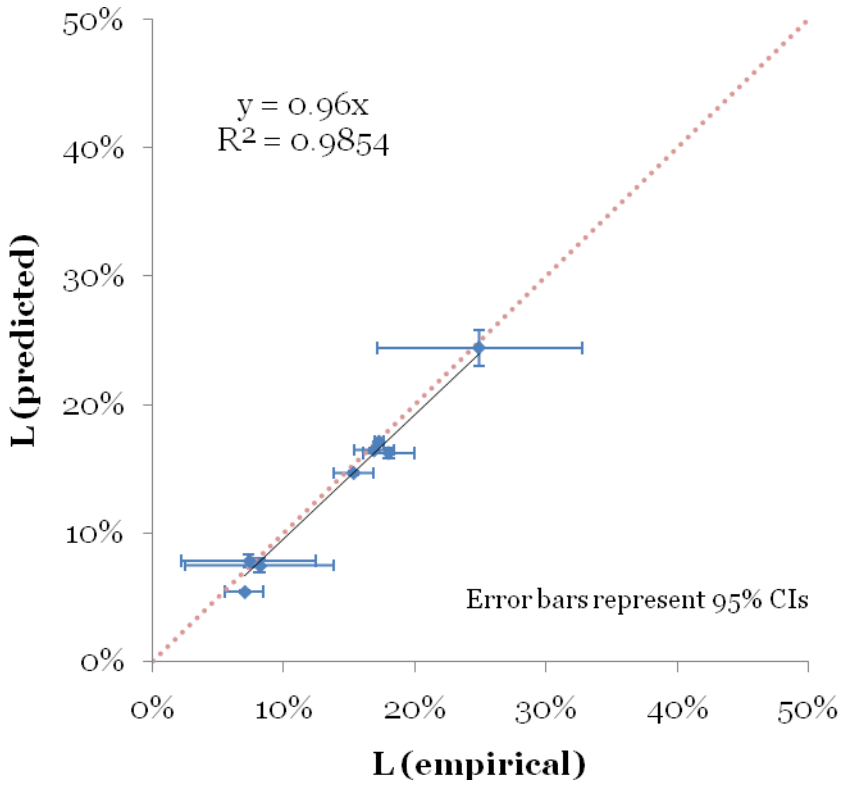


Table and Figure 2p: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Permutations of 10 mm tumor size bins and number of positive lymph nodes

Group† Size	Nodes	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
1-10 mm	0	63286	5.52%	(0.46%)	4.64%	(0.02%)	-0.88%
11-20 mm	0	93107	10.60%	(0.46%)	10.19%	(0.03%)	-0.40%
21-30 mm	0	35154	17.30%	(0.85%)	17.84%	(0.06%)	0.54%
31-40 mm	0	10632	22.23%	(1.72%)	25.28%	(0.14%)	3.05%
41-50 mm	0	4393	22.91%	(2.54%)	30.97%	(0.27%)	8.06%
1-10 mm	1	5356	11.47%	(1.87%)	12.17%	(0.06%)	0.69%
11-20 mm	1	17856	18.32%	(1.51%)	17.28%	(0.06%)	-1.05%
21-30 mm	1	10707	29.10%	(2.26%)	24.07%	(0.10%)	-5.03%
31-40 mm	1	3965	35.77%	(3.88%)	30.96%	(0.20%)	-4.80%
41-50 mm	1	1726	39.18%	(5.26%)	36.61%	(0.37%)	-2.58%
1-10 mm	2	1754	17.20%	(4.06%)	18.90%	(0.10%)	1.70%
11-20 mm	2	7653	22.12%	(2.18%)	23.65%	(0.08%)	1.53%
21-30 mm	2	5677	32.85%	(2.81%)	29.67%	(0.12%)	-3.18%
31-40 mm	2	2445	38.42%	(4.30%)	36.03%	(0.23%)	-2.39%
41-50 mm	2	1181	47.68%	(6.98%)	40.99%	(0.39%)	-6.69%
1-10 mm	3	709	19.75%	(6.16%)	24.92%	(0.13%)	5.17%
11-20 mm	3	3790	28.78%	(3.54%)	29.57%	(0.10%)	0.79%
21-30 mm	3	3531	39.27%	(3.84%)	35.15%	(0.14%)	-4.12%
31-40 mm	3	1712	44.75%	(4.89%)	41.18%	(0.27%)	-3.57%
41-50 mm	3	836	52.24%	(10.11%)	45.64%	(0.46%)	-6.60%
1-10 mm	4	427	26.91%	(8.53%)	30.68%	(0.18%)	3.77%
11-20 mm	4	2245	31.22%	(4.04%)	34.93%	(0.12%)	3.71%
21-30 mm	4	2287	45.14%	(4.49%)	39.94%	(0.16%)	-5.20%
31-40 mm	4	1185	48.29%	(6.36%)	45.44%	(0.29%)	-2.85%
41-50 mm	4	660	52.18%	(8.04%)	49.83%	(0.47%)	-2.35%
11-20 mm	5	1460	38.42%	(6.00%)	39.86%	(0.14%)	1.44%
21-30 mm	5	1717	41.88%	(4.67%)	44.49%	(0.17%)	2.60%
31-40 mm	5	833	54.36%	(7.24%)	49.31%	(0.31%)	-5.04%
11-20 mm	6	1096	36.35%	(5.31%)	44.34%	(0.15%)	7.99%
21-30 mm	6	1256	47.10%	(6.43%)	48.85%	(0.19%)	1.75%
31-40 mm	6	682	60.15%	(8.55%)	52.90%	(0.30%)	-7.25%
11-20 mm	7	793	42.41%	(6.31%)	48.54%	(0.17%)	6.13%
21-30 mm	7	978	54.87%	(6.51%)	52.66%	(0.21%)	-2.21%
31-40 mm	7	557	65.56%	(9.42%)	56.89%	(0.35%)	-8.67%
						Mean (std. dev.)	-0.76% (4.33%)
						Mean weighted by N (std. dev.)	-0.47% (2.18%)
						Root Mean Square (std. dev.)	4.33% (4.59%)
						Root Mean Square weighted by N (std. dev.)	2.20% (3.21%)

†for the combinations of 1-10 mm or 41-50mm with 5, 6, and 7 positive nodes, the groups are not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

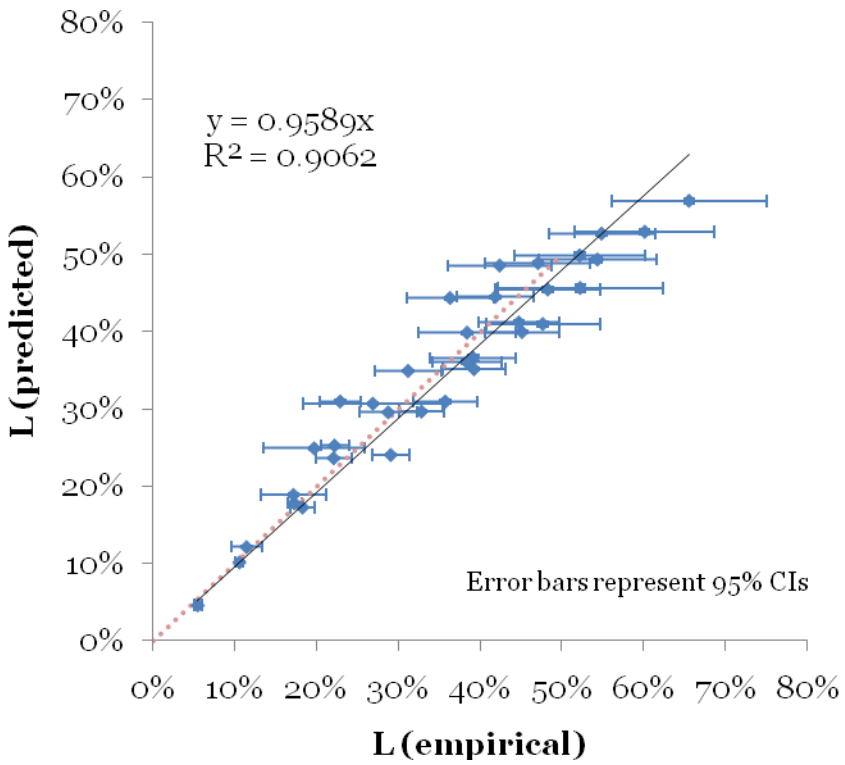

y = 0.9589x
R² = 0.9062
Error bars represent 95% CIs
L (predicted)
L (empirical)
0%
10%
20%
30%
40%
50%
60%
70%
80%

Table and Figure 2p: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Permutations of 10 mm tumor size bins and tumor grade

Group† Size	Grade	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
1-10 mm	1	21282	3.41%	(1.00%)	3.19%	(0.05%)	-0.22%
11-20 mm	1	20982	7.09%	(1.40%)	6.95%	(0.08%)	-0.14%
21-30 mm	1	5092	13.32%	(3.08%)	12.10%	(0.23%)	-1.22%
31-40 mm	1	1315	22.54%	(10.7%)	16.34%	(0.52%)	-6.20%
41-50 mm	1	673	11.25%	(4.81%)	19.67%	(0.79%)	8.42%
1-10 mm	2	27446	6.29%	(0.82%)	6.63%	(0.06%)	0.35%
11-20 mm	2	50953	13.01%	(0.91%)	13.18%	(0.07%)	0.17%
21-30 mm	2	20759	22.88%	(1.81%)	21.55%	(0.13%)	-1.33%
31-40 mm	2	6401	30.19%	(3.26%)	29.30%	(0.25%)	-0.90%
41-50 mm	2	2924	33.95%	(5.18%)	35.29%	(0.37%)	1.34%
1-10 mm	3	11735	10.96%	(1.40%)	9.51%	(0.12%)	-1.45%
11-20 mm	3	35984	18.47%	(0.95%)	18.23%	(0.09%)	-0.24%
21-30 mm	3	25495	28.21%	(1.26%)	27.82%	(0.11%)	-0.39%
31-40 mm	3	10460	36.01%	(2.26%)	36.45%	(0.18%)	0.44%
41-50 mm	3	4695	41.82%	(3.16%)	44.08%	(0.28%)	2.25%
						Mean (std. dev.)	0.06% (2.96%)
						Mean weighted by N (std. dev.)	-0.20% (0.65%)
						Root Mean Square (std. dev.)	2.86% (4.46%)
						Root Mean Square weighted by N (std. dev.)	0.66% (0.84%)

†grade 4 also exists in the dataset, but is not included in the calculation of the mean or displayed on graph; grade 4 no longer exists

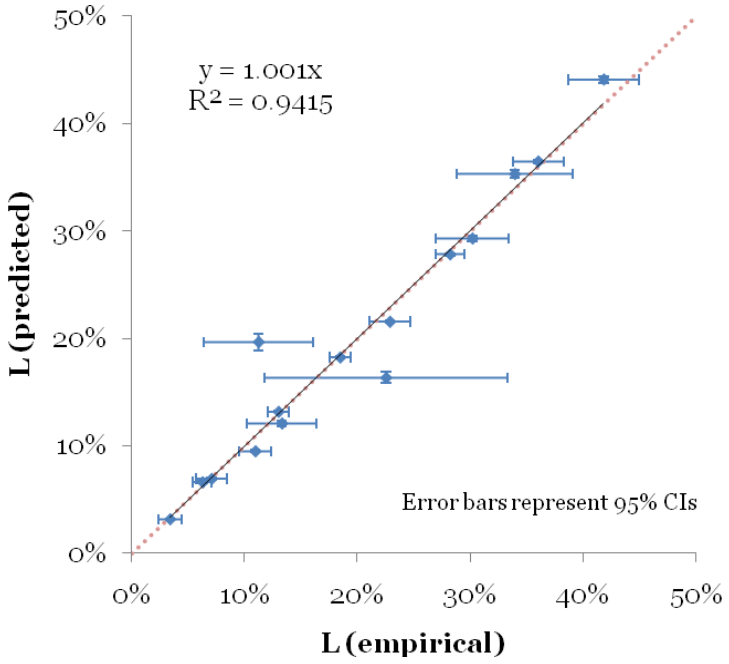


Table and Figure 2q: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Permutations of 10 mm tumor size bins and estrogen/progesterone receptor status

Group†		N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
Size	ER/PR						
1-10 mm	ER+/PR+	40037	6.48%	(1.48%)	5.51%	(0.05%)	-0.97%
11-20 mm	ER+/PR+	69671	12.79%	(1.42%)	12.41%	(0.06%)	-0.37%
21-30 mm	ER+/PR+	28800	22.27%	(1.47%)	21.63%	(0.12%)	-0.64%
31-40 mm	ER+/PR+	9105	30.89%	(2.95%)	29.69%	(0.23%)	-1.19%
41-50 mm	ER+/PR+	4129	33.59%	(4.56%)	35.71%	(0.37%)	2.12%
1-10 mm	ER+/PR-	7515	6.48%	(1.71%)	6.73%	(0.13%)	0.26%
11-20 mm	ER+/PR-	12295	14.32%	(1.59%)	15.24%	(0.15%)	0.92%
21-30 mm	ER+/PR-	5864	31.65%	(5.29%)	25.99%	(0.26%)	-5.66%
31-40 mm	ER+/PR-	2208	35.22%	(4.73%)	34.95%	(0.46%)	-0.28%
41-50 mm	ER+/PR-	998	39.33%	(6.92%)	41.65%	(0.71%)	2.32%
1-10 mm	ER-/PR+	1196	6.29%	(2.03%)	7.11%	(0.34%)	0.83%
11-20 mm	ER-/PR+	2378	14.62%	(2.18%)	15.33%	(0.35%)	0.72%
21-30 mm	ER-/PR+	1250	27.53%	(4.07%)	24.67%	(0.55%)	-2.86%
31-40 mm	ER-/PR+	486	36.97%	(7.77%)	33.08%	(0.91%)	-3.89%
1-10 mm	ER-/PR-	7058	11.98%	(1.96%)	8.95%	(0.15%)	-3.04%
11-20 mm	ER-/PR-	17516	18.12%	(1.60%)	18.03%	(0.12%)	-0.09%
21-30 mm	ER-/PR-	12393	26.50%	(1.87%)	27.93%	(0.16%)	1.43%
31-40 mm	ER-/PR-	5283	33.44%	(2.62%)	36.71%	(0.26%)	3.26%
41-50 mm	ER-/PR-	2422	39.72%	(4.16%)	43.99%	(0.40%)	4.27%
						Mean (std. dev.)	-0.15% (2.47%)
						Mean weighted by N (std. dev.)	-0.36% (1.34%)
						Root Mean Square (std. dev.)	2.41% (2.90%)
						Root Mean Square weighted by N (std. dev.)	1.36% (1.68%)

† the group that is the combination of 41-50 mm with ER-/PR+ is not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

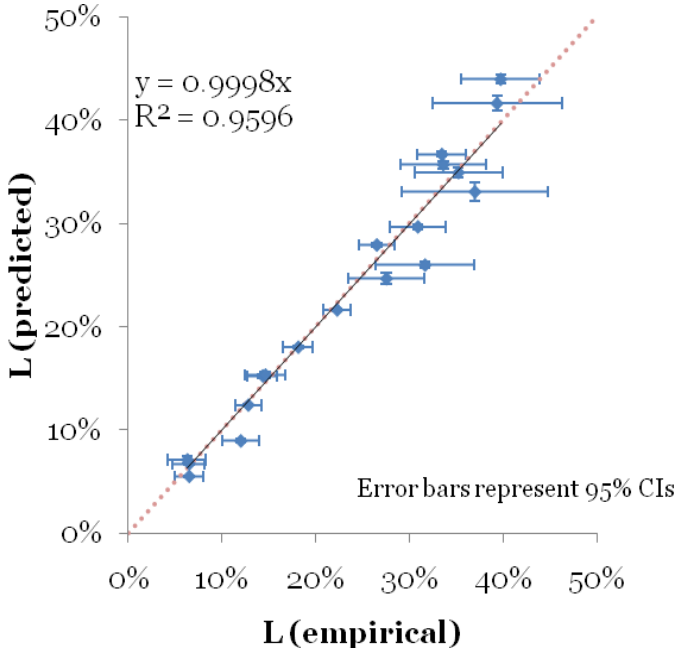


Table and Figure 2r: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Permutations of 10 mm tumor size bins and histological type

Group†		N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
Size	Histology						
1-10 mm	Ductal	53119	7.07%	(0.53%)	6.59%	(0.05%)	-0.49%
11-20 mm	Ductal	95130	14.67%	(0.51%)	14.63%	(0.05%)	-0.04%
21-30 mm	Ductal	44982	26.04%	(0.87%)	25.09%	(0.09%)	-0.94%
31-40 mm	Ductal	15876	34.00%	(1.54%)	34.13%	(0.16%)	0.14%
41-50 mm	Ductal	6813	39.18%	(2.37%)	41.35%	(0.25%)	2.17%
1-10 mm	Lobular	4225	7.57%	(2.14%)	6.81%	(0.17%)	-0.76%
11-20 mm	Lobular	9070	12.56%	(2.08%)	13.10%	(0.16%)	0.53%
21-30 mm	Lobular	4794	22.41%	(3.47%)	21.53%	(0.28%)	-0.89%
31-40 mm	Lobular	1907	38.49%	(7.86%)	29.53%	(0.47%)	-8.96%
41-50 mm	Lobular	1173	37.18%	(8.43%)	34.91%	(0.64%)	-2.28%
1-10 mm	Intraductal+LCIS	4913	6.84%	(2.58%)	6.03%	(0.16%)	-0.81%
11-20 mm	Intraductal+LCIS	10239	13.38%	(2.18%)	12.32%	(0.16%)	-1.06%
21-30 mm	Intraductal+LCIS	4783	21.83%	(3.92%)	20.80%	(0.29%)	-1.03%
31-40 mm	Intraductal+LCIS	1514	30.51%	(5.76%)	28.47%	(0.54%)	-2.04%
41-50 mm	Intraductal+LCIS	833	30.49%	(10.93%)	34.39%	(0.82%)	3.90%
1-10 mm	Tubular	775	9.36%	(3.15%)	6.37%	(0.39%)	-3.00%
11-20 mm	Tubular	1170	14.63%	(2.72%)	13.65%	(0.44%)	-0.98%
21-30 mm	Tubular	664	27.16%	(5.54%)	21.79%	(0.70%)	-5.37%
31-40 mm	Tubular	299	24.96%	(5.82%)	28.43%	(1.04%)	3.47%
41-50 mm	Tubular	170	31.84%	(8.21%)	35.77%	(1.42%)	3.93%
						Mean (std. dev.)	-0.72% (3.00%)
						Mean weighted by N (std. dev.)	-0.39% (0.91%)
						Root Mean Square (std. dev.)	3.01% (4.28%)
						Root Mean Square weighted by N (std. dev.)	0.97% (1.55%)

†all other permutation groups are not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

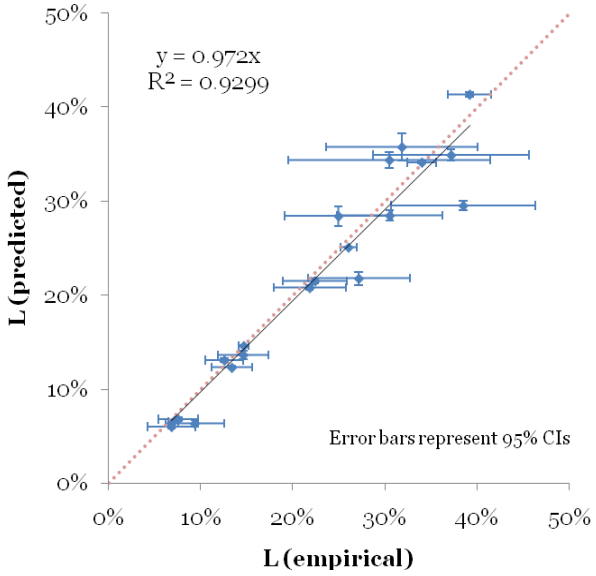


Table and Figure 2s: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Permutations of number of positive lymph nodes and estrogen/progesterone receptor status

Group†		N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
Nodes	ER/PR status						
0	ER+/PR+	106881	9.74%	(1.00%)	8.81%	(0.04%)	-0.93%
1	ER+/PR+	20980	19.56%	(1.96%)	18.36%	(0.09%)	-1.19%
2	ER+/PR+	9637	26.18%	(3.57%)	25.58%	(0.12%)	-0.60%
3	ER+/PR+	5446	31.56%	(4.78%)	32.15%	(0.16%)	0.59%
4	ER+/PR+	3369	33.45%	(4.28%)	37.62%	(0.19%)	4.17%
5	ER+/PR+	2369	40.57%	(6.90%)	42.64%	(0.21%)	2.07%
6	ER+/PR+	1742	41.08%	(6.04%)	47.11%	(0.24%)	6.03%
7	ER+/PR+	1318	50.92%	(8.27%)	51.56%	(0.25%)	0.64%
0	ER+/PR-	20228	13.15%	(2.10%)	12.02%	(0.11%)	-1.13%
1	ER+/PR-	3994	24.56%	(5.62%)	22.50%	(0.26%)	-2.06%
2	ER+/PR-	1817	28.79%	(4.50%)	29.96%	(0.36%)	1.17%
3	ER+/PR-	1040	33.66%	(6.01%)	36.50%	(0.46%)	2.84%
0	ER-/PR+	3754	11.97%	(1.68%)	12.52%	(0.24%)	0.55%
1	ER-/PR+	749	25.92%	(5.50%)	21.75%	(0.54%)	-4.18%
0	ER-/PR-	30538	15.62%	(1.04%)	17.53%	(0.10%)	1.91%
1	ER-/PR-	6119	30.28%	(3.59%)	26.66%	(0.22%)	-3.63%
2	ER-/PR-	2959	35.07%	(3.60%)	33.12%	(0.30%)	-1.94%
3	ER-/PR-	1781	42.41%	(4.23%)	39.25%	(0.36%)	-3.16%
4	ER-/PR-	1195	47.14%	(5.58%)	43.93%	(0.41%)	-3.21%
5	ER-/PR-	843	51.27%	(8.97%)	47.99%	(0.47%)	-3.28%
						Mean (std. dev.)	-0.27% (2.76%)
						Mean weighted by N (std. dev.)	-0.47% (2.47%)
						Root Mean Square (std. dev.)	2.70% (2.98%)
						Root Mean Square weighted by N (std. dev.)	2.45% (4.21%)

† the groups that are the combinations of 4,5,6,7 positive nodes with ER+/PR-, and 2,3,4, 5,6,7 positive nodes with ER-/PR+, and 6,7 positive nodes with ER-/PR-, are not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

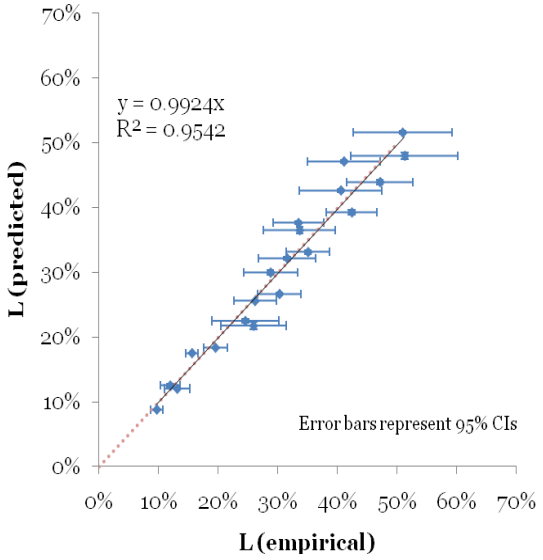


Table and Figure 2t: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Permutations of tumor grade and estrogen/progesterone receptor status

Group† ††		N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
Grade	ER/PR status						
1	ER+/PR+	33359	5.97%	(1.55%)	6.12%	(0.07%)	0.15%
2	ER+/PR+	67774	13.81%	(1.97%)	14.00%	(0.08%)	0.19%
3	ER+/PR+	32077	22.11%	(1.87%)	21.21%	(0.13%)	-0.90%
1	ER+/PR-	5260	9.38%	(4.57%)	6.89%	(0.19%)	-2.49%
2	ER+/PR-	11461	16.86%	(3.89%)	16.39%	(0.20%)	-0.48%
3	ER+/PR-	8407	25.80%	(2.92%)	25.34%	(0.28%)	-0.46%
						Mean (std. dev.)	-0.66% (0.99%)
						Mean weighted by N (std. dev.)	-0.21% (0.55%)
						Root Mean Square (std. dev.)	1.12% (1.56%)
						Root Mean Square weighted by N (std. dev.)	0.54% (0.67)

† grade 4 also exists in the dataset, but is not included in the calculation of the mean or displayed on graph; grade 4 no longer exists

†† all groups with ER- status are are not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

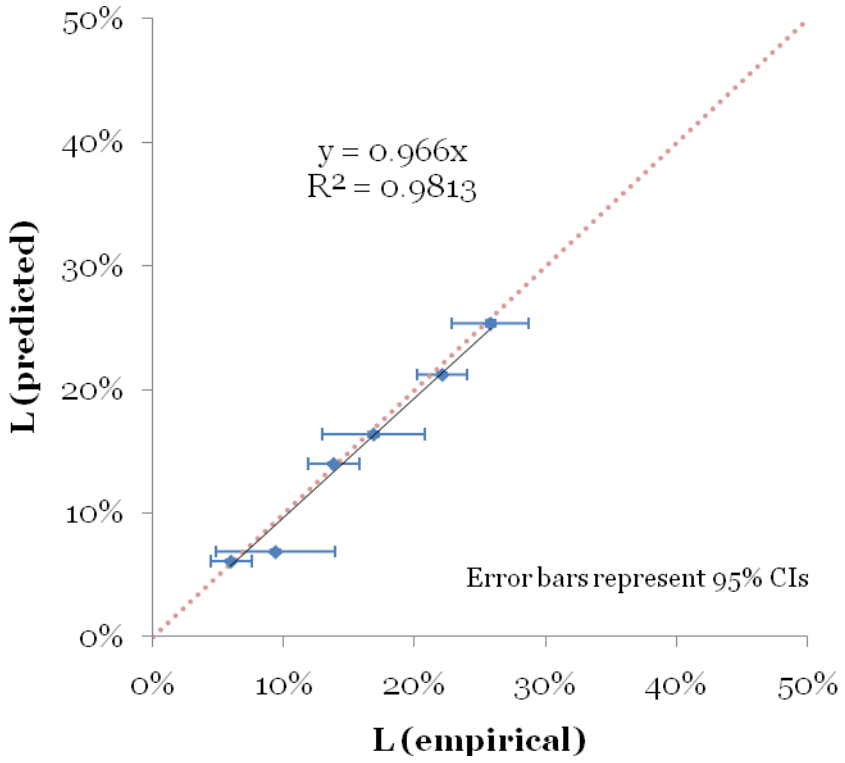


Table and Figure 2u: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Permutations of tumor grade and histological type

Group†		N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
Grade	Histology						
1	Ductal	33086	6.64%	(1.00%)	6.61%	(0.07%)	-0.03%
2	Ductal	83705	14.79%	(0.69%)	14.81%	(0.07%)	0.02%
3	Ductal	73999	23.45%	(0.68%)	23.79%	(0.09%)	0.34%
1	Lobular	3131	6.60%	(3.67%)	8.96%	(0.29%)	2.36%
2	Lobular	6511	14.54%	(5.10%)	16.21%	(0.26%)	1.67%
3	Lobular	2174	24.75%	(5.78%)	22.51%	(0.54%)	-2.24%
1	Intraductal+ LCIS	4001	6.62%	(3.80%)	7.37%	(0.24%)	0.75%
2	Intraductal+ LCIS	10507	11.77%	(2.09%)	14.27%	(0.20%)	2.50%
3	Intraductal+ LCIS	4600	24.74%	(4.72%)	20.80%	(0.36%)	-3.95%
						Mean (std. dev.)	0.16% (2.11%)
						Mean weighted by N (std. dev.)	0.23% (0.57)
						Root Mean Square (std. dev.)	2.00% (2.25%)
						Root Mean Square weighted by N (std. dev.)	0.58% (0.68%)

† all other permutation groups are not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

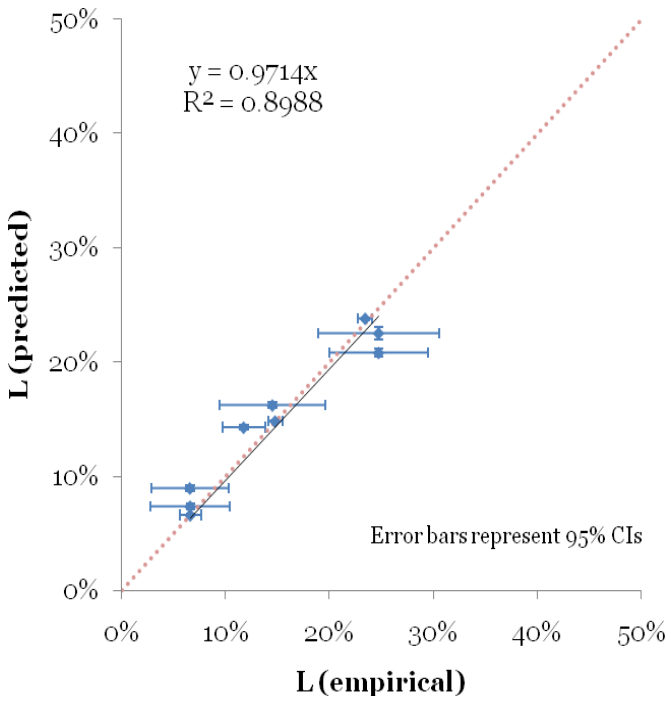


Table and Figure 2v: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Grouped by race

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
Black	22458	25.36%	(1.28%)	20.06%	(0.17%)	-5.30%
White	248716	16.01%	(0.33%)	15.93%	(0.05%)	-0.08%
Unknown/ Other	22402	13.59%	(0.98%)	16.61%	(0.16%)	3.02%
					Mean (std. dev.)	-0.79% (4.2%)
					Mean weighted by N (std. dev.)	-0.24% (0.95%)
					Root Mean Square (std. dev.)	3.52% (3.79%)
					Root Mean Square weighted by N (std. dev.)	0.82% (0.86%)

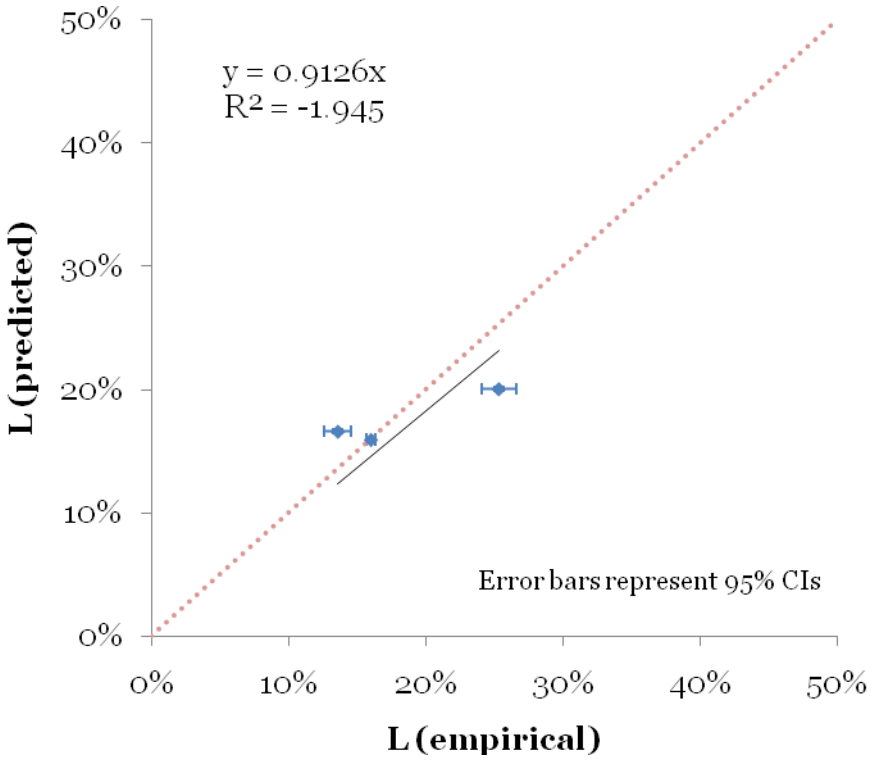
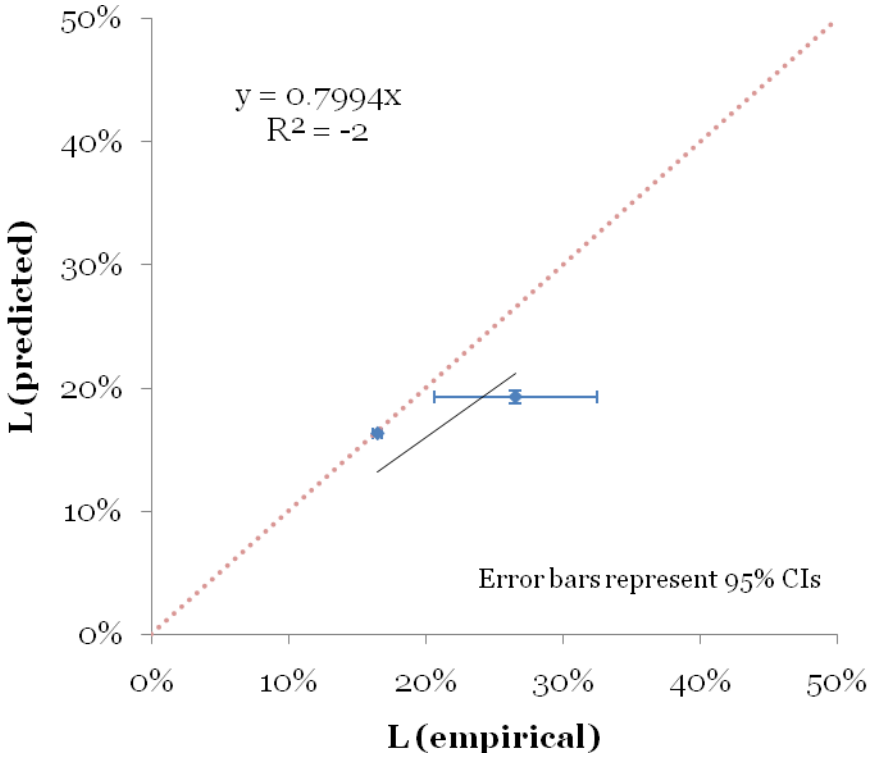


Table and Figure 2w: **Verification of the SNAP method on the SEER dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Grouped by sex

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
Female	291713	16.48%	(0.3%)	16.28%	(0.04%)	-0.20%
Male	1863	26.51%	(5.93%)	19.26%	(0.56%)	-7.25%
				Mean (std. dev.)		-3.73% (4.98%)
				Mean weighted by N (std. dev.)		-0.25% (0.22%)
				Root Mean Square (std. dev.)		5.13% (6.09%)
				Root Mean Square weighted by N (std. dev.)		0.29% (0.33%)



II Validation on the Partners dataset

Table and Figure 2aa: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 10% predicted lethality bins using the Size+Nodes+PrognosticFactors equation

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0-10%	2923	6.15%	(1.86%)	5.17%	(0.1%)	-0.98%
11-20%	1737	13.22%	(2.79%)	14.55%	(0.13%)	1.34%
21-30%	937	26.96%	(4.72%)	24.63%	(0.18%)	-2.33%
31-40%	499	37.36%	(7.32%)	34.12%	(0.24%)	-3.25%
40+%	319	39.61%	(8.56%)	47.34%	(0.65%)	7.73%
					Mean (std. dev.)	0.50% (4.39%)
					Mean weighted by N (std. dev.)	-0.29% (2.00%)
					Root Mean Square (std. dev.)	3.96% (4.99%)
					Root Mean Square weighted by N (std. dev.)	1.81% (1.11%)

†indiviual groups of 41-50%, 51-60%, … 81-90% not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

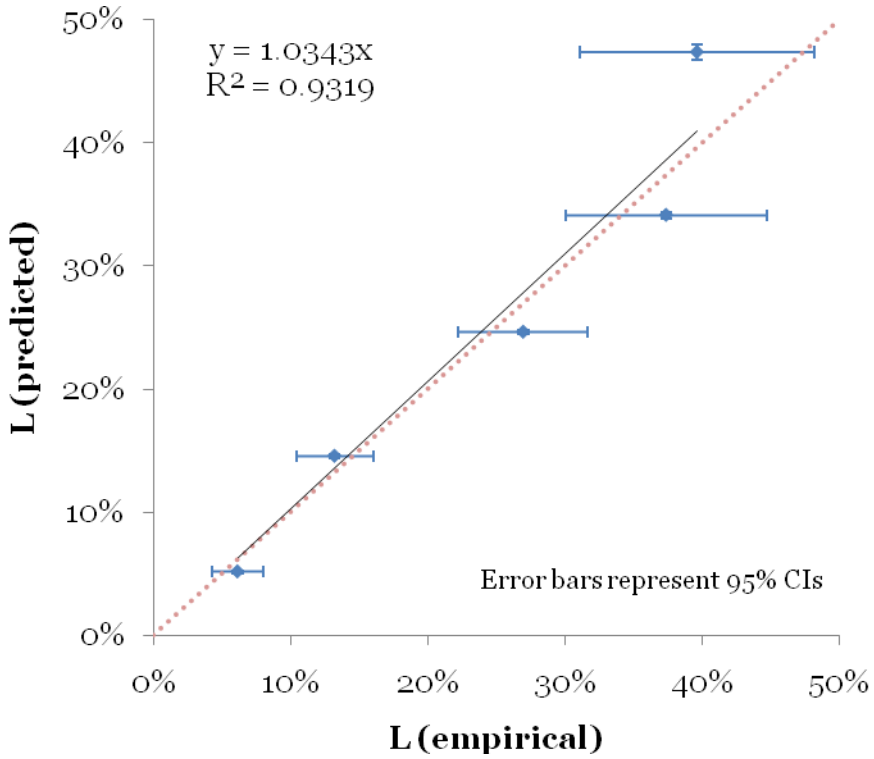


Table and Figure 2bb: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 5% predicted lethality bins using the Size+Nodes+PrognosticFactors equation

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0-5%	1402	4.47%	(2.98%)	2.72%	(0.07%)	-1.75%
6-10%	1521	7.39%	(2.36%)	7.42%	(0.07%)	0.03%
11-15%	959	12.38%	(4.79%)	12.39%	(0.09%)	0.00%
16-20%	778	14.79%	(3.7%)	17.23%	(0.1%)	2.44%
21-25%	507	22.64%	(6.47%)	22.36%	(0.12%)	-0.28%
26+%	1248	36.35%	(4.43%)	35.15%	(0.47%)	-1.20%
					Mean (std. dev.)	-0.13% (1.44%)
					Mean weighted by N (std. dev.)	-0.34% (1.4%)
					Root Mean Square (std. dev.)	1.32% (1.54%)
					Root Mean Square weighted by N (std. dev.)	1.32% (1.47%)

†individual groups of 26-30% … 81-85%, 86-90% not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

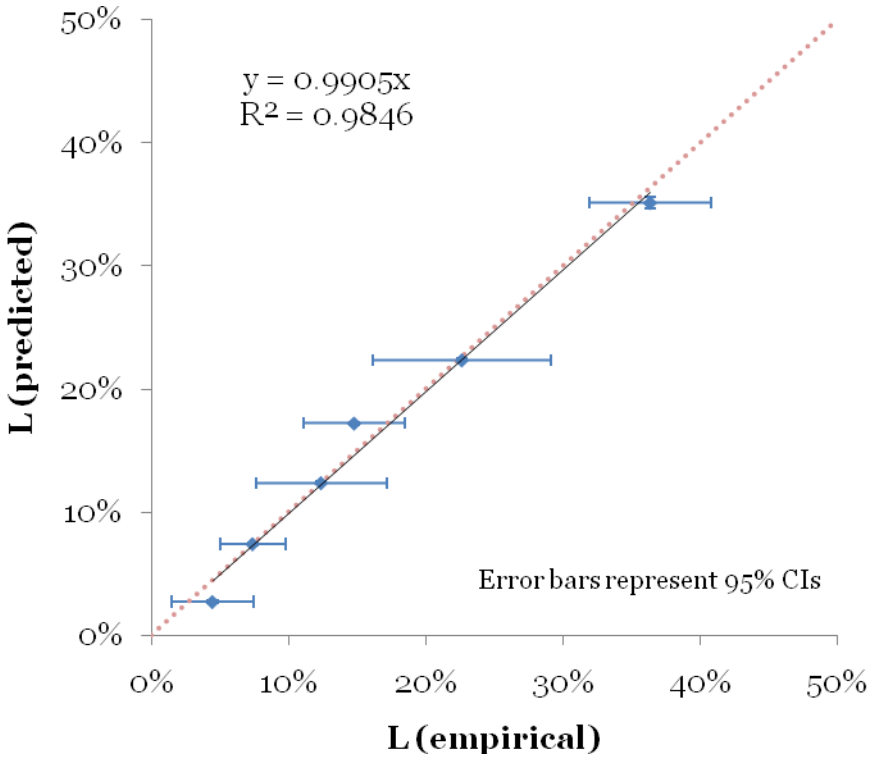


Table and Figure 2cc: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 2% predicted lethality bins using the Size+Nodes+PrognosticFactors equation

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0-2%	467	2.79%	(2.11%)	1.21%	(0.05%)	-1.57%
3-4%	646	4.39%	(5.6%)	3.01%	(0.04%)	-1.37%
5-6%	621	6.98%	(3.83%)	5.04%	(0.05%)	-1.95%
7-8%	575	11.15%	(6.3%)	7.04%	(0.05%)	-4.11%
9-10%	614	5.11%	(2.4%)	8.82%	(0.04%)	3.72%
11-12%	406	8.83%	(4.93%)	10.95%	(0.06%)	2.12%
13-14%	400	15.45%	(8.87%)	13.02%	(0.05%)	-2.44%
15-16%	403	10.76%	(4.02%)	15.22%	(0.06%)	4.46%
17-18%	276	18.62%	(8.4%)	17.01%	(0.06%)	-1.60%
19-20%	252	17.52%	(7.59%)	19.04%	(0.07%)	1.52%
21-22%	212	25.46%	(11.13%)	21.06%	(0.08%)	-4.41%
23-24%	222	19.71%	(7.69%)	22.88%	(0.07%)	3.17%
25+%	1321	36.01%	(4.44%)	34.56%	(0.46%)	-1.44%
					Mean (std. dev.)	-0.3% (2.95%)
					Mean weighted by N (std. dev.)	-0.53% (2.8%)
					Root Mean Square (std. dev.)	2.85% (2.65%)
					Root Mean Square weighted by N (std. dev.)	2.74% (2.8%)

†individual groups of 25-26%, 27-28% … 85-86%, 87-88% not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

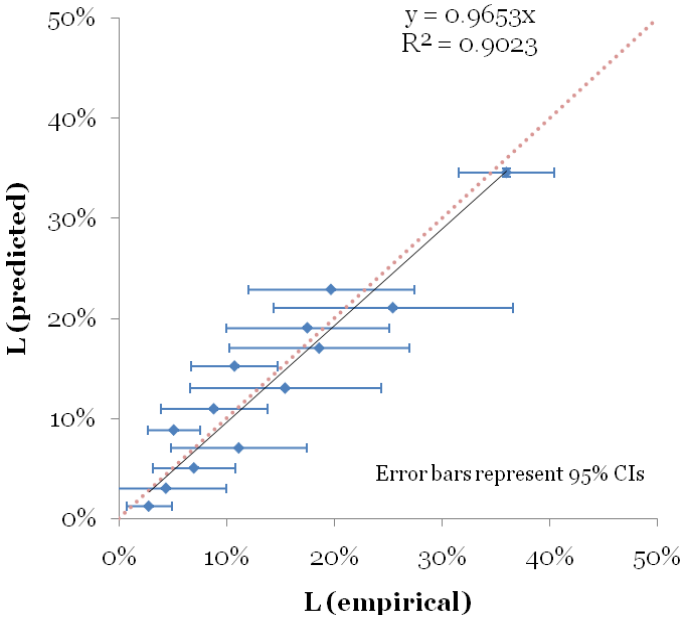


Table and Figure 2dd: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 20% predicted lethality percentiles using the Size+Nodes+PrognosticFactors equation

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0-4.5%	1282	4.04%	(2.97%)	2.53%	(0.06%)	-1.52%
4.6-8.5%	1285	7.67%	(2.58%)	6.68%	(0.07%)	-0.99%
8.6-14.9%	1283	10.64%	(4.07%)	11.44%	(0.1%)	0.80%
15-25%	1282	16.77%	(3.09%)	18.99%	(0.15%)	2.22%
26-77%	1282	36.62%	(4.49%)	34.84%	(0.46%)	-1.78%
					Mean (std. dev.)	-0.25% (1.71%)
					Mean weighted by N (std. dev.)	-0.25% (1.71%)
					Root Mean Square (std. dev.)	1.55% (1.32%)
					Root Mean Square weighted by N (std. dev.)	1.55% (1.32%)

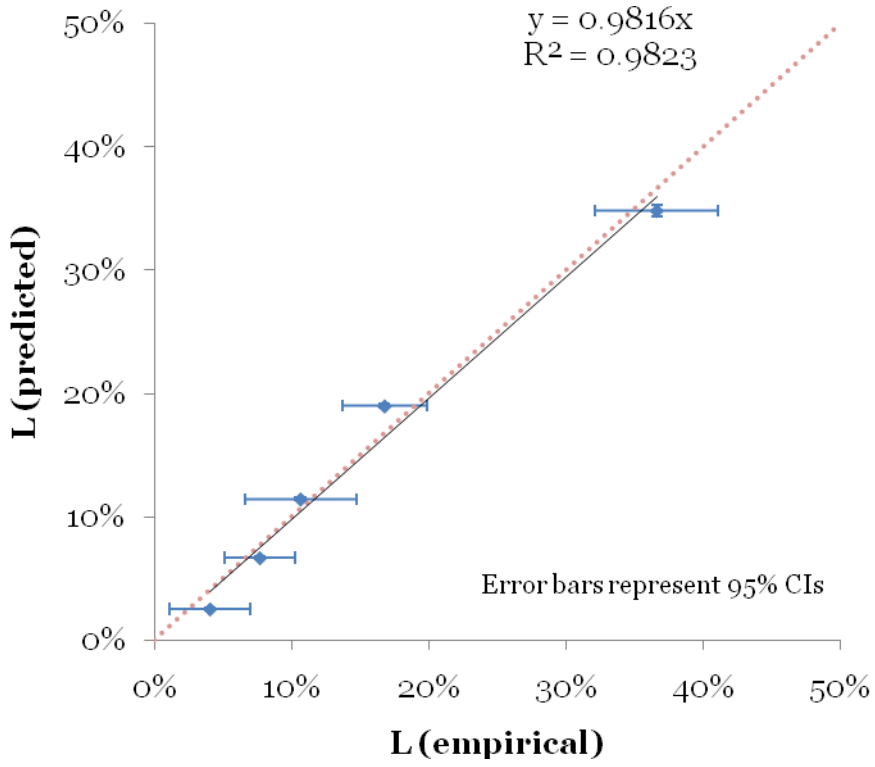


Table and Figure 2ee: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 10% predicted lethality percentiles using the Size+Nodes+PrognosticFactors equation

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0-2.6%	635	2.19%	(1.56%)	1.51%	(0.05%)	-0.68%
2.7-4.6%	647	5.37%	(5.03%)	3.53%	(0.04%)	-1.84%
4.7-6.7%	643	7.10%	(3.77%)	5.63%	(0.05%)	-1.47%
6.8-8.5%	642	8.03%	(3.22%)	7.73%	(0.04%)	-0.30%
8.6-11.3%	641	6.79%	(3.46%)	9.85%	(0.06%)	3.06%
11.4-14.9%	642	14.34%	(6.94%)	13.03%	(0.07%)	-1.31%
15.0-18.6%	640	14.09%	(3.81%)	16.54%	(0.08%)	2.45%
18.7-24.6%	642	19.87%	(5.01%)	21.44%	(0.13%)	1.56%
24.7-32.6%	642	34.54%	(6.4%)	28.25%	(0.18%)	-6.29%
32.7-77%	640	38.67%	(6.32%)	41.44%	(0.56%)	2.77%
					Mean (std. dev.)	-0.21% (2.84%)
					Mean weighted by N (std. dev.)	-0.21% (2.84%)
					Root Mean Square (std. dev.)	2.7% (3.43%)
					Root Mean Square weighted by N (std. dev.)	2.7% (3.43%)

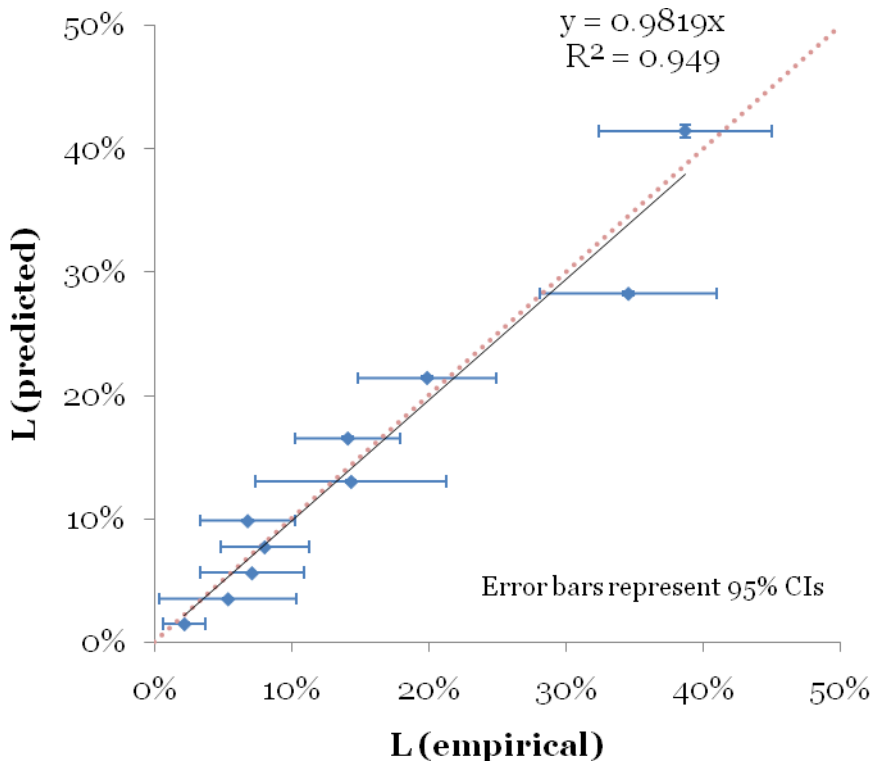


Table and Figure 2ff: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 10 mm tumor size bins

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
1-10 mm	2283	7.15%	(2%)	6.13%	(0.24%)	-1.03%
11-20 mm	2508	15.13%	(2.65%)	14.44%	(0.35%)	-0.69%
21-30 mm	1080	24.16%	(4.47%)	24.48%	(0.59%)	0.32%
31-40 mm	381	38.00%	(8.41%)	33.29%	(1.09%)	-4.71%
41-50 mm	163	32.33%	(11.4%)	38.35%	(1.74%)	6.02%
					Mean (std. dev.)	-0.02% (3.87%)
					Mean weighted by N (std. dev.)	-0.71% (1.15%)
					Root Mean Square (std. dev.)	3.47% (4.06%)
					Root Mean Square weighted by N (std. dev.)	1.25% (1.13%)

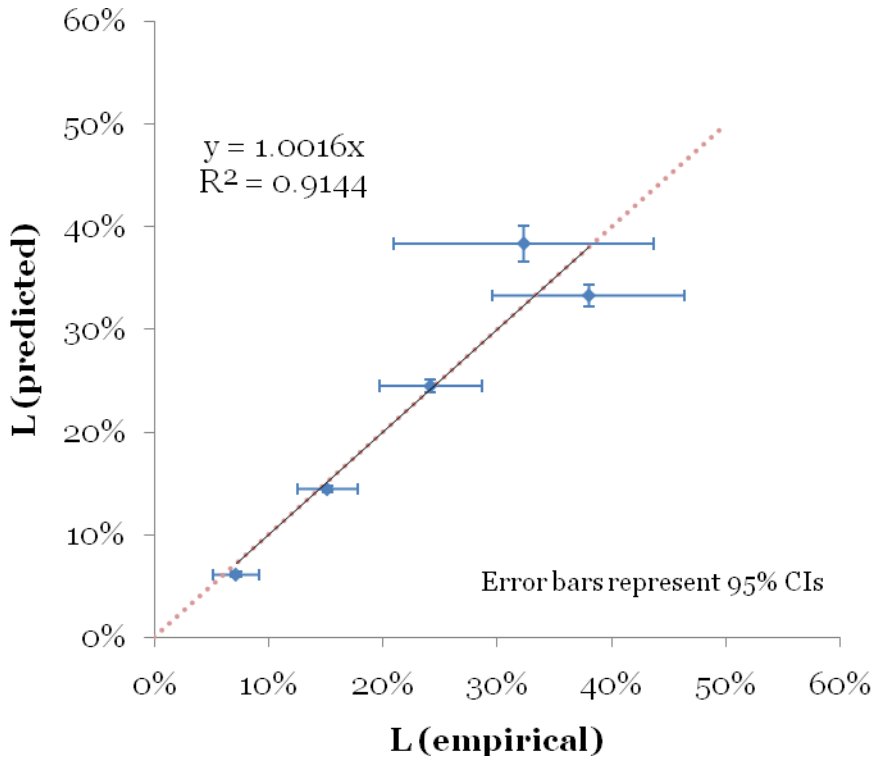


Table and Figure 2gg: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 5 mm tumor size bins

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
1-5 mm	666	7.70%	(7.31%)	3.30%	(0.37%)	-4.40%
6-10 mm	1617	7.50%	(2.15%)	7.29%	(0.29%)	-0.21%
11-15 mm	1493	11.58%	(2.84%)	12.14%	(0.4%)	0.56%
16-20 mm	1015	18.66%	(3.96%)	17.84%	(0.56%)	-0.82%
21-25 mm	640	26.01%	(8.57%)	23.05%	(0.74%)	-2.96%
26-30 mm	440	24.24%	(5.78%)	26.56%	(0.95%)	2.32%
					Mean (std. dev.)	-0.92% (2.43%)
					Mean weighted by N (std. dev.)	-0.71% (1.58%)
					Root Mean Square (std. dev.)	2.4% (2.74%)
					Root Mean Square weighted by N (std. dev.)	1.6% (1.84%)

†groups of 30-35, 35-40, 40-45, and 45-50mm are not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

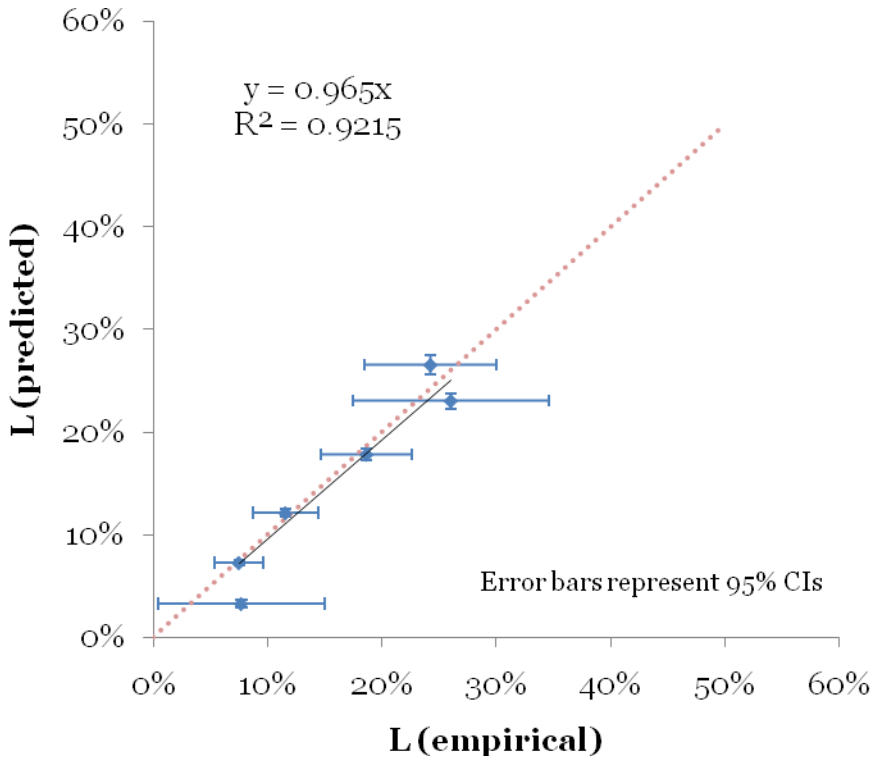


Table and Figure 2hh: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by 10% tumor size percentiles

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
1-6 mm	666	7.70%	(7.31%)	3.30%	(0.37%)	-4.40%
7-8 mm	716	5.76%	(2.98%)	5.72%	(0.38%)	-0.04%
8-10 mm	901	8.32%	(2.73%)	8.54%	(0.4%)	0.22%
11-12 mm	578	13.47%	(6.31%)	10.27%	(0.54%)	-3.20%
13-15 mm	375	7.07%	(4.08%)	12.36%	(0.84%)	5.29%
16-18 mm	656	12.38%	(3.66%)	14.22%	(0.65%)	1.84%
19-20 mm	899	19.26%	(4.09%)	18.17%	(0.6%)	-1.10%
21-27 mm	640	26.01%	(8.57%)	23.05%	(0.74%)	-2.96%
28-32 mm	440	24.24%	(5.78%)	26.56%	(0.95%)	2.32%
33-50 mm	544	35.95%	(6.74%)	34.81%	(0.95%)	-1.15%
					Mean (std. dev.)	-0.32% (2.92%)
					Mean weighted by N (std. dev.)	-0.61% (2.43%)
					Root Mean Square (std. dev.)	2.79% (3.05%)
					Root Mean Square weighted by N (std. dev.)	2.39% (2.54%)

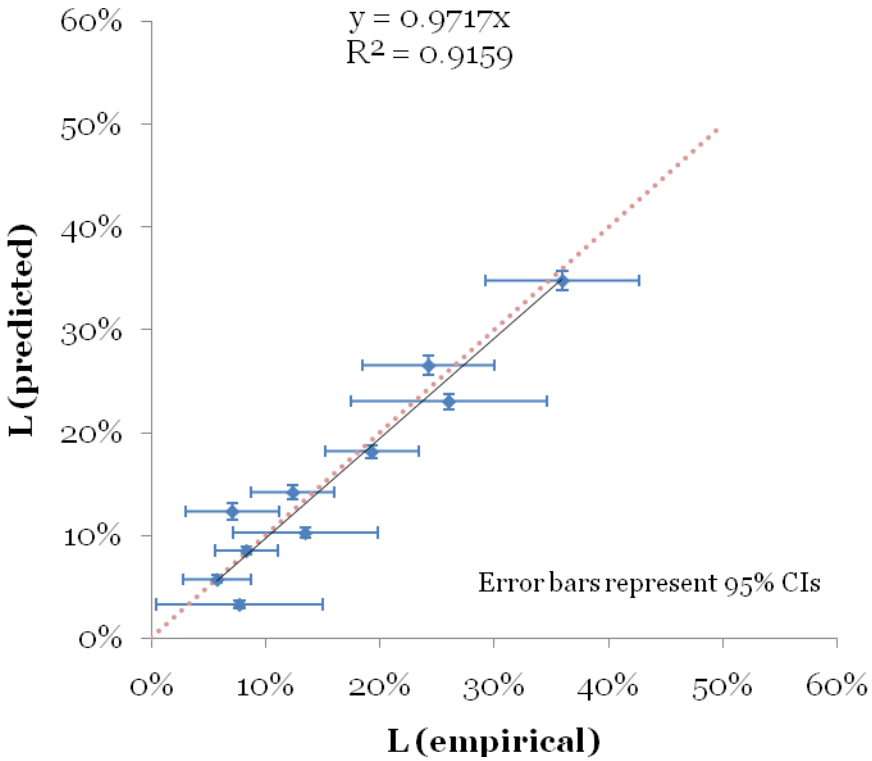


Table and Figure 2ii: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by number of positive lymph nodes

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
0	4441	10.31%	(1.64%)	9.53%	(0.22%)	-0.79%
1	917	21.48%	(4.92%)	18.83%	(0.48%)	-2.65%
2	452	29.05%	(7.55%)	26.88%	(0.63%)	-2.18%
3	254	36.03%	(10.74%)	33.05%	(0.83%)	-2.98%
4-7	351	36.23%	(8.48%)	44.08%	(0.88%)	7.85%
					Mean (std. dev.)	-0.81% (3.18%)
					Mean weighted by N (std. dev.)	-1.88% (3.18%)
					Root Mean Square (std. dev.)	3.06% (3.49%)
					Root Mean Square weighted by N (std. dev.)	3.49% (5.24%)

†the group of patients with 4,5,6 and 7 positive nodes is not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

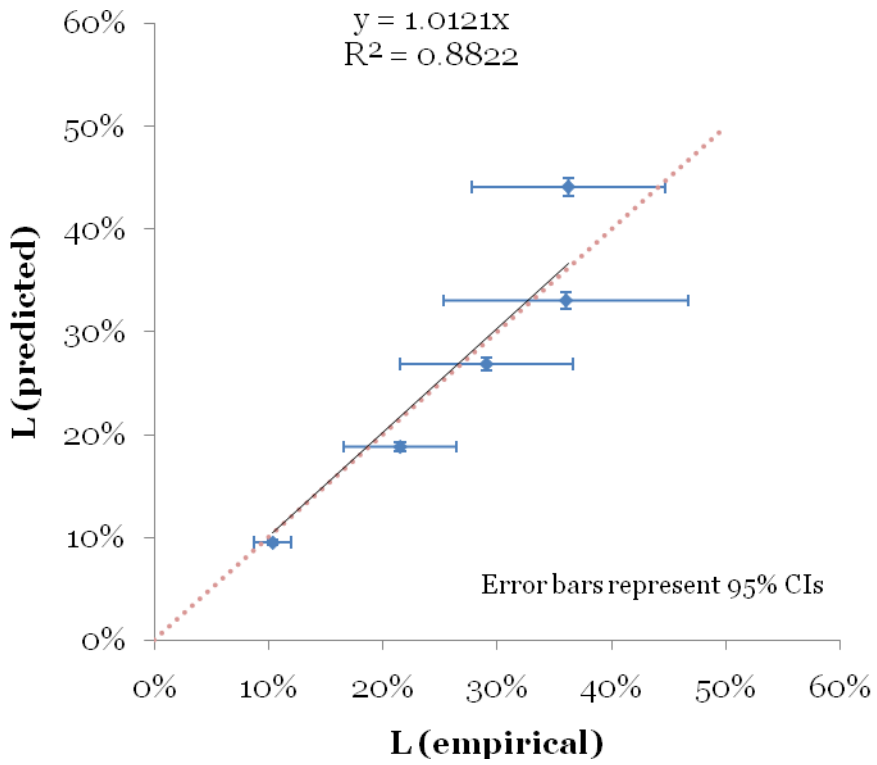


Table and Figure 2ijj: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by tumor grade

Group†	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
grade 1	1043	3.23%	(2.07%)	5.58%	(0.4%)	2.36%
grade 2	2427	14.77%	(3.22%)	13.14%	(0.41%)	-1.63%
grade 3	1821	22.87%	(3.6%)	21.62%	(0.59%)	-1.25%
					Mean (std. dev.)	-0.17% (2.2%)
					Mean weighted by N (std. dev.)	-0.71% (1.89%)
					Root Mean Square (std. dev.)	1.8% (1.44%)
					Root Mean Square weighted by N (std. dev.)	1.7% (1.37%)

†grade 4 also exists in the dataset, but is not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group; in addition, grade 4 no longer exists

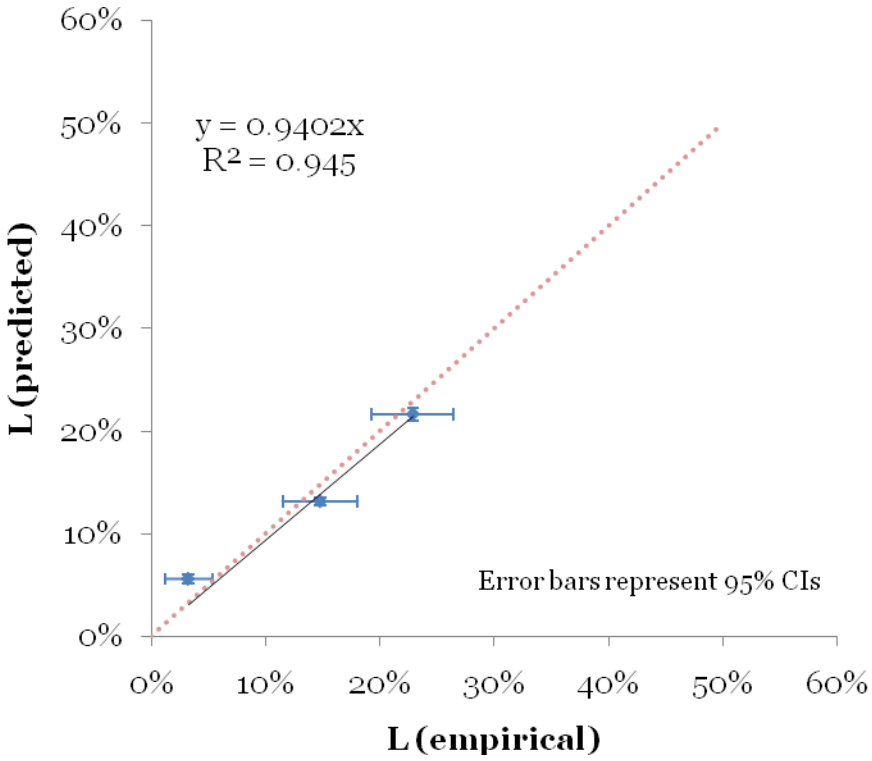


Table and Figure 2ikk: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by estrogen receptor status

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
Positive	3681	11.96%	(2.59%)	13.23%	(0.37%)	1.27%
Negative	993	26.12%	(12.04%)	20.75%	(0.84%)	-5.37%
					Mean (std. dev.)	-2.05% (4.69%)
					Mean weighted by N (std. dev.)	-0.14% (3.02%)
					Root Mean Square (std. dev.)	3.9% (4.39%)
					Root Mean Square weighted by N (std. dev.)	2.14% (0.93%)

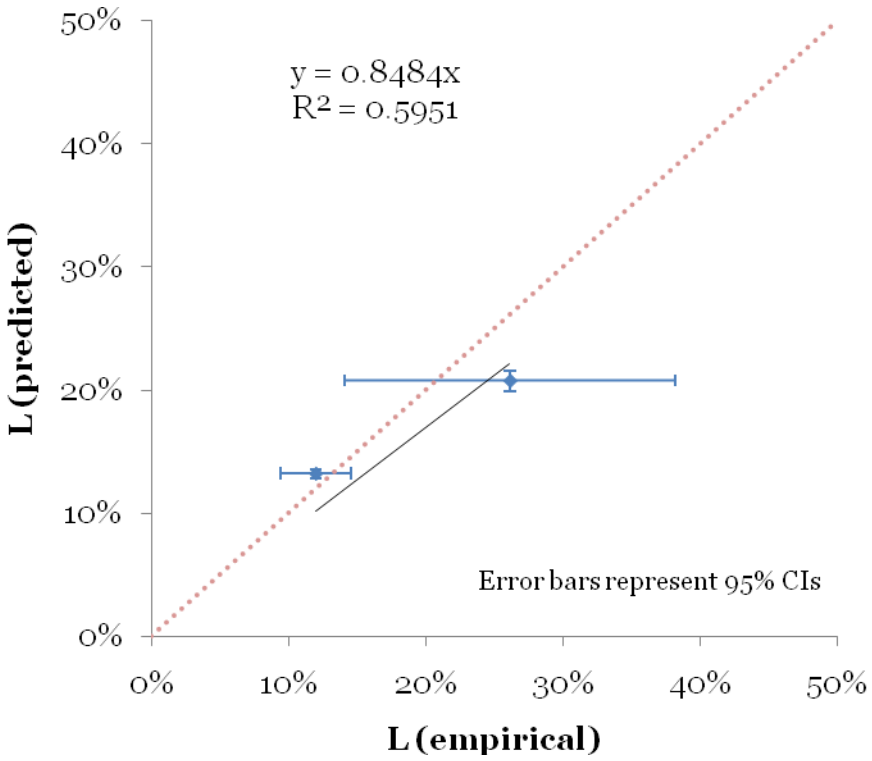


Table and Figure 2ill: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Patients grouped by histological type

Group	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
Ductal	4909	16.59%	(1.78%)	15.57%	(0.34%)	-1.01%
Lobular	448	9.59%	(4.95%)	13.93%	(1.02%)	4.34%
Ductal and Lobular	722	10.88%	(6.32%)	13.65%	(0.88%)	2.77%
Mucinous	81	7.56%	(7.48%)	5.24%	(1.37%)	-2.33%
					Mean (std. dev.)	0.94% (3.13%)
					Mean weighted by N (std. dev.)	-0.2% (2.13%)
					Root Mean Square (std. dev.)	2.87% (2.75%)
					Root Mean Square weighted by N (std. dev.)	1.85% (2.18%)

†the other histology groups are not included in the calculation of the mean or displayed on graph; 95% CI is greater than 20% or there is insufficient follow-up data for the group

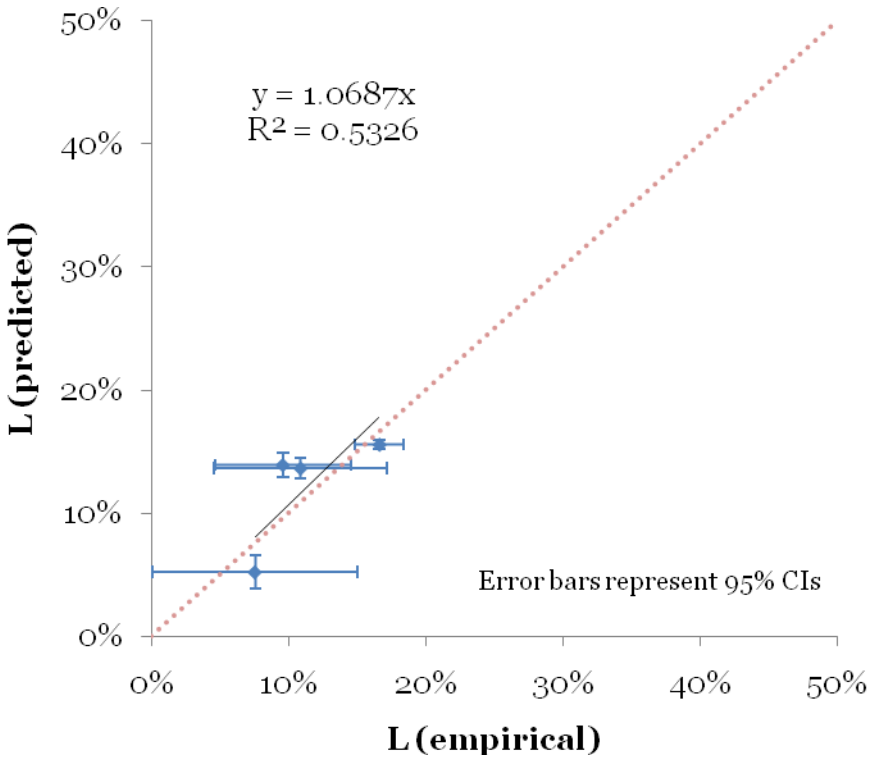
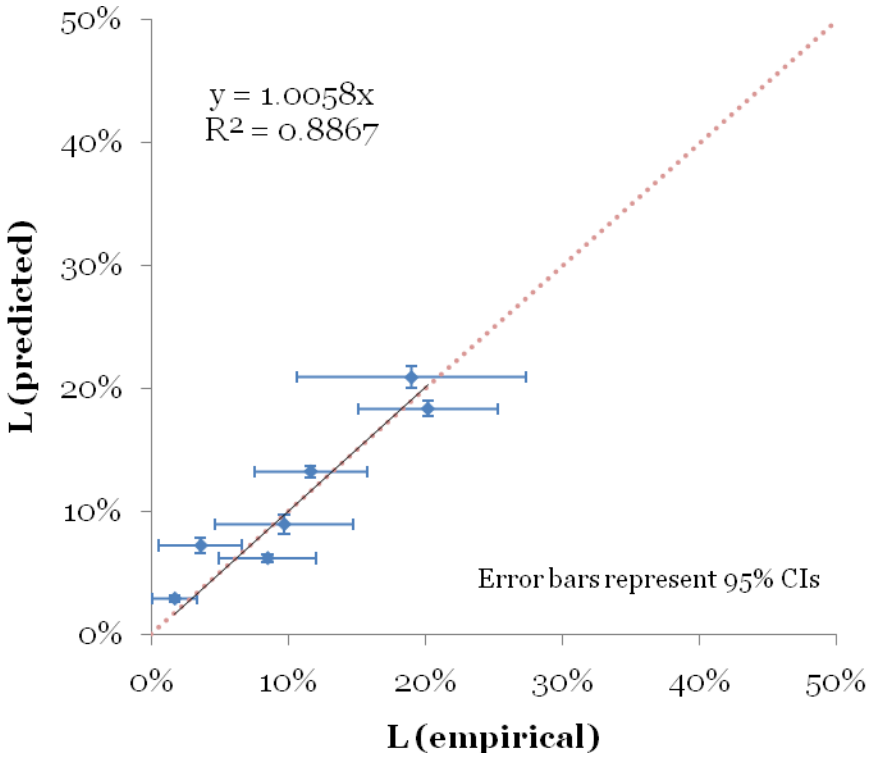


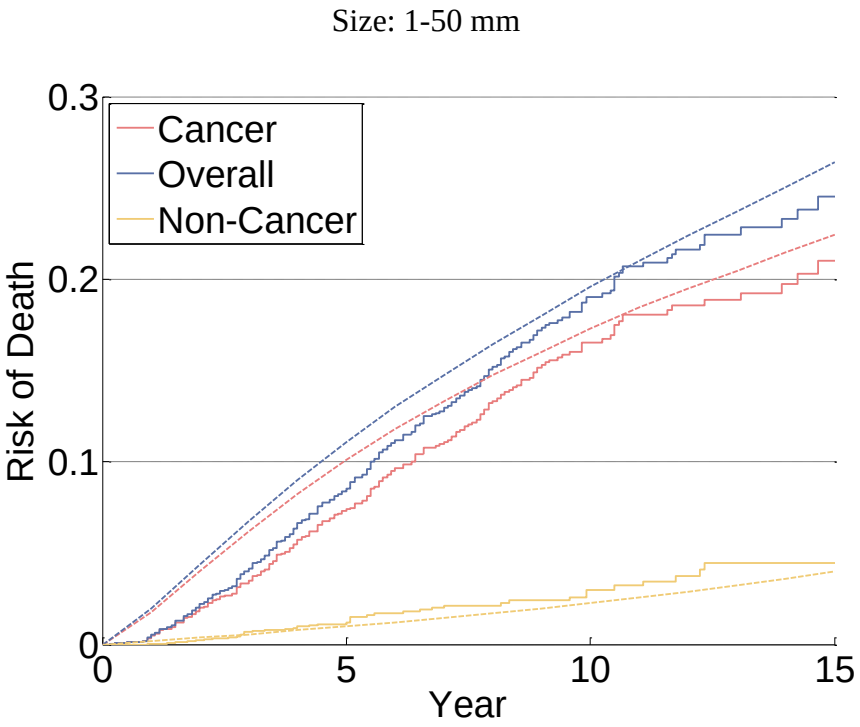
Table and Figure 2mm: **Verification of the SNAP method on the Partners dataset** (patients after 1987, parameters for the outcome calculator, see TABLE I). Permutations of 10 mm tumor size bins and tumor grade

Group† Size	Grade	N	L empirical (SEM)		L predicted (SEM)		Difference (pred – emp)
1-10 mm	1	571	1.69%	(1.65%)	2.89%	(0.26%)	1.20%
11-20 mm	1	373	3.60%	(3.02%)	7.22%	(0.66%)	3.62%
1-10 mm	2	861	8.49%	(3.57%)	6.19%	(0.36%)	-2.30%
11-20 mm	2	1034	11.63%	(4.09%)	13.24%	(0.49%)	1.61%
21-30 mm	2	369	18.99%	(8.35%)	20.91%	(0.87%)	1.92%
1-10 mm	3	387	9.69%	(5.05%)	8.94%	(0.74%)	-0.75%
11-20 mm	3	725	20.19%	(5.09%)	18.36%	(0.63%)	-1.84%
						Mean (std. dev.)	0.49% (2.17%)
						Mean weighted by N (std. dev.)	0.19% (2.22%)
						Root Mean Square (std. dev.)	2.07% (2.04%)
						Root Mean Square weighted by N (std. dev.)	2.06% (1.91%)

† other permutation groups are also not included, because 95% CI is greater than 20% or there is insufficient follow-up data for the group; in addition, grade 4 no longer exists



A



B

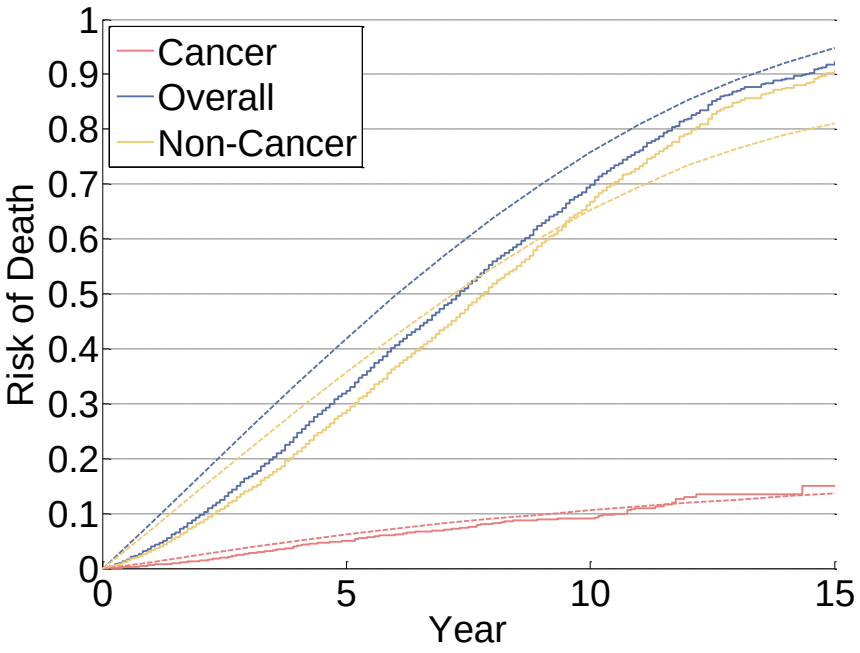


Figure 3. Comparison of predicted vs. observed mortality curves for two representative cases. Dashed lines represent calculator output; solid lines represent empirical Kaplan-Meier curves. **(A)** Patient group 1 (N=2,821): year of diagnosis 1988+, age 31-50 (mean: 43.61), tumor size 11-20 mm (mean: 16.36 mm), 2 positive nodes. Calculator input: age 44, size 1.6 cm, 2 nodes. **(B)** Patient group 2 (N=7,301): year of diagnosis 1988+, age 81-90 (mean: 83.80), tumor size 11-20 mm (mean: 15.57 mm), 0 positive nodes. Calculator input: age 84, size 1.6 cm, 0 nodes.

CODE IN THE TREATMENT CALCULATOR

The javascript code performing the life expectancy and mortality calculations for www.cancermath.net is provided below. The code begins by initializing several lengthy arrays (such as the life expectancy tables)—the code corresponding to Step 1 described above begins at the top of page 20.

```
/*~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
  This web calculator estimates the risk of breast cancer death,
  and the impact that adjuvant treatment will have on that risk of death.
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~*/
/*~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
   © James Michaelson PhD
   May be freely used for any scientific purpose; For permission to use commercially or in a
website,
   contact Dr. Michaelson at james.michaelson@gmail.com
   For the mathematical essentials used below, see:
     Michaelson JS et al. "The effect of tumor size and lymph node status on breast carcinoma
lethality"
     Cancer. 2003;98:2133-2143.
  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~*/


/*********************************************************************
* Array of probability of dying between year x and x+1, where x is age, starting at age x=0,
taken from:
*       -National Vital Statistics Reports Vol 54 No 14, April 19, 2006, United States Life
Tables 2003,
*       Table 3. Life table for females: United States, 2003.
* adjusted to exclude the probability of dying from breast cancer using data from:
*  -National Vital Statistics Reports Vol 55, No 19, August 21, 2007, Deaths: Final Data for
2004,
*  Table 3. Number of deaths and death rates by age, race, and sex: United States, 2004
*  Table 10. Number of deaths from 113 selected causes by age: United States, 2004
* q(x)
*********************************************************************************/
var nvsr_death_prob_yearly = new Array(0, 0.006083, 0.000414, 0.000301, 0.000218, 0.000172,
0.000158, 0.000141, 0.000128, 0.000117, 0.000108, 0.000105, 0.000110, 0.000132, 0.000173,
0.000228, 0.000292, 0.000355, 0.000407, 0.000440, 0.000456, 0.000471, 0.000489, 0.000501,
0.000509, 0.000515, 0.000506, 0.000516, 0.000531, 0.000552, 0.000578, 0.000609, 0.000646,
0.000693, 0.000753, 0.000827, 0.000864, 0.000949, 0.001047, 0.001157, 0.001273, 0.001393,
0.001514, 0.001639, 0.001774, 0.001919, 0.002045, 0.002211, 0.002383, 0.002560, 0.002746,
0.002949, 0.003176, 0.003431, 0.003718, 0.004039, 0.004462, 0.004859, 0.005304, 0.005809,
0.006384, 0.007060, 0.007817, 0.008596, 0.009350, 0.010100, 0.011305, 0.012297, 0.013426,
0.014706, 0.016123, 0.017603, 0.019196, 0.021033, 0.023175, 0.025585, 0.028661, 0.031295,
0.034315, 0.037906, 0.042094, 0.046670, 0.051554, 0.057062, 0.063411, 0.070761, 0.079054,
0.087065, 0.095796, 0.105294, 0.115605, 0.126771, 0.138833, 0.151829, 0.165787, 0.180734,
0.196684, 0.213644, 0.231608, 0.250560, 0.270467,
0.295466253,0.317857963,0.341365799,0.365989493,0.39172287,0.418553461,0.44646215,0.475422854,0.5
0540224,0.53635949,0.568246115,0.601005825,0.634574456,0.66887997,0.703842512,0.73937455,0.775381
082,0.811759916,0.848402039,0.88519205,0.92200867,0.958725335,0.995210852);

/*********************************************************************
* A 15-part step function (L_breastcancer_distribution) to represent the fraction
* of deaths that occurs in each of the 15 years after diagnosis.
* This is the probability density function of the SEER cohort from which we derived our
parameters,
* normalized to 1 at 15 years.
*********************************************************************************/
var L_breastcancer_distribution = new Array(0.031545, 0.076709, 0.094483, 0.091899, 0.083154,
0.079074, 0.073931, 0.069022, 0.066901, 0.059433, 0.061200, 0.058224, 0.054612, 0.052216,
0.047595);

/*********************************************************************
*  Expectation of life in years at age x, starting at age x=0, taken from:
*       -National Vital Statistics Reports Vol 54 No 14, April 19, 2006, United States Life
Tables 2003,
*       Table 3. Life table for females: United States, 2003.
* adjusted to exclude the probability of dying from breast cancer using data from:
*  -National Vital Statistics Reports Vol 55, No 19, August 21, 2007, Deaths: Final Data for
2004,
*  Table 3. Number of deaths and death rates by age, race, and sex: United States, 2004
*  Table 10. Number of deaths from 113 selected causes by age: United States, 2004
```

```
* e(x)
*************************************************************************/
var nvsr_life_expect = new Array(80.5, 80.0, 79.0, 78.0, 77.1, 76.1, 75.1, 74.1, 73.1, 72.1,
71.1, 70.1, 69.1, 68.1, 67.2, 66.2, 65.2, 64.2, 63.2, 62.3, 61.3, 60.3, 59.4, 58.4, 57.4, 56.4,
55.5, 54.5, 53.5, 52.6, 51.6, 50.6, 49.7, 48.7, 47.7, 46.8, 45.8, 44.8, 43.9, 42.9, 42.0, 41.1,
40.1, 39.2, 38.3, 37.3, 36.4, 35.5, 34.6, 33.6, 32.7, 31.8, 30.9, 30.0, 29.2, 28.3, 27.4, 26.5,
25.7, 24.8, 24.0, 23.1, 22.3, 21.5, 20.7, 19.9, 19.1, 18.4, 17.6, 16.9, 16.1, 15.4, 14.7, 14.0,
13.3, 12.7, 12.0, 11.4, 10.8, 10.2, 9.6, 9.0, 8.5, 8.0, 7.5, 7.0, 6.6, 6.2, 5.8, 5.4, 5.0, 4.7,
4.4, 4.1, 3.8, 3.5, 3.3, 3.1, 2.8, 2.6, 2.5, 2.3, 2.1, 2.0, 1.8, 1.7, 1.5, 1.4, 1.2, 0.8, 0.4,
0.2, 0.1, 0.05, 0.03, 0.01, 0.006, 0.003, 0.001, 0.002, 0.0008);



//The following function performs all of the the web calculations
function updateGraph(){

if ( typeof updateGraph.counter == 'undefined' ) { //show the disclaimer the first time function
is called
        tb_show(null,
"/cancer/limitations.html?placeValuesBeforeTB_=savedValues&TB_iframe=true&height=380&width=330&mo
dal=true", false);
        updateGraph.counter = 1;
     }

/*********************************************************************
STEP 1  collect user input
Acquire all the user inputs and assigns the appropriate j_primary and g-values to:
Age; Tumor Diameter; Nodal Status; # of Positive Nodes; ER Status; PR Status;
HERS Status; Histologic Type; Grade; Endocrine Therapy Type (if applicable); Chemotheraphy Type
(if applicable)
*************************************************************************/
age = document.form.ageInput.value*1;
dia = document.form.diameter.value*1;
nnum=document.form.nodenumber.value*1;
er =document.form.erstatus.value*1;
pr =document.form.prstatus.value*1;
her =document.form.her2status.value*1;
his =document.form.histology.value*1;
grade =document.form.grade.value*1;
endo=document.form.endotherapy.value*1;
chemo=document.form.chemotherapy.value*1;



/*********************************************************************
STEP 2 calculate yearly and cumulative cancer and total death rates
*************************************************************************/
    /*********************************************************************
    Initizes all arrays necessary for calculations
    * L_cancer_percentage: contains the distribution of breast cancer lethality for 15 years
    * L_cancer_death_yearly: describes to the total number of deaths due to breast cancer at a
single year
    * L_cancer_death_cumm: describes the cummulative number of deaths due to breast cancer for a
given year and all its previous years
    * L_noncancer_prob: contains the probability of survival from the nsvr data adjusted to
exclude the probability of dying from breast cancer as mentioned previously
    * L_noncancer_death_yearly: describes to the total number of deaths due to non-breast cancer
causes at a single year
    * L_noncancer_death_cumm: describes the cummulative number of deaths due to non-breast cancer
causes for a given year and all its previous years
    * L_overall_death_yearly: describes to the total number of deaths due to all causes at a
single year
    * L_overall_death_cumm: describes the cummulative number of deaths due to all causes for a
given year and all its previous years
    * remaining_percentage: describes the percentage of the initial population surviving to a
specific year
    * calc_num_surviving: calculated number of women with breast cancer surviving
    * calc_person_years_lived_between_years: calculated number of person years lived by women
with breast cancer between two consecutive years
```

```
    * calc_total_num_of_person_years_lived_above: caculated number of person years lived by women
with breast cancer above a certain age
    ************************************************************************/

    var L_breastcancer_percentage=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_nonbreastcancer_prob=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_breastcancer_death_yearly=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_cancer_death_cumm=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_nonbreastcancer_death_yearly=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_noncancer_death_cumm=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_overall_death_yearly=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_overall_death_cumm=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var remaining_percentage=new Array(1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var cancer_death_hazard = new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var cancer_death_dist_cumm = new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);

    var L_cancer_percentage_therapy=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_noncancer_prob_therapy=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_cancer_death_yearly_therapy=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_cancer_death_cumm_therapy=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_noncancer_death_yearly_therapy=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_noncancer_death_cumm_therapy=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_overall_death_yearly_therapy=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var L_overall_death_cumm_therapy=new Array(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);
    var remaining_percentage_therapy=new Array(1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0);


    /*********************************************************************
    *Initialize Qs, R, and j_primary variables depending on if treatments are selected
    * Variable       Standard Therapy         No Therapy
    *   Qs            0.010054                0.014751
    ************************************************************************/
    Qs=0.014751;
    j_primary=0.8057;
    Z=1; //initializes Z value
    R=0.07581; //initializes R value
    endotherapyEffect=0; //Resets endotherapy effect before every calculation to no endotherapy
    chemotherapyEffect=0; //Resets endotherapy effect before every calculation to no chemotherapy
    totaltherapyEffect=0; //Resets total therapy effect before every calculation to no therapy

    /*********************************************************************
    * If nodal status is unknown, j_primary is set to 1
    ************************************************************************/
    if (document.form.nodesKnown.value *1 == 0)  {
       j_primary=1;
       nnum=0;
    } else {
       nnum=nnum;
    }
/*********************************************************************
 STEP 2.a      The program loads the g parameters determined by the user input, and computes the
product of all of them
************************************************************************/

    g_parameter = 1; //Resets g_parameter before every calculation

    /*********************************************************************
    * Multiply g-parameter with g-value based on age
    ************************************************************************/
    if(age == 0) {
       g_parameter=g_parameter*1; //if user age is 0 then g-value is 1
      } else if(age >= 21 && age <= 30) {
        g_parameter=g_parameter*1.2035;     //if user age is 21-30 then g-value is 1.2035
      } else if(age >= 31 && age <= 40) {
         g_parameter=g_parameter*1.0705;  //if user age is 31-40 then g-value is 1.0705
      } else if(age >= 41 && age <= 50) {
         g_parameter=g_parameter*0.85655;  //if user age is 41-50 then g-value is 0.85655
      } else if(age >= 51 && age <= 60) {
         g_parameter=g_parameter*1.0228;  //if user age is 51-60 then g-value is 1.0228
      } else if(age >= 61 && age <= 70) {
         g_parameter=g_parameter*1.0248;  //if user age is 61-70 then g-value is 1.0248
```

```
   } else if(age >= 71 && age <= 80) {
      g_parameter=g_parameter*1.01945;  //if user age is 71-80 then g-value is 1.01945
   } else if(age >= 81 && age <= 90) {
      g_parameter=g_parameter*1.17735;  //if user age is 81-90 then g-value is 1.17735
   } else if(age >= 91 && age <= 100) {
      g_parameter=g_parameter*1.32845;  //if user age is 91-100 then g-value is 1.32845
   } else {
      g_parameter=g_parameter*1; //if user age is <=20 or >=101 then g-value is 1
   }

/*****************************************************************
* Multiply g-parameter with g-value based on ER/PR status
*****************************************************************/
switch (er) {
    case 0: //ER status Unknown
           if (pr==0){ //PR Unknown
                        //if ER is UNKNOWN and PR is UNKNOWN then g-value is 1
                        g_parameter=g_parameter*1;
           } else if(pr==1) { //PR Positive
                        //if ER is UNKNOWN and PR is POSTIVE then g-value is 0.9166
                        g_parameter=g_parameter*0.9166;
           } else if(pr==2) { //PR Negative
                        //if ER is UNKNOWN and PR is NEGATIVE then g-value is 1.1701
                        g_parameter=g_parameter*1.1701;
           }
    break;
    case 1: //ER status Positive
           if (pr==0){ //PR Unknown
                        //if ER is POSITIVE and PR is UNKNOWN then g-value is 0.953
                        g_parameter=g_parameter*0.953;
           } else if(pr==1) { //PR Positive
                        //if ER is POSITIVE and PR is POSITIVE then g-value is 0.91685
                        g_parameter=g_parameter*0.91685;
           } else if(pr==2) { //PR Negative
                        //if ER is POSITIVE and PR is NEGATIVE then g-value is 1.15415
                        g_parameter=g_parameter*1.15415;
           }
    break;
    case 2: //ER status Negative
           if (pr==0){ //PR Unknown
                        //if ER is NEGATIVE and PR is UNKNOWN then g-value is 1.1753
                        g_parameter=g_parameter*1.1753;
           } else if(pr==1) { //PR Positive
                        //if ER is NEGATIVE and PR is POSITIVE then g-value is 1.0131
                        g_parameter=g_parameter*1.0131;
           } else if(pr==2) { //PR Negative
                        //if ER is NEGATIVE and PR is NEGATIVE then g-value is 1.1904
                        g_parameter=g_parameter*1.1904;
           }
    break;
}  //end Switch for erpr

/*****************************************************************
* Multiply g-parameter with g-value based on HER2 status
*****************************************************************/
switch (her) {
    case 0:
           //if HER2 is UNKNOWN then g-value is 1
           g_parameter=g_parameter*1;
    break;
    case 1:
           //if HER2 is POSITIVE then g-value is 1.515
           g_parameter=g_parameter*1.515;
    break;
    case 2:
           //if HER2 is NEGATIVE then g-value is 0.9662
           g_parameter=g_parameter*0.9662;
    break;
}  //end switch for HER2

/*****************************************************************
```

```
* Multiply g-parameter with g-value based on histology status
*************************************************************************/
switch (his) {
    case 0:
           //if HISTOLOGY is UNKNOWN then g-value is 1
           g_parameter=g_parameter*1;
    break;
      case 1:
           //if HISTOLOGY is DUCTAL then g-value is 1.04495
           g_parameter=g_parameter*1.04495;
    break;
    case 2:
           //if HISTOLOGY is LOBULAR then g-value is 0.97825
           g_parameter=g_parameter*0.97825;
    break;
    case 3:
           //if HISTOLOGY is DUCTALandLOBULAR then g-value is 0.8624
           g_parameter=g_parameter*0.8624;
    break;
    case 4:
           //if HISTOLOGY is MUCINOUS then g-value is 0.42355
           g_parameter=g_parameter*0.42355;
    break;
      case 5:
           //if HISTOLOGY is MEDULLARY then g-value is 0.55305
           g_parameter=g_parameter*0.55305;
    break;
    case 6:
           //if HISTOLOGY is TUBULAR then g-value is 0.2639
           g_parameter=g_parameter*0.2639;
    break;
    case 7:
           //if HISTOLOGY is COMEDO then g-value is 0.84305
           g_parameter=g_parameter*0.84305;
    break;
    case 8:
           //if HISTOLOGY is SCIRRHOUS then g-value is 1.51235
           g_parameter=g_parameter*1.51235;
    break;
    case 9:
           //if HISTOLOGY is INFLAMMATORY then g-value is 3.1544
           g_parameter=g_parameter*3.1544;
    break;
    case 10:
           //if HISTOLOGY is PAGETS then g-value is 1.42765
           g_parameter=g_parameter*1.42765;
    break;
    case 11:
           //if HISTOLOGY is PAPILLARY then g-value is 0.49
           g_parameter=g_parameter*0.49;
    break;
    case 12:
           //if HISTOLOGY is CRIBIFORM then g-value is 0.70395
           g_parameter=g_parameter*0.70395;
    break;
    case 13:
           //if HISTOLOGY is APOCRINE then g-value is 0.8505
           g_parameter=g_parameter*0.8505;
    break;
    case 14:
           //if HISTOLOGY is PHYLLODES then g-value is 0.14972
           g_parameter=g_parameter*0.14972;
    break;
}  //end Switch for histology

/*****************************************************************
* Multiply g-parameter with g-value based on grade
*************************************************************************/
switch (grade) {
    case 0:
           //if GRADE is UNKNOWN then g-value is 1
```

```
                g_parameter=g_parameter*1;
        break;
        case 1:
                //if GRADE is 1 then g-value is 0.41345
                g_parameter=g_parameter*0.41345;
        break;
        case 2:
                //if GRADE is 2 then g-value is 0.8267
                g_parameter=g_parameter*0.8267;
        break;
        case 3:
                //if GRADE is 3 then g-value is 1.11584
                g_parameter=g_parameter*1.11584;
        break;
        case 4:
                //if GRADE is UNDIFFERENTIATED then g-value is 1.23275
                g_parameter=g_parameter*1.23275;
        break;
    } //end Switch for grade


/*********************************************************************
*  STEP 2.b    The program calculates the 15-year Kaplan-Meier cancer death rate, L, using the
SNAP method and the product of the g parameters
* Calculates lethality of primary breast cancer tumor (L_primary), lethality of nodes (L_nodes),
and 15-year Kaplan Meier cancer * specific death rate (L_breastcancer_KM)
* Second L_breastcancer_KM function adjusts for threatment effects
*********************************************************************/
    L_primary = 1 - Math.exp(-Qs*g_parameter*j_primary*Math.pow(dia*10,Z));
    L_nodes = 1 - Math.exp(-nnum*R);
    L_breastcancer_KM = L_primary + L_nodes - (L_primary*L_nodes);


/*********************************************************************
* STEPs 2.c, 2.d, & 2.e calculate cancer death rate in each of the 15 years following diagnosis
* Calculates yearly lethalities due to breast cancer and other causes
*********************************************************************/
       for (i=1; i<=15; i++) {
                //STEP 2.c calculates cancer death distribution by multiplying 15yr KM cancer
death rate by expected BRCA yearly lethality
                //percentage of overall cancer deaths occuring in the given year is computed, and
cumulatively summed
            cancer_death_dist_cumm[i] = cancer_death_dist_cumm[i-1] +
L_breastcancer_distribution[i-1]*L_breastcancer_KM;
            //cancer-specific hazard is computed as the chance of cancer death divided by cancer-
specific survival to that point
            cancer_death_hazard[i] = L_breastcancer_distribution[i-1]*L_breastcancer_KM / (1-
cancer_death_dist_cumm[i-1]);

            L_breastcancer_death_yearly[i]=remaining_percentage[i-1] * cancer_death_hazard[i];
            //STEP 2.d calculates non-BRCA death rate by multiplying the fraction of patients not
dying of cancer by the yearly risk of death due to non-cancer causes for the given age
            if (age==0){
                                L_nonbreastcancer_prob[i]=0;
                        } else {
                                L_nonbreastcancer_prob[i]=nvsr_death_prob_yearly[i+age];
                        }
            L_nonbreastcancer_death_yearly[i]=(remaining_percentage[i-1] -
L_breastcancer_death_yearly[i]) *L_nonbreastcancer_prob[i];
            //STEP 2.e calculates overall death rate by adding breast cancer deaths to non-breast
cancer deaths

L_overall_death_yearly[i]=L_breastcancer_death_yearly[i]+L_nonbreastcancer_death_yearly[i];
                remaining_percentage[i]=remaining_percentage[i-1]-L_overall_death_yearly[i];
       }

/*********************************************************************
* STEP 2.f  Calculate 15 values for cumulative breast cancer, non-breast cancer, and total death
rates by summing the respective yearly values computed in the steps above.
*********************************************************************/
```

```
    for(i=1;i<=15;i++) {
       L_cancer_death_cumm[i]=L_cancer_death_cumm[i-1]+L_breastcancer_death_yearly[i];
       L_noncancer_death_cumm[i]=L_noncancer_death_cumm[i-1]+L_nonbreastcancer_death_yearly[i];
       L_overall_death_cumm[i]=L_overall_death_cumm[i-1]+L_overall_death_yearly[i];
    }
/*****************************************************************
* STEP 3   Calculate the mean number of years of life left that can be expected for the cancer
patient
*****************************************************************************/
/*****************************************************************
* STEP 3.a   Calculate the life expectancy for the cancer patient by multiplying the chance of
dying in each of the years 1-15 by the number of years survived to that point.  Then add the NVSR
life expectancy for people 15 years older than the patient's current age, multiplied by the
patients chance of surviving 15 years.
*****************************************************************************/
    calc_life_expectation = 0;
    for (i=1; i<=15; i++){
        calc_life_expectation = calc_life_expectation + L_overall_death_yearly[i] * (i-0.5);
    }
    calc_life_expectation = calc_life_expectation + (1 - L_overall_death_cumm[15]) *
(nvsr_life_expect[age + 15] +15)

/*****************************************************************
* STEP 3.b   The program calculates the expected years of life lost due to cancer, by subtracting
the calculated life expectancy (step 3.a) from the NVSR-given life expectancy for the specified
age.
*****************************************************************************/

    expect_years_life_lost = nvsr_life_expect[age] - calc_life_expectation;

/*****************************************************************
* Determine whether projections exceed 100 years of age, and remove such projections-- data is
not projected to ages above 100
*****************************************************************/
    age_difference = 100-age;

    if (age_difference<15){
       for (i=age_difference; i<=15; i++) {
              L_cancer_death_cumm[i]=L_cancer_death_cumm[age_difference];
              L_noncancer_death_cumm[i]=L_noncancer_death_cumm[age_difference];
              L_overall_death_cumm[i]=L_overall_death_cumm[age_difference];
       }
    }

/*****************************************************************
* Determine whether cumulative death rate exceeded 1, and terminate projections
*****************************************************************/
    j=0;
    for (i=0; i<15; i++) {
       if(L_overall_death_cumm[i]<1){
              j=j+1;
       } else{
       }
    }

    if(L_overall_death_cumm[j]>=1) {
       for (k=j;k<=15;k++) {
              L_overall_death_cumm[k]=1;
              L_cancer_death_cumm[k]=L_cancer_death_cumm[j];
              L_noncancer_death_cumm[k]=L_noncancer_death_cumm[j];
       }
    }


/*************************************************
STEP 4  Calculate death rates with a specific therapy type
*************************************************/
/*************************************************
STEP 4.a  Calculate the "risk-reduction" value based on the combination of therapies entered by
the user and the information collected in steps 1.c and 1.d, consistent with the assumptions of
Adjuvant!Online
```

```
**************************************************/
    /*********************************************************************
    * The following code gives the effect of endocrine (hormonal) therapy.
    **********************************************************************/
    if (endo==0) {
        endotherapyEffect=0;
    } else {
    switch(er) {
        case 0: //If ER unknown
                if(age<50){
                        endotherapyEffect=0.20;
                } else if(age >=50 && age < 60) {
                        endotherapyEffect=0.21;
                } else {
                        endotherapyEffect=0.23;
                }
        break;
        case 1: //If ER+
                if(age<50){
                        endotherapyEffect=0.32;
                } else if(age >=50 && age < 60) {
                        endotherapyEffect=0.32;
                } else {
                        endotherapyEffect=0.32;
                }
        break;
        case 2: //If ER-
                if(age<50){
                        endotherapyEffect=0;
                } else if(age >=50 && age < 60) {
                        endotherapyEffect=0;
                } else {
                        endotherapyEffect=0;
                }
        break;
    } //end switch
    } //end if for endo

    /*********************************************************************
    * The following code gives the effect of chemotherapy.
    **********************************************************************/

    switch (chemo) {
        case 0: //no chemo
                chemotherapyEffect=0;
        break;
        case 1: //CMF
                switch(er) {
                        case 0: //If ER unknown
                                if(age<50){
                                        chemotherapyEffect=0.30;
                                } else if(age >=50 && age < 60) {
                                        chemotherapyEffect=0.18;
                                } else {
                                        chemotherapyEffect=0.10;
                                }
                        break;
                        case 1: //If ER+
                                if(age<50){
                                        chemotherapyEffect=0.30;
                                } else if(age >=50 && age < 60) {
                                        chemotherapyEffect=0.16;
                                } else {
                                        chemotherapyEffect=0.08;
                                }
                        break;
                        case 2: //If ER-
                                if(age<50){
                                        chemotherapyEffect=0.30;
                                } else if(age >=50 && age < 60) {
                                        chemotherapyEffect=0.22;
```

```
                        } else {
                                chemotherapyEffect=0.15;
                        }
                break;
        } //end switch er
break;

case 2: //Anthra
        switch(er) {
                case 0: //If ER unknown
                        if(age<50){
                                chemotherapyEffect=0.41;
                        } else if(age >=50 && age < 60) {
                                chemotherapyEffect=0.31;
                        } else {
                                chemotherapyEffect=0.24;
                        }
                break;
                case 1: //If ER+
                        if(age<50){
                                chemotherapyEffect=0.41;
                        } else if(age >=50 && age < 60) {
                                chemotherapyEffect=0.29;
                        } else {
                                chemotherapyEffect=0.23;
                        }
                break;
                case 2: //If ER-
                        if(age<50){
                                chemotherapyEffect=0.41;
                        } else if(age >=50 && age < 60) {
                                chemotherapyEffect=0.34;
                        } else {
                                chemotherapyEffect=0.29;
                        }
                break;
        } //end switch er
break;

case 3: //1st gen
        switch(er) {
                case 0: //If ER unknown
                        if(age<50){
                                chemotherapyEffect=0.30;
                        } else if(age >=50 && age < 60) {
                                chemotherapyEffect=0.18;
                        } else {
                                chemotherapyEffect=0.10;
                        }
                break;
                case 1: //If ER+
                        if(age<50){
                                chemotherapyEffect=0.30;
                        } else if(age >=50 && age < 60) {
                                chemotherapyEffect=0.16;
                        } else {
                                chemotherapyEffect=0.08;
                        }
                break;
                case 2: //If ER-
                        if(age<50){
                                chemotherapyEffect=0.30;
                        } else if(age >=50 && age < 60) {
                                chemotherapyEffect=0.22;
                        } else {
                                chemotherapyEffect=0.15;
                        }
                break;
        } //end switch er
break;
```

```
        case 4: //2nd gen
                switch(er) {
                        case 0: //If ER unknown
                                if(age<50){
                                        chemotherapyEffect=0.44;
                                } else if(age >=50 && age < 60) {
                                        chemotherapyEffect=0.34;
                                } else {
                                        chemotherapyEffect=0.28;
                                }
                        break;
                        case 1: //If ER+
                                if(age<50){
                                        chemotherapyEffect=0.44;
                                } else if(age >=50 && age < 60) {
                                        chemotherapyEffect=0.33;
                                } else {
                                        chemotherapyEffect=0.26;
                                }
                        break;
                        case 2: //If ER-
                                if(age<50){
                                        chemotherapyEffect=0.44;
                                } else if(age >=50 && age < 60) {
                                        chemotherapyEffect=0.38;
                                } else {
                                        chemotherapyEffect=0.32;
                                }
                        break;

        } //end switch er
        break;

        case 5: //3rd gen
                switch(er) {
                        case 0: //If ER unknown
                                if(age<50){
                                        chemotherapyEffect=0.55;
                                } else if(age >=50 && age < 60) {
                                        chemotherapyEffect=0.47;
                                } else {
                                        chemotherapyEffect=0.42;
                                }
                        break;
                        case 1: //If ER+
                                if(age<50){
                                        chemotherapyEffect=0.55;
                                } else if(age >=50 && age < 60) {
                                        chemotherapyEffect=0.45;
                                } else {
                                        chemotherapyEffect=0.40;
                                }
                        break;
                        case 2: //If ER-
                                if(age<50){
                                        chemotherapyEffect=0.55;
                                } else if(age >=50 && age < 60) {
                                        chemotherapyEffect=0.49;
                                } else {
                                        chemotherapyEffect=0.45;
                                }
                        break;
                } //end switch er
        break;
    } //end switch for chemo

    /*****************************************************
    * Combine effect of endocrine therapy and chemotherapy
    *****************************************************/
    totaltherapyEffect = endotherapyEffect + chemotherapyEffect - (endotherapyEffect *
chemotherapyEffect);
```

```
    L_breastcancer_KM_therapy=L_breastcancer_KM*(1-totaltherapyEffect);

/*********************************************************************
* STEP 4.b calculates 15 values for the breast cancer death rate with therapy in each of the 15
years after diagnosis by multiplying the 15-year Kaplan-Meier cancer death rate, L, (calculated
in step 1) by the "risk-reduction" value computed above, and by the fraction of the total
lethality which can be expected in each year(the 15-part step function described in step 2.a that
captures the breast carcinoma hazard function).
*********************************************************************/
       for (i=1; i<=15; i++) {
              //percentage of overall cancer deaths occuring in the given year is computed, and
cumulatively summed
        cancer_death_dist_cumm[i] = cancer_death_dist_cumm[i-1] + L_breastcancer_distribution[i-
1]*L_breastcancer_KM_therapy;
        //cancer-specific hazard is computed as the chance of cancer death divided by cancer-
specific survival to that point
        cancer_death_hazard[i] = L_breastcancer_distribution[i-1]*L_breastcancer_KM_therapy / (1-
cancer_death_dist_cumm[i-1]);

        L_cancer_death_yearly_therapy[i]=remaining_percentage_therapy[i-1] *
cancer_death_hazard[i];
              if (age==0){
                     L_noncancer_prob_therapy[i]=0;
              } else {
                     L_noncancer_prob_therapy[i]=nvsr_death_prob_yearly[i+age];
              }
              L_noncancer_death_yearly_therapy[i]=(remaining_percentage_therapy[i-1]-
L_cancer_death_yearly_therapy[i]) *L_noncancer_prob_therapy[i];

       L_overall_death_yearly_therapy[i]=L_cancer_death_yearly_therapy[i]+L_noncancer_death_year
ly_therapy[i];
              remaining_percentage_therapy[i]=remaining_percentage_therapy[i-1]-
L_overall_death_yearly_therapy[i];

       } //end of yearly lethality calculator

/*********************************************************************
* STEP 4.c & 4d   Calculate 15 values for the cumulative breast cancer death rate and cumulative
overall death rate in each of the 15 years after diagnosis by summing the respective yearly risks
of cancer death, with therapy, (step 2) from the time of diagnosis.
*********************************************************************************/
    for(i=1;i<=15;i++) {
       L_cancer_death_cumm_therapy[i]=L_cancer_death_cumm_therapy[i-
1]+L_cancer_death_yearly_therapy[i];
       L_noncancer_death_cumm_therapy[i]=L_noncancer_death_cumm_therapy[i-
1]+L_noncancer_death_yearly_therapy[i];
       L_overall_death_cumm_therapy[i]=L_overall_death_cumm_therapy[i-
1]+L_overall_death_yearly_therapy[i];
    }  //end of lethality summation

/*********************************************************************
* STEP 5 Calculates the life expectancy gained from therapy
*********************************************************************************/
/*********************************************************************
* STEP 5.a   Calculate the life expectancy for the cancer patient by multiplying the chance of
dying in each of the years 1-15 by the number of years survived to that point.  Then add the NVSR
life expectancy for people 15 years older than the patient's current age, multiplied by the
patients chance of surviving 15 years.
*********************************************************************************/
    calc_life_expectation_therapy = 0;
    for (i=1; i<=15; i++){
       calc_life_expectation_therapy = calc_life_expectation_therapy +
L_overall_death_yearly_therapy[i] * (i-0.5);
    }
    calc_life_expectation_therapy = calc_life_expectation_therapy + (1 -
L_overall_death_cumm_therapy[15]) * (nvsr_life_expect[age + 15] +15)

/*********************************************************************
* STEP 5.b   calculates the life expectancy gained from therapy by subtracting the mean life
expectancy with therapy (step 2.e) from the mean life expectancy for the cancer patient (step 3).
*********************************************************************************/
```

```
    expect_years_life_lost_therapy = nvsr_life_expect[age] - calc_life_expectation_therapy;

    expect_life_saved_years=expect_years_life_lost-expect_years_life_lost_therapy;
    expect_life_saved_days=expect_life_saved_years*365.25;

/*************************************************************
* Data is not projected to ages above 100
*************************************************************/
age_difference = 100-age;

if (age_difference<15){
        for (i=age_difference; i<=15; i++) {
                L_cancer_death_cumm_therapy[i]=L_cancer_death_cumm_therapy[age_difference];
                L_noncancer_death_cumm_therapy[i]=L_noncancer_death_cumm_therapy[age_difference];
                L_overall_death_cumm_therapy[i]=L_overall_death_cumm_therapy[age_difference];
        }
}

/*************************************************************
* Calculations ends if L_overall_death_cumm_therapy equals or becomes greater than 1
*************************************************************/
j=0;
for (i=0; i<15; i++) {
        if(L_overall_death_cumm_therapy[i]<1){
                j=j+1;
        } else{
        }
}

if(L_overall_death_cumm_therapy[j]>=1) {
        for (k=j;k<=15;k++) {
                L_overall_death_cumm_therapy[k]=1;
                L_cancer_death_cumm_therapy[k]=L_cancer_death_cumm_therapy[j];
                L_noncancer_death_cumm_therapy[k]=L_noncancer_death_cumm_therapy[j];
        }
}

/*****************************
* Throws alert messages if age is 0 or over 100
*****************************/
var alertOldText=document.getElementById("alertmess");

if (age==0){
alertOldText.firstChild.nodeValue="Calculations do not incorporate \n non breast cancer risk.";
} else if (age>85){
alertOldText.firstChild.nodeValue="Calculations end at age 100 due to the lack of viable data for
people over the age of 100."+j;
} else {
alertOldText.firstChild.nodeValue=" ";
}

if (document.form.distantMetastasis.value == '1') {
}
/*************************************************************************
* Displays results on the website
*************************************************************************/
document.getElementById("death_reduction").firstChild.nodeValue=
Math.round(totaltherapyEffect*1000) / 10;
document.getElementById("life_expect").firstChild.nodeValue=Math.round(nvsr_life_expect[age]*10)/
10;
document.getElementById("expect_life_lost").firstChild.nodeValue=Math.round(expect_years_life_los
t*10)/10;
document.getElementById("expect_saved_years").firstChild.nodeValue=Math.round(expect_life_saved_y
ears*10)/10;
document.getElementById("expect_saved_days").firstChild.nodeValue=Math.round(expect_life_saved_da
ys);
document.getElementById("l_km").firstChild.nodeValue=Math.round(L_breastcancer_KM_therapy*1000)/1
0;
document.getElementById("life_expect_with_cancer").firstChild.nodeValue=Math.round((nvsr_life_exp
ect[age] - expect_years_life_lost)*10)/10;
```

```
document.getElementById("l_expected").firstChild.nodeValue=Math.round(L_cancer_death_cumm_therapy
[15]*1000)/10;
document.getElementById("ageText").firstChild.nodeValue=age;


/*******************************
* STEP 6  graphs the risk curves for cancer (step 2.c), cancer with therapy (step 1), non-cancer
(step 2.d), overall (step 2.e), and overall with therapy (step 5.b) in the user-specified mode,
either as mortality curves, survival curves, a bar graph, a pie chart, or a pictogram. For the
outcome calculator, the program displays the life expectancy (step 3.a), the life expectancy lost
to cancer (step 3.d), and the 15-year Kaplan-Meier cancer-specific death rate (step 1). For the
treatment calculator, the program displays these values, as well as the risk reduction value from
therapy (step 1) and the life expectancy gained from therapy (step 5.d).
*******************************/


plotGraph(new Array(L_cancer_death_cumm,L_noncancer_death_cumm ,L_overall_death_cumm ,
L_cancer_death_cumm_therapy, L_noncancer_death_cumm_therapy, L_overall_death_cumm_therapy),
document.getElementById('display_options').value, null, null);

/*********************************************************
* STEP 6b  Display staging information
**********************************************************************/
//updateStaging();
tnm_stage = getTNM_Stage();
document.getElementById("tnm").firstChild.nodeValue= tnm_stage[0];
document.getElementById("stageNum").firstChild.nodeValue = tnm_stage[1];
```

REFERENCES

[1] Naik AD, Petersen LA. The neglected purpose of comparative-effecttivenss research. N Engl J Med. 2009 May 7;360(19):1929-31.
[2] Avorn J. Debate about funding comparative-effecttivenss research. N Engl J Med. 2009 May 7;360(19):1927-9.
[3] Garber AM, Tunis SR. Does comparative-effecttivenss research threaten personalized medicine? N Engl J Med. 2009 May 7;360(19):1925-7
[4] Alexander GC, Stafford RS. Does comparative effectiveness have a comparative edge? JAMA. 2009 Jun 17;301(23):2488-90.
[5] Lundin J, Lundin M, Isola J, et al: A web-based system for individualised survival estimation in breast cancer. BMJ 326:29, 2003
[6] Ravdin PM, Siminoff LA, Davis GJ, et al: Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 19:980-991, 2001
[7] Ivo A. Olivotto, Chris D. Bajdik, Peter M. Ravdin, Caroline H. Speers, Andrew J. Coldman, Brian D. Norris, Greg J. Davis, Stephen K. Chia, Karen A. Gelmon Population-Based Validation of the Prognostic Model ADJUVANT! for Early Breast Cancer *J Clin Oncol* 23(12):2716-25 2005.
[8] Stella Mook, Marjanka K Schmidt, Emiel J Rutgers, Anthonie O van de Velde, Otto Visser, Sterre M Rutgers, Nicola Armstrong, Laura J van't Veer, Peter M Ravdin Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: a hospital-based retrospective cohort study The Lancet Oncology, Volume 10, Issue 11, November 2009, Pages 1070-1076
[9] Huobera J and Thürlimanna B Adjuvant! When the new world meets the old world The Lancet Oncology, Volume 10, Issue 11, Pages 1028 - 1029, November 2009
[10] 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. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). *Lancet*. 2005 May 14-20;365(9472):1687-717. Links
[11] Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials *Lancet* 371: 29-40 2008
[12] Technical Report #1 - Mathematical Methods (March 9, 2009) at http://cancer.lifemath.net/about/techreports/index.php
[13] Michaelson JS, et al: How Cancer at the Primary Site and in the Nodes Contributes to Lethality. *CANCER* in press 2009
[14] Michaelson JS, et al: Why Cancer at the Primary Site and in the Nodes Contributes to Lethality. *CANCER* in press 2009
[15] Michaelson JS, et al: The Impact of Primary Tumor Size, Nodal Status, and Other Prognostic Factors on the Risk of Cancer Death. *CANCER* in press 2009
[16] Rusby JE, Brachtel EF, Othus M, Michaelson JS, Koerner FC and Smith BL, Development and validation of a model predictive of occult nipple involvement in women undergoing mastectomy *British Journal of Surgery* 95: 1356–1361 2008
[17] CancerMath.net Technical Report #1, Mathematical Methods. http://cancer.lifemath.net/about/techreports/technical_report_1.pdf
[18] Mariotto A, Feuer EJ, Harlan LC, Wun LM, Johnson KA, Abrams J. Trends in use of adjuvant multi-agent chemotherapy and tamoxifen for breast cancer in the United States: 1975-1999. *J Natl Cancer Inst.* Nov 6;94(21):1626-34. 2002
[19] Chen and Michaelson, Technical Report #9 - Adjuvant Multi-agent Chemotherapy and Tamoxifen Usage Trends for Breast Cancer in the United States (March 27, 2009), http://cancer.lifemath.net/about/techreports/technical_report_9.pdf
[20] Karrison TG Ferguson DJ Meier P Dormancy of mammary carcinoma after mastectomy. *J Natl Cancer Inst* 91:80-5 1991

[21] Berry DA, Cirrincione C, Henderson IC, Citron ML, Budman DR, Goldstein LJ, Martino S, Perez EA, Muss HB, Norton L, Hudis C, Winer EP. Estrogen-receptor status and outcomes of modern chemotherapy for patients with node-positive breast cancer. *JAMA.* 295(14):1658-67. 2006

[22] CancerMath.net Technical Report #7b, SEER Breast Cancer Database. http://cancer.lifemath.net/about/techreports/technical_report_7b.pdf

[23] National Vital Statistics Reports Vol. 54 No. 14, April 19, 2006, United States Life Tables 2003

[24] Singletary SE, Connolly JL Breast Cancer Staging: Working With the Sixth Edition of the AJCC Cancer Staging Manual CA Cancer J Clin, Jan 2006; 56: 37 - 47.

[25] Ozanne EM, Braithwaite D, Sepucha K, Moore D, Esserman L, Belkora J. Sensitivity to input variability of the Adjuvant! Online breast cancer prognostic model. *J Clin Oncol*. Jan 10;27(2):214-9. 2009

[26] Bush DM, Chen LL, Michaelson JS, Laboratory for Quantitative Medicine Technical Report #12 Validation: Breast Cancer Web Calculator August 27, 2009. http://www.lifemath.net/devcancer/about/techreports/technical_report_12a_draft.pdf

[27] Michaelson J. The role of molecular discreteness in normal and cancerous growth. *Anticancer Res.* 1999;19:4853-4867.

[28] Hughes KS, Schnaper LA, Berry D, Cirrincione C, McCormick B, Shank B, Wheeler J, Champion LA, Smith TJ, Smith BL, Shapiro C, Muss HB, Winer E, Hudis C, Wood W, Sugarbaker D, Henderson IC, Norton L; Cancer and Leukemia Group B; Radiation Therapy Oncology Group; Eastern Cooperative Oncology Group Lumpectomy plus tamoxifen with or without irradiation in women 70 years of age or older with early breast cancer. *N Engl J Med*. Sep 2;351(10):971-7. 2004

[29] Fyles AW, McCready DR, Manchul LA, Trudeau ME, Merante P, Pintilie M, Weir LM, Olivotto IA. Tamoxifen with or without breast irradiation in women 50 years of age or older with early breast cancer. Engl J Med. Sep 2;351(10):963-70. 2004

[30] 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

[31] Michaelson, J, Halpern, E, Kopans, D. A Computer Simulation Method For Estimating The Optimal Intervals For Breast Cancer Screening. *Radiology*. 212:551-560 1999

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

[33] 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

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