



Original article

Conditional survival in gastric cancer: a SEER database analysis

SAMUEL J. WANG^{1,2}, RACHEL EMERY³, CLIFTON D. FULLER^{4,5}, JONG-SUNG KIM⁶, DEAN F. SITTING^{2,7},
and CHARLES R. THOMAS JR.¹

¹Department of Radiation Medicine, MC KPV4, Oregon Health and Science University, 3181 SW Sam Jackson Park Road, Portland, OR 97239-3098, USA

²Department of Medical Informatics and Clinical Epidemiology, Oregon Health and Science University, Portland, OR, USA

³School of Medicine, Oregon Health and Science University, Portland, OR, USA

⁴Department of Radiation Oncology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

⁵Graduate Division of Radiological Sciences, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

⁶Department of Mathematics and Statistics, Portland State University, Portland, OR, USA

⁷Applied Research in Medical Informatics, Northwest Permanente, PC, Portland, OR, USA

Abstract

Background. Gastric cancer survival is typically reported in terms of survival from the time of diagnosis. Conditional survival is a more relevant measure of prognosis for patients who have already survived 1 or more years since diagnosis.

Methods. Using the Surveillance, Epidemiology, and End Results (SEER 17) database from the National Cancer Institute, we analyzed data from 20018 gastric cancer patients diagnosed between 1988 and 1998. Using the life-table method, we computed 5-year relative conditional survival, grouped by summary stage, age, sex, and ethnicity, for patients who had already survived up to 5 years from diagnosis.

Results. Relative conditional survival improves over time for all groups of gastric cancer patients who survive a period of time after diagnosis. The largest gains in conditional survival were seen in patients with advanced stage disease. In general, females showed better survival than males. When grouped by ethnicity, Asians continued to have improved survival compared to other ethnic categories, even at 5 years out from diagnosis.

Conclusion. For gastric cancer patients who survive a period of time after diagnosis, the largest increases in conditional survival are seen for patients with advanced stage disease and for those less than 65 years old. Conditional survival can provide more relevant prognostic information than survival from the time of diagnosis for gastric cancer patients who survive a period of time after diagnosis.

Key words Survival analysis · Epidemiological methods · Stomach neoplasms

Introduction

In 2006, the American Cancer Society estimated that there were 22280 new cases of gastric cancer diagnosed

in the United States, and 11430 gastric cancer deaths [1]. The incidence of gastric cancer increases with age, with a median age at diagnosis of 70 years for males and 74 for females. Gastric cancer is almost twice as common in men than in women, and incidence rates are 2.2 times higher in blacks compared to whites [2]. While the overall incidence of gastric cancer has been decreasing over time since the 1930s, the incidence rates for black males has not changed significantly, and rates actually increased for white females between 1974 and 1994 [2].

Gastric cancer is an aggressive malignancy that is difficult to detect at an early stage. Survival estimates for gastric cancer patients are usually only reported in the literature as survival from the time of diagnosis, and 5-year overall survival rates from diagnosis are reported as about 22% [2,3]. Survival probability changes, however, for gastric cancer patients who survive for a period of time after diagnosis, and their prognosis is more accurately described using conditional survival (CS) [4]. CS is taken from the concept of conditional probability, and is a more accurate estimate of survival probability for patients who have survived for 1 or more years after diagnosis. CS accounts for the fact that hazard rates can change over time. Similar to many other cancer types, hazard rates for gastric cancer are relatively higher in the first few years after diagnosis, signifying greater risk, but then the rates decrease markedly after the first few years. For those patients who are fortunate to survive beyond this initial period of time, prognosis can be substantially improved, and the initial estimates of survival made at the time of diagnosis no longer apply. CS is a more relevant measure of survival probability for these cancer survivors.

The concept of CS has important practical clinical value for patients, providers, and researchers [4,5]. Cancer patients who are seen in follow-up clinics several years after their diagnosis may wish to know how their

Offprint requests to: S. J. Wang

Received: December 19, 2006 / Accepted: May 2, 2007

prognosis is changing over time. If a cancer patient's CS has risen high enough to essentially match the expected survival of the general population, it allows a more objective basis to deem a patient "cured" of their disease. Providers can also make use of CS information to more objectively determine an appropriate frequency of follow-up visits and aggressiveness of surveillance testing, based on the patient's current risk profile. When designing clinical trials, clinical researchers may also find CS useful in helping to determine sufficient follow-up times for trial endpoints.

Several authors have previously published studies on CS for various disease sites, including breast [6], colon [7], central nervous system (CNS) [8,9], lung [10,11], and other advanced carcinomas [12]. We have also previously presented our work on CS for various disease sites [13–15]. However, to our knowledge, no formal literature has explored the CS patterns of gastric carcinomas. The specific aim of this study was to investigate the prognostic factors that affect the CS of gastric cancer patients, using the Surveillance, Epidemiology, and End Results (SEER) database [16].

Methods

Definition of conditional survival

Conditional survival (CS) is derived from the concept of conditional probability in biostatistics. CS can be calculated from traditional Kaplan-Meier or actuarial life-table survival data. The mathematical definition of CS can be expressed as follows: CS, $CS(y|x)$, is the probability of surviving an additional y years, given that the person has already survived x years. Let $S(t)$ be the traditional actuarial life-table survival at time t . CS can be expressed as:

$$CS(y|x) = \frac{S(x+y)}{S(x)}$$

For example, to compute the 5-year CS for a patient who has already survived 2 years, the survival at $5 + 2$ years, $S(7)$, is divided by the survival at 2 years, $S(2)$. When a survival curve has a changing hazard rate over time, this will be reflected as a change in CS as more time elapses from the time of diagnosis.

Data selection criteria

The SEER Program [16] from the National Cancer Institute is a population-based cancer registry covering approximately 26% of the United States population across several disparate geographic regions, the largest publicly available domestic cancer dataset. SEER program registries collect data on patient demograph-

ics, cancer type and site, stage, and first course of treatment, and they follow up vital status.

Using the November 2005 release of the SEER 17 database [16] with the SEER*Stat 6.2.4 software [17] we analyzed survival data from all patients diagnosed with gastric cancer between 1988 and 1998, with follow-up to December 2003, to ensure a minimum of 5 years of follow-up data. Patients were selected who had a "site recode" data field of "Stomach" with ICD-O-3 histology codes for adenocarcinoma (or unspecified carcinoma) (8000-8001, 8010, 8020-8021, 8050, 8140-8221, 8255-8560, 8570-8576).

Case selection options in SEER*Stat were used to restrict cases to "actively followed", "malignant behavior", "exclude all death certificate only and autopsy only", and "exclude second and later primaries".

Because American Joint Committee on Cancer (AJCC) staging information was not available in SEER for gastric cancer, we used the SEER summary stage (1977) [18]. Using a combination of both the clinical and the pathological documentation of the extent of disease, summary staging categorizes how far a cancer has spread from its point of origin. The stages included: "localized only", "regional by direct extension only", "regional by lymph node involvement only", "regional by both direct extension and lymph node(s) involvement", and "distant site/node involvement".

Data analysis

Conditional survival probabilities were calculated using SEER*Stat 6.2.4 [17], using the actuarial life-table method. All CS rates are reported as relative survival, defined as the observed CS rate divided by the expected CS rate for similar individuals matched for age, sex, ethnicity, and date at which the age was coded. This is a method for accounting for competing causes of death when exact cause-of-death information is incomplete or unknown. We computed 5-year CS by ethnicity and stage, and further grouped stage results by age (>65 vs <65 years), and sex.

Results

After eliminating 3195 cases that did not meet the case selection criteria, a total of 20018 gastric cancer patients were included in this analysis. When grouped by summary stage, 21% had localized disease, 10% had regional disease by direct extension only, 10% had regional disease by lymph node involvement only, 18% had regional disease by both direct extension and lymph node involvement, and 41% had distant site/nodal disease. Patients were further subsetted by sex and age, as shown in Table 1, and the number of patients in each ethnic group is shown in Table 2.

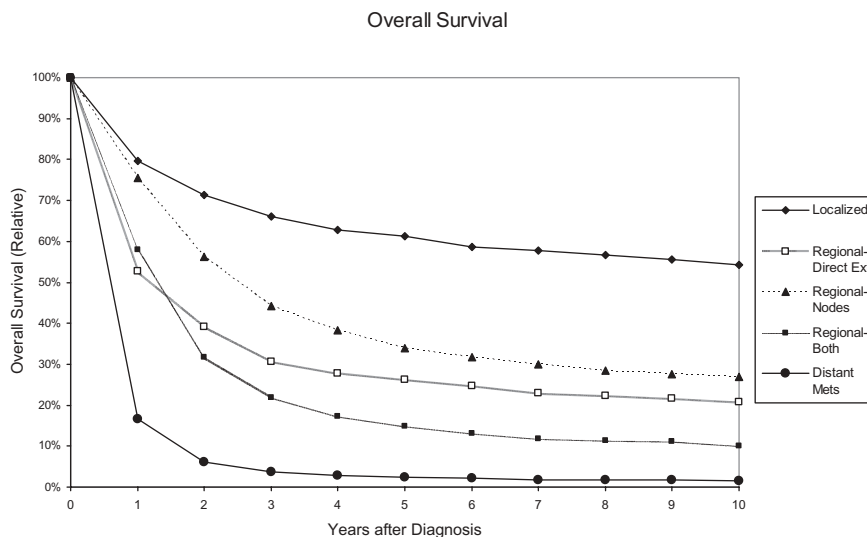


Fig. 1. Ten-year (relative) overall survival curves by Surveillance, Epidemiology, and End Results (SEER) summary stage. These data were used to calculate the 5-year relative conditional survival probabilities. *Ext*, extension; *mets*, metastases

Table 1. Number (%) of patients, grouped by age, sex, and summary stage ($n = 20018$)

	Localized	Regional-direct extension	Regional-lymph node(s)	Regional-both	Distant mets
Age < 65	1126 (27)	568 (29)	807 (38)	1408 (38)	3277 (40)
Age > 65	3013 (73)	1386 (71)	1304 (62)	2270 (62)	4859 (60)
Men	2527 (61)	1187 (61)	1357 (64)	2410 (66)	5247 (64)
Women	1612 (39)	767 (39)	754 (36)	1268 (34)	2889 (36)
Totals	4139 (100)	1954 (100)	2111 (100)	3678 (100)	8136 (100)

Table 2. Number (%) of patients in each ethnic group

White	13950 (70)
Black	2368 (12)
American Indian/Alaskan Native	196 (1)
Asian/Pacific Islander	3476 (17)
Other/Unknown	28 (<1)
Total	20018 (100)

The 10-year relative overall survival data (Fig. 1) were used to calculate 5-year relative CS, shown in Figs. 2 and 3. Error bars depict the 95% confidence intervals for each CS statistic. Overall, while higher stage (distant disease) patients had lower CS at diagnosis compared to other stages, these patients also saw the greatest increases in CS as more time elapsed from diagnosis (Fig. 2). When grouped by ethnicity (Fig. 3), Asians had the quickest sustained increase in CS, but the largest differences in CS from year 0 to year 5 were in American Indians/Alaskan Natives. When grouped by age and stage, patients over 65 years of age with localized and distant disease did worse than those under 65 for all times following diagnosis. For patients over age 65 with distant disease, this discrepancy in CS appeared to widen over time (at 5 years, CS was 77% for age <65,

and 46% for age >65). When examined by sex and stage, CS for females was generally higher than for males, with the largest gains over time seen in patients with distant disease — for females, CS rose from 3% at diagnosis to 67% after 5 years, and for males, CS rose from 2% at diagnosis to 61% after 5 years.

Discussion

In this study, we found that various prognostic factors affect the conditional survival (CS) of patients with gastric cancer. From other studies of CS, it is known that patients with poorer initial prognoses often see the largest increases in CS over time [6,8,10,11]. Our study confirmed this finding, and also found that CS appeared to increase the most for American Indians/Alaskans, women, and those less than 65 years of age with advanced stage disease. Our data appear to correspond with previously published data regarding the influence of ethnicity, sex, and age on gastric cancer survival at diagnosis. Asian ethnicity is associated with distinct characteristics at presentation and has previously been shown to exhibit more favorable survival as compared with that in non-Asians, even when adjusted for age,

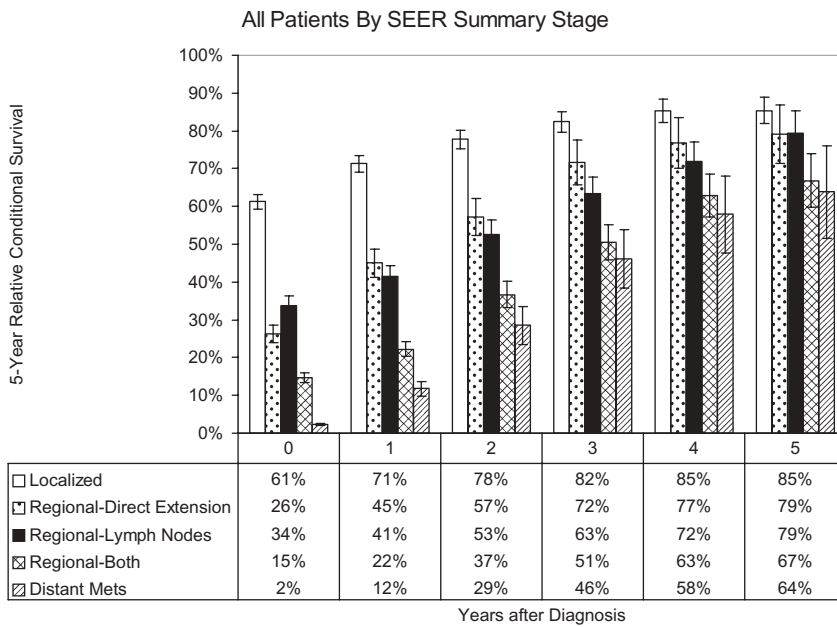


Fig. 2. Five-year relative conditional survival by SEER summary stage, as a function of elapsed time since diagnosis. Each bar represents the probability of surviving an additional 5 years, after having already survived for 0 to 5 years since diagnosis

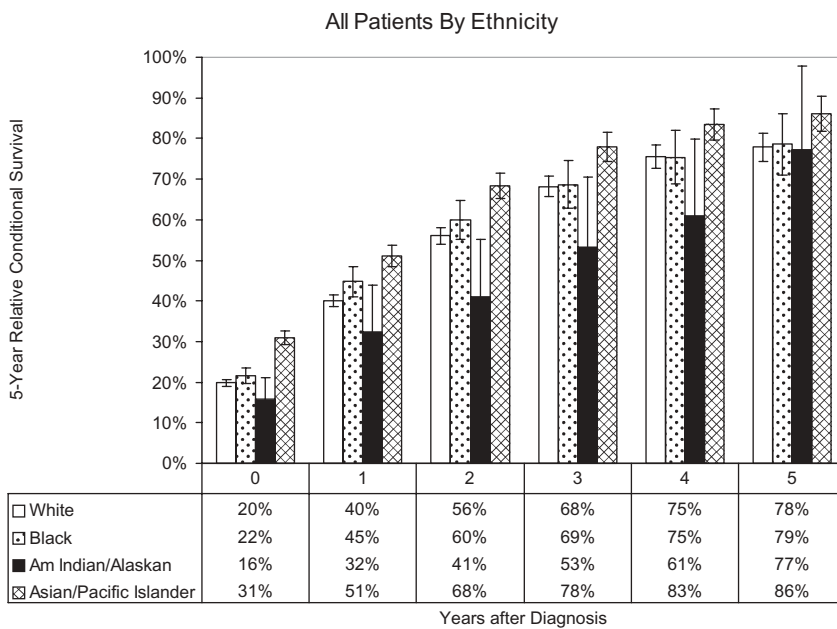


Fig. 3. Five-year relative conditional survival by ethnicity, as a function of elapsed time since diagnosis

and location [19,20]. Our findings show that Asians continue to have more favorable CS compared to non-Asians, even out to 5 years from the time of diagnosis.

As time progresses from diagnosis, CS provides more relevant prognostic information compared to more commonly reported static survival statistics, such as 5-year overall survival. CS information is potentially of great interest to patients, their clinicians, and researchers. When patients who are seen in follow-up inquire about their current prognosis, they should be given an accurate risk assessment that accounts for time already survived since diagnosis. Psychologically, the ability of

patients to more accurately quantify their improvement in prognosis over time may be of great benefit. Accordingly, every effort should be made to communicate this changing risk profile in terms that are accessible to the lay patient. Five-year CS probability is an easily understandable measure that can be used to accurately portray to a patient their current risk profile.

Clinicians can also make use of CS data to implement more evidence-driven approaches to plan an appropriate post-therapy surveillance schedule based on a patient's changing risk. Many physicians arbitrarily taper follow-up visit frequency after 2–3 years, but often

without evidence to justify whether this is an appropriate timeframe. The determination of optimum follow-up testing frequency and duration should ideally be based on the patient's disease risk, rather than simply on custom or tradition. For instance, the CS data in the present study show that, for patients with regional lymph node involvement who survive for more than 3 years from diagnosis, subsequent mortality outcomes are comparable to those in patients with localized disease at diagnosis. Consequently, if one uses a particular follow-up surveillance schedule for patients with localized disease for the first year after diagnosis and treatment, it stands to reason that patients with regional lymph node involvement should be similarly followed at 3 years out from diagnosis.

When researchers design clinical trials, the study duration is predicted based on the estimated length of follow-up needed to see a significant result in a given endpoint. These study duration estimates have important implications in terms of the economic costs of a trial and the timeliness in which trial results can be reported. Gill and Sargent [21] have proposed that alternative surrogate endpoints, such as disease-free survival, may be appropriate in certain situations, as a means to accelerate the completion of adjuvant clinical trials. Examination of how CS data change over time may be another method to ascertain an appropriate trial follow-up time, because these data effectively quantify the remaining risk to a patient after a given survival time.

It is imperative for physicians to more specifically characterize, utilize, and communicate patient risk profiles as accurately as possible. Models such as Adjuvant! Online [22] or the Memorial Sloan-Kettering Cancer Center (MSKCC) risk nomograms [23–27] are especially inviting as a mechanism for initial risk stratification. We are currently investigating the incorporation of a CS component into a similar predictive model that could be used to allow patients and physicians to re-evaluate a patient's prognosis as it changes over time. Such tools could potentially afford more appropriate management of disease as a part of the physician's surveillance algorithm.

Utilization of the SEER dataset represents an effort to determine CS parameters based on large population cohorts by using the single largest domestic cancer case data registry. The SEER dataset, while geographically limited, represents the best large-scale pool of patient data, and allows us to make reasonable estimates of CS that are generally applicable for the United States population. Nonetheless, several limitations must be noted. Increased cohort size comes at the cost of treatment homogeneity. Treatment data reported in SEER are limited (i.e., no chemotherapy regimen or radiotherapy dose is recorded), so we did not attempt a subanalysis

by treatment modality. Also, the sample size for some subgroups is small, and is reflected in the larger confidence interval bars seen in some groups, such as for patients with distant disease and those in the American Indian/Alaskan subgroup. This precludes our ability to make more definite generalizations regarding the significance of differences between these groups. Also, the small number of patients in the minority ethnic categories precluded our ability to subanalyze ethnicity by stage. Because SEER does not include information on disease recurrence, and because cause-of-death information is not always reliable, we are unable to analyze other endpoints, such as time to recurrence, or disease-free survival. Finally, historical survival data collected over an extended period of time may not reflect current practices in oncology. CS should be considered within the individual context of patient care, where specific risk factors must be carefully considered.

Throughout this analysis, relative survival rates are reported, defined as observed survival divided by the expected survival for that patient. Because cause-of-death information is often unreliable in the SEER database [28], relative survival is a useful alternative measure that assesses the proportion of excess deaths that occur in gastric cancer patients compared to the general population. It serves to adjust for differences in observed survival that may be due to competing causes of death [5]. Expected survival rates were obtained from the US SEER 1970, 1980, and 1990 expected rate tables by matching patients for age, sex, ethnicity, and the date at which age was coded.

In summary, our data show that CS for gastric cancer patients changes over time, and CS can be used as an important adjunct to traditional survival statistics. We have found that CS in gastric cancer shows the largest increases over time in patients under age 65, and in those with advanced stage disease. We hope that CS calculations will become more common in cancer reporting, and that extrapolation from other epidemiological and clinical trial datasets may allow future optimization of risk stratification for gastric cancer patients. Conditional survival data provide patients, clinicians, and researchers with more accurate prognostic information about how risk changes over time for cancer survivors.

References

1. American Cancer Society. Cancer facts and figures 2006 [online] 2006 [cited 12-12-2006]. <http://www.cancer.org>.
2. Blanke CD, Coia LR, Schwarz RE, Bonin SR. Gastric cancer. In: Pazdur R, Coia LR, Hoskins WJ, Wagman LD, editors. Cancer management: a multidisciplinary approach. 9th ed. Lawrence, KS: CMP Media, LLC; 2005. p. 279–92.
3. Pisters PWT, Kelsen DP, Powell SM, Tepper JE. Cancer of the stomach. In: Devita VT, Hellman S, Rosenberg SA, editors.

- Cancer. 7th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2004.
4. Henson DE, Ries LA. On the estimation of survival. *Semin Surg Oncol* 1994;10:2–6.
 5. Gloeckler Ries LA, Reichman ME, Lewis DR, Hankey BF, Edwards BK. Cancer survival and incidence from the Surveillance, Epidemiology, and End Results (SEER) program. *Oncologist* 2003;8:541–52.
 6. Henson DE, Ries LA, Carriaga MT. Conditional survival of 56268 patients with breast cancer. *Cancer* 1995;76:237–42.
 7. Merrill RM, Henson DE, Ries LA. Conditional survival estimates in 34963 patients with invasive carcinoma of the colon. *Dis Colon Rectum* 1998;41:1097–106.
 8. Davis FG, McCarthy BJ, Freels S, Kupelian V, Bondy ML. The conditional probability of survival of patients with primary malignant brain tumors: surveillance, epidemiology, and end results (SEER) data. *Cancer* 1999;85:485–91.
 9. Hwang SL, Yang YH, Lieu AS, Chuang MC, Chang SJ, Chang YY, et al. The conditional survival statistics for survivors with primary supratentorial astrocytic tumors. *J Neurooncol* 2000;50:257–64.
 10. Merrill RM, Henson DE, Barnes M. Conditional survival among patients with carcinoma of the lung. *Chest* 1999;116:697–703.
 11. Skuladottir H, Olsen JH. Conditional survival of patients with the four major histologic subgroups of lung cancer in Denmark. *J Clin Oncol* 2003;21:3035–40.
 12. Kato I, Severson RK, Schwartz AG. Conditional median survival of patients with advanced carcinoma: surveillance, epidemiology, and end results data. *Cancer* 2001;92:2211–9.
 13. Wang SJ, Fuller CD, Thomas CR. Ethnic disparities in conditional survival of patients with non-small cell lung cancer. *J Thorac Oncol* 2007;2:180–90.
 14. Fuller CD, Wang SJ, Thomas CR, Hoffman HT, Goldwein JW, Weber RS, Rosenthal DI. Conditional survival in head and neck squamous cell carcinoma: Results from the SEER Dataset 1973–1998. *Cancer* 2007;109:1331–43.
 15. Wang SJ, Fuller CD, Emery R, Thomas CR. Conditional survival in rectal cancer: a SEER database analysis. *Gastrointest Cancer Res* 2007;1.
 16. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence — SEER 17 Regs Public-Use, Nov 2005 Sub (1973–2003 varying), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2006, based on the November 2005 submission.
 17. Surveillance Research Program, National Cancer Institute SEER*Stat software (www.seer.cancer.gov/seerstat) version 6.2.4.
 18. Young JL Jr, Roffers SD, Ries LAG, Fritz AG, Hurlbut AA, editors. SEER summary staging manual — 2000: codes and coding instructions. National Cancer Institute, NIH Pub. No. 01-4969. Bethesda, MD: 2001. <http://seer.cancer.gov/tools/ssml>.
 19. Gill S, Shah A, Le N, Cook EF, Yoshida EM. Asian ethnicity-related differences in gastric cancer presentation and outcome among patients treated at a Canadian cancer center. *J Clin Oncol* 2003;21:2070–6.
 20. Theuer CP, Kurosaki T, Ziogas A, Butler J, Anton-Culver H. Asian patients with gastric carcinoma in the United States exhibit unique clinical features and superior overall and cancer specific survival rates. *Cancer* 2000;89:1883–92.
 21. Gill S, Sargent D. End points for adjuvant therapy trials: has the time come to accept disease-free survival as a surrogate end point for overall survival? *Oncologist* 2006;11:624–9.
 22. Ravdin P. Adjuvant! Online [online] 2006 [cited 6-20-2006]. <http://www.adjuvantonline.com/>.
 23. Memorial Sloan-Kettering Cancer Center. Prediction tools [online] 2006 [cited 6-20-2006]. Available from: URL: www.mskcc.org/nomograms.
 24. Kattan MW. Nomograms are superior to staging and risk grouping systems for identifying high-risk patients: preoperative application in prostate cancer. *Curr Opin Urol* 2003;13:111–6.
 25. Van Zee KJ, Manasseh DM, Bevilacqua JL, Boolbol SK, Fey JV, Tan LK, et al. A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. *Ann Surg Oncol* 2003;10:1140–51.
 26. Diblasio CJ, Kattan MW. Use of nomograms to predict the risk of disease recurrence after definitive local therapy for prostate cancer. *Urology* 2003;62(Suppl 1):9–18.
 27. Brennan MF, Kattan MW, Klimstra D, Conlon K. Prognostic nomogram for patients undergoing resection for adenocarcinoma of the pancreas. *Ann Surg* 2004;240:293–8.
 28. Hoel DG, Ron E, Carter R, Mabuchi K. Influence of death certificate errors on cancer mortality trends. *J Natl Cancer Inst* 1993;85:1063–8.