NOMOGRAM PREDICTIVE OF 10-YEAR CAUSE-SPECIFIC MORTALITY IN DIFFERENTIATED THYROID CANCER

By

Renee E. Park

A THESIS

Presented to the Department of Public Health and Preventive Medicine
and the Oregon Health & Science University
School of Medicine
in partial fulfillment of
the requirements for the degree of

Master of Public Health

May 2009

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the Master's thesis of

Renee E. Park

has been approved

Motomi Mori, Ph.D., Chair

Neil D. Gross, M.D., Mentor

Donald F. Austin, M.D., M.P.H., Member

Table of Contents

Acknowledgements

Thank you to

Dr. Neil Gross for his project, support, and mentorship, without which I would be starting residency in internal medicine instead of otolaryngology

Dr. Motomi Mori and Dr. Donald Austin for their counsel and support

Michael Lasarev, Dr. Samuel Wang, and Dr. Frank Harrell for their statistical prowess and generous giving of their time and wisdom

Dr. John Stull and the MD/MPH Class of 2009

The Tartar Trust Fellowship from the OHSU Foundation

My family and friends who have supported and encouraged me to pursue things worth pursuing

ABSTRACT

Background:

The application of appropriate treatment for differentiated thyroid cancer (DTC), including extent of surgery and adjuvant therapy, is predicated on accurate patient risk stratification. Although risk factors for mortality from DTC have been well-described on the population level, they have not been unified into a single algorithm to predict individual risk. This study aimed to develop a nomogram for estimating 10-year cause-specific mortality in well to poorly DTC.

Methods:

A historical cohort of 9,654 patients with DTC recorded in the SEER national cancer registry from 1985 to 1995 was used to identify and quantify all clinically relevant predictors of 10-year cancer-specific mortality. Multivariable Cox proportional hazards regression was used for model selection and nomogram development. The predictive accuracy of the nomogram was internally validated using bootstrapping methods and quantitated using the area under the receiver operating characteristic curve (AUC).

Results:

Ten-year cause-specific mortality was 3.3%. Significant predictors of mortality included age, gender, extracapsular extension, tumor size, nodal status, distant metastasis and histology. The nomogram successfully estimated an individualized risk of mortality from DTC by assigning relative weights to each of these risk factors. Model discrimination was excellent with an AUC of 0.93, with good calibration.

Discussion & Conclusions:

This nomogram is the first prognostic model developed to predict the likelihood of mortality for an individual patient with DTC. More accurate patient risk stratification using the nomogram has practical applications for clinical care and research.

CHAPTER 1: BACKGROUND

Part 1: Introduction

In 2008, it is estimated that 37,340 people were newly diagnosed with thyroid cancer in the United States.[1] Thyroid cancer accounts for 3.4% of cancers in the US.[2] The age-adjusted annual incidence of malignant thyroid cancer has increased in recent decades from 4.3 (1980) to 5.5 (1990) to 7.6 (2000) cases per 100,000 people.[3] Although incidence has been rising, possibly due to improved detection,[4] mortality has remained stable at approximately 4.6 deaths per million cases.[5] Survival is likewise encouraging, with 10-year relative survival from differentiated thyroid cancer estimated to be 96.5%.[1] Despite these population statistics, a diagnosis of thyroid cancer is not uniformly reassuring for individual patients.

Thyroid carcinomas show heterogeneous clinical behavior, ranging from indolence to rapid lethality. Several staging and scoring systems have been developed to prognosticate survival in thyroid cancer, such as AGES,[6] AMES,[7] MACIS,[8] among others. Tumor size/extension, lymph node involvement, metastasis, histology, age and gender have been used in these various systems, and their prediction of mortality validated by other groups. The American Joint Committee on Cancer TNM (tumor, lymph node, distant metastasis) staging system[9] is a widely accepted system in the

description of thyroid tumors. TNM stages range from I to IV, with worse outcomes found in higher stages. One study demonstrated 1.7% 25-year cancer-specific mortality for stage I, 15.8% in stage II, 30% in stage III and 60.9% in stage IV well-differentiated carcinomas.[10] Multiple studies and scoring systems have reinforced the significant predictive value of extracapsular extension, tumor size and distant metastasis.[11-13] Any evidence of extracapsular invasion distinguishes the tumor as T4 in the TNM classification. An estimated 15% of differentiated thyroid carcinomas have extracapsular extension, with 10-year survival being nearly half that of intrathyroid carcinoma patients.[14] The prognostic value of nodal involvement is controversial, with data describing both no effect on mortality, and a statistically significant OR of 1.9 over patients with no lymph node metastasis.[15,16]

Histology largely impacts survival in thyroid cancer. A recent study on selective US populations reported descriptive statistics of thyroid carcinomas. In this study, papillary and follicular carcinomas, also known as well-differentiated carcinomas, accounted for approximately 80% and 11% of malignant thyroid cancers, respectively. The overall relative survival of papillary carcinoma was 93%, while follicular carcinoma relative survival was 85%. Hurthle cell carcinomas compose approximately 3% of thyroid carcinomas, with 76% overall relative survival.[17] Undifferentiated (anaplastic) thyroid carcinomas represent only a small fraction of all thyroid cancers, and result in nearly complete 5-year mortality.[18] Poorly differentiated thyroid cancers are considered to be intermediate in stage, with higher rates of recurrence and mortality than well-differentiated tumors.[19]

In addition to stage and histology, several other factors have been found to be associated with higher mortality in thyroid cancer. Age has consistently been found to be a strong predictor of mortality, and it is included in nearly all prognostic scoring systems.[6-9] Older age is associated with lower relative survival in papillary, follicular and medullary carcinomas.[20] Thyroid carcinoma TNM staging is unique among other TNM classifications in accounting for age, with patients greater than 45 years old having a significantly higher stage, and associated higher risk of mortality.[9,15] Thyroid malignancies are less common in young patients (<18), and have better survival despite the higher risk of nodal metastasis.[21] Additionally, several studies have found that in well-differentiated thyroid cancer, male sex was associated with multiple recurrences and higher mortality, although incidence is much greater in women.[22,23] Advanced stage and larger primary tumor diameter have been found to be greater in men than women.[24]

Racial disparities in cancer are well documented, and have likewise been seen in thyroid cancer. In comparison to non-Hispanic White populations, Asian-Americans have been found to have improved survival, while African-Americans had worse survival.[20,25] In contrast, incidence has been reported to be lower in Blacks subjects, and higher in Chinese, Japanese, Hawaiian and Filipinos in comparison to White populations.[26-28] The significant association between socioeconomic status/position (SES) and cancer incidence and survival has also been extensively studied.[29-32] Economic deprivation has been associated with increased risk of death in various cancers.[30] SES variables recorded in the US census have been linked to patient addresses to identify measures such as education, working class, and poverty. This census-based methodology has been validated, and is increasingly utilized in investigations of cancer incidence and survival.

[30,31,33] Census block and tract levels have been found to be effective measures of socioeconomic position.[34] However, county-level SES information has also been found to be significantly associated with prostate cancer treatment choice, as well as cervical cancer incidence and survival.[35,36] The limited-use SEER dataset has linked each case to US Census variables at the county level, with data that includes median income, percent under the poverty line, and rural versus urban identification.

One study specifically addresses SES in thyroid cancer. Among 327 patients, the investigators found lower 10-year overall survival in the lowest income quartile (median income by zip-code), and worse stage at diagnosis associated with lower income, but no thyroid cancer-specific survival difference, and no difference in survival based on occupational prestige. This study also found no survival difference based on ethnicity, insurance status and marital status. However, this was a small study, with over10% loss to follow-up in a select geographic region. [37] In contrast, the study presented here includes a much larger population using the national cancer registry. The large population results in greater power to detect differences in survival. SEER data also have minimal loss to follow-up and is more representative of the US population.

Part 2: Significance

Mortality due to thyroid cancer varies greatly. Though the majority of cases will have a high likelihood of survival, not all outcomes are easily predictable. Several prognostic variables have been well characterized for differentiated thyroid cancer. This study will expand current knowledge in part by including poorly differentiated carcinomas and SES associations on a population basis. As thyroid cancer incidence

continues to rise, it will become increasingly important to determine the risk of mortality based on factors known at diagnosis. This information may help clinicians titrate surgical and non-surgical treatment to better match the aggressiveness of the disease. Nomograms also provide a means of educating clinicians-in-training regarding relative predictive strength of various known risk factors. Finally, risk stratification using a consistent and reliable algorithm such as a nomogram provides researchers with standardized case classification, which improves internal and external consistency among clinical studies.

While several studies have estimated overall thyroid cancer mortality, this information has not been applied to the individual. Nomograms are increasingly utilized and practical tools that allow for the prediction of individual risk, which have been used for multiple cancers such as prostate cancer or oral cavity squamous cell carcinoma.[38-40] A prediction tool that estimates risk of mortality based on individual variables will provide consistent prognostic information to the clinician and newly diagnosed thyroid cancer patient, and improve therapeutic management.

The use of census data linked to cases by SEER will afford specific investigation into the association of socioeconomic variables and thyroid cancer mortality. Independent predictors of mortality will be unified into a nomogram. Successful completion will result in a tool that can provide reliable prognostic information to the clinician and newly diagnosed differentiated thyroid cancer patient at the individual level. This information can be used to better inform treatment decisions by allowing an accurate estimation of risk of death from disease.

Part 3: Objectives

1. Confirm and quantify the association between predictor variables and 10-year cause-specific mortality in differentiated thyroid cancer cases from a historical cohort selected from the SEER national cancer registry.
2. Develop a nomogram that will predict 10-year cause-specific mortality in differentiated thyroid cancer patients based on analysis of confirmed prognostic variables using SEER data.
3. Internal validation of nomogram performance.

Part 4: SEER Cancer Registry:

SEER Introduction

Surveillance Epidemiology and End Results (SEER)[41] is a program of the National Cancer Institute that measures incidence of cancers in the United States. It is a population-based registry, which was established after the National Cancer Act of 1971. SEER functions to collect, analyze and distribute cancer incidence information with the goal of improving cancer prevention and outcomes. All cancer cases in 18 regions are reported to SEER from local cancer registries.

Cancer statistics began to be collected in 1973 in seven regions. Since that time, the program has expanded to include greater geographic and ethnic populations (American Indians, Native Alaskans, rural African-Americans). The current Limited-Use SEER Dataset contains cases diagnosed from 1973 to 2005. The population is a nonrandom sample that represents 26.2% of the total US population, collected from Connecticut, New Jersey, Atlanta, Kentucky, Louisiana, Rural Georgia, Detroit, Iowa,

Hawaii, New Mexico, Seattle-Puget Sound, Utah, San Francisco-Oakland, San Jose-Monterey, Los Angeles, Remainder of California, Arizona, and Alaska. Over-representation of minority groups is intentional, with the purpose of improving understanding of racial disparities in health. All minorities other than Blacks are proportionally more represented in SEER than would be expected if the sample reflected the true ethnic distribution of the US.[42]

SEER Case Selection

Mandatory reporting of inpatient and outpatient cancers has been legislated state by state, but national submission of cancer cases is voluntary. Various agencies are involved in the reporting of cases, including clinics, hospitals, labs, nursing homes, and other treatment centers or organizations. Submission of patient information to state cancer registries are exempt from the requirements of informed consent defined in the Privacy Rule of the Health Insurance Portability and Accountability Act of 1996.[43] Active case finding occurs at the local registry level to ensure completeness of each registry. Reportable tumors are limited to new primary cancers. The World Health Organization (WHO) has published the International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3), which lists reportable cancer categories.[44] SEER recodes all tumors with ICD-O-3 codes, and publishes updated validation lists for reference.[45]

SEER Data Collection

Mortality is strictly recorded and confirmed to identify the patient and the cause of death. These data are primarily collected from the National Center for Health Statistics, through the National Vital Statistics System, which collects all legally registered deaths

within the 50 States, 2 cities (Washington, DC, and New York City), and 5 territories (Puerto Rico, the Virgin Islands, Guam, American Samoa, and the Commonwealth of the Northern Mariana Islands). State agencies may also submit death information directly to SEER. Death certificate records are linked with SEER through the process of Death Clearance, which confirms cause of death, as well as assures reporting of new cases identified at death or autopsy. If data are incomplete, SEER conducts physician follow-back or contacts the associated facility to satisfy database requirements, including cause of death.

SEER performs active follow-up to update patient information. This includes identification of out-of-date patient information, and contacting the patient, family members, providers or others to confirm information. Passive follow-up occurs when databases are linked, usually at the state level. This includes linking of information from the department of motor vehicles, voter registration, the Centers for Medicare and Medicaid Services, and others. Survival analysis was used to right censor survival greater than 10 years, as well as any cases lost to follow up. Mortality data are obtained from the National Center for Health Statistics, through the National Vital Statistics System, which collects all legally registered deaths within the United States. The robust nature of mortality records and cause of death records in SEER will result in minimal loss to follow-up.

Population data are supplied by the United States Census Bureau at the county level, and are linked to SEER data in the supplied limited use dataset. This provides the opportunity to study measures such as percent below poverty, median family or household income level, education level and urban/rural classification.

SEER Data Entry

Data are submitted on a secure electronic Web-based application (SEER*DMS) by participating organizations. SEER has collaborated with the North American Association of Central Cancer Registries (NAACCR) to establish uniform data reporting.[46] New records are reviewed by multiple layers of automated and manual processing to ensure completeness and uniqueness. New cases are matched against the existing database to eliminate duplicate records. Incorrect records are manually evaluated and reconfirmed with the submitting party. Errors and edits are recorded in an audit log. Data collection was obtained at diagnosis and follow-up using strict privacy assurances at the regional and national registry levels. Data are de-identified and sensitive material is limited from access before distribution from the National Cancer Institute. No new data was required for this proposed study. Cancer reporting is excluded from informed consent requirements, but the general Privacy Rule is applied.

SEER Quality Control

Staff members of regional SEER registries conduct quality control studies in even number calendar years to evaluate case finding, coding and reliability. Training occurs in odd number calendar years. Conferences to address problems are conducted annually by the National Cancer Registrars Association. In addition, each registry has a Data Quality Profile that measures standards of data submission to the SEER program. The registry is stored in an Oracle database, and managed by the information technology staff that ensures its integrity and security.

CHAPTER 2: METHODS

Part 1: Study Design

A historical cohort of differentiated thyroid cancer cases identified at diagnosis, recorded in the SEER national cancer database from 1985 to 1995, was evaluated for an association between various prognostic variables with subsequent 10-year cancer-specific mortality. The variables to investigated included age, gender, tumor size, lymph node involvement, metastasis, extracapsular extension, and histology. Several independent or county-level socioeconomic measures, such as race, marriage status, median income, percent with high school education, and percent below poverty were also evaluated. The variables with the strongest prognostic values identified in survival analysis were developed into a nomogram, a tool that will enable patients and providers to predict 10-year cancer-specific mortality using individual patient risk factors.

Part 2: Case Selection

The SEER dataset that was available at the time of this study contained cases diagnosed from 1973 to 2005. To obtain 10 year follow-up information, the last year of diagnosis for inclusion was 1995. Within the census-linked SEER limited-use dataset, socioeconomic measures such as median income were only available starting with the 1990 census. Census data are collected decennially, and data should be considered valid within 5

Table 1: Inclusion & Exclusion Criteria

Inclusion	Exclusion
Primary thyroid cancer cases	Cases of thyroid cancer identified by autopsy or death certificate only
Cases diagnosed from 1985 to 1995	Anaplastic, Undifferentiated, Medullary
Received primary surgical treatment	Did not receive surgery, or Unknown if received surgery
Histologically-confirmed Diagnosis	

years of collection due to population shifts.[34] Therefore, an 11-year cohort approximately centered on the 1990 census with 10-year follow-up was selected.

Cases with primary thyroid cancer that were pathologically proven via histology were selected. Treatment of differentiated thyroid tumors includes surgical resection. Although there is some variability in the extent of surgery and adjuvant radiation, surgery is widely accepted as the initial treatment of choice. To ensure both pathology and appropriate treatment, all included cases were surgically treated. In order to avoid inclusion of incidental tumors that were not identified prior to death, and thus provide minimal survival information, those cases that were reported from autopsy or death certificate alone were excluded. Among the 16,816 cases of thyroid cancer, 15,412 received surgery and histologic confirmation prior to autopsy or death.

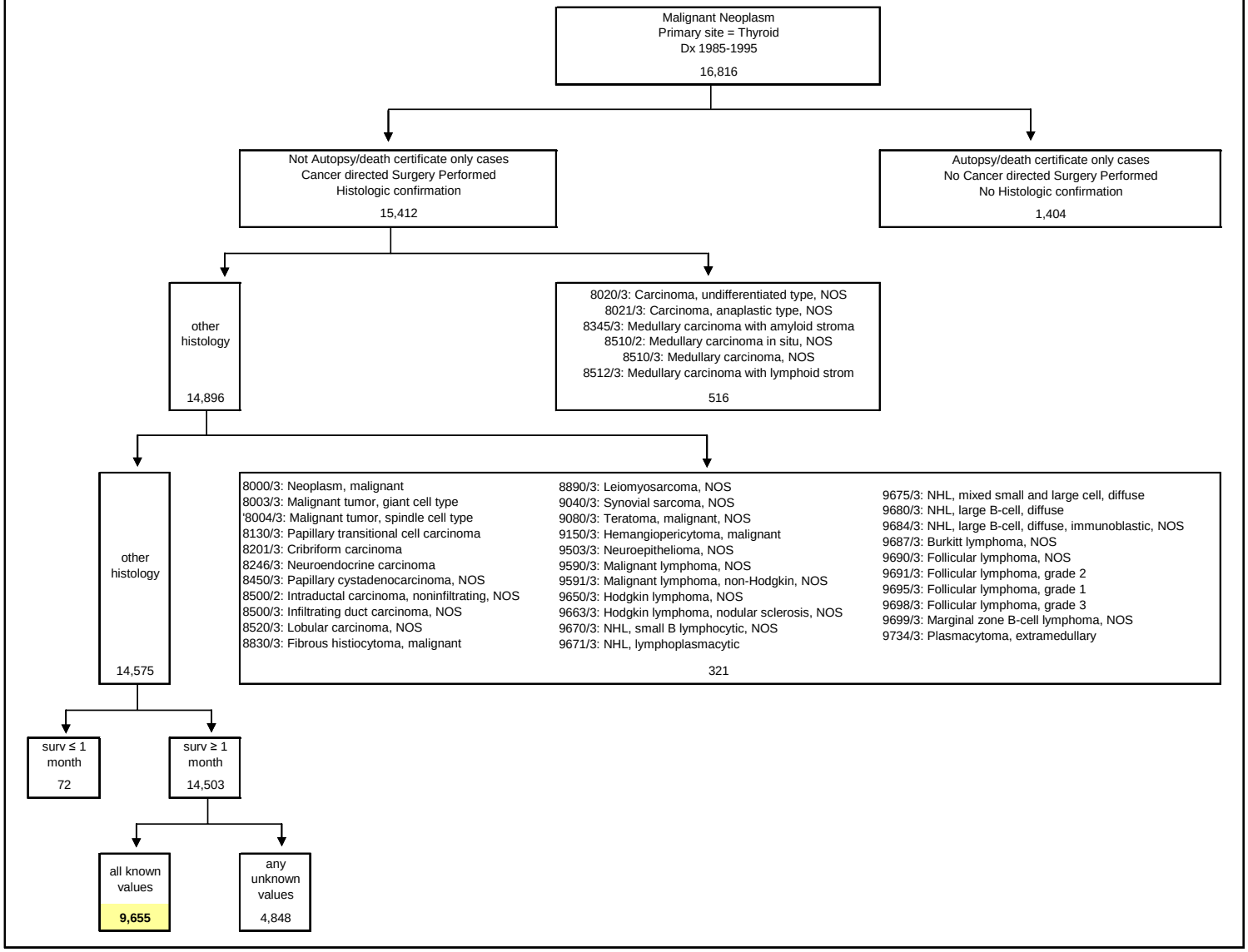
The scope of this study included well-differentiated and poorly-differentiated tumors. The clinical behavior of anaplastic and medullary tumors varies significantly from that of other differentiated tumors. Therefore, 516 anaplastic, undifferentiated and medullary tumors were excluded from analysis. Among the remaining 14,896 cases, 321 more were excluded based on histological classification review by an expert thyroid pathologist identified tumors that were unlikely to be of primary thyroid cellular origin. Beyond this, 72 patients that died within the first month of diagnosis were excluded from analysis to avoid possible surgery-related mortality. Finally, all cases with significant missing data for values of interest, such as nodal involvement, tumor size, or metastasis, were excluded. The final population available for analysis was 9,655 (Table 1, Figure 1).

Part 3: Predictors (Independent Variables)

Seven main predictors known to be associated with higher thyroid cancer mortality were evaluated to identify the independently predictive variables. These included age, gender, tumor size, lymph node involvement, distant metastasis, extracapsular extension, and histology. Additionally, five secondary socioeconomic predictor variables were also assessed for predictive strength. Individual level variables of race, marital status, and county-level measures of income, high school education, and poverty, were evaluated.

Figure 1: Case Selection

SEER*Stat Database: Incidence - SEER 17 Regs Limited-Use + Hurricane Katrina Impacted Louisiana Cases, Nov 2007 Sub (1973-2005 varying)



Part 4: Outcome (Dependent Variable)

The outcome of interest was 10-year cancer-specific mortality. Deaths were coded using the SEER recorded survival time and cause of death. Survival time is calculated in months by SEER using the date of diagnosis and either the date of death, date last known to be alive, or follow-up cutoff date December 31, 2005. Patients lost to follow-up will be censored at the time of last contact. All cases were followed to December 2005, unless they died of causes other than thyroid cancer. All cases with a cause of death that was not thyroid cancer were right censured at the time of death, or censored at December 2005. Subjects with a cause of death due to thyroid cancer were identified as 3.8% of the study population, while those that died of other causes composed 12% of the population.

Part 5: Statistical Analysis

Univariable Analysis of continuous predictor variables was achieved using simple Cox proportional hazards regression. Kaplan Meier curves and log-rank tests for used to estimate predictive significance of categorical variables. A Cox proportional hazards regression with Breslow method for ties was used to identify a set of independent predictors for thyroid cancer mortality using backwards stepwise elimination. Significance was determined as a $p<0.05$. Schoenfeld residual correlations were tested to verify proportional hazards assumptions in conjunction with Kaplan-Meier observed-versus-predicted curves. Overly influential observations were identified by identifying DFBETAS>u for a u=0.2 of the standard error. A nomogram was constructed based on the results from the stratified Cox proportional hazards regression analysis. Bootstrapped bias-corrected estimates of the AUC c-index assessed the performance of the nomogram.

An estimated bootstrapped calibration curve was created to inspect predictive accuracy. Validation and calibration were conducted on random samples selected with replacement using computerized bootstrapping procedures with 200 replications. Primary statistical analysis was conducted using STATA 10.1 produced by StataCorp LP (College Station, Texas), DFBETAS and Nomogram construction was conducted on R 2.8.0 produced by the R Foundation for Statistical Computing (Vienna, Austria), and the Design Package 2.1-2 for R by Frank E. Harrell, Jr (Nashville, Tennessee).

CHAPTER 3: RESULTS

Part 1: Case Selection & Cause of Death

After case selection, one observation was excluded as an outlier as described in the results section, resulting in the identification of 306 thyroid cancer-specific deaths amongst the final 9654 cases within 10 years of diagnosis (3.2%). The survival function is seen in Figure 2, demonstrating a 10-year cause-specific survival of 96.7%. Thyroid cancer was the cause of disease in 3.7% of the study population, while 12% of cases were right censored for a cause of death due to another disease process (Table 2). Twenty-two percent of those with other causes of death were due to heart disease. The next most common alternate cause of death was lung and bronchus disease (8.5%), followed by cerebrovascular disease (6.5%). Among the 1158 cases with other COD, 745 died within 10 years of diagnosis (7.7% of total study population).

Evaluation of the excluded observations with unknowns in comparison to the included observations is seen in Table 3. The mean age for the 4848 excluded observations was 46.8 years, while the mean age of cases was 44.3 years. This difference of 2.5 years was found to be statistically significant with a two-sample test of means ($p<0.001$). A two-sample test of proportions showed that gender distribution was not significantly different between the two groups ($p=0.064$). Lastly, a Pearson's chi-squared test of independence showed that histology was significantly different in the two populations with a $p<0.001$. Included observations contained proportionally more papillary cases, and fewer follicular and other cases.

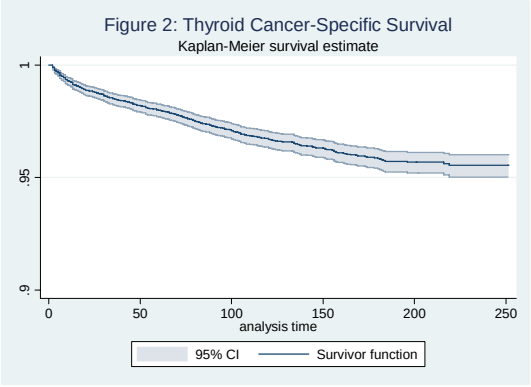


Table 2: CAUSE OF DEATH

SEER COD	Frequency	Percent	COD	Frequency	Percent
Thyroid	362	3.7	Death due to thyroid cancer	362	3.7
Alive	8,134	84.3	No TC Death	9,292	96.3
Other COD	1,158	12.0			
total	9,654	100			

SEER cause of death (COD) was used to identify those subjects with a COD due to thyroid cancer. Patients alive at the end of the study were censored at the end of the study, while cases with other cause of death were censored at the time of death.

Table 3: Excluded Cases With Unknown Values Compared With Included Cases

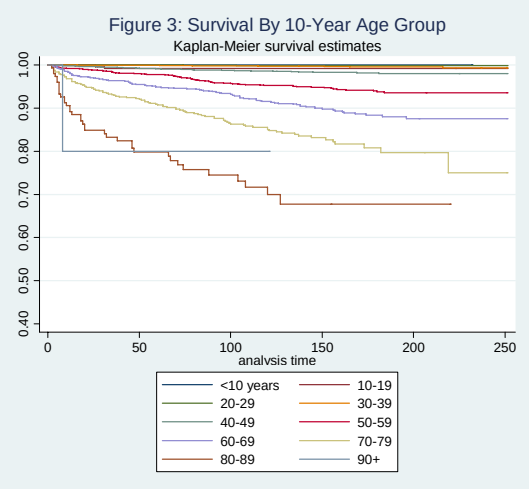
Variable	Excluded	Included	T / Z / X^2	p	Excluded KM Survival	Included KM Survival
N	4848	9655	-	-	-	-
Mean Age	46.8 (17.1 sd)	44.3 (15.8 sd)	8.74	0.000	-	-
Gender	Female 3617 (74.6%)	7343 (76%)	-1.85	0.064	95.3	97.6
	Male 1231 (25.4%)	2312 (24%)			92.5	93.9
Histology	Papillary 3848 (79.4%)	8308 (86.1%)	124.04	0.000	96.5	97.5
	Follicular 878 (18.1%)	1249 (12.9%)			89.8	94.3
	Other 122 (2.5%)	98 (1%)			68.0	62.8

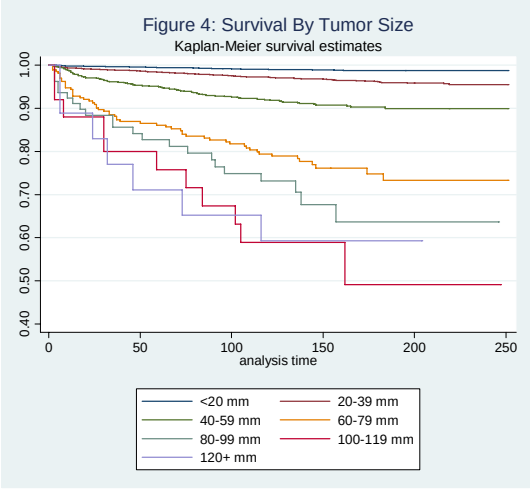
Two-sample mean comparison t-test was conducted for age. A two-sample test of proportions was conducted for gender. A Pearson's chi-squared test of independence was conducted on the tri-level histology variable. Kaplan-Meier (KM) 120-month survival probability is listed.

Part 2: Univariable Analysis

Main Predictors

Seven main predictor variables were investigated, including age, gender, tumor size, nodal involvement, distant metastasis, extracapsular extension and histology. Initial descriptive analysis and inspection of frequency distributions was conducted. All continuous variables were found to be normally distributed. A univariable Cox proportional hazards regression was used to test significance of continuous predictors, while a log rank test for equality was used to evaluate significance of categorical variables. Age by year was entered as a continuous variable, and shown to have a trend of increasing mortality with increasing age (Figure 3). Of note, Kaplan-Meier survival curves in women of child-bearing age did not show notable deviation from the overall trend of increasing mortality. Tumor size by millimeter (mm) was also found to be normally distributed with higher risk of mortality trended in those with larger tumors (Figure 4). Both age and size were found to be a significant predictors of mortality, with a hazard ratio (HR) of 1.09 ($p<0.001$) and 1.03 ($p<0.001$), respectively.





Three dichotomous variables, gender, nodal involvement, and distant metastasis, were analyzed. Nodal involvement was dichotomized from SEER categories as listed in Table 4 SEER did not contain a variable describing distant metastasis for the years of interest. Therefore, the SEER variables describing tumor extension and nodal involvement were used to create a new variable that identified cases with either metastatic tumor extension or distant lymph node involvement (Figure 5). Male gender, lymph node involvement and distant metastasis were shown to be predictive of thyroid cancer mortality with respective HR of 2.5 (Figure 6), 3.2 (Figure 7), and 17.0 (Figure 8), all with $p<0.001$.

Table 4: Lymph node involvement coded from SEER categories

SEER Lymph Node Categories	Node
No lymph node	No lymph node (N0)
ipsilateral cervical node	Any lymph node (N1)
bilateral, contralateral or midline cervical	
tracheoesophageal (posterior medistinum), upper anterior mediastinum, mediastinum NOS	
Region lymph node NOS	
distant - submandibular, submaxillary, submental	
distant other	

Figure 6: Survival By Gender

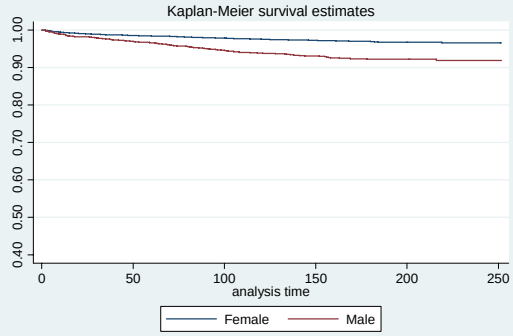
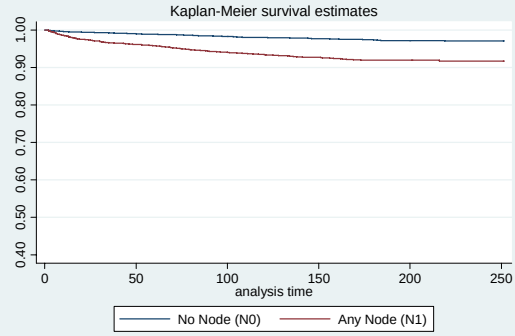
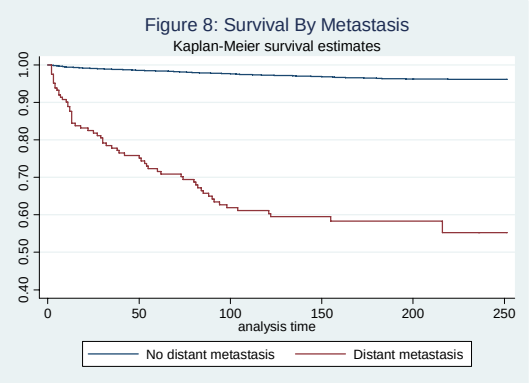
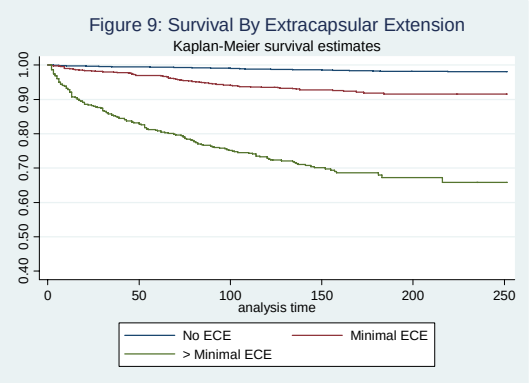


Figure 7: Survival By Nodal Involvement







Extracapsular extension (ECE) and histology were categorized with three levels. The SEER variable, Extension, details the extent of tumor invasion. Due to SEER categorization, minimal ECE was defined was pericapsular extension, which includes invasion into the capsule, strap muscles (sternothyroid, omohyoid, sternohyoid, sternocleidomastoid), and nerves including the recurrent laryngeal and vagus. Extension beyond this, involving any further soft tissue, vessels or bone was defined as greater than minimal extracapsular extension. (Table 5. Greater ECE was associated with increased risk of mortality ($p<0.001$, Figure 9).

Table 5: Categorization of Extracapsular Extension

SEER Extension	Extracapsular Extension (ECE)
confined to thyroid	No extracap ext
multifocal confined to thyroid	
localized NOS	
through capsule - not beyond	
pericapsular	Minimal extracapsular extension
major blood vessels, esophagus, larynx	> Minimal extracapsular extension
trachea, skeletal muscle, bone	
further contiguous extension, mediastinal ti	
further extension or metastasis	
metastasis (1988+)	

Figure 5: Categorization of Metastasis variable from SEER categories

SEER Extension
in situ - noninvasive
confined to thyroid
multifocal confined to thyroid
localized NOS
through capsule - not beyond
pericapsular
major blood vessels, esophagus, larynx
trachea, skeletal muscle, bone
further contiguous extension, mediastinal tissue
further extension or metastasis
metastasis (1988+)

Metastasis
no distant metastasis
distant metastasis

SEER Node
ipsilat cervical node
bilata, contralat or midline cerv
mediastinal
Region LN NOS
distant - submand subment
distant other

Histology was categorized as Papillary, Follicular, or Other. The last category contained a wide variety of less-common tumors that included poorly differentiated thyroid cancer, as delineated in Table 6, which lists the SEER recorded ICD-3 codes. The list of thyroid tumor histologies were reviewed by a pathologist to confirm appropriate categorization into the Papillary, Follicular, or Other groups. Papillary histology was associated with the best survival. Follicular increased risk of mortality by a small degree, but mortality was significantly higher in the Other category (Figure 10). Histologic categorization into these 3 groups was a significant predictor of mortality ($p<0.001$).

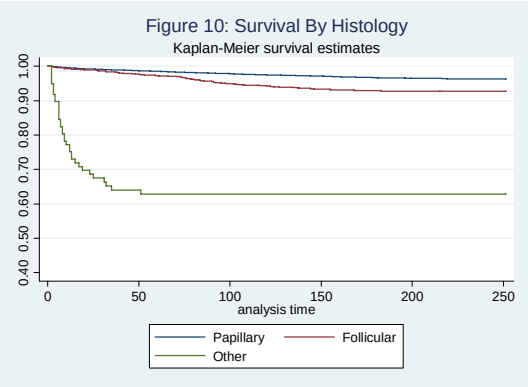


Table 6: Histologic Categories		
SEER Histology	differentiation	Histology
Papillary carcinoma, NOS	well	papillary
Papillary carcinoma, follicular variant	well	
Papillary adenocarcinoma, NOS	well	
Nonencapsulated sclerosing carcinoma	well - PTC variant	
Intracystic carcinoma, NOS	well - PTC variant	
Follicular adenocarcinoma, NOS	well	follicular
Follicular adenocarcinoma well differentiated	well	
Follicular adenocarcinoma trabecular	well	
Follicular carcinoma, minimally invasive	well	
Oxyphilic adenocarcinoma	well - FTC variant	
Carcinoma, NOS	poor	other
Squamous cell carcinoma, NOS	range	
Adenocarcinoma, NOS	poor	
Giant cell carcinoma	poor	
Clear cell adenocarcinoma, NOS	poor	
Spindle cell carcinoma	poor	
Large cell carcinoma, NOS	poor	
Giant cell and spindle cell carcinoma	poor	
Small cell carcinoma, NOS	poor	
Acinar cell carcinoma	poor	
Pleomorphic carcinoma	poor	
Papillary squamous cell carcinoma	range	
Squamous cell carcinoma, spindle cell	range	
Mucoepidermoid carcinoma	range	
Mucinous adenocarcinoma	poor	
Carcinosarcoma, NOS	poor	

Secondary Predictors (Socioeconomic Variables)

SEER includes the variables of race and marital status. In addition to these individual variables, county-level measures were available using census data linked to SEER. These included median household income, percent with high school education, and percent below poverty. White race composed the vast majority of the study cohort, accounting for 82%, followed by asian or pacific-islanders who composed 12.2% of the population, while blacks accounted for only 4.6% of the cases. There was no difference in survival seen between White and Non-white cases with thyroid cancer (HR=1.09; p=0.509). Marital status was significantly associated with survival, with 10-year predicted survival probability of single (never-married) cases estimated at 98.7% in comparison to 96.8% in married cases, 96.4 in divorced or separated subjects, and 87.4% in the widowed ($p<0.001$).

Median household income within a county, as estimated in the 1990 census, ranged from $12,990 to $54,800, with a mean of $35,561. The quartile with the lowest household income had worse survival as compared with the other 3 quartiles. The mean in the lowest quartile was $26,364, with a predicted 10-year cause-specific survival probability of 96.1%. In contrast, the mean in the other quartiles was $38,591, with a predicted survival probably of 96.9%. This difference was found to be significant (p= 0.043). The percent of people in a county below poverty (%Pov) was identified using 1990 census data, and determined to be 150% below the poverty level by age for population. The range of %Pov was 2.4-43.48%, with a mean of 11.1%. In the quartile with the highest poverty, the mean %Pov was 20.4% while the mean %Pov of the other quartiles combined was 9.6%. This difference was not found to be significant (p=0.094). Finally, the percent of people with less than a high school education (%HS) in a county was found to be significantly associated with survival. The range of %HS was 5.3-50.3%, with a mean of 20.6%. The combined quartiles with the lowest education had a hazard ratio of 1.6 times that of the highest educated quartile (p<0.001) with a mean %HS of 23.3% as compared with 13.2%. Predicted survival for the most educated quartile was 97.5% as compared with 96.4% within the combined lower quartiles. The univariable tests of significance are summarized in Table 7. All seven main predictors were found to be significant predictors of mortality in differentiated thyroid cancer.

Table 7: Univariable test of significance for independent variables

Variable	Z / X^2	p
Age (year)	23.04	0.000
Gender	82.26	0.000
ECE	1375.63	0.000
Tumor Size (mm)	26.63	0.000
Node	135.57	0.000
Distant metastasis	768.94	0.000
Histology	473.27	0.000
race	0.44	0.509
household income (quartiles)	4.11	0.043
% people below poverty (quartiles)	2.81	0.094
% people with < high school education (quartile)	12.65	0.000
married	175.8	0.000

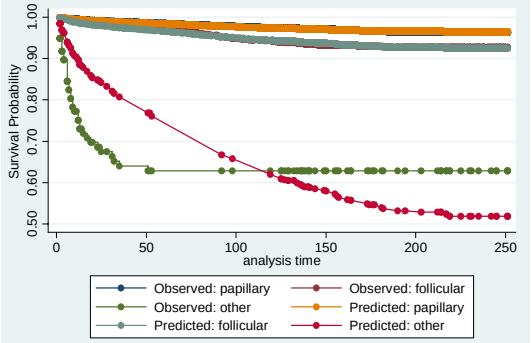
Continuous variables were assessed using univariable Cox proportaional hazards regression. Log-rank test evaluated significance for categorical data using a X^2 test.

Part 3: Model Fit

Table 8: Schoenfeld residual correlation test of PH assumption				
Variable	rho	chi2	df	Prob>chi2
Age	0.06935	1.65	1	0.199
Gender (male)	0.05112	1	1	0.3183
Size	0.11763	5.02	1	0.0251
Node	-0.01564	0.11	1	0.7424
Metastasis	0.0482	1.17	1	0.2786
ECE (minimal)	-0.03128	0.4	1	0.5264
ECE (> minimal)	-0.0407	0.72	1	0.3956
Histology (follicular)	0.03307	0.43	1	0.5124
Histology (other)	-0.21446	27.08	1	0.0000
global test		39.97	9	0.0000
unstratified Cox proportional hazards model				

To confirm the appropriateness of the Cox proportional hazards model, Schoenfeld residuals correlations were evaluated. Additionally, observed-versus-expected Kaplan-Meier curves were also graphed, showing the same results. The Schoenfeld residuals for both size and "Other" histology were found to be significantly associated with survival time (p=0.025 & p<0.001, Table 8). The violation of the PH assumption by the "Other" category of the histology variable is corroborated by the Kaplan-Meier curve of observed and expected survival in Figure 11. The observed and predicted lines cross, and the observed line does not appear to have a proportional hazard across categories over time, showing violation of the PH assumption. Therefore, the model was stratified on histology.[47] Goodness of model fit was tested with the Schoenfeld residual correlation again, after stratification.

Figure 11: Observed Versus Expected KM Curves for Histology

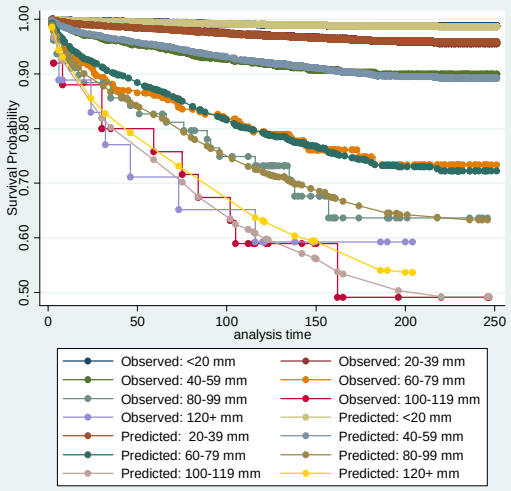


The Schoenfeld residual for Size was found to be correlated with time (p=0.032) after stratification (Table 9). In reviewing the KM curve of observed versus expected survival in size grouped by 2 cm increments, the majority of tumors appear to meet the PH assumptions (Figure 12). Size may present some violation of the PH assumption in very large sized tumors, but overall it does appear to meet the PH assumption in the most

common size range. Therefore, Size remained in the model without adjustment for time-dependence or stratification.

Table 9: Schoenfeld residual correlation test of PH assumption				
Variable	rho	chi2	df	Prob>chi2
Age	0.06678	1.53	1	0.2168
Gender (male)	0.05882	1.3	1	0.2539
Size	0.11645	4.58	1	0.0323
Node	-0.01731	0.13	1	0.7204
Metastasis	0.04147	0.82	1	0.3648
ECE (min)	-0.03148	0.4	1	0.5255
ECE (>min)	-0.03597	0.56	1	0.4552
global test		9.74	7	0.2035
Cox proportional hazards model stratified on histology				

Figure 12: Observed Versus Expected KM Curves for Size



Part 4: Multivariable Analysis

All predictors with a p>0.25 were combined into the preliminary multivariable model. Race was also included to account for any potential confounding, although the p=0.509. The preliminary main effects model is written below.

$$\mathbf{h_{histology}(t,X) = h_o(t) \cdot e^{y}}$$

***where y* = β1Age+β2Gender + β3Size + β4Node + β5 Metastasis + β6ECE + β7Race + β8(%Pov) + β9(household income) + β10(%HS) + β11(marital status)**

A stratified Cox proportional hazards (PH) regression was performed, using the seven main predictors and five secondary predictors (Table 10). Using backwards stepwise elimination, significant terms were selected. All seven main predictors were found to be significantly associated with mortality, however, all secondary predictors fell out. Interactions between Age and the other variables, as well as between Node, Metastasis , and ECE were also evaluated. Backwards stepwise elimination was conducted to identify the significant terms which are summarized in Table 11. Three interactions, Age*Size,

Age*ECE, and ECE*Node, were found to be significant ($p<0.001$, $p=0.049$, $p<0.001$ respectively). Although several significant interactions were identified interactions are cumbersome in nomograms. Additionally, the addition of these interactions did not change the predictive performance of the final nomogram and where therefore not included in the final model. The final model is listed below, with the main predictors of age, gender, size, metastasis, nodal involvement, extracapsular extension and histology (Table 12).

$$h_{histology}(t,X) = h_o(t) \cdot e^{\beta 1 Age + \beta 2 Gender + \beta 3 Size + \beta 4 Node + \beta 5\ Metastasis + \beta 6 ECE}$$

Table 10: Preliminary Main Effects Multivariable Model

Variable	category	HR	β	β SE.	Z / X^2	p	β 95% CI	
Age	year	1.0691	0.0669	0.0042	15.96	0.000	0.0586	0.0751
Size	mm	1.0175	0.0174	0.0019	9.38	0.000	0.0137	0.0210
Gender	Male	1.3819	0.3234	0.1171	2.76	0.006	0.0939	0.5529
ECE	Minimal	2.9631	1.0863	0.1529	145.75	0.000	0.7866	1.3859
	> Minimal	6.0738	1.8040	0.1504			1.5093	2.0987
Node	Any Node	2.3850	0.8692	0.1235	7.04	0.000	0.6271	1.1114
Metastasis	Metastasis	1.9196	0.6521	0.1753	3.72	0.000	0.3086	0.9957
Race	Non-white	0.9953	-0.0047	0.1434	-0.03	0.974	-0.2857	0.2763
% people below poverty	highest % of people below poverty	0.9432	-0.0585	0.2005	-0.29	0.771	-0.4514	0.3344
household income	higher income quartiles	0.7593	-0.2754	0.1621	-1.7	0.089	-0.5931	0.0423
% high school educated	lower educated	1.0031	0.0031	0.1471	0.02	0.983	-0.2852	0.2914
Marital Status	Never married	0.8052	-0.2166	0.2028	5.53	0.137	-0.6142	0.1809
	Divorced/Separated	1.4945	0.4018	0.2056			-0.0011	0.8047
	Widowed	1.0166	0.0164	0.1568			-0.2909	0.3238

Z-score determined for continuous variables, while a Wald test (X^2) was used for categorical data

Table 11: Multivariable model with significant interactions (stratified by histology)

Variable	category	HR	β	β SE	Z / X2	p	β 95% CI	
Age	year	1.1055	0.1003	0.0075	13.4	0.000	0.0856	0.1150
Size	mm	1.0512	0.0500	0.0077	6.53	0.000	0.0350	0.0650
Gender	Male	1.3210	0.2784	0.1077	2.59	0.010	0.0674	0.4894
Node	Any Node	5.7660	1.7520	0.1882	9.31	0.000	1.3830	2.1209
Metastasis	Distant Metastasis	1.9809	0.6835	0.1696	4.03	0.000	0.3511	1.0159
ECE	Minimal	43.3043	3.7683	0.5881	169.37	0.000	2.6156	4.9209
	> Minimal	0.9995	-0.0005	0.0001			-0.0007	-0.0003
Age x Size	Age x Size	0.9869	-0.0132	0.0099	-4.18	0.000	-0.0327	0.0062
Age x ECE	Age x ECE (min)	0.9789	-0.0213	0.0087	6.03	0.049	-0.0384	-0.0042
	Age x ECE (>min)	9.3823	2.2388	0.6824			0.9013	3.5764
Node*ECE	Node x ECE (min)	0.4214	-0.8642	0.2989	31.63	0.000	-1.4502	-0.2783
	Node x ECE (>min)	0.2538	-1.3711	0.2442			-1.8496	-0.8925

Z-score determined for continuous variables, while a Wald test (X^2) was used for categorical data

Table 12: Final multivariable model stratified by histology								
Variable	category	HR	β	β SE.	Z / X2	p	β 95% CI	
Age	year	1.0697	0.0674	0.0039	17.4	0.000	0.0598	0.0750
Size	mm	1.0176	0.0175	0.0018	9.54	0.000	0.0139	0.0210
Gender	Male	1.3751	0.3185	0.1089	2.92	0.003	0.1050	0.5320
ECE	Minimal	2.8679	1.0536	0.1522	148.02	0.000	0.7552	1.3519
	> Minimal	6.0100	1.7934	0.1480			1.5034	2.0834
Node	Any Node	2.4340	0.8895	0.1225	7.26	0.000	0.6494	1.1296
Metastasis	Metastasis	1.8759	0.6291	0.1738	3.62	0.000	0.2885	0.9696
Z-score determined for continuous variables, while a Wald test (X^2) was used for categorical data								

Part 5: Outliers

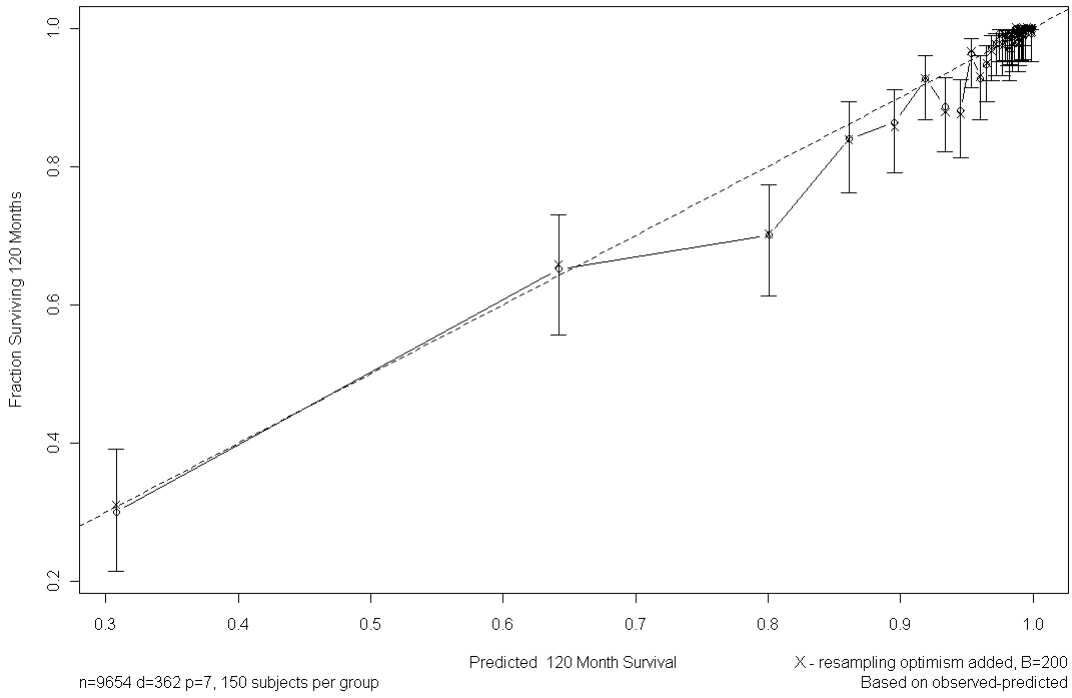
DFBETAS, or the difference in betas, were estimated to ensure that the model is not unduly influenced by unusual observations (see Appendix II). Observation 9109 had large, distinct DFBETAS in nearly all the predictor variables, and observation 1019 was notable when evaluating Size. Case 1019 had a reported tumor size of 55cm. Considering the extreme rarity of such a tumor, as well as the loss of discrepancy in the nomogram, this observation was removed from analysis. All other observations remained in the analysis to improve predictive accuracy.

Part 6: Model Validation

Model discrimination was assessed by the average calculated c-index, or the area under the curve (AUC), over 200 bootstrapped replications. Model calibration was evaluated by inspecting a calibration plot of the observed survival times against the predicted survival time (Figure 13). C-index for the stratified model with 9654 observations was 0.925. Bootstrapped calibration suggests that the model mildly overestimates survival when actual survival is approximately 75%. Otherwise, the model

has good predictive accuracy. Additionally, bias-correction shows that the model was not over-fit.

Figure 13: Calibration of Regression Model



CHAPTER 4: DISCUSSION

Part 1: Case Selection & Cause of Death

Estimated ten-year cause-specific survival of 96.7% is consistent with previously reported 10-year survival estimates of 96.5%.[1] Patients who died of other causes within the 10-year period were no longer at risk for death due to thyroid cancer. Their risk of thyroid cancer-specific death is then zero, although they did add time at risk prior to death. This results in a potential overestimation of predicted survival using Cox PH instead of cumulative incidence. However, the relatively small number of deaths due to other causes in the context of 8,133 subjects who lived to the end of the study (84%)

minimizes this effect. Additionally, the Cox PH model assumes independence of competing risks, while a cumulative incidence model does not. Although no formal test was done, lung and cardiovascular diseases are not known to increase risk of mortality in thyroid cancer, and are likely independent from the outcome of interest. Cox PH may be sufficient in modeling cause-specific mortality.

Excluded observations with unknown values were found to be significantly different in age and histology when compared with the included cases. The 2.5 year difference in age may be statistically significantly, however, there is little clinical significant between 44 and 46 years. It is interesting to note that the TNM staging system categorizes cases differently for patients 45 years of age and older, and the mean age of the included and excluded cases also seemed to split at age 45. The mean age of excluded cases was 46.8, which could suggest that those with worse prognosis were less likely to have complete data entry. Included cases did appear to have more Papillary histology, which has the best survival amongst the three histology types. There were proportionally fewer Follicular and Other histology cases in the analyzed dataset. Furthermore, the Kaplan-Meier survival probabilities for each histology level showed that included Papillary and Follicular cases had better survival than those that were excluded, and that included Other cases had worse survival than those that were excluded. These differences in age and histology may result in a model that overestimates predicted survival in cases with Papillary and Follicular histology or older age, and underestimates predicted survival in Other or younger cases. However, a model using imputed data for all unknown values using the hotdeck method resulted in a model with similar predictive discrimination, with a c-index of 0.928. This suggests that the differences noted in

histology and age are not of predictive significance in the final model, and do not effect the overall performance of the nomogram.

Part 2: Univariable Analysis

Univariable analysis showed that all seven of the main independent variables were significantly associated with survival in cause-specific differentiated thyroid cancer mortality with p-values <0.001. This is consistent with previously published literature. Among the secondary independent variables evaluated, race and percent of population below poverty were not significant predictors of survival, but marital status, median household income, and percent with less than high school education were significant. Although racial disparities in health and cancer-related mortality are well known, this was not reflected in the study results. Even when comparing White with Asian, or White with Black, no difference in survival was observed. Although SEER over-represents minority populations, incidence of thyroid cancer is highest in Asian and White populations who have better predicted survival than other races. With the population distribution of the United States, any racial difference in mortality may be obscured by the high incidence in the large proportion of White subjects in the cancer cohort.

Being widowed was associated with poorer survival, while being single was associated with the best survival. Most existing studies show that single status is associated with higher morbidity and mortality. Marital status may be associated with age, with younger people with better thyroid cancer survival being more likely to be single. Therefore, after reevaluating survival probability while accounting for age, the significance of marital status was no longer observed (p=0.709).

Median household income, percent of people with less than a high school education, and percent of people below poverty were found to be significant on the univariable level. This is also consistent with known socioeconomic disparities in health showing poorer survival in low income and low educated populations. The percent of people in a county below poverty (%Pov) was found to be insignificant. This is inconsistent with the association found between median household income and survival, however, unlike income, poverty is a normative value. These results suggest that 150% below the defined level of poverty does not reflect a meaningful measure of income in thyroid cancer mortality. Lastly, a difference in survival was observed in education quartiles (%HS). This is also consistent with studies showing improved health outcomes in various diseases in higher educated populations. The highest education quartile did have a 10-year predicted survival that was 1.1% greater than the lower educated quartiles.

Part 3: Model Fit

Proportional hazards assumes that predictors are linear and additive in relation to the log hazard. To confirm that Cox PH models would be appropriate, Schoenfeld residuals correlations and observed-versus-expected Kaplan-Meier curves were graphed to aid in clarifying violations of the PH assumption. The Schoenfeld residual is a measure of the difference between predictor values at each failure and the weighted average of the predictor for all subjects still at risk. A residual that is correlated with survival time suggests that the predictor violates the PH assumption. Although the Papillary and Follicular categories of histology met PH assumptions, the "Other" category in the Histology variable did not, thus the model was stratified on histology in order to obtain

independent baseline hazards for each category. The violation of the PH assumption by "Other" histology is not unexpected considering the heterogeneity of this group. A Cox proportional hazards model stratified on histology is an appropriate method of addressing violations of PH assumption by a single categorical variable, and results in a global Shoenfeld residual correlation of p=0.204 suggested overall model goodness of fit.

After stratification, Size continued to show potential violation of the PH assumption, which was not supported by the KM curves. Survival estimates for subjects with tumors greater than 12 cm in size showed less discrimination and greater divergence from predicted values, and crossing of KM lines. This may be due to the small number of observations with larger tumor size. More than 99.5% of cases had tumors <100 cm in size. Only 19 subjects (0.2%) reported a size of 12 cm or larger. The small number of observations in the largest size range increases the risk of distorting the Schoenfeld residual correlation. With generally concordant KM curves and overall goodness of fit, the stratified model was accepted to meet the PH assumption.

Part 4: Multivariable Model Selection

Multivariable regression with backwards step-wise elimination, with significance determined at $p<0.05$, revealed that the seven main predictors were significant after adjustment for covariates while all five SES measures were not. Age, gender, tumor size, nodal involvement, distant metastasis, extracapsular extension and histology were found to be significant independent predictors, which is consistent with previously published findings. Age is the largest predictor of mortality, with an increased hazard of mortality of 7% per year, or an increase in hazard of 96% for every 10 years. Size is also a

significant predictor, with a 9% increased risk of mortality per 5 mm, or 19% increased risk of mortality (HR 1.19) for each centimeter. For subjects who are young with small tumors, other risk factors may be more predictive than age or size, such as ECE, nodal involvement, metastasis, or gender. Cases with extracapsular extension are 6 times more likely to die in 10-year than those without any ECE, while those with pericapsular extension are 2.9 times more likely to die than those without ECE in the same time frame. Nodal involvement more than doubles the risk of morality (HR 2.43). The increased hazard is noted with any nodal involvement, including distant nodes. Although metastasis also includes distant nodes, nodal status was significant even after adjusting for the effect of distant metastasis. Surprisingly, after adjusting for node and ECE, distant metastasis only increases risk of mortality by 88%. Lastly, male gender increases risk of mortality by 38%.

Race and %Pov were not significant at the p=0.05 level in univariable analysis, and continued to be insignificant in the multivariable model. Although marital status, income and education were found to be independent predictors of survival on univariate analysis, the multivariable model shows that these SES measures are weaker predictors of survival than the seven main variables. Additionally, the use of county level information may obscure associations that may truly exist due to heterogeneity within counties. The use of census data, or aggregate information, may raise concerns of an ecological fallacy. However, Kreiger et al discuss the appropriateness of census SES data in such analysis.[34] Although census data is valid for individual SES analysis, the analysis is much improved with smaller regions of minimal heterogeneity such as census-tract level rather than zip-code or county level information.[30, 34]

Interactions of interest included all those with Age, because of the strong association with increased mortality known with Age. Interactions with Node, ECE and Metastasis were also of interest, because of the potential effect modification that may be seen in these related variables. After backward stepwise elimination, three interactions were found to be significant, including Age*Size, Age*ECE, and ECE*Node. It is notable that the effect of Age is modified by Size and ECE, and that ECE also interacts with Node. Predictive performance of the model with the significant interactions yielded no change in discrimination and more variability in calibration. While these are of interest, the construction of a nomogram limits the number of interactions that are reasonable to include. Due to the limitations of nomogram development, as well as the lack of improvement in predictive performance, these interactions were not included in the final model.

Part 5: Outliers

The DFBETAS (*difference of betas*) is a change of 0.2 standard errors in a regression coefficient with the removal of an observation. It identifies those observations that would effect the greatest change if removed from the model. One observation was repeatedly identified in all the predictor variables. This subject was an 83 year old female with a 2 cm thyroid carcinoma, NOS (Other histology), with distant lymph node involvement (metastatic), and greater than minimal extracapsular extension (major blood vessels, esophagus and/or larynx) who had survived the length of her time in the study period (>120 months). This case is an extremely unusual case, with poor prognostic factors in an elderly patient, who had good long-term survival. Although unusual, this

observation remained in the dataset with the expectation that it would improve model performance on the real data. Observation 1019 is a 28 year old female with a follicular variant of papillary carcinoma (papillary histology) with no lymph node involvement, extracapsular extension or metastasis, who also survived until the end of the study period. This case is less unusual except for the 55 cm tumor. The closest tumor to this was less than half the size (22 cm). The improbability of such a large papillary tumor in a young person and the unlikelihood that this point would add meaningfully to the predictive ability of the nomogram resulted in the exclusion of this observation.

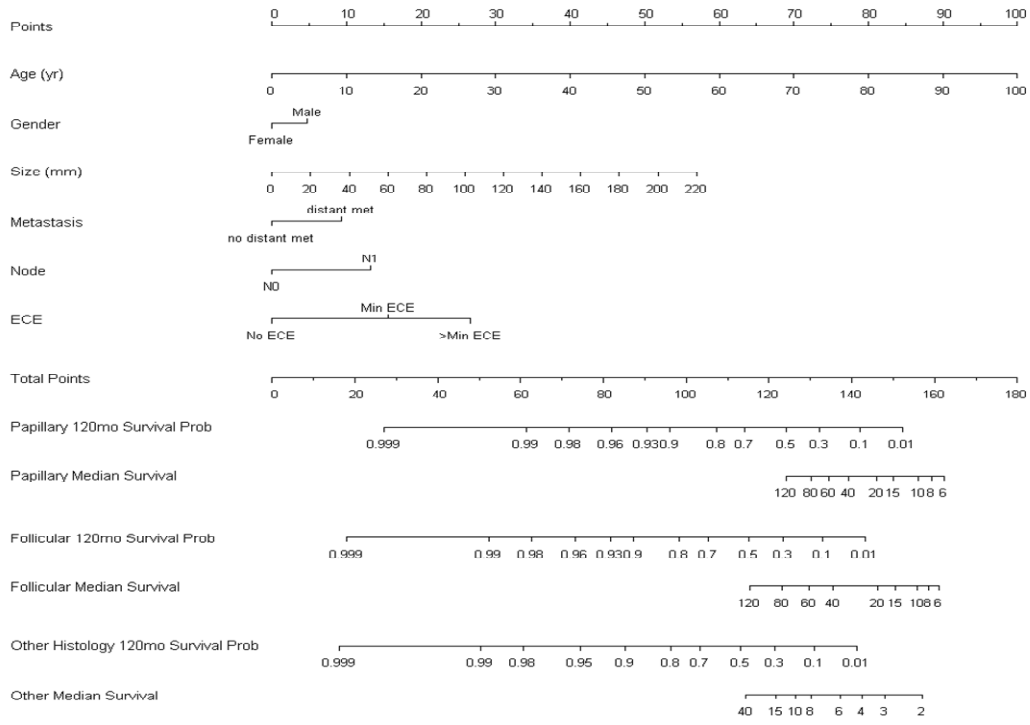
Part 6: Internal Validation

The model was validated by evaluating discrimination and calibration. Model discrimination was measured by the c-index, which is a measure of the concordance between observed and predicted survival. Perfect prediction is reflected by a c-index of 1.0, while a model with random prediction has a c-index of 0.5. Calibration displays agreement between predicted and Kaplan-Meier survival over fixed time intervals. This calculated difference in survival estimates is repeated on 200 bootstrapped samples to correct for over-optimism. A calibration plot was inspected to determine accuracy of the model. The model was found to have excellent discrimination, with a c-index of 0.925. There was evidence of overestimation of survival by the model when actual survival was approximately 75%. Although imperfect, the model otherwise shows generally good accuracy.

Part 7: Nomogram

The construction of a nomogram allows for the presentation of this predictive model in a concrete clinical tool. A nomogram is a graphical calculator that depicts a mathematical function. This study uses the regression coefficients obtained in the multivariable analysis to construct a nomogram that provides a visual representation of the relative size of the hazard associated with a predictor variable, as well as a means to

Figure 14: Nomogram Predictive of 10-Year Cause-Specific Mortality in Differentiated Thyroid Cancer



predict 10-year cause-specific mortality using individual characteristics (Figure 14, Appendix III). This is distinct from existing thyroid cancer classification systems because estimated risk of mortality is limited to large group estimates rather than individuals. For instance, a 14 year old female with a 2 cm papillary thyroid carcinoma without

extracapsular extension, lymph nodes, or metastasis is classified as T1N0M0 (stage I) by TNM staging. Based on TNM stage, her estimated 25-year mortality risk is approximately 1.7%. [10] Based on histology, her 10-year risk of mortality is estimated at 7%.[17] However, her very young age and lack of other risk factors results in an individualized probability of 10-year cause-specific mortality of <0.1% using the nomogram. In contrast, an 80 year old male with a 7cm insular carcinoma without distant metastasis, but with nodal involvement and greater than minimal extracapsular extension would have a T4N1M0 (stage IV) tumor. An estimate of 5-year mortality in poorly differentiated thyroid cancers with a TNM stage 3 or 4 is 52.6%, while estimates based on histology would suggest 33-66% 10-year mortality.[48, 19] However, based on the nomogram, this man has a > 99% probability of 10-year cause-specific mortality, with a median survival time of less than four months. These examples convey the predictive precision that is gained by using individualized risk factors to estimate risk of mortality in thyroid cancer.

Part 8: Limitations & Future Studies

Competing Risks

Amongst the 9,654 cases analyzed, 3.7% died of thyroid cancer, 84.3% were still alive at the end of the study period, and 12% had died of other causes. The largest proportion of other COD cases were due to cardiac disease (22.2%), followed by lung and bronchus disease (8.5%). The use of Cox PH assumes independence of competing risks, in this case, independence of death due to thyroid cancer, and death due to lung or cardiac disease. The estimate of cause-specific mortality in the context of competing risks

introduces the potential for a biased estimate of the effect of the covariates on survival time when using a Cox PH model. A cumulative incidence function is an alternative to the survival function used in Cox PH that does not assume independence. Formal tests are available to evaluate for independence of competing risks, which were not conducted in this study.

Generalizability

The selected model has very good predictive performance using a well represented cohort of the United States. However, generalizability of the nomogram may be a limitation. This nomogram predicts mortality presuming surgical excision, which is part of standard treatment. The effect on survival by other interventions, such as extent of surgery, radioactive iodine or external beam radiation therapy will not be included. Additionally, the degree of hormone replacement after treatment cannot be assessed, and its effect is unknown. External validation of the nomogram, which is addressed in the following section, will not be confirmed in our study.

Selection Bias

The excluded observations with unknown values did appear to have some difference in histology from the included cases. This difference may result in a model that overestimates survival in the Papillary and Follicular groups, while underestimating the survival in cases with Other histology. Model selection and validation using an imputed dataset showed no difference in predictive performance suggesting that the selection bias, if present, is non-differential. Additionally, the majority of patients who did not receive surgery were likely not surgical candidates, possibly as a result of higher cancer stage and higher morbidity. There are an additional 694 cases that fit the inclusion

criteria if surgical treatment is not considered. Treatment through surgical excision will decrease the risk of mortality. Inclusion of only those who received surgery may bias the results towards the null hypothesis.

Misclassification Bias

The use of county-level SES measures to predict individual mortality will also be a limitation. Median income of county of residence was used as a SES measure. Cancer mortality has been associated with economic deprivation at the census tract and block levels.[30] Median income was found to have a stronger association with cancer survival when compared with percent below poverty level at the census tract level.[49] Census SES measures are only available after 1990, which limits inclusion in this study. Previous studies using census-based methodology to geocode individual SES measures have been found to be applicable within 5 years of the closest census. The use of census-based methodology also has specific limitations. The census requires a physical address to be counted, tends to undercount both minorities and poor populations, and does not account for population fluctuations between the decennial data collection.[33]

Future research

In the future, assessment of the independence of competing risks should be conducted. A cumulative incidence model should be created to compare differences in covariate hazard ratios obtained with the competing risks method and the Cox PH method. Additionally, external validation using regional or international cancer populations should be conducted to assure model discrimination and calibration in alternate populations.

CHAPTER 5: CONCLUSIONS

Thyroid cancer accounts for 3.4% of the cancers in the United States.[2] Although 10-year survival of all thyroid cancer patients is greater than 95%, risk of mortality varies greatly depending on several factors. While several studies have estimated mortality risk due to thyroid cancer in large stratified groups, this risk estimate has not been applied to the individual. Nomograms are increasingly utilized and practical tools that allows for the prediction of risk. The strength of the nomogram is the ability to inform clinical decisions using individual information in real-time. It also provides a teaching tool that summarizes the amount of risk associated with each covariate of interest. Lastly, classification using the nomogram can provide the clinical researcher with a consistent means of risk stratification for use in clinical trials or other comparative studies.

The nomogram presented here shows very strong internal validity with excellent discrimination and good calibration. The adoption of this nomogram predictive of 10-year risk in the newly diagnosed patient with differentiated thyroid cancer patient is a concrete and practical tool for the clinician and patient.

REFERENCES

1. Horner MJ, Ries LAG, Krapcho M, Neyman N, Aminou R, Howlader N, Altekruse SF, Feuer EJ, Huang L, Mariotto A, Miller BA, Lewis DR, Eisner MP, Stinchcomb DG, Edwards BK (eds). SEER Cancer Statistics Review, 1975-2006, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2006/, based on November 2008 SEER data submission, posted to the SEER web site, 2009.
2. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov). Prevalence database: "US Estimated Complete Prevalence Counts on 1/1/2004". National Cancer Institute, DCCPS, Surveillance Research Program, Statistical Research and Applications Branch, released April 2007, based on the November 2006 SEER data submission. Includes any person alive on January 1, 2004 who had been diagnosed with cancer of the thyroid at any point prior to January 1, 2004, including persons with active and cured disease.
3. SEER*Stat Database: Incidence - SEER 9 Regs Limited-Use, Nov 2006 Sub (1973-2004), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2007, based on the November 2006 submission.
4. Davies L, Welch HG. Increasing Incidence of Thyroid Cancer in the United States, 1973-2002. JAMA. 2006 May 10;295(18):2164-7.
5. SEER*Stat Database: Mortality - All COD, Public-Use With State, Total U.S. (1969-2004), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2007. Underlying mortality data provided by NCHS (www.cdc.gov/nchs).
6. Hay ID, Grant CS, Taylor WF, McConahey WM. Ipsilateral lobectomy versus bilateral lobar resection in papillary thyroid carcinoma: a retrospective analysis of surgical outcome using a novel prognostic scoring system. Surgery. 1987 Dec;102(6):1088-95.
7. Cady B, Rossi R. An expanded view of risk-group definition in differentiated thyroid carcinoma. Surgery. 1988 Dec;104(6):947-53.
8. Hay ID, Bergstralh EJ, Goellner JR, Ebersold JR, Grant CS. Predicting outcome in papillary thyroid carcinoma: development of a reliable prognostic scoring system in a cohort of 1779 patients surgically treated at one institution during 1940 through 1989. Surgery. 1993 Dec;114(6):1050-7
9. Greene FL, Page DL, Balch CM, Fleming ID, Fritz A, Balch CM, Haller DG, Morrow M (Eds.): AJCC Cancer Staging Manual, 6th ed., Springer-Verlag, New York. 2002.
10. Loh KC, Greenspan FS, Gee L, Miller TR, Yeo PP. Pathological tumor-node-metastasis (pTNM) staging for papillary and follicular thyroid carcinomas: a retrospective analysis of 700 patients. J Clin Endocrinol Metab. 1997 Nov;82(11):3553-62.
11. D'Avanzo A, Ituarte P, Treseler P, Kebebew E, Wu J, Wong M, Duh QY, Siperstein AE, Clark OH. Prognostic scoring systems in patients with follicular thyroid cancer: a comparison of different staging systems in predicting the patient outcome. Thyroid. 2004 Jun;14(6):453-8.

12. Kingma G, van den Bergen HA, de Vries JE. Prognostic scoring systems in differentiated thyroid carcinoma: which is the best? Neth J Surg. 1991 Jun;43(3):63-6.
13. Gulcelik MA, Gulcelik NE, Kuru B, Camlibel M, Alagol H. Prognostic factors determining survival in differentiated thyroid cancer. J Surg Oncol. 2007 Dec 1;96(7):598-604.
14. Morton RP, Ahmad Z. Thyroid cancer invasion of neck structures: epidemiology, evaluation, staging and management. Curr Opin Otolaryngol Head Neck Surg. 2007 Apr;15(2):89-94
15. Shaha AR, Shah JP, Loree TR. Low-risk differentiated thyroid cancer: The need for selective treatment. Ann Surg Oncol 1997;4:328-33.
16. Lundgren CI, Hall P, Dickman PW, Zedenius J. Clinically significant prognostic factors for differentiated thyroid carcinoma: a population-based, nested case-control study. Cancer. 2006 Feb 1;106(3):524-31.
17. Hundahl SA, Fleming ID, Fremgen AM, Menck HR. A National Cancer Data Base report on 53,856 cases of thyroid carcinoma treated in the U.S., 1985-1995. Cancer. 1998 Dec 15;83(12):2638-48
18. Patel, KN; Shaha, AR. Poorly Differentiated and Anaplastic Thyroid Cancer. Cancer Control. 2006 Apr;13(2):119-128
19. Sanders EM Jr, LiVolsi VA, Brierley J, Shin J, Randolph GW. An evidence-based review of poorly differentiated thyroid cancer. World J Surg. 2007 May;31(5):934-45
20. Gilliland FD, Hunt WC, Morris DM, Key CR. Prognostic factors for thyroid carcinoma. A population-based study of 15,698 cases from the Surveillance, Epidemiology and End Results (SEER) program 1973-1991. Cancer. 1997 Feb 1;79(3):564-73.
21. Kowalski LP, Gonçalves Filho J, Pinto CA, Carvalho AL, de Camargo B. Long-term survival rates in young patients with thyroid carcinoma. Arch Otolaryngol Head Neck Surg. 2003 Jul;129(7):746-9
22. Palme CE, Waseem Z, Raza SN, Eski S, Walfish P, Freeman JL. Management and outcome of recurrent well-differentiated thyroid carcinoma. Arch Otolaryngol Head Neck Surg. 2004 Jul;130(7):819-24.
23. Witt RL, McNamara AM. Prognostic factors in mortality and morbidity in patients with differentiated thyroid cancer. Ear Nose Throat J. 2002 Dec;81(12):856-63.
24. Machens A, Hauptmann S, Dralle H. Disparities between male and female patients with thyroid cancers: sex difference or gender divide? Clin Endocrinol. 2006 Oct;65(4):500-5.
25. Cunningham MP, Duda RB, Recant W, Chmiel JS, Sylvester JA, Fremgen A. Survival discriminants for differentiated thyroid cancer. Am J Surg. 1990 Oct;160(4):344-7.
26. Spitz MR, Sider JG, Katz RL, Pollack ES, Newell GR. Ethnic patterns of thyroid cancer incidence in the United States, 1973-1981. Int J Cancer. 1988 Oct 15;42(4):549-53
27. Mulla ZD, Margo CE. Primary malignancies of the thyroid: epidemiologic analysis of the Florida Cancer Data System registry. Ann Epidemiol. 2000 Jan;10(1):24-30.

28. Haselkorn T, Stewart SL, Horn-Ross PL. Why are thyroid cancer rates so high in southeast asian women living in the United States? The bay area thyroid cancer study. Cancer Epidemiol Biomarkers Prev. 2003 Feb;12(2):144-50.
29. Kogevinas M, Pearce N, Susser M, Boffetta P, eds. Social Inequalities and Cancer. Lyons, France: International Agency for Research on Cancer/World Health Organization; 1997. IARC scientific publication 138.
30. Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter?: the Public Health Disparities Geocoding Project. Am J Epidemiol. 2002 Sep 1;156(5):471-82
31. Boyd C, Zhang-Salomons JY, Groome PA, Mackillop WJ. Associations between community income and cancer survival in Ontario, Canada, and the United States. J Clin Oncol. 1999 Jul;17(7):2244-55.
32. Krieger N, Quesenberry C Jr, Peng T, Horn-Ross P, Stewart S, Brown S, Swallen K, Guillermo T, Suh D, Alvarez-Martinez L, Ward F. Social class, race/ethnicity, and incidence of breast, cervix, colon, lung, and prostate cancer among Asian, Black, Hispanic, and White residents of the San Francisco Bay Area, 1988-92 (United States). Cancer Causes Control. 1999 Dec;10(6):525-37.
33. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. Am. J. Public Health 1992 May;82:703–10
34. Krieger N, Williams D, Moss N. Measuring social class in US public health research: concepts, methodologies and guidelines. Annu Rev Public Health. 1997;18: 341-378.
35. Krupski TL, Kwan L, Afifi AA, Litwin MS. Geographic and socioeconomic variation in the treatment of prostate cancer. J Clin Oncol. 2005 Nov 1;23(31):7881-8
36. Singh GK, Miller BA, Hankey BF, Edwards BK. Persistent area socioeconomic disparities in U.S. incidence of cervical cancer, mortality, stage, and survival, 1975-2000. Cancer. 2004 Sep 1;101(5):1051-7.
37. Ghori FY, Gutterman-Litofsky DR, Jamal A, Yeung SC, Arem R, Sherman SI. Socioeconomic factors and the presentation, management, and outcome of patients with differentiated thyroid carcinoma. Thyroid. 2002 Nov;12(11):1009-16.
38. Memorial Sloan-Kettering Cancer Center, Prediction Tools. Created 2001 Dec 11. Accessed 11/19/2007 from www.nomograms.org
39. Karakiewicz PI, Benayoun S, Kattan MW, Perrotte P, Valiquette L, Scardino PT, Cagiannos I, Heinzer H, Tanguay S, Aprikian AG, Huland H, Graefen M. Development and validation of a nomogram predicting the outcome of prostate biopsy based on patient age, digital rectal examination and serum prostate specific antigen. J Urol. 2005 Jun;173(6):1930-4.
40. Gross ND, Patel SG, Carvalho AL, Chu PY, Kowalski LP, Boyle JO, Shah JP, Kattan MW. Nomogram for Deciding Adjuvant Treatment after Surgery for Oral Cavity Squamous Cell Carcinoma. Head & Neck. 2008 Aug 21;30(10):1352-1360
41. SEER*DMS Users Manual, version 4. 2007 Aug. Accessed 2007 Nov 14 from http://seer.cancer.gov/seerdms/manual/

42. SEER Number of Persons by Race and Hispanic Ethnicity for SEER Participants (2000 Census Data), modified 2006 Apr 6. Accessed 2007 Nov 19 from http://seer.cancer.gov/registries/data.html
43. Thacker SB. HIPAA Privacy Rule and Public Health Guidance from CDC and the U.S. Department of Health and Human Services. MMWR. 2003 2 May;52(S-1):1-12.
44. SEER Training Module, modified 2007 Nov 18. Accessed 2007 Nov 18 2007 from http://training.seer.cancer.gov/module_icdo3/icd_o_3_lists.html
45. SEER ICD-O-3 Coding Materials, modified 2007 Aug 15. Accessed 2007 Nov 18 2007 from http://seer.cancer.gov/icd-o-3/
46. SEER Data Items, created 2007 Oct 4. Accessed 2007 Nov 19 from http://seer.cancer.gov/tools/seer07.dataitems.pdf
47. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996 Feb 28;15(4):361-87.
48. Pulcrano M, Boukheris H, Talbot M, Caillou B, Dupuy C, Virion A, De Vathaire F, Schlumberger M. Poorly differentiated follicular thyroid carcinoma: prognostic factors and relevance of histological classification. Thyroid. 2007 Jul;17(7):639-46.
49. Zhang-Salomons J, Qian H, Holowaty E, Mackillop WJ. Associations between socioeconomic status and cancer survival: choice of SES indicator may affect results. Ann Epidemiol. 2006 Jul;16(7):521-8

Appendix I: TNM Staging

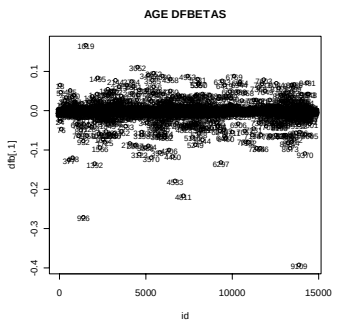
Tumor (T), Lymph Node (N), and Distant Metastasis (M) Classification and Staging of Thyroid Cancer

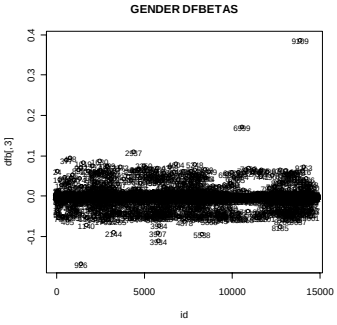
Greene FL et al (editors): AJCC Cancer Staging Manual, 6th ed., Springer, 2002.

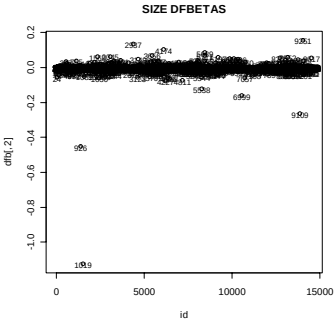
TNM CLASSIFICATION	
Primary Tumor (T)	
Note: All categories may be subdivided: (a) solitary tumor, (b) multifocal tumor (the largest determines the classification).	
TX	Primary tumor cannot be assessed
T0	No evidence of primary tumor
T1	Tumor 2 cm or less in greatest dimension limited to the thyroid
T2	Tumor more than 2 cm but not more than 4 cm in greatest dimension limited to the thyroid
T3	Tumor more than 4 cm in greatest dimension limited to the thyroid or any tumor with minimal extra-thyroid extension (eg, extension to sternothyroid muscle or perithyroid soft tissues)
T4a	Tumor of any size extending beyond the thyroid capsule to invade subcutaneous soft tissues, larynx, trachea, esophagus, or recurrent laryngeal nerve
T4b	Tumor invades prevertebral fascia or encases carotid artery or mediastinal vessels
Anaplastic Carcinomas	
T4a	Intrathyroidal anaplastic carcinoma-surgically resectable
T4b	Extrathyroidal anaplastic carcinoma-surgically unresectable
Regional Lymph Nodes (N)	
Regional lymph nodes are the central compartment, lateral cervical, and upper mediastinal lymph nodes	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Regional lymph node metastasis
N1a	Metastasis to Level VI (pretracheal, paratracheal, and prelaryngeal/Delphian lymph nodes)
N1b	Metastasis to unilateral, bilateral, or contralateral cervical or superior mediastinal lymph nodes
Distant Metastasis (M)	
MX	Distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis

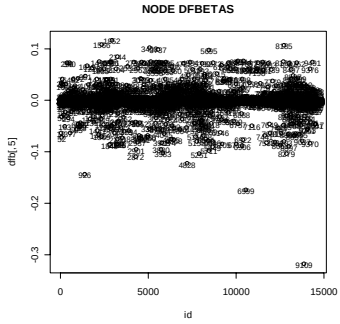
TNM STAGE			
Separate stage groupings are recommended for papillary or follicular, medullary, and anaplastic (undifferentiated) carcinoma.			
Papillary or Follicular Under 45 Years			
Stage I	Any T	Any N	M0
Stage II	Any T	Any N	M1
Papillary or Follicular 45 Years and Older			
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
	T1	N1a	M0
	T2	N1a	M0
	T3	N1a	M0
Stage IVA	T4a	N0	M0
	T4a	N1a	M0
	T1	N1b	M0
	T2	N1b	M0
	T3	N1b	M0
	T4a	N1b	M0
Stage IVB	T4b	Any N	M0
Stage IVC	Any T	Any N	M1
Medullary Carcinoma			
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
	T1	N1a	M0
	T2	N1a	M0
	T3	N1a	M0
Stage IVA	T4a	N0	M0
	T4a	N1a	M0
	T1	N1b	M0
	T2	N1b	M0
	T3	N1b	M0
	T4a	N1b	M0
Stage IVB	T4b	Any N	M0
Stage IVC	Any T	Any N	M1
Anaplastic Carcinoma			
All anaplastic carcinomas are considered Stage IV			
Stage IVA	T4a	Any N	M0
Stage IVB	T4b	Any N	M0
Stage IVC	Any T	Any N	M1

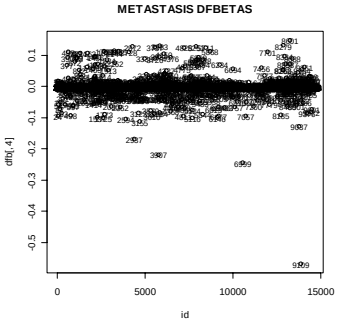
Appendix II: Scatterplot of case id by DFBETAS for each predictor variable

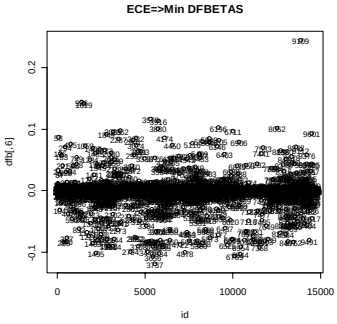


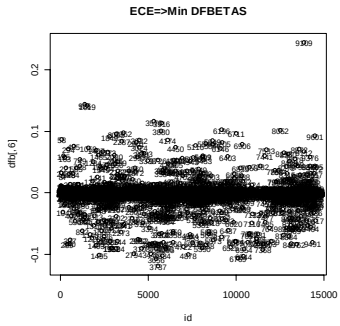












Appendix III: Nomogram Predictive of 10-Year Cause-Specific Mortality in Differentiated Thyroid Cancer