Revisiting Overdiagnosis and Fatality in Thyroid Cancer

Robin T. Vollmer, MD

From VA and Duke University Medical Centers, Durham, NC.

Key Words: Overdiagnosis; Survival; Fatality of thyroid cancer; Hazard

ABSTRACT

Objectives: To examine the rates of incidence and fatality in cohorts of patients diagnosed with thyroid cancer from 1975 to 1999.

Methods: This study uses National Cancer Institute’s Surveillance, Epidemiology and End Results data and derives hazard functions in order to examine the fatality in thyroid cancer.

Results: The study documents forms of rapidly evolving and fatal tumors as well as forms of tumor that evolve more slowly to cause death. It demonstrates that the incidences of nonfatal forms of thyroid cancer have risen dramatically in the years from 1975 to 1999—mostly due to papillary carcinomas—but that the incidences of fatal forms of thyroid cancer have remained nearly constant.

Conclusions: The results of this study support the notion that many thyroid cancers are part of a reservoir of nonfatal tumors that are increasingly being overdetected and overdiagnosed.

Prognosis and survival for most carcinomas of the thyroid has long been known to be excellent. For example, data collected by the American Joint Committee on Cancer (AJCC) demonstrate that 98% of papillary carcinomas are in stages 1 to 3 and have a weighted mean 5-year relative survival of 99%. AJCC data for follicular carcinomas demonstrate that 86% are in stage 1 or 2 and have a weighted mean 5-year relative survival of 99%. Although there is general agreement about these good outcomes, agreement disappears when it comes to the defining histologic features of well-differentiated thyroid cancers, especially papillary carcinoma. and such disagreements have been repeatedly documented as significant interobserver differences in diagnosis. In this setting, some pathologists have lowered their thresholds for diagnosing papillary carcinomas. Others have suggested that diagnostic categories be expanded to include

 Upon completion of this activity you will be able to:

• describe the shape of a typical hazard function for thyroid cancer and relate this to the timing of fatalities.
• describe how incidences of fatal and nonfatal forms of thyroid cancer have changed in the past decades.
• describe how changes in incidence of fatal and nonfatal forms of thyroid cancer might relate to the phenomenon of overdiagnosis.

The ASCP is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians. The ASCP designates this journal-based CME activity for a maximum of 1 AMA PRA Category 1 Credit™ per article. Physicians should claim only the credit commensurate with the extent of their participation in the activity. This activity qualifies as an American Board of Pathology Maintenance of Certification Part II Self-Assessment Module.

The authors of this article and the planning committee members and staff have no relevant financial relationships with commercial interests to disclose. Questions appear on p 143. Exam is located at www.ascp.org/ajcpCME.
tumors with uncertain outcomes or some assessment of risk. Still others have advocated for prospective long-term studies. The problem, however, with doing long-term studies of outcome for tumors with such good outcomes is lack of statistical power. For example, to document a statistically significant difference in survival between follicular adenoma and the encapsulated, noninvasive follicular variant of papillary carcinoma, one would need at least 120,000 patients and at least 1,200,000 patient-years if the study were to include just 10 years of follow-up. Because such numbers are unlikely to be collected, another approach to studying these issues is to use prospectively collected data in the public sector.

For example, Davies and Welch used incidence data collected from the National Cancer Institute’s Surveillance, Epidemiology and End Results (SEER) program and mortality data from the Centers for Disease Control and Prevention’s National Vital Statistics System to document increasing disparity between incidence and mortality and concluded that we are overdiagnosing thyroid cancer, particularly papillary carcinoma. Because Davies and Welch’s incidence data came from one cohort and the mortality data came from another, and because these two cohorts were diagnosed at different times, some uncertainty remains regarding the effects of lead times and changes in pathologic diagnostic criteria. Herein, I use SEER data to derive incidence and fatality results from the same cohorts with thyroid carcinoma, that is, those diagnosed during the same year (or short time interval) and the same geographic regions.

Materials and Methods

Study Data

The main data for this study comprise incident rates and relative survivals. Relative survival probabilities are relative to the general population. Both were obtained for cohorts matched by year of diagnosis for the following years: 1975 to 1979 (median at 1977), 1980 to 1984 (median at 1982), 1985 to 1989 (median at 1987), 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, and 1999. The data were obtained from the SEER database, available at http://www.seer.cancer.gov/csr/1975_2009 (last accessed on April 12, 2013), and the follow-up time for these cohorts was such that there were 10 to 20 years of documented survival probabilities for each year of diagnosis. Data regarding the incidence of papillary carcinomas come from Davies and Welch.

Estimating Overall Survival and Fatality

If survival probability at any time \( t \) is symbolized as \( S(t) \), then \( S(t) \) gives the fraction of nonfatal cases by time \( t \), and the corresponding fraction of fatal cases is \( 1 - S(t) \). The overall nonfatal fraction is \( S(\infty) \), and the overall fatal fraction is \( 1 - S(\infty) \). Values for \( S(t) \) and \( 1 - S(t) \) were obtained from the SEER data. Then, exploiting the calculus of survival given by

\[
S(t) = \exp\left(-\int_0^t h \, dx\right)
\]

a hazard function was derived for each diagnostic cohort. Here, \( \exp \) stands for exponentiation, \( h \) symbolizes the hazard function, and the integration limits for \( x \) are from 0 to \( t \). Hazard functions were modeled as the sum of two \( \gamma \) functions as before:

\[
h = \alpha \, t \, \exp(-\beta \, t) + \lambda \, t^2 \, \exp(-\delta \, t)
\]

and the coefficients were derived via a least squares fitting algorithm so that right-hand side of equation (1) fit the observed values of \( S(t) \) in the SEER data. Finally, values of \( S(\infty) \) and \( 1 - S(\infty) \) were obtained by integrating the hazard function from 0 to the limit of \( t \to \infty \).

Results

Derived Hazard Functions for Thyroid Cancer Survival Curves

The derived hazard functions provided excellent fits to the original SEER survival curves, and this is shown in Figure 1, which illustrates a plot of the observed values of \( S(t) \) on the x-axis vs values of \( S(t) \) derived from the hazard function.
functions on the y-axis. The line shows where perfect fits should occur, and the fact that the points are consistently close to the line implies that the derived hazard functions accurately reflect observed survival probabilities. Altogether, the mean difference between observed and fitted survival probabilities was $4.5 \times 10^{-6}$ (units of probability), with a median difference of 0 and a range of –.005 to .005.

Figure 2 shows an example of a typical derived hazard function: the one for those diagnosed in 1990. The hazard function for this cohort rose rapidly to peak at approximately 4 months after diagnosis and then rapidly fell. A smaller rise occurred at 10 years, but eventually the hazard function fell to nearly zero. Among all the diagnostic cohorts, the mean time for peak hazard was 4.1 months, implying that an average of 4.4% died by two years. Furthermore, over all the cohorts, 56% of fatalities occurred by two years. Thus, the forms of these hazard functions imply that more than half of fatal thyroid cancers are characterized by rapidly evolving tumor kinetics—undoubtedly due to undifferentiated, medullary, and squamous tumor types or a few follicular and papillary tumors of higher initial stages.

For those who survive two years, the hazard function drops dramatically, and by 10 years, its value averages just 5% of its peak. The drop implies that approximately 80% of tumor-related fatalities have occurred by 10 years. After 10 years, a further drop in the hazard implies that subsequent survival approximates that of the general population.

Delayed Fatality in Thyroid Cancers

The above results suggest that a small fraction of patients have rapidly evolving and fatal forms of thyroid cancer that cause death within two years of diagnosis. What happens to the larger group of patients who survive two years? One can address this by using the SEER data to calculate the survival probability $S(t > 2)$, which gives the probability of long-term survival, given survival past two years. Mathematically, $S(t > 2)$ is computed from the hazard functions as follows:

$$S(t > 2) = \exp(-\int_{t}^{\infty} h \, dx) = \exp(-[\int_{0}^{\infty} h \, dx - \int_{0}^{2} h \, dx]) = \exp(-\int_{0}^{\infty} h \, dx) / \exp(-\int_{0}^{2} h \, dx) = S(\infty) / S(2)$$

For the SEER data, the mean long-term survival probability, given survival of at least two years, was 0.96 (range, .94-.98). Thus, after two years, approximately 4% will have a delayed death due to a more slowly evolving form of cancer. The remaining 96% should expect subsequent survival to match those without thyroid cancer.

Incidence of Fatal and Nonfatal Thyroid Cancers

Figure 3 shows plots of the incidences of fatal (lower line) and nonfatal (upper line) carcinoma (units of cases per 100,000) vs the year of diagnosis. From 1977 to 1999, the incidence of fatal cases remained nearly constant at a mean of 0.47 per 100,000 ($P > .3$ for linear regression analysis with time); however, the incidence of nonfatal cases rose from 4.8 to 6.7 per 100,000 ($P = .0003$ for linear regression analysis with time). Figure 4 demonstrates that the rise of nonfatal carcinomas was directly and linearly related to the incidence of papillary carcinomas. The points on the plot are the observed values, and the line is a fit by linear regression analysis, which explained 97% of the variance in the data and yielded a $P$ value of approximately 0. The slope of the line
was 0.91, suggesting that on average, 91% of papillary carcinomas were nonfatal. Furthermore, the fact that this value remained constant over the diagnostic years from 1977 to 1999 suggests that the criteria for the diagnosis of papillary carcinomas—including the follicular variant—did not change in a way that affected outcomes.

Discussion

Using some of the same SEER data, the results of this study validate and extend those obtained by Davies and Welch. Namely, both studies document a steady increase in the diagnosis of nonfatal forms of thyroid cancer in the past two decades. This study differs from Davies and Welch’s study by deriving incidence and fatality from the same cohorts. It does this by using the hazard functions of each diagnostic cohort, and it relies on patients with at least 10 years of follow-up. Regarding nonfatal cancers, lead time bias seems unlikely to explain their rising incidence, because fatality was derived from the same cohorts as incidence. Furthermore, this study demonstrates that the relationship between the diagnosis of papillary forms of thyroid cancer and outcomes has not changed over two decades, implying that more recent diagnoses of papillary forms of thyroid cancer are not diluted by other lesions. Nevertheless, it is clear that more tumors are being detected, undoubtedly due to increases in ultrasound-guided fine-needle aspirate biopsies, which now represent an increasing part of most pathologists’ practices. Ultrasound analyses are uncovering smaller and smaller lesions, which we are increasingly recognizing as variants of papillary cancers. Is this “overdiagnosis”? If one examines the fatality of the thyroid carcinomas that comprise this “epidemic,” then both Davies and Welch’s results and the results of this study suggest that the answer is yes.

Others have opined about this or related issues. For example, speaking of follicular carcinomas, DeMay21 wrote, “Most are minimally invasive (and minimally malignant) tumors.” Considering the encapsulated follicular variant of papillary thyroid carcinoma (PTC), Chan7 suggested that “when there are uncertainties in the diagnosis, it would not be a disservice to the patient if a genuine follicular variant of PTC were misdiagnosed as follicular adenoma, because simple excision of the lesion is already curative.” Addressing small asymptomatic and dormant thyroid cancers, Mazzaferri22 suggested that they be “best left undiagnosed.” Speaking of encapsulated, noninvasive follicular patterned tumors that contain some papillary-type nuclei, Rosai12 questioned, “Why call them cancer?” Addressing encapsulated well-differentiated follicular patterned thyroid carcinomas, Piana et al13 questioned the “wisdom of calling them carcinomas.” Finally, referring to Welch and Black’s article on overdiagnosis in cancer,16 Esserman and Thompson23 issue a call to “acknowledge the spectrum of cancer behavior and the need to reclassify ‘indolent’ lesions with a term other than ‘cancer.’” Given the above results and opinions, as well as variances in attitudes between pathologists about the diagnoses of papillary carcinomas, it may be time to assemble a consensus conference regarding diagnosis, terminology, and outcomes for thyroid carcinomas.

Despite results documenting probable overdiagnosis of thyroid carcinomas, this study also demonstrates the continued presence of rapidly evolving and fatal forms of thyroid cancer—ones characterized by an early peak in the hazard function and by a high probability for death within two years of diagnosis—as well as fewer more slowly evolving forms of fatal thyroid carcinomas. The constancy of incidence found for fatal thyroid cancers suggests that these tumors may be independent of endemic goiter and that they evolve too quickly to be susceptible to newer diagnostic procedures. Finally, this study and several others15-17 demonstrate once again how important SEER data are for the study of outcomes in thyroid cancer, for which it is difficult and expensive, if not impossible, to collect sufficient numbers of uncensored patients for prospective studies. The SEER data’s derivation of and reliance on relative survivals compensates for limited data on outcomes. Furthermore, reliance on relative survivals directly addresses patients’ concerns about how they will fare relative to those without their tumors. Thus, continued funding and support of the SEER project are important.

Address reprint requests to Dr Vollmer: Laboratory Medicine 113, VA Medical Center, 508 Fulton St, Durham, NC 27705; e-mail Robin.Vollmer@med.va.gov.
Vollmer / Overdiagnosis and Fatality in Thyroid Cancer

References


© American Society for Clinical Pathology