Published in final edited form as:

Med Care Res Rev. 2011 December ; 68(6): 627–649. doi:10.1177/1077558711408324.

STRUCTURAL ESTIMATES OF TREATMENT EFFECTS ON OUTCOMES USING RETROSPECTIVE DATA: AN APPLICATION TO DUCTAL CARCINOMA IN SITU

Heather Taffet Gold, PhD,

Weill Cornell Medical College, New York, NY

Melony E. S. Sorbero, PhD, MS,

RAND Corporation, Pittsburgh, PA

Jennifer J. Griggs, MD, MPH,

University of Michigan School of Medicine, Ann Arbor, MI

Huong T. Do, MA, and

Hospital for Special Surgery, New York, NY

Andrew W. Dick, PhD

RAND Corporation, Pittsburgh, PA

Abstract

Analysis of observational cohort data is subject to bias from unobservable risk selection. We compared econometric models and treatment effectiveness estimates using the linked Surveillance, Epidemiology, and End Results (SEER)-Medicare claims data for women diagnosed with ductal carcinoma in situ. Treatment effectiveness estimates for mastectomy and breast conserving surgery (BCS) with or without radiotherapy were compared using three different models: simultaneous-equations model, discrete-time survival model with unobserved heterogeneity (frailty), and proportional hazards model. Overall trends in disease-free survival (DFS), or time to first subsequent breast event, by treatment are similar regardless of the model, with mastectomy yielding the highest DFS over 8 years of follow-up, followed by BCS with radiotherapy, and then BCS alone. Absolute rates and direction of bias varied substantially by treatment strategy. DFS was underestimated by single-equation and frailty models compared to the simultaneous-equations model and RCT results for BCS with RT and overestimated for BCS alone.

Corresponding author: Heather Taffet Gold, PhD, Department of Public Health, Weill Cornell Medical College, 402 E. 67th Street, New York, NY 10065; work phone: 646-962-8040; fax: 646-962-0281; home phone: 914-630-1237, heg2001@med.cornell.edu. Other contact information: Melony ES Sorbero, RAND, 201 North Craig Street, Pittsburgh, PA 15213, phone: 412-683-2300 ext. 4628, msorbero@rand.org

Jennifer J Griggs, North Ingalls Building, 300 North Ingalls, Room NI3A22, Ann Arbor, MI 48109-5419, phone: (734) 647-9210, jengrigg@umich.edu

Huong T Do, Hospital for Special Surgery, 535 E. 70th St., Belaire 9F, New York, NY 10021, phone: 646-797-8208, huong.do@alumni.brown.edu

Andrew W Dick, RAND, 201 North Craig Street, Pittsburgh, PA 15213, phone: 412-683-2300 ext. 4474, andrew_dick@rand.org

The authors have no potential conflicts of interest to disclose.

¹Distance to radiation facility, a variable used by others (Hadley et al., 2003), may have been a strong instrument, but we did not have access to this information.

Keywords

endogeneity bias; econometrics; retrospective cohort study; ductal carcinoma in situ

INTRODUCTION

Many studies of treatment effectiveness rely on randomized controlled trials (RCTs) and biostatistical analysis of patient cohorts. Both of these methods have limitations. The strict selection of patients into RCTs reduces the generalizability of their results to patients obtaining care in the community. In traditional biostatistical or epidemiologic analysis of observational cohorts the results may be influenced by selection (endogeneity) bias. In addition, all possible treatments may never have been compared directly in RCTs. Finally, treatment choice outside of RCTs often is guided by factors unrelated to clinical trial results, such as community norms, diffusion of scientific information, population density, physician supply, and patient characteristics such as race, age, and educational attainment (Gold & Dick, 2004; Sakorafas & Farley, 2003; Ernster, Barclay, Kerlikowske, Grady, & Henderson, 1996; Baxter, Virnig, Durham, & Tuttle, 2004).

In the case of breast cancer and ductal carcinoma in situ (DCIS), there are three principal treatment options, mastectomy, breast-conserving surgery (BCS) with radiotherapy, and BCS alone, and together they have not been compared head-to-head in a RCT. Published DCIS treatment trials show that BCS with radiotherapy is more efficacious than BCS alone (Fisher et al., 2001; Fisher et al., 1993; Holmberg et al., 2008; Bijker et al., 2006; Julien et al., 2000), but no randomized trials compare mastectomy to both BCS strategies (Fonseca et al., 1997; National Cancer Institute, 2006), nor do they attempt to compare the effectiveness of treatment as delivered in the community. Cohort studies, such as the one presented here, evaluate treatment effectiveness in a broader patient population than RCTs, but analyses of retrospective cohorts using population-based tumor registry data may be compromised by potential selection bias (Ernster, Barclay, Kerlikowske, Wilkie, & Ballard-Barbash, 2000; Paszat et al., 1998; Hershman et al., 2006; Smith et al., 2006; Solin et al., 1996; Cutuli et al., 2001; Warren et al., 2005; Habel et al., 1998; Wamberg, Yuen, & Holmberg, 2000). If key information that influences treatment choice, such as disease or patient characteristics (including patient preferences), is missing, the comparison groups may not be equivalent. If the groups are not equivalent on unmeasured factors, the observed relationships between treatment and outcomes may not be causal. Such selection bias may be a particular problem for studies of DCIS because treatment often is guided by pathology and histology data, which are not all available in large, secondary data sources such as the Surveillance, Epidemiology and End Results (SEER) data. For example, SEER does not indicate if a patient has clear margins. In the case of DCIS, patients with more extensive or aggressive disease may undergo more intensive treatment, resulting in an increased likelihood of choosing mastectomy over BCS. Alternatively, if BCS is chosen for more aggressive disease, radiotherapy then may be preferred to no radiotherapy. Thus, estimates of the relative benefits of mastectomy and radiotherapy will be understated due to unmeasured tumor characteristics.

A previous analysis of a cohort of Medicare beneficiaries' claims data (with verification of cancer diagnosis from physicians) estimated three-year overall survival after treatment for early stage (localized) breast cancer (not DCIS) by using an instrumental variables, two-stage least squares approach run separately for each treatment option (Hadley et al., 2003). In that model, the first equation for treatment choice depended linearly on measured health and demographic factors and variables unrelated to a patient's underlying health status. The authors used instruments including Medicare's average fees for breast surgery and geographic adjustment factor and patient distance to a hospital with a radiation facility. The second equation for three-year survival depended linearly on measured health and demographic factors. The authors found that women undergoing BCS had the poorest overall 3-year survival (79.2%), followed by mastectomy (85.3%) and BCS with radiotherapy (93%). The analyses attempted to overcome endogeneity bias but were limited by the short time horizon and linear-model methodology.

New Contribution

This study explicitly confronts the problem of endogenous treatment choice by estimating two alternative econometric models designed to address this issue. The first type, referred to as a frailty model, was originally proposed by Heckman and Singer (Heckman & Singer, 1984) to account for unobserved heterogeneity across subjects, for example, unobserved differential health status. The second model is an adaptation of a simultaneous-equations approach to a discrete-time duration model. The purpose is to compare the models to each other, to a naïve survival model, and to RCT results that compare two of the three treatment arms in our study (BCS with and without radiotherapy). From these comparisons, we will identify the magnitude of selection bias in the absence of RCTs, determine the extent to which the econometric methods successfully correct the estimates, and better understand the effectiveness of treatments in the imperfect real-world setting.

CONCEPTUAL FRAMEWORK

The conceptual framework for this study links outcomes of care to the process and structure of health care, characteristics of individuals involved in care, and factors related to the health system (Andersen, 1995). Outcomes can be directly or indirectly linked to these factors. For example, the structure and process of care include availability of physicians in a community or treatment patterns in a geographic area that directly could affect receipt of particular treatments. Individual and community characteristics, such as age, ruralness, or area-based poverty may influence treatment choice and the primary outcome of interest in this study, disease-free survival. However, not all factors that influence treatment choice are bound to affect disease-free survival. Therefore, we developed analytic models that take advantage of this fact and disentangle the effects of factors that only would affect treatment from those that might directly and indirectly affect outcomes. This framework guides the analytic models.

DATA

We used the linked SEER-Medicare database for women age 66 or older with unilateral, surgically-treated DCIS diagnosed from 1991 to 1999 in the 9 SEER-registry database and

their Medicare inpatient, outpatient, and physician claims from 1991 to 2001. The nine registries represented approximately 14% of the US population (Ries et al., 2000). In the data linkage, 93% of persons age 65 or older in the SEER database were matched to Medicare claims (Fahey, 2000). We excluded subjects if the reason for Medicare eligibility was due to end-stage renal disease status, DCIS was diagnosed at autopsy, they were not eligible for both Parts A and B of Medicare, or they were enrolled in an HMO for 6 months prior to or 12 months after their month of diagnosis. Subjects without Medicare Parts A and B or who were members of an HMO would not have complete administrative claims available for analysis. Our final analytic sample included 3,888 subjects out of 5,843 total subjects who were potentially eligible for inclusion.

Only month and year of DCIS diagnosis are available from the SEER data, thus we assumed the diagnosis was made on the first day of the month of diagnosis. Claims data provided the treatment information for our study, except when treatment claims were missing. The most extensive surgery performed, i.e., BCS or mastectomy, in the six months post-diagnosis was identified in the claims. Subjects who underwent BCS also may have had radiotherapy, so we searched 6 months post-surgery for radiotherapy claims. [Subjects undergoing mastectomy with radiotherapy (n=44) were excluded.] In those cases, SEER information was used to supplement the claims data. Twenty-nine percent (n=1,136) of subjects had both SEER and Medicare claims indicating radiotherapy, 3% (n=126) had only Medicare claims indicating radiotherapy, and 0.5% (n=19) had only SEER data indicating radiotherapy. The only histopathological information available was whether a subject had comedo-type necrosis, which tends to have worse prognosis than other forms of DCIS (Silverstein et al., 1995).

The primary outcome was disease-free survival, which ended with a subsequent breast event, defined as recurrence of DCIS or development of any invasive breast cancer. We chose this outcome because we were interested in a subject's next breast cancer event, regardless of type. The SEER data cannot be used to study the key outcome of recurrence, because SEER captures only primary cancers. We therefore used the Medicare claims data to define subsequent breast events from the claims data as adapted from a published algorithm (Warren, Riley, McBean, & Hakim, 1996; Gold, Do, & Dick, 2008). Subjects were considered to have a subsequent breast event if they had a Medicare claim at least 183 days (6 months) after diagnosis based on the first instance of either 1) an inpatient claim with a primary or secondary diagnosis of DCIS, axillary node involvement, or breast cancer, with a procedure of BCS, mastectomy, or axillary node dissection; 2) an outpatient claim with a primary diagnosis of DCIS, axillary node involvement, or breast cancer, with a procedure of BCS, mastectomy, or axillary node dissection; or 3) a physician-surgeon claim for BCS, mastectomy, or axillary node dissection. If no claims met these criteria, but a subject's cause of death was breast cancer, the subject was considered to have had a subsequent breast event at the date of death. The follow-up period started six months after diagnosis and was categorized into discrete annual increments. Because patients with DCIS should not receive chemotherapy (National Institutes of Health, 2009), only 6 months was allowed for treatment before starting the follow-up period. This lag period was varied from 6 to 12 months and showed virtually no difference in rates of subsequent breast events. In our approach, a woman could not have more than one event because the period of disease-free

survival ends with the occurrence of the first subsequent event. Subjects were censored at death from any cause other than breast cancer, upon entering an HMO, or end of follow-up.

In addition to treatment information, we included patient and area-based variables in the model. The SEER data included geographic region, comedo necrosis histology (which tends to be more aggressive, also called comedo), marital status, age, and race. The geographic region was based on the SEER cancer registry location as there were too few observations to use an analytic unit smaller than the cancer registry. We also wanted this study to be comparable to others that used the SEER registries as geographic indicators. Two comorbidity indices were calculated from Medicare claims based on the method of Klabunde (Klabunde, Potosky, Legler, & Warren, 2000), one for inpatient claims and one for outpatient and physician claims during the year prior to diagnosis. Covariates for area-based poverty level, educational attainment, and rural population were taken from the 1990 Census data at the zip-code level and matched to the SEER-Medicare data. Finally, we included a hospital-service-area level variable for the number of radiation oncologists per 100,000 population from the 1999 Dartmouth Atlas of Health Care in the United States. The study was approved by the first author's Institutional Review Board.

ANALYTIC METHODS

We compared three methods for estimating the effects of treatments for DCIS in a cohort. The first model is a “naïve” discrete-time duration model that does not account for unobserved heterogeneity; the second is a discrete-time semi-parametric hazard model with unobserved heterogeneity (frailty) as originally proposed by Heckman and Singer (Heckman & Singer, 1984) and fully developed by Jenkins (Jenkins, 1995); and the third is an adaptation of a simultaneous-equations model to a discrete-time survival model framework.

Naïve Model

We began by estimating the “naïve” model as a standard discrete-time duration model. The basic element of the discrete-time model is the hazard function, $h(t)$, that determines the probability of exiting the disease-free state (i.e., having a subsequent breast event) during period t , given survival until time t . We modeled $h(t)$ in a latent variable framework given by

$$Y_i^* = \alpha_1 T_{1i} + \alpha_2 T_{2i} + X_i \beta + g(\tau) + \varepsilon, \quad Y_i = 1 \text{ if } Y_i^* > 0, \text{ and } Y_i = 0 \text{ if } Y_i^* \leq 0, \quad \text{Model I (Naïve)}$$

where the subscript i represents the i th woman, Y^* is an unobserved latent variable that represents the propensity for a subsequent breast event (i.e., exit from the disease-free state), T_1 is an indicator for breast conserving surgery (BCS), T_2 is an indicator for radiotherapy (can only be received conditional on receipt of BCS), X is a vector of observable subject and area-based characteristics, $g(\bullet)$ is a function of time since treatment (duration, τ) that captures duration dependence of the health outcome (a subsequent breast event), and ε is a normally distributed error term. Because we assumed ε is normal, we operationalized the model as a probit regression. We began by estimating specifications of $g(\bullet)$ that are nonparametric in τ and tested down to more parsimonious parametric functions of τ . We also tested for interactions between $g(\tau)$ and T_1 , T_2 , and X , which could be important if, for

example, mastectomy has a relatively larger effect on survival than BCS in later years of follow-up. Admitting interactions between $g(\bullet)$ and other variables allowed us to formally test the proportional hazards assumption. We also estimated a simple Kaplan-Meier model for each treatment regimen.

The unit of observation in Model I was an annual transition (or annual disease-free survival (DFS) status). Thus, each subject in our sample contributed multiple observations to the analysis with the number of observations depending on her length of DFS and length of follow-up. If a subject was lost to follow-up (i.e., survived beyond the observation period, entered an HMO, or died from non-breast cancer cause), her subsequent observations were treated as censored. We estimated robust standard errors (Huber-White) to account for intra-person correlation in the data, clustering on a unique individual identifier (Huber, 1964). We fit Model I with nonparametric baseline hazards stratified by treatment, tested down to a model in which duration dependence was a quadratic function in time, and finally tested down to a proportional hazards specification of duration dependence.

For the probit estimates of α_j to be unbiased, $\text{Cor}(T_j, \varepsilon) = 0$. Unmeasured factors related to both T_j and ε , however, made this unlikely. For example, if the pathology showed extensive high nuclear-grade DCIS with comedo histology and close margins, the woman would be more likely to undergo mastectomy and, all else equal, would be more likely to have a recurrence than women without these characteristics (Silverstein et al., 1995; Silverstein & Lagios, 1997). Because we did not observe all tumor characteristics (e.g., margin status), we could not control for them. Thus, a subject who underwent mastectomy was likely to have worse risk than a subject who underwent the other treatment options, generating a bias against the effectiveness of mastectomy.

Frailty Models

Frailty models, introduced into survival analyses by Prentice and Gloeckler (Prentice & Gloeckler, 1978) in a breast cancer application, explicitly modeled individuals' unobserved heterogeneity in health status (i.e., frailty) in the hazard function. Later, Heckman and Singer (Heckman & Singer, 1984) proposed modeling the frailty distributions non-parametrically in continuous-time hazard models, and these ideas subsequently were applied to discrete-time survival models by Lancaster (Lancaster, 1991) and Jenkins (Jenkins, 1995), who also developed Stata programs to operationalize the model estimation (-hshaz-). We adopted the discrete-time models and estimated specifications with parametric and non-parametric frailty using Jenkins' programs (Jenkins, 2006).

Simultaneous Equations Model (SEM)

In our third approach, we adapted a familiar econometric simultaneous equations approach to a discrete-time survival application. (Stata programs available from lead author upon request.) This approach addresses the problem of unobserved heterogeneity (frailty) by identifying exogenous sources of variation in treatment. We formulated a three-equation model in which the first equation represented whether a subject received BCS or mastectomy, the second equation estimated whether the subject received radiotherapy

conditional on receiving BCS, and the final equation was the probability of a subsequent breast event during follow-up. Let

$$T_{1i}^* = X_i\beta_1 + Z_{1i}\gamma_1 + \varepsilon_{1i}, \quad (\text{Equation 1})$$

$$T_{2i}^* = X_i\beta_2 + Z_{21i}\gamma_2 + \varepsilon_{2i}, \quad (\text{Equation 2})$$

$$Y_i^* = \alpha_1 T_{1i} + \alpha_2 T_{2i} + X_i\beta_3 + g(\tau) + \varepsilon_{3i}, \quad (\text{Equation 3})$$

$$T_{1i} = 1 \text{ if } T_{1i}^* > 0, 0 \text{ otherwise, } T_{2i} = 1 \text{ if } T_{1i}^* > 0 \text{ and } T_{2i}^* > 0, 0 \text{ otherwise, and } Y_i = 1 \text{ if } Y_i^* > 0, 0 \text{ otherwise.}$$

Model III

(3-

Equation)

We assumed the error terms ($\varepsilon_1, \varepsilon_2$, and ε_3) had a trivariate normal distribution,

$f(\varepsilon_{1i}, \varepsilon_{2i}, \varepsilon_{3i}) \sim N(0, 0, 0, \rho_{12}, \rho_{13}, \rho_{23}, \sigma_1^2, \sigma_2^2, \sigma_3^2)$. We assumed $\sigma_1 = \sigma_2 = \sigma_3 = 1$, formulated the likelihood function, and estimated the parameters ($\beta, \gamma, \alpha, \rho_{12}, \rho_{13}, \rho_{23}$) using maximum likelihood methods. Our model was similar to a standard trivariate probit (Goldman et al., 2001) except that we did not observe T_2 unless $T_1 = 1$ because radiotherapy occurred only in combination with BCS. Therefore, T_2 (radiotherapy) was only observed conditional on BCS being performed. We then integrated out this area of the trivariate density so that we had a cohesive statistical model. Estimates of the parameters ρ_{13} and ρ_{23} allowed us to test whether estimates of α_1 and α_2 in Model I were unbiased; the estimates gave us the direction and magnitude of the unmeasured variables' effects on T_1 , T_2 , and Y . Estimates of Equation 3 parameters restricting $\rho_{13} = 0$ and $\rho_{23} = 0$ would produce identical estimates to those of Model I. We compared results from Models I and III in light of estimates of ρ_{13} and ρ_{23} .

Simultaneous equations models produce unbiased estimates of the structural parameters if the distributional assumptions are correct, and the models are quite robust to misspecification as long as there are adequate exclusion restrictions in the treatment equations. Without exclusion restrictions, however, Monte Carlo experiments have found that the models are highly sensitive to errors in distributional assumptions (Bhattacharya, Goldman, & McCaffrey, 2006). Because the distributional assumptions are typically ad hoc, the acceptability of the exclusion restrictions are crucial to evaluation of the strength of the evidence from the models. Our main empirical strategy was to identify plausible exclusion restrictions and estimate specifications that were identified by these restrictions. We also estimated specifications without the exclusions in order to examine the robustness of the results and provide evidence about the quality of the exclusion restrictions.

We identified measures to use as exclusion restrictions that were important determinants of treatment choice, but not outcomes (recurrence or subsequent invasive breast cancer). It has been theorized and empirically verified that the density of radiation oncologists and the urban-rural residence status of a subject would affect treatment choice (Gold & Dick, 2004). We conjectured, however, that conditional on treatment choices (including the use of radiation therapy), the availability of radiation oncologists would not affect outcomes directly, making this measure an ideal candidate for exclusion in the outcomes equation. We

made similar assumptions for urban-rural status because the additional travel burden may be associated with the selection of radiotherapy (Nattinger, Kneusel, Hoffmann, & Gilligan, 2001), but not outcomes directly, after controlling for observable socioeconomic status and histopathology.⁽¹⁾ We also excluded marital status from the outcomes equation because arguments that marital status explains outcomes are weak, particularly after controlling for other observable measures, but marital status could affect treatment choice because of differences in preferences (Hawley et al., 2009; Berz et al., 2009). Finally, we also excluded SEER registry indicators (i.e., geographic regions) from the outcomes equation because, although there is substantial variation in treatment across regions, conditional on observable characteristics (including treatments), arguments for why region or residence explains outcomes are weak. We found no evidence in our data to contradict this, but we did estimate alternative specifications that relaxed these assumptions and included each of the instruments (separately and together) in the outcomes equations. We found that our candidate instruments (radiation oncologist density, urban-rural status, geographic regions, and marital status) effectively predicted treatment choice (a requirement for good instruments). In addition, as described below, we found our estimated treatment effects to be robust to the various specifications of exclusion restrictions.

Standardized Predictions

Direct interpretation of the models' parameter coefficient estimates can be difficult because they are in the latent variable metric (standard deviation shifts in the standard normal latent variable) and, because the models were nonlinear, effect sizes depended on the values of the other variables. Thus, we generated standardized population survival functions to characterize substantive effect sizes. For each set of model estimates, we assigned subjects one at a time to each treatment alternative. Then, we predicted the hazard and survival functions from each subject and accumulated the survival functions across the entire sample. Thus, we eliminated any differences in population characteristics of actual treatment groups when we compared the survival effects of BCS alone, BCS with radiotherapy, and mastectomy on DFS.

RESULTS

In the sample of 3,888 treated women, 1,133 (29%) underwent BCS only, 1,281 (33%) underwent BCS with radiotherapy, and the remaining 1,474 (38%) underwent mastectomy. (Table 1) There were significant differences in treatment groups by many model covariates.

Differences in treatment use by registry ($\chi^2=240.3$, $p<0.0001$) indicated geographic variation in treatment patterns. Women in lower poverty areas were more likely to get radiotherapy, while those in higher poverty areas were more likely to get mastectomy ($\chi^2=71.4$; $p<0.0001$). Other significant differences by treatment group included year of diagnosis (indicating changing trends over time), marital status, age, race, rural versus urban residence, comedo necrosis histology, area-based educational attainment, and comorbidities.

Figure 1 shows the empirical DFS functions (Kaplan-Meier) for each treatment alternative. They take roughly the same shape, with mastectomy yielding better expected DFS than BCS

with radiotherapy, which is better than BCS alone ($p<0.005$ based on bootstrapped standard errors of differences in expected DFS).

Table 2 shows results for the outcomes models, including the naïve model (Model I), the frailty model (Model II), and the simultaneous three-equation model (SEM) (Model III). Model I estimates met the proportional hazards assumption (joint test of duration-treatment interactions, $p=0.3115$; results not shown). The results shown are for the model in which we excluded the measures that were candidates for exclusion in the SEM (Model III). The results, however, were substantively similar in the alternative specifications of each of these models that included these (ultimately excluded) measures in the outcomes equations.

Naïve Model Estimates (Model I)

There was evidence that outcomes were improving over time (calendar-year variables). Increased age was strongly associated with longer DFS, but there was no evidence that DFS differed by other SES or demographic variables. In an alternative specification that included SEER regions, the outcomes did not differ significantly across SEER regions ($\chi^2=9.36$, $p=0.31$). In addition, the histologic subtype comedo necrosis did not appear to be related to DFS. The point estimate was both small and precisely estimated. Finally, the estimated treatment effects indicated that there was a large and statistically significant difference favoring mastectomy over BCS with or without radiotherapy ($p<0.001$), and there was evidence indicating that BCS with radiotherapy was associated with better outcomes than BCS alone ($p=0.049$). The test statistic indicates that mastectomy is statistically significantly better than BCS with radiotherapy ($p<0.0001$).

Frailty Model Estimates (Model II)

The results of the models with parametric and non-parametric frailty were similar. Here we have shown estimates of the non-parametric frailty model because it had slightly better fit than the parametric models (either with gamma frailty or normal frailty). A comparison of the log likelihood functions showed the frailty model was superior to the naïve model. We found that 2 points of support were sufficient to characterize the frailty distribution, representing two types of individuals: those with more serious disease and those with less serious disease. The probability of membership in the more serious disease group was found to be just over 5%, and DFS was considerably worse for this group ($p<0.001$). The results were similar to those of the naïve model except that mastectomy's advantage over BCS alone and BCS with radiotherapy was substantially increased. This is consistent with the concern that women who received mastectomy had more serious disease (unobservable in the data). The benefit of radiotherapy was slightly larger than in the naïve model, though the difference was not statistically significant.

Simultaneous Equations Model (Model III)

The measures excluded from the outcomes equation (density of radiation oncologists, marital status, urban-rural status, and SEER region) were important predictors of treatments. (Results available from lead author on request.) We tested their joint significance in each treatment equation and found them to have substantial explanatory power ($p<0.0001$). There were substantively large and statistically significant differences in treatment choice by

SEER region. Alternative specifications that included SEER regions in the outcomes equations found regions to be insignificant (both individually and jointly) in the outcomes equation. Marital status was a strong predictor of radiotherapy receipt ($p < 0.001$), but not of BCS. The supply of radiation oncologists strongly affected (positively) choosing BCS, but given BCS, the supply density had little effect on the use of radiotherapy.

There were other important predictors of treatment choice, as well. There was a strong substantive time trend towards use of BCS, a weaker but still strong trend towards the use of radiotherapy, but little evidence of a time trend for outcomes. Age was a significant predictor of radiotherapy, with older women less likely to receive radiotherapy than younger women ($p < 0.001$ for each of the older age groups relative to the youngest age group). Age was not predictive of BCS, however. Comorbid conditions appeared (weakly) to reduce the likelihood of receiving radiotherapy, but there was little evidence that comorbidity was related to the choice to have BCS or mastectomy. Finally, comedo histology appeared to have a large substantive and statistically significant effect on treatment choice. Women with comedo necrosis were far more likely to have mastectomy, and conditional on receiving BCS, were far more likely to have radiotherapy than women with other histologic subtypes.

Model III SEM outcomes equation estimates yielded a substantively similar difference between BCS and mastectomy to that of the Naïve model, but a much larger benefit from radiotherapy. Thus, BCS alone was found to be considerably worse than mastectomy ($p < 0.05$), but DFS for the combination of BCS and radiotherapy was not statistically significantly worse than for mastectomy ($p = 0.434$). The remaining model estimates did not differ substantively from either the naïve or frailty models. Estimates of the correlation coefficients in the 3-equation model (ρ_{12} , ρ_{13} , and ρ_{23}) were modest in magnitude but jointly significantly different from 0 ($p = 0.035$). The correlation in the errors between receipt of radiotherapy and DFS (ρ_{23}) was positive (0.14, $p < 0.05$) and substantively important, indicating that women who had unmeasured worse disease were also more likely to receive radiotherapy conditional on having BCS. This accounts for the large difference in estimates of the effect of BCS alone in Models I and III.

We investigated the robustness of the treatment effect estimates by estimating models in which we relaxed the exclusion restrictions (including a specification with no exclusion restrictions). We found the model estimates were robust to the specification of instruments.

Table 3 presents the substantive effect sizes for each treatment option (expected years of DFS through 8 years for a standardized population). The table shows the relative treatment effect sizes across the 3 models. The frailty model predictions (combining predictions from both the more-serious and less-serious disease types) were similar to those of the naïve model. The 3-equation model estimates differed markedly, however, particularly for BCS alone and BCS with radiotherapy. As with Models I and II, the Model III estimates of mastectomy generated the best outcomes in terms of DFS up to 8 years following treatment (expected DFS=7.42 years). The Model III estimates of BCS with radiotherapy (expected DFS=7.23 years; not statistically significantly different from mastectomy) were substantially higher than the other models, and the Model III estimates of BCS alone (expected DFS =

6.53, $p < 0.01$ compared to mastectomy; $p < 0.01$ compared to BCS with radiotherapy) were substantially lower than the other models.

Figure 2 presents the predicted standardized population DFS estimates from Models I, II, and III, by treatment. Both Figure 2 and Table 3 indicate that the Model III method of addressing the treatment-selection effects resulted in substantially lower DFS for BCS alone than either of the alternative treatments (or models). A comparison of the predictions from Models I, II, and III in Table 3 shows that the estimated biases induced by selection (the differences between Model I and either Model II or Model III estimates) differed markedly by the estimation method. The frailty model estimates indicated that unobserved frailty was present, but the substantive conclusions were similar to those of the naïve model. The 3-equation model estimates, however, identified substantively important differences in the estimated outcomes by treatment.

Comparison to RCT Results

Although there are no RCT studies that compare mastectomy to BCS for DCIS, we identified 4 RCT studies of BCS with radiotherapy versus BCS without radiotherapy for DCIS, which allowed us to examine the performance of our models for these two treatments. Table 4 presents the results. We estimated the hazard ratio (HR)(BCS with radiotherapy versus BCS without radiotherapy) for each trial based on published data regarding the number of observations and events for each treatment. The RCT results were quite similar, ranging from 0.455 to 0.622. A comparison to the RCT shows that the Naïve model was biased substantially in favor of BCS without radiotherapy (HR much closer to 1.0). The Model III estimates (3-equation model) fell squarely within the RCT estimates. The frailty models fell in between. Although the model fit was better for the non-parametric frailty model than for the parametric frailty model, the parametric model estimates of the HRs, though still higher than the RCT estimates, were much closer than the estimates from the nonparametric model.

DISCUSSION

This study used econometric models to generate causal estimates of the effects of treatments for DCIS on DFS. By comparing the results of a naïve model that did not account for risk-based selection into treatment to estimates both from models that accounted for the selection and to some limited RCTs, we were able to assess the extent to which selection generates bias in traditional models and the extent to which the econometric methods (frailty and simultaneous equations approaches) successfully accounted for this source of bias. We found that the overall trends in DFS estimates by treatment were largely the same regardless of the model, with mastectomy yielding the highest DFS over 8 years of follow-up, followed by BCS with radiotherapy, and then BCS alone. The absolute rates and direction of bias, however, varied substantially by treatment strategy. DFS was underestimated by the single-equation (Model I) and frailty models (Model II), compared to the simultaneous equations model (Model III) and RCT results for BCS with radiotherapy, and overestimated for BCS alone, as indicated in Tables 3 and 4.

These results imply that there were important systematic, unobservable (to researchers) differences in the disease and health characteristics or preferences of women who received BCS alone compared to women who received either BCS with radiotherapy or mastectomy for DCIS. Although the frailty models, which are designed to address this kind of unobserved heterogeneity, improved the estimates somewhat, they were still considerably different from the RCT results. Our analyses showed that these differences led to significant understatements of the relative benefits of BCS with radiotherapy compared BCS alone.

Our simultaneous-equations model results are consistent with evidence from RCTs for DCIS treatment that show improvement in DFS with the addition of radiotherapy to BCS (Holmberg et al., 2008; Fisher et al., 1993; Fisher et al., 2001; Bijker et al., 2006) and high recurrence rates seen following BCS alone (Wong et al., 2006). The methods we used build on a published analysis of similar observational data for invasive breast cancer (Hadley et al., 2003). Our approach expands the analytic framework by allowing for nonlinear survival models, extending the time horizon, and estimating the treatment and outcomes equations simultaneously.

We focused considerable attention on the inclusion of comedo status in the outcomes equation, because a goal of the analysis was to determine which histopathologic factors generate the greatest risk for disease progression. Many surgeons believe comedo status is an important predictor of future recurrence (Silverstein et al., 1995), and indeed, our results confirmed that comedo histology was a very important predictor of mastectomy. We found no evidence, however, that comedo status was an important predictor of DFS. In each of our alternative specifications, we found the relationship between comedo histology and DFS to be both small and precisely estimated. Because we lack pathology and genetic data, one could suspect that comedo status is endogenous in the treatment-choice equations. However, comedo status does not depend on surgery type, and does not appear related to genetic information, so we think it is implausible that comedo status is endogenous in the treatment-choice equations. Having a mastectomy should not yield more comedo diagnoses. Comedo is thought to be a more aggressive subtype, therefore requiring more aggressive treatment. Our results confirm this treatment behavior, but we find no evidence that it is related to worse outcomes. It is worth noting that comedo is related to nuclear grade and necrosis. Previous work (Dick et al., 2010) has found that these pathologic measures are highly correlated, and that it is not possible to include all of the information in regression equations. In our case, therefore, comedo may be a proxy for a combination of unobserved pathologic measures. While this may complicate the interpretation of the role of comedo, it does not change the interpretation of the key measures in our models. We do note, however, that the measure of comedo histology may have been underreported because there were subjects for whom histologic information was incomplete. Our findings suggest that further research about the consequence of comedo histology on health outcomes is necessary, particularly because of the strong evidence that it shapes treatment choice. If our estimates are correct, this aggressive treatment might not be necessary. Our data lack detail regarding pathology, margin status, and family history, and because of the strong prevailing opinion that comedo histology matters, more research is needed to confirm our findings (Silverstein et al., 1995; Silverstein & Lagios, 1997).

Our study has some methodological and data-related limitations. First, identification assumptions were important in simultaneous equations models. Although our results were robust to specifications – and therefore the source of identification – they must be interpreted with the proper caution. Monte Carlo simulations have shown that such estimation works well when the parametric assumptions about the error terms (in this case, normality) are violated as long as there are “good” exclusion restrictions (Goldman et al., 2001). Our exclusion restrictions clearly met the first criterion for good instruments as strong predictors of treatment. Our assumptions that these measures do not belong in the outcomes models are plausible, but they are not testable. As noted above, our data did not contain detailed information on pathology, margin status, or family history, all of which may have affected both treatment and outcomes (Silverstein & Lagios, 1997). For these omitted variables to compromise our exclusion assumptions, they would have to be related to radiation oncologist supply, urban-rural status, region indicators, and marital status. We found our key results to be robust to the various choices of these exclusion restrictions. We also found the results to be robust to the inclusion of all of these measures, a specification that was identified only by the nonlinearities in the model. The similarity of the results to these alternative specifications lends credibility to the assumptions both about the instruments and the choice of normality for parameterization of the error distribution.

The linked SEER-Medicare database did not include information on younger women or tamoxifen use, which is known to be effective in reducing DCIS recurrence (Fisher et al., 1999). The cohort in this study underwent treatment prior to publication of the evidence on tamoxifen for DCIS (Fisher et al., 1999), so no patients during this time period were likely taking tamoxifen. This makes any unobserved differential effects of tamoxifen unlikely, if for example, patients receiving BCS alone would be more likely to take tamoxifen for its protective effect. Additionally, this analysis only included women ages 66 and older in fee-for-service Medicare, which may not be representative of all women diagnosed with DCIS (Nattinger, McAuliffe, & Schapira, 1997).

Our simultaneous-equations econometric approach to estimating disease-free survival in an observational cohort of DCIS patients showed promise in overcoming selection bias to provide better estimates of treatment effects. These methods could be adopted in future work by applying them to cohort data in other diseases and populations. The results also have important clinical implications; the intensity of mammography screening leads to increased incidence of DCIS (Ernster et al., 2002) yet there is a strong trend towards less intensive treatment (Gold & Dick, 2004). While “over-diagnosis” may be a reality, in fact, “over-treatment” may not be, given our findings that mastectomy yields the best DFS. If subjects do undergo BCS, they certainly should have radiotherapy to achieve better outcomes than BCS alone. This study used large retrospective cohort data coupled with methods to eliminate selection effects; future similar analyses, even for other cancers, also could eliminate endogeneity bias in cohort studies.

Acknowledgements

The interpretation and reporting of the Linked SEER-Medicare Database are the sole responsibility of the authors. The authors acknowledge the efforts of the Applied Research Program, NCI; the Office of Information Services, and the Office of Strategic Planning, CMS; Information Management Services (IMS), Inc.; and the Surveillance,

Epidemiology, and End Results (SEER) Program tumor registries in the creation of the SEER-Medicare database. This work was supported by DHHS/HCFA 30-P-91277/2-01 (HTG); NIH/NCI R01 CA922444-01A1 (AWD and JGG); and American Cancer Society MRSCT-04-002-01-CPHPS (HTG). Presented in part at the 24th Annual Meeting of the Society for Medical Decision Making, October 2002. The authors appreciate the thoughtful comments of the editor and two anonymous reviewers.

References

- Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? *Journal of Health and Social Behavior*. 1995; 36:1–10. [PubMed: 7738325]
- Baxter NN, Virnig BA, Durham SB, Tuttle TM. Trends in the treatment of ductal carcinoma in situ of the breast. *Journal of the National Cancer Institute*. 2004; 96:443–448. [PubMed: 15026469]
- Berz JP, Johnston K, Backus B, Doros G, Rose AJ, Pierre S, et al. The influence of black race on treatment and mortality for early-stage breast cancer. *Medical Care*. 2009; 47:986–992. [PubMed: 19648837]
- Bhattacharya J, Goldman D, McCaffrey D. Estimating probit models with self-selected treatments. *Statistics in Medicine*. 2006; 25:389–413. [PubMed: 16382420]
- Bijker N, Meijnen P, Peterse J, Bogaerts J, Van H, I Julien J, et al. Breast-conserving treatment with or without radiotherapy in ductal carcinoma-in-situ: ten-year results of European Organisation for Research and Treatment of Cancer randomized phase III trial 10853--a study by the EORTC Breast Cancer Cooperative Group and EORTC Radiotherapy Group. *Journal of Clinical Oncology*. 2006; 24:3381–3387. [PubMed: 16801628]
- Cutuli B, Cohen-Solal-Le NC, De LB, Mignotte H, Fichet V, Fay R, et al. Ductal carcinoma in situ of the breast results of conservative and radical treatments in 716 patients. *European Journal of Cancer*. 2001; 37:2365–2372. [PubMed: 11720829]
- Dick AW, Sorbero ME, Ahrendt GM, Hayman JA, Gold HT, Schiffhauer LM, et al. Comparative Effectiveness of the Management of Ductal Carcinoma in Situ and the Roles of Margins and Surgeons. *Journal of Clinical Oncology*. 2011; 103:92–104.
- Ernst VL, Ballard-Barbash R, Barlow W, Zheng Y, Weaver DL, Cutter G, et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. *Journal of the National Cancer Institute*. 2002; 94:1546–1554. [PubMed: 12381707]
- Ernst VL, Barclay J, Kerlikowske K, Grady D, Henderson IC. Incidence of and treatment for ductal carcinoma in situ of the breast. *JAMA*. 1996; 275:913–918. [PubMed: 8598618]
- Ernst VL, Barclay J, Kerlikowske K, Wilkie H, Ballard-Barbash R. Mortality among women with ductal carcinoma in situ of the breast in the population-based surveillance, epidemiology and end results program. *Archives of Internal Medicine*. 2000; 160:953–958. [PubMed: 10761960]
- Fahey, A. Documentation for the Patient Entitlement and Diagnosis Summary File. Maryland: Information Management Systems, Inc. & Applied Research Program, NCI; 15-. 2000
- Fisher B, Costantino J, Redmond C, Fisher ER, Margolese R, Dimitrov N, et al. Lumpectomy compared with lumpectomy and radiation therapy for the treatment of intraductal breast cancer. *New England Journal of Medicine*. 1993; 328:1581–1586. [PubMed: 8292119]
- Fisher B, Dignam J, Wolmark N, Wickerham DL, Fisher ER, Mamounas E, et al. Tamoxifen in treatment of intraductal breast cancer: National Surgical Adjuvant Breast and Bowel Project B-24 randomised controlled trial. *Lancet*. 1999; 353:1993–2000. [PubMed: 10376613]
- Fisher B, Land S, Eleftherios M, Dignam J, Fisher ER, Wolmark N. Prevention of invasive breast cancer in women with ductal carcinoma in situ: an update of the national surgical adjuvant breast and bowel project experience. *Seminars in Oncology*. 2001; 28:400–418. [PubMed: 11498833]
- Fonseca R, Hartmann L, Petersen IA, Donohue JH, Crotty TB, Gisvold JJ. Ductal carcinoma in situ of the breast. *Annals of Internal Medicine*. 1997; 127:1013–1022. [PubMed: 9412283]
- Gold HT, Dick AW. Variations in treatment for ductal carcinoma in situ in elderly women. *Medical Care*. 2004; 42:267–275. [PubMed: 15076826]
- Gold HT, Do HT, Dick AW. Correlates and effect of suboptimal radiotherapy in women with ductal carcinoma in situ or early invasive breast cancer. *Cancer*. 2008; 113:3108–3115. [PubMed: 18932243]

- Goldman DP, Bhattacharya J, McCaffrey DF, Duan N, Leibowitz AA, Joyce GF, et al. Effect of insurance on mortality in an HIV-positive population in care. *Journal of the American Statistical Association*. 2001; 96:883–894.
- Habel LA, Daling JR, Newcomb PA, Self SG, Porter PL, Stanford JL, et al. Risk of recurrence after ductal carcinoma in situ of the breast. *Cancer Epidemiology, Biomarkers and Prevention*. 1998; 7:689–696.
- Hadley J, Polsky D, Mandelblatt JS, Mitchell JM, Weeks JC, Wang Q, et al. An exploratory instrumental variable analysis of the outcomes of localized breast cancer treatments in a Medicare population. *Health Economics*. 2003; 12:171–186. [PubMed: 12605463]
- Hawley ST, Griggs JJ, Hamilton AS, Graff JJ, Janz NK, Morrow M, et al. Decision involvement and receipt of mastectomy among racially and ethnically diverse breast cancer patients. *Journal of the National Cancer Institute*. 2009; 101:1337–1347. [PubMed: 19720966]
- Heckman J, Singer BA. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*. 1984; 52:271–320.
- Hershman DL, Wang X, McBride R, Jacobson JS, Grann VR, Neugut AI. Delay in initiating adjuvant radiotherapy following breast conservation surgery and its impact on survival. *International Journal of Radiation Oncology*Biophysics*. 2006; 65:1353–1360.
- Holmberg L, Garmo H, Granstrand B, Ringberg A, Arnesson LG, Sandelin K, et al. Absolute risk reductions for local recurrence after postoperative radiotherapy after sector resection for ductal carcinoma in situ of the breast. *Journal of Clinical Oncology*. 2008; 26:1247–1252. [PubMed: 18250350]
- Huber PJ. Robust estimation of a location parameter. *Annals of Mathematical Statistics*. 1964; 35:73–101.
- Jenkins SP. Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics*. 1995; 57:129–138.
- Jenkins, SP. Boston College Department of Economics. Boston, MA: Statistical Software Components S444601; 2006. HSHAZ: Stata module to estimate discrete time (grouped data) proportional hazards models [Computer software].
- Julien JP, Bijker N, Fentiman IS, Peterse JL, Delledonne V, Rouanet P, et al. Radiotherapy in breast-conserving treatment for ductal carcinoma in situ: first results of the EORTC randomised phase III trial 10853. *Lancet*. 2000; 355:528–533. [PubMed: 10683002]
- Klabunde CN, Potosky AL, Legler JM, Warren JL. Development of a comorbidity index using physician claims data. *Journal of Clinical Epidemiology*. 2000; 53:1258–1267. [PubMed: 11146273]
- Lancaster, T. *The econometric analysis of transition data*. Cambridge: Cambridge University Press; 1991.
- National Cancer Institute. [Accessed 06/2006] Breast Cancer (PDQ): Treatment, Ductal Carcinoma In Situ. 2006. http://www.cancer.gov/cancertopics/pdq/treatment/breast/HealthProfessional/page5#Section_75
- National Institutes of Health. [Accessed 10/2009] State-of-the-Science Conference Statement: Diagnosis and Management of Ductal Carcinoma In Situ (DCIS). 2009. http://consensus.nih.gov/2009/DCIS%20images/dcis_draftstmt.pdf.
- Nattinger AB, Kneusel RT, Hoffmann RG, Gilligan MA. Relationship of distance from a radiotherapy facility and initial breast cancer treatment. *Journal of the National Cancer Institute*. 2001; 93:1344–1346. [PubMed: 11535710]
- Nattinger AB, McAuliffe TL, Schapira MM. Generalizability of the surveillance, epidemiology, and end results registry population: factors relevant to epidemiologic and health care research. *Journal of Clinical Epidemiology*. 1997; 50:939–945. [PubMed: 9291879]
- Paszat LF, Mackillop WJ, Groome PA, Boyd C, Schulze K, Holowaty E. Mortality from myocardial infarction after adjuvant radiotherapy for breast cancer in the surveillance, epidemiology, and end-results cancer registries[published erratum appears in *J Clin Oncol* 1999 Feb;17(2):740]. *Journal of Clinical Oncology*. 1998; 16:2625–2631. [PubMed: 9704712]
- Prentice RL, Gloeckler LA. Regression analysis of grouped survival data with application to breast cancer data. *Biometrics*. 1978; 34:57–67. [PubMed: 630037]

- Ries, LA.; Eisner, MP.; Kosary, CL.; Hankey, BF.; Miller, BA.; Clegg, L., et al. SEER Cancer Statistics Review, 1973–1997. Bethesda, MD: National Cancer Institute; 2000.
- Sakorafas GH, Farley DR. Optimal management of ductal carcinoma in situ of the breast. *Surgical Oncology*. 2003; 12:221–240. [PubMed: 14998563]
- Silverstein MJ, Lagios MD. Use of predictors of recurrence to plan therapy for DCIS of the breast. *Oncology (Williston.Park)*. 1997; 11:393–10. [PubMed: 9109132]
- Silverstein MJ, Poller DN, Waisman JR, Colburn WJ, Barth A, Gierson ED, et al. Prognostic classification of breast ductal carcinoma-in-situ. *Lancet*. 1995; 345:1154–1157. [PubMed: 7723550]
- Smith BD, Haffty BG, Buchholz TA, Smith GL, Galusha DH, Bekelman JE, et al. Effectiveness of radiation therapy in older women with ductal carcinoma in situ. *Journal of the National Cancer Institute*. 2006; 98:1302–1310. [PubMed: 16985249]
- Smith BD, Haffty BG, Buchholz TA, Smith GL, Galusha DH, Bekelman JE, et al. Effectiveness of Radiation Therapy in Older Women With Ductal Carcinoma In Situ. *JNCI Cancer Spectrum*. 2006; 98:1302–1310.
- Solin LJ, Kurtz J, Fourquet A, Amalric R, Recht A, Bornstein BA, et al. Fifteen-year results of breast-conserving surgery and definitive breast irradiation for the treatment of ductal carcinoma in situ of the breast. *Journal of Clinical Oncology*. 1996; 14:754–763. [PubMed: 8622021]
- Wamberg F, Yuen J, Holmberg L. Risk of subsequent invasive breast cancer after breast carcinoma in situ. *Lancet*. 2000; 355:724–725. [PubMed: 10703809]
- Warren JL, Riley GF, McBean AM, Hakim R. Use of Medicare data to identify incident breast cancer patients. *Health Care Financing Review*. 1996; 18:237–246. [PubMed: 10165033]
- Warren JL, Weaver DL, Bocklage T, Key CR, Platz CE, Cronin KA, et al. The frequency of ipsilateral second tumors after breast-conserving surgery for DCIS: a population based analysis. *Cancer*. 2005; 104:1840–1848. [PubMed: 16136599]
- Wong JS, Kaelin CM, Troyan SL, Gadd MA, Gelman R, Lester SC, et al. Prospective Study of Wide Excision Alone for Ductal Carcinoma in Situ of the Breast. *Journal of Clinical Oncology*. 2006; 24:1031–1036. [PubMed: 16461781]

Figure 1.
Empirical Disease-Free Survival

Figure 2.
Predicted disease-free survival, By Treatment

Table 1

Sample Characteristics, SEER 1991 – 1999, N = 3,888 Women

Variable	n	%	Test for treatment differences by group
Treatment			
Breast-conserving surgery (BCS)	1133	29.1	
BCS with radiotherapy	1281	33	
Mastectomy	1474	37.9	
Calendar Year			$X^2=120.0$; $p<0.0001$
1991	261	6.7	
1992	385	9.9	
1993	401	10.3	
1994	415	10.7	
1995	435	11.2	
1996	426	11	
1997	496	12.8	
1998	542	14	
1999	527	13.6	
Registry			$X^2=240.3$; $p<0.0001$
San Francisco	396	10.2	
Connecticut	703	18.1	
Detroit	889	22.9	
Hawaii	150	3.9	
Iowa	607	15.6	
New Mexico	159	4.1	
Seattle-Puget Sound	516	13.3	
Utah	171	4.4	
Atlanta	297	7.6	
Married	1905	49	$X^2=37.3$; $p<0.0001$
Age group			$X^2=225.5$; $p<0.0001$
Ages 66–69	1162	29.9	
Ages 70–74	1170	30.1	
Ages 75–79	871	22.4	
Ages 80 & over	685	17.6	
Race Group			$X^2=27.6$; $p<0.006$
White	3363	86.5	
Black race	247	6.4	
Other race	278	7.2	
Rural/Urban			$X^2=12.7.5$; $p=0.013$
Percentage living in rural zip code 0 & <10	2787	71.7	

Variable	n	%	Test for treatment differences by group
Percentage living in rural zip code 10 & <50	615	15.8	
Percentage living in rural zip code 50	486	12.5	
Comedonecrosis	921	23.7	
Variable	Mean	Std dev	Test for treatment differences by group
Physician claims comorbidity index	0.164	0.398	$X^2=6.4$; $p=0.04$ by categories
Inpatient claims comorbidity index	0.045	0.249	$X^2=13.5$; $p=0.009$ by categories
Radiation oncologists per 100,000 population	1.28	0.675	
Percentage living below poverty level in zip code	12.1	10.8	$X^2=71.4$; $p<0.0001$ by quartiles
Percentage without high school diploma in zip code	20.5	14.3	$X^2=12.7$; $p=0.048$ by quartiles

Table 2

Disease-Free Survival Model Estimates

Variable	Model I (Naïve)		Model II (Gamma Frailty)		Model II (Discrete Frailty)		Model III (3-Equation)	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
BCS	0.429	0.051 ***	0.814	0.156 ***	0.875	0.150 ***	0.452	0.211 **
RT	-0.107	0.049 **	-0.208	0.102 **	-0.145	0.081 *	-0.323	0.106 ***
Duration	-0.350	0.033 ***	-0.289	0.092 ***	-0.164	0.092 *	-0.353	0.033 ***
(Duration) ²	0.025	0.004 ***	0.023	0.007 ***	0.008	0.009	0.026	0.003 ***
Calendar year								
1991	0.186	0.101 *	0.295	0.186	0.302	0.194	0.125	0.104
1992	0.173	0.094 *	0.319	0.171 *	0.269	0.190	0.123	0.095
1993	0.180	0.093 *	0.198	0.163	0.236	0.185	0.173	0.092 *
1994	0.175	0.093 *	0.279	0.165 *	0.227	0.185	0.143	0.094
1995	0.133	0.094	0.180	0.161	0.161	0.192	0.108	0.093
1996	0.209	0.092 **	0.320	0.164 *	0.283	0.182	0.190	0.092 **
1997	0.050	0.095	0.088	0.160	0.048	0.190	0.054	0.093
1998	-0.095	0.099	-0.162	0.164	-0.245	0.212	-0.075	0.097
Ages 70-74	-0.047	0.048	-0.048	0.093	-0.055	0.083	-0.034	0.049
Ages 75-79	-0.146	0.056 ***	-0.202	0.106 *	-0.312	0.115 ***	-0.167	0.057 ***
Ages 80 & over	-0.240	0.073 ***	-0.394	0.143 ***	-0.571	0.192 ***	-0.280	0.076 ***
Black race	-0.056	0.091	-0.069	0.174	-0.008	0.148	-0.033	0.088
Other race	-0.035	0.082	-0.137	0.155	-0.161	0.166	-0.011	0.083
Physician Claims Comorbidity Index	0.014	0.057	0.052	0.109	-0.036	0.114	-0.004	0.059
Inpatient Claims Comorbidity Index	0.039	0.086	0.143	0.180	0.103	0.141	0.029	0.086
2nd lowest poverty quartile	-0.034	0.057	-0.142	0.119	0.002	0.106	-0.053	0.057
2nd highest poverty quartile	-0.051	0.063	-0.135	0.125	-0.014	0.116	-0.080	0.065
Highest poverty quartile	-0.076	0.066	-0.232	0.142	-0.008	0.123	-0.111	0.067
2nd lowest education quartile	-0.027	0.057	-0.064	0.110	-0.067	0.104	-0.030	0.057

Variable	Model I (Naïve)		Model II (Gamma Frailty)		Model II (Discrete Frailty)		Model III (3-Equation)	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
2nd highest education quartile	-0.010	0.060	0.008	0.115	-0.095	0.115	0.000	0.059
Highest education quartile	0.022	0.063	0.068	0.125	-0.057	0.119	0.022	0.062
Comedo necrosis histology	-0.009	0.048	-0.061	0.095	0.030	0.091	-0.002	0.054
Intercept	-1.510	0.106 ***	-1.594	0.188 ***	-2.466	0.330 ***	-1.404	0.174 ***
Gamma variance			3.519 *					
ρ_{12}							-0.044 *	
ρ_{13}							0.059	
ρ_{23}							0.140 **	
Discrete frailty 2nd mass point								
Location					2.537 ***			
Probability of "healthy" type					0.948 ***			
Probability of "unhealthy" type					0.052 ***			
Tests:	Statistic	P		Statistic				Statistic
BCS+RT = 0, chi2 (1)	37.50	<0.0001		22.460		24.100		0.610
Gamma variance = 0, chi2 (1)				5.123				
$\rho_{12}=\rho_{13}=\rho_{23} = 0$, chi2 (3)								8.622
N (Number of Women)	3888			3888		3888		3888
N (Number of Transitions)	20095			20095		20095		20095
ln(LF) (Outcomes)	-2108.2			-2102.6		-2098.4		
ln(LF) (Treatment & Outcomes)	-6046.3							- 6042.0

p<0.01

**
p<0.05

*
p<0.10

Table 3

Expected Disease-Free Survival, 8 Year Horizon*

	Empirical	Model 1 (Naïve)	Model 2 (Frailty)	Model 3 (3-Equation)
Mastectomy	7.51	7.51	7.49	7.42
BCS with Radiotherapy	6.72	7.00	6.96	7.23
BCS	6.95	6.76	6.78	6.53

Naïve Model: $P < 0.01$ for differences by treatment.

Frailty Model: $P < 0.01$ for differences between mastectomy and other treatments;

$P < 0.05$ for difference between BCS and BCS with radiotherapy.

3-Equation Model: $P < 0.01$ for differences between BCS alone and other treatments;

Difference between mastectomy and BCS with radiotherapy is not statistically significant.

* Predictions are based on standardized population predictions (except empirical), which eliminates the effects of differences in observed population characteristics across treatment groups. BCS = breast conserving surgery.

Table 4

Estimated Hazard Ratios: BCS with Radiotherapy versus BCS without Radiotherapy

Randomized Clinical Trial	HR	Model Estimate	HR
Holmberg, 2008	0.511	Model 1 (Naïve)	0.824
Fisher, 2001	0.557	Model 2a (Parametric Frailty)	0.686
Fisher, 1993	0.455	Model 2b (Nonparametric Frailty)	0.768
Bijker, 2006	0.622	Model 3 (3-Equation)	0.557

Notes: Randomized clinical trial results are based on outcome of local DCIS and invasive carcinoma, contralateral DCIS and invasive carcinoma, and regional and distant recurrence HRs for each RCT are approximate based on sample sizes and event counts as published for each trial.

HR = Hazard Ratio; RCT = Randomized Clinical Trial.