Published in final edited form as:

Med Care Res Rev. 2014 June ; 71(3): 224–242. doi:10.1177/1077558713510359.

Does Cancer Reduce Labor Market Entry? Evidence for Prime-Age Females

John R. Moran and

Department of Health Policy and Administration, Penn State University, 604 Ford Building,
University Park, PA 16802, Phone: 814-865-8893, Fax: 814-863-2905

Pamela Farley Short

Department of Health Policy and Administration, Penn State University, 604 Ford Building,
University Park, PA 16802, Phone: 814-863-8786, Fax: 814-863-2905

John R. Moran: jrm12@psu.edu; Pamela Farley Short: pxs46@psu.edu

Abstract

Existing studies of the labor market status of cancer survivors have focused on the extent to which cancer disrupts the employment of individuals who were working when diagnosed with cancer. We examine how surviving cancer affects labor market entry and usual hours of work among females age 28-54 who were not working when first diagnosed. We find that prime-age females have employment rates two to six years after diagnosis that are 12 percentage points lower than otherwise similar women who were initially out of the labor force, full-time employment rates that are 10 percentage points lower, and usual hours of work that are 5 hours per week lower. These estimates are somewhat larger than estimates for prime-age women employed at the time of diagnosis and highlight the importance of considering non-working females when assessing the economic and psychosocial burden of cancer.

Keywords

cancer; survivorship; employment; matching

Introduction

In the United States, where the average 5-year cancer survival rate has reached 65% (Siegel et al., 2012), improvements in early detection and treatment have made cancer survivorship an emerging chronic condition (Rowland, 2004). These trends, although welcome news, have led to an increased focus on the possibly deleterious effects of cancer and cancer treatment on the long-term well-being of survivors, a group now approaching 14 million in number (Siegel et al., 2012). One of the concerns is the long-run effect on the employment of cancer survivors (Steiner, Cavender, and Main, 2004). Employment is particularly relevant to survivors' quality of life because of the need to maintain a source of income and the possibly heightened need to maintain access to affordable health insurance (Bradley,

Neumark, and Luo, 2007; Tunceli, Short, and Moran, 2009), and because of the links between employment and personal identity, self-esteem, and other aspects of psychological and social well-being (Peteet, 2002; Holahan, Holahan, and Wonacott, 1999).

Existing studies of the labor market status of cancer survivors have focused on those working at the time of diagnosis, in an effort to better understand how cancer disrupts the employment trajectories of working survivors, both in the short run (Short, Vasey, and Tunceli, 2005; Bradley, Neumark, and Bednarek, 2005a; Bradley, Neumark, and Luo, 2005b; Kessler, Greenberg, and Mickelson, 2001) and longer term (Short, Vasey, and Moran, 2008). An important, but little studied, component of the overall economic and psychosocial burden of cancer is its role in reducing entry into the labor force among non-working individuals who might have otherwise sought employment in the future. This is especially important for female cancer survivors. Female labor force participation rates are roughly 15 percentage points lower than those of males (Bureau of Labor Statistics, 2009), so women are more likely to be out of work at the time they are diagnosed, and women often enter or re-enter the labor force as their children get older.

In this paper, we augment the literature on survivorship and employment by estimating how cancer affects employment, full-time employment, and usual hours of work among female cancer survivors age 28-54 who were not working when diagnosed with cancer two to six years earlier.¹ Although one might expect that cancer would not exert as much influence on the subsequent employment of non-workers, we find evidence of long-term employment reductions that are somewhat larger than those for women who were working at diagnosis (Moran, Short, and Hollenbeak, 2011).

New Contribution

To the best of our knowledge, this is the first study to examine the effect of surviving cancer on labor market entry. Our estimates illustrate the pervasive nature of cancer's effect on employment, and the need to consider entry into the labor force when assessing the overall economic burden of cancer. Preliminary analysis of the data sources used in this study, which allowed us to consider men and older women up to age 65, suggest that cancer's effect on entry is mainly an issue for somewhat younger women represented by the age group reported on here. Given the concentration of cancer incidence at older ages and generally high rates of female labor force participation, relatively few studies of cancer's longer term effects on employment have presented an opportunity to consider labor market entry in a group of younger survivors. A number have relied on data from the Health and Retirement Study, which does not enroll subjects before age 51, while others were not designed to capture a sufficiently large sample of cases diagnosed at younger ages while out of work.

Conceptual Framework

The standard economic model of labor supply, in which utility-maximizing agents allocate time between hours of market work and leisure, based on their labor-leisure preferences,

¹The number of prime-age males in our sample who were not employed at baseline (12 cancer survivors and 233 non-cancer controls) was too small to support a separate analysis for men.

relative prices (in this case, the market wage), and their wealth and time constraints, provides the theoretical backdrop for thinking about the effects of cancer on labor supply.² In the present case, one can view surviving cancer as a health shock that has the potential to alter decisions about entering the labor force in a number of ways: through productivity-related reductions in the available wage; through expenditures on health care or health insurance that reduce wealth; through a tightening of the time constraint due to the demands of treatment for long-term or late-term health effects; and through changes in labor-leisure preferences. Because several of these mechanisms potentially operate in opposite directions, the predictions of economic theory are ambiguous with regard to the net effect of cancer on labor market entry. As a result, the question motivating this study can only be resolved empirically.

Methods

Regression and propensity score matching methods were used to compare employment outcomes two to six years after diagnosis for prime-age female cancer survivors from the Penn State Cancer Survivor Survey, who were not working at the time of diagnosis, to a comparison group of similarly aged females from the Panel Study of Income Dynamics who were also not working at a comparable baseline. Our empirical methodology closely follows Short et al. (2008) and Moran et al. (2011), as described below.

Cancer Sample

The Penn State Cancer Survivor Survey (PSCSS) was a longitudinal study of nearly 1800 cancer survivors of both genders who were identified from the tumor registries at three large medical centers in central and northeastern Pennsylvania and at Johns Hopkins Hospital in Baltimore, Maryland. The research protocol was approved by Institutional Review Boards at Penn State's University Park campus and each hospital. The subjects were first diagnosed with cancer during the 3-year period from 1997 through 1999. They were surveyed in the first of four annual interviews from October 2000 to December 2001. For the analyses in this article, we compared employment data from the second PSCSS interview, conducted in 2002, with PSID data for 2002.

Patients with all types of cancer (except superficial skin cancers) were eligible for the cancer survey. However, most cases diagnosed at Stage 4 were excluded, except for patients with Stage 4 leukemias, lymphomas, and plasma cell cancers who had a good chance of surviving to the end of the study. Eligibility was further restricted to adults of working age (25 to 62 years old at diagnosis, implying that they were born between 1935 and 1974, inclusive) and to subjects who could be interviewed in English.

A total of about 5000 cases, 2000 from Johns Hopkins and 3000 from the Pennsylvania registries, met the eligibility criteria for sampling and were contacted by hospital employees who solicited participation in the study and obtained informed consent. Of the women young enough to be included in the analyses reported here (those aged 28-54 at follow-up), 48% percent of the initially sampled cases who were alive when contacted agreed to be

²See Bradley et al. (2002) for a theoretical discussion of how health shocks affect employment decisions.

interviewed (848 subjects, Table 1). Eighty-six percent of eligible cases who gave consent and survived were successfully interviewed (703 subjects). Eighty-eight percent of those who completed the first interview and survived until contacted for the second interview completed the second interview (598 cases). After restricting the 2002 sample to females who were not working at the time of diagnosis, there were 116 female cancer survivors in our target age range.

Concerns about bias raised by the low participation rate were investigated by using de-identified data for non-participants to estimate a logit model predicting completion of the first interview based on age, race, diagnosis year, facility, cancer types, and stage for women in the target age group. The only significant differences identified ($p < .05$) were lower participation rates for non-whites, at the one facility where written consent was required, and for un-staged cancers. A chi-square test showed that working at the first PSCSS interview was unrelated to attrition between the first and second interviews ($p = .561$) among females in the target age group.³

Our employment analyses rely on the date of diagnosis, along with cancer site and stage, recorded in the cancer registries. Retrospective questions in the first interview asked about employment and job characteristics, health insurance, and marital status at diagnosis. Current employment status and usual hours per week were determined at each of the four interviews. Time from diagnosis ranged from 26 to 70 months (mean = 47, standard deviation = 11).

Comparison Group

The Panel Study of Income Dynamics is an ongoing longitudinal study conducted by the Institute for Social Research at the University of Michigan (Institute for Social Research, 2009a). It has been collecting data from the same families and individuals since 1968. The PSID is designed to maintain a nationally representative sample of individuals and families in the U.S. population over time by following the children of initially sampled families as they leave home and form new households, thereby mimicking the process of family formation in the population. A sample of immigrant families was added in 1997 to maintain the representativeness of the sample in the face of increased immigration into the U.S. Since 1997, interviews have been conducted every two years (in odd years) to collect information about the previous (even) year.

We used the 2003 PSID family and individual files to obtain information on employment and other data elements for 2002.⁴ We restricted the PSID comparison group to the same age group as the cancer sample (28 to 54 years old in 2002). Although the PSID questionnaire contains extensive information for the householder and, if applicable, the

³Nonrandom survey participation (which includes attrition in longitudinal surveys) has the potential to result in sample selection bias when the factors driving nonparticipation are unobservable and correlated with unobserved determinants of employment or hours. Standard corrections require either a credible instrument for survey participation (i.e., a variable that affects participation but has no direct effect on employment), or strong assumptions about the joint distribution of the unobserved factors influencing participation and employment/hours.

⁴Some of the data used in this analysis are derived from Sensitive Data Files of the Panel Study of Income Dynamics, obtained under special contractual arrangements designed to protect the anonymity of respondents. These data are not available from the authors. Persons interested in obtaining PSID Sensitive Data Files should contact the PSID at psidhelp@isr.umich.edu.

householder's spouse, there is very little information about other adults in the household (beyond adult children eligible for the survey in their own right). Consequently, our PSID sample excludes the small proportion of adult females under age 55 who live in the household of a non-parent relative. We also dropped PSID subjects with any history of cancer (ignoring superficial skin cancers). Finally, because we required information for the PSID sample over the same time period as the cancer sample, we restricted our analytic PSID sample to primary adults with data in the 1997, 1999, 2001, and 2003 “family files” (corresponding to reference years 1996, 1998, 2000, and 2002 respectively).

To ensure a comparable distribution of follow-up intervals in the PSID relative to the PSCSS, we randomly chose from the distribution of diagnosis months in the cancer sample (January 1997 through December 1999) and used these draws to assign a “baseline” month for each respondent in the PSID for the purpose of determining initial employment status. A similar procedure was used by Bradley, Bednarek, and Neumark (2002) for comparing cancer survivors to a non-cancer comparison group. Job characteristics in the baseline month were compiled from monthly questions in the PSID about individual employment in the reference year, along with job beginning and ending dates. Employment information for the non-reference years of 1997 and 1999 was acquired from the monthly earnings data reported in the supplemental “t-2 individual income files” constructed by the Institute for Social Research from the 1999 and 2001 interviews (Institute for Social Research, 2009b).

When the PSID sample was restricted to females aged 28-54 who were not working during their assigned baseline month, we were left with 861 non-working women without a history of cancer in our comparison group.

Employment Outcomes

To estimate the effect of cancer on subsequent employment, we studied three employment outcomes observed in 2002: the probability of working; the probability of working full time (35+ hours in a usual week); and usual hours per week. In the PSCSS, working was defined by respondents selecting the first option when asked, “Now I am going to ask you some questions about your current employment situation. Are you working for pay, temporarily laid off or on leave, unemployed and looking for work, retired, disabled and unable to work, a homemaker, a student or something else?” In the PSID, working was defined based on answers to the question, “In which months during [year] were you working for [each employer identified during the year]?”

Identification Strategy

The estimand of interest is the average effect of treatment on the treated (ATT) which, in our case, corresponds to the average effect on employment and hours of having survived cancer, among those who have survived cancer. Strictly speaking, the ATT measures the effect of having survived cancer for current cancer survivors, but it can also be used to estimate the impact of cancer on future survivors if those who contract and survive cancer in the future are similar to those who survive cancer today.

In the absence of a viable instrumental variable, we base our analysis on a “selection-on-observables” assumption (Heckman and Robb, 1985): specifically, we assume that after conditioning on a set of observable covariates, one's propensity for developing cancer is independent of potential employment outcomes. This assumption is difficult to assess. One concern is that variation in preferences, e.g., in the rate at which individuals discount the future, could influence both one's willingness to invest in work-related human capital and, through lifestyle choices, one's risk of developing cancer, at least for the subset of cancers known to be influenced by personal behavior.⁵ Under this scenario, our estimates would overstate the negative impact of cancer on the labor market outcomes we study. An additional concern when estimating employment effects among those not working at baseline is that we have fewer proxies for laborforce attachment to include in our propensity score model. For example, in contrast to studies that focus on initially employed survivors, we cannot condition on baseline hours of work, occupational classification, or job tenure, variables that may provide valuable information about tastes for market work. If laborforce attachment varies systematically between our cancer and non-cancer comparison groups, the absence of proxies for this important preference parameter will result in biased estimates of the impact of surviving cancer on employment and hours.

The ATT is most commonly estimated using one of several matching estimators. Matching offers two potential advantages over linear regression. First, because the ATT is based on averaged differences among matched individuals with similar observed characteristics, no assumption regarding functional form is imposed in calculating the ATT. Second, matching provides a means of ensuring that only comparable individuals are compared. This contrasts with ordinary least squares regression, which uses all of the untreated observations to form an (implicit) counterfactual outcome for each treated individual.

In practice it is rare to match directly on the covariates because of the potentially large number of dimensions across which matches would need to be made. Rosenbaum and Rubin (1983) show that if the selection-on-observables assumption holds conditional on the covariates, then it also holds conditional on the predicted probability of treatment given the covariates, which allows one to construct matching estimates based on a set of unidimensional propensity scores rather than the multidimensional covariate vector.

When estimating the ATT, propensity score matching requires that for each distinct value of the propensity score in the treatment group there must be some untreated individuals available for matching. This requirement, known as the “common support” condition, can either be assumed or imposed, and there are several methods available for implementing it.⁶ We imposed the support condition prior to matching using a method known as “trimming” that removes q percent of the treated observations whose corresponding control observations

⁵Presumably because of low incidence and survival rates, we only had one case of lung cancer and two cases of head/neck cancer in our sample. Thus, there were only three cases of the malignancies most closely linked to behavior through their association with smoking and heavy drinking. Not surprisingly, omitting these individuals from our analysis had virtually no effect on the resulting estimates.

⁶In comparing a variety of matching estimators to experimental estimates based on randomization, Heckman, Ichimura, and Todd (1997) find that violating the support condition accounts for a substantial portion of the bias arising from the use of non-experimental estimators.

have an estimated density below an endogenously determined cut-off (Smith and Todd, 2005). We used a trimming level of 2 percent.⁷

We focused on two matching estimators, kernel matching and k -nearest neighbor matching. The methods differ in the way counterfactual outcomes are estimated. Kernel matching estimates the counterfactual using a distance-weighted mean of the comparison group observations.⁸ Observations with more similar propensity scores receive greater weight, with the extent of differential weighting determined by the chosen kernel and bandwidth. The matching estimates presented in Table 4 are based on the Epanechnikov kernel and a bandwidth of 0.06, which are the defaults for kernel matching using PSMATCH2.⁹ As a robustness check, we also experimented with a Gaussian (Normal) kernel and obtained similar results. K -nearest neighbor matching constructs counterfactual outcomes by using the k comparison group observations with propensity scores most similar to the treated observation, with equal weighting of the observations. The estimates presented below are based on $k = 10$. Similar estimates were obtained when $k = 5$ or $k = 1$.¹⁰

Propensity Score Model

The PSCSS and PSID samples differ along several dimensions; most notably, in terms of urbanicity and socioeconomic status (Table 2). This likely reflects the geographic concentration of the PSCSS sample. To adjust for these differences, we estimated propensity score models that include socio-demographic characteristics such as age, race, marital status, the presence of children under 18 in the home, and educational attainment, along with measures of urbanicity and the strength of each respondent's local labor market. We controlled for age with indicators for each year and measured schooling using indicators for less than high school (the omitted category), high school completion, some college, college completion, and any post-college education. The models also include indicators for five common chronic conditions (diabetes, chronic lung disease, heart disease, stroke, and arthritis) at follow-up.¹¹ Chronic conditions were identified in both surveys by asking, "Has a doctor ever told you that you had [condition]?" Although we assigned baseline dates to PSID respondents in a manner designed to mimic the distribution of diagnosis dates in the PSCSS, we also included a covariate measuring the number of months from diagnosis/baseline to the 2002 interview to ensure balance in the length of the follow-up periods across the two samples.

To control for differences in urbanicity and local labor market conditions, we constructed three county-level variables. The first variable was a set of three rural-urban indicators based on groups of Beale codes: (a) counties in metropolitan areas of 1 million population or more (code 1); (b) counties in smaller metropolitan areas (codes 2 and 3); and (c) all

⁷Smith and Todd (2005) use a trimming level of 2 percent.

⁸Frolich (2004) presents Monte Carlo evidence that kernel matching is often the best estimator in settings where the ratio of comparison to treated observations is 4:1 or better, as is the case for our sample.

⁹Heckman et al. (1997) use a bandwidth of 0.06.

¹⁰While the chosen number of neighbors is necessarily ad hoc, numbers in the range from $k=1$ to $k=10$ are common in the evaluation literature (Heckman et al., 1997; Smith and Todd, 2005).

¹¹In some cases, it seemed more appropriate to measure a variable at follow-up, coincident with when our employment outcomes were measured, rather than at baseline. To check for possible endogeneity, we re-estimated each of our matching models including only variables measured at baseline. Results were similar in all cases.

nonmetropolitan counties (codes 4 – 9). The second variable is population density, which we computed as the population per square mile in the respondent's county of residence. In addition to measuring where counties fall on the rural-urban continuum, these variables may also proxy for the size and extent of the local labor market. A more direct measure of labor market strength - the unemployment rate in the respondent's county of residence - serves as our third county-level variable. Each of these variables was calculated for calendar year 2002, to capture local labor market conditions at the time employment outcomes were evaluated.

The analyses do not distinguish non-working women who were unemployed at baseline from those who were out of the labor force. This distinction could not be made in the PSID for baseline months in 1997 and 1999, which were not survey reference years. For married women, who constituted a large majority of non-working women in both samples, the percentage in the target age group who were unemployed in the month of the 1999 PSID interview (10% of non-workers) was similar to the percentage in the PSCSS who were unemployed at diagnosis (12%). The fraction unemployed was less similar for unmarried women (35% in the PSID compared to 8% in the PSCSS), but there were only 12 unmarried women who were not working at diagnosis in the PSCSS.

The dependent variable in each propensity score model was equal to “one” if the respondent was a cancer survivor, and “zero” otherwise. All of the propensity score models were estimated as Probits. The estimated propensity score models (not presented) were consistent with the cross-sample comparisons depicted in first two columns of Table 2.¹² For comparison purposes, the third and fourth columns of Table 2 contain analogous summary statistics for the matched PSCSS and PSID samples that reflect the reweighting of the PSID sample used by the kernel estimator and the imposition of the common support condition. Paired *t*-tests for equality of means in the treatment and matched comparison groups were performed for every covariate in each specification, using the PSTEST procedure in PSMATCH2.¹³ In no case was the null hypothesis of equality of the covariates across samples rejected.

Regression and Matching Models

We regressed each of the three outcome variables on an indicator for cancer survivorship and the same set of covariates used in the propensity score models. In the case of discrete outcomes, such as the probability of working or the probability of working full-time, we estimated Probit models, reporting marginal effects. Estimates for usual weekly hours were based on an ordinary least squares regression. Robust standard errors were estimated for all regression models. For the matching estimators, we display two sets of standard errors: the analytic standard errors from the PSMATCH2 program and bootstrapped standard errors

¹²According to the authors' calculations based on Census Bureau population estimates and age- and gender-specific prevalence rates from SEER (Altekruse et al., 2010) for 2007, about 36% of female cancer survivors from age 30 to 49 nationally had breast cancer, similar to the 40% in our sample. As in our sample, thyroid cancer is next most common malignancy among female survivors in this age group nationally (14%). Cervical and uterine cancers account for another 13% of survivors nationally, compared to 9% grouped under “uterine” in our sample.

¹³In the interest of space, we do not present the balancing tests, which are available upon request from the authors. For information on the PSMATCH2 program, see Leuven and Sianesi (2003).

that incorporate the estimation error in the propensity scores. Each of the 500 bootstrap replications was based on the combined estimation of the propensity scores and the associated matching estimator.¹⁴

Results

Unadjusted employment outcomes for cancer survivors and PSID respondents without a history of cancer, among women who were not working at baseline, are presented in Table 3. At follow-up, survivors are less likely to work, less likely to work full-time, and work fewer hours per week than otherwise similar women with no history of cancer. Without adjusting for observable differences between the PSCSS and PSID samples, prime-age females who have survived cancer have employment rates that are 16.7 percentage points lower two to six years after diagnosis than similarly aged females from the PSID comparison group, full-time employment rates that are 16.1 percentage points lower, and usual weekly hours that are lower by 7.2 hours per week. For comparison with our regression and matching estimates, these mean differences are reproduced in the first column of Table 4.

The regression-based estimates and the ATT estimates derived from matching are generally a bit smaller than the unadjusted mean differences. Relative to the females in our non-cancer comparison group, female cancer survivors have employment rates that are 12 percentage points lower and full-time employment rates that are 10 percentage points lower. Female cancer survivors work about 5 hours per week less than comparable females without a history of cancer.

The control variables used in our regression and matching models influence employment and hours in the expected way. For example, employment rates in 2002 are 27 percentage points higher among college graduates than among those without a high school diploma, and college graduates work an average of 6 hours per week more in 2002 than those who didn't finish high school. Older females have lower employment and full-time employment rates, with age exerting a consistently negative effect on weekly hours beginning at age 42. Diabetes, arthritis, and chronic lung disease all significantly reduce the probability of working, while heart disease has the largest influence on full-time employment and hours (-19 percentage points and -5 hours per week, respectively). Finally, the presence of children in the home significantly lowers female labor market activity along all three dimensions: by roughly 10 percentage points for any employment, by 16 percentage points for full-time employment, and by 4 hours per week in the usual week.

Sensitivity Analysis

Ichino, Mealli, and Nannicini (2008) have developed methods for assessing the sensitivity of matching estimates to the selection-on-observables assumption. Their methodology allows researchers to simulate the effect of an omitted binary variable based on assumptions about the distribution of the omitted variable conditional on treatment assignment and outcome.¹⁵

¹⁴Bootstrapping may not produce valid standard errors for “nonsmooth” matching estimators, such as the nearest neighbor estimator (Abadie and Imbens, 2008), but should lead to valid inferences for “smooth” estimators, such as those based on kernel matching.

They propose two approaches for examining the robustness of matching estimates: the first simulates the effect of an omitted variable whose statistical properties mimic a given (included) covariate, while the second calculates the size of the selection and outcome effects required to produce an ATT estimate of zero.

We believe the first approach provides a more intuitive way of thinking about the influence of potential confounders and present a sensitivity analysis based on it in Table 5. Using the SENSATT command in Stata (Nannicini, 2007), separate simulations were conducted for all binary covariates whose mean values differed significantly across the treatment and comparison groups. The estimated ATT for each employment outcome and simulated confounder was based on 1000 replications of the simulation procedure. As shown in Table 5, our ATT estimates appear to be robust to the existence of omitted variables that are statistically similar to these covariates. When analyzing labor market entry, it is not possible to incorporate markers for laborforce attachment into our empirical model and, consequently, into the sensitivity analysis performed here.

Conclusions

Most previous studies of the employment effects of adult cancers have focused on people who were working when they were diagnosed with cancer. These studies have asked whether survivors return to work following cancer treatment (Spelten et al., 2002; Bouknight, Bradley, and Luo, 2006), considered absenteeism from work during treatment (Bradley, Oberst, and Schenk, 2005), and investigated the possibility that survivors are more likely to retire or quit their jobs than other workers (Short et al., 2005; Bradley et al., 2005a; Bradley et al., 2005b; Kessler et al., 2001; Short et al., 2008; Moran et al., 2011). However, in this study we call attention to the potential long-term effects of cancer on the employment of people who were not working when they were diagnosed, and who may have been more likely to enter the workforce if they had not had cancer. As a practical matter, the effect of cancer on labor market entry is mainly an issue for women in the younger to middle-aged group we consider here, because they are more likely to be out of the workforce caring for their families and to enter the workforce at a later point in their lives.

We find that female cancer survivors are significantly less likely to enter, or re-enter, the labor market than other women. Moreover, the long-term effects of cancer on the rate of any employment (-12% versus -7%), full-time employment (-10% versus -6%), and usual hours per week (-5 versus -3) seem larger here for non-workers than comparable estimates for females in the same age group who were working at diagnosis (Moran et al., 2011). Bradley and colleagues similarly found that employment rates for long-term breast cancer survivors aged 51-61 were 7 percentage points lower than for comparable women without a history of breast cancer (Bradley et al., 2002). However, if employed, women with a history of breast cancer in their study worked 3.4 more hours per week. The short term effects of breast cancer on employment are much larger. Six months after diagnosis, employment rates for women with breast cancer were 18-25 percentage points lower than for comparable women

¹⁵For details, see Ichino et al. (2008) or Nannicini (2007).

from a non-cancer control group (Bradley et al., 2005a). In the same study, employed women with breast cancer worked 7 fewer hours per week than other working women.

In comparing our estimates to those for survivors working at the time of diagnosis, two points merit consideration. First, because there is no baseline variation in employment and hours among the non-working women in our sample, we were unable to utilize a difference-indifferences estimation strategy to remove bias attributable to time-invariant unobservables, such as differences in the rate of time preference that might, under plausible assumptions, lead us to overstate the negative employment consequences of having survived cancer. Second, our selection-on-observables assumption is more tenuous when studying labor market entry than when studying employment disruption because we are unable to condition on baseline job and worker characteristics that may convey information about tastes for market work. Our inability to distinguish between women who were unemployed versus out of the labor market at diagnosis, which may serve as an indicator of laborforce attachment among non-working women, is another reason for interpreting our estimates with caution.

Ours and other longer-term estimates for women with cancer are fairly similar to estimates of the effects of other chronic illnesses on the employment of men and women. For example, respondents with chronic obstructive pulmonary disease (COPD) had employment rates that were 3.9 percentage points lower than respondents without COPD (Sin et al., 2002). Tunceli et al. (2005) found that diabetic women had employment rates that were 4.4 percentage points lower than non-diabetic women, while for men the corresponding reduction in employment was 7.1 percentage points. Changes in weekly hours were not significantly different across diabetics and non-diabetics of either gender. Obesity took a similar toll on labor market activity: meeting the clinical definition of obesity in 1986 lowered employment rates in 1999 by 4.8 percentage points among men and by 5.8 percentage points among women (Tunceli, Li, and Williams, 2006). Even larger effects have been documented for mental illness. Ettner, Frank, and Kessler (1997) found that employment rates were approximately 11 percentage points lower for both men and women with a psychiatric disorder, which is similar to the 11 percentage point reduction in employment rates among male Latinos found by Chatterji, Alegria, and Mingshan (2007), but substantially smaller than the 22 percentage point reduction found for Latino females. In the case of heart disease, labor market impacts appear to vary by age. In a recent study based on Australian data, the effect of cardiovascular disease (CVD) on employment rates was small and statistically insignificant among both men and women in the 18-49 age group, perhaps reflecting less advanced disease in younger respondents, whereas employment rates for men and women aged 50-64 with CVD were about 4 percentage points lower than among comparable individuals without heart disease, although the difference didn't rise to the level of statistical significance among women (Zhang, Zhao, and Harris, 2009).

Several limitations of our analysis should be noted. First, because our study is based on survivors treated at four medical centers in the mid-Atlantic region, our estimates may not be representative of cancer survivors generally. Also, because of the difficulties inherent in obtaining a series of interviews with cancer survivors, the survivors who agreed to be interviewed for our survey may differ from other survivors in ways that we cannot account for. Second, as discussed earlier, our analysis relies on a selection-on-observables

identification strategy that leaves open the possibility of biases from unobserved factors that potentially influence both employment and the probability of developing cancer (or of participating in a survey of cancer survivors). Moreover, in contrast to studies that examine the effect of cancer on initially employed individuals, we have a less informative set of covariates due to the absence of baseline job and worker characteristics among the sample of non-working women we use to analyze cancer's effects on labor market entry.

Finally, like most other investigators who have studied the labor market effects of chronic illness, we are unable to disentangle the underlying mechanisms through which cancer survivorship reduces labor market entry among prime-age females. The observed reductions in employment and hours among cancer survivors could be prompted by lingering physical and mental health effects of cancer and its treatment, changing tastes for market work in the aftermath of a serious illness, or discrimination on the part of employers. Each of these explanations has different implications for whether and what kind of clinical or policy responses are appropriate.

To the best of our knowledge, no other study of the employment effects of cancer, or chronic illness more generally, has considered labor market entry. Instead, most research in this area either estimates the impact of a chronic medical condition on those employed at the time of diagnosis or utilizes a sample comprised of both workers and nonworkers in the baseline period. Because the incidence of breast cancer among younger women is higher than the incidence of many other chronic diseases, cancer is especially likely to affect women at a stage of life when entering or re-entering the labor force is a consideration. Admittedly, given the relative numbers of workers and non-workers, the effect of cancer on women working at diagnosis contributes more to the total burden of cancer on survivors and society. Nevertheless, a full accounting must take into consideration survivors who do not pursue paid employment because of their encounter with cancer. Our study also serves as a reminder to oncology professionals and survivor groups to look out for non-working female survivors who may need special support in dealing with the twin challenges of transitioning into the workforce and living with a cancer history.

Acknowledgments

The authors are indebted to the Penn State Cancer Survivor Survey's Patient Advisory Committee; Witold Rybka and Margaret Davitt (Milton S. Hershey Medical Center), James Evans and Sharon Reeder (Geisinger Medical Center), Gregory Harper and Andrea Geshan (Lehigh Valley Hospital and Health Network), and James Zabora and Barbara Abdullah (The Johns Hopkins Hospital) for advice and patient recruiting; the Survey and Epidemiology Services division of Social and Scientific Systems, Inc. of Silver Spring, MD for conducting the interviews; and Cindy Mitchell for programming assistance.

Funding: This work was supported by grants from the American Cancer Society (Grant No. RSGHP-08-140-01-CPHPS) and the National Cancer Institute (Grant No. R01 CA82619), and research services funded by a grant from the National Institutes of Health to Penn State's Population Research Institute (Grant No. 2 R24 HD041025).

References

- Abadie A, Imbens G. On the failure of the bootstrap for matching estimators. *Econometrica*. 2008; 76(6):1537–1557.

- Altekruse, SF.; Kosary, CL.; Krapcho, M.; Neyman, N.; Aminou, R.; Waldron, W., et al. SEER cancer statistics review, 1975-2007. National Cancer Institute; Bethesda, MD: 2010. Accessed online at: http://seer.cancer.gov/csr/1975_2007/
- Bouknight RR, Bradley CJ, Luo Z. Correlates of return to work for breast cancer survivors. *Journal of Clinical Oncology*. 2006; 24(3):345–353. [PubMed: 16421415]
- Bradley CJ, Bednarek H, Neumark D. Breast cancer survival, work, and earnings. *Journal of Health Economics*. 2002; 21(5):757–779. [PubMed: 12349881]
- Bradley CJ, Neumark D, Bednarek HL, Schenk M. Short-term effects of breast cancer on labor market attachment: results from a longitudinal study. *Journal of Health Economics*. 2005a; 24(1):137–160. [PubMed: 15617792]
- Bradley CJ, Neumark D, Luo Z, Bednarek H, Schenk M. Employment outcomes of men treated for prostate cancer. *Journal of the National Cancer Institute*. 2005b; 97(13):958–965. [PubMed: 15998948]
- Bradley CJ, Neumark D, Luo Z, Bednarek H. Employment-contingent health insurance, illness, and labor supply of women: evidence from married women with breast cancer. *Health Economics*. 2007; 16(7):719–737. [PubMed: 17177273]
- Bradley CJ, Oberst K, Schenk M. Absenteeism from work: the experience of employed breast and prostate cancer patients in the months following diagnosis. *Psycho-Oncology*. 2005; 15(8):739–747. [PubMed: 16374893]
- Bureau of Labor Statistics. Civilian labor force participation rates by age, sex, race, and ethnicity. 2009. Accessed online at: http://www.bls.gov/emp/ep_table_303.htm
- Chatterji P, Alegria M, Mingshan L, Takeuchi D. Psychiatric disorders and labor market outcomes: evidence from the national latino and asian american study. *Health Economics*. 2007; 16(10): 1069–1090. [PubMed: 17294497]
- Ettner SL, Frank RG, Kessler RC. The impact of psychiatric disorders on labor market outcomes. *Industrial and Labor Relations Review*. 1997; 51(1):64–81.
- Frolich M. Finite-sample properties of propensity-score matching and weighting estimators. *Review of Economics and Statistics*. 2004; 86(1):77–90.
- Heckman, J.; Robb, R. Alternative methods for evaluating the impact of interventions. In: Heckman, J.; Singer, B., editors. *Longitudinal analysis of labor market data*. Cambridge: Cambridge University Press; 1985. p. 145-246.
- Heckman J, Ichimura H, Todd P. Matching as an econometric evaluation estimator: evidence from evaluating a job training program. *Review of Economic Studies*. 1997; 64(4):605–654.
- Holahan C, Holahan C, Wonacott N. Self appraisal, life satisfaction, and retrospective life choices across one and three decades. *Psychology of Aging*. 1999; 14(2):238–244.
- Ichino A, Mealli F, Nannicini T. From temporary help jobs to permanent employment: what can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics*. 2008; 23(3):305–327.
- Institute for Social Research. An overview of the panel study of income dynamics. 2009a. Accessed online at: <http://psidonline.isr.umich.edu/Guide/Overview.html>
- Institute for Social Research. T-2 individual income files: 1999-2001. 2009b. Accessed online at: <http://simba.isr.umich.edu/Zips/zipSupp.aspx#indsr>
- Kessler RC, Greenberg PE, Mickelson K, Meneades L, Wang P. The effects of chronic medical conditions on work loss and work cutback. *Journal of Occupational and Environmental Medicine*. 2001; 43(3):218–225. [PubMed: 11285869]
- Leuven, E.; Sianesi, B. PSMATCH2: stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. 2003. Accessed online at: <http://ideas.repec.org/c/boc/bocode/s432001.html/>
- Moran J, Short PF, Hollenbeak C. Long-term employment effects of surviving cancer. *Journal of Health Economics*. 2011; 30(3):505–514. [PubMed: 21429606]
- Nannicini T. Simulation-based sensitivity analysis for matching estimators. *The Stata Journal*. 2007; 7(3):334–350.
- Peteet JR. Cancer and the meaning of work. *General Hospital Psychiatry*. 2002; 22(3):200–205. [PubMed: 10880715]

- Rowland J. Cancer survivorship – United States, 1971-2001. *MMWR Weekly*. 2004; 53(24):526–529.
- Short PF, Vasey J, Tunceli K. Employment pathways in a large cohort of adult cancer survivors. *Cancer*. 2005; 103(6):1292–1301. [PubMed: 15700265]
- Short PF, Vasey J, Moran J. Long-term effects of cancer survivorship on the employment of older workers. *Health Services Research*. 2008; 43(1):193–210. [PubMed: 18211525]
- Siegel R, DeSantis C, Virgo K, Stein K, Mariotto A, Smith T, et al. Cancer treatment and survivorship statistics, 2012. *CA: A Cancer Journal for Clinicians*. 2012; 62(4):220–241. [PubMed: 22700443]
- Sin DD, Stafinski T, Ng YC, Bell NR, Jacobs P. The impact of chronic obstructive pulmonary disease on work loss in the United States. *American Journal of Respiratory Critical Care Medicine*. 2002; 165(5):704–707. [PubMed: 11874818]
- Smith J, Todd P. Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*. 2005; 125(1-2):305–353.
- Spelten ER, Sprangers MAG, Verbeek JHAM. Factors reported to influence the return to work of cancer survivors: a literature review. *Psycho-Oncology*. 2002; 11(2):124–131. [PubMed: 11921328]
- Steiner JF, Cavender TA, Main D, Bradley CJ. Assessing the impact of cancer on work outcomes. *Cancer*. 2004; 101(8):1703–1711. [PubMed: 15386303]
- Tunceli K, Bradley C, Nerenz D, Williams LK, Pladevall M, Lafata JE. The impact of diabetes on employment and work productivity. *Diabetes Care*. 2005; 28(11):2662–2667. [PubMed: 16249536]
- Tunceli K, Li K, Williams LK. Long-term effects of obesity on employment and work limitations among U.S. adults, 1986 to 1999. *Obesity*. 2006; 14(9):1637–1646. [PubMed: 17030975]
- Tunceli K, Short PF, Moran J, Tunceli O. Cancer survivorship, health insurance, and employment transitions among older workers. *Inquiry*. 2009; 46(1):17–32. [PubMed: 19489481]
- Zhang X, Zhao X, Harris A. Chronic diseases and labour force participation in Australia. *Journal of Health Economics*. 2009; 28(1):91–108. [PubMed: 18945504]

Table 1
Recruiting and interviewing outcomes for female survivors eligible for the PSCSS who would have been age 28-54 at follow-up

Selected for recruiting	1886
<i>Discovered ineligible for study (includes deaths)</i>	104
<i>Eligible, but no consent</i>	934
Unable to locate or contact	74
Refusal	724
Recruiting ended before consent was obtained	136
Consent obtained	848
<i>Ineligible for interview</i>	28
Discovered ineligible for study	27
Died before interview	1
<i>Eligible, but no interview</i>	117
Language problem	1
Physically incapacitated	2
Unable to locate or contact	58
Refusal	56
First Interview completed	703
<i>Ineligible for second interview (died)</i>	20
<i>Eligible for second interview, but no interview</i>	85
Physically incapacitated	2
Unable to locate or contact	37
Refusal	46
Second interview completed	598

Table 2
Characteristics of females age 28-54 not working at baseline/diagnosis, by survey

Characteristic	Before Matching		After Matching	
	Cancer	PSID	Cancer	PSID
Sample size	113	803	111	721
Cancer Type				
Blood	4%	--	5%	--
Breast	40%	--	39%	--
Central nervous system	4%	--	4%	--
Colorectal	4%	--	4%	--
Head/neck	1%	--	1%	--
Lymphoma	4%	--	5%	--
Other	9%	--	9%	--
Respiratory	2%	--	2%	--
Skin	2%	--	2%	--
Thyroid	20%	--	20%	--
Urinary	3%	--	3%	--
Uterus	9%	--	9%	--
Months since diagnosis / baseline	47	45	47	47
Any new cancers	11%	--	11%	--
Chronic conditions				
Diabetes	6%	7%	6%	9%
Chronic lung disease	12% *	3%	11%	10%
Heart disease	19% *	6%	18%	16%
Stroke	1%	2%	1%	1%
Arthritis	29% *	17%	30%	30%
Age: 28-33	5% *	20%	5%	6%
Age: 34-38	8% *	19%	8%	9%
Age: 39-43	15% *	24%	15%	14%
Age: 44-48	26% *	19%	25%	26%
Age: 49-54	46% *	19%	46%	45%
Nonwhite	5% *	46%	5%	8%
Married/partner (baseline)	90% *	70%	90%	91%
Children < 18	45% *	71%	46%	44%
Education				
Less than high school	12% *	21%	13%	14%
High school	35%	35%	34%	37%
Some college	22%	26%	23%	20%
College	21% *	12%	22%	20%
Post college	9%	7%	9%	10%
Urbanicity				

Characteristic	Before Matching		After Matching	
	Cancer	PSID	Cancer	PSID
Urban	31% *	48%	32%	29%
Suburban	45% *	32%	44%	46%
Not urban	24%	20%	24%	25%
County population density (pop / mile ²)	767*	1803	771	729
County unemployment rate	5.3% *	6.2%	5.3%	5.3%

PSID: Panel Study of Income Dynamics. An asterisk denotes a significant difference between the Cancer and PSID samples at the 0.05 level. Summary statistics prior to matching are unweighted. After matching, all summary statistics are based on the kernel weights used to reweight the PSID comparison sample. Because we are estimating the ATT, the Cancer sample was not reweighted, but it is approximately 2% smaller due to the imposition of the common support condition. The sample size for the (matched) PSID comparison sample is the number of PSID observations used by the kernel estimator.

Table 3
Employment, full-time employment, and usual hours per week in 2002 for females age 28-54 not working at baseline/diagnosis, by survey

	Cancer	PSID
Sample size	116	861
Working (percent)	32.8%	49.5%
Working full-time (percent)	13.9%	30.0%
Usual hours per week (mean)	9.2	16.4

Table 4
Effect of cancer survivorship on employment outcomes for females age 28-54 not working at baseline/diagnosis, 2-6 years post-diagnosis

	Unadjusted Mean Difference	Regression Adjusted Difference	Kernel Matching ATT	NN-10 Matching ATT
Working (% points) (N = 916)	-0.167** (0.047)	-0.155** (0.055)	-0.121* (0.060) [0.070]	-0.140** (0.060) [0.070]
Working Full Time (% points) (N = 908)	-0.161** (0.036)	-0.107** (0.045)	-0.102** (0.049) [0.051]	-0.121** (0.048) [0.053]
Usual Weekly Hours (N = 908)	-7.215** (1.597)	-4.560** (1.848)	-5.133** (2.121) [2.267]	-6.127** (2.086) [2.346]

Notes: Analyses included the following covariates: age, race, education, marital status, presence of children under 18 in the household, presence of specific chronic conditions, urbanicity, county population density, the county unemployment rate, and time since baseline/diagnosis. Regression-adjusted differences are marginal effects. For regression-based estimates, robust standard errors are shown in parentheses. For matching estimates, analytic standard errors are in parentheses; bootstrapped standard errors are in brackets. Bootstrapped standard errors are based on 500 replications of the combined propensity score - matching estimation routine. Estimates are statistically significant at $p < 0.10^*$, $p < 0.05^{**}$. Statistical significance of the matching estimates is based on the bootstrapped standard errors.

Table 5
Sensitivity analysis for matching estimates

	Working	Full Time	Hours
ATT: PSMATCH2	-0.121	-0.102	-5.133
ATT: SENSATT	-0.125	-0.100	-4.980
ATT: omitted binary variable similar to:			
Lung Disease	-0.107	-0.089	-4.374
Heart Disease	-0.107	-0.079	-4.055
Stroke	-0.127	-0.100	-5.033
Arthritis	-0.113	-0.090	-4.471
Nonwhite	-0.152	-0.078	-4.979
Married/partner	-0.131	-0.086	-4.808
Children < 18	-0.120	-0.101	-4.723
College	-0.136	-0.095	-5.004
Urban	-0.123	-0.091	-4.648
Suburban	-0.125	-0.091	-4.770

Notes: This table displays results of the sensitivity analysis developed by Ichino et al. (2008) using the SENSATT command in Stata. We simulate the effect on the estimated ATT parameters of an omitted binary confounder with the same statistical properties as each covariate listed in the first column of the table. We focus on binary covariates whose mean values differed significantly across the treatment and comparison groups. The estimated ATT parameter for each employment outcome and simulated omitted confounder is based on 1000 replications of the simulation procedure.