Quality

Was Breast Conserving Surgery Underutilized for Early Stage Breast Cancer? Instrumental Variables Evidence for Stage II Patients from Iowa

John M. Brooks, Elizabeth A. Chrischilles, Shane D. Scott, and Shari S. Chen-Hardee

Objective. To estimate the average survival effects of breast conserving surgery plus irradiation relative to mastectomy for marginal stage II breast cancer patients in Iowa from 1989–1994.

Data Sources/Data Setting. Secondary linked Iowa SEER Cancer Registry—Iowa Hospital Association discharge abstract data for women in Iowa with stage II breast cancer from 1989–1994.

Study Design. Observational instrumental variables (IV) analysis.

Data Collection/Extraction Methods. Women with stage II breast cancer from the Iowa SEER Cancer Registry 1989–1994 who received all of their inpatient care in Iowa were linked with their respective hospital discharge abstracts.

Principal Findings. Breast conserving surgery plus irradiation decreased survival relative to mastectomy for marginal stage II breast cancer patients in Iowa during the early 1990s. In this study marginal patients were those whose surgery choices were affected by differences in area treatment rates and access to radiation facilities.

Conclusions. If marginal patients are representative of patients whose treatment choices would be affected by changes in treatment rates, an increase in the breast conserving surgery plus irradiation rate for stage II early stage breast cancer patients would have decreased survival in Iowa during the early 1990s. Further research with newer data and broader samples is needed to make more current and specific assessments.

Key Words. Breast cancer, instrumental variables, treatment effectiveness

The initial treatment decision for patients with early stage breast cancer (stages I, IIa, and IIb) is the surgical approach for local tumor control—mastectomy (MAS) or breast conserving surgery plus irradiation (BCSI). In its 1991 Consensus Statement on this decision, the National Institutes of Health (NIH) recommended BCSI for most women with early stage breast cancer (ESBC) based on randomized controlled trial (RCT) evidence showing equivalent survival benefits of BCSI and MAS and the cosmetic superiority of BCSI

(National Institutes of Health Consensus Conference 1991). The NIH statement suggested that patients should be educated on the choices and they should make decisions consistent with their preferences. The BCSI rates increased in the United States afterward, but many ESBC patients still do not receive BCSI, and BCSI rates vary well beyond what would be expected based solely on differences in patient preferences (Benedict et al. 2001; Du et al. 2000; Guadagnoli et al. 1998; Keleman et al. 2001; Morrow et al. 2001; Riley et al. 1999).

The slow and varied rate of BCSI diffusion in the United States puzzled researchers and policymakers (Morrow et al. 2001; Adams 2001, Keating et al. 2001). This phenomenon has been attributed to a lack of provider knowledge of the RCT evidence. Interventions to increase the BCSI rate have been suggested (Keleman et al. 2001; Benedict et al. 2001; Nold et al. 2000; Stafford et al. 1998). An alternative explanation for the slow and varied BCSI diffusion rate, though, necessitates further research prior to initiating efforts to increase BCSI rates. Providers may have been aware of the NIH Consensus Statement and its supporting RCT evidence and they may have had beliefs consistent with the NIH for patients clinically similar to those in the RCTs. However, providers may have differed with the NIH in whether the RCT evidence could be extrapolated to ESBC patients that were clinically different from the average patients in the RCTs. Providers may have believed that MAS offered survival benefits for many ESBC patients and BCSI rate variation may have resulted from regional differences in these beliefs. If providers sorted patients between MAS and BCSI appropriately, an increase in the BCSI rate would have worsened survival for patients whose surgery choices were affected (i.e., ESBC patients who would have otherwise received MAS). Therefore, before initiating efforts to increase BCSI rates, policymakers need estimates of the

Address correspondence to John M. Brooks, Ph.D., Associate Professor, Health Effectiveness Research Center, College of Pharmacy and the College of Public Health, S-515 Pharmacy Bldg. University of Iowa, Iowa City, IA 52242. With the Health Effectiveness Research Center at the University of Iowa are Elizabeth A. Chrischilles, Ph.D., Professor, College of Public Health and College of Pharmacy; Shane D. Scott, PharmD., Associate Professor, College of Medicine; and Shari S. Chen-Hardee, M.S., Research Assistant, College of Public Health.

relative survival benefits of BCSI and MAS for the set of patients who would switch from MAS to BCSI as a result.

In this study we assume that the set of patients whose surgery choice would be more likely affected by a BCSI rate increase would be those patients for whom the RCT evidence supporting the survival equivalence between BCSI and MAS is the least certain. The RCT evidence supports the survival equivalence of BCSI and MAS for the "average" ESBC patient in each RCT (Veronesi et al. 1990; Arriagada et al. 1996; Jacobson et al. 1995; Fisher et al. 1995; van Dongen et al. 2000). However, if the survival benefits of BCSI and MAS are heterogeneous across ESBC patients, it is not clear whether this evidence can be generalized to patients with clinical circumstances differing from the average patients in the RCTs. Patients with ESBC are classified with stage I disease if they have localized tumors less than 2 cm with no lymph node involvement; stage IIa if they have either a localized tumor less than 2 cm with positive lymph node metastasis on the same side, or a tumor between 2 and 5 cm with no lymph node involvement; and stage IIb if they have either a localized tumor between 2 and 5 cm with positive lymph node metastasis on the same side, or a tumor greater than 5 cm with no lymph node involvement. Two RCTs included only stage I patients and provided compelling evidence that MAS offers no survival benefit for these patients (Veronesi et al. 1990; Arriagada et al. 1996). The evidence for stage II patients is less certain. The remaining studies contained a mix of stage I and stage II patients and each estimated a single average treatment effect (Jacobson et al. 1995; Fisher et al. 1995; van Dongen et al. 2000). Because no studies contained only stage II patients, it is less certain whether these estimates can be generalized to stage II patients. In fact, in the study with the most stage II patients (mainly stage IIa) tumor size and nodal involvement were related to an increased risk of local recurrence for BCSI patients but not for MAS patients (van Dongen et al. 1992).

As a result of this uncertainty, we focused this study on ESBC patients with stage II disease. We apply instrumental variable (IV) methods to obtain estimates of BCSI survival effects relative to MAS for stage II ESBC patients in Iowa from 1989–1994. Instrumental variable methods group using measured instrumental variables (instruments) that have the following two properties: (1) they are related to treatment choice and (2) they are assumed related to outcomes only through their effect on treatment choice (no direct effect on outcome and no indirect effect on outcome through unmeasured confounders). Instrumental variable estimates are obtained by exploiting the treatment variation across patient groups defined by the instruments. The inferences

made from IV estimates are conditional on the assumption that instruments essentially ex post randomize unmeasured confounders across patient groups. Given this assumption, IV methods yield consistent estimates of treatment effects for *marginal patients* (Angrist, Imbens, and Rubin 1996; Imbens and Angrist 1994; McClellan, McNeil, and Newhouse 1994; Harris and Remler 1998; Brooks, McClellan, and Wong 2000; McClellan and Newhouse 2000) that are defined as the subset of patients whose treatment choices varied with the instrument. In addition, previous IV research assumed that IV estimates can be generalized to the set of patients affected by treatment rate changes as marginal patients are also theorized to come from the set of patients for whom the best treatment is least certain (McClellan, McNeil, and Newhouse 1994; Harris and Remler 1998; Brooks et al. 2000a). However, marginal patients associated with a single instrument may not be representative of all patients potentially affected by a treatment rate change. We assess the validity of this assumption by using different instrument specifications. The nature of IV estimation limits our ability to generalize the findings beyond the set of marginal patients defined by our instruments, but this empirical scenario provides an opportunity to demonstrate the applicability of IV methods to policy-based research questions.

INSTRUMENTAL VARIABLES APPROACH

Previous research attributed variation in ESBC surgical choice across geographic areas to regional differences in "surgical philosophy" or "surgeon propensity" (Guadagnoli et al. 1998; Sainsbury et al. 1995; Foster, Farwell, and Costanza 1995; Answini et al. 2001; Iscoe et al. 1994; Mandelblatt et al. 2001). We theorize that there are regionally distinct surgical philosophies that lead to regional differences in BCSI rates. Empirically, we measure surgical philosophy in the area surrounding each stage II patient as the BCSI percentage of ESBC surgeries for *all other* ESBC patients (stage I and II) in a 50-mile radius around each patient's residence in their diagnosis year. Our IV approach groups ESBC stage II patients based on these rates and exploits the BCSI rate differences for stage II patients across these groups to estimate the average survival benefits of BCSI relative to MAS for marginal stage II patients.

   Our theory relies on the notion that differences in patient access to providers with different surgical philosophies *at the time of diagnosis* lead to different surgery choices for marginal ESBC patients. We theorize that

patients initially seek care from providers closer to their residences and the surgical philosophies of these providers weigh in the surgical decisions made by ESBC patients (even if surgery is eventually performed outside the patients local area). With respect to unmeasured confounders, there is little theoretical basis linking patient residence decisions made *prior* to an ESBC diagnosis with unmeasured patient severity found *after* diagnosis. Without more data collection it is impossible to validate directly whether instruments are unrelated to unmeasured confounders, and spurious correlations between instruments and unmeasured confounders may remain that bias our estimates. For example, area BCSI rates may be correlated with patient socioeconomic status or the general access of patients to health care. We mitigated this risk to our estimates by specifying area poverty percentages and the distance to the nearest hospital in our IV analysis.

In addition, the marginal patients defined by a single instrument may not fully describe the set of patients affected by an increase in BCSI rates. Alternative instruments may affect the surgical choices of distinct subsets of patients. Accordingly, if surgical effects are heterogeneous, variation in IV estimates across instruments suggests that the set of marginal patients varies with the instrument and that models specifying more than one instrument may yield estimates more representative of the average stage II ESBC patient affected by an increase in BCSI rates. To evaluate this possibility, we also grouped patients by the distance from their residence to the nearest radiation treatment center based on the theory that travel cost affects surgery choice. Patients with residences further from radiation treatment centers at the time of diagnosis have been less likely to receive BCSI (Hadley and Mitchell 1997). We specified models with each instrument separately and with both instruments.

METHODS

Data

Our data came from the Iowa Surveillance, Epidemiology, and End Results (SEER) Program Cancer Registry, the Iowa Hospital Association (IHA) inpatient discharge abstract files, and the Census Bureau's 1990 Zip Code Summary Tape File 3B. The Iowa SEER Registry provided the universe of first primary ESBC patients (stages I, IIa, and IIb) diagnosed during 1989–1994 that had either MAS or BCSI indicated within their first course of treatment ($n = 8,143$). To collect comorbidity and payer data we excluded

patients that received inpatient care outside of Iowa ($n = 848$) and linked the remaining patients ($n = 7{,}296$) to their respective inpatient discharge abstracts from the Iowa Hospital Association database. A previous paper provides a summary of the linkage approach and linkage validation statistics (Brooks et al. 2000). Inpatient discharge abstracts were linked to 84 percent of stage I patients and 89 percent percent of stage II patients. Of the linked patients, 2,905 were either stage IIa or IIb.

For each of the 2,905 stage II patients, Iowa SEER data were used to create binary variables defining treatment choice (MAS or BCSI), survival (alive 1 year, 2 years, 3 years, 4 years after diagnosis), cancer stage (T,N,M), cancer grade, tumor location, and age (younger than 50, 50–64, 65–69, 70–74, 75–79, 80–84, older than 84). We used the discharge abstracts to specify binary variables for payer (Medicare, Medicaid, Blue Cross/Blue Shield, other private, other government, no third party), and a Charlson comorbidity index based on the diagnoses within their initial inpatient discharge after diagnosis (Charlson et al. 1987). We created binary variables for the poverty percentage in the patient's zip code using data from the 1990 zip code summary file (#7, 7–10, 10–13, 13–20, > 20), and distance from residence to nearest hospital (#2.83, 2.83–9, 9–15, > 15) in the patient's year of diagnosis using hospital zip code data from the Iowa Hospital Association. Distances between zip code centroids were used to calculate the BCSI percentage of ESBC surgeries for *all other* ESBC patients (stage I and II) in a 50-mile radius around each patient's residence in their diagnosis year. The distance from each patient to the nearest radiation treatment center was calculated using the miles from each patients residence zip code centroid to the centroid of the zip code containing the nearest radiation treatment center in the year the patient was diagnosed that we obtained from Iowa SEER.

Empirical Model and Estimation Approach

We employed a nonparametric two-stage least squares (2SLS) variant of IV estimation that uses a minimum of distributional assumptions. This approach was used in previous IV research in healthcare (McClellan, McNeil, and Newhouse 1994; McClellan and Newhouse 1997; Brooks, McClellan, and Wong 2000) and in questions of labor supply (Angrist and Evans 1998, Angrist 2001). Consistent estimates are yielded by 2SLS regardless of the underlying error distributions, whereas alternative estimators that rely on error term distributional assumptions are inconsistent if the assumptions are wrong (Angrist 2001). In the first stage of the 2SLS approach we used ordinary least

squares (OLS) to estimate the following model of surgery choice:

$$R_i = a_0 + a_1 X_i + a_2 A_i + (c_i + e_i), \qquad (1)$$

where R_i is the surgical choice for patient i (1 = BCSI, 0 = MAS), X_i is a vector of binary variables containing measured confounding variables (diagnosis year, age groups, tumor size, nodal status, tumor grade, tumor location, comorbiditity index, insurance status, distance to nearest hospital, zip code poverty level), c_i is the effect of unmeasured confounders that affect both surgery choice and patient survival, and e_i is the net impact of the set of unmeasured factors that affect surgery choice only. A_i is a vector of binary variables that group patients based on each patient's instrument value. The distribution of each instrument was assessed across the sample, and cutoff values for each instrument were determined to divide patients into groups of similar size. A Chow F-test (Chow 1960) of whether A_i describes a significant portion of the variation in R_i (i.e., whether the estimates of a_2 are simultaneously equal to zero) provides a natural test of whether the instruments affect treatment choice.

In the second stage of 2SLS, we estimated survival models using four different survival measures (1 year, 2 years, 3 years, 4 years). Each survival model was specified as follows:

$$S_i = b_0 + b_1 X_i + b_2 R_i + (c_i + u_i), \qquad (2)$$

where S_i is binary variable equal to 1 if patient i survives beyond a certain time interval past their diagnosis (1 year, 2 years, 3 years, 4 years), X_i, R_i, and c_i are defined as in equation (1), and u_i is the set of unmeasured factors that affect patient survival and not surgery choice. The average survival effect of BCSI relative to MAS is represented by b_2. Estimating equation (2) using OLS or another analysis-of-variance method (ANOVA) will yield a biased estimate of b_2 if c_i is not equal to zero.

When estimating equation (2), 2SLS avoids this bias by replacing the actual surgery variable in equation (2) R_i with the predicted BCSI surgery probability from equation (1) for each patient $\hat{R}_i$. Bias is avoided because X_i and A_i are the only sources of variation for $\hat{R}_i$ from equation (1), and because X_i is also specified in equation (2), the only variation in $\hat{R}_i$ that is used to estimate b_2 in equation (2) is the variation in $\hat{R}_i$ that is attributable to A_i. Because A_i is assumed to be unrelated to c_i, the IV estimate of b_2 provides a consistent estimate of the change in the survival rate from a one-unit change in the BCSI rate that is *only generalizable to the group of patients whose treatment choices were affected by the instruments—the marginal patients.*

If A_i is specified using a single dummy variable, the sample is divided into two groups and b_2 is estimated using the surgery rate differences between the two groups (e.g., Earle et al. 2001). If A_i is specified with several dummy variables that divide the sample into several groups, the empirical model is *overidentified* and b_2 is estimated as the weighted average of the many two-group estimates that are available. When an empirical model is overidentified, a Hausman statistic (Hausman 1983) for overidentifying restrictions can be used to test the null hypothesis that the exclusion of A_i from the outcome equation was appropriate (i.e., A_i affects S_i only through R_i). A large value of the Hausman statistic rejects the null hypothesis.

Little theoretical guidance exists to specify the number of binary variables (A_i) for each instrument. Adding groups increases the number of two-group comparisons used in estimation, but lowers the number of patients in each group and increases the risk of introducing spurious relationships that violate the IV assumptions (e.g., membership in a particular group may be perfectly correlated with an unmeasured confounder). We assessed the robustness of our findings by varying the number of patient groups for each instrument (2, 4, 8, and 12 groups).

RESULTS

Table 1 provides univariate comparisons of BCSI rates and measured confounders by grouping method to assess the properties of our instruments. A comparison of the third (MAS) and fourth (BCSI) columns shows that ESBC stage II patients receiving BCSI had characteristics related to lower survival risk regardless of surgery choice (smaller tumors, younger, fewer comorbidities). Columns 4 and 5 compare patients grouped by whether the BSCI rates in the area around their residences were lower or greater than the median, respectively. Columns 6 and 7 compare patients by whether they lived farther or nearer from a radiation treatment center, respectively. Stage II patients were less likely to receive BCSI if they lived in an area with lower BCSI rates across all ESBC patients (7.3 percent versus 12.1 percent) or were farther from a radiation treatment facility (6.5 percent versus 13.0 percent). For each instrument there were no distinct differences in tumor size or stage between groups. Significant differences in age and tumor grade were observed. Tumor grade differences appear attributable to differences in the percentage of patients with an unknown tumor grade (9) and Iowa SEER Registry officials suggest these probably reflect regional differences in reporting practices and

Table 1: Comparison of Patient Characteristics across ESBC Stage II Patients in Iowa Grouped by Treatment and Instruments, 1989–1994

	Full Sample	Treatment		Area BCSI Percentage[a]		Radiation Distance[b]	
		MAS	*BCSI*	*Lower*	*Greater*	*Far*	*Near*
BCSI (%)	9.7	0.0	100.0	7.3***	12.1	6.5***	13.0
Age (%)							
<65	50.1	48.0***	69.3	48.5*	51.6	44.7***	55.4
65–74	24.4	24.7	21.6	24.7	24.1	25.8*	23.0
75+	25.5	27.3***	9.2	26.8	24.3	29.5***	21.7
Tumor Size (%)							
T1 (2cm)	28.5	27.3***	39.6	29.4	27.7	28.7	28.4
T2 (2–5 cm)	68.7	69.7***	59.7	67.9	69.6	68.9	68.6
T3 (>5cm)	2.8	3.0**	0.7	2.8	2.8	2.5	3.1
Positive Nodes (%)	59.2	59.2	60.1	59.4	59.2	59.2	59.4
Stage IIB (%)	33.6	34.9***	21.2	32.8	34.3	33.0	34.1
Grade (%)							
1	3.6	3.5	5.3	2.6***	4.7	3.5	3.8
2	20.9	20.6	24.0	16.7***	25.0	22.4**	19.4
3	32.4	32.4	32.5	31.3	33.5	32.1	32.7
4	6.4	6.6	4.6	8.4***	4.4	8.7***	4.1
9 (unknown)	36.7	37.0	33.6	41.1***	32.7	33.3***	40.0
Charlson Comorbidity Index >0 (%)[c]	20.4	21.4***	11.3	19.5	21.3	21.4	19.4
Number of Patients in Group	2,905	2,622	283	1,437	1,468	1,445	1,460

***, **, *Statistically different rate across groups at .99, .95, and .90 confidence, respectively.

[a]Patient in "less" group if less then 20% of all ESBC surgeries (stage I and II) in the 50-mile radius around the patient's residence in the year of diagnosis were BCS; [b]Patient in "near" group if distance to radiation treatment center in year of diagnosis is less than 19 miles.; [c]Modified Charlson comorbidity indices developed by mapping Clinical Classifications Software (CCS) diagnosis and procedure groups available on each HCUP discharge abstract into Charlson index groups.

Table 2:   F-Statistics Testing Whether Patient Groups Defined by Model Variables Are Related to BCSI Choice for Stage II ESBC Patients in Iowa, 1989–1994[a]

| | Surgery Choice Model Specification | | | |
| | Patient Groups per Instrument | | | |
	2	4	8	12
Instrument Groups				
Area BCSI rate	8.57***	5.19***	3.43**	3.00***
Covariate Groups				
Age[b]	6.55***	6.49***	6.38***	6.78***
Tumor size[b]	14.67***	15.10***	15.73***	15.52***
Positive nodes[b]	12.86***	12.76***	13.15***	12.99***
Grade[b]	0.98	0.93	0.98	0.94
Tumor location[b]	5.77***	5.59***	5.58***	5.52***
Comorbidity index[b]	0.96	0.91	0.96	0.76
Poverty[b]	4.02***	3.78***	3.96***	3.37***
Hospital access[b]	0.65	0.55	0.56	1.07
Payer[b]	3.12***	3.32***	3.30***	3.29***
Year[b]	2.46**	2.31**	1.79	2.06*

***,**,*Statistically significant at .99, .95, and .90 confidence, respectively.
[a]Against the null hypothesis that the groups defined by the instrumental variables do not describe a significant portion of treatment variation.
[b]Binary variables for age groups (<50, 50–64, 65–69, 70–74, 75–79, 80–84, 85+), tumor sizes (<2 cm, 2–5 cm, 5+cm), positive lymph node involvement, tumor grade groups (1, 2, 3, 4, 9–unknown), tumor location groups (nipple, central portion, upper-inner quad, lower-inner quad, upper-outer quad, lower-outer quad, axillary tail, overlapping lesion, not stated), Charlson comorbidity index (0, 1, 2, 3+), residence zip code poverty percentage (#7, 7–10, 10–13, 13–20,>20), distance from residence to nearest hospital (#2.83, 2.83–9, 9–15, >15), payer (Medicaid, Medicare, Blue Cross/Blue Shield, other private, other government, self-pay), and year of diagnosis (1989, 1990, 1991, 1992, 1993, 1994).

not differences in disease severity. Differences in patient age reflect the pockets of rural elderly in Iowa. We control directly for tumor grade, age, and distance to hospitals in our IV analysis, but our results are conditional on the assumption that differences in measured covariates are not symptomatic of differences in unmeasured confounders across instrument groups.

Table 2 contains the estimates of the Chow F-statistics testing the statistical significance of the instruments in equation (1) for 2-, 4-, 8-, and 12-instrument group models using the area BCSI rate. Local-area overall BCSI rates describe a statistically significant portion of the variation in surgery choice for stage II patients across model specifications. Stage II ESBC patients who lived in areas with higher overall BCSI rates were more likely to receive

BCSI. Tumor size, tumor location, and lymph node involvement also explained a statistically significant portion of the variation in surgery choice. Stage II ESBC patients with smaller tumors, negative lymph nodes, and with tumors located in the lower-inner quadrant, the upper-outer quadrant, and the axillary tail were more likely to obtain BCSI. Additionally, younger patients and patients living in higher poverty areas were more likely to obtain BCSI.

Table 3 contains IV and OLS estimates of the effect of BCSI relative to MAS on patient survival. The interpretation of these estimates varies with the estimation approach. Row 1 contains unadjusted OLS estimates (estimates of equation [2] without specifying X_i), and row 2 contains adjusted OLS estimates (estimates of equation [2] with X_i as described in the table). Using unadjusted OLS, patients receiving BCSI appeared to have a higher probability of surviving three and four years after diagnosis. Adjusted OLS estimates directly controlled for measured confounders and no survival advantage from BCSI remained. Rows 3 though 6 contain IV estimates using the area BCSI rate instrument at various grouping levels. These estimates are consistently negative and often statistically significant, implying survival disadvantages from BCSI. For example, the 8-group, two-year-survival model suggests that increasing the BCSI rate by 2 percentage points among marginal patients (e.g., from 7.3 to 9.3 using the BCSI rates in Table 1) would have decreased two-year survival for that group by 1 percentage point. The estimates remained fairly consistent across the 2-, 4-, and 8-group specifications, but the 12-group estimates fell in absolute value. Violation of IV assumptions may be the source of inconsistency for the 12-group estimates. The Hausman test statistics (not shown) were all statistically insignificant for the 2-, 4-, and 8-group models. The Hausman test statistics were greatest for the 12-group specifications and the statistic was statistically significant in the four-year survival model, which suggests that our sample size limited the number of groups we can use without introducing spurious relationships that violate IV assumptions.

Instrumental variable estimates were consistently negative across instrument specifications. Estimates obtained using the area BCSI rate instrument were generally larger in absolute value and significantly different from zero more often than estimates found using the radiation distance instrument (rows 7–10). The IV estimates found specifying both instruments (rows 11–14) appear as averages between the estimates found using the individual instruments. These results suggest that the sets of marginal patients differ between instruments. To investigate this in a post hoc analysis we divided our sample into patients younger than age 65 and patients older than age 64 and reestimated equation (1) for both subsamples using both

Table 3:   Instrumental Variable and OLS Estimates of the Effectiveness of Breast Conserving Surgery Plus Irradiation on Survival Relative to Mastectomy

Row	Analysis Method	Instruments Specified	Number of Groups per Instrument	Instrument F-Statistic	After Diagnosis, Effect of BCSI on Patient Survival:			
					1 year	2 years	3 years	4 years
1	Unadjusted OLS	None	na	na	0.01	0.02	0.04**	0.06**
2	Adjusted OLS[a]	None	na	na	−0.001	−0.003	−0.0001	0.007
3	Instrumental Variable Estimates[a]	BCSI Rate	2	8.57***	−0.32	−0.68	−0.57	−0.51
4			4	5.19***	−0.37**	−0.54**	−0.45	−0.65*
5			8	3.43***	−0.33**	−0.50**	−0.46*	−0.52*
6			12	3.00***	−0.23**	−0.41**	−0.33	−0.11
7		Radiation Distance	2	21.79***	−0.21*	−0.12	−0.33	−0.23
8			4	7.52***	−0.14	−0.22	−0.39	−0.38
9			8	3.30***	−0.14	−0.19	−0.35	−0.28
10			12	2.94***	−0.05	−0.14	−0.33	−0.40*
11		BCSI Rate and Radiation Distance	2	13.08***	−0.24**	−0.25	−0.38*	−0.30
12			4	4.99***	−0.24**	−0.32*	−0.39*	−0.45*
13			8	2.76***	−0.24**	−0.31**	−0.34*	−0.27
14			12	2.74***	−0.12*	−0.23**	−0.30**	−0.15

***, **, *Statistically significant at .99, .95, and .90 confidence, respectively.

[a]Models also specified binary variables for age groups (<50, 50–64, 65–69, 70–74, 75–79, 80–84, 85+), tumor sizes (<2 cm, 2–5 cm, 5+ cm), positive lymph node involvement, tumor grade groups (1, 2, 3, 4, 9–unknown), tumor location groups (nipple, central portion, upper-inner quad, lower-inner quad, upper-outer quad, lower-outer quad, axillary tail, overlapping lesion, not stated), Charlson comorbidity index (0, 1, 2, 3+), residence zip code poverty percentage (#7, 7–10, 10–13, 13–20, >20), distance from residence to nearest hospital (#2.83, 2.83–9, 9–15, >15), payer (Medicaid, Medicare, Blue Cross/Blue Shield, other private, other government, self-pay), and year of diagnosis (1989, 1990, 1991, 1992, 1993, 1994).

instruments at the 4-group level. Differences in marginal patients were confirmed as we found that radiation distance had a relatively greater impact on surgery choice for patients younger than age 65; patient and area BCSI rates had a relatively greater impact on patients older than age 64.

DISCUSSION

The inability of the 1991 NIH Consensus Statement to substantially increase the rate of BCSI has often been attributed to a lack of provider knowledge. Alternatively, BCSI may not have been more widely and consistently adopted because many providers may not have shared NIH's beliefs that RCT evidence could be extrapolated to ESBC patients with more extensive disease. If it is unclear whether RCT results can be extrapolated to ESBC patients with more extensive disease, justifying a BCSI rate increase needs estimates of the survival impacts for the patients whose surgery choices would be affected by a rate increase. We used IV methods to estimate the survival effects of BCSI relative to MAS for stage II ESBC patients from Iowa whose surgery decisions varied with the practice styles of local providers and access to radiation treatment. We argued that these estimates are naturally representative of patients whose surgery choice would be affected by an increase in BCSI rates.

Our IV estimates are based on the surgery variation revealed by our instruments and their consistency is conditional on the assumption that our instruments are not systematically related to unmeasured confounding variables. We supported this assumption in Table 1 by comparing the distribution of measured confounders across patients grouped by surgery choice and their instrument values. Patients with measured clinical characteristics associated with less-extensive disease were more likely to obtain BCSI. If unmeasured confounders share the same relationship with surgery choice as measured confounders (patients with unmeasured confounder values related to lower survival risk are more likely to receive BCSI), then OLS estimates of BCSI survival adjusted for measured confounders will remain biased in favor of BCSI. When patients were grouped by instruments, the distributions of tumor size and comorbidities appeared similar across groups. Patient age and tumor grade varied with the instruments. We directly controlled for these and other measured confounders in our IV analysis, but our results are conditional on the assumption these differences are not symptomatic of differences in unmeasured confounders

across instrument groups. Fuller assumption validation requires additional data. We next estimated a surgery choice model that revealed that stage II patients with more extensive disease were less likely to obtain BCSI, and patients grouped by our instruments had statistically significant different BCSI rates.

Using the BCSI variation revealed by our instruments in IV analysis, we found consistently negative one-, two-, three-, and four-year average survival impacts of BCSI relative to MAS for marginal stage II ESBC patients—patients whose surgery choices were affected by variation in our instruments. If marginal patients are closely aligned to the set of patients whose surgery choices would have been changed with an increase in BCSI rates, these results suggest that an increase would have decreased survival rates. Previous authors (e.g., Harris and Remler 1998) justify this alignment by assuming that the set of patients whose treatment choices are affected by instruments and those patients whose treatment choices change with an increase in treatment rates are similar—those for whom optimal treatment is the least certain. While our IV estimates were consistently negative, their magnitude and significance varied by instrument. This suggests that patients defined as "marginal" vary with the instrument and that the set of patients potentially affected by an increase in BCSI rates is broader than the marginal patients from a single instrument. As a result, our combined instrument estimates may be more closely aligned to the average survival effects that would have resulted from an effort to increase BCSI rates.

Readers should be careful in generalizing our estimates beyond the stage II ESBC patients in Iowa during 1989–1994 whose treatment choices were associated with our instruments. If treatment effects are heterogeneous across ESBC patients, it would be risky to generalize our estimates to stage II patients in other states or other time periods that had markedly different BCSI rates because characteristics of the patients defined as marginal will differ. In addition, our results cannot distinguish whether BCSI was inappropriate treatment for the marginal patients in our sample or whether the BCSI techniques that were used *during this period in Iowa* were inappropriate for the marginal patients. This is important to distinguish because as BCSI rates increased during the 1990s, it is likely that the set of marginal patients in the late 1990s would have had more extensive disease and, *all else being equal*, would be more likely to have survival disadvantages from BCSI. However, increased use of BCSI also may have been associated with better skills, which may have reduced the survival advantages of MAS in the later time period for these patients.

This research demonstrates the value of IV methods to help policymakers assess the impacts associated with proposed treatment rate changes. To use IV estimates for this purpose, policymakers must assume that the patients whose treatment choices are affected by instruments are similar to the patients whose treatment choices are affected by rate changes. If this assumption holds here, our results suggest that during the early 1990s in Iowa, BCSI may have been overused, and efforts to increase BCSI rates at this time in Iowa may have been inappropriate. Application of IV methods to more current data from across the country with larger samples that allows for subgroup analysis will provide policymakers with the information required to make more current and specific assessments.

ACKNOWLEDGMENTS

REFERENCES

Adams, E. K. 2001. "Going beyond Costs When Evaluating Surgical Options for Women Newly Diagnosed with Breast Cancer." *Medical Care* 39 (11): 1143–5.

Angrist, J. D. 2001. "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice." *Journal of Business and Economic Statistics* 19 (1): 2–16.

Angrist, J. D., and W. N. Evans. 1998. "Children and Their Parent's Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review* 88 (3): 450–77.

Angrist, J. D., G. W. Imbens, and D. B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–54.

Answini, G. A., W. L. Woodard, H. J. Norton, and R. L. White. 2001. "Breast Conservation: Trends in Major Metropolitan Area Compared with Surrounding Rural Counties." *American Surgeon* 67 (10): 994–8.

Arriagada, R., M. G. Le, F. Rochard, and G. Contesso. 1996. "Conservative Treatment versus Mastectomy in Early Breast Cancer: Patterns of Failure with 15 Years of Follow-Up Data." *Journal of Clinical Oncology* 14 (5): 1558–64.

Begg, C. B., L. D. Cramer, W. J. Hoskins, and M. F. Brennan. 1998. "Impact of Hospital Volume on Operative Mortality for Major Cancer Surgery." *Journal of the American Medical Association* 280 (20): 1747–51.

Benedict, S., D. J. Cole, L. Baron, and P. Baron. 2001. "Factors Influencing Choice between Mastectomy and Lumpectomy for Women in the Carolinas." *Journal of Surgical Oncology* 76 (1): 6–12.

Brooks, J. M., M. McClellan, and H. Wong. 2000. "The Marginal Benefits of Invasive Treatment for Acute Myocardial Infarction: Does Insurance Coverage Matter?" *Inquiry* 37 (1): 75–90.

Brooks, J. M., E. Chrischilles, S. Scott, J. Ritho, and S. Chen-Hardee. 2000. "Information Gained from Linking SEER Cancer Registry Data to State-Level Hospital Discharge Abstracts." *Medical Care* 38 (11): 1131–40.

Charlson, M. E., P. Pompei, K. L. Ales, and C. R. MacKenzie. 1987. "A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation." *Journal of Chronic Diseases* 40 (5): 373–83.

Chow, G. C. 1960. "Tests of Equality between Sets of Coefficients in Two Linear Models." *Econometrica* 28 (3): 591–605.

Du, X., J. L. Freeman, J. L. Warren, A. B. Nattinger, D. Zhang, and J. S. Goodwin. 2000. "Accuracy and Completeness of Medicare Claims Data for Surgical Treatment of Breast Cancer." *Medical Care* 38 (7): 719–27.

Earle, C. C., J. S. Tsai, R. D. Gelber, M. C. Weinstein, P. J. Neumann, and J. C. Weeks. 2001. "Effectiveness of Chemotherapy for Advanced Lung Cancer in the Elderly: Instrumental Variable and Propensity Analysis." *Journal of Clinical Oncology* 19 (4): 1064–70.

Fisher, B., S. Anderson, C. K. Redmond, N. Wolmark, D. L. Wickerham, and W. M. Cronin. 1995. "Reanalysis and Results after 12 Years of Follow-Up in a Randomized Clinical Trial Comparing Total Mastectomy with Lumpectomy with or without Irradiation in the Treatment of Breast Cancer." *New England Journal of Medicine* 333 (22): 1456–61.

Foster, R. S., M. E. Farwell, and M. C. Costanza. 1995. "Breast-Conserving Surgery for Breast Cancer: Patterns of Care in a Geographic Region and Estimation of Potential Applicability." *Annals of Surgical Oncology* 2 (3): 275–80.

Guadagnoli, E., J. C. Weeks, C. L. Shapiro, J. H. Gurwitz, C. Borbas, and S. B. Soumerai. 1998. "Use of Breast-Conserving Surgery for Treatment of Stage I and Stage II Breast Cancer." *Journal of Clinical Oncology* 16 (1): 101–6.

Harris, K. M., and D. K. Remler. 1998. "Who is the Marginal Patient? Understanding Instrumental Variables Estimates of Treatment Effects." *Health Services Research* 33 (5): 1337–60.

Hausman, J. 1983. "Specification and Estimation of Simultaneous Equations Models." In *Handbook of Econometrics*, edited by Z. Griliches, and M. Intriligator, Amsterdam: North Holland pp. 391–448.

Hadley, J., and J. M. Mitchell. 1997. "Breast Cancer Treatment Choice and Mastectomy Length of Stay: A Comparison of HMO and Other Privately Insured Woman." *Inquiry* 34 (4): 288–301.

Imbens, G. W., and J. D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467–75.

Iscoe, N. A., V. Goel, K. Wu, G. Fehringer, E. J. Holowaty, and C. D. Naylor. 1994. "Variation in Breast Cancer Surgery in Ontario." *Canadian Medical Association Journal* 150 (3): 345–52.

Jacobson, J. A., D. N. Danforth, K. H. Cowan, T. D. Angelo, S. M. Steinberg, L. Pierce, M. E. Lippman, A. S. Lichter, E. Glatstein, and P. Okunieff. 1995. "Ten-Year Results of a Comparison of Conservation with Mastectomy in the Treatment of Stage I and II Breast Cancer." *New England Journal of Medicine* 332 (14): 907–11.

Keating, N. L., J. C. Weeks, M. B. Landrum, C. Borbas, and E. Guadagnoli. 2001. "Discussion of Treatment Options for Early-Stage Breast Cancer: Effect of Provider Specialty on Type of Surgery and Satisfaction." *Medical Care* 39 (7): 681–91.

Keleman, J. J., T. Poulton, M. T. Swartz, and I. Jatoi. 2001. "Surgical Treatment of Early-Stage Breast Cancer in the Department of Defense Healthcare System." *Journal of the American College Surgeons* 192 (3): 293–7.

Mandelblatt, J. S., C. D. Berg, N. J. Meropol, S. B. Edge, K. Gold, Y. Hwang, and J. Hadley. 2001. "Measuring and Predicting Surgeons' Practice Styles for Breast Cancer Treatment in Older Women." *Medical Care* 39 (3): 228–42.

McClellan, M., B. J. McNeil, and J. P. Newhouse. 1994. "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality: Analysis Using Instrumental Variables." *Journal of the American Medical Association* 272 (11): 859–66.

McClellan, M., and J. P. Newhouse. 1997. "The Marginal Cost-Effectiveness of Medical Technology: A Panel Instrumental-Variables Approach." *Journal of Econometrics* 77 (1): 39–64.

———. 2000. "Introduction: Overview of Special Supplement Issue." *Health Services Research* 35 (5, part 2): 1061–9.

Morrow, M., J. White, J. Moughan, J. Owen, T. Pajack, J. Sylvester, J. F. Wilson, and D. Winchester. 2001. "Factors Predicting the Use of Breast-Conserving Therapy in Stage I and II Breast Carcinoma." *Journal of Clinical Oncology* 19 (8): 2254–62.

National Institutes of Health Consensus Conference. 1991. "Treatment of Early-Stage Breast Cancer." *Journal of the American Medical Association* 265 (3): 391–5.

Nold, R. J., R. L. Beamer, S. D. Helmer, and M. F. McBoyle. 2000. "Factors Influencing a Woman's Choice to Undergo Breast-Conserving Surgery versus Modified Radical Mastectomy." *American Journal of Surgery* 180 (6): 413–8.

Riley, G. F., A. L. Potosky, C. N. Klabunde, J. L. Warren, and R. Ballard-Barbash. 1999. "Stage at Diagnosis and Treatment Patterns among Older Women with Breast Cancer: An HMO and Fee-for-Service Comparison." *Journal of the American Medical Association* 281 (8): 720–6.

Sainsbury, R., L. Rider, A. Smith, and A. MacAdam. 1995. "Does It Matter Where You Live? Treatment Variation for Breast Cancer in Yorkshire." *British Journal of Cancer* 71 (6): 1275–8.

Stafford, D., R. Szczys, R. Becker, J. Anderson, and S. Bushfield. 1998. "How Breast Cancer Treatment Decisions Are Made by Women in North Dakota." *American Journal of Surgery* 176 (6): 515–9.

van Dongen, J. A., H. Bartelink, I. S. Fentiman, T. Lerut, F. Mignolet, G. Olthuis, E. van der Schueren, R. Sylvester, D. Tong, J. Winter, and K. van Zijl. 1992. "Factors Influencing Local Relapse and Survival Results of Salvage Treatment after Breast-Conserving Therapy in Operable Breast Cancer: European Organization for Research and Treatment of Cancer Trial 10801, Breast Conservation Compared with Mastectomy in TNM Stage I and II Breast Cancer." *European Journal of Cancer* 28A (415): 801–5.

van Dongen, J. A., A. C. Voogd, I. S. Fentiman, C. Legrand, R. J. Sylvester, D. Tong, E. vander Schueren, P. A. Helle, K. van Zijl, and H. Bartelink. 2000. "Long-Term Results of a Randomized Trial Comparing Breast-Conserving Therapy with Mastectomy: European Organization for Research and Treatment of Cancer 10801 Trial." *Journal of the National Cancer Institute* 92 (14): 1143–50.

Veronesi, U., A. Banfi, B. Salvadori, A. Luini, R. Saccozzi, R. Zucali, E. Marubini, M. Del Vecchio, P. Boracchi, S. Marchini, M. Merson, V. Sacchini, G. Riboldi, and G. Santoro. 1990. "Breast Conservation Is the Treatment of Choice in Small Breast Cancer: Long-Term Results of a Randomized Trial." *European Journal of Cancer* 26 (6): 668–70.