Comparative Effectiveness of Oxaliplatin Versus 5-flourouricil in Older Adults

An Instrumental Variable Analysis

Christina DeFilippo Mack,[a,b] *M. Alan Brookhart,*[a] *Robert J. Glynn,*[c] *Anne Marie Meyer,*[d] *William R. Carpenter,*[d,e] *Robert S. Sandler,*[f] *and Til Stürmer*[a]

Background: Oxaliplatin was rapidly adopted for treatment of stage III colon cancer after FDA approval in November 2004, thus providing an opportunity to use calendar time as an instrumental variable in nonexperimental comparative effectiveness research.

Assuming instrument validity, instrumental variable analyses account for unmeasured confounding and are particularly valuable in sub-populations of unresolved effectiveness, such as older individuals.

Methods: We examined stage III colon cancer patients ages 65+ years initiating chemotherapy between 2003 and 2008 using US population-based cancer registry data linked with Medicare claims (N = 3,660). Risk differences for all-cause mortality were derived from Kaplan–Meier survival curves. We examined instrumental variable strength and compared risk differences with propensity score estimates.

Results: Calendar time greatly affected oxaliplatin receipt. The calendar time instrument compared patients treated from January 2003 through September 2004 (N = 1,449) with those treated from March 2005 through May 2007 (N = 1,432), resulting in 54% compliance. The 1-, 2-, and 3-year local average treatment effect of the risk differences per 100 patients in the "compliers" (95% confidence intervals) were −4.6 (−8.2, −0.44), −6.3 (−12, −0.16), and −9.2 (−15, −2.5), respectively. Corresponding propensity score-matched results were −1.9 (−4.0, 0.2), −3.4 (−6.2, −0.05), and −4.3 (−7.5, −0.96).

Conclusions: Instrumental variable and propensity score analyses both indicate better survival among patients treated with oxaliplatin. As these results are based on different populations and assumptions, the instrumental variable analysis adds to evidence of oxaliplatin's effectiveness in older adults, who bear the greatest burden of colon cancer yet were underrepresented in clinical trials. In nonexperimental comparative effectiveness research of rapidly emerging therapies, the potential to use calendar time as an instrumental variable is worth consideration.

(*Epidemiology* 2015;26: 690–699)

Colon cancer is the third most common cancer with nearly 100,000 incident cases per year in the US, and the third most deadly with over 50,000 deaths per year.[1] Based on superior efficacy demonstrated in the 2003 MOSAIC trial[2,3] and subsequent FDA approval in 2004, the drug oxaliplatin was rapidly adopted as treatment for stage III colon cancer as part of multiagent chemotherapy. This innovation was broadly adopted and has become the stan-

dard of care despite gaps in knowledge regarding drug benefits for specific subpopulations, including older patients. Indeed, the MOSAIC trial did not include patients over the age of 75 years and had a median patient age of 60 years,[2] much younger than the median age of colon cancer diagnosis, 72 years.[1] Subsequent studies have produced inconclusive or contradictory results among diverse subpopulations and this drug's effectiveness in the older population remains unknown, prompting a need for additional research and robust methodologies.[4–7]

Nonexperimental comparative effectiveness research can address gaps and unanswered questions that commonly accompany new treatments. Comparative effectiveness research often relies on secondary data such as administrative or linked databases that have many strengths, including large, diverse study populations, and long periods of follow-up.[8] A common weakness of these data is that key confounders such as functional status, frailty and disease severity are often missing or are recorded with differential fastidiousness between providers.[9] These factors can bias analyses that assume no unmeasured confounding, including those relying on propensity score adjustment.[10,11]

When a suitable instrument exists, instrumental variable methods can address these challenges. An instrumental variable is an observed variable that predicts treatment based on the context surrounding treatment receipt, which effectively pseudo-randomizes the population. This replaces the assumption of no unmeasured confounding with assumptions around the instrument's association with treatment and its relationship with confounders and the outcome.

The case of rapid innovation diffusion and oxaliplatin effectiveness among older cancer patients provides an excellent platform for employing instrumental variable methods. New treatments often experience rapid dissemination upon arrival to the market, and treatment decisions may be driven by external factors rather than patient-centric characteristics. In this context, calendar time proximal to FDA approval or other drug lifecycle events may serve as a good instrument for treatment receipt.

OBJECTIVE

The objective of this study is to demonstrate the utility of calendar time as an effective instrumental variable, and in doing so examine the effectiveness of an innovative cancer treatment, oxaliplatin, in older patients with colon cancer at the time of the drug's initial adoption. We focused on the period before and during oxaliplatin's dissemination, with attention on FDA approval for stage III colon cancer as a pivotal time point. We compared instrumental variable estimates for the effect in the compliers with a diverse set of propensity score-adjusted estimates to contrast the effect observed in different subsets of the study population.

METHODS

Data Source

Patients were drawn from Surveillance, Epidemiology and End Results (SEER)-Medicare linked data.[12,13] The cohort included individuals ages 65 years and older from 12 US states who were diagnosed with primary stage III colon cancer between 2003 and 2007, with follow-up through April 2010. To be eligible, patients had to receive surgical resection within 90 days of diagnosis, survive longer than 30 days, and initiate either oxaliplatin or 5-fluorouracil (5-FU)/capecitabine without oxaliplatin within 110 days of surgery and 120 days of diagnosis (SEER does not collect diagnosis day; rationale described elsewhere[12,14]). Patients were excluded if they received radiation, were diagnosed at autopsy, or had health maintenance organization coverage or incomplete Medicare claims during the 12 months pre- and postdiagnosis (or until death).

Instrumental Variable

Instrumental variable methods replace the assumption of no unmeasured confounding with the following conditions: the instrument (1) is associated with treatment receipt, (2) is related to the outcome only through treatment (exclusion restriction), and (3) does not share any causes with the outcome (independence restriction). An additional assumption of monotonicity allows us to estimate the average treatment effect in the "compliers."[15,16]

To specify the instrumental variable, we defined a binary measure of calendar time based on the month and year of first treatment receipt anchored around oxaliplatin's FDA approval for stage III colon cancer and subsequent rapid adoption. In our study, monotonicity holds if there are no patients who would have received oxaliplatin if diagnosed with cancer and treated before the date of FDA approval, yet would not have received it if treated after approval ("defier"). Analogously, "compliers" are patients whose initial treatment is determined by the calendar time in which their treatment was received: 5-FU if diagnosed and treated before FDA approval and oxaliplatin if treated after approval.

We identified the "optimal" instrumental variable measure through evaluation of two criteria: (1) the compliance percentage (i.e., strength of the instrument's effect on treatment receipt)[17] and (2) the shortest overall timespan, to reduce the potential for violating instrumental variable assumptions. To achieve the latter, we excluded patients treated several years after FDA approval by truncating cohort enrollment while using all follow-up time. This minimized the effects of calendar time on survival, directly or indirectly, through changes in diagnostic paradigms or improvements in care unrelated to chemotherapy (assumption 2). In addition, we tested the effect of excluding those treated in the months immediately surrounding FDA approval, when information dissemination and drug

access may have been ambiguous. We examined the instrument in relation to instrumental variable assumptions using measured covariates, falsification tests, expert knowledge, and time trends.[16,18] These time trends were examined relative to the inception of oxaliplatin-based treatment options as well as other possible changes in colon cancer care that may have created an association between time and mortality. Calendar-time intervals for the instrument were decided before examination of effect estimates.

Exposure and Outcome

First treatment receipt was defined as the date of first 5-FU/capecitabine claim with no oxaliplatin claim within 30 days (unexposed) or the date of first oxaliplatin claim with or without the presence of 5-FU/capecitabine (exposed). We ignored oxaliplatin claims that occurred greater than 30 days after 5-FU/capecitabine receipt, because we are only interested in the first treatment received. However, because late receipt of oxaliplatin may suggest that these individuals had a recurrence or were too sick to receive oxaliplatin initially, we performed a sensitivity analysis excluding these patients (n = 46). All-cause mortality information was based on date of death according to Medicare via the US Social Security Administration.[19]

Analysis

We derived risk of mortality from Kaplan–Meier survival curves and estimated 1-, 2-, and 3-year mortality risk differences (RD). Instrumental variable RDs were scaled by the compliance percentage to estimate the average treatment effect among "compliers" (main analysis).[17] The Balke–Pearl method[20] was used to place bounds around the instrumental variable point estimate for the average treatment effect in the population. We generated covariate-adjusted instrumental variable estimates of the RDs using inverse-probability of treatment-weighted Kaplan–Meier methods to estimate the effect of the IV on the outcome and then scaled this by the strength of the instrument (full detail: eFigure 4; http://links.lww.com/EDE/A946). For comparison, we generated propensity score estimates using both matching and inverse-probability of treatment-weighted methods to account for all measured confounders.[21,22] These methods are based on different assumptions and resulted in appreciably different patient populations.

Previous literature guided identification of potential confounders: age, sex, race, tumor grade, tumor substage at diagnosis (IIIA–IIIC), urban/rural status, socioeconomic status, physician organizational affiliation with a National Cancer Institute Cooperative Group,[23,24] and 13 prevalent comorbidities from the Charlson comorbidity index (listed in Table 1).[25] Follow-up began on the date of first treatment receipt for 5-FU and 1 day after for oxaliplatin (based on observed median oxaliplatin start time after 5-FU to avoid systematic differences between exposure groups). Propensity score models used the full study population and we

performed sensitivity analyses in the reduced instrumental variable cohort to evaluate selection differences that may have been induced based on instrumental variable exclusions. All analyses were conducted using SAS version 9.2. (SAS Institute, Cary, NC). The University of North Carolina Office of Human Research Ethics (IRB study number 12–0139) approved this study.

RESULTS

Calendar time greatly affected treatment receipt (Figure 1). The "optimal" 2-level calendar time instrument compared patients treated from January 2003 to September 2004 (n = 1,449) with those treated from March 2005 to May 2007 (n = 1,432). This excluded patients treated during two separate time periods in the oxaliplatin lifecycle: (1) an interim period, when oxaliplatin use first exceeded 40% until it was used in the majority of patients, indicating a change in the standard of care and (2) 2.5 years after FDA approval, when the market for innovation was likely functionally saturated and calendar time unlikely to further dictate treatment choice. This instrumental variable definition produced oxaliplatin treatment rates of 11% and 65% in the early versus late arms of the instrument and thus yielded 54% compliance.

Measured patient characteristics were well balanced between instrument levels (Table 1), thereby supporting the assumption that the instrumental variable is unrelated to patient risk factors for the outcome. Prevalence differences for covariates stratified by the instrumental variable compared with treatment assignment were greatly attenuated, which further indicates a strong instrument that may be independent of unmeasured covariates (assumptions 1, 3).[17] The instrumental variable particularly improved balance for age, substage, and cooperative group. For example, the difference in the percentage of 65–69 year old patients was 13.6 between treatment groups but only 3.7 across the instrumental variable.

In the reduced instrumental variable population, we observed 11.7 deaths per person-year in the 5-FU group and 8.7 in the oxaliplatin group (Table 2). The 1-, 2-, and 3-year mortality risks were 11%, 22%, and 29% in the 5-FU treatment group and 7.9%, 17%, and 23% in the oxaliplatin group, respectively. The instrumental variable estimate of the 3-year RD per 100 patients for all-cause mortality was −9.2 (−15, −2.5), which suggests that for every 100 compliant patients treated with oxaliplatin, nine additional patients survived to 3 years compared with those treated with 5-FU or capecitabine alone. One- and 2-year instrumental variable RDs per 100 patients were −4.6 (−8.2, −0.44) and −6.3 (−12, −0.16), with bounds of (−16, 30), (−20, 26), and (−23, 24) for 1–3 year RDs, respectively. Covariate-adjusted instrumental variable estimates were virtually identical (Table 3).

TABLE 1. Characteristics of Stage III Colon Cancer Patients by Treatment Received (n = 3,660) and Calendar Time Instrument (n = 2,881)

| | Treatment | | Calendar Time[a] (Instrument) | |
| | 5-FU (n = 2,095) | Oxaliplatin (n = 1,565) | Pre-FDA Approval (n = 1,449) | Post-FDA Approval (n = 1,432) |
Patient Characteristics	N (%)	N (%)	N (%)	N (%)
Treatment				
5-FU			1,291 (90)	503 (35)
Oxaliplatin			158 (11)	929 (65)
Year treated[b]				
2003–2004	1,397 (67)	228 (15)	1,449 (100)	0 (0)
2005	312 (15)	391 (25)	0 (0)	573 (40)
2006	192 (9)	400 (26)	0 (0)	592 (41)
2007–2008	194 (9)	546 (35)	0 (0)	267 (19)
Race				
Caucasian American	1,822 (87)	1,363 (87)	1,266 (87)	1,249 (87)
African American	137 (7)	107 (7)	99 (7)	97 (7)
Other	135 (6)	95 (6)	83 (6)	86 (6)
Age				
65–69	348 (17)	473 (30)	292 (20)	341 (24)
70–74	564 (27)	508 (33)	434 (30)	410 (29)
75–79	607 (29)	428 (28)	408 (28)	405 (28)
80–84	435 (21)	143 (9)	241 (17)	226 (16)
85+	141 (7)	13 (1)	74 (5)	50 (4)
Mean (SD)	75.7 (5.6)	73.0 (4.8)	74.8 (5.5)	74.3 (5.4)
Sex				
Female	1,155 (55)	849 (54)	783 (54)	789 (55)
Male	940 (45)	716 (46)	666 (46)	643 (45)
Urbanity				
Metro	1,741 (83)	1,338 (86)	1,220 (84)	1,196 (84)
Urban	298 (14)	206 (13)	199 (14)	204 (14)
Rural	56 (3)	21 (1)	30 (2)	32 (2)
Census median income:[c] mean (SD)	49.9 (22)	54.1 (25)	50.7 (23)	52.2 (24)
Socioeconomic status: state buy-in (yes)	241 (0)	173 (1)	168 (0)	160 (1)
Census percent nonhigh school graduation: mean (SD)	18.7 (13)	17.3 (13)	18.4 (13)	18.0 (12)
Cooperative group[d]				
Yes	926 (44)	743 (48)	651 (45)	668 (47)
No	1,162 (56)	808 (52)	791 (55)	757 (53)
Substage				
A	255 (12)	154 (10)	159 (11)	153 (11)
B	1,233 (59)	810 (52)	832 (57)	780 (55)
C	607 (29)	600 (38)	458 (32)	498 (35)
Grade				
Well differentiated; differentiated, NOS	101 (5)	82 (5)	80 (6)	64 (5)
Moderately/intermediately differentiated	1,339 (64)	959 (61)	925 (64)	886 (62)
Poorly or undifferentiated; anaplastic	613 (29)	485 (31)	411 (28)	451 (32)
Cell type not determined, stated or applicable	42 (2)	39 (3)	33 (2)	31 (2)

(Continued)

TABLE 1. (Continued)

Patient Comorbidities	Treatment		Calendar Time[a] (Instrument)	
	5-FU (n = 2,095)	Oxaliplatin (n = 1,565)	Pre-FDA Approval (n = 1,449)	Post-FDA Approval (n = 1,432)
	N (%)	N (%)	N (%)	N (%)
CHF	160 (8)	72 (5)	99 (7)	82 (6)
MI	31 (2)	15 (1)	24 (2)	15 (1)
Old myocardial infarction	47 (2)	31 (2)	32 (2)	31 (2)
COPD	212 (10)	166 (11)	130 (9)	157 (11)
CVD	112 (5)	59 (4)	76 (5)	60 (4)
Diabetes (with or without sequelae)	479 (23)	339 (22)	311 (22)	321 (22)
PVD	93 (4)	50 (3)	59 (4)	51 (4)
Ulcer	30 (1)	15 (1)	21 (1)	12 (1)
Chronic renal failure	45 (2)	19 (1)	23 (2)	21 (2)
RA	241 (11)	173 (10)	168 (11)	160 (10)

Due to SEER-Medicare confidentiality requirements, treatment years are combined and presence of paralysis, dementia, and cirrhosis are not shown.

[a]Instrumental variable definition is focused around FDA action in November 2004. Patients treated from October 2004 through February 2005 and after June 2007 were excluded (n = 779, or 21% of N = 3,660, excluded).

[b]Year treated was based on date qualifying treatment (5-FU without oxaliplatin within 30 days or oxaliplatin) was received. Median follow-up time in years for the 5-FU and oxaliplatin groups was 4.5 and 3.3 (full cohort, N = 3,660) and 4.7 and 3.7 years (IV cohort, N = 2,881), respectively.

[c]Median household income in 1,000 US Dollars, based on 2000 data.

[d]Physician organizational affiliation with a National Cancer Institute (NCI) Cooperative Group containing a colon cancer research portfolio.[23] Includes the American College of Surgeons Oncology Group (ACOSOG), Eastern Cooperative Oncology Group (ECOG), Cancer and Leukemia Group B (CALGB), Southwest Oncology Group (SWOG), and the National Surgical Adjuvant Breast and Bowel Project (NSABP).

CHF, congestive heart failure; MI, myocardial infarction; COPD, chronic obstructive pulmonary disease; CVD, cerebrovascular disease; PVD, peripheral vascular disease; RA, rheumatoid arthritis.

Instrumental variable results were consistent with propensity score comparators, although slightly less precise and further from the null; all suggested a protective effect (Figures 2 and 3). Propensity score-matched 1–3 year RDs were −1.9 (−4.0, 0.21), −3.4 (−6.2, −0.05), and −4.3 (−7.5, −0.96); propensity score results agreed regardless of adjustment method.

Sensitivity Analyses

We conducted four sensitivity analyses as follows: (1) Two additional sets of propensity score-matched[26] estimates which included calendar time alone and as an interaction term with each measured covariate,[21] (2) propensity score-matched and weighted analyses in reduced IV population rather than full study population, (3) instrument defined by a cutpoint rather than interim exclusion, and (4) removed 46 patients initiating oxaliplatin more than after 30 days of chemotherapy inception. None of these results substantially differed from our main findings; absolute changes in RDs per 100 patients were ≤2.2 (eAppendix A; http://links.lww.com/EDE/A947).

DISCUSSION

In this study, we used methods within nonexperimental data to study comparative effectiveness of oxaliplatin-based therapy versus the prior standard of care for older stage III colon cancer patients. The instrumental variable analyses controls, at least in part, for unmeasured confounding, which is especially problematic for studies of older patients who are disproportionately affected by the disease yet poorly represented in randomized controlled trials. Across diverse analytic approaches, we consistently found that oxaliplatin reduces all-cause mortality compared with 5-FU alone in this population. Results were generally consistent with effects observed in the younger population of the MOSAIC randomized controlled trial, which reported 2- and 3-year RDs (95% confidence interval) per 100 patients of −2 (−5, 2) and −3 (−7, 1; derived from Kaplan–Meier survival curves; A. de Gramont, written communication, December 2012). The stability of these effectiveness estimates within different populations and in the presence of differing assumptions provides important information about oxaliplatin effectiveness in older adults, which could aid in decision making among patients, providers, and policymakers.

Results across the various methods generalize to different populations and therefore apply to potentially dissimilar subgroups.[15] The instrumental variable estimates the average treatment effect in 54% of the IV cohort (the "compliers"). The inverse-probability of treatment-weighted-adjusted propensity score also estimates the average treatment effect, which applies to all patients indicated for treatment. Propensity score matching estimates the treatment effect in the oxaliplatin-treated patients who were successfully matched to 5-FU-treated patients. An examination of patient characteristic distributions between the key populations[22] showed general similarity, although the compliers were slightly
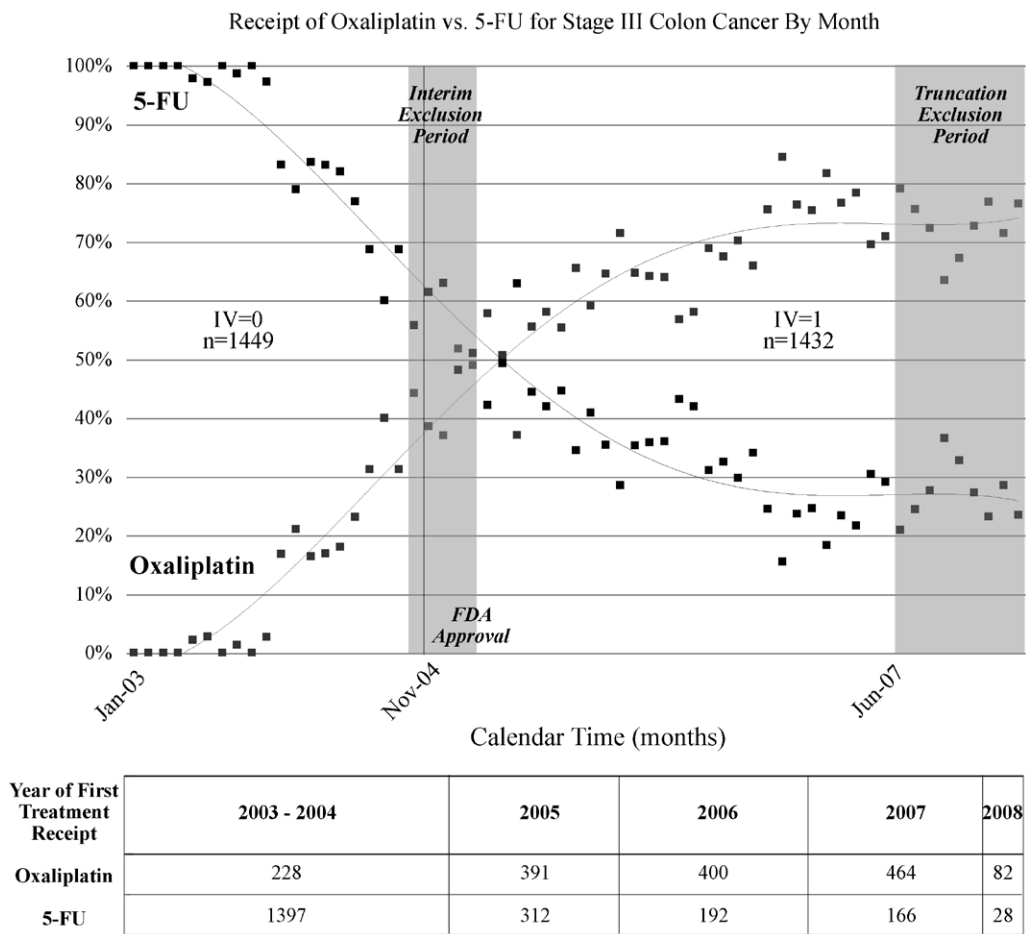
FIGURE 1. Dissemination of oxaliplatin: receipt of oxaliplatin versus 5-FU for stage III colon cancer by month and definition of calendar time instrumental variable (N = 3,660). Points indicate the percentage of patients in each month receiving oxaliplatin or 5-FU. *Grey shading* Excluded patients due to interim period (October 2004–February 2005) and the truncation period of June 2007 and later. For illustrative purposes, diffusion patterns for each treatment are fitted with a fourth-order polynomial trendline. The intersection point of lines is not statistically meaningful in terms of dissemination activity. Due to SEER-Medicare confidentiality requirements, treatment years 2003 and 2004 are combined.

TABLE 2. Person Time, All-cause Mortality Rate and 1-, 2-, and 3-year Mortality Risk Comparison Among Treatment Groups in the Full and Instrumental Variable Populations

Treatment Group	Person Time (Years)[a]	Events Within 3-years	Mortality Rate (Per 100 Person-years)	1-Year Risk	2-Year Risk	3-Year Risk
Full population (n = 3,660)						
5-FU	5,234	602	11.5	11.3%	21.3%	29.0%
Oxaliplatin (with 5-FU/capecitabine)	4,008	354	8.8	7.9%	18.1%	24.7%
Instrumental variable population (n = 2,881)						
5-FU	4,495	526	11.7	11.3%	21.6%	29.4%
Oxaliplatin (with 5-FU/capecitabine)	2,839	248	8.7	7.9%	16.9%	23.0%

[a]Patients administratively censored 3 years after treatment initiation.

younger than the overall population and more likely to have tumors identified as substage B than substage C (eAppendix B; http://links.lww.com/EDE/A948). These differences make clinical sense, as older age and the accompanying

reduced functional status are major reasons that a patient would receive 5-FU in the time period when oxaliplatin would be predicted by the instrument (i.e., noncompliance). That oxaliplatin-treated patients would be more likely to

TABLE 3. Comparison of Risk Differences Using Instrumental Variable and Propensity Score Analyses

Model	1-Year Risk Difference[a]			2-Year Risk Difference[a]			3-Year Risk Difference[a]		
	RD	95% CI	NNT	RD	95% CI	NNT	RD	95% CI	NNT
Unadjusted	−3.41	−5.71, −1.36	29.3	−4.69	−7.74, −1.31	21.3	−6.39	−9.77, −3.05	15.6
IV estimator[b]	−4.58	−8.17, −0.44	21.8	−6.30	−11.89, −0.16	15.9	−9.24	−14.67, −2.47	10.8
Adjusted IV estimator[b]	−5.01	−8.88, −0.80	20.0	−6.45	−11.31, 0.06	15.5	−9.42	−14.93, −2.04	10.6
PS-matched[c]	−1.90	−4.01, 0.21	52.6	−3.37	−6.23, −0.05	29.4	−4.25	−7.53, −0.96	23.3
PS-IPTW[c]	−1.89	−4.23, 0.54	52.9	−3.87	−6.41, −1.52	25.8	−3.69	−6.78, −0.71	27.1

[a]Risks per 100 patients taken from Kaplan–Meier survival curve; RDs calculated assuming binomial distributions and independent observations, using a standard error of $\sqrt{(SE_{unexposed}^2 + SE_{exposed}^2)}$. Nonparametric bootstraps with 250 resamples were used to compute CIs. NNT calculated by 1/RD.
[b]IV estimator is run in reduced IV population (n = 2,881).
[c]Estimates shown here are based on the full population (n = 3,660) and are not adjusted for the instrument, calendar time; see sensitivity analyses (eAppendix A) for additional propensity score estimates.
RD, risk difference per 100 patients; CI, confidence interval; NNT, number needed to treat; IV, instrumental variable; PS, propensity score; IPTW, inverse-probability of treatment-weight.
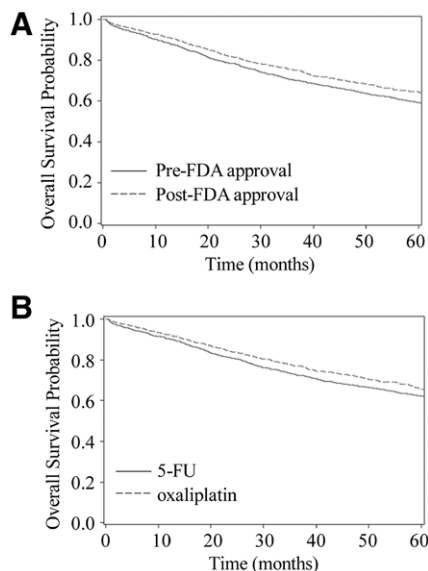

FIGURE 2. Probability of overall survival (A) by calendar time instrumental variable (N = 2,881) and (B) with 5-FU versus oxaliplatin in matched propensity score analysis (N = 2,732). A, Patient assignment to instrumental variable category is based on month treatment was first received. January 2003–September 2004 (pre-FDA approval, referent) is compared with March 2004–May 2007 (post-FDA approval). B, Propensity score-matched analysis adjusts for age, sex, race, tumor grade, tumor substage at diagnosis (IIIA–IIIC), urban/rural status, socioeconomic status measured using number of months of state buy-in and census percentage of high school graduates, physician organizational affiliation with a NCI Cooperative Group, and comorbidities. NCI indicates National Cancer Institute.

have a more aggressive substage (substage C) is also consistent with clinical practice.

While estimates generally agreed, the magnitude of oxaliplatin effectiveness in this older population cannot be

confirmed due to the inability to empirically test assumptions required by these analyses. Validity of propensity score estimates could be compromised because of unmeasured confounding, which could occur through lack of data on patient frailty, and disease severity not captured by staging. Although this would be less pronounced due to comparison with an active treatment, the increased toxicity and cost associated with oxaliplatin may contribute to unmeasured confounding, particularly in this older population. If time is an instrument rather than a confounder, propensity score comparators that adjust for calendar year may be more biased in the presence of unmeasured confounding.[27,28]

Instrumental variable assumptions could also be violated. While we were able to verify that calendar time's relation to oxaliplatin receipt was strong during the study years, time could affect mortality in ways other than through treatment and, as in all instrumental variable analyses, the exclusion restriction is not empirically verifiable. We mitigated this possibility by carefully considering the means through which calendar time might affect mortality other than through treatment changes and truncated the cohort accordingly. Stage migration or improvements in surgical techniques and other nonchemotherapeutic treatments could also create an association between time and mortality. However, American Joint Committee on Cancer tumor staging guidelines[29] and oncologist interviews suggested that this association is unlikely between January 2003 and May 2007. Five-year relative survival in colon cancer patients improved from 1975 to 2004 according to more inclusive national statistics[30]; this trend flattens after 2000, however, and is not strong enough to explain our observed results over a narrow time interval. The percentage of stage III patients who received no chemotherapy did not change from 2003 to 2007 for most patients, although those over age 80 years became slightly less likely to receive adjuvant chemotherapy in later study years. It is possible that
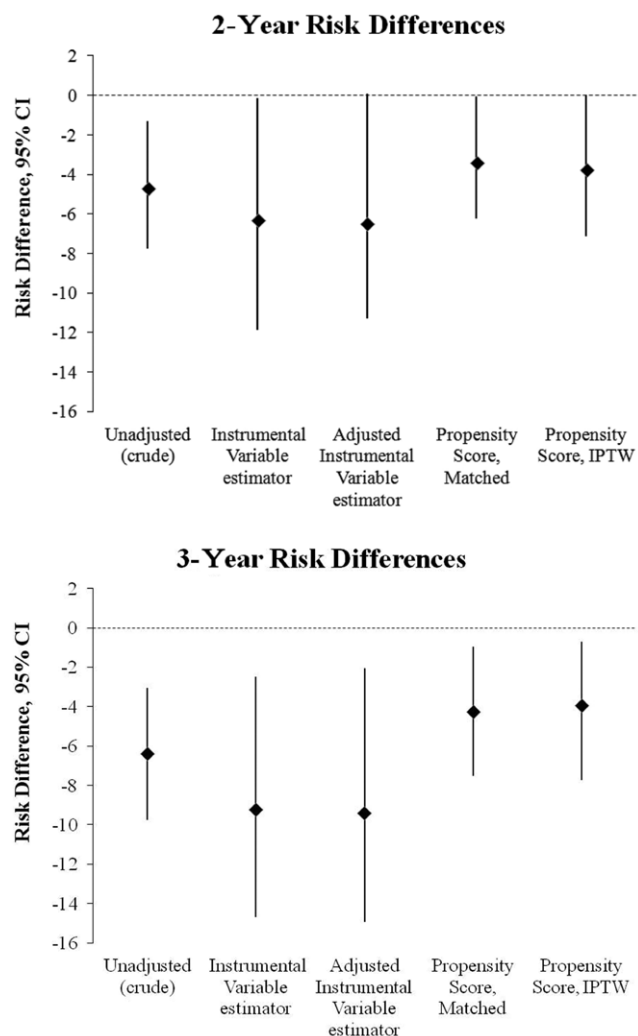
FIGURE 3. Comparison of risk differences estimating comparative effectiveness of oxaliplatin versus 5-fluorouracil. Estimates of RD are based on risks per 100 patients taken from Kaplan–Meier survival curves. The instrumental variable estimator is scaled by a compliance percentage of 54%. Adjusted estimates account for the variables presented in Table 1 except year. IPTW = inverse probability of treatment weight.

in the oldest age groups, sicker patients who may have been included in the study in early years may not have qualified in later years, thereby indirectly associating calendar time with decreased mortality. The improved balance of measured confounders by instrumental variable level shown in Table 1 supports that the independence restriction, also not verifiable, may be upheld. The following falsification tests did not find violations of the instrumental variable conditions[16,31]: (1) comparison of instrumental variable results with conventional estimates (Table 3) and (2) comparison of instrumental variable strength and subgroup estimates for potentially modifying factors, as identified in prior investigations[21] (eAppendix C; http://links.lww.com/EDE/A949).

The width of the bounds for the instrumental variable suggests that our estimate relies heavily on the assumption of monotonicity; we did not compute equivalent bounds for unmeasured confounding for the propensity score estimates. Monotonicity is clinically reasonable in this setting, as it is improbable that a patient would receive oxaliplatin off-label before FDA approval while (holding all other considerations constant) an identical patient would receive only 5-FU after FDA approval. Adverse event reports that would preclude an early oxaliplatin patient from receiving oxaliplatin in a later calendar month were unlikely to be an issue over this time period. While it is possible that physician-observed neuropathy may eventually have deterred an oncologist from prescribing oxaliplatin to a diabetic patient in later years, it is very unlikely that such patients would be preferentially treated with oxaliplatin in 2003 and 2004.[32]

The overall consistency of results between these methods suggests oxaliplatin is effective among older adults, a finding which is robust to the absence of measured or unmeasured confounding. While chance is a plausible explanation for the slight differences between the instrumental variable and propensity score point estimates, the divergence in direction from the unadjusted estimate may provide insight into differing abilities to control for measured versus unmeasured confounding. Instrumental variable effect estimates may reflect control of unmeasured confounders that increase both mortality and probability of receiving oxaliplatin. For example, tumor pathology information regarding extent and aggressiveness of the cancer is not entirely captured by the relatively coarse grade and substage variables. This unmeasured disease severity information would be accounted for in slightly stronger instrumental variable estimates, while biasing propensity score estimates toward the null (in this example). Removal of tumor substage and grade from the propensity score model moved the results closer to the null, thereby supporting this theory. Treatment effect heterogeneity is another possibility, as oxaliplatin may be more effective in the medium-health "compliers" than the healthier always-treated or not contraindicated patients, who may have survived up to 3 years regardless of treatment.

Some limitations of claims data, in general and specific to SEER-Medicare, apply to this study.[12] Medicare has an estimated 75% sensitivity for 5-FU,[33] and therefore a proportion of the referent group may have been missed. Comorbidity assessed through claims may be underestimated in this population, as older age is associated with less aggressive treatment and coding for a number of diseases.[34] We used Kaplan–Meier survival curves to allow us to use all available patients and to avoid conditioning on follow-up; however, exclusion of a small number of patients with incomplete claims or health maintenance organization coverage after diagnosis could introduce selection bias. This exclusion was

necessary to obtain treatment data. The group of patients that were excluded from this analysis due to lack of chemotherapy receipt may have included patients that died before their planned treatment was initiated. If treatment delays were more common in earlier years, this could have biased our instrumental variable estimates; however, the overall number of patients who would be eligible to fall into this category (a subset of n = 293) is small. We did not use calendar time as a continuous instrumental variable, as there is not a practical interpretation of compliers in this setting. Finally, we cannot exclude chance as an alternative explanation for our findings.

In the presence of emerging therapies, consideration should be given to treatment variability by calendar time and the contribution of dissemination patterns to treatment receipt.[21] When clinical uptake of treatment occurs quickly over a narrow time interval, calendar time is a plausible instrument that can account for bias due to unmeasured confounding. In this case, clinical paradigms outside of the new treatment should be closely investigated to ensure stability during the study time period. Over longer timeframes, calendar time may be a confounder, and is often treated as such by default.[35] The utility of calendar time as an instrumental variable has been shown by Cain et al.,[36] Johnston et al.,[37] and Shetty et al.[38] when, similar to this setting, trends in medication use create a natural experiment that can be used to strengthen clinical evidence. In studies, where sample size is small, the inefficiencies of instrumental variable methods may result in confidence intervals that are uninformative; however, in pharmacoepidemiology studies using large databases, the primary concern is often bias rather than imprecision.

Nonexperimental research is necessary to answer questions of treatment effectiveness but requires careful attention to methods and examination of potential biases. This study exemplifies this while answering a high priority clinical question on the comparative effectiveness of oxaliplatin versus 5-FU in older patients, who are the most affected by colon cancer yet were underrepresented in clinical trials. The methods presented address different biases and assumptions which cannot be directly quantified, such as the effect of unmeasured confounding or the exact relationship of a natural instrument with exposures and outcomes. The presentation of a consistent set of results based on different methods and assumptions builds needed confidence regarding the benefit of oxaliplatin in older adults in a real-world population.

ACKNOWLEDGMENTS

REFERENCES

1. Altekruse SF, Kosary CL, Krapcho M, et al. *SEER cancer statistics review, 1975–2007*. Bethesda, MD: National Cancer Institute; 2010. Available at: http://seer.cancer.gov/csr/1975_2007/. Accessed October 29, 2012.
2. André T, Boni C, Mounedji-Boudiaf L, et al.; Multicenter International Study of Oxaliplatin/5-Fluorouracil/Leucovorin in the Adjuvant Treatment of Colon Cancer (MOSAIC) Investigators. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *N Engl J Med*. 2004;350:2343–2351.
3. De Gramont A, Banzi M, Navarro M, et al. *Oxaliplatin/5-FU/LV in adjuvant colon cancer: results of the international randomized mosaic trial [abstract]*. Presented at 2003 ASCO Annual Meeting, Chicago, IL, May 31–June 3, 2003. *Proc Am Soc Clin Oncol*. 22: 2003 (abstr 1015).
4. Sanoff HK, Carpenter WR, Martin CF, et al. Comparative effectiveness of oxaliplatin vs non-oxaliplatin-containing adjuvant chemotherapy for stage III colon cancer. *J Natl Cancer Inst*. 2012;104:211–227.
5. Tournigand C, André T, Bonnetain F, et al. Adjuvant therapy with fluorouracil and oxaliplatin in stage II and elderly patients (between ages 70 and 75 years) with colon cancer: subgroup analyses of the Multicenter International Study of Oxaliplatin, Fluorouracil, and Leucovorin in the Adjuvant Treatment of Colon Cancer trial. *J Clin Oncol*. 2012;30:3353–3360.
6. McCleary NAJ, Meyerhardt J, Green E, et al. Impact of older age on the efficacy of newer adjuvant therapies in >12,500 patients with stage II/III colon cancer: findings from the ACCENT database [abstract]. *J Clin Oncol*. 2009;27(15s):4010.
7. Marshall JL, de Gramont A, Köhne CH, et al. Accomplishments in 2008 in the adjuvant treatment of colon cancer. *Gastrointest Cancer Res*. 2009;3(suppl):S2–S7.
8. Stürmer T, Jonsson Funk M, Poole C, Brookhart MA. Nonexperimental comparative effectiveness research using linked healthcare databases. *Epidemiology*. 2011;22:298–301.
9. Meyer AM, Carpenter WR, Abernethy AP, Stürmer T, Kosorok MR. Data for cancer comparative effectiveness research: past, present, and future potential. *Cancer*. 2012;118:5186–5197.
10. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70:41–55.
11. Glynn RJ, Schneeweiss S, Stürmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. *Basic Clin Pharmacol Toxicol*. 2006;98:253–259.
12. Warren JL, Klabunde CN, Schrag D, et al. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care*. 2002;40(Suppl):IV-3–18.
13. National Cancer Institute. *About the SEER Program*. Bethesda, MD: National Cancer Institute; 2012. Available at: http://seer.cancer.gov/about/. Accessed September 28, 2012.
14. Mack CD, Carpenter W, Meyer AM, Sanoff H, Stürmer T. Racial disparities in receipt and comparative effectiveness of oxaliplatin for stage III colon cancer in older adults. *Cancer*. 2012;118:2925–2934.
15. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. *Pharmacoepidemiol Drug Saf*. 2010;19:537–554.
16. Swanson SA, Hernán MA. Commentary: how to report instrumental variable analyses (suggestions welcome). *Epidemiology*. 2013;24:370–374.
17. Brookhart MA, Schneeweiss S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. *Int J Biostat*. 2007;3:Article 14.
18. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology*. 2006;17:360–372.
19. Bach PB, Guadagnoli E, Schrag D, et al. Patient demographic and socioeconomic characteristics in the SEER-Medicare database applications and limitations. *Med Care*. 2002;40(Suppl):IV-19–25.
20. Balke A, Pearl J. Bounds on treatment effects from studies with imperfect compliance. *J Am Stat Assoc*. 1997;92:1171–1176.
21. Mack CD, Glynn RJ, Brookhart MA, et al. Calendar time-specific propensity scores and comparative effectiveness research for stage III colon cancer chemotherapy. *Pharmacoepidemiol Drug Saf*. 2013;22:810–818.
22. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc*. 1996;81:444–455.

23. Carpenter WR, Reeder-Hayes K, Bainbridge J, et al. The role of organizational relationships and research networks in the diffusion of breast cancer treatment innovation. *Med Care*. 2011;49:172–179.

24. McFall SL, Warnecke RB, Kaluzny AD, Ford L. Practice setting and physician influences on judgments of colon cancer treatment by community physicians. *Health Serv Res*. 1996;31:5–19.

25. Klabunde CN, Legler JM, Warren JL, Baldwin LM, Schrag D. A refined comorbidity measurement algorithm for claims-based studies of breast, prostate, colorectal, and lung cancer patients. *Ann Epidemiol*. 2007;17:584–590.

26. Parsons LS. *Reducing bias in a propensity score matched-pair sample using greedy matching techniques*. Ovation Research Group, Seattle, WA. Paper 214–226.

27. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. *Am J Epidemiol*. 2006;163:1149–1156.

28. D'Agostino RB Jr, D'Agostino RB Sr. Estimating treatment effects using observational data. *JAMA*. 2007;297:314–316.

29. American Joint Committee on Cancer. *Summary of changes: understanding the changes from the sixth to the seventh edition of the AJCC cancer staging manual*. Chicago, IL: American Joint Committee on Cancer; 2011. Available at: http://www.cancerstaging.org/products/comparison.html. Accessed November 10, 2012.

30. Howlader N, Noone AM, Krapcho et al. SEER cancer statistics review, 1975–2011, National Cancer Institute. Bethesda, MD. Available at: http://seer.cancer.gov/csr/1975_2011/, based on November 2013 SEER data submission, posted to the SEER web site, April 2014. Accessed December 21, 2014.

31. Glymour MM, Tchetgen Tchetgen EJ, Robins JM. Credible Mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. *Am J Epidemiol*. 2012;175:332–339.

32. Sanofi-Aventis US, LLC ELOXATIN: *Highlights of prescribing information*. Bridgewater, NJ: Sanofi-Aventis U.S. LLC; 2009. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/label/2009/021492s011,021759s009lbl.pdf. Accessed November 20, 2011.

33. Warren JL, Harlan LC, Fahey A, et al. Utility of the SEER-Medicare data to identify chemotherapy use. *Med Care*. 2002;40(8 Suppl):IV–55.

34. Glynn RJ, Knight EL, Levin R, Avorn J. Paradoxical relations of drug treatment with mortality in older persons. *Epidemiology*. 2001;12:682–689.

35. Rassen JA, Doherty M, Huang W, et al. *Pharmacoepidemiology toolbox*. Boston, MA. Available at: http://www.hdpharmacoepi.org. Accessed July 25, 2012.

36. Cain LE, Cole SR, Greenland S, et al. Effect of highly active antiretroviral therapy on incident AIDS using calendar period as an instrumental variable. *Am J Epidemiol*. 2009;169:1124–1132.

37. Johnston KM, Gustafson P, Levy AR, Grootendorst P. Use of instrumental variables in the analysis of generalized linear models in the presence of unmeasured confounding with applications to epidemiological research. *Stat Med*. 2008;27:1539–1556.

38. Shetty KD, Vogt WB, Bhattacharya J. Hormone replacement therapy and cardiovascular health in the United States. *Med Care*. 2009;47:600–606.