Published in final edited form as:
*Pharmacoepidemiol Drug Saf*. 2013 August ; 22(8): . doi:10.1002/pds.3386.

# Calendar Time-Specific Propensity Scores and Comparative Effectiveness Research for Stage III Colon Cancer Chemotherapy

**Christina DeFilippo Mack**[1], **Robert J. Glynn**[2], **M. Alan Brookhart**[1], **William R. Carpenter**[3,4], **Anne Marie Meyer**[4], **Robert S. Sandler**[5], and **Til Stürmer**[1]

[1]Department of Epidemiology, UNC Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[2]Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard, Medical School, Boston, MA, USA

[3]Department of Health Policy and Management, Gillings School of Global Public Health UNC, Chapel Hill, NC, USA

[4]Lineberger Comprehensive Cancer Center and Cecil G. Sheps Center for Health Services Research, University of North Carolina, Chapel Hill, NC, USA

[5]Division of Gastroenterology and Hepatology, University of North Carolina, Chapel Hill, NC, USA

## Abstract

**Purpose—**Nonexperimental studies of treatment effectiveness provide an important complement to randomized trials by including heterogeneous populations. Propensity scores (PS) are common in these studies, but may not adequately capture changes in channeling experienced by innovative treatments. We use calendar time-specific (CTS) PSs to examine the effect of oxaliplatin during dissemination from off-label to widespread use.

**Methods—**Stage III colon cancer patients aged 65+ initiating chemotherapy between 2003–06 were examined using cancer registry data linked with Medicare claims. Two PS approaches for receipt of oxaliplatin vs. 5-flourouricil were constructed using logistic models with key

Corresponding author: Christina Mack, University of North Carolina at Chapel Hill, 2101 McGavran-Greenberg Hall, CB #7435, Chapel Hill, NC 27599-7435, T:574-276-7958, F:919-966-2089, cdmack@unc.edu.

**Funding Disclosures/Conflict of Interest:**
Christina Mack: No perceived conflicts of interest. Fellowship, Merck; NIH grants R01 AG023178 and T32KD07634; Consultant to Roche and AHRQ through Quintiles | Outcome.
Robert J. Glynn: Consulting to Merck; Data & Safety Monitoring Board for Tryton Medical trial; Independent academic statistician for Novartis trial
Alan Brookhart: Received research support from Amgen and has served as a scientific advisor for Amgen, Rockwell Medical, and Pfizer (honoraria declined, donated, or paid to institution) and has consulted for Crimson Health, DaVita Clinical Research, the Foundation for the National Institutes of Health, and World Health Information Consultants.
William R. Carpenter: No perceived conflicts of interest. Supported by NCI Grant 5R01CA124402, and the Integrated Cancer Information and Surveillance System (ICISS), UNC Lineberger Comprehensive Cancer Center, through the University Cancer Research Fund through the State of North Carolina.
Anne-Marie Meyer: No perceived conflicts of interest. Supported by NCI Grant 5R01CA124402, and the Integrated Cancer Information and Surveillance System (ICISS), UNC Lineberger Comprehensive Cancer Center, through the University Cancer Research Fund through the State of North Carolina.
Robert Sandler: No perceived conflicts of interest. R01DK094738, R01CA136887, R01CA44684, P30 Dk034987, U01CA 93326, the Crohn's and Colitis Foundation of American and the Leona and Helmsley Charitable Trust.
Til Stürmer: R01 AG023178 from the National Institute on Aging at the NIH; UNC-DEcIDE center from AHRQ; salary support from Center for Pharmacoepidemiology, Department of Epidemiology and unrestricted research and other grants from pharmaceutical companies (GSK, Merck, Sanofi) to UNC.

components of age, sex, substage, grade, census level income, and comorbidities: 1) a conventional, year-adjusted PS and 2) a CTS PS constructed and matched separately within 1-year intervals, then combined. We compared PS-matched hazard ratios (HR) for mortality using Cox models.

**Results—**Oxaliplatin use increased significantly; 8%(n=86) of patients received it in the first time period vs. 52%(n=386) in the last. Channeling by comorbidities, income, and age appeared to change over time. The CTS PS improved covariate balance within calendar time strata and yielded an attenuated estimated benefit of oxaliplatin (HR=0.75) compared with the conventional PS (HR=0.69).

**Conclusion—**In settings where prescribing patterns have changed and calendar time acts as a confounder, a CTS PS can characterize changes in treatment choices and estimating separate PSs within specific calendar time periods may result in enhanced confounding control. To increase validity of CER, researchers should carefully consider drug lifecycles and effects of innovative treatment dissemination over time.

## Introduction

Propensity scores (PS) are widely used to control confounding in comparative studies of medical products. A PS is an estimate of the probability that a patient receives one treatment over another, given characteristics of the patient and his/her condition at the time the treatment decision is made.[1,2] PSs are routinely estimated as averages of the effect of patient characteristics on treatment choice over multiple study years. However, many drugs have a dynamic lifecycle, experiencing changes in prescribing based on events and dissemination. A patient characteristic that was once associated with treatment selection may become less relevant over time, or vice versa.

The key assumptions underlying PS methods are that all confounders are accurately measured and the model of treatment receipt given confounders is correct. If calendar time (as a proxy for other changes) is a confounder and prescribing patterns are dynamic, calendar time and its relations to other confounders must be correctly modeled. To our knowledge, few studies consider specific lifecycle events for the drug of interest and incorporate potentially heterogeneous effects of time into PS analyses.[3] This may violate the assumption that the score correctly reflects the underlying propensity for treatment given the confounders.[1] A direct comparison of PS approaches for handling calendar time in dynamic settings has not been performed.

Oxaliplatin, an innovative chemotherapeutic, is a drug that saw striking uptake among stage III colon cancer patients over a short time period, from off-label to widespread use. In this setting, we construct and examine a calendar time-specific (CTS) PS within policy-based time periods of a study cohort to understand possible validity benefits of accounting for changes in confounding by indication over calendar time based on specific patient characteristics. The CTS PS allows the effect of each covariate on the propensity for treatment receipt to be non-uniform over time, taking into account changes in channeling (used here to denote any degree of confounding by indication) relevant to a specific multi-year cohort. Examination of the CTS PS provides insight into prescribing variations and barriers to treatment receipt across calendar years.

## Methods

We examined the CTS PS in the context of a CER study of oxaliplatin versus 5-fluorouracil (5-FU) and all-cause mortality in patients with stage III colon cancer, focusing on the early years of oxaliplatin adoption. Based on efficacy results from the MOSAIC trial[4,5] and

subsequent FDA approval in November 2004,[6] FOLFOX, defined as the addition of oxaliplatin to 5-FU and folinic acid, replaced 5-FU monotherapy as the standard of care.

Patients were drawn from the Surveillance, Epidemiology and End Results (SEER)-Medicare linked data (described elsewhere)[7,8] and included those diagnosed with stage III colon cancer between 2003 and 2005, with follow-up through 31 December 2006. All patients were traditional Medicare subscribers aged 65+ who received curative surgery and initiated either oxaliplatin or 5-FU without oxaliplatin within 90 days of surgical resection.

We defined three study time periods, each one year in duration beginning in May 2003, the month MOSAIC trial results were released (Figure 1). The second time period encompasses FDA approval and spans the six months pre-approval as well as immediate post-approval dissemination. Patients were categorized into time periods based on their receipt date of 5-FU (referent) or oxaliplatin (exposed). Directed Acyclic Graph methodology[9] and expert knowledge were used to identify potential confounders of age, sex, race, tumor grade, tumor substage at diagnosis (IIIA–IIIC), urban/rural status, income and 11 individual comorbidities.[10]

Multivariable logistic regression was used to estimate PSs. A conventional PS, the primary comparator, was estimated across all years, adjusting for time period. The CTS PS required a separate model for each period to estimate the time-specific propensity of treatment receipt per covariate. To understand if relations between patient characteristics and treatment preference changed for individual covariates over time using the CTS PS, changes in odds ratios (OR) and 95% confidence intervals (CI) for receipt of oxaliplatin were compared graphically over each successive time period (from the CTS PS models) and for the full cohort (from the conventional PS model).

Greedy 5-to-1 digit matching[11] was used for covariate adjustment. CTS PS matching was performed within each time period; matched pairs were pooled to create a full study cohort. For the conventional PS, patients were matched across all three years and the matched cohort was used for both overall and year-specific estimates. For the latter, matching was ignored (broken). To evaluate confounding control, we examined covariate balance between matched cohorts using the absolute difference in percentage by time period for each covariate, with focus on strong confounders. We also report the cumulative balance for each cohort, irrespective of the strength of covariate association with the outcome, and the percent of oxaliplatin-exposed patients retained.[12]

We compared effectiveness of oxaliplatin vs. 5-FU for prevention of all-cause mortality by constructing Cox models to estimate hazard ratios (HR) and 95% confidence intervals (CI), in an intent-to-treat approach. Cox models used an origin of 90 days after surgery to landmark the analysis.[13] This origin avoids immortal time bias and systematic differences in exposure time by treatment group, and excludes a small proportion of patients that likely died due to surgical complications. An HR for the full study and separate HRs per time period were generated using 4 PSs to adjust for confounding using matching: 1) the calendar time-specific PS, 2) the conventionally estimated PS that adjusts for calendar time (primary comparator), 3) a conventionally estimated PS with full interaction terms between calendar time period and each covariate and 4) a conventional PS with no adjustment for calendar time. For comparison, we also fitted unadjusted and adjusted Cox proportional hazards outcome models. We compared HRs graphically and with percent and absolute differences. The UNC Office of Human Research Ethics (Study number 12-0139) approved this study.

## Results

Oxaliplatin treatment increased significantly, from 8% (n=86) in 2003–2004 to 52% (n=386) in 2005–2006. Overall, 71% of patients received 5-FU and 29% received oxaliplatin. Exposure group characteristics were similar over the 3 years of the study with the exception of diabetes, which increased in prevalence (Table 1).

### Channeling

Comorbidities, race, income, urbanity, and age appeared to experience changes in channeling over time. Results for selected covariates are shown to illustrate patterns (Figure 2). The adjusted relative odds of oxaliplatin receipt increased over the later two time periods for patients with congestive heart failure (CHF) and decreased across all time periods for those with diabetes. Residence in a high income census area appeared to increase patients' odds of receiving oxaliplatin, particularly prior to FDA approval. Older age was consistently associated with decreased odds of oxaliplatin receipt, and those above 79 became slightly less likely over time to receive it. The effects of tumor grade, COPD and sex on channeling were relatively constant (adjusted OR = 1.0, 0.8/0.9, and 1.0/1.1).

### Cohort balance

The CTS PS retained 77% of oxaliplatin-exposed patients (100%, 91%, 59% for each time period, ascending) and the conventional PS retained 79%. Patients were excluded if a suitable match could not be found. In the first time period, the CTS PS was able to include all oxaliplatin-exposed patients because there were relatively few patients receiving oxaliplatin. In later years, the percentage of patients receiving oxaliplatin increased, and therefore oxaliplatin-exposed patients were excluded due to lack of available unexposed (5-FU) matches.

Variables were generally more balanced when using the CTS PS. For example, the balance between 5-FU and oxaliplatin for census income>$60,000 in 2003–2004 was 30% vs. 37% (balance=6.9) for the conventional PS cohort, compared with 36% vs. 37% for the CTS PS (balance=1.2) (Appendix A. Supplemental Table 1). Because imbalance of the strongest risk factors for mortality leads to more problematic confounding, we focused on the balance of tumor substage (HR=1.8 and 3.6 for substage IIIB and IIIC vs. IIIA), older age (HR=1.2 for ages 70–74, 1.3 for ages 75–79, and 1.8 for 80+, compared with ages 65–69), undifferentiated/unknown tumor grade (HR=1.3), lower income (HR=1.2), CHF (HR=1.5), COPD (HR=1.4) and diabetes (HR=1.2). Age, COPD, CHF, and income were more balanced across the study years for the CTS PS cohort, although substage and diabetes were less balanced. The balance statistics showing the distribution of the imbalance for each covariate (Figure 3) show lower means in each time period for the CTS PS. The entire distribution of balance (quartiles, means, medians) was lower for the CTS PS, showing less imbalance in this cohort, with the exception of 2005–2006. In this time period, the CTS PS cohort had a slightly higher median and maximum than the conventional PS (1.8 vs. 1.5; 8.3 vs. 6.5, respectively); however, it was lower in all other statistics (mean: 1.2 vs. 1.6; $25^{th}$/$75^{th}$ percentiles: 0.3/1.6 vs. 0.4/2.5). Cumulative imbalances of 125 vs. 158 for the CTS compared with the conventional PS further suggest improved balance in the calendar time-specific cohort.

### Comparison of hazard ratio estimates

For the CER analysis, patients receiving oxaliplatin (n=810) were compared with those on a non-oxaliplatin 5-FU regimen (n=1990). Over a median follow-up of 2.65 years, 860 patients (31%) died. The crude mortality rate was 83/1,000 person-years in patients receiving oxaliplatin and 129/1,000 person-years in patients receiving 5-FU. There was a

22% change in HR (HR=0.69 vs. 0.75) between the conventional and CTS PS-adjusted estimates (Table 2). Precision between the two methods was similar. The full interaction PS (HR=0.73) generated results similar to the CTS PS in both magnitude and precision. When comparing within time periods, CTS PS-estimated HRs differed more than conventional PS estimates and were closer to the null, with the exception of 2004–2005, in which both estimates moved farther from the null (HR=0.64 and 0.65, respectively) (Figure 4).

## Discussion

Set in the context of a CER examination of a new chemotherapeutic agent, this study examined a novel approach to propensity score estimation which addresses changes in intervention adoption over time. During oxaliplatin's first three years of rapid adoption for stage III colon cancer, the CTS PS method proved to more adequately address subtle changes in factors associated with treatment selection, under the assumption that calendar time is a confounder or a proxy for confounders, than the commonly applied PS model that assumes uniform effects of patient factors over multiple study years. The CTS PS characterized changes in treatment choices and likely resulted in enhanced confounding control.

Our research expands upon work by Seeger et al[14,15] and Rassen et al[16] by comparing estimates, quantifying bias and defining time periods using policy-related timepoints.[17] In research examining statins and MI, Seeger et al accounted for changes in statin use by estimating PSs and matching within half-year time blocks, then stratifying Cox models by these blocks. They found that allowing flexibility in PS estimation showed changes in drug use over time, but did not compare effect estimates with a conventional PS. Rassen et al assessed the performance of an overall PS within subgroups based on patient characteristics and found larger effect differences for small subgroups and few exposed outcomes. Schneeweiss et al[18] noted the particular importance of CER evidence immediately after FDA approval and highlighted the methodological challenge of bias due to confounding by indication of new medications. They describe a sequential cohort study which proposes matching within quarterly[19] or monthly blocks, but do not explicitly compare alternative approaches.

Our results suggest oncologists may initially have been reluctant to give oxaliplatin to patients with comorbidities such as CHF as they learned of this new drug's effect in patients with characteristics that may have been excluded from clinical trials. A decline in oxaliplatin use was observed in patients with diabetes, suggesting that physicians may have observed neurotoxicity[20,21] that shaped their decision-making in subsequent chemotherapy decisions for these already susceptible patients.[22,23] Similarly, consistent channeling away from older patients may suggest that age-correlated unmeasured variables such as frailty or age discrimination[24] were found to be increasingly relevant over time. Although all patients were covered exclusively by Medicare, higher income areas had increased access to the innovation. This difference dissipated after FDA approval but did not disappear.

The CTS PS produced results closer to the MOSAIC Randomized Controlled Trial (RCT) (HR=0.80, 0.65–0.97),[25] than did the conventional PS and recent observational study findings.[26] Because the true effect among older individuals is unknown, it is not possible to empirically evaluate bias reduction.[12,27] Increased validity of CTS PS estimates can be inferred, however, as there was evidence of changes in confounding by indication over time by individual patient characteristics. Because the CTS PS led to better overall balance of observed covariates within calendar periods and we assume here that calendar time is a confounder, the increased balance reduces confounding bias, at minimum within calendar time period and possibly for the overall estimate. The closeness of the full interaction and

CTS PS estimates was expected, as both account for changes in channeling; however, the CTS PS provided the benefit of easier-to-interpret evidence of calendar time's impact on treatment choice.

The differences observed between the year-specific estimates of the CTS PS could be attributed to several factors. In this population, there is evidence of modification by race,[28] and the proportion of African Americans on oxaliplatin ranges from 6.5% to 8.1% per year. Changes in population mix over time may lead to varied treatment effect estimates in the presence of effect modification. Additionally, unmeasured confounding may change over calendar years and estimating CTS PSs within time periods may allow better identification and management of observations treated contrary to prediction.[29]

Overall matching of exposed patients was similar for both the CTS and conventional PS groups, suggesting practical feasibility of this approach. In this examination, the CTS PS-matched cohort demonstrated greater balance within years than did the conventional PS, as measured by the statistical distribution of individual covariate imbalance, by cumulative absolute difference, and for most, but not all, of the strong confounders. An alternative approach would be to match on the conventional PS within calendar year, which could affect balance comparisons even in the presence of a misspecified model. We matched on the conventional PS across the full cohort after adjusting for calendar year under the assumption that this implementation is most common.

As new drugs are continually entering the market, comparative effectiveness questions of new versus old treatments will continue to arise. PS use has increased exponentially in the last 2 decades[12] and although variables are often liberally included in PS models, vigilance is required in selection.[30,31] Even in settings of dynamic prescribing, calendar time is often not considered in CER specific to its possible role as an instrument, confounder, or modifier of covariate effects on treatment choice. Our assumption is that most researchers would see calendar year as a confounder and thus include it in the PS model; for example, the high-dimensional propensity score algorithm documentation lists year of treatment as the common example of a predefined variable.[32] This is a reasonable assumption in many cases, as time can serve as a proxy for changes in tumor staging, improvements in surgical technique, increases in provider experience and the use of additional effective treatments that affect common CER outcomes such as mortality and disease recurrence. These factors are unmeasured in these data, as in many claims databases, and controlling for calendar time will limit their potential to confound treatment effects. CTS PSs should be considered in dynamic settings, when calendar time acts as a confounder between the exposure and the outcome and is also a potential modifier due to non-homogenous prescribing or treatment determinants. However, if time is not a confounder but instead an instrument for treatment receipt,[33,34,35] it should not be included in the propensity score model regardless of changes in channeling of the treatment over time. Doing so would result in inflation of the variance and bias if residual confounding is present.[36] As in other settings, the important distinction between a variable (here: calendar time) acting as a confounder or as an instrument cannot be confirmed based on observed data.

Thoughtful consideration of time periods is warranted and ideal choice of calendar time periods is not tested. In this specific example, it was most appropriate to anchor the time periods around efficacy results, when off-label use commenced (period 1), the months surrounding FDA approval (period 2), and the post-approval year when wide dissemination had likely occurred (period 3). In general, drug lifecycle milestones or policy events (e.g., safety warnings) are good candidates for choosing calendar time periods. Providing buffer time around events of interest is needed to allow for dissemination of new information. There is also a need to have reasonable numbers of subjects in each of the treatment groups

within time intervals to support estimating a multivariable PS in that interval. In some settings, including pharmacovigilance, allowing for similar numbers of patients per time period may be preferred because such a strategy may be optimal to compare the effect of the treatment over time periods. In any setting, the ability to divide the full cohort into granular time periods depends on the number of events in each cohort and time period.

While RCTs remain the gold standard for assessing an intervention's effectiveness, they are not always feasible, and their findings often have limited generalizability to the broader population. Comparative effectiveness research using non-experimental data addresses many of the limitations of RCTs, and method development to strengthen CER is critical. Oxaliplatin provides a good practical example for investigating a CTS PS in a non-experimental CER setting. The nature of chemotherapeutic use among oncologists is particularly dynamic; due to rapid disease progression and high mortality, chemotherapies are commonly used off-label, quickly approved for new indications, and rapidly disseminated. These drugs are then used widely, despite unknown effects in populations not included in RCTs such as the elderly and patients with high comorbidity.[37] Age has been associated with receipt of both chemotherapy in general as well as oxaliplatin specifically.[38,39] However, although the median age at colon cancer diagnosis is 72 years, the key RCT establishing oxaliplatin's efficacy had a median age of 60[40,41] and these results cannot be generalized to the older population, especially those over 75. The CTS PS method allows us to not only examine age and other specific characteristics of the general non-RCT population and their association with treatment decisions, but also to see how these things may have changed as the health care community adopted this novel drug and became more familiar with its side effects and clinical use over time.

Limitations of claims data such as lack of information on frailty,[7,42] census-level socioeconomic data, and inexact dates of diagnosis and service apply to these effectiveness results. Medicare is estimated to have 75% sensitivity for picking up 5-FU,[43] and therefore a proportion of the referent group may have been missed. Comorbidity assessed through claims may also be underestimated for this population, as older age is associated with less aggressive treatment for a number of diseases.[44]

This examination was performed in a single setting and results could be due to chance. The CTS PS should be examined in other settings and over more calendar years. If few potential matches for treated observations exist, the CTS PS may decrease efficiency by diminishing match options. Summary balance measures for PS matching are limited, as they may upweight multi-level variables and ignore individual covariate effects on the outcome, a prerequisite for confounding.

The construct of the calendar time-specific propensity score in the first years of a new drug or after a policy event is likely beneficial to confounding control and validity of estimates in non-experimental CER. The CTS PS allows transparent examination of changes in channeling over time for many covariates at once and is thus useful for understanding determinants of treatment receipt over a drug lifecycle. Creating a CTS PS also prompts researchers to start on the drug life year, which is sensitive to changes in drug prescription patterns, rather than the standard calendar year. Wider implementation of the CTS PS and comparison of estimates with conventional methods is needed in order to further understand the effects of accounting for time in studies of dynamic therapies.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This research was supported in part by NIH research grants R01 AG023178 from the National Institute on Aging, R01 CA124402 from the National Cancer Institute and T32KD07634. We thank Bryan J. Weiner for his support in association with this NCI grant.

## References

1. Rosenbaum PR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika. 1983; 70:41–55.
2. Stürmer T, Schneeweiss S, Brookhart MA, Rothman KJ, Avorn J, Glynn RJ. Analytic strategies to adjust confounding using exposure propensity scores and disease risk scores: nonsteroidal antiinflammatory drugs and short-term mortality in the elderly. Am J Epidemiol. 2005 May 1; 161(9):891–8. [PubMed: 15840622]
3. Rubin DB. Estimating causal effects from large data sets using propensity scores. Ann Intern Med. 1997; 127:757–763. [PubMed: 9382394]
4. National Cancer Institute. [accessed 30 October 2010] Highlights from ASCO 2003. May 31. 2003 http://www.cancer.gov/asco2003/highlights
5. André T, Boni C, Mounedji-Boudiaf L, Navarro M, Tabernero J, Hickish T, Topham C, Zaninelli M, Clingan P, Bridgewater J, Tabah-Fisch I, de Gramont A. 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; 350:2343–51. [PubMed: 15175436]
6. National Cancer Institute. [accessed 27 December 2012] FDA Approval for Oxaliplatin. http://www.cancer.gov/cancertopics/druginfo/fda-oxaliplatin
7. Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. Med Care. 2002 Aug; 40(8 Suppl):IV-3–18.
8. National Cancer Institute. [accessed 24 Sept 2010] SEER Program Overview. http://seer.cancer.gov/
9. Glymour, MM.; Greenland, S. Causal Diagrams. In: Rothman, KJ.; Greenland, S.; Lash, TL., editors. Modern epidemiology. 3. Philadelphia, PA: Lippincott Williams & Wilkins; 2008. p. 183-194.
10. Sanoff HK, Carpenter WR, Martin CF, Sargent DJ, Meyerhardt JA, Stürmer T, Fine JP, Weeks J, Niland J, Kahn KL, Schymura MJ, Schrag D. Comparative Effectiveness of Oxaliplatin vs Non-Oxaliplatin-containing Adjuvant Chemotherapy for Stage III Colon Cancer. J Natl Cancer Inst. 2012 Feb 8; 104(3):211–27. Epub 2012 Jan 20. [PubMed: 22266473]
11. Parsons, LS. Reducing bias in a propensity score matched pair sample using greedy matching techniques. Proceedings of the 26th annual SAS users group international conference; Cary, NC. http://www2.sas.com/proceedings/sugi26/p214-26.pdf
12. Stürmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. J Clin Epidemiol. 2006; 59:437–47. [PubMed: 16632131]
13. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. J Clin Oncol. 1983 Nov; 1(11):710–9. [PubMed: 6668489]
14. Seeger JD, Walker AM, Williams PL, Saperia GM, Sacks FM. A propensity score-matched cohort study of the effect of statins, mainly fluvastatin, on the occurrence of acute myocardial infarction. Am J Cardiol. 2003; 92:1447–1451. [PubMed: 14675584]
15. Seeger JD, Williams PL, Walker AM. An application of propensity score matching using claims data. Pharmacoepidemiol Drug Saf. 2005; 14:465–476.10.1002/pds.1062 [PubMed: 15651087]
16. Rassen JA, Glynn RJ, Rothman KJ, Setoguchi S, Schneeweiss S. Applying propensity scores estimated in a full cohort to adjust for confounding in subgroup analyses. Pharmacoepidemiol Drug Saf. 2012 Jul 21.7:697–709. Epub 2011 Dec 8. 10.1002/pds.2256 [PubMed: 22162077]
17. Mack CD, Glynn RJ, Stürmer T. Calendar Time-Specific Propensity Scores and Hazard Ratio Estimation. Pharmacoepidemiol Drug Saf. 2011; 20(suppl 1):S296.Oral presentation, 27th

International Conference on Pharmacoepidemiology and Therapeutic Risk Management; Chicago, IL. 2011.

18. Schneeweiss S, Gagne JJ, Glynn RJ, Ruhl M, Rassen JA. Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. Clin Pharmacol Ther. 2011 Dec; 90(6):777–90. Epub 2011 Nov 2. 10.1038/clpt.2011.235 [PubMed: 22048230]
19. Gagne JJ, Glynn RJ, Rassen JA, Walker AM, Daniel GW, Sridhar G, Schneeweiss S. Active safety monitoring of newly marketed medications in a distributed data network: application of a semi-automated monitoring system. Clin Pharmacol Ther. 2012 Jul; 92(1):80–6. Epub 2012 May 16. 10.1038/clpt.2011.369 [PubMed: 22588606]
20. Burakgazi AZ, Messersmith W, Vaidya D, Hauer P, Hoke A, Polydefkis M. Longitudinal assessment of oxaliplatin-induced neuropathy. Neurology. 2011 Sep 6; 77(10):980–6. Epub 2011 Aug 24. [PubMed: 21865571]
21. Weickhardt A, Wells K, Messersmith W. Oxaliplatin-induced neuropathy in colorectal cancer. J Oncol. 2011; 2011:201593. Epub 2011 Dec 12. [PubMed: 22203844]
22. Dyck PJ, Kratz KM, Karnes JL, Litchy WJ, Klein R, Pach JM, Wilson DM, O'Brien PC, Melton LJ 3rd, Service FJ. The prevalence by staged severity of various types of diabetic neuropathy, retinopathy, and nephropathy in a population-based cohort: the Rochester Diabetic Neuropathy Study. Neurology. 1993 Apr; 43(4):817–24. [PubMed: 8469345]
23. Ramanathan, RK.; André, T.; Rothenberg, ML.; de Gramont, A.; Tournigand, C.; Goldberg, RM.; Gupta, S. Diabetes mellitus (DM) and the incidence and time to onset of oxaliplatin induced peripheral sensory neurotoxicity (PSN) in patients with colorectal cancer. A pooled analysis of 3 randomized studies. American Society of Clinical Oncology; Gastrointestinal Cancers Symposium; 2008. p. Abstract 347http://www.asco.org/ascov2/Meetings/Abstracts?&vmview=abst_detail_view&confID=53&abstractID=10742
24. Haas JS, Brawarsky P, Iyer A, Fitzmaurice GM, Neville BA, Earle C. Association of area sociodemographic characteristics and capacity for treatment with disparities in colorectal cancer care and mortality. Cancer. 2011 Sep 15; 117(18):4267–76. Epub 2011 Mar 16. 10.1002/cncr.26034 [PubMed: 21413000]
25. André T, Boni C, Navarro M, et al. Improved Overall Survival With Oxaliplatin, Fluorouracil, and Leucovorin As Adjuvant Treatment in Stage II or III Colon Cancer in the MOSAIC Trial. J Clin Oncol. 2009 Jul 1; 27(19):3109–3116. [PubMed: 19451431]
26. Sanoff HK, Carpenter WR, Martin CF, Sargent DJ, Meyerhardt JA, Stürmer T, Fine JP, Weeks J, Niland J, Kahn KL, Schymura MJ, Schrag D. Comparative effectiveness of oxaliplatin vs non-oxaliplatin-containing adjuvant chemotherapy for stage III colon cancer. J Natl Cancer Inst. 2012 Feb 8; 104(3):211–27. Epub 2012 Jan 20. [PubMed: 22266473]
27. Schneeweiss S, Patrick AR, Stürmer T, Brookhart MA, Avorn J, Maclure M, Rothman KJ, Glynn RJ. Increasing levels of restriction in pharmacoepidemiologic database studies of elderly and comparison with randomized trial results. MedCare. 2007 Oct; 45(10 Supl 2):S131–42.
28. 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 Jun 1; 118(11):2925–34. Epub 2011 Nov 9. 10.1002/cncr.26622 [PubMed: 22072441]
29. Stürmer T, Rothman KJ, Avorn J, Glynn RJ. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution--a simulation study. Am J Epidemiol. 2010 Oct 1; 172(7):843–54. Epub 2010 Aug 17. [PubMed: 20716704]
30. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. Am J Epidemiol. 2006 Jun 15; 163(12):1149–56. Epub 2006 Apr 19. [PubMed: 16624967]
31. Austin PC, Grootendorst P, Anderson GM. A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a Monte Carlo study. Stat Med. 2007 Feb 20; 26(4):734–53. [PubMed: 16708349]
32. Rassen, JA.; Doherty, M.; Huang, W.; Schneeweiss, S. Pharmacoepidemiology Toolbox. Boston, MA: http://www.hdpharmacoepi.org [accessed 25 July 2012]

33. Mack CD, Brookhart MA, Glynn RJ, Stürmer T. Calendar Time as an Instrumental Variable in Nonexperimental Comparative Effectiveness Research of Dynamic Therapies. Am J Epidemiol. 2012; 175(suppl 11):181-S.Oral presentation, 45th Annual meeting Society for Epidemiologic Research; Minneapolis, USA. 2012.
34. Cain LE, Cole SR, Greenland S, Brown TT, Chmiel JS, Kingsley L, Detels R. Effect of highly active antiretroviral therapy on incident AIDS using calendar period as an instrumental variable. Am J Epidemiol. 2009 May 1; 169(9):1124–32. Epub 2009 Mar 24. [PubMed: 19318615]
35. Johnston KM, Gustafson P, Levy AR, et al. 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. [PubMed: 17847052]
36. Brookhart MA, Stürmer T, Glynn RJ, Rassen J, Schneeweiss S. Confounding control in healthcare database research: challenges and potential approaches. Med Care. 2010 Jun; 48(6 Suppl):S114–20. [PubMed: 20473199]
37. Bugeja G, Kumar A, Banerjee AK. Exclusion of elderly people from clinical research: a descriptive study of published reports. BMJ. 1997 Oct 25.315(7115):1059. [PubMed: 9366735]
38. Sanoff HK, Carpenter WR, Freburger J, Li L, Chen K, Zullig LL, Goldberg RM, Schymura MJ, Schrag D. Comparison of adverse events during 5-fluorouracil versus 5-fluorouracil/oxaliplatin adjuvant chemotherapy for stage III colon cancer: A Population-based analysis. Cancer. 2012 Jan 31. [Epub ahead of print]. 10.1002/cncr.27422
39. Abraham A, Habermann EB, Rothenberger DA, Kwaan M, Weinberg AD, Parsons HM, Gupta P, Al-Refaie WB. Adjuvant chemotherapy for stage III colon cancer in the oldest old: Results beyond clinical guidelines. Cancer. 2012 Jul 17. [Epub ahead of print]. 10.1002/cncr.27755
40. Altekruse, SF.; Kosary, CL.; Krapcho, M.; Neyman, N.; Aminou, R.; Waldron, W.; Ruhl, J.; Howlader, N.; Tatalovich, Z.; Cho, H.; Mariotto, A.; Eisner, MP.; Lewis, DR.; Cronin, K.; Chen, HS.; Feuer, EJ.; Stinchcomb, DG.; Edwards, BK., editors. SEER Cancer Statistics Review, 1975–2007. National Cancer Institute; Bethesda, MD: 2010. http://seer.cancer.gov/csr/1975_2007/, based on November 2009 SEER data submission, posted to the SEER web site [accessed 20 Mar 2010]
41. André T, Boni C, Mounedji-Boudiaf L, Navarro M, Tabernero J, Hickish T, Topham C, Zaninelli M, Clingan P, Bridgewater J, Tabah-Fisch I, de Gramont A. 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; 350:2343–51. [PubMed: 15175436]
42. Cooper GS, Virnig B, Klabunde CN, Schussler N, Freeman J, Warren JL. Use of SEER-Medicare data for measuring cancer surgery. Med Care. 2002; 40:IV-43–8.
43. Warren JL, Harlan LC, Fahey A, et al. Utility of the SEER-Medicare data to identify chemotherapy use. Med Care. 2002; 40:IV-55–61.
44. Glynn RJ, Knight EL, Levin R, Avorn J. Paradoxical relations of drug treatment with mortality in older persons. Epidemiology. 2001 Nov; 12(6):682–9. [PubMed: 11679797]

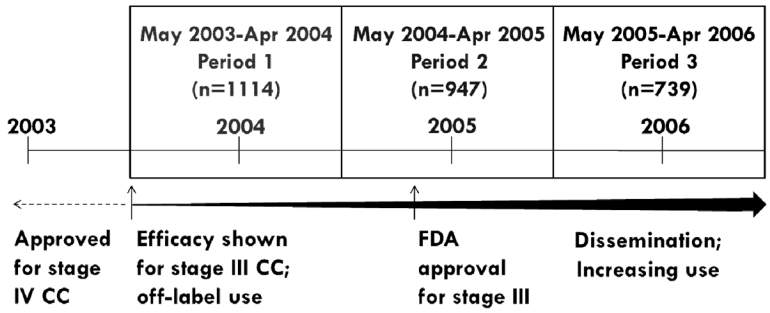


**Figure 1. Calendar time periods for stage III colon cancer patients based on first date of 5-FU or oxaliplatin receipt (N=2800) and FDA approval history of oxaliplatin**

5-FU=5-Fluorouracil; FDA=Food and Drug Administration; CC=colon cancer

Efficacy results based on the MOSAIC clinical trial, presented in May 2003; FDA approval for stage III colon cancer granted in November 2004.

Time period 1 (May 2003-April 2004) was used as referent time period in the conventional PS model that adjusts for calendar time as a confounder.

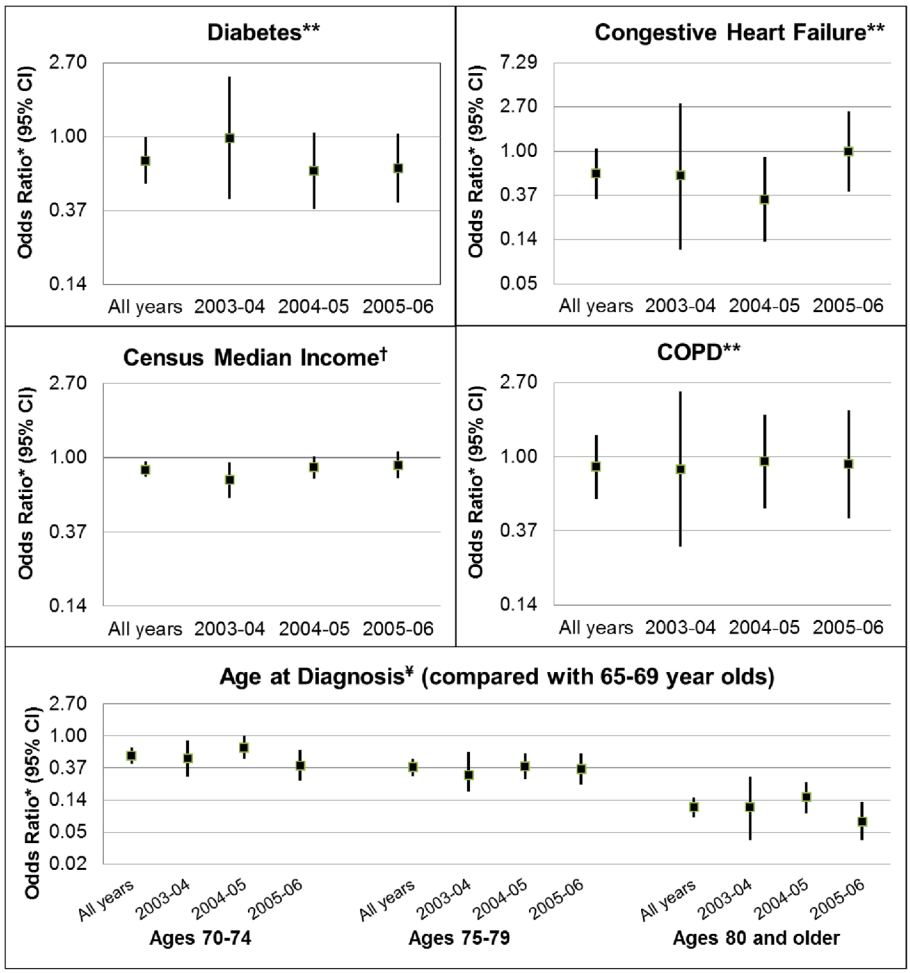


**Figure 2. Changes in channeling by covariate over study time periods (adjusted OR, 95% CI) comparing reciept of oxaliplatin with 5-FU**

OR=Odds Ratio; CI=Confidence Interval. Time periods are May through April of the years noted; time period specific estimates are from the calendar time-specific propensity score. Estimates for all years encompass May 2003-April 2006 and are from the conventional propensity score, adjusted for calendar time.

OR scale for CHF and Age at diagnosis is expanded due to wide confidence intervals or extreme values.

* Odds ratios are adjusted for all variables included in the propensity score

** Referent patients for diabetes, congestive heart failure and COPD are those without the condition

†Census median income modeled in quartiles as a continuous covariate, with highest income level as referent

¥Age modeled as categorical variable, with age group 65–69 as referent

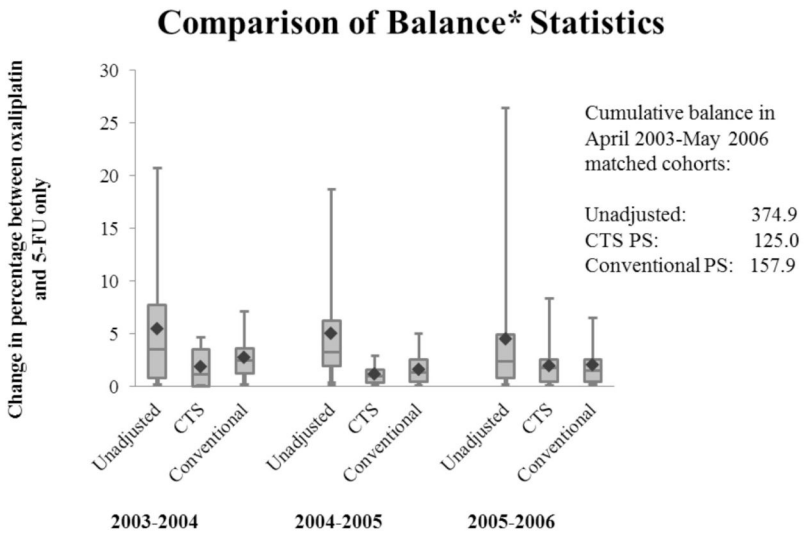


**Figure 3. Comparison of covariate balance between full (unmatched) population and matched cohorts generated by the conventional and CTS PS from April 2003 through May 2006**

CTS=Calendar Time-Specific; Time periods are May through April of the years noted. ♦ =Mean; Center line=median; Bottom/Top of box=25$^{th}$ / 75$^{th}$ percentile; Bottom/Top lines=Minimum/Maximum;

Time periods are May through April of the years noted.

*Balance measured by absolute difference in percentage between exposed and unexposed within covariate level for each time period.

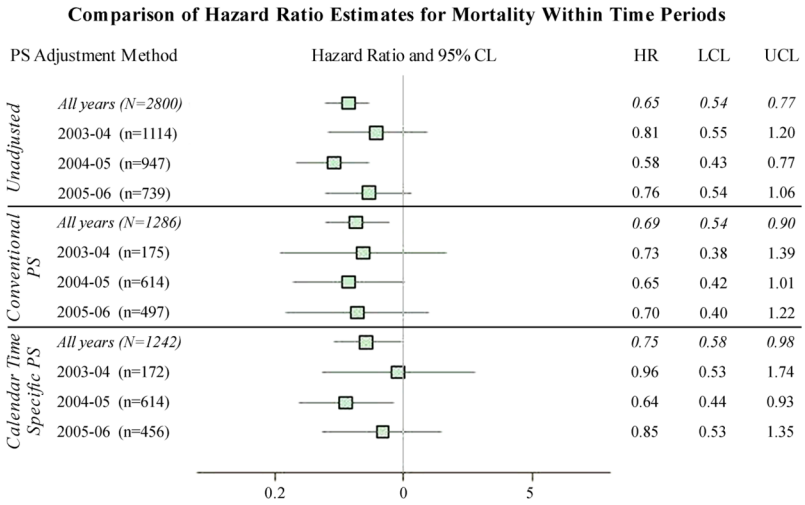


**Figure 4. Estimated hazard ratios for different PS adjustment methods comparing oxaliplatin with 5-FU for prevention of all-cause mortality, across all study years and within calendar-specific time periods**

PS=Propensity Score; HR=Hazard Ratio; CL=Confidence Limit; LCL=Lower Confidence Limit; UCL=Upper Confidence Limit; CTS=Calendar Time-Specific.

Time periods are May through April of the years noted.

For conventional PS year-specific estimates, the matched cohort was used but matching was broken for pairs that received first chemotherapy treatment in different time periods.

## Table 1

Characteristics and treatment receipt of Stage III Colon Cancer Patients in SEER-Medicare Study Population by Time Period (N=2800)

| Characteristic (N=2800) | May 2003–Apr 2004 N (%) | May 2004–Apr 2005 N (%) | May 2005–Apr 2006 N (%) |
|---|---|---|---|
| Treatment: * | | | |
| 5-FU without oxaliplatin | 1028 (92.3) | 609 (64.3) | 353 (47.8) |
| Oxaliplatin | 86 (7.7) | 338 (35.7) | 386 (52.2) |
| Race: | | | |
| Caucasian American | 975 (87.7) | 832 (87.9) | 639 (86.6) |
| African American | 83 (7.5) | 60 (6.3) | 53 (7.2) |
| Other | 54 (4.9) | 54 (5.7) | 46 (6.2) |
| Age: | | | |
| Mean (sd) | 75.1 (5.5) | 74.5 (5.6) | 74.8 (5.6) |
| Sex: | | | |
| Female | 602 (54.0) | 517 (54.6) | 419 (56.7) |
| Male | 512 (46.0) | 430 (45.4) | 320 (43.3) |
| Urbanity: | | | |
| Metro | 931 (83.6) | 781 (82.5) | 616 (83.4) |
| Urban | 166 (14.9) | 146 (15.4) | 108 (14.6) |
| Rural | 17 (1.5) | 20 (2.1) | 15 (2.0) |
| Substage: | | | |
| A | 99 (8.9) | 115 (12.1) | 80 (10.8) |
| B | 662 (59.8) | 542 (57.2) | 412 (55.8) |
| C | 346 (31.3) | 290 (30.6) | 247 (33.4) |
| Grade: | | | |
| Differentiated | 749 (67.2) | 635 (67.1) | 490 (66.3) |
| Undifferentiated/Unk | 365 (32.8) | 312 (32.9) | 249 (33.7) |
| Census Median Income: ** | | | |
| Mean (sd) | 49.8 (22.3) | 50.5 (24.2) | 50.6 (23.6) |
| Congestive heart failure (CHF): | 37 (3.3) | 42 (4.4) | 30 (4.1) |
| Myocardial Infarction (MI): | 32 (2.9) | 21 (2.2) | 15 (2.0) |
| Chronic obstructive pulmonary disease: | 66 (5.9) | 60 (6.3) | 43 (5.8) |
| Cerebrovascular Disease (CVD): | 31 (2.8) | 17 (1.8) | 16 (2.2) |
| Diabetes: | 96 (8.6) | 98 (10.3) | 115 (15.6) |
| Peripheral Vascular Disease (PVD): | 29 (2.6) | 24 (2.5) | 17 (2.3) |

Abbreviations SD, standard deviation; SEER, Surveillance, Epidemiology, and End Results.

Peptic ulcer disease, mild liver disease, paraplegia/hemiplegia, chronic renal failure, and rheumatologic disease are not shown but were also included in propensity score.

* First chemotherapy treatment received by newly diagnosed patients (intent-to-treat)

** USD, thousands

## Table 2

Mortality hazard ratios for stage III colon cancer patients treated with oxaliplatin versus 5-FU from May 2003 to April 2006

| Outcome Models | HR (95% CI)* | Percent difference** (absolute change) from CTS PS |
|---|---|---|
| Outcome model, unadjusted | 0.65 (0.54, 0.77) | −35% (0.15) |
| Outcome model, adjusted for time† | 0.71 (0.59, 0.87) | −16% (0.05) |
| Outcome model, not adjusted for time† | 0.67 (0.55, 0.80) | −30% (0.12) |
| **Propensity Score Models** | | |
| Conventional PS, adjusted for time | 0.69 (0.54, 0.90) | −22% (0.08) |
| Calendar Time-Specific (CTS) PS | 0.75 (0.58, 0.98) | -- |
| Full interaction PS | 0.73 (0.56, 0.95) | −10% (0.03) |
| Conventional PS, not adjusted for time | 0.67 (0.53, 0.84) | −30% (0.12) |

HR=Hazard Ratio; CI=Confidence Interval; CTS=Calendar Time-Specific; PS=Propensity Score;

*Although a true estimate is unknown, results can be indirectly compared with MOSAIC RCT results of HR (95% CI)=0.80 (0.65, 0.97).

**Percent difference calculated by: $\frac{(\ln(\text{CTS PS HR}) - \ln(\text{comparison HR}))}{\ln(\text{comparison HR})}$

†Outcome model comparators used conventional Cox proportional hazards regressionfor three estimates: unadjusted, adjusted for all covariates included in the propensity scores including calendar time, and adjusted for all covariates included in the propensity scores excluding calendar time