© Health Research and Educational Trust
DOI: 10.1111/1475-6773.12170
METHODS ARTICLE

Incident User Cohorts for Assessing Medication Cost-Offsets

Bruce Stuart, F. Ellen Loh, Pamela Roberto, and Laura Miller

Objective. To develop and test incident drug user designs for assessing cost savings from statin use in diabetics.

Data Source. Random 5 percent sample of Medicare beneficiaries, 2006–2008.

Study Design. Seven-step incident user design to assess impact of statin initiation on subsequent Medicare spending: (1) unadjusted pre/post initiation test; (2) unadjusted difference-in-difference (DID) with comparison series; (3) adjusted DID; (4) propensity score (PS)-matched DID with static and dynamic baseline covariates; (5) PS-matched DID by drug adherence strata; (6) PS-matched DID for high adherers controlling for healthy adherer bias; and (7) replication for ACE-inhibitor/ARB initiators.

Data Collection/Extraction Methods. Subjects with prevalent diabetes and no statin use (January–June 2006) and statin initiation (July 2006–January 2008) compared to nonusers with a random “potential-initiation” month. Monthly Medicare spending tracked 24 months pre- and post-initiation.

Principal Findings. Statistically significant savings in Medicare spending were observed beginning 7 months post-initiation for statins and 13 months post-initiation for ACEIs/ARBs. However, these savings were only observed for adherent patients in steps 5 and 6.

Conclusions. Drug initiator designs are more robust to confounding than prevalent user designs in assessing cost-offsets from drug use but still require other adjustments and sensitivity analysis to ensure proper inference.

Key Words. Drug initiator design, statins, cost-offsets

The advent of large, population-based computerized claims and medical records databases has led to an explosion of observational studies of medication effects in real-world settings. Yet, despite significant efforts to codify “good research practices” in observational designs (Moher et al. 2003; Owens et al. 2009; ISPOR, 2012), policy makers and many researchers remain skeptical about the validity and reliability of even the most carefully developed observational studies (Concato, Shah, and Horwitz 2000; Ioannidis et al. 2001; Ioannidis 2005, 2013). Recent advancements in propensity score

analysis, instrumental variables, and new user designs may help change that attitude. New user or initiator designs are particularly appealing in that the basic idea is to mimic clinical trials by measuring a drug's impact on patient outcomes from first use onward and comparing the results to a sample of non-users (or users of an alternative therapy) who share the same indication. Initially described by Ray (2003), and systematically evaluated by Schneeweiss et al. (2007), the approach has recently been recommended for comparative effectiveness studies in an AHRQ Effective Health Care Program Research Report (Johnson et al. 2012).

New initiator designs have several important advantages over alternative designs that compare prevalent drug users to nonusers (or high adherers to low adherers). Foremost, incident cohort designs make it possible to track changes in drug effectiveness from a given starting point, thus assuring proper temporal order between cause and effect. Comparisons among prevalent users are also likely to systematically underestimate early impacts of drug use (Schneeweiss et al. 2007) and may introduce an upward "survivor" bias to the extent that prevalent users are less likely to have discontinued use if they perceive beneficial effects from the drug. Long-term prevalent users may possess other unobserved characteristics associated with reduced risk of adverse outcomes, including healthier lifestyles and better adherence with other drugs in their medication regimens. Even when such characteristics are observable, unless the analyst knows the initiation date of the drug in question, statistical controls may inject unwarranted mediating effects into the causal pathway,¹ thus underestimating the drug's true impact (Johnson et al. 2012). Of course, these analytic advantages involve tradeoffs, some of which can be daunting. These include the need for very large study samples and the possible loss of generalizability with successive data reductions designed to improve internal validity.

In this article, we present a step-by-step tutorial in the application of incident user designs, taking as our example the impact of statin treatment on Medicare Part A and B expenditures for beneficiaries with diabetes enrolled in Part D prescription drug plans. There is a growing literature showing savings

---

Address correspondence to Bruce Stuart, Ph.D., The Peter Lamy Center on Drug Therapy and Aging, Department of Pharmaceutical Health Services Research, University of Maryland School of Pharmacy, 220 Arch Street, Room 12-212, Baltimore, MD 21201; e-mail: bstuart@rx.umaryland.edu. F. Ellen Loh, M.B.A., B.S.Pharm., is with the Department of Pharmaceutical Health Services Research, University of Maryland School of Pharmacy, Baltimore, MD. Pamela Roberto, M.P.P., is with the Pharmaceutical Research and Manufacturers of America, Washington, DC. Laura Miller, Ph.D., is with the National Association of Chain Drug Stores, Arlington, VA.

associated with evidence-based drug use in chronic disease (Shang and Goldman 2007; Zhang et al. 2009; Stuart et al. 2011; CBO 2012), but the studies are based almost exclusively on prevalent disease cohorts. In addition to highlighting the relative advantages and limitations of the incident user approach in modeling economic endpoints, we present several innovative applications intended to strengthen the internal validity of such designs and extend their reach to other analytic challenges faced in observational studies of drug effects. These innovative applications include the use of pre/post comparison series designs, dynamic controls for confounding by indication, visual time plots for evaluating drug indications and study outcomes, and explicit controls for healthy adherer bias.

METHODS AND RESULTS

Unlike traditional health services research articles, we present our methods in a series of steps, each followed by results. At every step, we provide a justification for the analytical procedure and then judge the findings in terms of potential threats to causal inference. Subsequent steps are designed to address threats identified in the preceding analysis. In this way, we hope to help the reader appreciate both the magnitude and direction of potential biases frequently seen in observational studies of drug outcomes. We then summarize the lessons learned in a checklist of good research practices for analyzing incident drug user cohorts. An online appendix contains complete results for all models estimated in the course of the study.

Our study cohorts were selected from a random 5 percent sample of Medicare beneficiaries with prevalent diabetes continuously enrolled in fee-for-service Part D plans between January 2006 and December 2008. Diabetes was identified from the following ICD-9 codes on hospital and medical claims: 250.xx, 357.2, 362.01, 362.02, or 366.41. Our primary drug of interest was initiation with any medication in the class of HMG-CoA reductase inhibitors (statins). Numerous clinical trials have demonstrated the efficacy of statins in reducing cardiovascular complications among patients with diabetes (Ross et al. 1999; Leiter 2005). As such, they are almost universally recommended in diabetes treatment guidelines for individuals with even mildly elevated LDL cholesterol levels (American Diabetes Association 2008).

We identified initiators with no evidence of statin use between January and June 2006 and at least one refill following their initial statin fill between July 2006 and January 2008. The sample was restricted to individuals with at

least two statin fills in order to exclude cases where a physician might discontinue the medication based on early adverse side effects (Sinzing, Wolfram, and Peskar 2002). We required at least 11 months of follow-up after initiation to provide sufficient time to observe short and intermediate effects associated with this drug class (Aubert et al., 2010). We excluded enrollees in capitated Medicare Advantage plans because they generate no Medicare Part A and B claims, which were necessary for assessing expenditures. For the same reason, we excluded beneficiaries with noncontinuous coverage for Medicare Part A and B. Finally, we excluded any beneficiary who died before December 31, 2008.² Our final sample comprised 15,169 diabetics enrolled in stand-alone Part D plans who initiated a statin between June 2006 and January 2008 together with a control group of 35,505 diabetics with no evidence of statin use over the entire study period. Because of our rolling enrollment criterion, we tracked subjects up to 24 months prior to drug initiation and up to 24 months post-initiation. Characteristics of these two cohorts are shown in the upper panel of Table 1.

Step 1: Pre/Post Drug Initiator Design

Typically, incident drug user designs focus on the probability and timing of events occurring after drug initiation, relying upon baseline preinitiation information as control variables. For some outcomes like mortality and rare adverse events, such post-only designs are the only feasible alternative. However, if the outcome measure of interest is continuous and the research question seeks to determine whether the drug reduces the magnitude of that outcome, then pre/post designs are superior because subjects act as their own controls. Monthly Medicare spending meets that criterion and is an appropriate choice based on evidence that statin use not only reduces risk of macrovascular disease in diabetics (Ross et al. 1999) but may also reduce the development of microvascular complications of the disease, including nephropathy, retinopathy, and neuropathy (Almuti et al. 2005; Leiter 2005). And although statins are generally prescribed to reduce long-term cardiovascular risk, they have been shown to have short-term benefits (Tsiara, Elisaf, and Mikhailidid 2003), which suggests that the economic returns to statin therapy could be evident within months of initiation.

Based on this literature, we hypothesized that mean monthly Medicare spending among study subjects would decrease following statin initiation with savings growing over time. To assess this expected trend, we stratified the follow-up period into three segments, 1–6, 7–12, and 13–24 months. But

Table 1: Characteristics of Medicare Beneficiaries with Diabetes Who Newly Initiate Statins and Nonusers of Statins

<i>Characteristics</i>	<i>New Statin Initiators (N = 15,169)</i>	<i>Nonusers (N = 35,505)</i>	<i>Standardized Difference* (%)</i>
Low-income subsidy (LIS) (%)	64.5	69.2 [†]	−10.0
Age (%)			
<65	23.8	24.6	−1.8
65–74	34.1	25.0 [†]	20.0
75–84	32.2	30.2 [†]	4.1
85+	10.0	20.2 [†]	−28.9
Sex (%)			
Male	34.8	33.1	3.5
Female	65.2	66.9	−3.5
Race/ethnicity (%)			
White	70.6	71.1	−0.7
Black	17.9	19.0 [†]	−2.1
Hispanic	5.6	4.7 [†]	2.8
Other	5.9	5.2 [†]	2.1
Region (%)			
Northeast	17.4	18.6 [†]	−3.1
North Central	21.8	23.6 [†]	−4.5
South	44.7	41.8 [†]	5.7
West	16.2	16.0	0.7
RxHCC comorbidity count (SD)	7.7 (3.5)	7.3 [†] (3.5)	11.4
Diagnosis of hyperlipidemia	37.6	23.2 [†]	31.8
Follow-up period (%)			
11–18 months	47.17	55.62 [†]	−13.8
19–24 months	52.83	44.38 [†]	13.8
Medicare spending comparisons			
Mean preperiod monthly Medicare spending (\$, SD)	1,111 (2,069)	1,097 (1,889)	0.7
Mean postperiod monthly Medicare spending, 1–6 months (\$, SD)	1,213 (2,358)	1,108 [†] (2,236)	6.6
Pre/post difference	102 [†]	10 [†]	
Mean postperiod monthly Medicare spending, 7–12 months (\$, SD)	1,208 (2,507)	1,154 [†] (2,453)	2.2
Pre/post difference	97 [†]	57	
Mean postperiod monthly Medicare spending, 13–24 months (\$, SD)	1,259 (1,955)	1,260 (2,686)	0.0
Pre/post difference	148 [†]	162	

Note. *The standardized difference in % is the mean difference as a percentage of the average standard deviation: $100(\bar{X}_I - \bar{X}_N)/\sqrt{(S_I^2 + S_N^2)/2}$, where for each covariate $\bar{X}_I$ and $\bar{X}_N$ are the sample means in the initiator and nonuser groups, respectively, and S_I^2 and S_N^2 are the corresponding sample variances.

[†]Difference between initiators and nonusers significant at $p < .05$.

instead of observing declining costs in our initial descriptive comparison, we found that spending was higher throughout the follow-up period, ranging from around \$100 more per month in the first year and then rising to \$148 more in months 13–24 (Table 1, lower panel). These differences were all statistically significant at $p < .05$ (Table 2, Step 1). Two possible reasons for such counterintuitive findings are maturation and history. Maturation might be responsible for higher post-initiation spending if the underlying condition worsens over time. History refers to all events that occur between the beginning and the end of the study period that could affect the study outcome in the absence of treatment (Shadish, Cook, and Campbell 2002).

Step 2: Pre/Post with Control Group Design

A standard approach to controlling for maturation and history is to compare time trends of treated subjects to individuals who resemble the treatment group but are untreated. Our control group included all Medicare beneficiaries with diabetes who met the study inclusion criteria but had no evidence of statin use over the entire study period. We considered these individuals “potential statin initiators.”³ To obtain pre/post values for nonusers, we selected a random “potential-initiation” month between July 2006 and January 2008 for each person in the control group. The results showed that nonusers also had higher monthly Medicare spending in the postperiod rising by \$10 in months 1–6, \$57 in months 7–12, and \$162 in months 13–24. This pattern is consistent with the maturation/history hypothesis, but the differences were smaller than for new statin initiators. We estimated a simple unadjusted pre/post control group design (also known as a difference-in-difference or DID model) with ordinary least squares (OLS)⁴ and found that the association of statin initiation controlling for the experience of nonusers remained positive in months 1–6 (\$91, $p < .05$) and months 7–12 (\$41, ns) but turned negative (\$–13, ns) in months 13–24 (Table 2, Step 2 and Table A1).

Step 3: Adjusting for Sample Differences

Of course, the true drug effect could also be confounded by differences in characteristics of the two groups (as can be seen in Table 1, the two cohorts differ on most factors, age in particular). To assess that concern, we estimated a multivariate DID model, including baseline preinitiation measures for demographic factors, comorbidities, and prevalence of hyperlipidemia. While most of the covariates in this model were significant predictors of Medicare

Table 2: Estimated Change in Monthly Medicare Spending Associated with Statin Initiation among Medicare Beneficiaries with Diabetes Derived from Alternative Models

Model Step Number	Model	Sample Size		Estimated Treatment Effect on Change in Monthly Medicare Spending (SE)
		New Initiators	Nonusers	
(1)	Unadjusted difference pre/post initiation for new drug initiators			
	1st to 6th months post-initiation	15,169	-	\$102* (25)
	7th to 12th months post-initiation	15,169	-	\$97* (26)
	13th to 24th months post-initiation	14,233	-	\$148* (27)
(2)	Unadjusted difference-in-difference model			
	1st to 6th months post-initiation	15,169	35,505	\$91* (23)
	7th to 12th months post-initiation	15,169	35,505	\$41 (25)
	13th to 24th months post-initiation	14,233	32,585	\$-13 (26)
(3)	Adjusted difference-in-difference model			
	1st to 6th months post-initiation	15,169	35,505	\$91* (23)
	7th to 12th months post-initiation	15,169	35,505	\$41 (25)
	13th to 24th months post-initiation	14,233	32,585	\$-19 (26)
(4)	Basic propensity score matched model			
	1st to 6th months post-initiation	13,176	13,176	\$-2 (26)
	7th to 12th months post-initiation	13,176	13,176	\$-1 (30)
	13th to 24th months post-initiation	12,379	12,054	\$-59 (32)
(5)	New drug initiators with post-period PDC values over and under 80%			
(5a)	PDC ≥ 80%			
	1st to 6th months post-initiation	4,012	4,012	\$-20 (45)
	7th to 12th months post-initiation	4,012	4,012	\$-70 (50)
	13th to 24th months post-initiation	3,721	3,688	\$-99 (56)

continued

Table 2. *Continued*

Model Step Number	Model	Sample Size		Estimated Treatment Effect on Change in Monthly Medicare Spending (SE)
		New Initiators	Nonusers	
(5b)	PDC <80%			
	1st to 6th months post-initiation	4,012	4,012	\$90 (46)
	7th to 12th months post-initiation	4,012	4,012	\$67 (54)
	13th to 24th months post-initiation	3,801	3,688	\$42 (58)
(6)	Propensity score matched model restricted to beneficiaries with PDC ≥80% adjusted for healthy adherer bias			
	1st to 6th months post-initiation	5,066	5,066	\$-77 (44)
	7th to 12th months post-initiation	5,066	5,066	\$-160* (47)
	13th to 24th months post-initiation	4,691	4,632	\$-223* (50)

Note. *Difference significant at $p < .05$.

spending (Table A2), the treatment effects from the post-period/new initiator interaction terms were virtually unchanged from step 2 (Table 2, Step 3).

Step 4: Adjusting for Confounding by Indication

The fact that the multivariate DID model still produced counterintuitive results led us to believe that some unobserved dynamic factor(s) might be confounding the true impact of statin therapy initiation. To assess that possibility, we plotted mean monthly Medicare costs over the entire observation period for both groups (Figure 1). Altogether, there were 726,482 observations pre-initiation (195,805 for new initiators and 531,677 for nonusers) and 1,000,664 observations post-initiation (314,438 for new initiators and 686,226 for nonusers). The obvious difference between the two time plots is the spike in spending in the initiation month for statin users. In fact, in that month statin initiators spent about two-and-a-half times as much as nonusers. Most of the difference could be traced to differential hospitalization rates. Over the course of the study period, between 4 and 5 percent of the two groups were hospitalized in any given month, but 12 percent of new initiators were hospitalized during their initial month of statin use and the hospitalization rates for this group were also elevated in the 2 months prior to initiation (results not shown).

Figure 1: Mean Pre/Post Monthly Part A and B Spending Comparing Medicare Beneficiaries with Diabetes Who Newly Initiate Statins to Nonusers of Statins

It thus appears that for a nontrivial portion of initiators, statin use coincided with some adverse health event that could in turn influence post-initiation Medicare costs—in short, confounding by indication.

To address this confound, we turned to propensity score (PS) analysis. Various PS techniques have been used in prior studies to compare drug users and nonusers (Wang and Donnan 2001; Weitzen et al. 2004; Glynn, Schneeweiss, and Sturmer 2006). Our application of this technique is unique in that in addition to matching statin initiators and nonusers on the baseline measures shown in Table 1, we also included five preinitiation Medicare cost segments (27 to 4 months, 3 months, 2 months, 1 month, and the initiation month) in the matching algorithm.⁵ In this way, we could implicitly control for the underlying dynamic adverse health events associated with drug initiation for certain statin users. After matching, we had 13,176 individuals in each group with virtually identical average values for the demographic variables, comorbidity score, and prevalence of hyperlipidemia, as well as very similar values on the preinitiation Medicare cost segments (Table A3). Time plots for the two matched groups are shown in Figure 2 (note the scale in this figure has been expanded to make monthly difference in spending more legible). Note that

Figure 2: Mean Pre/Post Monthly Part A and B Spending Comparing Matched Cohorts of Medicare Beneficiaries with Diabetes Who Newly Initiate Statins to Nonusers of Statins

the preinitiation spike among statin users is attenuated and there is a parallel uptick in Medicare spending prior to the index month for noninitiators. When we replicated the DID model using this PS-matched sample (Table 2, Step 4),⁶ we found that statin initiation was now associated with lower Medicare spending throughout the entire follow-up period but none of the savings estimates achieved statistical significance.⁷

Step 5: Adjusting for Statin Adherence

The Step 4 result might be considered a reasonable estimate of the association of statins on Medicare costs for the entire sample. Whether it represents a good estimate of the treatment effect on the treated depends on statin utilization patterns post-initiation. Our inclusion criteria specified at least two statin fills but did not impose any requirement that initiators be good adherers. One plausible explanation for the null finding in Step 4 is that some new initiators were poor adherers or discontinued use altogether. In Step 5, we took adherence explicitly into consideration. To measure adherence, we used the proportion of days covered (PDC) metric, which divides the sum of days' supply for

all statin fills during the post-initiation period by the number of days between the initial fill and December 31, 2008. We removed inpatient hospital and postacute nursing home days from the PDC denominator as facilities are responsible for providing all inpatient drug therapy.

Our Step 4 adherence results showed a mean statin PDC of 59 percent with just 36.4 percent of initiators in Step 4 achieving a level of $\text{PDC} \geq 80$ percent over the follow-up period. A PDC value of 80 percent is commonly used as a threshold for good adherence (Osterberg and Blaschke 2005). Based on this finding, we concluded that it is likely that the Step 4 results mask important subgroup associations when the subgroups are defined by adherence. To obtain better estimates of these parameters, we divided the new initiator PS-matched cohort into those with PDC values over and under 80 percent and then compared each group to the nonusers from Step 4 using a three-way propensity matching procedure following Solomon et al. (2010). First, we matched new statin initiators with postindex $\text{PDC} \geq 80$ percent to new initiators with $\text{PDC} < 80$ percent. Next, we matched nonusers to the high adhering statin group. Finally, because the nonuser cohort was the largest group, we only used pairs of low and high adherers that also matched nonusers. The time plots for pre- and post-initiation monthly average Medicare costs for the three groups are presented in Figure 3. High statin adherers had consistently lower Medicare costs compared to low adherers and nonusers of statins.

Results from the DID estimates of statin effects for those with $\text{PDC} \geq 80$ percent were $-\$20$ in months 1–6, $-\$70$ in months 7–12, and $-\$99$ in month 13–24 as shown in Table 2, Step 5a (full results shown in Table A4). None of these results was statistically significant. On the other hand, the estimates of drug effects from statin initiators with PDC values below 80 percent were all positive as shown in Table 2, Step 5b, which supports the old adage that drugs only work if they are taken.

Although a PDC cutoff at 80 percent is the conventional way to define good adherence, it is nonetheless an arbitrary threshold and thus raises the question of whether lower adherence levels might also generate Medicare savings. Prior research has suggested that beneficiaries with both high and moderate levels of adherence experience lower costs and better clinical outcomes in relation to those with less optimal use (Briesacher et al. 2007; Esposito et al. 2009; Pittman et al. 2011). To assess that possibility, we re-estimated our Step 5 model by matching nonusers to statin initiators with a PDC above and below 70 and 60 percent, respectively. As expected, the results (not shown) indicated lower net savings associated with the lower PDC thresholds.

Figure 3: Mean Pre/Post Monthly Part A and B Spending Comparing Matched Cohorts of Medicare Beneficiaries with Diabetes Who Newly Initiate Statins Stratified by Post-Initiation PDC to Nonusers of Statins

Step 6: Adjusting for Healthy Adherer Bias

One additional confound required our attention; namely, the possibility that high statin adherers were high adherers with other unmeasured medications that also contributed to lower Medicare costs. Such a “healthy adherer bias” (Humphrey, Chan, and Sox 2002; Brookhart et al. 2007) would overstate the true statin effect. To correct for this potential bias, we re-estimated our propensity score model including variables for preinitiation period use and adherence with two other drug classes recommended in diabetes treatment guidelines: oral antidiabetic drugs and ACE-inhibitors/ARBs. The logic here is that any healthy adherer effect would be shared among the drug measures; that is, that high adherers with one medication are likely to be high adherers with other drugs in the regimen. In that case, any shared variance associated with overall high rates of adherence will be eliminated from the estimated impact of each individual drug variable, thus expunging healthy adherer bias from our measure of statin effectiveness.

The matched sample for Step 6 included 5,066 individuals in each study arm. As hypothesized, the statin effect on monthly Medicare spending fell slightly to $-\$77$ in months 1–6 (ns), $-\$160$ in months 7–12 ($p < .05$), and $-\$223$ in months -13 to 24 ($p < .05$) (Table 2, Step 6 with full results in

Table A5). The time plots in Figure 4 are similar to those in Figure 3 with consistently lower mean monthly Medicare costs for statin users beginning 4 months after drug initiation. Overall, the estimated savings attributable to high statin adherence reached 20 percent of preinitiation monthly Medicare spending by this group (\$1,108) in the second year after drug initiation.

Step 7: Replication

A final challenge was the potential for model overfitting. Because we developed the new user design as a sequence of steps within the same sample, we risked the possibility of inadvertently selecting methods that confirmed our hypotheses. To assess the robustness of our methods, we replicated the six steps in a cohort of 13,948 diabetes patients who newly initiated therapy with ACE-inhibitors or ARBs between July 2006 and January 2008. We selected ACEI/ARBs because this class is also recommended in diabetes treatment based on strong evidence of effectiveness from clinical trials (American Diabetes Association 2008). The sample selection criteria (except for drug class) and

Figure 4: Mean Pre/Post Monthly Part A and B Spending Comparing Matched Cohorts of Medicare Beneficiaries with Diabetes Who Newly Initiate Statins with Post-Initiation PDC ≥ 80 percent to Nonusers of Statins Adjusted for Healthy Adherer Bias

analytic steps duplicated those described above for the statin initiator analysis. Full results are presented in the Appendix.

As in the statin analysis, we discovered that mean monthly Medicare costs post-ACEI/ARB initiation were significantly higher than before initiation: \$297 in months 1–6, \$264 in months 7–12, and \$335 in months 13–24 (all significant at $p < .05$). Comparing new initiators to nonusers had no substantive effect on this relationship in either unadjusted or adjusted DID models. Charting monthly Medicare costs over the entire study period revealed a sharp spike in spending prior to initiation, but unlike the statin case, post-initiation spending never declined to the level of nonusers. Even after matching initiators and nonusers to four decimal points in a propensity score model, we still estimated that ACEI/ARB use *added* \$117 in Medicare costs in months 1–6, \$96 in months 7–12, and \$87 in months 13–24 (all significant at $p < .05$). Adherence with ACEIs/ARBs was marginally higher than for statins with a mean PDC = 63 and 42.2 percent having a PDC ≥ 80 percent. Only when we compared matched samples of new initiators with PDC < 80 percent to initiators with PDC ≥ 80 percent and nonusers did savings appear (–\$50 in months 1–6, –\$38 in months 7–12, and –\$89 in months 13–24), albeit none of the estimates was statistically significant. Adjusting for healthy adherer bias actually raised the savings estimates to –\$60 in months 1–6 (ns), –\$84 in months 7–12 (ns), and –\$159 in month 13–24 ($p < .05$). Statistically significant savings associated with ACEI/ARB initiation were observed in the second year following initiation and averaged approximately 15 percent of preinitiation Medicare spending. There were no savings associated with ACEI/ARB adherence rates below 80 percent; indeed, the estimated Medicare costs for low adherers exceeded those of nonusers throughout the post-initiation period.

DISCUSSION

What We Learned about Drug Initiators

Perhaps the most important methodological lesson we learned from examining the behavior of drug initiators is that for many individuals the decision to begin therapy occurred in the context of some unfolding adverse health event (s). While we did not observe the actual clinical factors involved, spikes in Medicare spending just prior to initiation with both drug classes provide strong a priori evidence for such an interpretation. This is problematic for researchers because such confounding by indication can lead to downward

bias in estimates of drug effectiveness when differences in costs post-initiation are compared to preinitiation costs.

A related lesson is the importance of being able to statistically control for dynamic trends in factors predictive of drug initiation in observational studies. In clinical trials, this is not an issue as everyone selected presumably has an active indication for the treatment at the time of randomization. However, in our case we observed rising medical expenditures prior to initiation. This situation is not amenable to traditional propensity score algorithms that match on static variables during the baseline, preinitiation period. Because Medicare expenditures are a continuous variable, we could control for this confound by matching on segments of preinitiation Medicare spending designed to mimic observed trends among drug initiators. Visual analysis of spending trends postmatching confirmed that the spikes were substantially reduced but not totally eliminated by matching with tight calipers. We thus conclude that our methods produce lower bound estimates of the true drug effect.

A third lesson of note is the powerful influence that drug adherence plays in generating positive outcomes from drug therapy. This is no surprise of course. The literature is awash in studies supporting the importance of good drug adherence (DiMatteo et al. 2002; Sokol et al. 2005; Salas et al. 2009). However, the magnitude of the effect was so large in our analysis that on balance drug initiation did not appear to reduce subsequent Medicare spending. The greatest savings accrued to the highest adherers (PDC ≥ 80 percent) who represented just 36 percent (statins) and 42 percent (ACEIs/ARBs) of all drug initiators. We found some evidence of savings among beneficiaries with PDC levels as low as 60 percent.

One possible explanation for why we did not see aggregate Part A and B savings across all new initiators could be related to the high rate of therapy discontinuation among beneficiaries with poor adherence. Premature discontinuance is by far the most common form of poor drug adherence (World Health Organization 2003; Andrade, Kahler, and Frech 2006) and more than 40 percent of beneficiaries with PDC < 80 percent in our sample discontinued use of statins for a minimum of 6 months during follow-up. Not surprisingly, beneficiaries who prematurely cease to use recommended medications are likely to remain at a higher risk of future adverse health events; in effect, they are non-users as opposed to poor adherers.

Finally, we found mixed evidence of healthy adherer bias from our two initiator cohorts. For statin initiators, adding preinitiation drug adherence measures to the propensity score matching criteria resulted in a 4 percent decline in estimated savings in the second year following drug initiation, but

for ACEI/ARB initiators, our savings estimates nearly doubled with the inclusion of these control variables during that time span.

Strengths and Limitations of New Initiator Designs

The major advantage of new initiator designs over more traditional methods of assessing drug effectiveness in observational studies is the ability to assure a strong temporal association between drug use and outcomes measured over time. Specifically, it permits researchers to identify when drug impacts first become apparent and how long they last. Other advantages include more robust controls for confounding by indication and healthy adherer bias.

A limitation of drug initiator designs is that they require large population databases to generate samples of sufficient size for precise estimates of drug effectiveness. Our study began with a sample of 50,674 Medicare beneficiaries diagnosed with diabetes of whom slightly more than a third (15,169) newly initiated statin therapy between July 2006 and January 2008. The initial PS matching algorithm (Step 4) reduced the sample of initiators to 13,176. Limiting the study to matched samples of high adherers, low adherers, and nonusers (Step 5) reduced it further to 4,012 initiators. While such sample restrictions increase the internal validity of cost comparisons to nonusers, they also limit the generalizability of results to the larger population. Schneeweiss et al. (2007) argue that the question of generalizability depends on how the results are to be used. Taking the perspective of the physician who wants to know whether therapy should be started for a particular patient, restricting the study to new initiators is perfectly appropriate. On the other hand, matching new initiators and nonusers on demographic characteristics, comorbidities, and drug indications does limit generalizability. This is a particular concern if one suspects that treatment effects will differ across patient population. Accounting for heterogeneous treatment effects is just as important in observational studies as in clinical trials (Rothwell 1995, 2005; Hayward et al. 2005).

The requirement of adequate sample sizes also dictated the use of a relatively long (19 months) window for capturing new drug initiators. But that decision had an unintended consequence of limiting the number of pre- and post-initiation monthly observations at both tails of the time series. In particular, smaller numbers of observations meant greater variation and less precise estimates of drug effectiveness at the end of the observation period. Of course, this limitation could be addressed by adding additional years of data to the sample (Abbas et al. 2012). In addition, adding more years of data would make it possible to extend the washout period for identifying new initiators

and thus increase our confidence that new initiators were truly that and not previous users with long gaps in use.

Our choice of dependent variable (cost) and unit of observation (the month) also contributed to high variances. We chose Medicare spending as the dependent variable (and primary preinitiation control variable) because it encompasses a broad spectrum of health factors in a single continuous metric. However, month-to-month spending among study subjects varied from zero (the modal value) to several thousands of dollars, thus creating wide fluctuations and statistical noise in the time series both pre- and post-initiation. We had sufficiently large samples to identify the main signal despite the noise, but it did limit our ability to draw strong inferences regarding short-term monthly changes in outcomes. An alternative strategy might be to use the log of costs as the dependent variable (giving zeros a value of \$1), which would dampen monthly spikes in costs while maintaining important trend values.

Finally, there is the question of whether nonusers represent the best comparison group for new initiators. We considered nonusers to be “potential-initiators,” but that is clearly not true for some and perhaps many of them. Only 23 percent of the statin nonusers in our sample had a baseline diagnosis of hyperlipidemia (then again, the rate for initiators was less than 38 percent). Matching on the presence of hyperlipidemia at baseline resulted in sample pairs with identical probabilities of having the diagnosis (33 percent), but the possibility remains that the two groups differed systematically on factors we could not assess. Instrumental variable techniques might help here, but finding good instruments for drug initiation could pose a challenge.

Estimated Savings from Drug Initiation

We estimate that Medicare Part D enrollees with diabetes who newly initiated therapy with statins and ACEIs/ARBs between 2006 and 2008 *and were adherent to therapy* saved the Medicare program between 15 and 20 percent per month in spending on traditional Part A and Part B services in the second year following initiation. Lower savings (14 percent) were observed in the first 6 months following statin initiation, but no significant savings were observed during this period following initiation of ACEI/ARBs. These estimated savings exceed the cost of drugs in both therapeutic classes. However, we also found a high proportion of new drug initiators were not adherent and that poor adherers did not experience the same spending reductions as high adherers.

These findings demonstrate the opportunities available to policy makers and Part D plan managers to reduce program costs (and presumably achieve

better health outcomes as well, although we did not study that) by improving adherence with evidence-based medications. Despite the benefits of drug therapy in maintaining or improving patient health, medication use in general remains suboptimal, with adherence rates declining shortly after initiation and worsening significantly over time (Benner et al. 2002; Karter et al. 2009). In an effort to optimize therapeutic outcomes through better medication use, the Medicare Modernization Act of 2003 required Part D prescription drug plans to offer medication therapy management (MTM) services to qualified beneficiaries; however, it remains unknown whether current eligibility criteria are optimally targeted to capture those most at risk of underuse and poor adherence to recommended medications (Stuart et al. 2014).⁸ By studying incident user cohorts, we believe that researchers will be able to more accurately predict costs associated with nonadherence and early discontinuation, potentially identifying problematic utilization patterns that Part D plans could use to target and evaluate the effectiveness of MTM interventions.

CONCLUSIONS

There is a large literature on potential cost-offsets from use of and adherence to evidence-based medications in treating common chronic diseases in the elderly. However, most studies use prevalent user designs that are subject to serious methodological limitations, including lack of temporal order between cause and effect, underestimates of early impacts of drug use, survivor bias, and inadequate controls for indication bias. Our study used an incident user design to determine if diabetics on statin therapy generate savings in Medicare costs compared to diabetics who do not use statins. We conducted a series of tests with increasing sample restrictions in order to maximize the internal validity of the comparison. We conclude that incident user designs have several advantages over prevalent user designs in medication costs studies, but they still require other adjustments and sensitivity analyses to ensure proper inference.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: This work was supported by funding from NACDS and PhRMA. We appreciate the two unidentified reviewers' effort to help us make this manuscript better.

Disclosures: None.

Disclaimers: None.

NOTES

1. In prevalent drug use studies, it is generally impossible to assure that all variables are measured before drug initiation because the date of first drug use is unknown. The problem arises when a variable affects the outcome of interest but is itself affected by drug use (e.g., blood pressure in studies of the impact of antihypertensive medications). Including values for that variable occurring subsequent to first drug use will lead to a systematic underestimate of the true drug effect. New initiator designs avoid that problem because the date of first use is known.
2. Decedents within the 11 months following their index dates were excluded pursuant to our requirement of a minimum postindex observation period for all study subjects. Given this restriction, it made sense to exclude decedents throughout the entire follow-up period. However, this exclusion is not strictly necessary to the design. In a series of sensitivity tests, we dropped the minimum required follow-up criterion and included subjects who died at any point after their index date. The results (not shown) were not substantively different from those reported in the text.
3. One might consider a diagnosis of hyperlipidemia to be a more precise indicator for potential statin initiation. However, as we note in Step 4, a minority of subjects had such a diagnosis in their Medicare claims (Table 1) whether they initiated statins or not.
4. The model was $Y = \alpha + \beta_1 I + \beta_2 P + \beta_3 (I + P) + \mu$, where Y = monthly Medicare expenditures; I = new initiators (nonusers as reference); P = the post-initiation period (preperiod as reference), α and β are parameters to be estimated, and μ is the error term. The coefficient, β_3 provides the DID estimate of interest. We used OLS in all of the regression models in order to preserve the linear relationship in Medicare spending reflected in the bivariate comparisons. This enabled us to directly compare unadjusted and adjusted estimates of Medicare savings.
5. The model used logistic regression with initiator status as the dependent variable (1 = statin initiators, 0 = nonusers) and the demographic variables, comorbidity counts, hyperlipidemia diagnosis, and preinitiation Medicare cost segments as regressors. For this model, we computed the probability (the propensity scores) that each observation was a statin initiator. We then used a greedy matching technique to create initiator/nonuser pairs with like propensity scores with the caliper set to four decimal points.
6. We used the same DID model as in Note 1 above. However, in this case it is technically a multivariate model as we have controlled for other factors through the PS matching algorithm.
7. An alternative approach would have been to stratify the sample according to evidence of hospitalization during the immediate preinitiation period. This approach could be combined with PS matching within the hospitalized and nonhospitalized

cohorts to achieve greater specificity of the causal path, albeit at the cost of smaller sample sizes and less precise estimates.

8. As of 2013, MTM services were most commonly offered to Part D beneficiaries with annual drug spending of at least \$3,144, a minimum of two or three chronic conditions, and taking between two and eight different Part D covered drugs.

REFERENCES

- Abbas, S., P. Ihle, I. Koster, and I. Schubert. 2012. "Estimation of Disease Incidence in Claims Data Dependent on the Length of Follow-Up: A Methodological Approach." *Health Services Research* 47 (2): 746–55.
- Almuti, K., R. Rimawi, D. Spevack, and R. Ostfeld. 2005. "Effects of Statins beyond Lipid Lowering: Potential for Clinical Benefits." *International Journal of Cardiology* 109 (1): 7–15.
- American Diabetes Association. 2008. "Standards of Medical Care in Diabetes." *Diabetes Care* 31 (Suppl 1): S12–54.
- Andrade, S. E., K. H. Kahler, and F. Frech. 2006. "Methods for Evaluation of Medication Adherence and Persistence Using Automated Databases." *Pharmacoeconomics and Drug Safety* 15 (8): 565–74.
- Aubert, R. E., J. Yao, F. Xia, and S. B. Garavaglia. 2012. "Is There a Relationship between Early Statin Compliance and a Reduction in Healthcare Utilization?" *The American Journal of Managed Care* 16 (6): 459–66.
- Benner, J. S., R. J. Glynn, H. Mogun, P. J. Neumann, M. C. Weinstein, and J. Avorn. 2002. "Long Term Persistence of Statin Therapy in Elderly Patients." *Journal of the American Medical Association* 288 (4): 455–61.
- Briesacher, B. A., S. E. Andrade, R. A. Yood, and K. H. Kahler. 2007. "Consequences of Poor Compliance with Bisphosphonates." *Bone* 41 (5): 882–7.
- Brookhart, M. A., A. R. Patrick, C. Dormuth, J. Avorn, W. Shrank, S. M. Cadarette, and D. H. Solomon. 2007. "Adherence to Lipid-Lowering Therapy and the Use of Preventive Health Services: An Investigation of the Healthy User Effect." *American Journal of Epidemiology* 166 (3): 348–54.
- Concato, J., N. Shah, and R. I. Horwitz. 2000. "Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs." *New England Journal of Medicine* 342: 187–1892.
- Congressional Budget Office. 2012. "Offsetting Effects of Prescription Drug Use on Medicare's Spending for Medical Services" [accessed on December 11, 2012]. Available at <http://www.cbo.gov/sites/default/files/cbofiles/attachments/43741-MedicalOffsets-11-29-12.pdf>
- DiMatteo, M. R., P. J. Giordani, H. S. Lepper, and T. W. Croghan. 2002. "Patient Adherence and Medical Treatment Outcomes: A Meta-Analysis." *Medical Care* 40 (9): 794–811.
- Esposito, D., A. D. Bagchi, J. M. Verdier, D. S. Bencio, and M. S. Kim. 2009. "Medicaid Beneficiaries with Congestive Heart Failure: Association of Medication

- Adherence with Healthcare Use and Costs." *The American Journal of Managed Care* 15 (7): 437–45.
- Glynn, R. J., S. Schneeweiss, and T. Sturmer. 2006. "Indications for Propensity Scores and Review of Their Use in Pharmacoepidemiology." *Basic & Clinical Pharmacology & Toxicology* 98 (3): 253–9.
- Hayward, R., D. Kent, S. Vijan, and T. Hofer. 2005. "Reporting Clinical Trial Results to Inform Providers, Payers, and Consumers." *Health Affairs* 24 (6): 1571–81.
- Humphrey, L. L., B. K. S. Chan, and H. C. Sox. 2002. "Postmenopausal Hormone Replacement Therapy and the Primary Prevention of Cardiovascular Disease." *Annals of Internal Medicine* 137 (4): 273–84.
- International Society for Pharmacoeconomics and Outcomes Research [ISPOR]. 2012. ISPOR Good Practices for Outcomes Research [accessed on November 28, 2012]. Available at http://www.ispor.org/workpaper/practices_index.asp
- Ioannidis, P. 2005. "Contradicted and Initially Stronger Effects in Highly Cited Clinical Research." *Journal of the American Medical Association* 294 (2): 218–28.
- . 2013. "Are Mortality Differences Detected by Administrative Data Reliable and Actionable?" *Journal of the American Medical Association* 309 (13): 1410–1.
- Ioannidis, P., A. Haidich, M. Pappa, N. Pantazis, S. I. Kokori, M. G. Tektonidou, D. G. Contopoulos-Ioannidis, and J. Lau. 2001. "Comparison of Evidence of Treatment Effects in Randomized and Nonrandomized Studies." *Journal of the American Medical Association* 286 (7): 821–30.
- Johnson, E. S., B. A. Bartman, B. A. Briesacher, N. S. Fleming, T. Gerhard, C. J. Kornegay, P. Nourjah, B. Sauer, G. T. Schumock, A. Sedrakyan, T. Stuermer, S. L. West, and S. Schneeweiss. 2012. *Effective Health Care Program Research Reports, No. 32. The Incident User Design in Comparative Effective Research*. Rockville, MD: AHRQ.
- Karter, A. J., M. M. Parker, H. H. Moffet, A. T. Ahmed, J. A. Schmittiel, and J. V. Selby. 2009. "New Prescription Medication Gaps: A Comprehensive Measure of Adherence to New Prescriptions." *Health Services Research* 44 (5): 1640–61.
- Leiter, L. 2005. "The Prevention of Diabetic Microvascular Complications of Diabetes: Is There a Role for Lipid Lowering?" *Diabetes Research and Clinical Practice* 68S (2): S3–14.
- Motheral, B., J. Brooks, M. A. Clark, W. Crown, P. Davey, D. Hutchins, B. C. Martin, and P. Stang. 2003. "A Checklist for Retrospective Database Studies -Report of the ISPOR Task Force on Retrospective Databases." *Value in Health* 6 (2): 90–7.
- Osterberg, L., and T. Blaschke. 2005. "Adherence to Medication." *New England Journal of Medicine* 353 (5): 487–97.
- Owens, D. K., K. N. Lohr, D. Atkins, J. R. Treadwell, J. T. Reston, E. B. Bass, S. Chang, and M. Helfand. 2009. *Methods Guide for Comparative Effectiveness Research: Grading the Strengths of a Body of Evidence When Comparing Medical Interventions*. Rockville, MD: AHRQ.
- Pittman, D. G., W. Chen, S. J. Bowlin, and J. A. M. Foody. 2011. "Adherence to Statins, Subsequent Healthcare Costs, and Cardiovascular Hospitalizations." *The American Journal of Cardiology* 107 (11): 1662–6.

- Ray, W. A. 2003. "Evaluating Medication Effects Outside of Clinical Trials: New-User Designs." *American Journal of Epidemiology* 158 (9): 915–20.
- Ross, S. D., I. E. Allem, J. E. Connelly, B. M. Korenblat, M. E. Smith, D. Bishop, and D. Luo 1999. "Clinical Outcomes in Statin Treatment Trials: A Meta Analysis." *Archives of Internal Medicine* 159 (15): 1793–802.
- Rothwell, P. 1995. "Can Overall Results of Clinical Trials Be Applied to All Patients?" *Lancet* 345 (8965): 1616–9.
- . 2005. "External Validity of Randomised Controlled Trials: 'To Whom Do the Results of This Trial Apply?'" *Lancet* 365 (9453): 82–93.
- Salas, M., D. Hughes, A. Zuluaga, K. Vardeva, and M. Lebmeier. 2009. "Costs of Medication Nonadherence in Patients with Diabetes Mellitus: A Systematic Review and Critical Analysis of the Literature." *Value in Health* 12 (6): 915–22.
- Schneeweiss, S., A. R. Patrick, T. Sturmer, M. A. Brookhart, J. Avorn, M. Maclure, K. J. Rothman, and R. J. Glynn. 2007. "Increasing Levels of Restriction in Pharmacoepidemiologic Database Studies of Elderly and Comparison with Randomized Trial Results." *Medical Care* 45 (10 Suppl 2): S131–42.
- Shadish, W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Chicago, IL: Wadsworth Publishers.
- Shang, B., and D. P. Goldman. 2007. *Prescription Drug Coverage and Elderly Medicare Spending*. NBER Working Paper No. 13358. Cambridge, MA: NBER.
- Sinzinger, H., R. Wolfram, and B. Peskar. 2002. "Muscular Side Effects of Statins." *Journal of Cardiovascular Pharmacology* 40 (2): 163–71.
- Sokol, M. C., K. A. McGuigan, R. R. Verbrugge, and R. S. Epstein. 2005. "Impact of Medication Adherence on Hospitalization Risk and Healthcare Cost." *Medical Care* 43 (6): 521–30.
- Solomon, D. H., J. A. Rassen, R. J. Glynn, J. Lee, R. Levin, and S. Schneeweiss. 2010. "The Comparative Safety of Analgesics in Older Adults with Arthritis." *Archives of Internal Medicine* 170 (22): 1968–78.
- Stuart, B. C., A. Davidoff, R. Lopert, T. Shaffer, J. S. Shoemaker, and J. Lloyd. 2011. "Does Medication Adherence Lower Medicare Spending among Beneficiaries with Diabetes?" *Health Services Research* 46 (4): 1180–99.
- Stuart, B., F. E. Loh, L. Miller, and P. Roberto. 2014. "Should Eligibility for Medication Therapy Management Be Based on Drug Adherence?" *Journal of Managed Care Pharmacy* 20 (1): 66–75.
- Tsiara, S., M. Elisaf, and D. Mikhailidid. 2003. "Early Vascular Benefits of Statin Therapy." *Current Medical Research and Opinion* 19 (6): 540–56.
- Wang, J., and P. T. Donnan. 2001. "Propensity Score Methods in Drug Safety Studies: Practice, Strengths, and Limitations." *Pharmacoepidemiology and Drug Safety* 10 (4): 341–4.
- Weitzen, A., K. L. Lapane, A. Y. Toledano, A. L. Hume, and V. Mor. 2004. "Principles for Modeling Propensity Scores in Medical Research: A Systematic Literature Review." *Pharmacoepidemiology and Drug Safety* 13 (12): 841–53.
- World Health Organization. 2003. *Adherence to Long-Term Therapies: Evidence for Action*. Geneva: WHO.

Zhang, Y., J. M. Donohue, J. R. Lave, G. O'Donnell, and J. P. Newhouse. 2009. "The Effect of Medicare Part D on Drug and Medical Spending." *New England Journal of Medicine* 361 (1): 52–61.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Tables A1–A5: Present Results from the Difference-in-Difference Models for the Statin Analysis.

Tables A6–A13 and Figures A1–A4: Present Results from the Replication Tests for ACE-Inhibitors and ARBs.