

Medicare Letters To Curb Overprescribing Of Controlled Substances Had No Detectable Impact On Providers

Adam Sacarny, David Yokum, Amy Finkelstein, and Shantanu Agrawal*

February 2016

Abstract

Inappropriate prescribing is a rising threat to the health of Medicare beneficiaries and a drain on Medicare's finances. In this study, we used a randomized controlled trial approach to evaluate a low-cost, light-touch intervention aimed at reducing the inappropriate provision of Schedule II controlled substances in the Medicare Part D program. Potential overprescribers were sent a letter explaining that their practice patterns were highly unlike those of their peers. Using rich administrative data, we were unable to detect an effect of these letters on prescribing. We describe ongoing efforts to build on this null result with alternative interventions. Learning about the potential of light-touch interventions both effective and ineffective will help produce a better toolkit for policy makers to improve the value and safety of health care.

* Sacarny: Columbia University; ajs2102@columbia.edu; Yokum: White House Social and Behavioral Sciences Team; david.yokum@gsa.gov; Finkelstein: Department of Economics, MIT, J-PAL North America, and NBER; afink@mit.edu; Agrawal: Center for Program Integrity, CMS; Shantanu.Agrawal@cms.hhs.gov. This research was presented at the Robert Wood Johnson Foundation Scholars in Health Policy Research program's annual meeting, in San Diego, California, June 4, 2015. The authors gratefully acknowledge funding through J-PAL North America from the Alfred P. Sloan Foundation and the Laura and John Arnold Foundation, as well as support from the Robert Wood Johnson Foundation. This study was approved by the Massachusetts Institute of Technology's Committee on the Use of Humans as Experimental Subjects and the Harvard Committee on the Use of Human Subjects, and was registered on the American Economic Association Randomized Controlled Trials Registry and ClinicalTrials.gov (see Note 32). The views expressed in this article represent those of the authors and not their respective organizations including the Centers for Medicare and Medicaid Services (CMS). The contents of this article were reviewed for compliance by CMS.

Inappropriate prescribing threatens patients' health and increases health care costs. A body of evidence, ranging from academic studies to investigative reports by the Government Accountability Office and the Office of Inspector General for the Department of Health and Human Services (HHS), has described overprescribing of many pharmaceuticals, including controlled substances (such as opioids), benzodiazepines, and antipsychotics.[1-4] These substances are associated with a host of adverse health consequences from accidental falls to overdose and death; their misuse also triggers costly health care use.[5-7]

The Centers for Medicare and Medicaid Services (CMS) is exploring a variety of innovative approaches to combat overprescribing behavior. This study evaluated one approach: an inexpensive intervention to affect questionable prescribing by sending an informative letter to health care providers suspected of improperly writing prescriptions for controlled substances through Medicare Part D, the prescription drug insurance program for Medicare beneficiaries. Alongside the evaluation of this initial effort, the study further considered how the letter approach could be continuously adapted and analyzed so that it may be more effective in the future.

Insurers frequently communicate with their providers to ensure billing is accurate and medically necessary. The Medicare program, for example, regularly sends billing reports to physicians and hospitals. Our study was the first attempt that we know of to rigorously evaluate an informative letter aimed at reducing potentially inappropriate medical practices. This approach was worth exploration given the existing literature showing that such letters can have effects on a wide range of outcomes, including health care outcomes such as physicians' vaccinating their patients and legal compliance outcomes such as individuals' payment of delinquent taxes.[8-10] In much of the literature, these effects are found even when the letters do not mention penalties for noncompliance, which matches our approach in this study.

Letters such as these are one potential tool in the arsenal that insurers might use to encourage provider compliance with appropriate practices. In the case of the Medicare program, these letters are aimed at reducing inappropriate prescribing behavior with dual objectives of saving costs and protecting the health of beneficiaries. They complement efforts such as audits and investigations to fight fraud and abuse. Informative letters, if effective, could offer a low-cost, collaborative approach to reducing improper prescribing behavior. They allow CMS to target a larger group of providers than would normally be practical with traditional methods like audits and investigations, an important advantage given that more than half a million practitioners are associated with Part D prescriptions for the most addictive controlled substances every year.

Medicare Part D is an ideal setting in which to test the impact of such letters. It is the largest single insurer for prescription drugs, and its data are updated often and with a minimal lag, allowing for fast evaluation. To implement this study, the White House Social and Behavioral Sciences Team facilitated a research effort with CMS and academic researchers at the Massachusetts Institute of Technology (MIT) and Columbia University. CMS used Part D data to identify “outlier prescribers”--physicians and other practitioners who prescribed vastly more controlled substances (more than 400 percent more, on average) than their peers. We then randomized these prescribers into a treatment group and a control group. The treatment group received a letter depicting their level of prescribing in comparison to their peers (see online Appendix Exhibit A1),^[11] while the control group received nothing. Using Part D administrative data, we tracked the effect of the letter on prescribing behavior during the following ninety days.

We were able to perform preliminary evaluations of the letters just months after they were sent. This article presents our full evaluation, which took less than one year. We found no evidence of an impact of the letter on prescribing behavior. Indeed, our estimates suggested a

statistically insignificant and substantively small increase in prescribing of the targeted drugs; our confidence intervals allowed us to reject effects bigger than a 1.4 percent reduction in prescribing of these drugs. However, given the low cost of this intervention and the success of similar letters in related domains like vaccinations and tax compliance,[8-10] we believe that additional randomized trials of alternative letter designs are warranted. These results have already informed a series of changes for future letters that are actively being tested, making this study part of a process that harnesses Medicare data to continuously improve CMS's efforts to lower improper payments.

Background

Inappropriate Prescribing And Schedule II Drugs

Overprescribing of pharmaceuticals has been found across a variety of substances, exposing patients--and seniors in particular--to unnecessary risks. For example, seniors often are prescribed benzodiazepines for extended periods of time even though this puts them at risk of debilitating falls,[3] and antipsychotics are often prescribed to seniors with dementia even though these drugs may increase the chance of falls and death.[4] Overprescription can raise health care expenditures as a result of the direct cost of the drugs and from the resulting avoidable health care use caused by adverse outcomes.

At the extreme, inappropriate prescribing may include outright fraud, such as taking kickbacks from patients in exchange for prescriptions or using a stolen prescribing pad to write prescriptions for drugs (or selling the prescribing pad for this purpose). But deliberate deception is not necessary for prescribing to be inappropriate. Other examples of inappropriate prescribing include physicians providing addictive drugs at the request of drug-seeking or -abusing patients without proper medical evaluation and physicians providing medications without properly

monitoring patients for adverse outcomes. Overprescribing may also result from misinformation if, for example, a doctor were influenced by biased marketing or continuing medical education to provide addictive drugs to patients who did not stand to benefit from them clinically.[12]

This study focused on inappropriate prescribing of Schedule II controlled substances, a set of medications that carry particularly large risks for patients and that policy makers widely believe are overprescribed. The Drug Enforcement Administration classifies substances on the basis of their potential for abuse and dependency; Schedule II is the highest-risk category for which a prescription is still legal. The category includes opioid pain relievers such as morphine and oxycodone (branded as OxyContin or Percocet), as well as stimulants such as amphetamines (branded as Dexedrine or Adderall) and methylphenidate (branded as Ritalin). The use and abuse of opioid pain relievers has risen dramatically since the late 1990s.[5,6] The opioid overdose death rate more than quadrupled between 1999 and 2014,[13] and more than one-third of Medicare Part D enrollees now fill an opioid pain reliever prescription each year.[14] Prescribing of these drugs is also widespread, with over 600,000 clinicians responsible for scrips for the drugs in Medicare Part D annually. However, the most frequent prescribers make up much of the total volume; the prescribers in our study accounted for about 10 percent of the Schedule II prescribing in Medicare in 2012 but represented just 0.2 percent of practitioners who wrote any prescription for these substances in the program that year.

Policy makers have become increasingly concerned that a rise in inappropriate prescribing has driven the increase in opioid pain reliever abuse. Two recent HHS reports identified pharmacies and physicians whose Schedule II prescribing practices appeared anomalous:[1,2] These reports highlighted the role of a small number of unusually high-volume prescribers and were the basis for CMS's approach to identifying prescribers of Schedule II substances. On the legislative side, the Affordable Care Act contains provisions raising penalties

for false medical claims and requiring state Medicaid and Children's Health Insurance programs to bar providers when Medicare does. Nearly all states and the District of Columbia have created or begun to roll out prescription drug monitoring programs to track the dispensing of certain drugs, facilitating the detection of inappropriate prescribing and drug abuse in real time.[15] Alongside informative letters, CMS has used its rulemaking authority in a variety of ways to curb drug abuse in the Part D program. Initiatives include giving the agency authority to bar abusive prescribers from the entire Medicare program and providing antifraud contractors with streamlined access to all prescribing records.[16]

Informative Letters

Medicare has been sending informative letters with peer comparisons to physicians regarding their Part B billing behavior since at least 2010 and similar reports to hospitals since at least 2003. The letters about Part B billing behavior are the most similar to the overprescribing letters studied here. These letters are called comparative billing reports, and they target providers who bill high volumes for a particular Medicare Part B (that is, outpatient) service. The letters are designed to educate providers and encourage them to "self-audit" to correct improper payments. To this end, they include information on how many units the target provider billed of the service and how the provider's billing compares with his or her peers (for example, other providers in that specialty or geography). Although receiving a comparative billing report is not a direct indication that a provider will be subject to an audit, it may act as a warning sign because antifraud investigators often employ similar approaches to selecting their targets.[17] We know of no rigorous evaluation of comparative billing report letters to date.

The letter we studied targeting questionable prescribers represents CMS's initial foray into expanding the comparative billing report approach to prescribing in Medicare Part D (see Appendix Exhibit A1 for a sample of the letter).[11] It depicts in several ways that the

prescriber's actions were highly dissimilar to his or her peers, using both text and graphics to show that the prescriber has supplied far more Schedule II controlled substances than peer group providers.

There are several mechanisms whereby this type of letter might induce providers to reduce their inappropriate prescribing of Schedule II substances. First, the letters may educate prescribers whose practice patterns have inadvertently drifted from medical standards. In this way, the letters could represent a form of continuing medical education, signaling that prescribers must reevaluate their treatment methods. Second, the letter may impose moral costs, by informing or reminding individuals that they have drifted away from medical norms.[9,18] Finally, the letter may serve as a signal that CMS is actively monitoring Schedule II prescribing and (although the letter does not say so explicitly) that it is willing to undertake audits or even more severe administrative actions. As a result, it may increase the prescriber's perception of the expected penalty from continuing his or her current prescribing behavior.

To our knowledge, these informative letters to outlier prescribers have not previously been evaluated. However, in other settings, randomized evaluations have found that such letters can be effective, particularly those involving peer comparisons. For example, letters showing a household's energy consumption relative to its neighbors cause people to use less energy.[19] In the legal compliance realm, a randomized evaluation in the United Kingdom sent a variety of messages to delinquent taxpayers and found the strongest effect from a statement that they were "in the very small minority of people who have not paid." [9] Several other studies in Europe and South America looked at the impact of letters on tax and fee collection and found that they improve reporting and payment behavior by individuals and firms; [10,20-22] one exception was a study that found letters raised firms' reporting of taxable income but that this was largely offset by more reporting of deductible costs.[23]

Our work also relates to the practice of providing automated feedback to physicians as a form of continuing medical education.[24] In a classic study from the 1990s that is particularly related to our context, CMS (then known as the Health Care Financing Administration) oversaw a randomized evaluation of peer-comparison letters to physicians. This intervention tested a standard letter that listed a physician's scores on several quality metrics (for example, the physician's influenza vaccination rate) against a modified letter that included comparisons of the physician's performance relative to his or her top-performing peers. The study found that adding the peer comparison to the letter raised the odds that the influenza vaccine was provided by a statistically significant 57 percent (the study's four other quality scores also showed improvements, half of which were statistically significant).[8]

Looking more broadly beyond peer-comparison letters, there is longstanding evidence that computer-based reminders and certain other methods of auditing and providing feedback can raise physicians' compliance with recommended practices.[25-27] A host of studies found clinically and economically meaningful effects of reminders, including effects on prescribing behavior (the subject of our study). In a central example from this literature, a randomized encouragement intervention provided computer-generated reminders to physicians to vaccinate their patients for influenza. The reminders induced physicians to double their influenza vaccination rates and may have reduced hospitalizations.[28-30] A recent systematic review of audit and feedback interventions found that they yielded "small but potentially important improvements" with greater effectiveness in certain contexts, such as when the targeted providers were performing poorly already.[27]

Study Data And Methods

Research Design And Implementation

In July 2014 CMS conducted an analysis to identify questionable--outlier--prescribers of Schedule II controlled substances in the Medicare Part D administrative data (analogous to an insurance claims file). (For a full accounting of this analysis, see Appendix A).[11] The analysis selected 1,525 individuals who prescribed much more of these substances than their peers (defined as providers in the same state with the same specialty) in at least two of the three years 2011, 2012, and 2013. The average prescriber from the selected group was responsible for 406 percent more prescription drug fills than comparable peers.

The idea behind this approach was that extremely high levels of prescribing most likely reflected inappropriate prescribing behavior, instead of properly monitored prescribing to patients actually in need of these substances. The threshold for identifying an outlier prescriber was set to select prescribers at or above approximately the 99.7th percentile of prescribing volume in each year among prescribers of Schedule II drugs in Medicare Part D. Consistent with the idea that this method was likely to identify prescribers engaged in questionable practices even without more complicated adjustments for patient characteristics, we observed that 21 percent of these prescribers had already been investigated for fraud or abuse by mid-2014. We view this finding as a significant cross-validation that our approach selected prescribers whose practice patterns had drifted from appropriate standards of care.

We received approval from the MIT Committee on the Use of Humans as Experimental Subjects (protocol no. 1409006595) and the Harvard Committee on the Use of Human Subjects (protocol no. IRB14-3112). In August 2014 we randomly assigned 50 percent of the 1,525 outlier prescribers to the treatment group that received a letter from CMS and the remainder of the prescribers to the control group that received no intervention (the randomization was performed in Stata software and is described in Appendix B).[11] Seven prescribers (two in the treatment

group and five in the control group) had died before the outlier analysis was conducted; therefore, they were removed from the analysis.

About 85 percent of the 1,518 non-deceased prescribers in our study population were outliers in 2013, and about 60 percent were outliers in all three years (Exhibit 1). The average prescriber was associated with about 1,444 Schedule II prescription fills in 2013, amounting to nearly \$200,000 in payments by Part D plans and beneficiaries for these drugs. Three-fifths of the prescribers were general care practitioners, one-fifth were nurse practitioners or physician assistants, and one-fifth were in the remaining six categories of physician specialists. Naturally, because of the randomization, these characteristics were very similar on average between the treatment and control arms, both including and excluding the seven deceased prescribers; on average, the treatment and control arms were statistically indistinguishable.

Peer comparison letters were mailed to the treatment group prescribers ($n = 760$) in mid-September 2014.[31] Of these letters, 131 were returned to sender. We resolved the addresses of letters that were returned and resent the letters to the new addresses in batches through mid-November.

Data

We tracked the behavior of prescribers through our access to the CMS Integrated Data Repository, the warehouse for Medicare data used by CMS and its program integrity contractors. The repository includes beneficiary enrollment information as well as administrative data on health care use in Medicare Parts A, B, C, and D. This study used the Part D data, which include records for all prescription drug fills that the program covers and note the prescriber who was responsible for each fill.

Analytical Approach

We publicly registered this trial and uploaded a plan for our analysis in January 2015, before we accessed any post-intervention study data.[32] Our approach, as stated in the plan and further described in Appendix C,[11] was to use linear regression to estimate the correlation between prescriber outcomes (Schedule II prescription fills) and assignment to the treatment group (being sent a letter). Some letters were returned to sender and delivered late, and some may not have been received at all. We analyzed the full set of providers to whom we mailed letters, which included those who may not have successfully received or opened the letter. The approach is what is often called an “intention to treat” analysis. In our case, the study revealed the effect that implementation of this type of intervention was likely to have in the real world, where address data are inevitably spotty. Our results are, therefore, policy relevant for future letters designed to fight improper payments.

Since we were using a randomized controlled trial approach, our results were unbiased estimates of the impact of our intervention, even without including control variables to try to capture potential differences between the treatment and control groups. However, controls can improve the power of our statistical analysis. Therefore, we show results both with and without controls. The analysis with a control variable controlled for the volume of Schedule II prescribing in the months before the letters were sent. This was our preferred specification since (as expected) the inclusion of the control variable improved the precision of our estimates.

We focused our analysis on the impact on total Schedule II prescription drug fills because it is through changes in prescribing volume that the letters might potentially drive improvements in patient outcomes and reductions in improper payments. We examined total Schedule II prescribing at thirty- and ninety-day time windows following the mailing of the letters. We defined this outcome as the number of Schedule II prescription fills attributed to the prescriber,

and we adjusted the fills by the days supply of the prescriptions--we counted a fill as its days supply divided by thirty, so that a thirty-day fill counted as 1 and a fifteen-day fill counted as 0.5, for example. This measure is called thirty-day-equivalent prescribing. Because it accounted for prescription duration, we selected thirty-day-equivalent prescribing over the ninety days following the mailing of the letter as our primary study outcome.

The average outlier prescriber in our study control group had 155.5 Schedule II prescription drug fills attributed over the thirty days following the mailing, and 461.1 Schedule II prescription drug fills over the ninety-day time window--more than five fills per day (Exhibit 2).

Study Results

Exhibit 3 presents our main results in table form, and Exhibit 4 presents them graphically. The lines in the graphic trace out the cumulative average prescribing for the treatment and control groups in the days following the mailing of the letters. The estimates of the effect of the letters in the columns of the table (Exhibit 3) marked “No Extra Control” equal the distances between the lines at the thirty- and ninety-day marks in Exhibit 4. In Exhibit 4, the lines for treatment and control groups appear to diverge slightly, but the regression results in Exhibit 3 show that this divergence is not statistically significant.

To improve our power to detect effects of the letters, in the columns of Exhibit 3 marked “Extra Control” we included a control variable: Schedule II prescribing in the thirty or ninety days prior to the mailing of the letters. The results remain statistically insignificant. In other words, we cannot reject the hypothesis that the letters had no effect on prescribing behavior.

Our thirty-day time window estimates indicated that the letters reduced Schedule II prescribing by a statistically insignificant 0.8 fills (0.5 percent of the average). At the ninety-day window, we estimated that the letters caused a statistically insignificant increase of 3.5 fills (0.8

percent of the average). Moreover, we were able to rule out substantial reductions in prescribing as a result of the letters. With a 95 percent confidence level, we rejected that they caused a decline of more than 6.4 fills (1.4 percent of the average) over the ninety days.

In advance of viewing the study data, we preregistered that the primary outcome was Schedule II prescription fills (adjusted for the days supply) over the ninety days following the mailing, meaning that our primary focus is on whether the letters reduce prescribing according to this metric. We were unable to detect an effect on this primary outcome. The results were unchanged even if we removed the adjustment for the days' supply of the fills (that is, counted all fills as one regardless of their duration; see Appendix Exhibit A2).[11] We also failed to detect effects in the secondary analyses we ran, including looking just at subgroups such as prescribers who were high or low outliers, or those who had or had not already been investigated for fraud.[33]

Discussion

In this study, we used a randomized controlled trial to evaluate whether letters can reduce prescribing of highly addictive substances by questionable prescribers in Medicare Part D. We were unable to detect a statistically significant effect of this intervention--the letters were associated with a (statistically insignificant) increase of 3.5 prescription fills over ninety days, or 0.8 percent. At the 95 percent confidence level, we could rule out that they lowered prescribing by more than 6.4 fills, or 1.4 percent, over that time window.

Given the low cost of this intervention, the potentially deleterious effects of inappropriate prescribing on beneficiary health and Medicare spending, and the existing track record of an impact of similar informative letters in health care and non-health care contexts, these results suggest further work is needed to investigate whether alternative informative letters can be more

effective in reducing overprescribing behavior. The current study has shown that these letters can be evaluated rigorously, cheaply, and rapidly in the Medicare context. CMS is now incorporating such evaluations as part of a process to continuously innovate, test, and improve its approach to reducing improper payments. Indeed, our study team is currently implementing additional letter-based randomized interventions targeting other substances at risk of inappropriate prescription. These subsequent interventions take an evidence-based approach to our goal of designing letters that cost-effectively target questionable prescribing. They address concerns that the initial study uncovered and consider new methods of intervening with prescribers based on academic literature.

To redesign the letters for subsequent rounds of the study, we considered the most likely reasons that we failed to detect an effect in this initial round. Broadly, one set of issues involves whether the letters reached their intended targets. A second class of issues involves whether the letters, even if appropriately targeted, were effective at altering behavior.

With the new letters, we have made a host of changes to address these concerns. One set of changes is operational: we are working across divisions in CMS to use the best provider address data possible, thereby reducing returns to sender. We are also taking advantage of the high-frequency data that CMS collects to identify the most recent outlier prescribers for targeting, instead of relying on data from previous years as in the current study.

Another set of changes takes advantage of insights from behavioral and psychological research to try to design *ex ante* more effective letters. For example, the design of a current informative letter campaign that is already in the field for questionable prescribers of quetiapine (branded Seroquel), an antipsychotic, drew on research that has found the effects of letters to become more persistent when individuals were repeatedly contacted.[19] The Seroquel letter campaign therefore follows up multiple times with prescribers to impress upon them that they are

being monitored. We have also altered the language of our letters to emphasize the negative consequences of inappropriate prescribing behavior. This change follows research that has found messages emphasizing penalties to have more dramatic effects on tax compliance than messages that mention the fairness and equity of making required payments.[20]

Conclusion

Through this ongoing collaboration, we will continuously innovate and evaluate low-cost, light-touch interventions designed to reduce inappropriate prescribing behavior and improper payments in the Medicare system. An approach that does affect prescribing or other health care provider behavior would be a useful tool for CMS, other insurers, and perhaps providers in financial risk arrangements. It could also significantly affect policy approaches to curtailing waste, abuse, and fraud.

Notes

1. Office of Inspector General. Retail pharmacies with questionable Part D billing. Washington (DC): Health and Human Services; 2012 May. (Technical Report No. OEI-02-09-00600).
2. Office of Inspector General. Prescribers with questionable patterns in Medicare Part D. Washington (DC): Health and Human Services; 2013 Jun. (Technical Report No. OEI-02-09-00603).
3. Olfson M, King M, Schoenbaum M. Benzodiazepine use in the United States. *JAMA Psychiatry*. 2015;72(2):136–42.
4. Government Accountability Office. Antipsychotic drug use: HHS has initiatives to reduce use among older adults in nursing homes, but should expand efforts to other settings. Washington (DC): GAO; 2015 Jan. (Technical Report No. GAO-15-211).
5. Centers for Disease Control and Prevention. Vital signs: overdoses of prescription opioid pain relievers--United States, 1999–2008. *MMWR Morb Mortal Wkly Rep*. 2011;60(43):1487–92.
6. Frenk SM, Porter KS, Paulozzi LJ. Prescription opioid analgesic use among adults: United States, 1999-2012. *NCHS Data Brief*. 2015;(189):1–8.
7. Hampton LM, Daubresse M, Chang HY, Alexander GC, Budnitz DS. Emergency department visits by adults for psychiatric medication adverse events. *JAMA Psychiatry*. 2014;71(9):1006–14.

8. Kiefe CI, Allison JJ, Williams OD, Person SD, Weaver MT, Weissman NW. Improving quality improvement using achievable benchmarks for physician feedback: a randomized controlled trial. *JAMA*. 2001;285(22):2871–9.
9. Hallsworth M, List J, Metcalfe R, Vlaev I. The behavioralist as tax collector: using natural field experiments to enhance tax compliance. Cambridge (MA): National Bureau of Economic Research; 2014 Mar. (NBER Working Paper No. 20007).
10. Kleven HJ, Knudsen MB, Kreiner CT, Pedersen S, Saez E. Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*. 2011;79(3):651–92.
11. To access the Appendix, click on the Appendix link in the box to the right of the article online.
12. Here is an example of the potential for biased materials to influence prescribing behavior away from best practices. Van Zee A. The promotion and marketing of OxyContin: commercial triumph, public health tragedy. *Am J Public Health*. 2009;99(2):221–7.
13. Centers for Disease Control and Prevention. Wide-ranging Online Data for Epidemiologic Research (WONDER), Multiple-Cause-of-Death file, 2000–2014 [Internet]. Atlanta (GA): CDC; 2015 [cited 2016 Feb 2]. Available from:
http://www.cdc.gov/nchs/data/health_policy/AADR_drug_poisoning_involving_OA_Heroin_US_2000-2014.pdf

14. Suzuki S. Potentially inappropriate opioid use in Medicare Part D [Internet]. Washington (DC): Medicare Payment Advisory Commission; 2014 Oct 9 [cited 2016 Feb 1]. Available from: <http://www.medpac.gov/documents/october-2014-meeting-presentation-potentially-inappropriate-opioid-use-in-medicare-part-d-.pdf>

15. Centers for Medicare and Medicaid Services. The role of a prescription drug monitoring program in reducing prescription drug diversion, misuse, and abuse [Internet]. Baltimore (MD): CMS; 2014 Jun [cited 2016 Feb 1]. Available from: <http://www.cms.gov/Medicare-Medicaid-Coordination/Fraud-Prevention/Medicaid-Integrity-Education/Downloads/prescriptiondrug-monitoring-factsheet.pdf>

16. Centers for Medicare and Medicaid Services. CMS strategy to combat Medicare Part D prescription drug fraud and abuse [Internet]. Baltimore (MD): CMS; 2014 Jan 6 [cited 2016 Feb 1]. Available from: <https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2014-Fact-sheets-items/2014-01-06-2.html>

17. Fehn AK, Colagiovanni J. Comparative billing reports: what providers should know [serial on the Internet]. St. Paul (MN): RAC Monitor; 2011 Aug 3 [cited 2016 Feb 1]. Available from: <http://www.racmonitor.com/news/3-feature-aritcles/625-comparative-billing-reports-what-providers-should-know.html>

18. Levitt SD, List JA. What do laboratory experiments measuring social preferences reveal about the real world? *J Econ Perspect.* 2007;21(2):153–74.

19. Allcott H, Rogers T. The short-run and long-run effects of behavioral interventions: experimental evidence from energy conservation. *Am Econ Rev.* 2014;104(10):3003–37.
20. Castro L, Scartascini C. Tax compliance and enforcement in the Pampas: evidence from a field experiment. Washington (DC): Inter-American Development Bank. 2013 Dec. (IDB Working Paper Series No. IDB-WP-472).
21. Fellner G, Sausgruber R, Traxler C. Testing enforcement strategies in the field: threat, moral appeal and social information. *J Eur Econ Assoc.* 2013;11(3):634–60.
22. Pomeranz D. No taxation without information: deterrence and self-enforcement in the value added tax. Cambridge (MA): National Bureau of Economic Research. 2013 Jul. (NBER Working Paper No. 19199).
23. Carrillo P, Pomeranz D, Singhal M. Dodging the taxman: firm misreporting and limits to tax enforcement. Cambridge (MA): National Bureau of Economic Research. 2014 Oct. (NBER Working Paper No. 20624).
24. This review article presents Catarina I. Kiefe and colleagues’s study (Note **Error! Bookmark not defined.**) as an example of the value of benchmarking and other forms of feedback in continuing medical education. Mazmanian PE, Davis DA. Continuing medical education and the physician as a learner: guide to the evidence. *JAMA.* 2002;288(9):1057–60.

25. Arditi C, Rège-Walther M, Wyatt JC, Durieux P, Burnand B. Computer-generated reminders delivered on paper to health care professionals; effects on professional practice and health care outcomes. *Cochrane Database Syst Rev.* 2012; 12. DOI: 10.1002/14651858.CD001175.pub3
26. Soumerai SB, McLaughlin TJ, Avorn J. Improving drug prescribing in primary care: a critical analysis of the experimental literature. *Milbank Q.* 1989;67(2):268–317.
27. There is a long history of attempting to use audits and feedback to raise quality of care; for an early systematic review and discussion of the literature, see Note 26. This article presents a more recent systematic review. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev.* 2012;6:CD00D259.
28. McDonald CJ, Hui SL, Tierney WM. Effects of computer reminders for influenza vaccination on morbidity during influenza epidemics. *MD Comput.* 1992;9(5):304–12.
29. McDonald CJ, Hui SL, Smith DM, Tierney WM, Cohen SJ, Weinberger M, et al. Reminders to physicians from an introspective computer medical record. A two-year randomized trial. *Ann Intern Med.* 1984;100(1):130–8.
30. Hirano K, Imbens GW, Rubin DB, Zhou XH. Assessing the effect of an influenza vaccine in an encouragement design. *Biostatistics.* 2000;1(1):69–88.

31. The prescribers were not notified that they were taking part in the study because such a message would need to indicate that the prescriber was an outlier, disclosing the information that we intended to test. The intervention was reviewed extensively by the Centers for Medicare and Medicaid Services to ensure that it was appropriate for the population.

32. The analysis plan is available at the study's page on the American Economic Association Randomized Controlled Trials Registry (ID: AEARCTR-0000525). Finkelstein A, Sacarny A, Yokum D. The effect of informative letters on the prescription and receipt of Schedule II controlled substances [Internet]. Nashville (TN): American Economic Association Randomized Controlled Trials Registry; 2015 Jan 20 [cited 2016 Feb 1]. Available from: <https://www.socialscienceregistry.org/trials/525>

33. We also ran subanalyses by prescriber specialty and geographic region, both of which failed to detect effects. We expanded the outcome window of Exhibit 3 to 180 days, which also did not uncover effects. Of all the analyses we ran, the only statistically significant result was the thirty-day time window prescribing for low outliers, which showed an *increase* at the 10 percent level.

Disclosures

Amy Finkelstein is a member of the Congressional Budget Office's Panel of Health Advisers. She has no relevant or material financial interests that relate to the research described in this article. Adam Sacarny, David Yokum, and Shantanu Agrawal have no relevant or material financial interests that relate to the research described in this article.

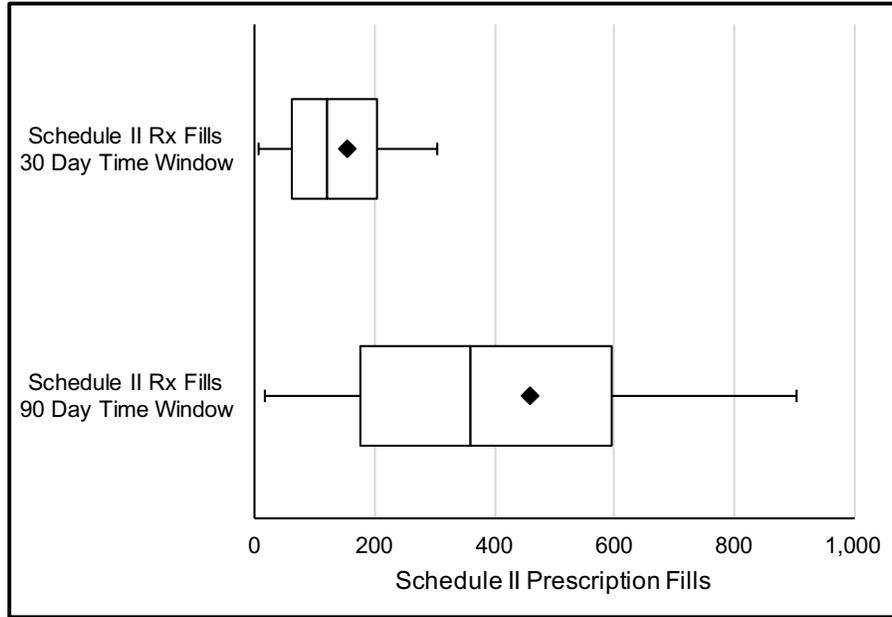
Exhibits

Exhibit 1: Summary Statistics On Outlier Schedule II Controlled Substance Prescribers From Medicare Study Group Population, 2011–13

Prescribers	Overall <i>N</i> = 1,518	Treatment group <i>n</i> = 760	Control group <i>n</i> = 758
Outliers in 2013	84.2%	83.7%	84.8%
Outliers in all three years, 2011–13	59.7%	59.7%	59.6%
Average number of Schedule II prescription fills, 2013	1,444	1,403	1,486
Average number of Schedule II prescription fills, 2011–13	4,205	4,160	4,251
Average Schedule II total dollars paid, 2013	\$198,076	\$189,914	\$206,259
Average Schedule II total dollars paid, 2011–13	\$596,553	\$586,063	\$607,070
General care practitioners	58.6%	57.8%	59.4%
Nurse practitioners or physician assistants	20.1%	18.8%	21.4%
Physician specialists	21.4%	23.4%	19.3%

SOURCE Authors' analysis of Medicare Part D administrative data on prescription drug fills in 2011–13. NOTES Prescription fills count each time a Part D beneficiary fills a Schedule II prescription associated with a prescriber in the study and are not adjusted for the days supply of the fills. Average total dollars paid refers to the total payments for the prescriptions, including payments from the Part D plan as well as out-of-pocket payments by the beneficiary, on average per prescriber. An F-test of the equality of means of all these variables between treatment and control groups is unable to reject the null of equality (p value = 0.42), as is expected when treatment is randomly assigned.

Exhibit 2: Prescription Drug Fill Outcomes For Control Group Of Outlier Schedule II Controlled Substance Prescribers in Days After Letters Sent, 2014



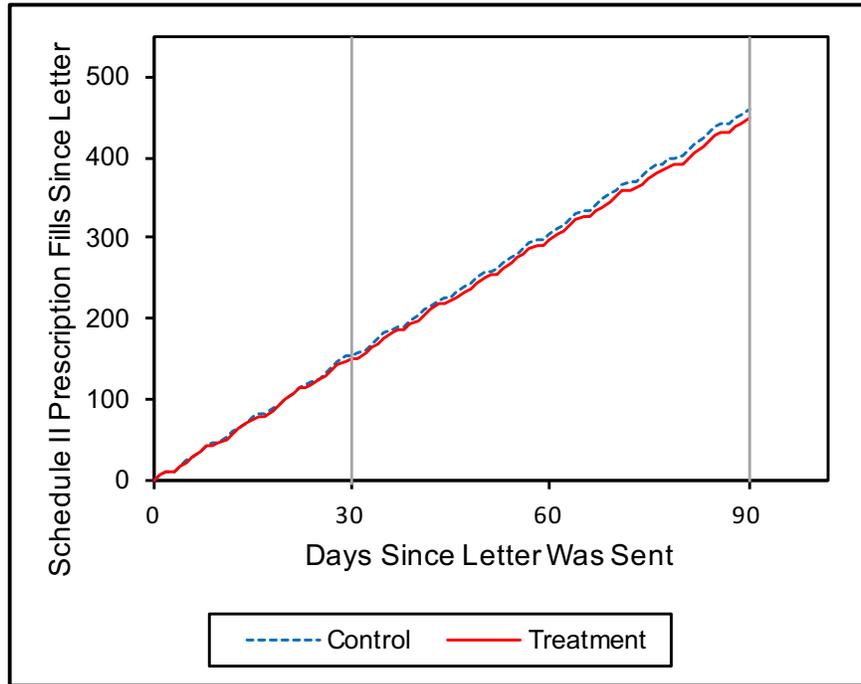
SOURCE Authors' analysis of Medicare Part D administrative data on prescription drug fills in 2014. NOTES The diamond indicates the mean prescribing level for control group prescribers, and the whiskers show the range of one standard deviation below to one standard deviation above the mean. The box spans the interquartile range of prescribing with the dividing line marking the median. Prescriptions are adjusted for the days supply of the fills. Control group ($n = 758$).

Exhibit 3: Estimated Effect Of Informative Letters On Prescribing Habits Of Outlier Schedule II Controlled Substance Prescribers, Within Thirty Days And Ninety Days After Letters Sent, 2014

Outcome time window	30 days		90 days	
	No Extra Control	Extra Control	No Extra Control	Extra Control
Effect of letter on Schedule II prescription fills	-4.41	-0.79	-11.05	3.53
Confidence interval of effect	(-19.46 to 10.64)	(-3.68 to 2.10)	(-56.07 to 33.97)	(-6.35 to 13.40)
Average Schedule II prescription fills in control group	155.5	155.5	461.1	461.1

SOURCE Authors' analysis of Medicare Part D administrative data on prescription drug fills in 2014. NOTES For specifications looking at thirty-day (ninety-day) outcomes, we used prescribing in the thirty days (ninety days) prior to the letters being sent as the extra control variable. The column headings indicate which estimates use this control variable. The confidence interval shows the margin of error for the estimates at the 95% confidence level and uses robust standard errors. None of the estimates are statistically significant at the 5 percent or 10 percent level. Prescriptions are adjusted for the days supply of the fills. $N = 1,518$

Exhibit 4: Comparison Of Schedule II Controlled Substance Prescribing In Days After Informative Letters Sent, By Treatment Group Versus Control Group, 2014



SOURCE Authors' analysis of Medicare Part D administrative data on prescription drug fills in 2014. NOTES Lines depict cumulative average prescribing for treatment and control prescribers, starting from the day the letters were sent. Prescriptions are adjusted for the days supply of the fills. The thirty- and ninety-day marks are analogous to the columns marked "No Extra Control" in Exhibit 3.

Appendix to:
Reducing Inappropriate Prescribing of Controlled Substances in
Medicare Part D: Evidence from a Randomized Intervention

February 2016

A Selection of Prescribers

First, a sample of prescribers with at least 100 schedule II prescription drug events (PDEs, or records in the Part D events file that are generated whenever patients fill prescriptions) or at least \$100,000 in total Part D payments for schedule II prescriptions was created. Any specialty that accounted for less than 1% of these prescribers was removed from the analysis. Prescribers with a specialty equal to “Specialist” were also removed as this description was considered too vague to permit analysis. The result was a sample of prescribers in 9 specialties: Anesthesiology, Emergency Medicine, General Care Prescriber (which includes General Practitioners, Family Practitioners, and Internal Medicine practitioners with no specialization), Nurse Practitioner, Orthopedic Surgery, Pain Medicine, Physical Medicine & Rehabilitation, Physician Assistant, and Psychiatry & Neurology.

Prescribers were then grouped by state and specialty (e.g. a prescriber’s peer group was other prescribers with his/her specialty in his/her state) and two outlier thresholds were calculated for each group. In order to be considered an outlier, the prescriber had to pass both thresholds. The first threshold was with respect to schedule II PDE, and it was set equal to the 75th percentile for prescribers within the state-specialty plus three times the interquartile range (called the Tukey method; see Tukey¹). The second threshold was with respect to schedule II 30-day equivalents – the total “days supply” of schedule II substances appearing in the prescribers’ PDE records, divided by 30. The threshold for 30-day equivalents was set by the same Tukey method.

When this analysis was conducted using 2011 PDE data, 1,529 outlier prescribers were identified. The 2012 data resulted in 1,656 outliers and the 2013 data resulted in 1,803 outliers. 1,525 prescribers were outliers in at least two of the three years, and these prescribers became the study sample.

B Re-Randomization Procedure

We pre-specified a re-randomization procedure that would automatically re-randomize the prescribers if an allocation failed a balance test based on observable prescriber characteristics.

The re-randomization procedure was as follows:

1. Random values are generated. Prescribers with values less than the median are assigned to the control group. All other prescribers are assigned to the treatment group.
2. We test balance for a set of covariates using the Mahalanobis distance. This balance criterion is recommended in Morgan and Rubin,² who note that in the two group case it is equivalent to a MANOVA F test. Since we have two groups, treatment and control, we implement the test using MANOVA. The covariates are:
 - (a) Outlier in 2013 (indicator)
 - (b) Outlier in 2012 (indicator)
 - (c) Outlier in 2011 (indicator)
 - (d) Schedule II PDE Count, 2013
 - (e) Schedule II PDE Count, 2012
 - (f) Schedule II PDE Count, 2011
 - (g) Schedule II Total \$ Paid from Part D, 2013
 - (h) Schedule II Total \$ Paid from Part D, 2012
 - (i) Schedule II Total \$ Paid from Part D, 2011
 - (j) Address in Census Northeast (CT, ME, MA, NH, RI, VT, NJ, NY, PA) (indicator)

(k) Address in Census Midwest (IL, IN, MI, OH, WI, IA, KS, MN, MO, NE, ND, SD)
(indicator)

(l) Address in Census West (AZ, CO, ID, MT, NV, NM, UT, WY, AK, CA, HI, OR,
WA) (indicator)

(m) Specialty is General Care Practitioner (indicator)

(n) Specialty is Nurse Practitioner or Physician Assistant (indicator)

(o) Specialty is Anesthesiology or Pain Medicine or Physical Medicine & Rehabilitation
(indicator)

3. If the p-value of the F test is < 0.4 , return to step 1 and restart the procedure. Otherwise accept the randomization.

The initial randomization yielded a p-value of 0.61, so the allocation was accepted and no re-randomization was performed.

C Regression Equation

The regression analyses that we discuss in the study can be represented by the following equation:

$$y_d = \rho \cdot treat_d + Z_d\Gamma + \eta_d$$

Where d indexes prescribers, y_d is the outcome (e.g. Schedule II prescriptions written), $treat_d$ is an indicator for assignment to the treatment group, Z_d is a set of control variables, and η_d is an error term. ρ , the coefficient of interest, is the effect of being assigned to the treatment group on the outcome.

Exhibits (Additional for Appendices)

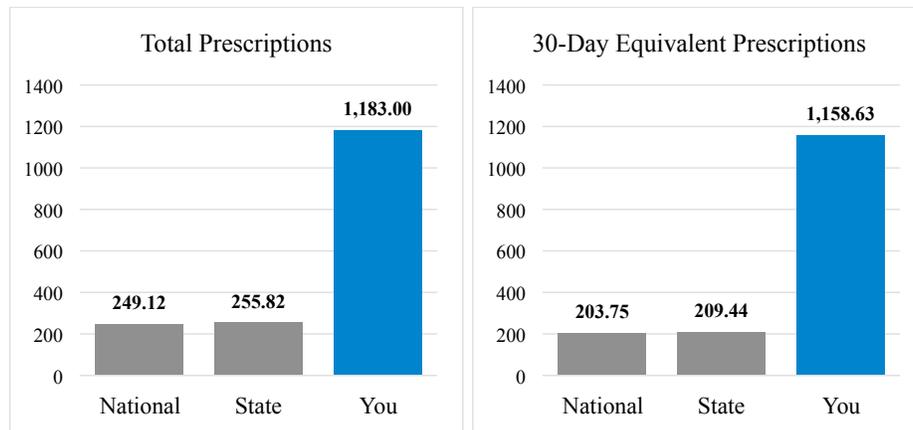
Department of Health & Human Services
7500 Security Boulevard, Mail Stop AR-18-50
Baltimore, Maryland 21244-1850



September 5, 2014

Pat Q. Provider MD
1234 Main St
Columbia, MD 21045
NPI: 1234567890
Specialty: General Care Practitioner

Re: You prescribed 362% MORE Schedule II controlled substances than your peers.



Dear Dr. Provider,

The figures above display the total count (left) and 30-day equivalent (right) of your Schedule II prescribing, compared to the national and state averages of those within your specialty. As can be seen, you prescribed far more – 362% more – than similar specialists within your state.

We hope that you will use the information provided to see if your high prescribing level is appropriate for your patient population. Read on for more information about the methodology used to analyze your prescribing behavior, and to learn what actions to take next.

Sincerely,



Medicare Program Integrity Group

Sensitivity of Estimated Effect of Letters on Schedule II Prescribing to Days Supply Adjustment

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Rx Fills (Unadjusted for Days Supply)				Rx Fills (Adjusted for Days Supply)			
Outcome Time Window	30 Days		90 Days		30 Days		90 Days	
Effect of Letter on Schedule II Prescription Fills	-4.936 (8.243)	-1.277 (1.580)	-13.81 (24.58)	2.543 (5.530)	-4.407 (7.679)	-0.794 (1.474)	-11.05 (22.97)	3.525 (5.039)
Lagged Dep Variable		0.972*** (0.0139)		0.983*** (0.0125)		0.980*** (0.0121)		0.994*** (0.0127)
Prescribers	1,518	1,518	1,518	1,518	1,518	1,518	1,518	1,518
R ²	0.000	0.964	0.000	0.950	0.000	0.964	0.000	0.952
Avg Outcome in Control Group	171.2		507.2		155.5		461.1	

SOURCE Authors' analysis of Medicare Part D administrative data on prescription (Rx) drug fills in 2014 NOTES This table shows how the results of Exhibit 3 are affected by adjusting the prescribing measure by the days supply of the fills. Columns 1-4 do not adjust prescribing for the days supply, while columns 5-8 make this adjustment (matching Exhibit 3 and the main text). Heteroskedasticity-robust standard errors in parentheses. The lagged dependent variable is the amount of prescribing (unadjusted for days supply in columns 2 and 4, adjusted for days supply for columns 6 and 8) in the 30 or 90 days (whichever the time window of the outcome variable) immediately prior to the mailing of the letters. *** p<0.01, ** p<0.05, * p<0.1

Appendix Exhibit A2

¹ Tukey JW. Exploratory Data Analysis. Addison-Wesley Publishing Company. 1977.

² Lock Morgan K, Rubin DB. Rerandomization to improve covariate balance in experiments. Ann Stat. 2012; 40(2): 1263–1282.