

Using Medicare Data for Comparative Effectiveness Research: Opportunities and Challenges

Vicki Fung, PhD; Richard J. Brand, PhD; Joseph P. Newhouse, PhD;
and John Hsu, MD, MBA, MSCE

Information on the safety, effectiveness, and value of medical care requires detailed clinical data from large numbers of patients receiving care in real-world settings. With the introduction of Medicare Part D prescription drug benefits in 2006, Medicare began to collect information on the use of prescription drugs for more than 27 million beneficiaries.¹ Previously, this information was not widely available for Medicare beneficiaries, although drug use data were available for dual-eligible Medicare beneficiaries via Medicaid claims. The more comprehensive collection of drug use data allows for linkage with previously available information from Parts A and B including inpatient and outpatient diagnoses, and major clinical events for millions of beneficiaries, including many persons over the age of 65 years. These data provide a promising resource for assessing the comparative effectiveness of many types of care across a range of settings and geographic areas in the United States.

Because Medicare collects data for payment and administrative purposes and not for research, there are several limitations to using these observational data. The care beneficiaries receive will vary depending on where they live, the types of Medicare plans they choose, and their physicians and hospitals. In other words, there may be factors associated with both the care beneficiaries receive and the outcomes of care; these factors confound assessments of care effectiveness and limit the validity of simple comparisons. However, exploiting certain program aspects and ongoing natural experiments within the Medicare program can mitigate some biases associated with purely observational data.

In this review article, we discuss these strengths and limitations of using Medicare data for comparative effectiveness research (CER) and propose policy recommendations for improving the usefulness of these data for patients, providers, and policy makers.

COMPARATIVE EFFECTIVENESS RESEARCH

There is a profound need for more evidence to guide clinical and policy decisions on drug treatments, devices, interventions, care delivery, payment models, and delivery systems. For example, while there is arguably substantial trial evidence

supporting the use of many prescription drugs, this evidence often provides limited guidance for actual clinical decisions.

In this article

Take-Away Points / p490
www.ajmc.com
Full text and PDF

Background: With the introduction of Part D drug benefits, Medicare began to collect information on diagnoses, treatments, and clinical events for millions of beneficiaries. These data are a promising resource for comparative effectiveness research (CER) on treatments, benefit designs, and delivery systems.

Objective: To explore the data available for researchers and approaches that could be used to enhance the value of Medicare data for CER.

Challenges and Opportunities: Using currently available Medicare data for CER is challenging; as with all administrative data, it is not possible to capture every factor that contributes to prescribing decisions and patients are not randomly assigned to treatments. In addition, Part D plan selection and switching may influence treatment decisions and contribute to selection bias. Exploiting certain program aspects could address these limitations. For example, ongoing changes in Medicare or plan policies and the random assignment of beneficiaries with Part D low-income subsidies into plans with different formularies could yield natural experiments.

Policy Implications: Refining policies for time to data release, provision of additional data elements, and linkage with more beneficiary-level information would improve the value and usability of these data. Improving the transparency and reproducibility of findings, and potential open access for qualified stakeholders are also important policy considerations. Data needs must be reconciled with current policies and goals.

Conclusions: Medicare data provide a rich resource for CER. Leveraging existing program elements, combined with some administrative changes in data availability, could create large data sets for evaluating treatment patterns, spending, and coverage decisions.

(Am J Manag Care. 2011;17(7):489-496)

For author information and disclosures,
see end of text.

Take-Away Points

Medicare now collects information on diagnoses, treatments, prescription drug use, and clinical events for millions of beneficiaries. These data are a promising resource for comparative effectiveness research (CER) on treatments, benefit designs, and delivery systems; however, there are a number of challenges to using these data for CER.

- We explore the data available for researchers and approaches that could be used to enhance the value of Medicare data for CER.
- Leveraging existing program elements, combined with some administrative changes in data availability, could create large data sets for evaluating treatment patterns, spending, and coverage decisions.

There are 2 main approaches for developing comparative effectiveness evidence: (1) clinical trials, including randomized clinical trials (RCTs) and pragmatic trials, and (2) studies using observational data from actual practice. While double-blinded RCTs represent the gold standard for generating clinical evidence, they have a number of practical limitations. Specifically, trials have historically compared single drugs with placebo rather than with existing alternative drugs, or examined them in combination with commonly used drug regimens. Trials are done under rigorous experimental conditions (efficacy) rather than real-world situations (effectiveness), and are not designed to evaluate costs or rare adverse events. Randomized clinical trials also tend to be expensive and examine relatively short-term effects. Moreover, trials may have limited generalizability to specific subgroups of patients (eg, the elderly, racial/ethnic minorities, those with severe diseases) because they tend to target relatively homogeneous patient groups rather than the broader mix seen in actual practice.^{2,4} Pragmatic trials attempt to overcome some of these limitations by focusing on more heterogeneous groups of patients and by evaluating effectiveness under routine care; however, existing evidence from pragmatic trials is in short supply and funding for these types of studies is limited.^{5,6}

Studies using observational, longitudinal data sets could provide complementary information that addresses many of the limitations associated with RCTs. For example, Medicare collects information on millions of beneficiaries and allows for linkages across a range of claims data, including inpatient, outpatient, and prescription drug data. Having a large sample of individuals is critical for ensuring adequate statistical power for studying rare conditions or specific patient subgroups. Use of observational data (including Medicare data), however, requires consideration of numerous factors such as the types of Medicare plans for inpatient, outpatient, and drug services; coverage/cost sharing for treatments; availability of physicians and hospitals; and clustering of patients by physician. In addition, there can be variations in practice patterns across geographic areas. Examinations of drug use within Part D should consider these various levels of analysis and account for factors that could affect drug use, adherence, and ultimately outcomes, including a range of patient, provider, and plan-level

characteristics. In short, the strengths and limitations of both clinical trials and observational data analyses should be considered carefully when evaluating the value of these approaches for addressing specific questions.

The following sections describe relevant structural aspects of the Medicare Part D program, the Medicare data available for researchers, and potential approaches that could be used to create quasi-experiments and enhance the value of historical Medicare data for CER.

MEDICARE PART D PROGRAM

Medicare currently collects diagnostic and treatment information through 4 programs: Part A (inpatient), Part B (outpatient), Part C (Medicare Advantage, which includes medical information for beneficiaries enrolled in managed care organizations), and Part D (prescription drugs). Part D is administered by private plans either as stand-alone Prescription Drug Plans (PDPs) that supplement traditional Medicare or Medicare Advantage Prescription Drug (MAPD) plans that bundle Part A, B, and D benefits. Part D is a voluntary benefit; in 2009 about 27 of 45 million Medicare beneficiaries were enrolled in a Part D plan, including 9.6 million low-income beneficiaries who received additional premium and cost-sharing subsidies from Medicare. The Centers for Medicare & Medicaid Services (CMS) randomly assigns low-income subsidy beneficiaries who have not chosen a Part D plan to qualified stand-alone drug plans.⁷

Beneficiaries choose their Part D plans; these plans have some autonomy in determining their benefit structures and formulary drug lists provided they meet basic Medicare requirements. For example, all plans must offer benefits at least as generous as the defined standard. Medicare also requires that plans cover at least 2 drugs within a therapeutic class. However, plans can determine coverage for specific drugs within a class, as well as tier placement and utilization management requirements. The use of utilization management tools such as prior authorization has grown since Part D's introduction; these tools are most often used for drugs that are newer, more expensive, or more risky, with greater potential for adverse effects or with less available evidence on the possible benefits or harms.⁸

MEDICARE DATA CURRENTLY AVAILABLE FOR RESEARCH

Since the introduction of the drug benefits, Part D plans have submitted detailed information on prescription drug events for

all Part D beneficiaries to CMS. Researchers can apply for access to Research Identifiable Files that include beneficiary-level information on Part A, B, and D claims. Proposals are reviewed by the CMS privacy board; if approved, researchers must sign a Data Use Agreement specifying the terms of use, including the destruction or return of data at the study's end.⁹ The final rule permitting release of the newly available Part D data to researchers was issued in May 2008. In addition to usual protections for beneficiary privacy, this rule includes additional protections for commercially sensitive plan information. Since the initial release of Part D data, CMS has rolled out an increasing number of data elements and linkages, and is continuing to assemble supplemental data files.

Part D Event Data

Table 1 outlines the available Part D research data files.¹⁰ The primary data source for Part D drug utilization is the Part D Event (PDE) files, which are currently available for 2006-2008. These files contain detailed information on each drug event for PDP and MAPD plan beneficiaries, and encrypted beneficiary, pharmacy, prescriber, and plan identifiers that allow linkage with other files such as inpatient and outpatient claims data, and Part D plan characteristics files. The PDE contains information on each drug dispensed including the National Drug Code, the quantity dispensed, and days of supply, allowing for the examination of therapy adherence and persistence based on dispensing data, which have been previously validated.^{11,12}

The PDE data also capture cost information such as total drug costs and patient payments. These data allow for examination of variation in spending patterns and cost-of-care analyses. The PDE data also specify the benefit phase during which each prescription was filled (eg, deductible or initial coverage phase) based on the benefit structure implemented by each beneficiary's plan, which affects costs for both patients and payers. This file also includes plan-specific information on the formulary coverage for each drug dispensed, including the tier and utilization management requirements. Plan-level information on formulary and benefit structures can be valuable for identifying quasi-experiments or instrumental variables for statistical analyses.

Beneficiary Summary Files

The PDE data can be linked with information on individual beneficiary characteristics. The Beneficiary Summary File contains beneficiary-level information on basic demographics, including age, sex, race, and geographic location. This file also describes beneficiaries' months of enrollment in Parts A, B, C, and D, including the type of coverage (eg, retiree, Part D stand-alone plan, Part D Medicare Advantage plan),

dual-eligibility status, and whether they are receiving the Part D low-income subsidy. The Beneficiary Annual Summary File contains additional information on patients' inpatient diagnosis-related groups, as well as 2 sets of chronic condition flags.

Supplemental Part D Files

Since the Part D data rule was originally issued, CMS has rolled out an increasing number of data elements that supplement the PDE and beneficiary data (Table 1). The Drug Characteristics File includes drug names (generic and brand) and strength and dosage form information by the National Drug Code. Plan characteristics files can be linked to the PDE data using the encrypted plan identifier to examine detailed information on plan type, cost-sharing levels, premium information, and service area. The Pharmacy Characteristics File includes information on the type of pharmacy where beneficiaries filled their prescriptions. Lastly, the Prescriber Characteristics File contains information on the prescriber's specialty, credentials, and geographic location.

Medicare Part A and B Data

Part D data can also be linked to traditional Medicare data on beneficiaries' other medical claims (eg, hospital, skilled nursing facility, hospice, physician). For example, the Inpatient Standard Analytic File includes claims for inpatient stays, including diagnosis and procedure codes (*International Classification of Diseases, Ninth Revision, Clinical Modification [ICD-9-CM]*), Diagnosis Related Groups (a classification system used for prospective Medicare payments), date, facility, and cost information. The Carrier File contains claims for noninstitutional providers, largely physicians. It also contains diagnosis and procedure codes (ICD-9-CM and CMS Healthcare Common Procedure Coding System codes), reimbursement amounts, and provider identifiers.

Medicare Advantage (Part C) Data

Information at the beneficiary level on inpatient and outpatient service use and diagnoses within the Medicare Advantage program is not currently available to researchers. Monthly enrollment information is available in the Beneficiary Summary File so beneficiaries' transitions between traditional fee-for-service Medicare and Medicare Advantage can be tracked. Prior to the introduction of Part D, beneficiaries could switch between traditional Medicare and Medicare Advantage plans monthly. In 2006, beneficiaries were restricted to switching only during the first 6 months of the year, and from 2007 onward, the first 3 months of the year.

Additional Linkable Data Sets

Medicare claims data also can be linked with Medicaid

■ **Table 1.** Part D Research Identifiable and Supplemental Data Files^a

File	Years Available	Types of Data Available	Value for Comparative Effectiveness Research
Part D Event data	2006-2008	<ul style="list-style-type: none"> • Encrypted identifiers (beneficiary, plan, prescriber, pharmacy) • Drug dispensed (eg, NDC, days of supply) • Benefit phase (eg, deductible, coverage gap, catastrophic coverage period) • Costs (eg, patient payments, Part D covered costs, gross drug costs) • Formulary coverage (eg, tier, prior authorization) 	<ul style="list-style-type: none"> • IDs allow for linkages with other Medicare data RIFs and supplemental files • Detailed information on drugs dispensed; ability to calculate adherence/persistence • Ability to conduct economic analyses and evaluate spending patterns • Variations in formulary coverage over time and across plans can be used to identify natural experiments and instrumental variables
Beneficiary Summary File	1999-2008	<ul style="list-style-type: none"> • Encrypted identifiers (beneficiary, Part D plan) • Geographic area (eg, state, county, ZIP code) • Basic demographics (eg, date of birth, sex, age, race) • Entitlement flags (eg, ESRD, disabled flags) • Coverage indicators for Parts A, B, C • Death indicator • Part D coverage flags (eg, Part D coverage, retiree coverage, dual-eligible status) 	<ul style="list-style-type: none"> • Monthly Part D plan identifiers provide information on enrollment and switching between plans • Geographic indicators allow for examination of and control for regional variation in treatment patterns • Basic demographic and entitlement information can be used for risk adjustment • Monthly enrollment information enables assessment of switching between TM and MA, start and termination of coverage • Mortality data important for censoring or as an outcome
Beneficiary Annual Summary File	1997-2006; 2007-2008 ^b	<ul style="list-style-type: none"> • Encrypted identifiers (beneficiary) • Geographic area (eg, ZIP code, metro status) • Basic demographics • Entitlement flags • Medicare enrollment • Diagnosis Related Groups for inpatient stays • Chronic condition category flags • Annual reimbursement amounts • Clinical event counts 	<ul style="list-style-type: none"> • Geographic indicators allow for examination of and control for regional variation in treatment patterns • Basic demographic and entitlement information can be used for risk adjustment • Clinical data can be used for study cohort identification, risk adjustment
Part D Drug Characteristics File	2006-2008	<ul style="list-style-type: none"> • Encrypted identifiers (Part D event number, beneficiary) • Drug name (brand, generic) • Drug strength, dosage form 	<ul style="list-style-type: none"> • Detailed information on drugs dispensed
Plan Characteristics File	2006-2008	<ul style="list-style-type: none"> • Encrypted identifiers (plan, formulary ID) • Plan type (eg, PDP, MA-HMO, MA-PPO), benefit type (eg, standard, enhanced) • Cost-sharing structures and levels (eg, gap threshold, gap coverage, initial coverage cost sharing, tiers) • Premium data (eg, Part D premium, LIS premium, benchmark plan flag) • Service area (eg, state, county, Medicare Part D region) 	<ul style="list-style-type: none"> • Variations in benefit design over time and across plans can be used to identify natural experiments and instrumental variables
Pharmacy Characteristics File	2006-2008	<ul style="list-style-type: none"> • Encrypted identifiers (pharmacy) • State • Pharmacy type (eg, retail, LTC, mail order) 	<ul style="list-style-type: none"> • Detailed information on pharmacy type
Prescriber Characteristics File	2006-2008	<ul style="list-style-type: none"> • Prescriber identifier • Prescriber taxonomy codes (eg, specialty) • Prescriber practice state • Prescriber credentials 	<ul style="list-style-type: none"> • Ability to adjust for clustering by prescriber and prescriber characteristics

ESRD indicates end-stage renal disease; LIS, low-income subsidy; LTC, long-term care; MA, Medicare Advantage; MA-HMO, Medicare Advantage HMO; MA-PPO, Medicare Advantage preferred provider organization; NDC, National Drug Code; PDP, Prescription Drug Plan; RIF, Research Identifiable File; TM, traditional Medicare.

^aSource: reference 10.

^bChanges in file between 2006 and 2007; 2007 data elements reported.

data for dual-eligible beneficiaries, as well as data sets focused on specific subpopulations or collected via surveys, such as the Surveillance Epidemiology and End Results Cancer Registry, the Long Term Care Minimum Data Set, the Home Health Outcome and Assessment Information Set, the Medicare Current Beneficiary Survey, and the Health and Retirement Survey. While the available sample size is more limited when linking to these data sources, they can provide a richer set of sociodemographic, health, and clinical characteristics than is available in the claims data alone.

RESEARCH DESIGN AND ANALYTIC STRATEGIES

A substantial body of literature describes the limitations of using observational data for CER, as well as potential strategies for addressing these limitations.¹³⁻¹⁷ Data and measurement quality, the formation and stability of comparison groups, and methods to deal with crossover effects or switching between therapies are particularly relevant for observational data studies and have been described elsewhere.¹⁸⁻²⁰

Despite the range of information available on Medicare beneficiaries' drug treatments and diagnoses, simple analyses using currently available Medicare data are likely to yield biased results (Table 2). Because patients are not randomly assigned to treatments, there may be important differences between the comparison groups with respect to the factors that contributed to patients receiving different treatments or therapies, such as disease severity, risk factors, and comorbidities (eg, confounding by indication).^{21,22} The lack of randomization creates analytic challenges because it is not possible using administrative data to capture all factors that contribute to physicians' prescribing decisions.

Plan selection and switching within the Part D program beneficiaries also create challenges. Beneficiaries choose to enroll in plans with varying characteristics that can influence treatment decisions, such as different delivery systems (eg, Medicare Advantage vs PDP), benefit designs, and formularies. Beneficiaries are encouraged to choose plans based on these characteristics to minimize treatment disruptions and lower out-of-pocket drug costs, so there is likely to be substantial self-selection into Part D plans.^{23,24} For example, there was evidence of substantial adverse selection into plans that offered full coverage during the standard coverage gap in 2006, prompting these plans to leave the market the following year.²⁵ In addition, the availability of employer-sponsored drug coverage for current or retired employees also influences the characteristics of beneficiaries choosing to enroll in Part D plans; an estimated 8.3 million beneficiaries received retiree drug coverage in 2010. Beneficiaries can also switch to

and from plans during the open enrollment phase of each year (eg, based on changes in clinical need). Beneficiaries' plan choice and geographic location may have important implications for their access to different treatment alternatives.^{26,27} While a number of patient, provider, and plan characteristics are available in Medicare data sets, simply adjusting for these factors in a multivariate model is unlikely to sufficiently address selection differences between patient groups that contribute to treatment decisions, plan choice, or health outcomes.

Structural Aspects of the Medicare Program and Natural Experiments

Certain features of the Medicare program present opportunities to design studies that mitigate the traditional limitations of observational data analyses and improve the relevancy of study findings. For example, ongoing changes in Medicare or Part D plan policies may provide opportunities to examine natural experiments. Policy changes that are determined at higher levels (eg, program, plan) are likely to result in changes in treatment patterns that are unrelated to underlying beneficiary characteristics. For example, within the Part D program, some plans may change their formulary requirements from year to year, while other plans' formularies remain stable. These variations could be used to create quasi-experimental treatment and control groups if researchers can identify comparable plans with stable enrollment.²⁸

Instrumental variables are another approach for accounting for unobserved differences between groups. Instrumental variables are related to treatment assignment, but not to other patient risk factors that are associated with the treatment choice or health outcome of interest. Numerous studies have demonstrated that the use of instrumental variables can decrease the bias associated with unmeasured confounders^{15-17,29}; however, this approach is dependent on finding a high-quality instrument, which can be difficult. Millions of beneficiaries who receive Part D low-income subsidies are randomly assigned into qualified stand-alone drug plans. Differences in plan formularies for these beneficiaries could be associated with variations in drug use that are unrelated to underlying risk factors, given that patients were randomly assigned to plans. Thus, plan formulary coverage for these beneficiaries represents a potential high-quality instrument because it removes the association between plan choice and patient risk factors.

Importantly, all of these approaches may not provide information that clinicians want when treating patients (ie, what is the best medical strategy for an individual patient vs the average or marginal effect for a group of patients). For example, the instrumental variable approach generally provides

■ **Table 2.** Medicare Program Features and Analytic Challenges

Program Characteristic	Analytic Challenge
1. Beneficiaries select their Medicare plan a. TM vs MA b. Supplemental insurance (Part B) vs none (if in TM) c. MA plan type, such as HMO, preferred provider organization, private fee-for-service (if in MA) d. Part D plan selection—variable premiums, cost sharing, formulary coverage, pharmacies	Selection bias
2. Beneficiaries can switch their plan choices during the calendar year a. Before 2006: Monthly switching allowed, no lock-in period b. 2006: Switching allowed between January and June, then locked in for remainder of year c. 2007 Onward: Switching allowed between January and March, then locked in for remainder of year	Selection bias
3. Beneficiary access to care depends on both plan and location a. Location (eg, rural vs urban), availability of services, volume for physicians and hospitals, regional variations in care patterns b. Plan (eg, formularies, physicians within network or accepting Medicare payment, disease management services, utilization management activities)	Different choice sets
4. Diagnostic and treatment decisions are based on clinical judgment a. Variable experiences and preferences across clinicians b. Decisions about treating and specific treatment choices depend on judgment of risk such that treated patients will differ from untreated patients, and patients receiving treatment A will differ from those receiving treatment B, in meaningful but often unmeasured ways	Confounding by indication

MA indicates Medicare Advantage; TM, traditional Medicare.

estimates of a marginal effect (eg, the effect on outcomes if the probability of receiving a treatment went from X percent to X + Y percent). However, if properly interpreted such information could still be valuable for informing policy or coverage decisions.

RECOMMENDATIONS/POLICY IMPLICATIONS

The release of Part D event data and related files was an important step; however, additional changes in the data policy could make Medicare data more useful for CER.

Release data faster. There currently is an approximate 21-month time lag between care received by beneficiaries and the availability of Part A, B, and D data to researchers. This lag hinders the ability to compare newer therapies, for which effectiveness and safety information could have the most value for patients and providers. New therapies also tend to be more expensive than older ones, and information on their use has value for policy makers making coverage decisions.³⁰

Release all of the data elements necessary to identify the care options available to individual beneficiaries. These data including information on drug formularies, cost sharing (including drug tiers), and linkages between patients, physicians, hospitals, and plans. These factors can vary substantially across

beneficiaries and influence treatment choices and adherence. These data could also be useful for developing valid instruments or identifying natural experiments. They can be released while protecting the identities of the patients, physicians, hospitals, and plans involved (eg, via coded identifiers and data use agreements). There have been promising developments in this area, particularly with respect to linkages across data sets (eg, establishment of the Chronic Condition Data Warehouse) and increases in available Part D data elements.

Release Part D plan assignment data for low-income subsidy beneficiaries. Medicare randomly assigns low-income subsidy beneficiaries to qualified plans, but a number choose other plans or remain in the same plan even if the plan no longer qualifies as a benchmark plan in the following year. The group that truly is randomly assigned to its Part D plan represents the most promising natural experiment; however, data on random assignment versus choice are not currently available to researchers.

Link data on care received for individual beneficiaries across all Medicare programs, including Medicare Advantage. About 10 million beneficiaries are now enrolled in Medicare Advantage plans, and there has been substantial growth in recent years.³¹ Providing researchers with comprehensive information on medical care received for both fee-for-service and Medicare Advantage beneficiaries would increase

the generalizability of comparisons and provide valuable information on how delivery systems and physician/hospital networks affect care. These data are particularly relevant to questions concerning medical spending, delivery systems, and workforce needs.

Create specific Medicare data sets for top comparative effectiveness priority topics identified by the Institute of Medicine.³² The Centers for Medicare & Medicaid Services is currently conducting a pilot project to create CER public use files for a 5% sample of Medicare beneficiaries.^{33,34} As part of this effort, they could also consider creating supplemental data sets tailored for CER priorities. For example, one priority topic is the comparison of treatment strategies for atrial fibrillation, including surgery, catheter ablation, and prescription drugs. The creation and release of a single data set including all Medicare beneficiaries with a diagnosis of or treatment for atrial fibrillation would permit numerous researchers to more rapidly evaluate treatments both overall and for clinically important subgroups of patients, apply innovative strategies to identify natural experiments or strong instruments, confirm or refute other evaluations, and identify remaining questions. Including detailed information on services available to patients (eg, treating surgeon, electrophysiologist, hospital, outpatient physicians, anticoagulation clinics, drug formularies, plans) could help account for the actual choices available to individual beneficiaries living in different geographic areas, potential volume-related outcomes, and other relevant clinical information. Incomplete data files could lead to less useful data sets or biased findings. The Centers for Medicare & Medicaid Services should convene groups of stakeholders, researchers, and CER analytic experts to refine the data elements and priorities. The Centers for Medicare & Medicaid Services could provide these data at moderate to no cost, but require researchers to apply to use these data sets as they currently apply to use other identifiable Medicare data, outlining their approach, methods, and mechanisms to protect the confidentiality of patient, physician, and plan identities, as well as any proprietary data. Any fees collected as part of this effort could be used to help build the infrastructure needed for more timely release of data. This approach would improve the transparency of comparative effectiveness research.

CONCLUSION

The Medicare program provides a rich data source for evaluating the comparative effectiveness of drug treatments. Simple analyses using currently available information, however, are unlikely to provide useful information. Leveraging program elements, combined with some changes in data

availability, could improve the value of these data sets and the transparency of CER in evaluating treatment patterns, spending, and coverage decisions.

Author Affiliations: From Mid-Atlantic Permanente Research Institute, Mid-Atlantic Permanente Medical Group (VF), Rockville, MD; Department of Epidemiology and Biostatistics (RJB), University of California, San Francisco; Department of Health Care Policy (JPN, JH), Harvard Medical School, Boston, MA; Department of Health Policy and Management (JPN), Harvard School of Public Health, Boston, MA; Kennedy School of Government (JPN), Harvard University, Cambridge, MA; and Mongan Institute for Health Policy (JH), Massachusetts General Hospital, Boston, MA.

Funding Source: The National Institute on Aging (R01 AG029316), the Commonwealth Fund, and the Alfred P. Sloan Foundation provided funding for the study.

Author Disclosures: Dr Newhouse reports board membership with Aetna and also reports holding stock in the company. Dr Hsu reports receiving grants from the National Institutes of Health. The other authors (VF, RJB) report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (VF, RJB, JPN, JH); analysis and interpretation of data (VF, JPN); drafting of the manuscript (VF); critical revision of the manuscript for important intellectual content (VF, RJB, JPN, JH); statistical analysis (RJB); obtaining funding (JH); and supervision (JH).

Address correspondence to: Vicki Fung, PhD, Kaiser Permanente, 2101 East Jefferson St, Rockville, MD 20852. E-mail: Vicki.Fung@kp.org.

REFERENCES

1. **The Kaiser Family Foundation.** *The Medicare Prescription Drug Benefit: Fact Sheet.* Washington, DC: The Henry J. Kaiser Family Foundation; October 2010. <http://www.kff.org/medicare/7044.cfm>. Accessed June 8, 2011.
2. **McNeil BJ.** Shattuck Lecture—hidden barriers to improvement in the quality of care. *N Engl J Med.* 2001;345(22):1612-1620.
3. **Gurwitz JH, Col NF, Avorn J.** The exclusion of the elderly and women from clinical trials in acute myocardial infarction. *JAMA.* 1992;268(11):1417-1422.
4. **Lee PY, Alexander KP, Hammill BG, Pasquali SK, Peterson ED.** Representation of elderly persons and women in published randomized trials of acute coronary syndromes. *JAMA.* 2001;286(6):708-713.
5. **Tunis SR, Stryer DB, Clancy CM.** Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA.* 2003;290(12):1624-1632.
6. **Luce BR, Paramore LC, Parasuraman B, Liljas B, de Lissvooy G.** Can managed care organizations partner with manufacturers for comparative effectiveness research? *Am J Manag Care.* 2008;14(3):149-156.
7. **The Kaiser Family Foundation.** *The Medicare Prescription Drug Benefit: Fact Sheet.* Washington, DC: The Henry J. Kaiser Family Foundation; November 2009. <http://www.kff.org/medicare/7044.cfm>. Accessed June 8, 2011.
8. **Medicare Payment Advisory Commission (MedPAC).** *Report to the Congress: Medicare Payment Policy.* www.medpac.gov/documents/mar10_entirereport.pdf. Published March 2010. Accessed June 8, 2011.
9. **Research Data Assistance Center.** Requesting CMS data. http://www.resdac.org/Medicare/requesting_data.asp. Published March 9, 2011. Accessed March 22, 2011.
10. **Research Data Assistance Center.** Medicare data documentation. <http://www.resdac.org/ddvh/index.asp>. Accessed June 10, 2010.
11. **Steiner JF, Prochazka AV.** The assessment of refill compliance using pharmacy records: methods, validity, and applications. *J Clin Epidemiol.* 1997;50(1):105-116.
12. **Hsu J, Price M, Huang J, et al.** Unintended consequences of caps on Medicare drug benefits. *N Engl J Med.* 2006;354(22):2349-2359.
13. **Crystal S, Akincigil A, Bilder S, Walkup JT.** Studying prescription drug use and outcomes with medicaid claims data: strengths, limitations, and strategies. *Med Care.* 2007;45(10 suppl 2):S58-S65.
14. **Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML.** Good research practices for comparative effectiveness research:

approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report—Part II. *Value Health*. 2009;12(8):1053-1061.

15. McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? analysis using instrumental variables. *JAMA*. 1994;272(11):859-866.

16. McClellan M, Newhouse JP. The marginal costs and benefits of medical technology: a panel instrumental variables approach. *J Econometrics*. 1997;77(1):39-64.

17. Newhouse JP, McClellan M. Econometrics in outcomes research: the use of instrumental variables. *Annu Rev Public Health*. 1998; 19:17-34.

18. Horwitz RI, Viscoli CM, Clemens JD, Sadock RT. Developing improved observational methods for evaluating therapeutic effectiveness. *Am J Med*. 1990;89(5):630-638.

19. McKee M, Britton A, Black N, McPherson K, Sanderson C, Bain C. Methods in health services research: interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ*. 1999; 319(7205):312-315.

20. Yusuf S, Held P, Teo KK, Toretzky ER. Selection of patients for randomized controlled trials: implications of wide or narrow eligibility criteria. *Stat Med*. 1990;9(1-2):73-83.

21. Seeger JD, Kurth T, Walker AM. Use of propensity score technique to account for exposure-related covariates: an example and lesson. *Med Care*. 2007;45(10 suppl 2):S143-S148.

22. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf*. 2006;15(5):291-303.

23. McCormack LA, Fox PD, Rice T, Graham ML. Medigap reform legislation of 1990: have the objectives been met? *Health Care Financ Rev*. 1996;18(1):157-174.

24. Riley GF, Levy JM, Montgomery MA. Adverse selection in the Medicare prescription drug program. *Health Aff (Millwood)*. 2009;28(6): 1826-1837.

25. Neuman P, Cubanski J. Medicare Part D update—lessons learned and unfinished business. *N Engl J Med*. 2009;361(4):406-414.

26. Wennberg JE, Freeman JL, Shelton RM, Bubolz TA. Hospital use and mortality among Medicare beneficiaries in Boston and New Haven. *N Engl J Med*. 1989;321(17):1168-1173.

27. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending, P I: the content, quality, and accessibility of care. *Ann Intern Med*. 2003; 138(4):273-287.

28. Soumerai S, Ross-Degnan D, Avorn J, McLaughlin T, Choodnovskiy I. Effects of Medicaid drug-payment limits on admission to hospitals and nursing homes. *N Engl J Med*. 1991;325(15):1072-1077.

29. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology*. 2006;17(3):268-275.

30. Wharam JF, Daniels N. Toward evidence-based policy making and standardized assessment of health policy reform. *JAMA*. 2007;298(6): 676-679.

31. The Kaiser Family Foundation. *Medicare Advantage Fact Sheet*. Washington, DC: The Henry J. Kaiser Family Foundation; September 2010. <http://www.kff.org/medicare/upload/2052-14.pdf>. Accessed June 23, 2011.

32. Institute of Medicine. *100 Initial Priority Topics for Comparative Effectiveness Research*. <http://www.iom.edu/-/media/Files/Report%20Files/2009/ComparativeEffectivenessResearchPriorities/Stand%20Alone%20List%20of%20100%20CER%20Priorities%20-%20for%20web.ashx>. Accessed June 10, 2010.

33. Federal Business Opportunities. CER-Public Use Data Pilot Project. June 3, 2010. https://www.fbo.gov/index?s=opportunity&mode=form&id=4841b0bcb3a66d848e5f9bca325d8083&tab=core&_cvview=0. Accessed September 7, 2010.

34. IMPAQ International LLC. Quantitative data analysis projects. Featured project: Recovery—Comparative Effectiveness Research (CER) Public Use Data Pilot Project. <http://www.impaqint.com/project-showcase/quantitative-data-analysis-projects/default.html>. Accessed September 7, 2010. ■