Facilitating Access to Medicare Part D Drug Claims Data

A study conducted by staff from Georgetown University and from NORC at the University of Chicago for the Medicare Payment Advisory Commission

Elizabeth Hargrave
NORC at the University of Chicago

Jack Hoadley
Georgetown University

•

MedPAC
601 New Jersey Avenue, NW
Suite 9000
Washington, DC 20001
(202) 220-3700
Fax: (202) 220-3759
www.medpac.gov

•

March 2008 • No. 08–2

NORC
A national organization for research
at the University of Chicago

Facilitating Access to Medicare Part D Drug Claims Data

FINAL REPORT

SUBMITTED TO:

MEDICARE PAYMENT ADVISORY COMMISSION
RACHEL SCHMIDT, PROJECT OFFICER

PRESENTED BY:

ELIZABETH HARGRAVE
NORC AT THE UNIVERSITY OF CHICAGO

JACK HOADLEY
GEORGETOWN UNIVERSITY

NATIONAL OPINION RESEARCH CENTER (NORC)
AT THE UNIVERSITY OF CHICAGO
4350 EAST-WEST HIGHWAY
BETHESDA, MD 20814
(301) 634-9327

MARCH 20, 2008

FACILITATING ACCESS TO MEDICARE PART D DRUG CLAIMS DATA

EXECUTIVE SUMMARY

Based on its current interpretation of law, the Centers for Medicare and Medicaid Services (CMS) is using drug claims data only for plan payment purposes. Under this strict interpretation of the allowable uses of the data, CMS, congressional agencies, and other researchers are not able to fully evaluate and oversee the Part D program that serves 24 million people at an annual cost of about $50 billion. In addition, the current restrictions prevent the Centers for Disease Control (CDC) and the Food and Drug Administration (FDA) from using the data to detect trends in epidemiology and drug safety.

In its March 2008 report to Congress, the Medicare Payment Advisory Commission (MedPAC) called on the Congress to take action making the Part D claims available "regularly and in a timely manner to congressional support agencies and selected executive branch agencies for purposes of program evaluation, public health, and safety." CMS and other federal agencies have an urgent need for these data, which are critical to understanding the successes and failures of the Medicare drug benefit. Access to the drug claims will provide researchers and policy analysts the opportunity to study carefully options for improving the benefit's design before making decisions about program changes. Furthermore, the claims data can be used to study whether access to drugs through Part D is improving the care provided to Medicare beneficiaries and how this care can be further improved with a closer look at whether drugs are being used safely and effectively.

This report examines how concerns of stakeholders about this potential data release can be addressed through many of the current protections in place for other Medicare claims data. It also explores additional protections that could address stakeholder concerns further while still making data available to improve the Medicare program and advance public health more generally. Current data policies in place at CMS are generally considered adequate for protecting other Medicare data, but the law establishing the Part D drug benefit created new challenges and requires new policy decisions regarding the disclosure of certain kinds of data. This report indicates how existing protections can be applied to drug claims data.

Research Questions that Require Part D Data

- Part D claims data offer a powerful research tool for a wide array of questions related to both the Medicare program and the public health.

- Basic questions about the Medicare drug benefit cannot be answered easily without access to claims data. They range from simple questions such as how many beneficiaries reach the Part D coverage gap and how many prescriptions are filled for certain types of drugs to broader questions such as whether the availability of drug coverage has reduced the morbidity and mortality of Medicare beneficiaries or has saved money either for the program or individual beneficiaries.

- Claims are also needed to answer policymakers' questions as they seek to make program changes. These questions may include issues of how different benefit designs or formularies influence utilization and out-of-pocket spending, whether beneficiaries change their use of drugs once reaching the coverage gap, and whether drug use varies geographically.

- Finally, the 24 million beneficiaries with Medicare drug coverage offer a significant opportunity to monitor both drug safety issues and potential public health concerns.

Protecting Beneficiary Privacy

- Like other Medicare claims data, drug claims include personal health care information. Although some stakeholders have questioned whether releasing claims data would raise new privacy concerns, protections already in place should address these concerns.

- Currently, researchers seeking to use Medicare claims data for Parts A and B must comply with a variety of requirements and go through a review process before data are released.

- Whenever possible, these data are released without individual identifying information, for example, by linking claims with an encrypted identifier before releasing them.

- Furthermore, researchers must demonstrate that their research is based on a strong research design and would advance the public interest, not a private commercial interest. Results of such research must be published, and any reports must ensure that results are not identifiable in any way at the individual level.

- Each request for data includes a signed data use agreement with CMS and includes significant penalties for any misuse.

Addressing Provider Privacy

- Under current Medicare policies, researchers may release data in ways that identify hospitals and other institutional providers. In fact, Medicare includes provider-specific measures on hospitals, home health, and nursing homes on its website as a tool for consumers.

- By contrast, individual physician profiles cannot currently be published, even though researchers may work with datasets that identify physicians in order to reach more general results.

- However, CMS is seeking legislative authority to allow release of physician-specific information. A lawsuit filed by Consumers' CHECKBOOK that would force the release of this information is also pending. The organization argues that the public interest would be served through a web-based resource to report on individual physicians. Furthermore, the Administration has proposed making available physician-specific quality and efficiency data.

- The status of allowing individual physicians to be profiled with drug claims data for public purposes will likely depend on the resolution of the question for other Medicare claims. If physician-specific information is released, however, the requirement that claims data be used only for non-commercial purposes would prevent pharmaceutical companies from obtaining the data to target detailing activities.

Proprietary Concerns of Drug Plans

- The private drug plans that offer drug benefits under Part D are generally different from other providers represented in Medicare claims, because they consider their prices and other practices to be proprietary.

- Much price and plan practice information is already publicly available. Plan formularies, benefit designs, and utilization management rules are currently available on Medicare's Plan Finder website and through formulary public use files. Information on prices is excluded from the public use files, although it is available on a drug-by-drug basis (with some imputed values) on the Plan Finder. Furthermore, some plans provide researchers access to their own data, creating a precedent for research use.

- Drug claims include slightly more price information than the Plan Finder: the full price of each prescription is broken out into an ingredient price and a dispensing fee, along with the amount paid by the beneficiary, the plan, and any third party payments.

- The prices that plans include on drug claims do not reflect the rebates paid to them by manufacturers. Plans are concerned that releasing claims data sets a precedent for releasing rebate data. But CMS has not announced any plans to make Part D rebates available, and has strongly protected rebate data that it collects in the Medicaid program.

- Drug-specific price disclosures are by far the greatest concern for plans. Some in the industry go further and argue more broadly that all plan practices, such as utilization management strategies, are fundamentally proprietary. They do not welcome plan-to-plan comparisons of these practices. Others respond that this position is overly protective. One of the most challenging policy questions is how evaluation and oversight research can examine actions of specific plans without revealing information that is truly proprietary.

- Regardless of whether Medicare decides to allow research that publicly identifies plans, plan identifiers and prices will be required as inputs to answer some important research questions about Part D. The public clearly has a compelling interest in understanding how taxpayer dollars are being spent in ways that require the use of these data.

How Data Protections Could Address Plan Concerns

- Access to drug claims data for federal agencies, including fully identified information on plans, seems critical for oversight and evaluation of the drug benefit. As such, the needs of the federal government seem to take precedence over plan concerns about protecting proprietary information, particularly since these federal agencies share concerns about not releasing information that would harm the program.

- Beyond data uses by federal agencies, Medicare policies that limit research to qualified researchers doing studies with a broad public interest would help to limit data availability to appropriate uses. Under current policy, CMS would not make proprietary information available to competing plans or to drug manufacturers.

- Furthermore, Medicare's policy of releasing only the "minimum necessary data" for any project should also help address stakeholder concerns. Many research projects do not require drug prices or plan names. For example, research addressing the impact of drug use on other health services or comparing the effectiveness of alternative therapies typically does not need prices at the individual drug level or the identity of plans.

- Research questions on the effectiveness of different plan or formulary designs might require that researchers know the identity of the drug plans, but might not require that results be published in a way that identifies plans. Still, the small number of organizations offering drug plans may allow a reader to deduce organizational identities even if not identified in published results.

Additional Approaches to Addressing Stakeholder Concerns

- Medicare could restrict some data elements to a subset of users, comparable to the current rules that restrict rebate data collected by Medicaid to limited uses by the Department of Health and Human Services and certain congressional support agencies.

- Medicare could increase current protections by requiring that researchers receive training on appropriate use of claims data and on the steps necessary for protecting sensitive data. The agency could also perform audits of selected data users or increase sanctions for misuse to help ensure appropriate data security.

- Medicare could institute a system of data centers or a web-based data enclave for use of claims data. Such steps provide stronger protection than shipping data to the researcher.

Introduction

Based on its current interpretation of the statutory language in the Medicare Modernization Act, the Centers for Medicare and Medicaid Services (CMS) is using drug claims data only for plan payment purposes. Under this strict interpretation of the allowable uses of the data, CMS, congressional agencies, and other researchers are not able to fully evaluate and oversee the Medicare Part D drug benefit.

The Medicare Payment Advisory Commission (MedPAC), in its June 2005 report to the Congress, included a recommendation that CMS should make Part D claims data available to congressional support agencies to enable them to report to the Congress on the drug benefit's impact on cost, quality, and access. In its March 2008 report, MedPAC reiterated and strengthened its call for claims data availability with the following recommendation for congressional action to make these data available for its use and use by other federal agencies:

> The Congress should direct the Secretary to make Part D claims available regularly and in a timely manner to congressional support agencies and selected executive branch agencies for purposes of program evaluation, public health, and safety.

CMS clearly needs the data for purposes that go beyond payment – to evaluate the Medicare drug benefit and make recommendations about the program. Likewise, congressional support agencies cannot conduct oversight and advise Congress regarding many essential questions about a program that serves 24 million people at an annual cost of about $50 billion without use of these data. Other government agencies, such as the Food and Drug Administration (FDA) and the Centers for Disease Control (CDC), could use the drug claims data to pursue larger goals of advancing public health and drug safety. Through the latter types of research, the claims data can be used to study whether access to drugs through Part D is improving the care provided to Medicare beneficiaries and how this care further benefit from information on the safe and effective use of drugs. This research is particularly important for the Medicare population, since few drug studies are conducted on elderly individuals or those with disabilities or multiple health conditions. But the research will also benefit populations of all ages.

In October 2006, CMS released a proposed rule that would allow the release of Medicare Part D drug claims data to other government agencies and to researchers in the general public. The agency proposed this rule to clarify statutory ambiguity regarding uses of these data beyond narrow payment-related purposes. As proposed, the rule would place Part D claims data under the same protections currently used for research that uses claims from the rest of Medicare. The agency is currently working on a final version of the rule. If the rule is published and incorporates the policies in the proposed rule, release of the data to both federal agencies and other researchers would follow directly. But in the absence of action by CMS or the Congress, broad use of claims data will be restricted.

Many comments on the rule supported the proposal to make claims data available to researchers outside the government pursuing these topics as well. However, some commenters expressed concerns that the release of drug claims data for research purposes might violate beneficiary privacy, prescriber privacy, and the privacy of proprietary information on plan practices and drug prices. Many commenters urged the agency to identify ways to support the research goals underlying the proposed rule while providing appropriate protections in these areas.

In this report, we examine some of these stakeholder concerns, and assess how current protections for claims data for the rest of Medicare will address those concerns. We also explore some additional protections that could address stakeholder concerns further while still making data available to improve the Medicare program and advance public health more generally.

Data Elements Included In Medicare Drug Claims

CMS currently collects information on every claim paid by a drug plan sponsor. Although the Medicare Part D program is generally a capitated system, not requiring CMS to pay on a transaction-by-transaction basis, there are several elements of the program that require CMS to track individual claims. These include the verification of out-of-pocket costs that count toward a beneficiary's catastrophic limit, payment of cost sharing for beneficiaries enrolled in the low-income subsidy, reinsurance payments for plans during the period when a beneficiary is eligible for catastrophic coverage, and risk adjustment of plan payments.[1]

Medicare Part D drug claims data include the following elements[2]:

Patient information:

- Health insurance claim number (used to identify the beneficiary and link to other files)
- Patient date of birth
- Gender
- Date of service
- Beneficiary coverage status

Provider information:

- Identification of prescribing health care professional
- Identification of pharmacy where the prescription was filled

---

[1] Greenwald, Leslie. 2007. "Medicare Part D Data: Major Changes on the Horizon." Medical Care • Volume 45, Number 10 Suppl 2, pp. S9-S12.

[2] Centers for Medicare and Medicaid Services. October 18, 2006. "Proposed rule: Medicare Program; Medicare Part D Data" [CMS-4119-P], *Federal Register* Volume 71, Number 201, pp. 61445-61455.

Plan information:

- Part D sponsor contract number and plan benefit package identification number

Information about the prescription drug:

- Identification of dispensed product, using national drug code (NDC) number
- Whether drug was compounded or mixed
- Prescriber's instructions on substitution of generic equivalents (*e.g.*, "dispense as written")
- Quantity dispensed (*e.g.*, number of tablets, grams, milliliters, or other unit)
- Days supply and fill number
- Dispensing status and whether the fully or partially filled

Drug price information:

- Ingredient cost of the product dispensed
- Dispensing fee paid to pharmacy
- Sales tax

Cost sharing information:

- Amount paid by the plan and date paid
- Beneficiary out-of-pocket costs (copayments, coinsurance, deductibles)
- Third party payments, and whether they count toward true out-of-pocket costs (TROOP)
- Whether beneficiary has reached catastrophic threshold; amount below and above threshold
- Low income cost sharing subsidy amount (if any)
- Whether unique pricing rules apply, such as out-of-network or Medicare as Secondary Payer

As outlined in this report, this wealth of information presents both opportunities to learn about the Medicare program – and about health care more generally – at the same time that it poses challenges to confidentiality of information about individuals, prescribers, and businesses. We will discuss the research opportunities, the confidentiality challenges, and ways to address those challenges so that the public can benefit from research using the data without threatening individual privacy, physicians' practice of medicine, or the confidentiality of proprietary information.

In addition to the claims data, CMS also collects a set of aggregate measures from the drug plans. These measures track such areas as exception requests, appeals, generic dispensing rates, and numbers of cases with prior authorization or step therapy. CMS also collects administrative measures such as rebate and claims processing information. These data are used internally by CMS for various administrative and oversight purposes, but are not made available to researchers. Because these measures are all aggregate, there are no individual privacy issues, but they may raise proprietary concerns for plans. This report does not address the research uses of these data, but some of the points of discussion would apply to them as well.

Research Questions That Require Drug Claims Data

There are many research questions that cannot be answered at all without access to drug claims data. Even some basic descriptive facts about the program cannot be learned from other administrative data, ranging from how many prescriptions Medicare Part D pays for each year, how many beneficiaries reach the coverage gap, or how much the average beneficiary is paying out of pocket to use the benefit. And while some of these basic facts could be obtained through aggregate data, other questions about whether the drug benefit has any impact on other Medicare expenditures requires linking individual drug claims to claims data on these other services.

There are still more topics on which drug claims data can contribute to a larger body of knowledge, such as using this large body of data to learn more about the safety and effectiveness of the drugs that beneficiaries are using. It is neither feasible nor desirable for CMS alone to undertake this body of research. Thus, if many of these questions are to be answered, data must be made available both inside and outside the agency. In this section, we briefly review some of the areas in which drug claims data are essential or could make a significant contribution. Exhibit 1 summarizes what elements of the claims data would be used to carry out each type of research described in this section.

Questions Related To Program Oversight

Basic statistics on Part D. Certain basic statistics about Part D would be highly useful for tracking the program but are only available through claims data. For example:

- How many people reach the coverage gap?
- How many prescriptions do Part D enrollees fill for a certain type of drug (e.g., statins)?
- What share of Medicare utilization is for generic drugs, whether measured in terms of the generic share for a particular chemical entity or for a particular drug class?

Evaluating the impact of Part D. Many policymakers have speculated that the costs of the drug benefit may be offset somewhat by reduced spending in other parts of Medicare. Answering some of the questions relating to Part D's impact will require beneficiary drug claims to be linked with records from other parts of Medicare to see the full picture of beneficiaries' spending. Questions include:

- How has access to drugs changed for different groups of beneficiaries (previously uninsured, Medicaid, supplemental)?
- Does availability of drug coverage improve the health status of beneficiaries and reduce their morbidity and mortality?

EXHIBIT 1. MEDICARE DRUG CLAIMS DATA NEEDED TO ANSWER SELECTED POLICY AND RESEARCH QUESTIONS

Policy or Research Question	Plan reported Aggregate Data	Claims Data				Links with A&B	Links with other data
		Beneficiary level data	Physician level data	Plan level data	Drug prices		
BASIC STATISTICS ON PART D							
People in coverage gap	•	•					
Prescriptions for a given drug		•					
IMPACT OF PART D							
How has access to drugs changed?		•					•
Does drug coverage reduce use of other health services or improve health status?		•				•	
Does spending on drugs affect other Medicare spending?		•				•	
ELEMENTS OF PART D DESIGN							
Effects of beneficiary choice		•		•			
Impact of the coverage gap on utilization or medical spending		•				•	
Impact of Medicaid-to-Medicare shift		•				•	•
Access to prescribed drugs via exceptions and appeals	•						
Achieving favorable drug prices				•	•		
PAYMENT ISSUES							
Use and spending by region		•					
Are payments and risk adjusters accurate?		•				•	
COST ESTIMATES AND IMPACT ANALYSIS	•	•	•	•	•	•	•
EVALUATING PLAN PERFORMANCE							
Impact of plan designs on use of generics or preferred drugs		•		•	•		
Best practices: encouraging/ discouraging certain drugs		•		•			
IMPROVING BENEFICIARY CARE							
Patterns reflecting opportunities to improve care		•				•	
Prescribing patterns of individual physicians		•	•			•	
COORDINATION OF BENEFITS							
What implications does utilization in Part D have for Medicaid and the VA?		•					•
DRUG SAFETY AND EFFECTIVENESS							
Adverse effects		•				•	
Effectiveness for Medicare population		•				•	
Comparative effectiveness		•				•	
PUBLIC HEALTH							
Health disparities and epidemiologic trends		•				•	

NOTE: Other types of data or data elements may be necessary in some cases for control purposes. Thus, for example, drug safety studies might require physician or plan information to control for the prescribing practices or plan management practices.

- Does availability of drug coverage reduce emergency room visits, physician office visits, and other use of health services? For example, has expanded access to drugs reduced the health services needed to treat complications from untreated medical conditions such as diabetes or hypertension? Or, is there an apparent increase in physician visits associated with getting and managing prescription medications?

- Do any of these trends differ across beneficiaries with different health conditions or different demographic characteristics?

Evaluating elements of Part D design. As policymakers consider how to manage the program in the future, they may have questions about whether particular features of the drug benefit's design are affecting the benefit's impact. For example:

- Selection among multiple plans: How often do beneficiaries take advantage of the opportunity during open seasons to switch plans? If they do, what effect does the switch have on their costs or their utilization of drugs? Do those not switching experience higher out-of-pocket costs as a result?

- The coverage gap: What is the impact of the benefit's coverage gap on utilization or out-of-pocket costs? Do beneficiaries in the gap stop taking medications, shift to less expensive alternatives, or simply absorb the full costs? Do any shifts in drug use persist after beneficiaries emerge from the gap or in a later year?

- Moving dual eligibles' drug coverage from Medicaid to Medicare: What impact has the shift of dually eligible beneficiaries from Medicaid to Medicare drug coverage had on either their utilization of drugs or on the out-of-pocket costs they incur? Have dually eligible beneficiaries experienced ongoing difficulties maintaining access to their drugs, and if so, have these problems increased physician visits to change medications or hospitalizations because of adverse health effects?

- Tiering, exceptions and appeals: To what extent do beneficiaries fill prescriptions for drugs that are not on plan formularies or that are placed on non-preferred tiers? To what extent do plans treat these as covered or preferred drugs, presumably because beneficiaries have gone through the exceptions or appeals process?

Payment issues. Claims data have clear uses for payment purposes, not only in the sense of ensuring accurate payments, but also in setting more general payment policies. While this task most clearly falls to CMS, outside researchers and oversight agencies also have an interest in pursuing questions related to payment policy. Questions could include:

- Is there significant regional variation in drug use or drug prices that would justify either making adjustments to government payments in a regionally based program or offering incentives to make drug utilization more uniform across the country?

- Do risk adjusters accurately compensate plans for the beneficiaries they enroll, on average and within particular groups of interest (by age, sex, Medicaid status)?

Cost estimates and impact analysis. As CMS and Congress consider regulatory and legislative changes to Part D, claims data can help make accurate estimates of the financial costs and other impacts of policy changes.

Evaluating plan performance and strategies. There is a great deal of variation in plan design within the Medicare drug benefit. As described in more detail below, plan-level price data are considered particularly sensitive by plans and drug manufacturers. But plans are also concerned that revealing information on their spending and utilization patterns will threaten their proprietary approaches to managing the benefit. Nonetheless, policymakers want to know whether the variations across plans and the competition among plans lead to different outcomes. For example:

- Do different cost sharing structures and other plan incentives lead to differential success in shifting beneficiaries' utilization to generic drugs, especially to similar generic drugs within the same drug class? Are some plan designs more successful than others in shifting utilization to preferred products for purposes of getting better prices?

- What is the effect of different approaches to monitoring and encouraging use of the clinically most appropriate drugs or avoiding drugs that are contraindicated for the Medicare population?

- Is enhanced coverage for drugs used in the gap selected by beneficiaries with expenses in the gap? Is there evidence of adverse selection? Does the higher premium for this coverage approximate the value received by the beneficiaries who select it? Is there a difference in drug use between beneficiaries with gap coverage and those without such coverage?

- Are there differences in the management of drug utilization and costs between standalone drug plans that sponsor only drug coverage and the Medicare Advantage plans that enroll beneficiaries for coverage of all health services?

Improving beneficiary care. As CMS explores the concept of "pay for performance" for physician payments in Medicare, prescribing may be one additional area in which measures could be designed. The large number of drug plans participating in Medicare makes it difficult for any one plan to have meaningful information about individual physicians. Bringing data together across health plans can provide more opportunities for this review.

- Can patterns of drug use be identified that indicate opportunities to improve the quality of care for individual beneficiaries?

- How can the prescribing patterns of individual physicians be used to improve quality of care?

- At the beneficiary level, how many different physicians typically prescribe drugs for one beneficiary? Is there evidence that multiple prescribers result in duplicative or contraindicated combinations of drugs?

Coordination of benefits. Medicare beneficiaries also receive health care coverage from other federal programs, including Medicaid and the Veterans Administration. These benefits can include coverage for some drugs excluded from the Medicare benefit. States and the VA have both expressed interest in accessing Part D claims data to better understand what is going on for the Medicare beneficiaries enrolled in their programs. Whenever multiple programs and plans are involved in financing the care of beneficiaries, both coordinating payments for care and managing health care services can be challenging.

Today states must negotiate data sharing agreements with each individual Medicare drug plan in order to gain access to these data for beneficiaries who are dually eligible for Medicare and Medicaid. Many states have disease management programs or other initiatives to manage the care of dual eligibles, and the large number of Medicare drug plans makes this process difficult.

OTHER QUESTIONS IN THE PUBLIC INTEREST

Drug safety and effectiveness. The roughly 24 million beneficiaries now enrolled in standalone Medicare drug plans and drug plans attached to Medicare Advantage plans represent a large segment of all drug use in the United States. As such, the introduction of the drug benefit offers a potentially large database for post-marketing surveillance of the safety and effectiveness of particular drugs. In fact, under the FDA Amendments Act of 2007, HHS is directed to develop a database of at least 25 million patients by July 1, 2010, and at least 100 million patients by July 1, 2012 to help monitor drug safety.[3] Medicare drug claims could play a major role in the development of this database.

Furthermore, most drug clinical trials today exclude elderly and disabled individuals and those with multiple health conditions. Although this makes sense for initial drug testing, it denies health care providers critical information for treating these types of patients. Access to Medicare drug claims could help fill this void.

Questions in this area could include:

- Is there evidence of adverse effects of the drugs being used by Medicare beneficiaries? In particular, is there any evidence of safety concerns for newly approved drugs?

- Do drugs appear to be effective and safe for older persons or those with multiple comorbidities, who likely were not included in the original clinical trials?

---

[3] Drug Benefit News. "New Law Now Gives FDA Sweeping New Authority over the Safety of Marketed Prescription Drugs." October 12, 2007. Accessed at http://www.aishealth.com/DrugCosts/DBN_FDA_Authority.html.

- Is there evidence of drug-drug interactions that may not have been detected in clinical trials?
- Are some medications more effective – or more cost-effective – than others?
- Are current lists regarding which drugs are high-risk in the elderly accurate? Are they being followed?
- Based on the evidence of effectiveness and safety, are there areas for which evidence-based guidelines should be developed specifically for the elderly?

Public health. Likewise, Part D claims data represent a database that could be used as a tool to study other public health questions. These could include:

- Do prescribing patterns show evidence of health disparities or epidemiologic trends, either by geography or by patient characteristic?

Concerns Raised By Stakeholders

CMS received comments from over 100 individuals or organizations to its proposed rule. Some comments were substantially in support of the rule, some were substantially opposed, and others raised issues that they believe need to be resolved before a final rule is published. Most commenters discussed issues of individual privacy because of its broad importance, but also because it offered some opponents of the rule a legal basis to raise objections. Most comments also offered suggestions on various aspects of general procedures under which data would be released. Some of the strongest comments against the rule came from health plan, pharmacy benefit manager (PBM), and physician stakeholders. These stakeholders typically raised strong concerns about uses of data to profile individual plans or physicians or to reveal prices or other data that plans believe are proprietary. This section elaborates further on the concerns of stakeholders, and the sections that follow discuss the safeguards that tend to provide appropriate protections for individual privacy, provider privacy, and proprietary information.

Like other Medicare claims, drug claims include personal information about the health care of individual Medicare beneficiaries. Stakeholders have raised concerns about protecting beneficiary privacy, particularly in an environment of competition among plans for beneficiary enrollment. In addition, commenters expressed concern that release of drug claims data might conflict with current procedures and protections under HIPAA and other federal and state law. The Medicare drug plans in particular raised the concern that the release of claims data should not expose them to liability or require procedural changes that would increase their administrative responsibilities (and costs), such as the distribution of revised notices of privacy practices.

Medicare drug claims also include the identity of the prescribing physician. Physicians are concerned that release of this information would allow questioning of their prescribing behavior by

plans or by the Medicare program. In addition, with unfettered access to claims data, physicians expressed the fear that pharmaceutical manufacturers could use physician prescribing data as a basis for marketing their drugs to individual physicians.

Each claim also includes data identifying the plan that the beneficiary is enrolled in. Most plans are concerned about disclosure of price data from the claims. The claim includes the exact amount paid by the beneficiary and the plan for each prescription but does not include the rebates that are paid by manufacturers directly to the plans.

Although the full retail price of each prescription (subject to some imputing of prices, especially for drugs not on a plan's formulary) is available on the Medicare Plan Finder website, claims data include further information about how that price is split between the actual ingredient cost and the dispensing fee paid to the pharmacy. These pieces of the retail drug price are considered proprietary information, as they are negotiated between each plan and the pharmacies serving its enrollees.

Of even greater concern to plans, however, is that CMS might reveal the nature and magnitude of plan discounts and rebates. Although the rebate amounts are not included on the claims, plans do report them to Medicare for specific uses such as making reinsurance payments, and they are concerned that the release of claims data sets a precedent for release of the rebate data. Plans also report some aggregate rebate information as part of their required quarterly reports to CMS, and some have indicated a concern that information from those reports, if released, could be linked to the claims and allow drug-level rebates to be determined.

Plans argue that public release of rebate information would result in fewer discounts and thus higher costs to both the individual beneficiary and the federal government. The Congressional Budget Office (CBO) has scored proposals to increase price transparency as increasing costs. The proposed rule, however, does not anticipate any release of rebate data.

In addition to their concerns about revealing drug pricing, Part D plan sponsors also expressed strong concerns about releasing claims data in a way that would allow people to examine management of the drug benefit in ways that could be used against plans. Examples of these concerns include:

- *Plan profiling.* Plans are concerned about the use of data for profiling individual plans. Beyond some of the specific concerns below, plans have broad concerns that analysis of the performance of individual plans could be used in ways that would influence fair marketplace competition or be used to target plan behavior in the political arena. Plans argue that if it does not take the diversity of plan designs into account, plan profiling could also lead to misleading comparisons among plans. Others respond that plan profiling can offer an important measure of performance in a market-based system.

- *Using data from early in implementation.* Plans are concerned that claims from the first year or two of the Medicare Part D program are not representative of normal plan operations. For example, extensive use of transition policies in 2006 meant that plans were not able to use

their formularies and other tools effectively to manage drug use as they intended. But properly designed research would take such circumstances into account.

- *Copying drug utilization management techniques.* Some plans are concerned that the effectiveness of their management techniques could be judged using claims data – and used by competitors seeking to copy effective techniques. But comparisons could also provide Medicare the opportunity to encourage plans to learn the best practices and improve the program as a whole.

- *Retroactive drug safety reviews.* Plans are concerned that claims data could be used to hold them accountable for covering a drug that harmed patients, even if the plan did not know about any safety concerns related to the drug. Existing information on formularies already provides this information, but plan-identified claims data would be needed for those studying drug safety or effectiveness to control for certain plan characteristics – such as plan coverage and utilization management of a given drug under review.

- *Proprietary ownership of the data.* Some plans want to control any research uses of the data from their plan, including the ability to sell that data. Some plans are already releasing some claims information to researchers, and some are obtaining revenue in exchange for use of the data. The government, however, could argue that the program is financed by public funds and individual beneficiary premiums, giving a compelling interest in making the data available for oversight and research in the public interest.

BENEFICIARY PRIVACY: CURRENT PROTECTIONS ARE STRONG

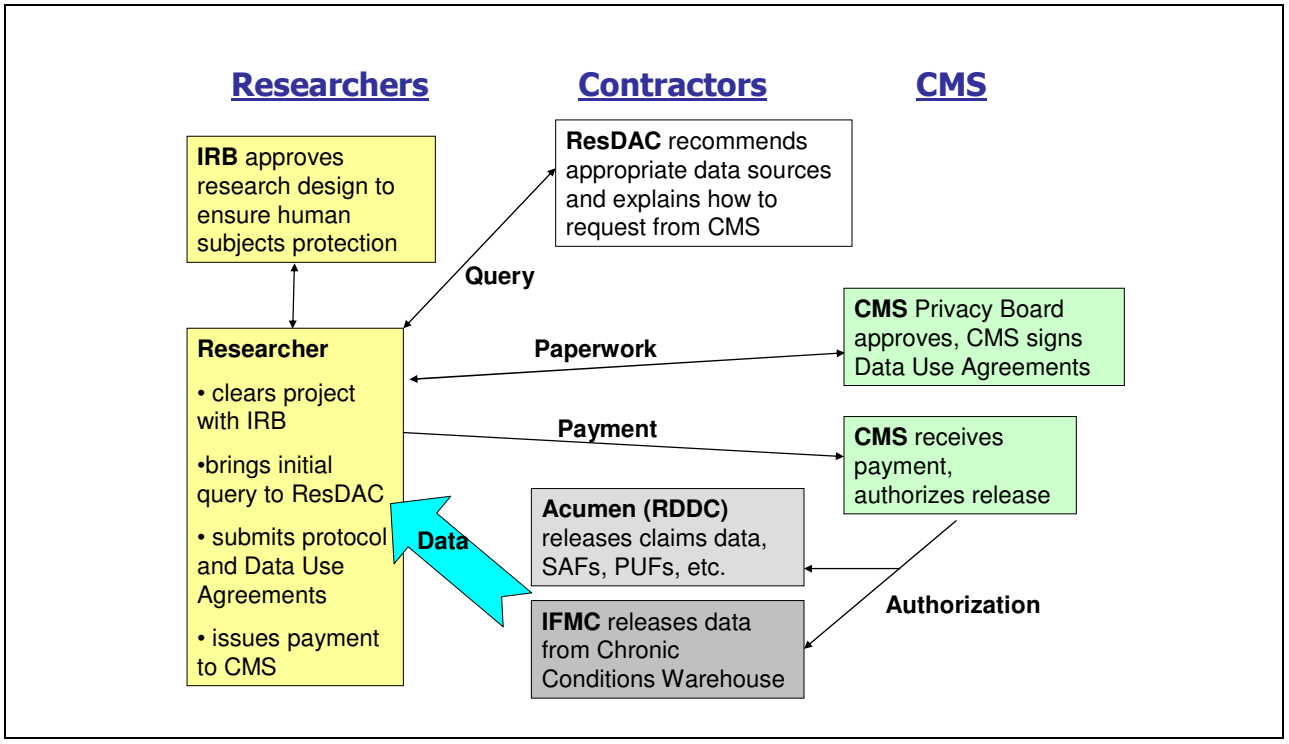
CMS currently makes Medicare claims data for Part A (the part of Medicare that pays hospitals and other institutional providers) and Part B (which pays physicians) available on two levels: a) public access to limited data, and b) restricted access to full data. Widely-available Public Use Files (PUFs) based on claims have been edited in ways that make it impossible to identify individuals in the data. In general the PUFs contain only aggregate-level information on enrollment, provider utilization and demographics, and expenditures. If these aggregate-level data are insufficient, researchers can access the full detail of Medicare claims if they apply for permission, as described in this section.

PROCESS FOR OBTAINING BENEFICIARY-LEVEL CLAIMS DATA

Research proposals seeking to use Medicare claims data generally must go through a significant review process before data are released (see Exhibit 2). All federally funded research must go

through the review of an Institutional Review Board (IRB), and CMS has begun requiring this level of review even for projects that are not federally funded.[4] An IRB is a group based at a research organization, medical facility, or university that has been formally registered with the Office for Human Research Protections in HHS to ensure that research protects the rights and welfare of research subjects. Among other things, the IRB will review a research proposal to determine whether the researcher should be required to obtain individual consent forms from each beneficiary whose claims data will be used in the project. For large scale projects, this is usually deemed impractical.

Exhibit 2. Process for Obtaining CMS Data



Source: Adapted from Saunders, William. 2006. "Centers for Medicaid and Medicare Services Research Agenda." Presentation to Academy Health Research Meeting. Accessed at http://www.academyhealth.org/2006/saundersw.ppt

With IRB approval in hand, researchers wishing to use Medicare claims then submit an application to CMS' Division of Privacy Compliance and Data Development. Proposals are reviewed first by an outside contractor, the Research Data Assistance Center (ResDAC). ResDAC may work with researchers to refine their request before sending it to CMS. For example, they may have

[4] Many research institutions already require IRB review for research, regardless of the source of funding.

recommendations about whether the request is likely to meet the Privacy Board requirements and whether all of the data requested are necessary for carrying out the study design.

The CMS Privacy Board, required by HIPAA, is the next review step for a data request. The Privacy Board includes representatives from all relevant components of CMS, as well as other components of HHS such as the National Institutes of Health and the Office of the Assistant Secretary for Planning and Evaluation (ASPE). The role of the Privacy Board is to determine whether a study adequately protects confidentiality. The requirements of the Privacy Board are outlined below in more detail.

The final approval for a data request comes from the "business owner" of the data within CMS. Each data set has a designated business owner, typically an organization within CMS that developed the data or that uses the data extensively. In the case of drug claims data, the business owner would be the Center for Beneficiary Choices. It is at this stage that CMS signs a Data Use Agreement with the researcher outlining the protections for the data about to be released.

Once a proposal has been fully approved, CMS' data contractors will prepare and release the dataset. This involves extracting only the data relevant to the project and encrypting them. The encrypted dataset is sent by mail to the researcher. The encryption key is sent separately, by electronic mail.[5]

CMS currently has two contractors responsible for sending data to researchers. The Iowa Foundation for Medical Care (IFMC) maintains the Chronic Care Data Warehouse (CCDW), which includes linked data from Medicare Part A, Medicare Part B, MDS (an assessment tool used for nursing home patients), OASIS (an assessment tool used for home health patients), eligibility, and denominator files. Another contractor, the Research Data Distribution Center run by Acumen, maintains standard public use files and limited data sets such as MedPAR. The CCDW would be the contractor responsible for drug claims. This would enable Part D claims to be linked to the other files the Warehouse maintains, so that researchers could obtain data on beneficiaries' use of services paid under all parts of Medicare in one file. The Warehouse links files by assigning a unique, unidentifiable, encrypted beneficiary link key that allows both cross-sectional and longitudinal research at the patient level. As a result, researchers have access to substantial information without requiring identified data.

REQUIREMENTS FOR OBTAINING MEDICARE CLAIMS

The CMS Privacy Board's existing requirements for access to identifiable data have strong protections for patient privacy.[6] Depending on how they are applied, many of the requirements could also be used to address stakeholder concerns about provider and plan data.

---

[5] This is one measure for ensuring data security, but there may new requirements at some point to protect data from inadvertent disclosure.

Relevant research in the public interest. The Privacy Act of 1974 generally prohibits the release of individually identifiable data, but allows disclosure for uses "compatible with the purpose for which the information was collected." Under CMS guidelines, this includes data used for a health-related research, evaluation, or epidemiologic project. The scope and subject matter of the project must assist CMS in monitoring, managing, and improving the Medicare and Medicaid programs or the services provided to beneficiaries. Each research proposal must clearly state the objectives and the significance of the study; in its review of each project, the CMS Privacy Board seeks to balance the potential risk to beneficiary confidentiality with the probable benefits gained from the completed research.

CMS expects any tool or result developed using claims data to be made available to the entire public, without charge. For example, CMS has historically denied data requests from a pharmaceutical company that wants to use the results in the development or marketing of a new drug, or an insurer seeking to improve the design or marketing of a Medicare Advantage plan. However, CMS may choose to permit research that is funded by a pharmaceutical company or insurer at a more arms-length relationship. Privately funded researchers must have total control over their results, and the private funder must never have direct access to the data. These researchers would still have to meet other requirements, such as undertaking a project with broad benefit to the Medicare program and publishing the results.

Certain stakeholders may pressure CMS to release Medicare drug claims data to a wider range of users. For example, several pharmaceutical companies submitted comments on the proposed rule supporting the release of the data, because they would like to use them. If CMS follows the same rules that it uses for other Medicare claims data, this will not be an allowed use of the drug claims.

Some drug plans have indicated interest in the data, particularly the claims history for their enrolled patients linked to other Medicare claims. For example, stand-alone Part D plans do not currently have information on the diagnoses for their enrolled beneficiaries. Although this could have advantages for patient care, it would not be an allowed use of the drug claims under the rules currently applied to claims for hospitals and physicians because CMS will not currently allow private plans to use the claims data to improve a product they are selling. Diagnostic information might be made available to all participating plans without providing more complete claims data.

Qualified researcher and strong design. The person requesting data from CMS must demonstrate the expertise and experience to conduct and complete the study. This includes submitting a research protocol that outlines a strong research design. Researchers are expected to describe how the requested data and the specific analytic techniques that will be used will help answer the research questions, addressing issues such as sample size and confounding variables. The application is

---

[6] Information in this section is based on CMS' "Criteria for Review of Requests for CMS Research Identifiable Data" at http://www.cms.hhs.gov/privprotecteddata/02_criteria.asp, ResDAC materials at http://www.resdac.umn.edu/Medicare/requesting_data_NewUse.asp, and conversations with CMS staff.

expected to include a breakdown of the analysis plan into specific tasks and a timeline for accomplishing them.

Data management. The researcher must also describe to CMS the measures he or she will use to safeguard the data and protect the privacy of CMS beneficiaries, and the process that will be followed for the destruction or return of the data to CMS at the conclusion of the study.[7] Additionally, the researcher must obtain specific permission to link any other data files to CMS databases.

Minimum identifying data. CMS asks researchers to consider how much identifying information is truly necessary for carrying out their research, and strives to release only the amount necessary. In general, files are divided into two levels of specificity. Research Identifiable Files (RIFs) do not include beneficiary or physician names, but they do include individual identifiers that would hypothetically permit the identity of a beneficiary or physician to be deduced (e.g., date of birth, age, race, sex, residence information). Limited Data Set (LDS) data still contain beneficiary-level health information, but they exclude identifiers such as service dates, birth dates, beneficiary zip code, or physician identification. If it is sufficient for the research to be conducted, researchers will be given this more limited type of dataset.

Under limited circumstances, even beneficiary names, addresses, and phone numbers may be released from the claims data. This is typically allowed for conducting surveys or focus groups of beneficiaries, when the researcher actually needs the beneficiary contact information to carry out research. In this case, however, the files are stripped of any claims information that is not directly relevant to the project. For example, CMS might give a researcher a sample of beneficiaries who have had claims for home health care, but would not provide any information about the particular health conditions of the beneficiaries in the sample unless it was directly relevant to the research.

Restrictions on publishable results. Although researchers are sometimes working with data that could be re-identified, they may only release results that are not identifiable in any way. CMS requires a minimum cell size of 11 in any tables that are released. Findings must also restrict the use of information such as geographic location, ages over 89, sex, diagnosis and procedure, admission and discharge date(s), and date of death, because they might allow someone to deduce the identity of an individual. CMS will review results if researchers have any questions about whether they should be revised to protect privacy. CMS staff also scan published literature to watch for inappropriate uses of CMS data.

Penalties. Each request for data includes a signed contract with CMS called a Data Use Agreement (DUA). The CMS DUA defines the restrictions on the use of the data, and outlines the penalties for using or disclosing data in any inappropriate way, including fines of up to $10,000 and imprisonment

---

[7] Safeguards are required to provide a level and scope of security at least as strong as set out in OMB Circular No. A-130, Appendix III--Security of Federal Automated Information Systems, which sets forth guidelines for security plans for automated information systems in Federal agencies http://www.whitehouse.gov/omb/circulars/a130/a130.html .

of up to 10 years.[8] Anecdotally, researchers have indicated that they take these agreements and restrictions quite seriously, and the IRB process to which most academic researchers must adhere provides further protection.

Beneficiary Notification

Drug plans raised some concerns about the increased burden that release of claims data might create if plans would need to send revised HIPAA compliance notices to all beneficiaries recognizing the possibility of data being released for research purposes. Similarly, they might be required to revise some administrative procedures around the collection and storage of these data. These concerns seem resolvable and might be best handled through some revision of guidelines to plans or issuance by CMS of a blanket notification to all beneficiaries.

Provider Privacy: Current Protections Are Under Debate

In addition to concerns about patient privacy, providers have concerns about their own privacy. Currently, CMS allows the use and release of data that identifies institutional providers, but not physicians. The current level of protection for physicians with regard to other Medicare claims would seem to address physician concerns if it were extended to drug claims. However, the restriction on the release of physician information is currently under litigation, because one group wants to create tools that would help consumers to select physicians based on their experience, closer to the model currently in place for identification of institutional providers.

Hospital and Other Institutional Provider Data

Because there are a limited number of hospitals in the United States, it is relatively difficult to mask the identity of hospitals when working with claims files. Furthermore, CMS has taken an interest in using claims data to monitor the quality of care provided by hospitals. Thus, CMS allows – and conducts – studies using claims data to report on individual hospitals in an identifiable way.

Hospital Compare, a tool on the Medicare website, provides hospital-specific mortality measures derived from claims files. A model uses 12 months of claims for each hospital patient admitted for heart attack or heart failure to predict that patient's mortality within 30 days, adjusted for other conditions and health factors. The website reports whether each hospital's actual 30-day mortality rate for heart attack and heart failure patients is better, the same, or worse than predicted by the model.

---

[8] For the DUA for RIFs see http://www.resdac.umn.edu/docs/CMS-R-0235-111507.pdf .

Similarly, Medicare publishes Home Health Compare and Nursing Home Compare. However, these sites use data specifically submitted to CMS for monitoring quality (e.g., OASIS), not claims data.

Private researchers have also published studies claims-based results that identify individual hospitals by name. For example, Wennberg *et al.* used Medicare claims to evaluate the management of chronically ill Medicare beneficiaries, comparing frequency of use of health care according to the hospital where patients receive most of their care. The study specified facility-specific results for the seven hospitals that ranked at the top of the *U.S. News and World Report* 2001 list for geriatric care.[9]

PHYSICIAN DATA

Although CMS allows researchers to work with data that identifies physicians in order to reach more general results, the agency does not allow the publication of results that refer to individual physicians. This policy is based on a 1979 court case that ruled the public interest was not advanced by revealing the identity of physicians and other individual providers.

However, the CMS policy of physician privacy is currently in flux. A CMS pilot project, Better Quality Information for Medicare Beneficiaries, is pooling public and private claims data to produce quality measures at the provider level. The goal of this pilot project is to provide web-based physician-level data to Medicare beneficiaries in six communities.[10] In addition, the Administration has developed a legislative proposal that would, among other things, allow CMS to provide Medicare beneficiaries with cost and quality data on individual physicians nationwide.[11] As written, the Administration's proposal appears to apply to all claims data, including drug claims.

In addition, there is pending litigation related to the use of physician-identifying claims data. In August 2007, Consumers' CHECKBOOK/Center for the Study of Services, a non-profit consumer research organization, won a Freedom of Information Act lawsuit that would require CMS to allow claims-based information about physicians to be available to the general public.[12] HHS appealed this ruling in October to seek resolution on the difference of opinion between the District of Columbia

---

[9] Wennberg, John E., Elliott S. Fisher, Thérèse A. Stukel, and Sandra M. Sharp. "Use Of Medicare Claims Data To Monitor Provider-Specific Performance Among Patients With Severe Chronic Illness." Health Affairs Web Exclusive, October 7, 2004. http://content.healthaffairs.org/cgi/reprint/hlthaff.var.5v1.pdf

[10] Department of Health and Human Services. "Better Quality Information Pilots." Accessed at http://www.hhs.gov/valuedriven/pilot/ on February 19, 2008.

[11] *Medicare Funding Warning Response Act of 2008*, Section 102. Accessed at http://www.hhs.gov/asl/medicarefundingwarninglegislation.pdf on February 19 2008.

[12] The court ruled, "The public interest at stake is the interest in obtaining information that would help the public make more informed Medicare decisions and the interest in more information of how government funds are spent," and argued that the privacy interest in protecting business affairs is outweighed by this public interest. *Consumers' Checkbook, Center For The Study Of Services* v. *United States Department of Health and Human Services* , (4thCir. 2007) Accessed at http://www.checkbook.org/Press/doc/Court%20Opinion.pdf on February 26, 2008.

court that decided the Consumers' CHECKBOOK case and the Florida court that decided the 1979 case restricting the use of the same data.[13]

If the data are released, Consumers' CHECKBOOK is planning a web-based resource that will report the number of various types of major procedures performed by each physician and reimbursed by Medicare, "so a consumer selecting a physician for a knee replacement or prostate surgery or other major procedure will be able easily to check that a physician has an appropriate level of experience."[14] The organization also argues that the public interest could be served by allowing physician-identified Medicare claims data to be used to measure physicians on how well they adhere to evidence-based care guidelines. While the current lawsuit does not directly concern Medicare drug claims data, the outcome will have strong parallels to potential uses for the drug claims.

Physician groups have opposed the Consumers' CHECKBOOK lawsuit, and their concerns in the case are similar to their concerns with drug claims data: they do not want claims data to be a tool used to question their practice of medicine. At the same time, consumers and the government have a strong interest in learning about differences in physician prescribing behavior. For example, consumers might want to know which physicians appear to adhere to evidence-based guidelines related to the prescription of certain drugs. To the extent that variations in prescribing have an impact on costs without an impact on quality or outcomes, the government has a strong interest in working with physicians to lower the costs of the Medicare drug benefit, and educating consumers could be a legitimate part of that effort.

If, either through litigation or legislation, physician-identifying research is allowed, CMS' policy on who is allowed to use data will be important in addressing one key prescriber concern. Under current policy, it appears that pharmaceutical companies would not be allowed to use the claims data to target their physician detailing practices. However, the government or groups working in the public interest might indeed use the data to pressure physicians to change their prescribing practices in ways that would be expected to improve beneficiary health.

Public concern over commercial use of prescribing data is evidenced by laws passed in three states to restrict the private-sector collection and use of pharmacy data, such as point-of-sale pharmacy transactions (normally without identification of individual patients). The goal of these laws has been to restrict the use of these data by pharmaceutical manufacturers as part of their marketing drugs to

---

[13] The case was *Florida Medical Association,Inc.* v. *Department of Health, Education, and Welfare,* 479 F.Supp. 1291 (M.D. Fla. 1979). According to the court in *Consumers Checkbook,* "The public interest asserted in favor of disclosure was 'knowing the amounts of public funds spent in reimbursing Medicare providers annually, especially in light of the ongoing legislative debate over national health insurance.' The court found, however, that this public concern was not advanced by revealing the identity of individual providers and their reimbursement amounts." *(ibid.)*

[14] Consumers' CHECKBOOK. "Consumers to Get Better Information for Choosing Physicians: Consumer Research Organization Wins Lawsuit Freeing Vast Government Data on Physicians." August 24, 2007. Accessed at http://www.checkbook.org/Press/doc/New%20Data%20to%20Measure%20Physician%20Quality.pdf on January 2, 2008.

individual physicians based on their prescribing profiles. All of these laws have been challenged in court.

Proprietary Concerns: What Are Appropriate Safeguards?

Plans' proprietary concerns about pricing and plan management strategies have no good parallel in claims for the other parts of Medicare. Medicare prices for hospitals, physicians, and other institutional providers are set by the government, and provider charges are not considered highly proprietary. These include claims for durable medical equipment, where suppliers compete with one another, bidding on the price of specific items in certain areas (scheduled to be implemented in about 80 markets in 2008 and 2009). Once the price is set, the claims are processed in the same way as other claims and would be available for research purposes. While there are differences between this competitive market and that in the Medicare drug benefit, the availability of DME claims creates a partial precedent for the drug claims.

Under Medicare Part D, each Medicare drug plan is responsible for negotiating prices with pharmacies and manufacturers. Plans are competing to achieve the lowest prices and to attract enrollment, and the discounted prices they achieve are considered proprietary. Likewise, each plan may have a different strategy for utilization management that it does not want to reveal to other plans. The current safeguards for Medicare claims data provide models for how to address many plan concerns, and the data element considered most sensitive by the plans – the rebate amount – is not part of the claim. CMS (or Congress) will have to make a key decision, however, as to whether to allow published studies to publicly identify individual plans or the transaction prices for individual drugs. CMS already publishes various performance measures on competing drug plans, so in that sense it has already made the case that the ability to compare plans is a key component in a market-based system. Claims-based analysis would create opportunities to develop many more plan performance measures.

Data currently available. Significant segments of the data that plans are concerned about releasing are already publicly available, either through CMS or through proprietary databases that are available for purchase or for use by a restricted set of researchers.

Information about plans' formulary placement and utilization management of individual drugs is already publicly available in two forms: on the Medicare Plan Finder website and in public use files available from CMS containing plans' formulary submissions. A significant amount of pricing information is also available on the Plan Finder website (subject to some imputing). For most drugs, Plan Finder users can obtain the full negotiated price of the drug as well as the beneficiary's copayment at various stages of the benefit. Both price and utilization management information are essential elements of a beneficiary's full understanding of which plan best covers his or her prescriptions, and this is the reason the data are available to the public on the web. However, with webcrawling technology, it has been possible for researchers and consulting firms to gather price

and utilization management information for large numbers of drugs at once. Plans and pharmaceutical companies have a strong incentive to collect this data to monitor the strategies of their competitors.

The price listed on either the Medicare Plan Finder or the drug claim is the retail or mail transaction price and thus excludes any rebate paid by the manufacturer directly to the plan. Plans may obtain discounts either on the transaction price or through rebates (or both), but many experts consider the rebate to be the more significant discount and the more sensitive one to plans and manufacturers. CMS collects information on rebates, but does not make this information available to the public.

There are already some proprietary databases that include claims data, available for purchase by the public. For example, IMS Health, Wolters Kluwer Health, and Verispan all collect data on pharmacy transactions in the United States, and some have developed databases aggregating these data at the person level. Likewise, Walgreens has made data on its pharmacy transactions available to some researchers. There has also been an effort to pool claims data from several Medicare drug plans. Each of these datasets is incomplete in various ways; most importantly, none includes data on every Medicare Part D enrollee as the Medicare claims data would. This limits the extent to which findings can be generalized for the entire program.

Minimum necessary data. With other Medicare claims, CMS strives to release only the identifying data necessary to answer a given question. A similar strategy could be applied to drug claims data. CMS has begun to think through the question of what summary measures could be provided to researchers when their research does not require specifically knowing the name of a beneficiary's drug plan. For example, the agency has explored the feasibility of establishing plan generosity measures that could be used to describe plans' benefit designs and formularies. Such measures could be used in lieu of plan-identifying information if plan identifiers were not made available as part of the claims data. Alternatively, they could supplement plan identifiers and offer an option under the principle of providing only the minimum necessary data.

For many questions that researchers would like to answer, potentially sensitive price and utilization management information is not necessary. Summary measures, such as total out-of-pocket costs per beneficiary, may be sufficient. Research on drug safety or effectiveness would not generally require any cost information, and utilization management information might only be needed as a control variable. Such studies are not looking at individual plans, but rather at patterns of utilization across all plans.

Research questions on the overall impact of the Medicare drug benefit similarly might not require identifying individual plans. Linked Medicare claims would be needed to see if patients who take different drugs – or different numbers of drugs – have better health outcomes. Plan information might be important to see whether utilization restrictions or formulary coverage is a factor affecting patient adherence to their medications, but these questions would not require knowing the plan in which a beneficiary was enrolled.

For other questions, it may be possible to provide specifics about price, formulary status, and utilization management for a particular prescription, but not plan identifiers. Although it would hypothetically be possible (though complicated) to re-identify plans based on these data elements, data users could be restricted by their Data Use Agreements from merging the dataset with such re-identifying information. (See the discussion below of a data enclave framework that would create stronger safeguards against re-identification.)

Other questions addressing the effectiveness of different formulary strategies or application of utilization management techniques would require extensive information about plan practices and would be difficult without plan identifiers, although measures of plan generosity (described above) or typologies of plan characteristics might offer an alternative. Even in these cases, reporting the research would not necessarily require identifying the plans by name. But with fewer than 50 drug plans and fewer than 20 organizations competing on a national basis, it might be difficult to report on such results in a way that could not allow plan identities to be revealed.

Public interest and qualified researchers. Although much information about plans' coverage and utilization management is already publicly available, claims will add the actual utilization associated with these practices. One key concern of plans is to prevent their competitors from unfairly learning information about the effectiveness of their practices. At the same time, the Medicare program has a strong interest in learning about the comparative effectiveness of various measures, both to create savings for the program and to ensure that beneficiaries maintain access to the drugs they need.

One safeguard against inappropriate commercial use of data about plan practices is the review by CMS of the researchers requesting the data and their research plans. Under current policy, use of data by other plans for private purposes would not be allowed. However, the government and independent researchers could legitimately study questions of access and effectiveness.

Although profiling is a concern for the health plans, beneficiary advocates and others argue that it is in the public interest to compare plans by name. Policymakers will have to decide whether the appropriate model is the profiling already undertaken for hospitals and other institutional providers or the current prohibition against such public profiling of physicians.

Furthermore, policymakers will have to decide whether the competitive market approach of the Medicare Part D drug benefit is sufficiently different from the government-administered approach in the rest of Medicare to support the plans' argument that they be judged on their premiums and certain other aggregate measures rather than profiling information based on claims. In fact, an argument can be made that public profiling of plans could actually help competition by helping the public determine which plans are doing a better job than their competitors managing the drug benefit. Plans would have a greater incentive to improve their performance in order to attract enrollment.

Study design. Another key concern by plans is that claims data could be used to inaccurately draw conclusions or comparisons. ResDAC and CMS could use their oversight of study design to mitigate this concern as they do today with research based on Medicare claims, pointing out to researchers some of the key flaws in study design that are of concern.

Rebate data. Plans are clearly most concerned that release of drug claims data would set a precedent for the release of drug rebate data. Indeed, an understanding of plan rebates in connection with pharmacy prices is the only way to get a full picture of plans' actual costs for medications. However, CMS has a strong history of protecting drug rebate data in the Medicaid program based on the argument that revealing these proprietary data would eliminate discounts and thus actually cost the government money. Pharmaceutical companies provide information on their rebates to CMS as part of assuring that Medicaid is paying the best price they offer to any commercial clients. These data are very closely held by the agency, although by statute it is available to DHHS, the Government Accountability Office (GAO) and CBO for research that supports the purpose of conducting cost estimates for legislative proposals.

Additional Approaches To Addressing Concerns

If concerns remain about the privacy of individuals or the secrecy of proprietary information, there are several steps that CMS could take to strengthen data security. Because the needs of the federal government for oversight and evaluation of the drug benefit take the highest priority, it is clear that CMS, congressional support agencies, and other federal agencies should obtain access to drug claims without the need for further steps. There should be no concern that federal agencies would conduct or publish research that would in any way jeopardize the interests of either individual beneficiaries or the plans that provide the drug benefit. The steps described in this section offer options to strengthen protections when data are released to non-government researchers.

Researcher education. In general, CMS requires researchers to obtain a review by an IRB or human-subject review committee. These review panels generally require a certain level of training for researchers in human subjects protection, such as completion of a web-based training module by NIH. However, this module is focused primarily on the collection of data (such as in surveys or clinical trials), not working with data that have already been collected (such as claims). CMS could require a similar level of training for individuals who want to use claims data. Training would include procedures for safeguarding data, the rules about what level of information can be published or disclosed, and the penalties for violating these rules. ResDAC currently offers training on some of these issues, but it is not required of researchers applying to use Medicare claims data.

Representing stakeholder interests on Privacy Board decisions. The CMS Privacy Board currently consists only of employees of the Department of Health and Human Services. Some stakeholders have suggested that inclusion of qualified outside stakeholders might increase confidence in the stakeholder community that CMS will not allow uses of the data that might threaten proprietary

concerns. Such a change would mark a significant departure from its current membership and would complicate the process considerably. It would also raise potential concerns that stakeholders might attempt to block legitimate research that might generate adverse results from their perspective.

User audits. One approach to stronger enforcement of the data protections outlined in the DUA would be to perform random audits of data users to investigate whether they are maintaining and using data appropriately. These audits could be targeted to the largest users of claims data, representing the highest disclosure vulnerabilities, or they could be truly random to encourage even smaller users to be vigilant about their data protections.

Increased sanctions. Although many researchers take the sanctions against inappropriate use of the data quite seriously, some have argued that increased sanctions could be another effective method for increasing data security. In effect, increasing sanctions might decrease the likelihood that someone will disclose data, while creating no increased burden on researchers who are already following the rules.[15]

Restrict access via data centers or a data enclave. Some federal agencies maintain datasets that they will not release on CD to external researchers, such as certain files from Census and MEPS that contain highly detailed and personal information. These agencies sometimes do allow use of some restricted datasets if researchers come to a designated facility to use the data. This system ensures that no data will leave the facility, but it is burdensome to researchers. There are only eight sites in the country where the restricted Census data are available, for example, creating access problems for many individuals with legitimate research interests.

An alternative approach uses secure remote access to create a similar level of protection while researchers use data from their own offices. For example, NORC currently maintains a data enclave for microdata data produced by the National Institute for Standards and Technology (NIST), the Ewing Marion Kauffman Foundation, and the Economic Research Service at the U.S. Department of Agriculture.[16] The data included are primarily from business surveys and include sensitive, proprietary information.

Once users have been approved for access to the data enclave, they log in via a secure internet connection. In this remote environment, users can use statistical and word-processing programs to work with the data and write up results, saving their work in their personal area of the data enclave. Users cannot transfer any information to their own computers without submitting that information for review, to ensure that it does not violate any of the disclosure guidelines for the dataset in question. Likewise, users can bring additional data into the secure environment only if they have it reviewed to ensure that they will not be able to inappropriately re-identify individuals or firms with

---

[15] Committee on National Statistics, National Research Council. *Improving Access to and Confidentiality of Research Data: Report of a Workshop*. Christopher Mackie and Norman Bradburn, Editors; The National Academies Press, 2000.

[16] For more information on NORC's data enclave, see http://dataenclave.norc.org/index.html .

the additional data. These protections provide a very proactive mode of enforcement to prevent disclosure. However, maintaining the system is more expensive than CMS' current system.

Pilot test. For any new system for releasing data, CMS could pilot test the system with a select number of well-respected, highly-trusted researchers. This might provide an opportunity for stakeholders to have a level of comfort with how data are being used while giving CMS the opportunity to make sure the system is working as intended.

Different Kinds Of Access For Different Kinds Of Users

As CMS works through some of the concerns and possible solutions outlined in this report, there are ways the agency could adjust the timing or expansiveness of data access based on the type of researcher accessing the data. There is an immediate need for government oversight agencies – including CMS – to evaluate the Medicare Part D drug benefit, without waiting for larger data policy issues to be resolved. Within the boundaries of CMS' current data protections, the public interest would be served by allowing legitimate researchers, both inside and outside government, access to as many data elements as possible. These options are available, however, if concerns about the sensitivity of certain data cannot be resolved.

Phasing. CMS could choose to roll out access to the data for different types of researchers in phases. For example, researchers in the Department of Health and Human Services and in oversight agencies such as CBO, GAO, and MedPAC (and their contractors) could have earlier access to the data, because they are least likely to violate the proprietary concerns held by the Medicare drug plans. Researchers in academic institutions with a history of work in the public interest might be granted access in a second wave. The latter concept could be implemented through a more rigorous review by the agency's Privacy Board of the researchers and research purposes that would be approved in early phases. In doing so, it would be important to ensure that the agency's procedures continue to avoid any political influences on decisions about what research is approved.

Restricting data elements. It may also be the case that while certain variables or types of data should not be given to the broadest array of users, other users may appropriately use them. There are currently several cases in which certain government agencies have access to prescription drug data that are not publicly available. For example, the bids of Part D plans contain proprietary data, but CMS has made them available to some congressional support agencies in their roles as advisors to Congress in overseeing the Part D program. Likewise, CMS has allowed CBO to use rebate information collected for the Medicaid program to conduct cost estimates.