Introduction and Overview

With the creation of the Medicare Plus Choice program, health plans were allowed to enroll Medicare beneficiaries. After an initial period of growth, enrollment in Medicare Plus Choice leveled and recently began to decline.

While some health plans have profited by enrolling Medicare beneficiaries, others have terminated or reduced participation in Medicare citing the high cost of caring for Medicare beneficiaries.[1] Many health plans have not been able to afford the high risks associated with very sick Medicare beneficiaries. There have been similar experiences with health plans enrolling Medicaid beneficiaries.[2] Most health plans started in the private sector with the enrollment of working populations in which the vast majority of enrollees are relatively healthy. The financial risks and clinical management challenges associated with the transition to Medicare and Medicaid populations in which a significant proportion of the population is relatively sick is proving to be difficult for some health plans.

One of the fundamental issue underlying the problems that health plans are experiencing relates to risk adjustment. MedPAC has defined risk adjustment as follows.[3]

> "The process used to adjust plan payments to compensate for differences in the health plan's health status of enrollees across plans."

Payers, in general, have done an inadequate job of risk adjusting payments to health plans. Since the success of payment on a capitated basis relies on health plans being able to manage the financial risk and respond to the incentive to be efficient, the failure to adequately adjust payments for the risk associated with the enrolled population represents a fundamental flaw. Kuttner has described the dangers of inadequate risk adjustment as follows:[4]

> "Thus failure to adjust compensation for patient's health status reinforces two of the more worrisome trends in the present healthcare system. First, it rewards plans for a business strategy of 'risk selection' in which they deliberately market their services to relatively healthy populations and avoid relatively sick ones. This strategy, in turn, punishes plans and physicians that do a good job of treating the sick, thus reinforcing the incentive to stint on care that is already present in a system that increasingly relies on payment by means of capitation rather than on fee-for-service reimbursement. Second, as risks are shifted to the individual physician, doctors with sicker patients must work longer hours or receive a reduced income or make unethical or clinically dangerous decisions to withhold necessary care."

Policy makers recognize that capitated payments under Medicare Plus Choice must be risk adjusted. Medicare's initial method of risk adjustment for Medicare Plus Choice was based on Principal Inpatient Diagnostic Cost Groups (PIP-DCGs). The diagnosis and procedures used in PIP-DCGs are

limited to those that occur in an inpatient setting. With a few exceptions, only diagnoses that are the principal diagnosis from a hospital stay are used in the PIP-DCGs. The PIP-DCGs used 16 diagnosis groups, age and sex, plus nonclinical factors (e.g., Medicaid eligibility or the reason for initial Medicare coverage) to establish risk adjusted payment rates.

The use of only inpatient encounter data for risk adjustment is problematic. Inpatient data tends to be high quality but of limited utility as hospital admissions are a relatively rare event for most populations. This is especially so for relatively healthy populations which have especially low admission rates. Further, it rewards health plans with inappropriately high hospital admission rates through higher risk adjusted payments.

Recognizing the problems with risk adjustment based only on inpatient data, Congress, in Section 603, of the Benefits Improvement and Protection Act (BIPA) required that a comprehensive risk adjustment system be incorporated into the method of determining capitated payments to health plans participating in the Medicare+Choice program. The comprehensive risk adjustment system would be based on inpatient and outpatient encounter data. Outpatient encounter data, when collected and maintained, are of lower quality than inpatient data. It is often not collected in a systemic fashion. Diagnostic coding practices for outpatient encounters are less than rigorous and are not subject to more than rudimentary edits (e.g., software ensures code validity but not reliability). Outpatient diagnoses are frequently speculative (i.e., rule-outs). In addition, most health plans have not established the necessary infrastructure to collect and code outpatient encounter data. Thus, there is substantial cost for health plans to collect outpatient encounter data.

Another source of data for risk adjustment is pharmacy data. Pharmacy data, or more correctly, data from pharmacy benefit management (PBM) programs can be a useful adjunct for risk adjustment. As Roblin states:

> "Computerized outpatient drug dispense data may be particularly valuable for case mix measurement in managed care organizations where diagnoses are not collected on physician visits or where data have not been verified against clinical medical records. Individuals with a diagnosed chronic disease – such as diabetes, hypertension, ischemic heart disease, asthma – will often include drug therapy as part of their overall treatment regimen."[5]

Even when outpatient data are available, PBM data offers the prospect of a more complete clinical picture than from claims data alone. At the very least pharmacy data offers proof of treatment. Indeed, on-going treatment can be verified with pharmacy data even in the absence of physician visits. For example, if an individual is receiving medication which controls his or her hypertension, he or she may only infrequently need to visit a doctor. In addition pharmacy data in some respects are arguably more clinically precise. A diagnosis on a claim form may be a "rule out" or even a clerical error. Most outpatient diagnoses are never subjected to an audit beyond simple validity checks. However, the presence of a documented prescription drug is reasonable proof of a diagnosis. For example, a diagnosis of diabetes without

any indication of treatment outside of the visit may be an error. However, treatment with oral hypolglycemics or insulin leaves little room for doubt that the individual has diabetes.

There are practical advantages to the use of PBM data. The data come from the vendor in machine readable form and explicitly identify each pharmaceutical. The fact that the data are machine readable means that the administrative costs of gathering these data are negligible. Because the pharmaceutical is reimbursed, the data are also likely to be scrutinized and subject to audit and other reviews.

However, pharmacy data has its limitations:

- Physicians may be using pharmaceuticals in ways unapproved by the Food and Drug Administration, i.e., off label use.
- Drugs can be prescribed based on erroneous diagnoses.
- As in other as aspects of physician practice, physicians will differ in prescription patterns.
- Drugs can have multiple applications such that a clear diagnosis is unknowable by the presence of the drug alone. For example, propanolol can be used as preventive medication for migraine headaches and as an anti-hypertensive.
- Some diseases lack specific drug treatments.
- Prescription drug programs do not cover all sites and providers.

Despite these limitations, pharmacy data has the potential to be very useful for risk adjustment. The purpose of this study is to evaluate the use of pharmacy data in risk adjustment.

Specifically, two primary issues will be examined.

- Can pharmacy data be a substitute for outpatient data in risk adjustment?
- Can the addition of pharmacy data improve the overall performance of risk adjustment systems?

Using Pharmacy Data

Pharmacy data is a by product of stand alone pharmacy coverage where pharmaceuticals are handled separately from other benefits. Each covered individual usually has a card and / or access to a mail order system. When a drug is prescribed or when a prescription needs to be renewed, the individual or his / her physician contacts the appropriate pharmacy or mail order house and the individual obtains his / her drugs. A record of that transaction is made which provides information about the transaction including the type and quantity of the drug, who prescribed it, the date the prescription was filled, renewals, etc.

Coverage for pharmacy plans is typically limited to self-administered drugs, i.e., those obtained at a pharmacy or by mail. Drugs administered by medical professionals or by institutional providers such as hospitals and home health

agencies are typically paid for outside of the PBM coverage. Clinically significant drugs such as chemotherapeutics, intravenous blood products, etc. will usually not be present in PBM data.

Pharmacy data is intrinsically more difficult to work with than encounter data because of a complex coding scheme. Pharmacy data is reported using the National Drug Codes (NDC).

> "Each drug product listed under Section 510 of the Federal Food, Drug, and Cosmetic Act is assigned a unique 10-digit, 3-segment number. This number, known as the National Drug Code (NDC), identifies the labeler/vendor, product, and trade package size. The first segment, the labeler code, is assigned by the FDA. A labeler is any firm that manufactures, repacks or distributes a drug product. The second segment, the product code, identifies a specific strength, dosage form, and formulation for a particular firm. The third segment, the package code identifies package sizes. Both the product and package codes are assigned by the firm. The NDC will be in one of the following configurations: 4-4-2, 5-3-2, or 5-4-1."[6]

The ten digit code assigned by the FDA is known as the "regulation code". The NDC codes in common usage, the "standardized" codes have a common format and consist of 11 digits. There are 75,000 NDC codes of which about 24,000 are obsolete. There are quarterly updates of individual codes. On sheer volume alone NDC codes are more complex than other medical coding schema. For example, there are about 12,000 ICD-9-CM diagnosis codes and 8,000 CPT-4 procedure codes.

More importantly, according to Feikema, NDC codes have the following limitations:

1. "It is difficult to find a definitive listing of all current NDC codes. Some sources question the comprehensiveness, timeliness, and accuracy of the FDA's quarterly list of NDC codes (available at: http://www.fda.gov/cder/ndc/index.htm). Several private pharmaceutical information services companies provide their own lists that they claim to be more timely and comprehensive. One company offers its database, with weekly updates, for free (available at: http://www.multum.com).
2. New NDC codes are supposed to be assigned when product packaging changes--even if the product itself remains the same. Thus, they must be monitored and updated regularly.
3. Regulation NDC codes cannot be automatically parsed (separated into their component sub codes) without knowledge of the applicable configuration to be applied. The configuration differs by manufacturer.
4. Product and package codes are meaningless without the accompanying labeler code. That is, there is no consistent meaning for the product code "1234." For one manufacturer, it refers to product x; for another manufacturer it refers to product y.

5. Product codes are not "clean." That is, they may reflect a given product alone, or, for products with multiple dose forms, they may reflect the given product plus the dose form. Some manufacturers even include package information in the product code. For example, Merck has 5 product codes for the pediatric dosage of its Hepatitis B vaccine. However, the products differ only in their packaging. This is not consistent with the NDC model laid out by the FDA.
6. For the purposes of clinicians and immunization registries, it is not necessary to identify package type. Using full NDC codes leads to multiple codes that identify the same vaccine type and dosage. Stripping the package code is not helpful when manufacturers embed the packaging information into the product code (as in the Merck example above). [Note that full NDC codes including package type are useful for inventory systems.]
7. Where a given product is marketed by two different firms, two unique NDC numbers exist for the exact same brand name product—to the packaging level. For example, Aplisol, a tuberculosis PPD skin test product, is manufactured and distributed in a 5-unit package by Parke-Davis (standardized NDC 00071-4525-08). The same product (brand name, strength, and packaging) is also distributed by Physician's Total Care (standardized NDC 54868-2328-01)."[7]

Knowledge of the subcomponents of NDC codes is useless, except perhaps to identify the labeler (who may or may not be the manufacturer). Instead, NDC codes are useful only in their full 10 (or 11) digit form where they identify the exact package form of a given branded product. With the appropriate corresponding information, these codes can be used to infer a specific brand name, the generic drug product, and, in some cases, a given dosage. However, multiple NDC codes can refer to the same product. Thus, an NDC code cannot be used as a unique identifier for a given product type.

To facilitate the use of pharmaceutical data and solve the methodological problems associated with NDC code, several systems have been developed which build on the NDC codes (e.g., FirstData Bank and Multum). Designed for pharmacies, they categorize drugs by active ingredient and classes of similar drugs, etc. For this study, the system developed by FirstData was used.

FirstData classifies drugs with a semi-hierarchical design. It is not completely hierarchical because drugs may have multiple active ingredients (e.g., codeine and acetaminophen are commonly combined in analgesics). The FirstData system also includes extensive non-pharmaceuticals (e.g., DME, disposable supplies, and OTC drugs) which offers the potential of including selected DME and supplies in the risk adjustment model. The FirstData system aggregates drugs depending on dosage and mode of administration, active ingredient, and categories of similar active ingredients. The first step is to give identical drugs from different labelers a single number. This number is referred to as the Generic Code Number SEQuence Number (GCNSEQNO). The GCNSEQNO is based on the active ingredient of a drug, its dosage and mode of administration. The next step is to assign a drug (GCNSEQNO) to a

hierarchy, or if there are multiple active ingredients, hierarchies. Each drug is assigned at least one Hierarchical Ingredient Code (HIC). The HIC6 level of aggregation identifies any variations of the active ingredient. The HIC4 level of aggregation identifies the active ingredient. The HIC3 level of aggregation identifies families of clinically comparable active ingredients. An example of the HIC hierarchy for penicillin is as follows:

- HIC-3 - Penicillins
- HIC-4 - Penicillin V
- HIC-6 - Penicillin V Potassium

Clinical Risk Groups (CRGs)

There are two general types of risk adjustment methodologies: categorical and regression. In a categorical model, each enrollee is assigned to a single mutually exclusive category. Associated with each category is a relative payment weight which provides a relative measure of the enrollees future need for medical resources. The capitated amount for an enrollee is risk adjusted by applying the relative payment weight as an adjustment to the capitated amount. In a regression method, a mathematical equation is used to produce a score which is unique for each enrollee. The capitated amount for an enrollee is risk adjusted by applying the score as an adjustment to the capitated amount for each enrollee.

The interpretation of pharmacy data in risk adjustment should be done in the context of the diseases profile of the enrollee. It is much more straightforward to make diagnosis-specific decisions using pharmacy data in a categorical model. Clinical Risk Groups (CRGs) are the most clinically precise categorical risk adjustment model. Therefore, the testing of the utility of the pharmacy data for risk adjustment was done by incorporating pharmacy data into the CRG model.

Before describing how pharmacy data was incorporated into CRGs, the CRG structure and logic will be described. CRGs make use of data from inpatient and outpatient encounters to assign an individual to a single severity adjusted risk group. The data include diagnoses, procedures and demographic factors along with data associated with the encounters such as date, site, and provider. There are four phases in assigning a CRG:

Phase 1: Create Disease Profile and History of Past Medical Interventions

- Diagnosis codes are categorized into mutually exclusive categories called Major Diagnostic Categories (MDCs) depending on the organ system or etiology of the disease.
- Diagnosis codes in each MDC are further categorized into mutually exclusive Episode Disease Categories (EDCs). There are six types of EDCs. There are three

chronic categories, dominant, moderate, and minor. There are two types of acute EDCs, significant acute and minor acute. The sixth type of EDC is the chronic manifestation. A chronic manifestation EDCs is typically a manifestation or acute exacerbation of an underlying chronic disease which identifies both the exacerbation and the illness. The presence of a chronic manifestation EDC results in the presence of the EDC for the underlying chronic disease.

- Procedure codes are categorized into mutually exclusive categories called Episode Procedure Categories (EPCs).
- Some EDCs and EPCs conditionally create additional EDCs.
- After all EDCs and EPCs have been created, some EDCs and EPCs are eliminated at this point based on temporal relationships with other EDCs or EPCs.
- At the end of Phase 1, the disease profile and history of past medical interventions of an individual is described by a list of EDCs and EPCs.

Phase 2: Select Primary Chronic Disease(s) and Assign Severity Level(s)

- The EDC that represents the most significant chronic disease under active treatment, referred to as the primary chronic disease (PCD), is identified for each organ system and etiology (i.e., MDC).
- The process of identification is hierarchical. It selects PCDs in order of the type of chronic EDC (dominant, moderate, or minor), then by a hierarchy of MDCs within EDC type, and finally by a set of tie break rules or an EDC hierarchy within the MDC.
- After identification each PCD is assigned a Severity of Illness Level.
- The process continues until no more PCDs can be identified.
- At the end of Phase 2, the most significant chronic diseases (if any) and their associated severity levels are identified.

Phase 3: Determination of Base CRG and Severity Level for the individual

- Based on the enrollee's PCDs and their associated severity levels, acute EDCs, and EPCs the individual is assigned to one of nine statuses.

1. Healthy: A healthy status is identified by the absence of any PCDs or significant acute EDCs or EPCs.
2. Recent History of Significant Acute Disease: A history of significant acute disease is identified by the presence within the most recent six month period of one or more Significant Acute EDCs or significant EPCs. There are no PCDs present.
3. Single Minor Chronic Disease: A single minor chronic disease is identified by the presence of a single Minor Chronic PCD. Minor chronic PCD with a severity of illness level of 1 are ignored if a dominant or moderate chronic PCD is present.
4. Minor Chronic Disease in Multiple Organ Systems: Minor chronic disease in multiple organ systems is identified by the presence of two or more Minor Chronic PCDs.
5. Single Dominant or Moderate Chronic Disease: Single dominant or moderate chronic disease is identified by the presence of a single Dominant or Moderate Chronic PCD.
6. Significant Chronic Disease in Multiple Organ Systems: Significant chronic diseases in multiple organ systems is identified by the presence of two or more PCDs of which at least one is a Dominant or Moderate Chronic PCD. PCDs that are a Severity Level 1 minor chronic disease are not considered a significant chronic disease and are not used to identify the presence of significant chronic disease in multiple organ systems, but Minor Chronic PCDs that are Severity Level 2 minor chronic diseases are used.
7. Dominant Chronic Disease in Three or More Organ Systems: Dominant chronic disease in three or more organ systems is identified by the presence of three or more dominant chronic or selected moderate chronic PCDs.
8. Dominant, Metastatic and Complicated Malignancies: A malignancy that dominates the medical care required (e.g., brain malignancy) or a nondominant malignancy (e.g., prostate malignancy) that is metastatic or complicated (e.g., requiring a bone marrow transplant).
9. Catastrophic Conditions: Catastrophic Conditions include long term dependency on medical technology (e.g., dialysis, respirator, and total parenteral nutrition (TPN)) and life-defining chronic diseases or conditions that dominate the medical care required (e.g., persistent vegetative state, cystic fibrosis, AIDS, history of heart transplant).

- Each status is further subdivided into mutually exclusive base CRGs. The base CRG is selected hierarchically using the enrollee's PCDs, acute EDCs, and EPCs.

- The enrollee is then assigned a Severity of Illness Level.

The combination of the base CRG and the Severity of Illness Level constitute the CRG. There are a total of 1075 CRGs. At the end of Phase 3, the enrollee has been assigned to a CRG. Table 1 contains a summary of the number of CRGs by status.

	CRG Status Hierarchy (Left to Right)								
Nine CRG Statuses	Cat	Dom/ Met Malig	Three Dom Chronic	Multiple Signif Chronic	Single Dom or Mod Chronic	Multiple Minor Chronic	Single Minor Chronic	Signif Acute	Healthy
269 Base CRGs	11	22	21	61	106	1	40	6	1
Example Base CRG	HX Major Organ Trans	Met Colon Malig	DM & CHF & COPD	DM & CHF	DM	Migraine & BPH	Migraine	Hx Acute ENT	Healthy
Severity Levels	4	4	6	2, 4 or 6	2 or 4	4	2	None	None
1075 Total CRGs	44	88	126	328	398	4	80	6	1

Table 1. Summary of the number of CRGs by status

Phase 4: Consolidation of CRGs into Three Successive Tiers of Aggregation

CRGs are consolidated into three tiers of aggregation. Each successive tier has fewer base CRGs but maintains the severity level. The aggregated CRGs are referred to as ACRGs and the successive tiers of aggregation are referred to as ACRG1, ACRG2 and ACRG3. At the end of Phase 4, the enrollee has been assigned an ACRG1, ACRG2, and ACRG3. There are 413, 149 and 37 ACRG1s, ACRG2s, and ACRG3s, respectively.

The Pharmacy Extension to CRGs (CR_xGs)

The pharmacy extension to CRGs is called CR_xGs. The clinical development of CR_xG consisted of several distinct parts.

- Several physicians using their clinical judgment and extensive literature review developed the original clinical model.
- The model was circulated to a wide range of clinicians and pharmacologists. The model was modified based on their comments.
- The model was then tested with the developmental database and modifications made. The modifications were in turn tested. This was repeated several times in an iterative process between the data and clinical judgment. When the two, clinical judgment and the data conflicted, clinical judgment prevailed.

The pharmacy logic was designed to interface with Phases 1 – 3 of the CRG logic. The following logical requirements were incorporated into the drug specifications.

- Specifications could be for individual drugs or for combinations of up to four different drugs.
- Drugs could be specified at either the HIC6, HIC4, or HIC3 levels of aggregation.
- Drug specifications would include conditional rules. The rules may be for demographic factors (e.g., pediatric vs. non-pediatric), occurrences over an extended period of time (90 or 180 days), or time and frequency of occurrence (e.g., four occurrences with the first and last occurrence being more that 180 days apart). The default rule is any occurrence of the drug within the last year.
- In order to avoid double counts, whenever a drug is used to meet a specification, that drug and its related aggregations can no longer be used again to meet a comparable specification.
- The pharmacy logic will attempt to identify the most significant disease before a less significant disease.
- The pharmacy logic will attempt to identify a more significant disease before a less significant disease.

The pharmacy logic performed three distinct tasks, create an EDC, assign a severity of illness level to a PCD, and assign a severity of illness level to a CRG.

- The first step in creating an EDC was to identify the most significant illnesses indicated by the presence of a specific drug or drugs which satisfied certain conditions. To identify the most significant illness an individual's pharmacy history was compared against a set of specifications. These specifications were in order of EDC type (dominant chronic, moderate chronic, minor chronic, chronic manifestation, significant acute, and minor acute), MDC hierarchy (within EDC type), EDC rank within MDC, and severity of illness level within EDC. Once a specification was met, all drugs used to meet the specification and all drugs which could have been used to meet the criteria for that PCD or CRG at or below the severity level which was satisfied were eliminated to prevent double counting, i.e., the same drug creating more than a single EDC or treatment for the same EDC creating more than a single EDC. The following are examples of EDC creation from pharmacy data.
    - Chest pain EDC is created by a single prescription for nitroglycerine.
    - Angina EDC is created when there are two or more prescriptions for nitroglycerine with at least 90 days between the first and last.
    - Epilepsy EDC is created when both phenytoin (dilantin) and gabapentin (neurontin) are prescribed.

- To assign a severity of illness level for a PCD, a similar process was used. The only difference was that once a drug was used it could only be used again for another potential PCD in the same MDC. As only a single PCD per MDC was allowed, no drug would be counted twice. If drugs were allowed to be used outside of the MDC, the same drug could assign a severity of illness level to two different PCDs which would then be reflected in the ultimate CRG assignment. Table 2 contains some examples of severity levels for congestive heart failure based on pharmacy data.
- The assignment of severity of illness levels for Status 8 (malignancy) and Status 9 (catastrophic) CRGs was identical to the process used for PCDs except that there was no elimination of drugs. As only a single CRG could be assigned, setting multiple minimum levels would have no effect on the ultimate CRG assignment.

Level I	Digoxin
Level II	Digoxin & Ace Inhibitor Ace Inhibitor & Diuretic
Level III	Digoxin & Lasic & Metolozone Digoxin & Lasic & HCTZ
Level IV	Digoxin & Diuretic & Amiodarone

Table 2. Examples of severity levels based on pharmaceutical data for congestive heart failure

Figure 1 contains an overview of the integration of the pharmacy data into CRGs.

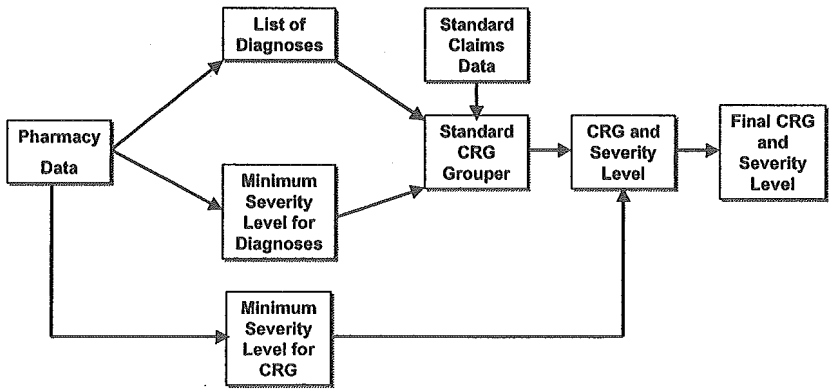

Figure 1. Integration of pharmacy data into CRGs

Data

The data used for this analysis consisted of two years of claim, demographic, and enrollment data from a commercial insurance data base. For inclusion in the analysis, an individual was required to have 24 months of continuous exposure except for newborns born in Year 1. Newborns were required to have at least 2 months exposure at the end of Year 1 and 12 months in Year 2. The final sample yielded 309,196 enrollees. Table 3 shows the distribution of enrollees by age and sex. This is a young population with about one-third (33.2%) of the population under 20 and slightly more than half (51.5%) between the ages of 20 and 50. Males outnumber females for all age ranges. This is unusual as it is contrary to the age distributions of the general population where females outnumber males for all but the younger age groups.

Age	Male	Female	Total	Percent
0 - 4	13,472	12,342	25,814	8.3%
5 - 9	14,238	13,770	28,008	9.1%
10 – 14	14,086	13,724	27,810	9.0%
15 – 19	10,586	10,397	20,983	6.8%
20 – 24	6,422	5,886	12,308	4.0%
25 – 29	10,164	9,444	19,608	6.3%
30 – 34	13,966	13,520	27,486	8.9%
35 – 39	18,210	17,657	35,867	11.6%
40 – 44	17,867	17,243	35,110	11.4%
45 – 49	14,882	13,993	28,875	9.3%
50 – 54	10,984	10,811	21,795	7.0%
55 – 59	7,741	7,318	15,059	4.9%
60 – 64	4,310	3,710	8,020	2.6%
65 +	1,441	1,012	2,453	0.8%
Total	158,369	150,827	309,196	100.0%

Table 3. Distribution of enrollees by age and sex

The distribution of expenditures for both years are typical of a healthy population (See Tables 4 and 5). The average expenditure for Year 1 was $1,627 and for Year 2 was $1,880, an increase of about 15%. For both years most individuals had minimal expenditures (less than $1,000). Expenditures were concentrated in a relatively small portion of the population. In Year 1, 7.0% of enrollees incurred 56.6% of expenditures. In Year 2, 8.2% of the enrollees incurred 64% of the expenditures.

As shown in Table 6, only 4.2% of enrollees were hospitalized in Year 1, while 81.1% had an outpatient visit, and 56.1% received a prescription drug. All but ten enrollees with an admission also had an outpatient visit. Most enrollees with a hospitalization also received a prescription drug.

Year1 Payments	Cases	Percent of Cases	Total Payments	Average Payment	Percent of Payments
0	48,512	15.7%	$0	$0	0.0%
0 < x <1000	166,859	54.0%	$59,947,651	$359	11.9%
1000 <= x <2000	39,420	12.7%	$55,897,377	$1,418	11.1%
2000 <= x <5000	32,770	10.6%	$102,351,124	$3,123	20.3%
5000 <= x <10000	13,645	4.4%	$94,556,499	$6,930	18.8%
10000 <= x <20000	5,438	1.8%	$73,397,387	$13,497	14.6%
20000 <= x <50000	1,961	0.6%	$57,731,681	$29,440	11.5%
50000 <= x <100000	423	0.1%	$28,318,589	$66,947	5.6%
100000 <= x <250000	143	0.0%	$21,727,235	$151,939	4.3%
250000 <= x <500000	23	0.0%	$7,547,541	$328,154	1.5%
500000 <= x	2	0.0%	$1,669,297	$834,649	0.3%
Total	309,196	100.0%	$503,144,381	$1,627	100.0%

Table 4. Distribution of expenditures in year 1

Year 2 Payments	Cases	Percent of Cases	Total Payments	Average Payment	Percent of Payments
0	46,468	15.0%	$0	$0	0.0%
0 < x <1000	160,439	51.9%	$58,669,455	$366	10.1%
1000 <= x <2000	40,872	13.2%	$58,143,179	$1,423	10.0%
2000 <= x <5000	36,082	11.7%	$113,336,876	$3,141	19.5%
5000 <= x <10000	15,453	5.0%	$107,178,194	$6,936	18.4%
10000 <= x <20000	6,532	2.1%	$88,638,747	$13,570	15.2%
20000 <= x <50000	2,598	0.8%	$77,118,655	$29,684	13.3%
50000 <= x <100000	522	0.2%	$35,798,408	$68,579	6.2%
100000 <= x <250000	196	0.1%	$28,466,505	$145,237	4.9%
250000 <= x <500000	28	0.0%	$9,172,278	$327,581	1.6%
500000 <= x	6	0.0%	$4,888,131	$814,689	0.8%
Total	309,196	100.0%	$581,410,428	$1,880	100.0%

Table 5. Distribution of expenditures in year 2

	Enrollees	Percent
Inpatient	13,102	4.2%
Outpatient	250,889	81.1%
Drug	173,388	56.1%
Inpatient & Outpatient	250,899	81.1%
Inpatient & Drug	176,733	57.2%
Outpatient & Drug	259,540	83.9%
Inpatient, Outpatient & Drug	259,549	83.9%
Total	309,196	100.0%

Table 6. Number of enrollees with inpatient, outpatient and pharmacy claims

Methodology

The data from year 1 was used to assign the CRG. The CRG based on year 1 was then used as the basis for the prediction of year 2 expenditures. In some of the analysis a stop loss was applied to year 2 expenditures as follows.

A_i = Actual plan expenditure for enrollee i

S = Stop loss cap

N = Number of enrollees

$$A'_i = min\left[A_i, S\right]$$

$$D = \sum_i A_i - \sum_i A'_i$$

$$A''_i = A'_i + D/\mathrm{N}$$

The impact of the stop loss is summarized in Table 7. Although the stop loss affected less than one percent of the enrollees, the expenditures over the cap constituted 7.31 and 13.47 percent of total expenditures for the $100,000 and the $50,000 cap, respectively.

Cap	Enrollees Over Cap	Percent Enrollees Over Cap	Expenditures Over Cap	Percent Expenditures Over Cap
$50,000	752	0.24%	$78,325,322	13.47%
$100,000	230	0.07%	$42,526,914	7.31%

Table 7. Impact of Cap

The CRG predicted year 2 expenditures were computed as follows.

$E_{g,y+1}$ = Predicted expenditures in year y+1 for enrollee assigned to CRG g in year y

$N_{g,y}$ = Number of enrollees assigned to CRG g in year y

$A''_{i,g,y+1}$ = Actual expenditure in year y+1 with stop loss for i^{th} enrollee assigned to CRG g in year y

$$E_{g,y+1} = \left(\sum_{i} A''_{i,g,y+1} \right) / N_{g,y}$$

$E_{g,y+1}$ is normalized to be budget neutral

$$\sum_{g} E_{g,y+1} N_{g,y} = \sum_{i,g} A''_{i,g,y+1}$$

The most common statistical measure used to compare risk adjustment systems is reduction of variance (R^2), which measures the proportion of variation that is explained by a risk adjustment system. R^2 provides a summary measure of the extent to which the risk adjustment system is able to predict the value of future expenditures. R^2 ranges between 0 and 100 and measures the percentage of variation in future expenditures explained by the risk adjustment system. Thus, an R^2 of 10.15 would mean that 10.15 percent of the variation in future expenditures is explained by the risk adjustment system. The average expenditures in year 2 for enrollees assigned to CRG g in year 1 are used to predict the actual expenditures in year 2 for each enrollee ($E_{g,y+1}$ predicts $A''_{i,g,y+1}$)

The predictive ratio can also be used as an evaluation statistic. The predictive ratio is computed as the predicted expenditures divided by the actual expenditures in the prediction year. Since the calculations are performed on a budget neutral basis, the predictive ratio across the entire database will be 1.0. However, the predictive ratio can be examined for subsets of the database. A predictive ratio greater than one means that in the prediction year, predicted expenditures are greater than actual expenditures. A predictive ratio of 1.2 means that predicted expenditures are 20 percent higher than actual expenditures. Conversely, a predictive ratio of 0.8 means that predicted expenditures are only 80 percent of actual expenditures. The predictive ratio is computed as the predicted CRG expenditures divided by the actual expenditures in the prediction year as follows.

$$\frac{\sum_{i \in x} E_{i,g,y+1}}{\sum_{i \in x} A''_{i,g,y+1}}$$

where x is any subset of enrollees (e.g., males)

The R^2 statistic and the predictive ratio are used to assess the impact of pharmacy data on the performance of the CRG risk adjustment methodology.

Results

The primary purpose of this study is to ascertain the feasibility of complementing or replacing outpatient data, which may be costly to obtain and which are frequently unavailable, with drug data from stand alone pharmacy programs. For the purposes of this analysis there are four key combinations of data inputs for risk adjustment:

- Inpatient only data
- Inpatient and drug data
- Inpatient and outpatient data
- Inpatient, outpatient, and drug data

For completeness another three combinations of data sources will also be presented.

- Outpatient only data
- Drug only data
- Outpatient and drug data

To analyze the impact of these different data sources upon the performance of CRGs, three general topics will be discussed:

- The movement between CRGs
- The predictive power associated with the different data sources (R^2)
- The performance of the different data sources with predictive ratios

Movement Between CRGs

Appendix A contains the CRG assignments for all seven models at the ACRG3 level of aggregation. The basic health of this population is evident. When only inpatient data are used, 303,306 (98.1%) enrollees are categorized as healthy, (i.e., they have no identifiable chronic conditions and no recent significant acute problems). The inclusion of drug data decreases the number of enrollees classified as healthy to 205,978 (66.7%). When inpatient and outpatient data are available 246,521 (79.7%) enrollees are classified as healthy. Adding drug data decreases the number of enrollees classified as healthy to 182,931 (59.2%). The movement from the healthy status when additional information is made available to the CRG grouper occurs across all

health statuses. Compared to the results from inpatient and outpatient data, the additional use of drug data (inpatient, outpatient, and drug) results in roughly a doubling the of the number of enrollees in all statuses except for Status 8 (malignancy) and Status 9 (catastrophic). This reflects the fact that drug data from pharmacy benefit programs will identify relatively few cases of metastatic malignancies or catastrophic cases that are not identified by standard claims data. In summary the inclusion of drug data substantially increases the number of enrollees identified with chronic or recent significant acute problems.

Appendix B displays the CRG assignment results for all seven models at the CRG level. It can be used to identify the specific CRGs which increased in volume the most when drug data are included in the CRGs. It should be noted that these data understate the number of new cases for many diseases. Due the hierarchical structure of CRGs the identification of previously unreported diseases increases the number of enrollees assigned to CRGs with multiple comorbid diseases (i.e., status 4, 6 or 7). Appendix C and D displays the movement of enrollees between base CRGs (CRGs without a severity of illness level) when drug data are added to inpatient and outpatient data. Table 8 shows some examples of the impact on risk group assignment when drug data are included:

Base CRG	Description	Inpatient Data	Outpatient Data	Drug Data	Inpatient & Outpatient Data	Inpatient & Drug Data	Outpatient & Drug Data	Inpatient, Outpatient, & Drug Data
3018	Migraine	34	795	1,556	791	1,547	1,568	1,569
3445	Hyperlipidemia	17	1,668	3,468	1,675	3,387	3,238	3,230
3525	Other Chronic Gyn. Diagnoses – Minor	144	919	8,238	818	8,902	8,331	8,238
3755	Depression	34	586	4,054	591	3,972	3,626	3,610
5138	Asthma	223	2,595	5,371	2,629	5,461	6,485	6,510
5179	Congestive Heart Failure	69	84	450	101	439	407	409
5192	Hypertension	11	3,948	15,547	3,894	15,006	14,682	14,491
5424	Diabetes	167	2,152	1,981	2,152	1,961	2,121	2,110
5743	Schizophrenia	10	9	467	17	454	424	420
6144	Diabetes & Hyper-tension	0	133	1,387	134	1,327	1,387	1,367
9030	HIV Disease	8	136	139	136	141	161	161

Table 8. Examples of changes in base CRG assignment

In the examples in Table 8, all but diabetes and HIV disease reveal a large number of enrollees being treated for conditions which are not reported with inpatient and outpatient data. In the case of migraines, large numbers of enrollees are receiving recurrent prescriptions for sumatriptan or butalbital (See Appendix E which identifies the most common drugs prescribed for each CRG). Similarly, the growth in the volume of enrollees in the CRG for Other Chronic Gynecological Diagnoses – Minor reflects the large numbers of women over 49 years old receiving estrogen supplements at some point during the year in which the CRG was assigned. Hyperlipidemia is another condition where significant numbers of enrollees are being treated for high cholesterol and triglyceride levels without having it reported for inpatient or

outpatient encounters. Similar statements can be made for asthma, hypertension, congestive heart failure, and schizophrenia. Diabetes does not appear to increase in volume but the data is deceptive. Many diabetics had additional illnesses also identified and were assigned to Status 6 or Status 7 for enrollees with multiple chronic diseases. For example, CRG 6144, Diabetes and Hypertension increased from 0 enrollees when only inpatient data were available to 1,327 enrollees when inpatient and drug data were available. In the examples in Table 8, only HIV Disease did not increase in volume appreciably. For HIV, however, it should be noted that while the inclusion of drug data resulted in the identification of relatively few new enrollees, inpatient and drug data worked about as well as inpatient and outpatient data and almost as well as using data from all three sources.

R^2 Results

R^2 statistics were calculated at the CRG and ACRG3 level for each data model and for three different stop-loss (cap) levels (See Table 9). The first thing that is apparent is that at the CRG level the highest R^2 without any stop loss cap are for inpatient only data (21.02%). Adding additional data, be it, outpatient, drug, or both reduces R^2. This is counterintuitive. More importantly it is misleading. The unexpected R^2 results reflects the sensitivity of R^2 to a few high cost enrollees. The health status of this population which as very few individuals with high cost illnesses and the categorical structure of CRGs which isolates high cost cases results in the high cost enrollees being in otherwise empty cells. When these cells gain additional cases, as happens when CRGs are aggregated to the ACRG3 level or when additional cases are assigned to the CRGs because of additional data the R^2 decreases. Putting a maximum (cap) on the total expenditures attributed to an individual ($50,000 or $100,000) significantly reduces the influence of a few extreme high cost enrollees and also decreases the R^2. It is also almost certain that if exogenous weights (estimates of year 2 expenditures based on another source of data) were used rather than actual average values for each CRG in year 2 that the R^2 for inpatient only data would be significantly reduced.

When capped expenditures are used in the calculation of R^2, a clearer picture emerges. Outpatient data are clearly the strongest of the three sources. Either alone or in combination with the other data types, outpatient data produces the highest R^2. The highest R^2 for both CRGs, (18.42% and 19.03%) and

Claim Type	$50,000 Cap		$100,000 Cap		No Cap	
	CRG	ACRG3	CRG	ACRG3	CRG	ACRG3
Inpatient	8.44	6.33	10.85	7.46	21.02	7.13
Outpatient	16.25	13.18	17.14	12.99	16.12	9.64
Drug	9.64	8.18	8.13	6.52	5.07	3.63
Inpatient & Outpatient	16.91	13.69	18.08	13.69	20.65	10.61
Inpatient & Drug	14.68	11.61	15.15	10.81	16.07	7.65
Outpatient & Drug	17.75	14.61	18.15	13.69	17.35	9.49
Inpatient, Outpatient & Drug	18.42	15.11	19.03	14.34	18.63	10.21

Table 9. R^2 results

ACRG3s (15.11% and 14.34%) for $50,000 and $100,000 cap, respectively, occurs when all three sources are used.

If outpatient data are not available, using inpatient and drug data results in an R^2 of 14.68% with a $50,000 cap and 15.15 with $100,000 cap. At the ACRG3 level, the comparable figures are 11.61 and 10.81.

Predictive Ratios

Predictive ratios are the ratio of predicted expenditures over the actual expenditures for a given segment of the population. The following analysis, looks at predictive ratios by expenditure group, demographic factors and selected disease categories.

Predictive Ratios – Quintiles

The quintile predictive ratio analysis was performed for all seven combinations of data, by CRG and ACRG3. The quintiles (See Tables 10 and 11) are computed based both on the prior year (year 1) and the predicted year (year 2).

At the CRG level the predictive ratios perform as is typical for expenditure quintiles and risk adjustment methodologies. Lower quintiles, composed of healthier individuals, are overpaid (predicted expenditures greater than actual expenditures) and the higher quintiles, composed of sicker individuals, tend to be underpaid (predicted expenditures less than the actual expenditures). This occurs across all data source and for both prior year and predicted year quintiles.

When quintiles based on prior year expenditures are compared to quintiles based on predicted year expenditures, the prior year predictive ratios are more accurate (closer to 1.00). This is to be expected as the prior year expenditure quintiles do not reflect the actual expenditures in the prediction year while the predicted year quintiles do reflect the actual expenditures in the prediction year.

The predictive ratios move closer to 1.00 as more data are included. For both prior year and predicted year quintiles, the most accurate predictive ratios occur when all three data sources are included. The least accurate occurs when only inpatient data are used to assign CRGs. It is interesting to note the differences between the drug and the outpatient data. The drug data tends to produce better results for Quintiles 1 and 2 than comparable outpatient data when used alone or in combination with inpatient data. The opposite is true for Quintiles 3,4, and 5 where outpatient data tends to be more accurate.

At the ACRG3 of aggregation, the accuracy of the predictive ratio is reduced. This is to be expected as the aggregation of the data by merging CRGs into ACRG3s reduces sensitivity to the differences between CRGs. This effect is weakest for the first quintile because this quintile is composed of healthiest people, many of whom incurred little or no expenses in either year. These individuals are assigned to the healthy CRG which is not affected by the aggregation of CRGs into ACRGs.

Quintile	Inpatient	Outpatient	Drug	Inpatient & Outpatient	Inpatient & Drug	Outpatient & Drug	Inpatient & Outpatient & Drug
Prior Year Expenditures							
1	2.511	1.903	1.839	1.896	1.782	1.550	1.545
2	1.988	1.554	1.589	1.547	1.539	1.378	1.372
3	1.355	1.174	1.273	1.167	1.230	1.169	1.162
4	0.854	0.910	1.038	0.902	1.002	1.014	1.007
5	0.528	0.737	0.652	0.746	0.700	0.782	0.789
Predicted Year Expenditures							
1	12.655	10.124	9.603	10.101	9.357	8.501	8.490
2	6.039	5.059	4.919	5.047	4.801	4.493	4.484
3	2.932	2.655	2.731	2.648	2.667	2.585	2.578
4	1.359	1.441	1.574	1.435	1.546	1.565	1.559
5	0.315	0.416	0.399	0.418	0.420	0.454	0.457

Table 10. CRG predictive ratios by expenditure quintiles and data source

Quintile	Inpatient	Outpatient	Drug	Inpatient & Outpatient	Inpatient & Drug	Outpatient & Drug	Inpatient & Outpatient & Drug
Prior Year Expenditures							
1	2.511	1.903	1.841	1.895	1.785	1.553	1.548
2	1.988	1.563	1.611	1.555	1.564	1.414	1.408
3	1.355	1.196	1.294	1.188	1.259	1.208	1.200
4	0.854	0.933	1.049	0.924	1.024	1.036	1.028
5	0.528	0.718	0.636	0.728	0.676	0.752	0.760
Predicted Year Expenditures							
1	12.685	10.177	9.656	10.160	9.416	8.572	8.560
2	6.055	5.108	4.980	5.097	4.868	4.577	4.568
3	2.946	2.706	2.768	2.699	2.714	2.645	2.638
4	1.370	1.468	1.588	1.462	1.570	1.587	1.582
5	0.310	0.403	0.390	0.405	0.408	0.440	0.442

Table 11. ACRG3 predictive ratios by expenditure quintiles and data source