ABSTRACT

OBJECTIVE: To examine challenges in forming accurate conclusions from claims data when comparing different patient groups, and to stress the need to adjust for confounding variables through application of statistical methods.

METHODS: Patients who were 60 or older and began amlodipine or felodipine in 1997 or 1998 were identified from a proprietary managed care claims database. Patients had 6 months of medical and outpatient pharmacy benefits coverage before and after initiating therapy. They were stratified on varying degrees of hypertension severity and on all-condition clinical complexity, and quantified with a proprietary Burden of Illness (BOI) score, defined as the expected resource use for the patient's combined conditions during the 6 months before therapy.

RESULTS: Of the 17,667 patients, 10,469 (59.3%) were female. The 12,389 amlodipine users were similar in age to felodipine users (n=5,278), 73 versus 74, respectively. More amlodipine users were assigned to the 2 most clinically complex quintiles of the BOI distribution (45.5% to 26.9%), with 24.3% of patients in the top quintile (9.7% for felodipine) and twice the prevalence of ischemic heart disease/angina (42.4% versus 19.4%). High-severity hypertension was found among 33.6% of amlodipine users, contrasted with 17.9% for felodipine. A significant difference in the mean average daily dose (ADD) between the amlodipine (5.93 mg) and felodipine (5.8 mg) groups (P=0.0007) was noted.

CONCLUSION: These results stress the importance of assessing the impact of potentially confounding variables and the need to address differing patient characteristics through either risk adjustment or stratification in making head-to-head comparisons.

KEYWORDS: Risk adjustment, Amlodipine, Felodipine, Burden of illness

JAUMCP 2002(8):5:353-59

A number of important studies have examined the methodological challenges inherent in retrospective analyses using administrative claims data4,5 and in conducting cohort studies following one or more groups.4,5 This body of literature describes the challenges of drawing broad conclusions from comparison groups with different underlying characteristics, and underscores the value and importance of rigorous risk adjustment6 in order to draw accurate and appropriate conclusions from the data.

For example, in a medical chart and administrative claims-based retrospective cohort study examining how well cardiologists, medical subspecialists, internists, and family practitioners adhered to established practice guidelines for an elderly population with acute myocardial infarction, Frances, et al., found, contrary to expectation, that differences in the use of recommended therapies by physician specialty are generally small and do not explain differences in patient outcomes. After adjusting for patient characteristics, medications, and revascularization procedures, it was observed that treatment by a cardiologist was not associated with greater short-term survival. However, patients treated by a cardiologist had lower one-year mortality, even after medications and revascularizations were included in the models.

This study noted that residual confounding from unmeasured patient or hospital characteristics, differences in unmeasured processes of care during or after hospitalizations, or both, were responsible for the observed survival advantage of patients seen by cardiologists. Adjusting for differences among patients and hospitals had a large effect on the association between physician specialty and patient mortality, whereas adjusting for differences in the use of recommended medications and revascularization had a relatively small effect, suggesting that residual confounding may account for much of the remaining survival benefit for patients treated by cardiologists.

A separate Framingham study comparing the prognosis of 2 groups appeared to demonstrate that the rate of stroke in patients with atrial fibrillation and rheumatic heart disease was 41 per 1,000 person years, similar to the rate for patients with atrial fibrillation and no rheumatic heart disease. Once adjustments were made for age, sex, and hypertensive status of the patients, however, the rate of stroke was 6-fold greater in patients with rheumatic heart disease and atrial fibrillation than in patients with atrial fibrillation who did not have rheumatic heart disease. These examples emphasize the need to ensure
similar patient demographics among comparison groups and to adjust for important prognostic factors and severity to preserve valid outcome results.

**Objectives**

This article presents an administrative claims-based case study approach to examine how a rigorous statistics methodology will support understanding of “real world” use of commonly prescribed drugs within the same class in a managed care plan. In this case, 2 calcium channel blockers (CCBs), amlodipine and felodipine, were selected, as these 2 agents have similar indications. The primary aim of this case study is to examine the challenges in forming accurate conclusions from claims data when comparing 2 different managed care patient groups using 2 medications with similar indications, and to stress the need to adjust for confounding variables through application of statistical methods. As this study demonstrates, a simple comparison of health care resource utilization for elderly patients treated with amlodipine and felodipine would lead to different conclusions than a comparison that also quantifies clinical complexity differences, a potential confounding variable that is not routinely assessed, between the 2 treatment groups.

**Methods**

**Data Source**

This study consisted of a retrospective administrative claims analysis derived from 3 million managed care plan members, representing a wide geographic distribution. Members resided in 22 states, concentrated in the South, Southwest, and Midwest, and 72% of the managed care database members belonged to health maintenance organization (HMO) or preferred provider organization (PPO) plans.

**Study Selection and Data Extraction**

All patients were age 60 or older and, by study design, first began using amlodipine or felodipine in 1997 or 1998. There was no requirement that the patients have a diagnosis code indicating hypertension in their claims records. Each patient was required to have a minimum of 6 months of medical and outpatient pharmacy benefits coverage prior to and after the date of amlodipine or felodipine initiation (the “index date”) to adequately support assessment of comorbidity profiles, concomitant medications used, and health care charges accrued. Persons using either amlodipine or felodipine during the 6 months prior to their index date were excluded to increase the likelihood that study patients were first-time users of amlodipine or felodipine. Patients using both amlodipine and felodipine in 1997 or 1998 were also excluded to create mutually exclusive treatment populations.

**Data Synthesis**

To assess clinical differences between the 2 treatment groups, patients were assigned to one of 3 separate hypertension severity categories. The severity-of-illness calculation starts by assigning each diagnosis within a condition to one of 3 severity levels: usual, moderate, and high. Examples of clinically assigned severity levels for diagnoses within the condition of hypertension are:

- **Usual** 401.9 unspecified essential hypertension
- **Moderate** 402.11 benign hypertension with congestive heart failure
- **High** 403.01 malignant hypertension with renal failure

In addition, the severity-of-illness calculation makes use of additional diagnoses belonging to different conditions—diagnoses that can represent complications or higher severity for the condition being analyzed. Thus, in the process of distinguishing “usual severity” hypertension from “moderate severity” or “high severity” disease, the severity calculation looks beyond the diagnoses that actually mention hypertension.

In the hypertensive patient, for example, the presence of the late effect of a stroke (code 438.20, hemiplegia) or peripheral vascular disease (code 443.9) can be used to indicate moderate-severity hypertension. If papilledema (code 377.00) or kidney transplant status (V42.0) is present in a hypertensive patient, the patient will be assigned to the high-severity level of hypertension, even though neither of these codes mentions hypertension.

In summary, a patient's diagnoses both within or beyond the hypertension diagnosis itself play a role in determining the severity level of hypertension. The severity level assigned to a patient will be the highest level that can be assigned, taking all diagnoses into account.

Hypertension was the clinical focus because it is the most common indication for these therapies, and nearly all study patients carried this diagnosis. To adjust for differences in patient case mix, a summary measure known as Burden of Illness (BOI) was employed to characterize each individual patient's history of illness over time. The BOI measure uses 189 diagnostic condition categories and the 3-level severity classification per category to assign an aggregated single illness weight to each patient. Hypertension severity is reflective of only a single condition, whereas clinical complexity is reflective of the patient's total burden of illness, which takes into account all the conditions a patient has, and, as part of that process, considers the severity of each of those conditions.

The BOI measure reflects a patient's expected consumption of health care resources. It is the sum of the expected resource needs for each of the conditions occurring in an individual patient's history and can be used to adjust for specific confounding variables. These adjusted variables serve as the basis for making comparisons between 2 or more groups. On average, patients with a higher BOI score should have greater resource need.

In a previous study, Solz and colleagues employed the BOI...
methodology as part of an analysis to demonstrate how health claims data can support the comprehensive goals of a disease management program by targeting conditions and the gaps in quality of care and variations of practice. BOI calculations and the identification of comorbidities were based on all claims during the 6 months prior to the drug-index date. Based on BOI score, each patient was assigned to one quintile, with the first quintile representing the lowest clinical complexity and the fifth quintile representing the highest. The average daily dose (ADD) was estimated for all patients for each month of use through 6 months post index date. In each month, the ADD was estimated by dividing the total number of milligrams dispensed by the total number of days supplied per prescription.

**Statistical Analysis**

Comparison testing for facility and professional services was performed using the Hypothesis Test for Sample Proportions. Comparison testing on the ADD and mean amlodipine and felodipine BOI scores was performed using t tests and Wilcoxon rank sum tests. The Wilcoxon tests are nonparametric tests that do not assume that the ADD or BOI scores are from a normal distribution. These tests use the ranks of the scores, not the scores themselves. Although the t tests are quite robust in this context (with 2 samples and where any skewness is likely in the same direction for both samples), the Wilcoxon results provide greater confidence in any significant differences found. Comparison testing on demographics, hypertension severity, and BOI quintiles between the 2 treatment groups was performed using chi-square tests.

**Results**

**Demographics of Study Population**

Of the 17,667 patients in the study population, 10,469 (59.3%) were women, and the mean age was 73 (Table 1). Mean age was similar between amlodipine users (73 years) and felodipine (74 years) users, however the age distribution was different between the 2 groups, as 39.2% of all amlodipine patients and 33.9% of felodipine patients were younger than 70 years. The utilization of amlodipine and felodipine was significantly different between women and men (P<0.0001), as well as significantly different across age groups for women (P=0.0008) and men (P=0.0012).

**Choice of Therapeutic Agent**

Of the 17,667 study patients, 12,389 (70.1%) were amlodipine users. There was a higher percentage of amlodipine users than felodipine users in every age group/gender category. During the 6 months prior to initiating amlodipine or felodipine use, almost 26% of the entire population used another CCB. 23% used beta-blockers, and 34% used angiotensin converting enzyme (ACE) inhibitors. During the 6 months following initiating amlodipine or felodipine therapy, 8.9% used other CCBs, 27% used beta-blockers, and 32% used ACE inhibitors. Felodipine users were slightly more likely to have used a non-amlodipine/non-felodipine CCB during the study period (9.1% and 8.8%, respectively). Amlodipine users were more likely to have used medications in the other 3 antihypertensive classes (Table 2).

In the 6-month post-index period, 54% of amlodipine users and 39% of felodipine users had at least one facility service, which included inpatient and outpatient hospitalizations, emergency room and ambulatory care visits, skilled nursing, independent lab, and home health services, and 86% of amlodipine users had at least one professional service, compared with 84% of felodipine users (P<0.05). Professional services included outpatient procedures, visits, and tests. By study design, all study patients used at least one pharmacy service.
Antihypertensive Use 6 Months Before And After the Index Date

<table>
<thead>
<tr>
<th>Antihypertensive</th>
<th>Utilization During the 6 Months Prior to the Index Date</th>
<th>Utilization During the 6 Months After the Index Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Amlodipine</td>
<td>Felodipine</td>
</tr>
<tr>
<td>Beta-blockers</td>
<td>25%</td>
<td>19.6%</td>
</tr>
<tr>
<td>Angiotensin converting enzyme (ACE) inhibitors</td>
<td>34.8%</td>
<td>31.7%</td>
</tr>
<tr>
<td>Angiotensin II inhibitors</td>
<td>1.6%</td>
<td>7%</td>
</tr>
</tbody>
</table>

Hypertension Severity

Of the 17,667 patients, 2,885 (16%) were not included in the hypertension severity subanalysis because their claims history did not explicitly include hypertension diagnoses, which was the basis for severity classification. Amlodipine users were more likely to have a moderate or high hypertension severity (7.6% of all moderate-severity patients and 81% of all high-severity patients) than felodipine users (25.4% of all moderate-severity patients and 19% of all high-severity patients). The difference in proportion of moderate severity, high severity, and moderate/high severity combined between amlodipine and felodipine users was significant in all cases (P<0.05). It is important to note that these statistics reflect the proportion of moderate/high patients compared to all patients in the sample(s).

Clinical Complexity Relative to Drug Choice

To assess potential differences in the clinical complexity and the impact on patient management and utilization between the 2 groups, comorbidity profiles were established for each patient.

Burden of Illness (BOI) Profiles

As calculated from information from the pretreatment period, the mean amlodipine BOI score was significantly higher than the mean felodipine BOI score (P=0.0001 for both t test and Wilcoxon). All patients were sorted into quintiles reflecting the distribution of BOI over the entire study population. A greater number of amlodipine patients were assigned to the higher quintiles (chi-square P<0.0001) whereas felodipine patients were more heavily represented in the lower quintiles (chi-square P<0.0001).

As seen in Figure 1, the fifth quintile, or the most clinically complex patients, contained 24.3% of the entire amlodipine population and 9.7% of the felodipine population. Amlodipine patients were also more common than felodipine patients in the fourth BOI quintile (21.2% of amlodipine users versus 17.2% of felodipine users). The observed distributions among the quintiles for both treatment groups differed significantly from the expected distributions (chi-square P<0.0001). Amlodipine patients were more likely to be assigned to the 2 most clinically complex quintiles, 45.5% to 26.9%, and felodipine users predominated in the 2 lowest BOI quintiles, demonstrating a lower level of expected resource consumption.

The presence of coronary artery disease, including angina, increased a patient's BOI and classified a patient in the moderate hypertension severity category, at a minimum. Thus all patients, regardless of study drug, with diagnoses of hypertension and angina were assigned to at least the moderate hypertension severity category; 42.4% of amlodipine patients and 19.4% of felodipine patients had angina.

Conditions affecting the circulatory system were the most common conditions among all study patients, with the percentage of patients with a circulatory system condition (including hypertension) increasing with each BOI quintile. That is, in the lowest quintile 30.6% of amlodipine users and 37.9% of felodipine users had a circulatory condition, and in the highest quintile, 98.7% of all amlodipine users and 96.7% of felodipine users had a circulatory condition. This distribution was significantly unlike the expected distribution (chi-square P<0.0001).

Clinical Profiles

In the 6 months preindex and postindex date, the 10 most frequently occurring conditions were hypertension, ischemic heart disease (IHD)/angina, chest pain, lipid disorder, diabetes, cough/respiratory signs, arrhythmia, other joint disease, other heart disease, and congestive heart failure (CHF). More than 92% of all patients had at least one of these conditions. Proportions in the amlodipine and felodipine populations (93% versus 92%, respectively) were statistically alike. During the study period, 86% of felodipine users and 83% of amlodipine users had hypertension; however, this difference in proportion was not statistically significant. Each of the other 9 conditions was seen in higher proportion among amlodipine users (P<0.05 or less), with amlodipine users having 2 or more times the observed prevalence of IHD/angina, chest pain, other heart disease, and CHF than felodipine users (Table 3).

The ADD of amlodipine and felodipine decreased with increasing age from between 5.9 mg (women) to 6.4 mg (men) mgs/day among those age 60 to 64, to between 5.0 mg (women) to 5.7 mg (men) mgs/day among those age 90 and older. Generally, men used higher doses than women, and amlodipine users had higher average doses than felodipine users. Overall, the ADDs increased by 0.4 to 0.5 mgs/day from the initial dose to the dose used in month 6. From a clinical perspective, numerically, the mean ADD between the 2 groups was minor. However, statistical testing revealed significance
with the mean ADD between the amlodipine (5.93 mg) and felodipine (5.8 mg) groups (P=0.0007).

**Discussion**

**CCB Utilization**

In this large sample of managed care patients, there was a higher percentage of amlodipine users than felodipine users in every age group/gender category. Within this elderly population, age 60 or older, the younger patients used higher doses than older patients. This may be due to younger patients being treated more aggressively with CCBs alone or the elderly using lower doses of CCBs along with concomitant antihypertensive or antianginal therapies. Age is one of the most commonly controlled-for variables in research studies of all kinds. However, this finding—that dose is inversely associated with age—clearly illustrates the importance of controlling for age in comparing either resource use or clinical outcomes in 2 different treatment groups.

In the current study, ADD increased following initiating amlodipine or felodipine therapy. This suggests that more patients had their initial dose titrated up than had their initial dose titrated down within 6 months following amlodipine or felodipine initiation. Amlodipine ADDs were slightly higher than felodipine. Reasons for this difference may be amlodipine’s potency and effectiveness in more clinically challenging patients, or amlodipine’s additional indication for angina, for which most patients take 10 mg/day.

**Clinical Profiles**

This study also demonstrates that there are other potential confounders at work. In particular, there are significant clinical differences between patients who are typically treated with amlodipine and felodipine. Amlodipine users were more clinically complex, as evidenced by their significantly higher BOI score, an overall measurement that takes into account both their cardiovascular and noncardiovascular conditions. Not only were these patients more likely to have a greater variety of heart disease conditions compared to felodipine users but those with explicitly reported hypertension (86% of felodipine users and 83% of amlodipine users) were also being treated for more severe manifestations of that disease.

The reasons for the observed prescribing preferences were not explored in this claims-based study, although it is true that amlodipine has an additional indication, i.e., for the treatment of angina pectoris. Channeling bias (i.e., drugs with similar therapeutic indications prescribed to groups of patients with prognostic differences) may be another potential reason for observed prescribing preferences. That is, assumed advantages of one particular drug may “channel” that drug to patients with a particular preexisting comorbidity, thus incorrectly attributing the use of the drug to the patient’s condition.
Exploring the Methodological Challenges of Investigating Comparison Groups with Different Underlying Characteristics: A Case Study

One taxing issue observed in the current study that must be addressed by managed care pharmacists in order to draw accurate conclusions from these results is ensuring similar comparison groups with respect to important determinants of outcomes and adjustment of residual differences. In a previous study, Naylor et al. examined the challenges of comparing outcomes of 2 or more sets of health care practitioners or delivery systems.13 When performing comparisons, selection biases may arise, resulting in differing patient populations that affect the likelihood of a good or bad outcome. Researchers must adjust for differences between patient groups or risk threatening the validity of research by case selection biases, creating noncomparable groups of patients and confounding any outcomes comparisons.13

Although the 2 populations in the current study were initially perceived to be similar based on their treatment as described in the data, application of statistical methods revealed that the amlodipine group was more resource-intensive than the felodipine group. Without mining the data further, one could conclude that the amlodipine patients are simply more expensive patients. However, the underlying question is why? After employing statistical methods that considered the overall illness burden and the severity of hypertension, the amlodipine group was shown to be significantly more clinically complex and had more severe hypertension than the felodipine group, providing a likely reason for the additional resources used for the management of hypertension and other comorbidities.

This interpretation underscores the relevance to managed care pharmacists in assessing the impact of potential confounding variables, particularly when performing cost, safety, and efficacy comparisons of various treatment regimens, to establish policies on the use of the most appropriate and effective pharmacologic agents. Managed care pharmacists should perform clinical and financial assessments cautiously, ensuring comparable patient populations, and stratify patients according to severity, as 2 products with perceived similarities may not, in fact, be similar.

From a broad perspective, results from this study provide concrete examples of potential confounders and the importance of addressing these factors through appropriate study design and statistical methods. This case study offered one statistical approach that considered confounding variables when comparing different patient groups. Opportunities exist to perform further research to determine if other conventional statistical methods may be applied to similar case studies to fully assess the impact of potential confounding variables. Similar to BOI, Charlson Comorbidity Index (CCI)14,15 is an alternative method that is widely used to adjust for comorbidities, yet it reduces the generalizability by focusing on a small number of comorbid conditions and may not be optimal when applied to the elderly group that tends to have many other comorbid conditions.

Moreover, propensity scoring and instrumental variables16,17 (use of variables that are not related to studied variables) may also be useful to control for channeling bias. Further research should focus on comparing the study results when using these methods alone or in combination in order to draw more valid conclusions from database analyses. From a narrow perspective, results suggest future opportunities to research features of a particular drug, (e.g., amlodipine, as opposed to a similar drug, e.g., felodipine) that may influence its apparently preferential use in a sicker, more complex population.

Limitations

When calculating BOI scores for each patient, 1,306 amlodipine and 482 felodipine patients (N=1,788) did not have sufficient claims history to calculate an accurate score (P=.0045). A statistically significant greater number of amlodipine patients (10.5%) than felodipine patients (9.1%) had insufficient data to calculate a BOI score and thus were assigned to the lowest possible BOI category, equivalent to a single well-care visit. Second, only 6 months of data were available to calculate BOI. A longer time period may have provided a more accurate glimpse of the claims history, with the potential to impact the calculated BOI of the 2 study groups. Third, other confounding factors such as patient compliance or socioeconomic factors (i.e., education and income) were unavailable and therefore not examined in the current study. Examination of these factors may provide further insight into the drug utilization patterns between the 2 groups.

Conclusions

Although the 2 populations in the current study were initially perceived to be similar based on their treatment as described in the data, application of statistical methods revealed that the amlodipine group was more resource-intensive than the felodipine group. Comparison of the underlying characteristics in

### TABLE 3

<table>
<thead>
<tr>
<th>Condition</th>
<th>% of Amlodipine Patients</th>
<th>% of Felodipine Patients</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td>82.9%</td>
<td>85.6%</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Ischemic heart disease/Angina</td>
<td>42.4%</td>
<td>19.4%</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Chest pain</td>
<td>34.7%</td>
<td>17.0%</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Lipid disorder</td>
<td>28.3%</td>
<td>23.2%</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Diabetes</td>
<td>27.1%</td>
<td>22.5%</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Cough or other respiratory signs</td>
<td>26.3%</td>
<td>14.4%</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Arrhythmia</td>
<td>24.4%</td>
<td>12.8%</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Other joint disease</td>
<td>24.2%</td>
<td>21.7%</td>
<td>0.0003</td>
</tr>
<tr>
<td>Other heart disease</td>
<td>22.0%</td>
<td>9.4%</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>21.4%</td>
<td>9.7%</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>
2 presumably similar groups of elderly managed care patients found the amlodipine group (n=12,389) compared to the felodipine group (n=5,278) to have more patients in the 2 most clinical complex quintiles of BOI distribution, twice the prevalence of ischemic heart disease or angina, and nearly twice the incidence of high-severity hypertension. Managed care researchers must be diligent in detecting and controlling for underlying characteristics of study populations that may explain differences in outcome measures.

DISCLOSURES
Funding for this research was provided by Pfizer, Inc., and was obtained by author Amy Solis. Authors Raul Arocho and Simon Tang are employed by Pfizer, Inc. Arocho served as principal author of the study. Study concept and design were contributed primarily by Arocho, Solis, and authors George Goldberg and Sally Wade. Analysis and interpretation of data were contributed primarily by Goldberg, Solis, and Wade. Drafting of the manuscript was primarily the work of Arocho, Solis, Goldberg, and Wade. Critical revision of the manuscript was the work of Arocho, Goldberg, Solis, Tang, and Wade. Statistical expertise was contributed by Goldberg, Solis, Tang, and Wade. Administrative, technical, and/or material support was provided by Protocare Sciences.

ACKNOWLEDGMENT
Kevin Gilbert, MeDecision, for providing the Burden of Illness (BOI) and hypertension severity calculation logic.

POSTER PRESENTATIONS

REFERENCES