Working paper CeDER-11-07
November 2011
http://hdl.handle.net/2451/31303

Comparative Effectiveness for Oral Anti-diabetic Treatments among Newly Diagnosed Type-2 Diabetics: Machine Learning Applied to a Large-Scale Claims Dataset[a]

Jon Maguire, Stern School of Business, NYU
Vasant Dhar, Stern School of Business, NYU

ABSTRACT

In this paper, we demonstrate how the US healthcare system can provide increased benefits per unit of spend, and earlier identification of and intervention in chronic diseases through better predictive data-based analytics applied to the increasingly available troves of healthcare claims data. Specifically, we demonstrate the effectiveness of data mining by applying machine learning methods to large-scale medical and pharmacy claims data for roughly 70,000 patients over six years on newly diagnosed with type-2 diabetes, a common disease in the US costing billions to treat. This analysis reveals important differences in cost and quality among the disease's common treatments some of which have been published in the American Diabetes Association, and others that are regarded as tentative or have not been considered at all. The study demonstrates the potential for using large scale data mining for better understanding other major diseases including coronary problems and cancers and for focusing further inquiry in these areas.

Keywords: diabetes, comparative effectiveness, healthcare, informatics, data mining, machine learning, predictive analytics, claims data, health insurance, prediction

[a] This paper is dedicated to Dr. Harry Pople, a pioneer in the research on intelligent medical diagnostic systems. Pople developed the INTERNIST and CADUCEUS systems during the 1970s and 80s. Pople passed away on March 26, 2011 in Pittsburgh, PA.

BACKGROUND

Healthcare systems are in the spotlight these days. According to the Congressional Budget Office spending on healthcare as a percentage of GDP has roughly tripled between 1960 and 2005[1]. It could more than triple again by 2082 to consume nearly half of GDP unless something is done to rein in costs.

Furthermore, according to the OECD, while the US spends more on healthcare per capita than any other country, only six OECD countries – Czech Republic, Poland, Mexico, Slovakia, Turkey and Hungary – have lower life expectancies[2].

Tremendous pressure has been placed on the US government to control costs while still providing quality healthcare outcomes. In 2009 the Obama Administration passed the American Recovery and Reinvestment Act of 2009[3] (ARRA) which included $145.7 billion for healthcare. Most of this funding is being used to support existing government programs that provide healthcare directly, such as Medicare, Medicaid and the Veteran's Administration. However, $25.8 billion has been set aside for accelerating the adoption of health IT. And $1.1 billion of these funds have been directed at Comparative Effectiveness Research (CER)[4]. Most of the $25.8 billion for accelerating the adoption of health IT is being used to subsidize healthcare providers' costs to acquire and implement Electronic Health Record (EHR) systems.

This could create a deep well of data from which to measure the effectiveness of treatments and reduce the impacts of disease. The $1.1 billion for CER attempts to jumpstart research using the healthcare system's existing disparate but more limited troves of data. With its support of EHRs

and CER, the ARRA could set the stage for the kind of huge advances in data-based decision-making that other industries have been able to achieve using data mining and machine intelligence technologies.

DATA MINING AND PATTERN DISCOVERY: AN INTELLIGENT PROACTIVE APPROACH TO COMPARATIVE EFFECTIVENESS

In order to leverage the expanding healthcare data assets in the US to measure and explore the fundamental concepts of CER, the traditional approaches to scientific discovery used in the healthcare arena can benefit immensely when augmented with state-of-the-art data mining and machine intelligence technology. Methods from data mining have the potential to maximize the benefits of the existing and proposed repositories of healthcare information to improve quality and access while reducing costs. By "letting the data" speak, instead of relying on clinical trials that are typically funded with a specific question—and often a desired outcome—in mind, data mining allows us, in effect, to be passive observers of the healthcare system, thereby allowing us to be neutral in comparing the effectiveness of alternative treatments. Equally importantly, since "patterns often emerge before reasons for them become apparent," this approach can be very useful in providing the research community with early patterns of costs and outcomes thereby bringing attention to these areas. In our study, for example, our findings both confirm some of treatment beliefs proposed by the American Diabetes Association (ADA) but reveal unforeseen patterns that merit consideration.

The scientific community is constantly looking to understand relationships among symptoms, diseases, and treatment. Beginning in the 1960s, there were several attempts at building "expert

systems" that attempted to codify the diagnostic experience of experts into knowledge representations and algorithms to assist physicians with differential diagnosis[5,6,7,8]. Given an initial list of symptoms, such systems engage in a dialog with a diagnostician to hypothesize the diseases that explain a given set of symptoms in a process characterized as abductive logic by Peirce[9] in the 19th century and mechanized by researchers in early medical diagnosis systems in the 1970s (i.e. Pople[8]). The first medical diagnostic system that covered the entire field of Internal Medicine was INTERNIST/CADUCEUS, that guided expert diagnosticians through an interactive dialog involving a differential diagnosis. This system recognized that diagnosis is a complex art, requiring the physician to combine heuristic medical knowledge and real-world experience into an inductive reasoning process, where the objective is to hone in on the correct differential diagnosis with minimal questioning and invasive procedures. Such systems have shown to be effective in helping with differential diagnosis when there are competing opinions among experts about the causes of symptoms, and also helping experts avoid tunnel vision or an inappropriate premature diagnostic conclusion[9].

There have been significant changes in the medical landscape since the development of the early expert systems. One major development is the computerization of medical records, where virtually every "transaction," (i.e. interaction with a physician or a pharmacy) is recorded in a database. Such a database, in effect, constitutes the collective experience of the medical system on the US population for the last few decades. The data provide an unprecedented opportunity to evaluate and improve healthcare. The question we address here is the following: is there any way to use this real-world data to improve healthcare? This is a promising area of inquiry in light of

the massive successes of data mining in extracting and using patterns from data to improve efficiency in other businesses and industries over the last decade[10].

In order to appreciate the compelling value proposition for data mining in healthcare, it is worth considering how it can boost the effectiveness of traditional approaches to scientific discovery which are based on hypothesis testing: experts formulate hypotheses based on past theory, and test these against available data. For example, an expert might conjecture that the selection of a particular medication by a diabetic patient leads to a better outcome than some other medication, and attempt to collect data on individuals with relevant backgrounds to study. Such studies are often expensive and time-consuming. The so-called "file-drawer effect", where the likelihood that a study is published in the literature depends on its results (Scargle, 2000)[11], suggests that many such studies turn out to be inconclusive. Equally significantly, many of these investigations end with the possibility that they might have ignored an important variable, which if considered, may have yielded a positive result. But such an approach is inherently fraught with difficulty. It requires too much from a theoretician in terms of well formulated hypotheses and data to test them. Is there an alternative?

Interestingly, the answer is yes because the transaction data provide an unprecedented history of medical phenomena as they occurred. The collection of this data is large enough for us to draw statistically significant conclusions about treatments and outcomes. Businesses including casinos, banks, manufacturers, and retailers have been using data mining methods successfully for over a decade now to draw actionable conclusions from the data they collect as a by-product

of commerce. Data mining is an exploratory approach to data analysis, where hypotheses can *emerge* from the data instead of having to be specified beforehand.

Consider the following comparison with traditional hypothesis testing, where one must formulate a question, and then query a database to extract and analyze the appropriate data. In other words, this approach asks: what are the data that satisfy a query, and what do they tell us about the evidence relative to the query? Contrast this with a situation where the computer, armed with data representing our collective experience of diagnoses and treatments, taunts us by saying: if only you knew thy right question to ask, I would show you some very interesting patterns in the data! For example, the computer might nudge the analyst into formulating a query that says "retrieve all cases of people under 35 who have been newly diagnosed with type 2 diabetes, observe their treatment selection, and examine the differences in outcomes." An analysis of the records returned by such a query or similar ones might reveal which treatments yield superior outcomes under disparate circumstances. It is important to appreciate that small variations in the query, such as "over 55" or "where treatment selection was not Metformin," might show very different cost and outcome patterns in the data. But even more importantly, it is unreasonable to expect even the most informed expert to come up with just the right query whereas data mining methods are designed to do just that: help us converge on the right questions.

THE STUDY

Traditional technologies are often measured by the cost per life saved or the cost per patient processed. For example, a recent study[12] estimated that the availability of electronic health

records (EHR) in hospitals reduced infant mortality from 500 to 484 per 100,000 births, where the cost of implementing the EHR system was estimated at $531,000 per birth. This presents an interesting dilemma to policy makers: is the expense justified? What is the quality of those the additional 16 lives saved? And so on.

In contrast, the use of data mining methods to the armamentarium of existing methods involves little cost, and the potential upside is tremendous. For example, even a 1% reduction is heart attacks, obesity, diabetes, etc., would be associated with huge reduction in healthcare costs with virtually no additional infrastructure investment, but rather, with a very modest investment in data and analytics resources.

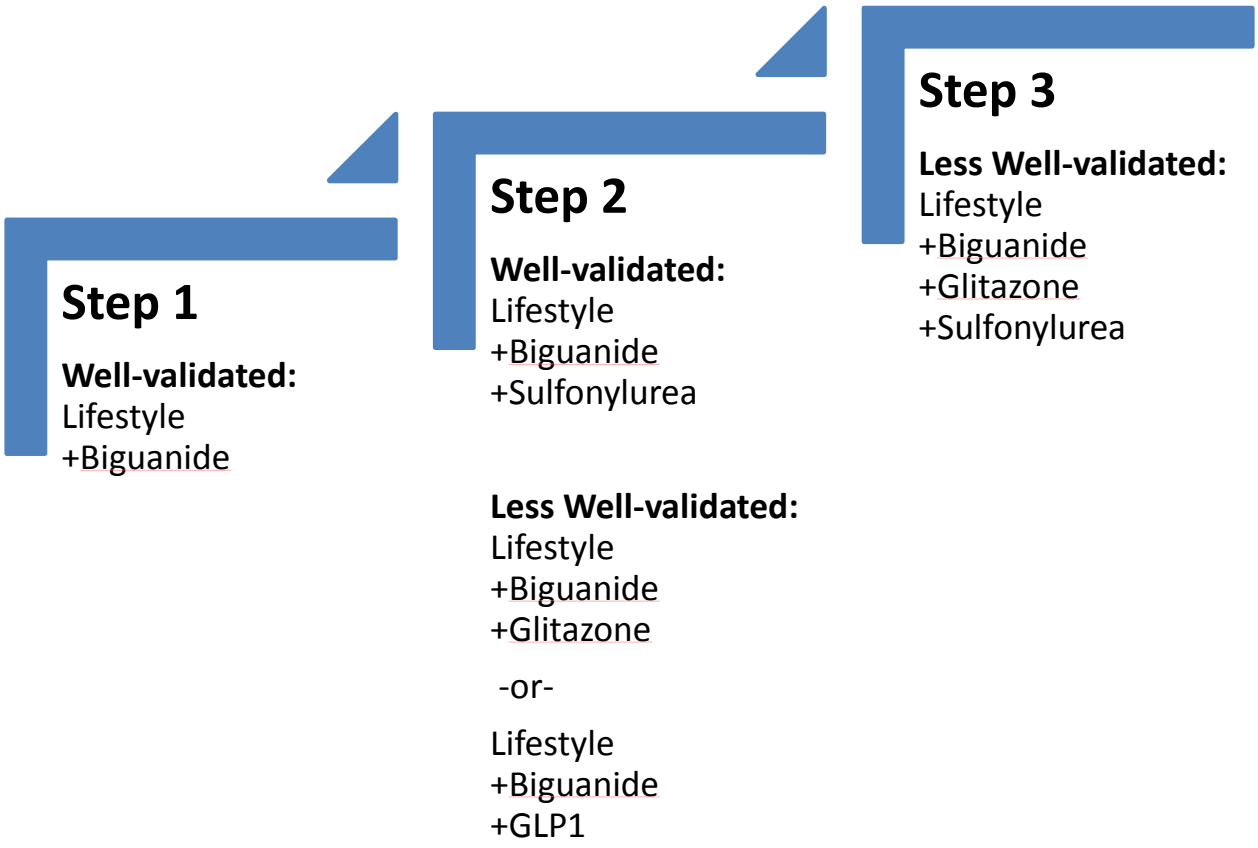
To provide a demonstration of how data mining techniques might be applied to CER, we obtained data from a large national health insurance company and used the data mining toolset available from SAS Institute's Enterprise Miner software[13]. The intention in our example was to see if machine learning could be used to shed light on available oral anti-diabetic medications used to treat type 2 diabetes.

This population is of particular interest due to the high costs society faces with the growing prevalence of diabetes. It is one of the leading causes of morbidity and mortality in the United States because of its role in the development of co-morbid conditions such as ophthalmic disease, kidney disease, and cardiovascular disease. Adults with diabetes have heart disease and stroke death rates 2-4 times higher than adults without diabetes[14]. According to the Centers for Disease Control (CDC)[15] 23.6 million people or 7.8% of the population had diabetes in 2007,

and if current trends continue then up to one third of Americans alive today will develop diabetes before they die. In the same report, the CDC also finds that $116 billion in direct medical costs were expended to treat diabetes in 2007, and the economy lost another $58 billion indirectly due to the negative effects of the disease on economic productivity. But the cost is not economic alone: those that suffer from the illness lose an average of 10-15 years of life and diabetes is the leading cause of blindness, kidney failure, and amputations.

Further complicating analysis of diabetes is the wide array of treatments available and the ability to combine multiple therapies to treat the condition. The American Diabetes Association (ADA) recommends specific therapy combinations be applies in a stepwise approach until the patient's blood A1C (glycated hemoglobin) is measured at less than 7%[16]. While the ADA's consensus algorithm includes options for adding insulin to patients' therapy regimens in later steps, we have excluded these patients (3,612 patients) to focus on oral anti-diabetic therapies. Adding the complexity of insulin use is left for further research. The ADA's consensus algorithm, without insulin-related steps, is summarized graphically in Figure 1. Each of these therapy regimens was observed in our study population, as were other regimens that are not part of the algorithm.

Figure 1: Graphical representation of the ADA's consensus algorthim (oral therapies only).



DATA & METHODS

Anonymized longitudinally linked medical claims, pharmacy claims, and eligibility records for commercially insured patients from January 2004 through March 2010 were used to construct patient-level records for this analysis. Selected commercially insured patients must have been 18 or older and have had a diagnosis for uncomplicated type 2 diabetes[17] (the "Trigger") between 2004 and 2009 such that there were at least 270 days of medical and pharmacy benefits eligibility prior to the trigger (the "Clean" period) and at least 450 days of medical and pharmacy

benefits eligibility following the trigger (the "Treatment Selection" and "Outcome Observation" periods) – see Figure 2.

Figure 2: Graphical representation of study period phases.

Clean Period	TRIGGER	Treatment Selection Period	Outcome Observation Period
270 Days		180 Days	270 Days

Source: the authors.

Diabetes is often treated with multiple medications simultaneously. The 180 day Treatment Selection period (Figure 2) was constructed to allow adequate time to observe which oral anti-diabetic medications newly diagnosed type 2 diabetic patients selected. Furthermore, patients must have had no diagnosis for type 1 diabetes[18] and no insulin[19] treatments at any time during the entire 720 day study period. They must have had no diabetes-related complications[20] during the Clean and Treatment Selection periods. And they must have had no insulin and no oral anti-diabetic medications[21] during the Clean period.

As a result of applying these inclusion and exclusion criteria to the data, 66,523 patients were included in the study. The resulting age-gender distribution and geographic distribution are given in Figures 3 and 4 respectively.

Figure 3: Age-gender distribution of selected patients.

Age Group	Female (N, row%, col%)	Male (N, row%, col%)	Total (N, row%, col%)
18-34	1,654 ($50.0%, 2.49%)	1,653 (50.0%, 2.48%)	3,307 (100.0%, 4.97%)
35-44	4,750 (41.0%, 7.1%)	6,828 (59.0%, 10.3%)	11,578 (100.0%, 17.4%)
45-54	9,946 (43.8%, 15.0%)	12,774 (56.2%, 19.2%)	22,720 (100.0%, 34.2%)
55-64	10,189 (45.1%, 15.3%)	12,396 (54.9%, 18.6%)	22,585 (100.0%, 33.95%)
65+	2,846 (44.9%, 4.3%)	3,487 (55.1%, 5.2%)	6,333 (100.0%, 9.5%)
Total	29,385 (44.2%, 100.0%)	37,138 (55.8%, 100.0%)	66,523 (100.0%, 100.0%)

Source: the authors.

Figure 4: Geographic distribution of selected patients.

Census Region[22]	Patients (N, %)
Midwest	16,255, 24.4%
Northeast	5,971, 9.0%
South	35,212, 52.9%
West	9,085, 13.7%

Source: the authors.

Both outcome variables and explanatory variables were measured in the relevant study sub-periods (Figure 5) for use in downstream data mining analyses. Outcome variables, measured

during the Outcome Observation period, included total charges (including drug charges), medical charges (excluding drug charges), and a binary variable indicating whether or not the patient developed a diabetes-related complication. Explanatory variables included patient age group, patient gender, flags for each type of oral anti-diabetic medication observed during the Treatment Selection period, and two measures of the patient's pre-existing health care status: the number of unique 3-digit ICD9 diagnosis groupings during the Clean period, and the number of unique drug subclasses during the Clean period.

Figure 5: Selected statistics (95% confidence intervals) for selected variables during each of the three study sub-periods.

	Clean Period (270 days)	**Treatment Selection Period (180 days)**	**Outcome Observation Period (270 days)**
Annualized Median Medical Charges	$1,515 ($1,492, $1,535)	$4,302 ($4,240, $4,360)	$2,740 ($2,699, $2,780)
Annualized Median Pharmacy (Drug) Charges	$592 ($580, $605)	$2,036 ($2,018, $2,056)	$1,947 ($1,927, $1,968)
Annualized Median Total Charges	$2,689 ($2,653, $2,721)	$7,296 ($7,218, $7,372)	$5,613 ($5,551, $5,676)
Annualized Mean Number of 3-digit ICD9 Diagnosis Codes	1.31 (1.28, 1.32)	3.24 (3.20, 3.26)	2.36 (2.33, 2.39)
Annualized Mean Number of Drug Subclasses	5.44 (5.40, 5.48)	12.66 (12.60, 12.72)	9.00 (8.96, 9.05)
% With Complications	n/a	n/a	1.99% (1,327) (1.89%, 2.10%)

Source: the authors; confidence intervals around medians using methods described by Hahn & Meeker (1991).

Dozens of therapy regimens were observed including monotherapy on a single compound, multiple therapies, and no therapies at all (which we interpreted to be a regimen of lifestyle changes only). A flag was constructed for each regimen seen in at least 0.5% of treated patients (i.e. excluding those with only lifestyle changes) using the drug utilization of the entire Treatment Selection period for each patient. In some cases regimens were grouped together by common components to form large enough categories. All other therapies were included in an "other" category. The results of this classification of patients are presented in Figure 6.

Figure 6: Classification of patients by mutually exclusive therapy regimens observed during the Treatment Selection period.

Therapy regimen	**ADA Step**	**Patients**	**Percent**
Lifestyle changes only		26,383	39.7%
Biguanide monotherapy	Step 1	20,497	30.8%
Biguanide+sulfonylurea	Step 2	5,183	7.8%
Biguanide+glitazone	Step 2	4,074	6.1%
Sulfonylurea monotherapy		3,274	4.9%
Glitazone monotherapy		2,355	3.5%
Biguanide+sulfonylurea+glitazone	Step 3	1,237	1.9%
Biguanide+DPP4		836	1.3%
Sulfonylurea+glitazone		561	0.8%
DPP4 monotherapy		401	0.6%
Other therapy NEC		371	0.6%

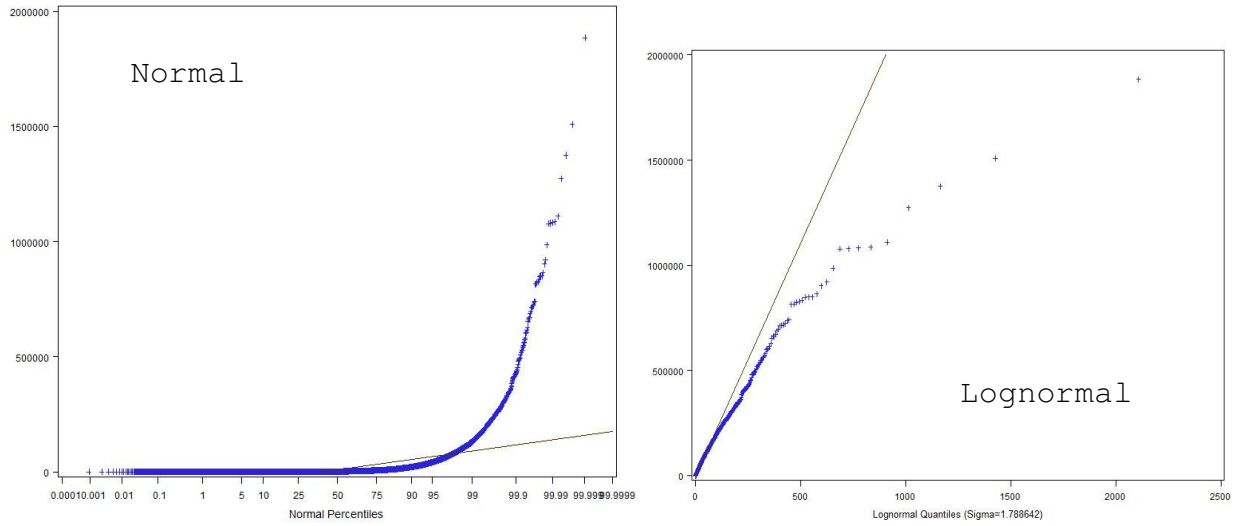
GLP1 monotherapy		304	0.5%
Biguanide+glitazone with 1 or more other therapy NEC		298	0.4%
Biguanide+sulfonylurea with 1 or more other therapy NEC		297	0.4%
Biguanide+GLP1	Step 2	262	0.4%
Biguanide with 1 or more other therapy NEC		190	0.3%

Source: the authors. Notes: NEC=not elsewhere classified; See Figure 1 for ADA reference.

RESULTS

We start by looking at medical charges as an outcome of treatment selection in that it serves as a proxy for the utilization of medical services. Drug charges are excluded because they are highly correlated with whether selected therapies still enjoy patent protection and therefore command higher prices than generically available alternatives. Medical charges are log-normally distributed with a long tail consisting of outliers, they are transformed by taking their natural logarithm, which results in a more normal distribution of costs than that of the untransformed values. This can be seen in the probability plots in Figure 7 where the observed distribution (crosses) is plotted against the theoretical distribution (solid line) constructed from distribution parameters estimated in the data.

Figure 7: Probability plots of medical charges against lognormal and normal distributions.



Source: the authors.

In addition to costs, we look also at the likelihood of developing a diabetes-related complication as an outcome of treatment selection, which - as discussed previously - can have a deleterious impact on the patient's quality of life.

Medical Charges Decision Tree

Using only the mutually exclusive flags indicating treatment regimens, a decision tree model was generated for medical charges. The tree algorithm partitions the data recursively to identify subsets of the data where the distribution of the dependent variable is as different as possible from the distribution of the variable in the data as a whole[10].

After iterating to select the best tree in terms of the average squared errors of the predicted values, the tree algorithm selected eight of 16 available therapy regimen flags for inclusion as

nodes in the tree and grouped the remaining therapies into a single final node. Figure 8 summarizes the retransformed results in tabular form.

Figure 8: Decision tree model results for Medical Charges

Therapy regimen	ADA Step	n	Expected Value	95% Lower Confidence Limit	95% Upper Confidence Limit
Lifestyle changes only		26,383	2,080	2,023	2,138
Biguanide+sulfonylurea	Step 2	5,183	1,353	1,264	1,448
Biguanide+glitazone	Step 2	4,074	1,274	1,181	1,375
GLP1 monotherapy		304	3,429	2,730	4,307
DPP4 monotherapy		401	3,103	2,577	3,735
Biguanide+GLP1	Step 2	261	2,951	2,338	3,725
Other therapies NEC		371	2,566	1,988	3,312
Biguanide with 1 or more other therapy NEC		190	3,103	2,371	4,060
None of the above		29,355	1,604	1,561	1,648

Source: the authors.

This model shows that there is no statistically significant difference (at the 95% level) in subsequent medical charges between patients in the biguanide+sulfonylurea or biguanide+glitazone groups (two of the three ADA Step 2 regimens). However, for patients in the biguanide+GLP1 group, costs are significantly higher (at the 95% level). This could be an important consideration because both glitazone and GLP1 products still retain market exclusivity, while sulfonylurea is generically available making it likely to be a less expensive option. It should be noted that the ADA algorithm[23] lists weight gain as disadvantages of both

sulfonylurea and glitazone therapy while it lists weight loss as an advantage of GLP1 therapy, and this may drive certain types of patients to the GLP1 therapy.

Perhaps more importantly, the expected medical charges of patients who select lifestyle changes only is statistically significantly higher (at the 95% level) than patients in the biguanide+sulfonylurea or biguanide+glitazone regimens. Patients who select lifestyle changes only have higher average medical charges (at the 95% level) than patients who are in none of the regimens listed in Figure 8 (i.e. "none of the above") which represents 20,497 (69.8%) patients in the biguanide monotherapy regimen – the only ADA Step 1 regimen. The result also suggest that patients can avoid significant subsequent medical charges by adding any one of three regimens recommended by the ADA (biguanide monotherapy, biguanide+sulfonylurea or biguanide+glitazone) to their lifestyle changes.

Medical Charges Regression

In addition to the decision tree model, a regression model was also generated for log medical charges. Using the regression model allows for the inclusion of gender, age, census region, year of index, and counts of pre-existing diagnosis and prescription drug groupings as fixed effects so that predicted values are adjusted for these effects. Figure 9 shows the coefficient estimates produced by the model and Figure 10 shows the predicted medical charges for each possible regimen selection.

Figure 9: Regression model results (log-linear, not retransformed into original scale)

Effect	N	Coeff. Est.	SE	t Stat.	P value	95% lower conf. limit	95% upper conf. limit
Intercept	66,523	6.90	0.03	251.91	<.0001	6.84	6.95
Female	29,385	0.17	0.01	19.58	<.0001	0.16	0.19
Age 18-34	3,307	-0.28	0.03	-8.74	<.0001	-0.34	-0.22
Age 35-44	11,578	-0.24	0.02	-12.35	<.0001	-0.28	-0.20
Age 45-54	22,720	-0.02	0.02	-0.95	0.34	-0.05	0.02
Age 55-64	22,585	0.22	0.02	13.44	<.0001	0.19	0.25
Midwest	16,255	-0.06	0.02	-3.78	0.00	-0.09	-0.03
Northeast	5,971	0.24	0.02	10.56	<.0001	0.20	0.29
South	35,212	-0.03	0.01	-2.53	0.01	-0.06	-0.01
Index Year 2005	19,965	-0.07	0.01	-4.70	<.0001	-0.10	-0.04
Index Year 2006	16,841	-0.02	0.02	-1.08	0.28	-0.05	0.01
Index Year 2007	15,405	0.00	0.02	-0.05	0.96	-0.03	0.03
# Pre-existing Diagnosis groups	66,523	0.15	0.01	25.24	<.0001	0.14	0.17
# Pre-existing Drug groups	66,523	0.14	0.00	53.30	<.0001	0.13	0.14
Biguanide monotherapy (Step 1)	20,497	-0.19	0.03	-6.92	<.0001	-0.25	-0.14

Biguanide with 1 or more other therapy NEC	190	0.54	0.15	3.57	0.00	0.25	0.84
Biguanide+DPP4	836	-0.07	0.08	-0.86	0.39	-0.21	0.08
Biguanide+GLP1 (Step 2)	262	0.19	0.13	1.42	0.15	-0.07	0.44
Biguanide+ sulfonylurea (Step 2)	5,183	-0.13	0.04	-3.58	0.00	-0.21	-0.06
Biguanide+ sulfonylurea with 1 or more other therapy NEC	297	0.09	0.12	0.75	0.46	-0.15	0.33
Biguanide+ sulfonylurea+ glitazone (Step 3)	1,237	-0.03	0.06	-0.44	0.66	-0.15	0.10
Biguanide+ glitazone (Step 2)	4,074	-0.24	0.04	-6.04	<.0001	-0.32	-0.17
Biguanide+ glitazone with 1 or more other therapy NEC	298	-0.10	0.12	-0.78	0.43	-0.34	0.14
DPP4 monotherapy	401	0.07	0.11	0.62	0.54	-0.14	0.27
Lifestyle changes only	26,383	-0.05	0.03	-1.71	0.09	-0.10	0.01
Other therapy NEC	371	0.10	0.11	0.92	0.36	-0.12	0.32
GLP1 monotherapy	304	0.27	0.12	2.23	0.03	0.03	0.51
Sulfonylurea monotherapy	3,274	-0.26	0.04	-5.89	<.0001	-0.34	-0.17
Sulfonylurea+ glitazone	561	-0.06	0.09	-0.70	0.48	-0.24	0.11

Source: the authors.

Figure 10: Retransformed predicted medical charges by therapy regimen

Regimen	N	Predicted Value
Biguanide monotherapy (Step 1)	20,497	$1,721
Biguanide w/ 1 or more other therapy NEC	190	$3,330
Biguanide+DPP4	836	$1,960
Biguanide+GLP1 (Step 2)	262	$3,083
Biguanide+sulfonylurea (Step 2)	5,183	$1,480
Biguanide+sulfonylurea with 1 or more other therapy NEC	297	$2,163
Biguanide+sulfonylurea+glitazone (Step 3)	1,237	$1,531
Biguanide+glitazone (Step 2)	4,074	$1,408
Biguanide+glitazone with 1 or more other therapy NEC	298	$1,782
DPP4 monotherapy	401	$3,139
Lifestyle changes only	26,383	$2,199
Other therapy NEC	371	$2,688
GLP1 monotherapy	304	$3,615
Sulfonylurea monotherapy	3,274	$1,605
Sulfonylurea+glitazone	561	$1,757
Glitazone monotherapy	2,355	$2,158

Source: the authors.

These results confirm those of the decision tree model where the addition of some therapy regimen to patients' lifestyle changes predicts lower subsequent medical charges than lifestyle changes alone, but we can go a step further. Patients are better off selecting any of the therapy regimens from the ADA algorithm (except biguanide+GLP1 as noted previously) than they are with lifestyle changes only. Furthermore, it appears that patients who select either sulfonylurea, glitazone or both in addition to biguanide are better off than patients who take only biguanide. Finally it appears that, though it is not part of the ADA algorithm, selection of sulfonylurea monotherapy has benefits similar to regimens that are part of the ADA algorithm in terms of predicted medical charges.

Diabetes Related Complications Decision Tree

While costs are certainly an important component of the current healthcare policy debate, the quality of patient's own outcome is of critical interest. To compare effectiveness among competing medications along this dimension, we classified patients according to whether or not they developed diabetes-related complications during the outcome observation period.

After iterating to select the best tree in terms of the average squared errors of the predicted values, the tree algorithm selected just two of the 16 available therapy regimen flags for inclusion as nodes in the tree and grouped the remaining therapies into a single final node. Figure 11 summarizes the results in tabular form.

Figure 11: Decision tree model results for Complications

Therapy regimen	ADA Step	n	Expected Value	95% Lower Confidence Limit	95% Upper Confidence Limit
Biguanide+GLP1 (Step 2)	Step 2	262	6.9%	6.7%	7.1%
Biguanide+ sulfonylurea (Step 2)	Step 2	5,183	2.9%	2.9%	2.9%
None of the above		61,078	1.9%	1.9%	1.9%

Source: the authors.

These results suggest that patients who develop diabetes-related complications are concentrated among those who also select either the biguanide+GLP1 regimen or the biguanide+sulfonylurea regimen. Earlier discussion of GLP1 therapy characteristics may explain part of this observation. However, it is not clear why we would expect a concentration of diabetes-related

complications among patients selecting the biguanide+ sulfonylurea regimen. Further research into this relationship is needed.

Diabetes Related Complications Regression

As with medical charges previously a regression model, in this case logistic regression, was added to the data mining analysis to model the binary (yes/no) variable indicating development of a diabetes related complication following treatment selection. Coefficient estimates from the model presented in Figure 12 were used to create predictions[24] in Figure 13 that are easier to understand since they are presented in terms of familiar percentages rather than odds ratios.

Figure 12: Decision tree model results for Complications

Effect	Coef. Est.	SE	Chi-Sq Stat.	P Value	Odds Ratio	SE O/R
Intercept	-3.944	0.086	2121.970	<.0001	0.019	0.002
Female	0.009	0.029	0.110	0.744	1.009	0.029
Age 18-34	-0.371	0.129	8.270	0.004	0.690	0.089
Age 35-44	-0.329	0.075	19.160	<.0001	0.720	0.054
Age 45-54	0.001	0.055	0.000	0.982	1.001	0.055
Age 55-64	0.054	0.055	0.980	0.323	1.055	0.058
Midwest	-0.052	0.054	0.930	0.334	0.949	0.051
Northeast	-0.098	0.078	1.570	0.210	0.907	0.071
South	0.011	0.044	0.070	0.798	1.011	0.045

Index Year 2005	-0.016	0.047	0.110	0.741	0.985	0.046
Index Year 2006	-0.131	0.050	6.820	0.009	0.877	0.044
Index Year 2007	0.036	0.049	0.530	0.469	1.036	0.051
# Pre-existing Diagnosis groups	0.075	0.015	23.710	<.0001	1.077	0.016
# Pre-existing Drug groups	0.034	0.007	22.280	<.0001	1.035	0.008
Biguanide monotherapy (Step 1)	-0.363	0.087	17.480	<.0001	0.695	0.060
Biguanide with 1 or more other therapy NEC	-0.447	0.551	0.660	0.418	0.640	0.353
Biguanide+DPP4	0.081	0.209	0.150	0.699	1.084	0.226
Biguanide+GLP1 (Step 2)	1.065	0.241	19.480	<.0001	2.901	0.700
Biguanide+ sulfonylurea (Step 2)	0.242	0.106	5.200	0.023	1.274	0.135
Biguanide+ sulfonylurea with 1 or more other therapy NEC	-0.608	0.477	1.620	0.203	0.545	0.260
Biguanide+ sulfonylurea+ glitazone (Step 3)	0.095	0.191	0.250	0.619	1.099	0.210
Biguanide+ glitazone (Step 2)	-0.152	0.128	1.400	0.237	0.859	0.110
Biguanide+ glitazone with 1 or more other therapy NEC	0.695	0.267	6.760	0.009	2.003	0.535
DPP4 monotherapy	-0.089	0.286	0.100	0.757	0.915	0.262
Lifestyle changes only	-0.394	0.084	22.090	<.0001	0.674	0.057
Other therapy NEC	0.448	0.252	3.170	0.075	1.565	0.394
GLP1 monotherapy	-0.473	0.429	1.220	0.270	0.623	0.267
Sulfonylurea monotherapy	-0.059	0.127	0.210	0.644	0.943	0.120
Sulfonylurea+ glitazone	0.312	0.237	1.740	0.188	1.366	0.324

Source: the authors.

Figure 12: Retransformed predicted complication rates by therapy regimen

Therapy Regimen	N	Expected proportion	95% conf. lower bound	95% conf. upper bound
Biguanide monotherapy (Step 1)	20,497	1.479%	1.477%	1.481%
Biguanide with 1 or more other therapy NEC	190	1.315%	1.293%	1.336%
Biguanide+DPP4	836	2.358%	2.333%	2.382%
Biguanide+GLP1 (Step 2)	262	6.089%	5.906%	6.272%
Biguanide+ sulfonylurea (Step 2)	5,183	2.503%	2.492%	2.514%
Biguanide+ sulfonylurea with 1 or more other therapy NEC	297	1.132%	1.118%	1.146%
Biguanide+ sulfonylurea+ glitazone (Step 3)	1,237	2.088%	2.071%	2.105%
Biguanide+ glitazone (Step 2)	4,074	1.724%	1.717%	1.731%
Biguanide+ glitazone with 1 or more other therapy NEC	298	4.102%	4.007%	4.197%
DPP4 monotherapy	401	2.253%	2.220%	2.286%
Lifestyle changes only	26,383	1.487%	1.485%	1.489%
Other therapy NEC	371	3.428%	3.363%	3.493%
GLP1 monotherapy	304	1.420%	1.401%	1.439%
Sulfonylurea monotherapy	3,274	2.015%	2.005%	2.025%
Sulfonylurea+ glitazone	561	2.726%	2.689%	2.764%
Glitazone monotherapy	2,355	2.150%	2.137%	2.162%

Source: the authors.

This model adds to the evidence that patients who develop diabetes-related complications are more concentrated among those who select sulfonylurea as part of their regimen. In all regimens

that include sulfonylurea except one (Biguanide+ sulfonylurea with 1 or more other therapy NEC), the subsequent rate of diabetes-related complications is over 2%. Within Step 2 and 3 therapies (and setting aside biguanide+GLP1), adding sulfonylurea alone to biguanide is associated with higher a diabetes-related complications rate than is adding either glitazone alone or glitazone+sulfonylurea to biguanide. As with the decision tree model before, these results suggest the need for additional research into the effects of sulfonylurea on subsequent development of diabetes complications.

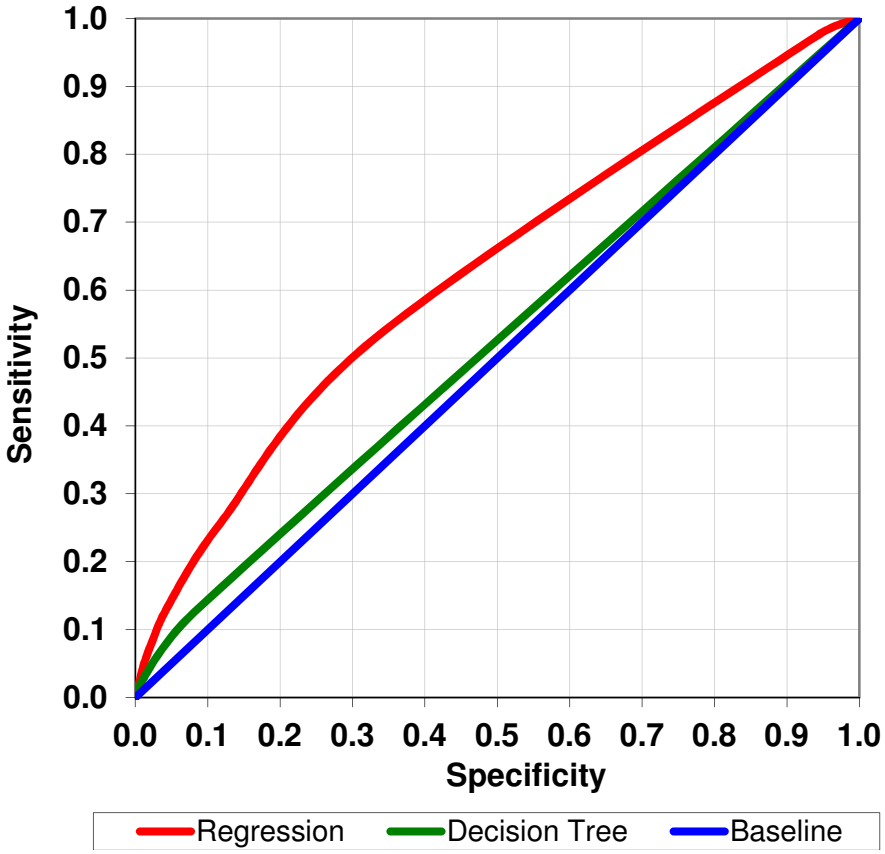
Model Evaluation

It is natural to want to compare models' predictive ability so that the best model is ultimately used to make predictions. An analysis of root mean squared errors (conceptually similar to a standard deviation as a measure of variability of model predictions) shows that the regression model has 11% lower variability than the decision tree model when predicting medical charges.

In the case of the models of diabetes-related complications, because the outcome is binary, it is also possible to compare the diabetes-related complications models' receiver operating characteristic (ROC) curves and the areas under the curves (AUC)[25]. The ROC curves graphically plot the true positive rate versus the false negative rate for a variety of model parameter thresholds that determine how often the models classify instances positively or negatively. The AUC measures the areas under these curves and represents the probability that the model will rank any given positive instance higher than any negative one.

In our case the ROC curves are plotted in Figure 14 below. The plot also includes a baseline that represents random chance. This provides the ability to visually inspect how much "better" each model is than essentially flipping a coin. The ROC curve for the regression model is higher at every point than the ROC curve for the decision tree model. This can be measured quantitatively using the AUCs for each model. For the regression model the AUC is 63.00% and for the decision tree model the AUC is 52.24%.

Figure 14: ROC curves comparing regression and decision tree models of diabetes-related complications rates



Different insights can be gained from both models. For example, although the decision tree model performs less well using these quantitative measures of model performance it can be easier to understand and explain to a wider audience. The regression model on the other hand, provides numeric estimates of the probabilities of developing a complication for each pharmaceutical therapy in the data. Hence, in addition to these quantitative evaluations, it is also important to consider the explanatory benefits that models appeal to common sense and can be put to practical use.

LIMITATIONS

The automated models derived from this data mining exercise offer several interesting and valuable insights into the comparative effectiveness of available treatments for type 2 diabetes. However, variables for medication adherence, persistence and compliance, which can also have an effect on outcomes, were not considered in these models. Also, while the health of patients prior to their diagnosis for diabetes was approximated by including the number of unique 3-digit ICD9 diagnosis codes and the number of unique drug classes during the Clean Period, formal case mix adjustments were not applied. These shortcomings, while important to bear in mind, offer opportunities for future refinements of this data mining system and are not expected to change the overall direction of results.

RELATIONSHIP TO PRIOR LITERATURE

Our results are supported in a few important ways by the literature, but they are unique in that they compare multiple therapies and multiple outcomes simultaneously while other studies found in the literature generally confine themselves to a particular therapy or to the comparison of a specific set of therapies. For example, the literature reveals several studies that show that sulfonylureas may be associated with higher risk of mortality and with higher risk of developing comorbidities such as cardiovascular disease and cancer risks[26,27,28,29,30,31,32,33]. It stands to reason that the poorer clinical outcomes of sulfonylurea users would be associated with increased diabetes-related complications rates compared to diabetics who don't use sulfonylureas, which is supported by both this and other studies. Another study looked at cases where patients who were initiated on a sulfonylurea that later required the addition of a second therapy and found that such patients, when augmented with metformin (biguanide), had 33% lower costs than those augmented with glitazones[34]. Still another study found that patients using metformin could be less adherent to their therapy regimens than users of either sulfonylureas or pioglitazone (a glitazone) and still achieve similar cost reductions compared with patients who were fully non-adherent to other therapy regimens[35]. None of the literature we reviewed provide as comprehensive a review of real-world therapy regimens in the context of the ADA's consensus algorithm.

Despite of the novelty of our findings, the study raises more focused questions for further investigation, such as *why* costs or complications associated with certain treatments are higher than others. Indeed, we consider the strengths of a study such as ours as being one of raising

questions and focusing further investigation that attempts to answer the "why" questions. Some of the specific questions are the following:

1. How would the results change by conditioning on the "A1C" blood sugar levels? The ADA recommends that A1C levels should be controlled to be under 7. It is possible that many of the treatments prescribed by physicians take into account the specific A1C levels and that higher levels are correlated with complications. This is an obvious area for further inquiry.
2. Another deeper investigation would be to consider weight levels of individuals and analyze whether costs and complications are related to obesity. Specifically, for individuals who choose "lifestyle only" changes to control diabetes, are there more complications when individuals are overweight?
3. The area of complications needs further inquiry in general. Specifically, for the tree shown in Table 8, what is the distribution of complications for the different clusters? What are the reasons for the differences in complications?
4. Finally, it is worth extending this study to insulin users, an area that is significant both in terms of healthcare costs and complications, and also extend our methodology to other costly and prevalent health conditions[36] such as heart disease, cancer, and mental disorders.
5. How accurate are the models extracted through the data mining process be used by physicians for prediction? In other words, given the current "state" of a patient, what is the accuracy of the model in predicting costs and complications of alternative treatments?

CONCLUSIONS

Diabetes is a complicated and costly disease and pharmacologic therapy has changed dramatically in the last decade. New chemical entities are altering type 2 diabetes mellitus treatment patterns and allowing glycemic control beyond the reach of previous medical therapy. The results offered in our study, produced using inexpensive automated techniques, are corroborated by other published literature and offer new, more holistic, insights into the effectiveness of competing diabetes therapy regimens and are further validated by the ADA's consensus algorithm.

An important general goal of this paper, other than addressing type 2 diabetes, is to highlight how data mining and machine intelligence technologies can be applied to the healthcare industry at large, and particularly to the questions posed by CER. We have relied significantly on machine automation to discover new predictive models yielding new hypotheses that seem qualified for further investigation, including the following.

Among adult newly diagnosed type 2 diabetics with commercial insurance who do not require insulin therapy:

- There is no difference in subsequent medical charges between patients selecting either biguanide+sulfonylurea or biguanide+glitazone (both ADA Step 2 regimens). This could be an important observation because glitazone products still retain market exclusivity, while sulfonylurea is generically available making it likely to be a less expensive option.

- Patients are better off selecting any of the therapy regimens from the ADA algorithm (except biguanide+GLP1 as noted previously) than they are with lifestyle changes only.
- Patients who select either sulfonylurea, glitazone or both in addition to biguanide are better off in terms of subsequent medical charges than patients who select only biguanide.
- Though it is not part of the ADA algorithm, selection of sulfonylurea monotherapy has benefits similar to regimens that are part of the ADA algorithm in terms of subsequent medical charges.
- Patients who develop diabetes-related complications are more concentrated among those who select sulfonylurea as part of their regimen.

We imagine implementing similar automated – even self-improving – techniques on a much larger scale to cull out possible hypotheses that warrant attention using traditional methods of scientific discovery. Such an approach would focus the valuable and expensive human and other resources being devoted to CER into areas that are more likely to yield actionable results than methods that rely solely on traditional hypothesis discovery techniques.

Often research is conducted at the behest of a particular set of interests, commercial or otherwise. By making the results of the data mining and machine intelligence activities public, this approach would have the further advantage of acting as a counterweight against whatever inherent biases exist in the existing hypothesis discovery process – machines are not subject to external interests or biases. This is particularly compelling when you consider the immense repository of data the US government is laying the groundwork for under ARRA. Not only will vastly more detailed information on individuals' healthcare utilization, diagnostic history,

demographic profile, and professional advice be captured and stored electronically in EHRs, but we will have the technology to comb through it all automatically without bias in search of the most effective treatments for a given illness or condition.

Appendix 1: Medical Claim Coding for Uncomplicated Type 2 Diabetes

ICD9-CM Diagnosis Codes

250.00	Diabetes mellitus without mention of complication, Type II, Controlled
250.02	Diabetes mellitus without mention of complication, Type II, Uncontrolled

Appendix 2: Medical Claim Coding for Type 1 Diabetes

ICD9-CM Diagnosis Codes

250.x1	Diabetes mellitus with or without mention of complication, Type I, Controlled
250.x3	Diabetes mellitus with or without mention of complication, Type I, Uncontrolled

Appendix 3: Medical Claim Coding for Diabetes Complications

ICD9-CM Diagnosis Codes

250.1x	Diabetes with ketoacidosis
250.2x	Diabetes with hyperosmolarity
250.3x	Diabetes with other coma
250.4x	Diabetes with renal manifestations
585.xx	Chronic kidney disease
58381	Nephritis and nephropathy, not specified as acute or chronic, in diseases classified elsewhere
58181	Nephrotic syndrome in diseases classified elsewhere
250.5x	Diabetes with ophthalmic manifestations
369.xx	Blindness and low vision
362.xx	Other retinal disorders
36641	Diabetic cataract
36544	Glaucoma associated with systemic syndromes
250.6x	Diabetes with neurological manifestations
337.1x	Peripheral autonomic neuropathy in disorders classified elsewhere
353.5x	Thoracic root lesions, not elsewhere classified
354.xx	Mononeuritis of upper limb and mononeuritis multiplex
355.xx	Mononeuritis of lower limb
357.2x	Polyneuropathy in diabetes
536.3x	Gastroparesis
713.5x	Arthropathy associated with neurological disorders
250.7x	Diabetes with peripheral circulatory disorders
44381	Peripheral angiopathy in diseases classified elsewhere
785.4x	Gangrene
250.8x	Diabetes with other specified manifestations
707.1x	Ulcer of lower limbs, except pressure ulcer
707.2x	Pressure ulcer stages
707.8x	Chronic ulcer of other specified sites
707.9x	Chronic ulcer of unspecified site
731.8x	Other bone involvement in diseases classified elsewhere
250.9x	Diabetes with unspecified complication

ICD9-CM Procedure Codes

84.0x	Amputation of upper limb
84.1x	Amputation of lower limb

Appendix 3 Continued: Medical Claim Coding for Diabetes Complications

CPT4 Procedure Codes

26910	Amputate Metacarpal Bone
26951	Amputation Of Finger/Thumb
26952	Amputation Of Finger/Thumb
27590	Amputate Leg At Thigh
27591	Amputate Leg At Thigh
27592	Amputate Leg At Thigh
27594	Amputation Follow-Up Surgery
27596	Amputation Follow-Up Surgery
27598	Amputate Lower Leg At Knee
27880	Amputation Of Lower Leg
27881	Amputation Of Lower Leg
27882	Amputation Of Lower Leg
27884	Amputation Follow-Up Surgery
27886	Amputation Follow-Up Surgery
27888	Amputation Of Foot At Ankle
27889	Amputation Of Foot At Ankle
28800	Amputation Of Midfoot
28805	Amputation Thru Metatarsal
28810	Amputation Toe & Metatarsal
28820	Amputation Of Toe
28825	Partial Amputation Of Toe

Appendix 4: Insulin Medications

Subclass	Medication
Diabetes Therapy, Animal Insulins	Iletin II NPH (Pork)
	Iletin II Regular (Pork)
Human Insulin, Analog Combinations	Humalog Mix 50-50
	Humalog Mix 50-50 Kwikpen
	Humalog Mix 75-25
	Humalog Mix 75-25 Kwikpen
	Novolog Mix 70-30
	Novolog Mix 70-30 Flexpen
Human Insulin, Analog Fast Acting	Apidra
	Apidra Solostar
	Humalog
	Humalog Kwikpen
	Humalog Pen
	Novolog
	Novolog Flexpen
Human Insulin, Analog Long Acting	Lantus
	Lantus Solostar
	Levemir
	Levemir Flexpen
Human Insulin, Combinations	Humulin 50/50
	Humulin 70/30
	Humulin 70/30 Pen
	Insulin Nph/Reg 70-30 Innolet
	Novolin 70/30
	Novolin 70/30 Innolet
	Novolin 70/30 Penfill
Human Insulin, Fast Acting	Exubera Combination Pack 12
	Exubera Combination Pack 15
	Exubera Kit
	Exubera Patient Pack
	Humulin R
	Humulin R U-500 "Concentrated"
	Novolin R
	Novolin R Innolet
	Novolin R Penfill
	Velosulin BR (RDNA)
	Velosulin Human BR
Human Insulin, Intermediate Acting	Humulin L
	Humulin N
	Humulin N Pen
	Insulin NPH Human Recomb
	Insulin NPH Innolet
	Novolin N
	Novolin N Innolet
	Novolin N Penfill
Human Insulin, Long Acting	Humulin U

Appendix 5: Oral Antidiabetic Medications.

Subclass	Medication
Alpha-Glucos Inhibitor, Alone	Acarbose
	Glyset
	Precose
Biguanides, Alone	Appformin
	Appformin-D
	Fortamet
	Glucophage
	Glucophage XR
	Glumetza
	Metformin
	Metformin HCL
	Riomet
Biguanides/Sulfonylurea Combination	Glipizide-Metformin
	Glucovance
	Glyburide Micronized-Metformin
	Glyburide-Metformin
	Metaglip
DPP-4 Inhibitor, Alone	Januvia
	Onglyza
DPP-4 Inhibitor/Biguanide Combination	Janumet
Glinide/Biguanide Combination	Prandimet
Glinides, Alone	Nateglinide
	Prandin
	Starlix
Glitazone/Biguanide Combination	Actoplus Met
	Actoplus Met XR
	Avandamet
	Duetact
Glitazone/Sulfonylurea Combination	Avandaryl
Glitazones, Alone	Actos
	Avandia
Human Amylin Analogs	Symlin
	Symlinpen 120
	Symlinpen 60
Human GLP-1 Analogs	Byetta
	Victoza
Sulfonylureas	Amaryl
	Chlorpropamide
	Diabeta
	Diabinese
	Glimepiride
	Glipizide
	Glucotrol
	Glucotrol XL
	Glyburide
	Glyburide Micronized
	Glycron

Glynase
Micronase
Tolazamide
Tolbutamide

FOOTNOTES

[1] "The Long-Term Outlook for Health Care Spending, Introduction and Summary, November 2007," Congressional Budget Office; retrieved 12/27/2010 from http://www.cbo.gov/ftpdocs/87xx/doc8758/MainText.3.1.shtml.

[2] Joumard, I., C. André and C. Nicq (2010), "Health Care Systems: Efficiency and Institutions", Figure 1 Panel C, OECD Economics Department Working Papers, No. 769, OECD Publishing. doi: 10.1787/5kmfp51f5f9t-en. Retrieved 12/27/2010 from http://dx.doi.org/10.1787/5kmfp51f5f9t-en.

[3] Actual text of the Act: http://fdsys.gpo.gov/fdsys/pkg/BILLS-111hr1ENR/pdf/BILLS-111hr1ENR.pdf. Retrieved 12/27/2010.

[4] Department of Health and Human Services, "Agency-Wide Plan [for ARRA]," Retrieved 12/27/2010 from http://www.hhs.gov/recovery/reports/plans/index.html.

[5] Peter Szolovits and Stephen G. Pauker. Categorical and probabilistic reasoning in medical diagnosis. *Artificial Intelligence*, 11:115-144, 1978. Reprinted in W. J. Clancey and E. H. Shortliffe, *Readings in Medical Artificial Intelligence: The First Decade*, Chapter 9, Addison-Wesley, 1984. http://groups.csail.mit.edu/medg/people/psz/ftp/Szolovits-Pauker78/Szolovits-Pauker78.pdf.

[6] Harry E. Pople: On the Mechanization of Abductive Logic. IJCAI 1973: 147-152

[7] Harry E. Pople: The Formation of Composite Hypotheses in Diagnostic Problem Solving: An Exercise in Synthetic Reasoning. IJCAI 1977: 1030-1037.

[8] Shortliffe, E.H.; Buchanan, B.G. (1975). "A model of inexact reasoning in medicine". *Mathematical Biosciences* **23** (3-4): 351–379.

[9] Peirce, C. S. (1883), "A Theory of Probable Inference" in *Studies in Logic*. http://ia600106.us.archive.org/10/items/studiesinlogic00peiruoft/studiesinlogic00peiruoft.pdf.

[10] Vasant Dhar and Roger Stein, *Seven Methods for Transforming Corporate Data into Business Intelligence*, Prentice-Hall, 1996.

[11] Jeffrey D. Scargle (2000). "Publication Bias: The "File-Drawer Problem in Scientific Inference". Journal of Scientific Exploration 14 (2): 94–106.

[12] Miller, Amalia R. and Tucker, Catherine (February 2009), "Privacy Protection and Technology Diffusion: The Case of Electronic Medical Records". NET Institute Working Paper No. 07-16. Available at SSRN: http://ssrn.com/abstract=960233.

[13] SAS Institute, "Enterprise Miner Product Documentation," Retrieved 12/27/10 from http://support.sas.com/documentation/onlinedoc/miner/.

[14] National Institute of Diabetes and Digestive and Kidney Diseases. National Diabetes Statistics, 2011. National Diabetes Information Clearinghouse. Retrieved 3/3/11 from http://diabetes.niddk.nih.gov/DM/PUBS/statistics/.

[15] Centers for Disease Control, "Diabetes Successes And Opportunities For Population-Based Prevention And Control, At-A-Glance 2010," Retrieved 12/27/10 from http://www.cdc.gov/chronicdisease/resources/publications/aag/pdf/2010/diabetes_aag.pdf.

[16] David M. Nathan, John B. Buse, Mayer B. Davidson, Ele Ferrannini, Rury R. Holman, Robert Sherwin, Bernard Zinman. Medical Management of Hyperglycemia in Type 2 Diabetes: A Consensus Algorithm for the Initiation and Adjustment of Therapy. Diabetes Care January 2009 vol. 32 no. 1 193-203. Retrieved 11/5/2011 from http://care.diabetesjournals.org/content/32/1/193.long.

[17] See Appendix 1: Medical Claim Coding for Uncomplicated Type 2 Diabetes.

[18] See Appendix 2: Medical Claim Coding for Type 1 Diabetes.

[19] See Appendix 4: Insulin Medications

[20] See Appendix 3: Medical Claim Coding for Diabetes Complications.

[21] See Appendix 5: Oral Antidiabetic Medications & Subclasses.

[22] U.S. Census Bureau. "Census Regions and Divisions of the United States", retrieved 11/5/2011 from http://www.census.gov/geo/www/us_regdiv.pdf

[23] David M. Nathan, John B. Buse, Mayer B. Davidson, Ele Ferrannini, Rury R. Holman, Robert Sherwin, Bernard Zinman. Medical Management of Hyperglycemia in Type 2 Diabetes: A Consensus Algorithm for the Initiation and Adjustment of Therapy. Diabetes Care January 2009 vol. 32 no. 1 193-203. Retrieved 11/5/2011 from http://care.diabetesjournals.org/content/32/1/193.long.

[24] "Applied Logistic Regression, Second Edition" by David W. Hosmer and Stanley Lemeshow

[25] Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. Peninsula Publishing, Los Altos, CA.

[26] Karter AJ, Ahmed AT, Liu J, Moffet HH, Parker MM. Pioglitazone initiation and subsequent hospitalization for congestive heart failure. Diabet Med. 2005;22:986-993.

[27] Nichols GA, Koro CE, Gullion CM, Ephross SA, Brown JB. The incidence of congestive heart failure associated with antidiabetic therapies. Diabetes Metab Res Rev. 2005;21:51-57.

[28] Tzoulaki I, et al. Risk of cardiovascular disease and all cause mortality among patients with type 2 diabetes prescribed oral antidiabetes drugs: retrospective cohort study using UK general practice research database. BMJ 2009;339:b4731.

[29] Bowker SL, Majumdar SR, Veugelers P, Johnson JA. Increased cancer-related mortality for patients with type 2 diabetes who use sulfonylureas or insulin. Diabetes Care. 2006 Feb;29(2):254-8.

[30] Simpson SH, Majumdar SR, Tsuyuki RT, Eurich DT, Johnson JA. Dose-response relation between sulfonylurea drugs and mortality in type 2 diabetes mellitus: a population-based cohort study. CMAJ. 2006 Jan 17;174(2):169-74.

[31] Eurich DT, Majumdar SR, McAlister FA, Tsuyuki RT, Johnson JA. Improved clinical outcomes associated with metformin in patients with diabetes and heart failure. Diabetes Care. 2005 Oct;28(10):2345-51.

[32] Eurich DT, Simpson SH, Majumdar SR, Johnson JA. Secondary failure rates associated with metformin and sulfonylurea therapy for type 2 diabetes mellitus. Pharmacotherapy. 2005 Jun;25(6):810-6.

[33] Johnson JA, Simpson SH, Toth EL, Majumdar SR. Reduced cardiovascular morbidity and mortality associated with metformin use in subjects with Type 2 diabetes. Diabet Med. 2005 Apr;22(4):497-502.

[34] Kalsekar I, Latran M. Economic effect of augmentation strategies in patients with type 2 diabetes initiated on sulfonylureas. Manag Care Interface. 2007 Sep;20(9):39-46.

[35] Hansen RA, Farley JF, Droege M, Maciejewski ML. A retrospective cohort study of economic outcomes and adherence to monotherapy with metformin, pioglitazone, or a sulfonylurea among patients with type 2 diabetes mellitus in the United States from 2003 to 2005. Clin Ther. 2010 Jul;32(7):1308-19.

[36] Soni, A. Top 10 Most Costly Conditions among Men and Women, 2008: Estimates for the U.S. Civilian Noninstitutionalized Adult Population, Age 18 and Older. Statistical Brief #331. July 2011. Agency for Healthcare Research and Quality, Rockville, MD.
http://www.meps.ahrq.gov/mepsweb/data_files/publications/st331/stat331.pdf