Effective Health Care
Research Reports

Number 18

Use of Electronic Medical Records and Administrative Claims Data for Assessing Type 2 Diabetes Care

Suzanne L. West, M.P.H., Ph.D.
Zhiwen Liu, M.D., M.S.
J. Nikki McKoy, B.A.
Maryann Oertel, Pharm.D., B.C.P.S.
Catherine Blake, Ph.D.
Bob Schwartz, B.A., M.A.
Frances Ochart, B.S.
Timothy S. Carey, M.D., M.P.H.

Research from the Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) Network

Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

January 2010

The DEcIDE (Developing Evidence to Inform Decisions about Effectiveness) network is part of AHRQ's Effective Health Care Program. It is a collaborative network of research centers that support the rapid development of new scientific information and analytic tools. The DEcIDE network assists health care providers, patients, and policymakers seeking unbiased information about the outcomes, clinical effectiveness, safety, and appropriateness of health care items and services, particularly prescription medications and medical devices.

This report is based on research conducted by the University of North Carolina at Chapel Hill DEcIDE (Developing Evidence to Inform Decisions about Effectiveness) Center under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-05-0040-1). The AHRQ Task Order Officer for this project was Chunliu Zhan, M.D., Ph.D.

None of the authors has a financial interest in any of the products discussed in this report.

A shortened form of this report has also been published: West SL, Blake C, Liu Z, et al. Reflections on the use of electronic health record data for clinical research. Health Informatics J 2009 Jun;15(2):108-21.

Suggested citation:

West SL, Liu Z, McKoy JN, Oertel M, Blake C, Schwartz B, Ochart F, Carey TS. Use of electronic medical records and administrative claims data for assessing type 2 diabetes care. Effective Health Care Research Report No. 18. (Prepared by University of North Carolina at Chapel Hill DEcIDE Center Under Contract No. 290-05-0040-1). Rockville, MD: Agency for Healthcare Research and Quality. January 2010. Available at: http://effectivehealthcare.ahrq.gov/reports/final.cfm.

Contents

Author affiliations:

Suzanne L. West, M.P.H., Ph.D.[a,b]
Zhiwen Liu, M.D., M.S.[b]
J. Nikki McKoy, B.A.[c]
Maryann Oertel, Pharm.D., B.C.P.S.[d]
Catherine Blake, Ph.D.[e]
Bob Schwartz, B.A., M.A.[f]
Frances Ochart, B.S.[f]
Timothy S. Carey, M.D., M.P.H.[f,g]

[a]RTI International, Research Triangle Park, NC
[b]Gillings School of Global Public Health, University of North Carolina at Chapel Hill, NC
[c]Institute for Medicine and Public Health, Vanderbilt University, Nashville, TN
[d]UNC Health Care and Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, NC
[e]School of Information and Library Science and UNC Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, NC
[f]Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, NC
[g]Department of Medicine, University of North Carolina at Chapel Hill, NC

Abstract

Background: Evaluating the adequacy of diabetic care requires access to clinical data not available from administrative claims data. Electronic medical records (EMRs) may be a valuable resource for studying diabetes and related complications. More medical systems are adopting EMRs, but we know relatively little about the opportunities and challenges of using these newer EMRs for clinical and health services research.

Objective: This project had 2 parallel goals. The informatics goal was to understand the challenges in conducting research using EMR data and to compare the usefulness of these data with that of administrative data from the North Carolina (NC) Medicaid program. The clinical goals were to evaluate medication use in patients with newly diagnosed diabetes mellitus type 2 (DMT2), some of whom may have comorbid conditions such as hypertension and dyslipidemia, and to assess the adequacy of diabetic care.

Design: Cohort analyses of patients with newly diagnosed DMT2.

Setting and Patients: Data from 2 populations were used: patients seen by clinicians affiliated with the University of North Carolina Health Care System (UNCHCS) and individuals covered by NC Medicaid. DMT2 was required to have been diagnosed after January 1, 2001 for the UNC cohort and after January 1, 2002 for the NC Medicaid cohort.

Methods: We conducted a literature search to identify publications focusing on the efficacy and effectiveness of medications for newly diagnosed DMT2. We used this review to determine which glycemic indicators should be assessed to determine the adequacy of medications used in the 2 study populations.

The EMR we used for this project was from the UNCHCS (WebCIS). Because UNC is an academic medical center, we developed inclusion and exclusion criteria to restrict the population to those who appeared to be seen regularly by UNC clinicians. We used the patient problem, laboratory test, medication prescribing, and transcription files to identify newly diagnosed DMT2 patients. We developed and tested a medical record abstraction form to guide clinical review of the EMR data, whereby clinicians evaluated the adequacy of glycemic, lipid, and hypertension control. For the NC Medicaid data, we used typical algorithms to identify new DMT2, which reflected eligibility for coverage and the absence of DMT2 disease and medication codes 1 year before diagnosis. In both patient cohorts, we described comorbidities present when DMT2 was diagnosed and medications used within 12 months after diagnosis. We also identified patients who died, had a stroke, or who had a myocardial infarction (MI) early in the course of their DMT2 and described their care in the 12 months before the outcome.

Results: We identified numerous challenges in meeting the informatics goal for this project. First, although structured data such as laboratory results had been deidentified by anonymizing the medical record number, full-text transcription notes from the visit files still contained identifiable patient information. These notes had to be manually deidentified before leaving the clinical site repository. We incorporated text data-mining procedures to use the free-text visit data most efficiently for the project.

For the clinical goal of assessing medication use in DMT2 patients, we focused on patients with adverse outcomes. In all, 78 of the 1664 WebCIS patients had an MI or stroke or died after DMT2 diagnosis, only 30 of whom had a truly new diagnosis of DMT2 based on manual record review. From the 2794 newly diagnosed DMT2 NC Medicaid patients, we identified 49 who had an MI, 173 who had a stroke, and 14 who had both events (death was not captured in this population).

Of the 30 newly diagnosed WebCIS diabetics who had an MI or stroke after their DMT2 diagnosis, most had hypertension (HT) and/or dyslipidemia (DL) in addition to their DMT2 (71% of MI patients and 60% of stroke patients). Only 20% of the patients who had both HT and DL in addition to DMT2 received adequate pharmacological treatment. Of the 19 WebCIS patients who had comorbid HT, only 42% were prescribed an angiotensin-converting enzyme inhibitor (ACEI) or angiotensin receptor blocker (ARB) soon after their DMT2 was diagnosed.

We could not assess the adequacy of glycemic, lipid, or hypertension control in the Medicaid population due to a lack of clinical information in the claims data. There were 49 patients who had an MI, 49% of whom had comorbid DM, HT, and DL, whereas only 30.6% of the patients who had a stroke had comorbid disease. Of the 1753 Medicaid patients who had DMT2 with comorbid HT, 61.5% were dispensed an ACEI or ARB within 12 months after the DMT2 diagnosis. The proportion of WebCIS diabetic patients who had an MI and/or stroke and comorbid HT and/or DL was larger than that in the Medicaid cohort, possibly reflecting a greater burden of illness at the tertiary care center.

Conclusions: By conducting similar analyses in both patient populations, we could discern the value of each data resource for conducting observational research on DMT2. We applied many of the principles of claims-based analysis to the EMR but faced many new challenges throughout the project.

1. With regard to the informatics goal, the UNC WebCIS EMR was a rich data source with good representation across all UNC care sites. However, the penetration and use of some of the specific functions (electronic prescribing) within the WebCIS system were variable and may have resulted in underreporting of medications. Although text-mining methods were helpful in addressing this issue, comparison of WebCIS data versus text-mining results indicated that neither data source provided complete reporting of patients' medications.
2. Given the fragmentation of healthcare in the US, we do not have access to longitudinal data sources that allow us to determine when a condition was first diagnosed. Identifying the onset of a chronic condition such as DMT2 is difficult when using any electronic database, whether EMRs or administrative claims databases. The challenge is to ensure high sensitivity and high specificity to reduce the number of false-positive cases, especially when dealing with relatively prevalent diseases such as diabetes. Validation of case status is critical to ensure accurate disease classification. Discrete EMR values may be validated using full-text information from the same EMR, whereas validation of administrative claims data requires access to outside sources.
3. Text-mining methods may be useful for ascertaining information critical to research studies and drug safety assessments. Greater focus should be placed on methods to

maximize automated extraction of clinically relevant text data regarding historical notations, negation, diagnostic issues, and adverse effects of medications.

4. Substantial technical challenges are present when using clinical data from an EMR for research. The promise of this activity is substantial and overall we were encouraged, but we underestimated the technical issues involved in carrying out research with an information-rich resource that has both discrete and free-text data.
5. With regard to the clinical goal of this project, patients who suffered adverse outcomes (MI, stroke, death) early in the course of their DMT2 had substantial preexisting and coexisting morbidities. Treatment of these coexisting conditions, specifically dyslipidemia and hypertension, might have reduced the complication rate. These patients had relatively high utilization of care in the year before the adverse outcome, suggesting ample opportunity for intensification of therapies.
6. Overall patterns of care and diabetes comorbidities in the Medicaid population were similar to those in the WebCIS population. The use of medications appeared to be somewhat more intensive in the Medicaid population. We caution that the populations were likely not directly comparable regarding factors such as demographics, socioeconomic status, and severity of disease.

Introduction

The estimated prevalence of diabetes mellitus type 2 (DMT2) in the United States was 20.8 million in 2005, representing about 7% of the US population.[1] Only 70% of these cases are diagnosed clinically.[1] Thus ~6.2 million Americans with undetected disease are not being monitored and treated for DMT2 and its complications. Even more alarming is that almost 21% of people over age 60 are estimated to have DMT2. The prevalence in the Latino population is higher than in non-Hispanic whites, which is of concern as the size of this ethnic group has increased greatly in the past 5 years.[2]

There are very few data regarding the natural history of newly diagnosed DMT2 after initiation of antidiabetic therapy. The United Kingdom Prospective Diabetes Study (UKPDS) enrolled newly diagnosed diabetics, treated them with diet therapy for 3 months, and then randomized them to therapy with diet alone, insulin or a sulfonylurea (chlorpropamide or glyburide), or, if the patient weighed 120% of ideal body weight, metformin. Because the metformin group was heavier, it was not comparable to the other 2 study groups.[3] Recently, Nichols and colleagues used data from Kaiser Permanente Northwest to assess secondary failure of initial metformin therapy, defined as the need to add or switch drugs.[4] Patients receiving metformin monotherapy whose glycosylated hemoglobin (HbA1c) level decreased to <8% within 1 year had a lower rate of secondary failure over time than treated patients whose HbA1c level never decreased to this level. Compared with patients whose HbA1c level decreased to <6.0%, the hazard ratio (95% confidence interval [CI]) for secondary failure was 3.3 (1.8–5.9) for patients whose HbA1c level decreased to 6%–6.9% and 6.5 (3.6–12.0) for those whose HbA1c level decreased to 7.0%–7.9%.

Cook and colleagues[5] studied patients in the UK General Practice Research Database (GPRD) who had an initial diagnosis of DMT2 and sulfonylurea or metformin monotherapy prescribed between January 1, 1998 and March 15, 2004. Among patients prescribed a sulfonylurea, the HbA1c level decreased to <6.5% at least once during the first year of therapy in 33% of patients, to <7.0% in 54%, and to <7.5% in 72%. Corresponding figures for metformin were 24%, 45%, and 64%. The pretreatment HbA1c level was the most important predictor of a decrease in HbA1c level to <7.0%. Of the patients whose HbA1c level had decreased to <7.0% with either therapy, about 30% had persistently higher levels after 1 year of follow-up. This proportion increased to slightly more than 50% after 2 years and to about 70% after 3 years.

Besides sulfonylureas and metformin, some clinicians consider the thiazolidinediones as an alternative first-line therapy for DMT2, although these drugs have not been approved for this indication. A descriptive study that provides the proportion of newly diagnosed diabetics who are receiving each type of antidiabetic drug would be a valuable addition to the literature.

DMT2 itself is a serious disease, but its consequences also include severe complications such as myocardial infarction (MI) and stroke. Selby and colleagues identified the prevalences of diabetes (DM), dyslipidemia (DL), and hypertension (HT), alone and in combination, using electronic medical records (EMRs) from the Northern California Kaiser Permanente Health Plan.[6] The authors compared their results with those from the third National Health and Nutrition Examination Survey (NHANES), which included self-reported disease information and laboratory test results for HbA1c and low-density lipoprotein cholesterol (LDL-C) levels. The prevalence of DMT2 alone in the Kaiser population was only 0.6%, whereas 1.2%, 1.1%, 3.7% of the population had DM and HT; DM and DL; and DM, DL, and HT, respectively. Similar

results were found from NHANES, except that, because screening for these conditions was complete in the NHANES population, a greater percentage of the NHANES population had DL alone and in combination with HT. This study and others confirm that a major challenge of diabetes care is not only to intervene according to the blood glucose level but also to treat these major comorbid conditions.

Diabetic patients are often undertreated with medications such as antidiabetic drugs, lipid-lowering agents, and angiotensin-converting enzyme inhibitors (ACEIs) or angiotensin receptor blockers (ARBs).[7-13] Several publications have evaluated the use of medications to treat DMT2 alone, whereas others have evaluated the use of therapies for DM, HT, and DL in diabetic populations. Some investigators have evaluated combinations of antidiabetic therapies, and others have looked at adherence to antidiabetic therapy. Many different factors can be evaluated to assess the adequacy of diabetic care, but the major underlying issue is having the appropriate data resource to address the research questions of interest.

Using an academic medical system's EMR (UNC's WebCIS) that dates to the mid-1980s, we identified patients with newly diagnosed DMT2 who had also been prescribed antidiabetic medication. WebCIS is an EMR that contains clinical measures, such as HbA1c level, lipid levels, and blood pressure, that are very useful for assessing the adequacy of medications in controlling DM, DL, and HT. From this population of diabetics, we determined which patients had an MI or stoke after the DMT2 diagnosis. We report on the clinical care and comorbidities occurring within the 12 months before the outcome event in these patients. We did not have the time or funds to evaluate similar patterns in diabetics in the WebCIS database who did not have an MI or stroke. The reasons for this will become apparent later in the report.

We then used an administrative claims database (NC Medicaid) to identify newly diagnosed diabetic patients who also received antidiabetic medications. We performed analyses similar to those done on WebCIS data, except that we also identified a cohort of individuals who had treated DMT2 but who did not have an MI or stroke. We looked at comorbidities such as HT and DL as well as the drugs used for their treatment.

Methods

This was a 3-phase project. The first phase was a literature search to identify the laboratory indicators used to assess glycemic control in newly diagnosed DMT2 patients. The second phase focused on using UNC's EMR, WebCIS, to identify the medications used for glycemic control in newly diagnosed DMT2 patients with and without dyslipidemia or hypertension. The third phase was similar to the second phase, except that we assessed medication use in the NC Medicaid population.

Literature Scan

Our task order was to focus on process outcomes related to DMT2, especially those regarding the misuse, overuse, or underuse of DMT2 medications. Thus, the initial Medline searches focused on studies addressing DMT2 process outcomes using administrative databases. After conducting this search and reviewing the abstracts of the retrieved articles, the project team realized that several studies had described process outcomes for DMT2 and that our study would not add significantly to this literature. More important, we realized that the available databases would not be appropriate to assess the usefulness and validity of new process measures we identified or developed. We did note that there was little information on the initial medications

that diabetic patients used and the resulting clinical outcomes. We discussed pursuit of this question with the Project Officer, Chunliu Zhan, and obtained approval to move forward using both the WebCIS and NC Medicaid datasets. Understanding the challenges and opportunities of generating practical findings from clinically based EMR data was seen as a significant opportunity that could inform future work commissioned by AHRQ as part of its Health Information Technology initiatives.

To address this revised objective, we developed 2 key questions, inclusion/exclusion criteria, and a search strategy to identify literature for the 2 key questions. We used established methods to conduct the literature search and to develop an evidence table on the use of medications to treat newly diagnosed DM. The 2 key questions for the literature search were:

Key Question 1: What are the primary DMT2 and cardiovascular outcomes and laboratory tests evaluated in comparative studies of antidiabetic therapy in newly diagnosed patients?

Key Question 2: Is there a difference in DMT2 outcomes such as HbA1c levels, fasting plasma glucose levels, or switching/augmentation of therapy depending on the initial medications DMT2 patients use?

We included 1 article[14] from the initial search on process outcomes in the second search addressing comparative studies of initial antidiabetic therapies. The inclusion and exclusion criteria for the second Medline search are listed in Table 1. The search strategy for addressing the 2 key questions is provided in Table 2. This search was conducted on February 7, 2006.

Table 1. Inclusion and exclusion criteria for Medline search

	Inclusion	Exclusion
Databases	MEDLINE	Other databases
Languages	English only	Other languages
Populations	Humans only	Animal studies
Time period	Publication from 1995 to the present	Publication before 1995
Study design	RCT, cross-sectional, cohort, case-control, systematic reviews, and meta-analysis	Letters, editorials, and non-systematic reviews
Study population	Patients with newly diagnosed diabetes or diabetes treated with diet therapy alone. Head-to-head studies only.	Use of other antidiabetic therapy, alone or in combination

Table 2. Literature search strategy to address the 2 key questions

Query	Search Terms	Citations
1	Yamanouchi T[au]	146
3	Diabetes mellitus, type 2/drug therapy [mh]	7690
4	"newly diagnosed"	11932
5	#3 AND #4	127
6	Limit to English language	116
17	Comparative Study [MeSH] OR (Randomized Controlled Trials"[MeSH] OR Randomized Controlled Trial"[Publication Type] OR "Single-Blind Method"[MeSH] or "Double-Blind Method"[MeSH] OR "Random Allocation"[MeSH]) Limit to English	1207009
18	#17 AND #3 Limit to English	2078
19	#18 Limit to English, Humans, Publication Date from 1995	1539
21	Insulin[MeSH] Limit to English, Humans, Publication Date from 1995	22329

22	#19 NOT #21 Limit to English, Humans, Publication Date from 1995	796
23	#22 NOT #7 Limit to English, Humans, Publication Date from 1995	783
24	#19 NOT #21 Limit to English, Humans, Publication Date from 1995	15
25	#19 NOT #21 Limit to English, Humans, Letter, Publication Date from 1995	41
26	Select 95 documents	95
27	#22 NOT #26 Limit to English, Humans, Publication Date from 1995	701
28	#27 NOT #24 OR #25 Limit to English, Humans, Publication Date from 1995	648
30	Diabetes mellitus, type 2/drug therapy Limit to English, Humans, Publication Date from 1995	4516
31	#30 AND #28 Limit to English, Humans, Publication Date from 1995	648
34	("Outcome Assessment (Health Care)"[MeSH] OR "Fatal Outcome"[MeSH] OR "Drug Utilization Review"[MeSH]) Limit to English, Humans, Publication Date from 1995	230962
36	("Treatment Outcome"[MeSH] OR "Outcome and Process Assessment (Health Care)"[MeSH]) Limit to English, Humans, Publication Date from 1995	217583
37	#36 OR #34 Limit to English, Humans, Publication Date from 1995	237717
38	#37 AND #31 Limit to English, Humans, Publication Date from 1995	132
45	Hemoglobin A, Glycosylated"[MeSH] OR "Creatinine"[MeSH]	42513
46	#31 AND #45	215
48	Drug Therapy, Combination [mh]	159175
49	#31 NOT #48	542
50	**#49 AND #45**	**163**
53	**Articles related to Yamanouchi T [au]**	**180**

The citations from Queries 50 and 53 were used for the literature scan. Dual abstract review was conducted by the DEcIDE project director (SW) and project pharmacist (MO). From Query 50 and using the inclusion and exclusion criteria above, we selected 13 abstracts for retrieval of the full article. Of these, 5 were included in the literature review table. From Query 53 (180 citations), we selected 31 abstracts for retrieval of the full article. Of these 31, 15 were included in the literature review table.

Simultaneously, the Johns Hopkins Evidence-based Practice Center was completing a systematic review of the comparative efficacy and safety of various antidiabetic medications.[15] We obtained their citation database and selected 36 additional abstracts for retrieval of the full-text article. Of these articles, 12 were included on the basis of the inclusion/exclusion criteria described above.

One of the authors (MO) initially reviewed all of the articles and developed the evidence tables. A second investigator (TC) reviewed the table for accuracy and clinical relevance. **Appendix A** contains the evidence table for the 26 studies (32 publications) reviewed and provides the study design, patient population, clinical outcomes evaluated, and limitations of each of the comparative studies. The evidence table is organized alphabetically by the last name of the first author and by publication date.

Data Sources

For Phases 2 and 3, which focused on the medications used for glycemic control in newly diagnosed DMT2 patients, we used 2 data sources: University of North Carolina Health Care System's (UNCHCS) EMR (WebCIS) and the NC Medicaid database. The focus of the WebCIS analysis was to use the clinical data describing initial treatment patterns in newly diagnosed DMT2 patients to assess the adequacy of glycemic control. In patients with comorbid HT and/or DL, we also evaluated maintenance of systolic and diastolic blood pressures and appropriate lipid levels. For the NC Medicaid administrative claims data analysis, we conducted similar

analyses using diagnoses, dispensed pharmaceuticals, and the use of medical procedures to assess the adequacy of treating DMT2 and its associated comorbidities in this population.

The UNC Health Care System

UNC Hospitals include North Carolina Children's Hospital, North Carolina Memorial Hospital, North Carolina Neurosciences Hospital, and North Carolina Women's Hospital. UNC Hospitals is a public, academic medical center affiliated with the UNC at Chapel Hill School of Medicine. Table 3 shows selected summary statistics for the Hospitals. All of the hospital facilities, their patients, and care providers are served by WebCIS.

Table 3. Basic statistics about UNC hospitals

Licensed beds	688
Average length of stay	6.4 days
Attending physicians	983
House staff (residents)	550
Discharges (not newborns)	30,212
Clinic visits (2002)	699,984
ER visits	41,829
Surgical cases	24,445
Births (deliveries)	2,919

UNCHCS's WebCIS

Below, we chronicle the history of WebCIS because it explains how certain of its features, such as the entry of prescribed medications, improved over time and why we needed to consider other ways of identifying medication information that was critical for project success.

The UNC School of Medicine and UNCHCS have built their computerized medical record over the past 20 years. Efforts began in the early 1980s through a general internal medicine initiative funded by the Robert Wood Johnson Foundation. This project had the overall goal of improving quality of care and education in the general medicine practice. It included an effort to partly computerize records, because outpatient records were unavailable for 10%–15% of visits. To overcome this problem, an outpatient "Mini-Medical Record" (MMR) was developed that contained essential information for practicing clinicians. It consisted of a computerized face sheet that included patient demographics linked to the registration system, a problem list, vital signs, a medication list that doubled as a prescription document for internal use, and limited health maintenance prompts.

During the mid- to late 1980s, the system was disseminated for use in essentially all outpatient practices at UNC. An evaluation from the early 1990s showed that the MMR was moderately useful for outpatient care; clinicians and administrators also recognized potential future promise.[16] There were several limitations to the system:

- The interval history, physical examination and plan, and laboratory data and discharge summaries from hospitalizations were unavailable.
- Medication information was inadequately coded, and the internal prescriptions were accepted only by the UNCHCS Pharmacy. This led to both duplicate and missing information on medication use.
- The MMR face sheet was a document printed on NCR pressure-sensitive paper at each visit. Clinicians wrote new information and changes on these sheets. Data entry by clerks was a substantial personnel cost.

- The system was written in FORTRAN and was becoming increasing difficult to maintain.

In 1990, the UNCHCS Information Services Division, working with medical staff, began to develop a fully computerized medical record that would replace the MMR.[17] UNCHCS concluded that none of the commercially available systems would be adequate and elected to develop its own system. The major driver of this decision was the desire to develop a single informatics system able to support an academic teaching practice, in both inpatient and outpatient sites of care. The initial version, the Clinical Information System (CIS) version 1.0, went live in 1991. It included the elements of the older MMR plus access to discharge summaries and radiology reports. Version 2.0 became available in early 1993 and included information from the inpatient wards. Version 2.1 allowed access to the Laboratory Information System so that laboratory tests could be viewed in real time. The final implementation step was the incorporation of a security system that recorded the date and time of each transaction and used a single user login for essentially all clinical data. This homegrown system was used until the Web-based version (WebCIS) was developed.

Development of WebCIS was guided by the functionality of the CIS and the following principles:

- Facilitate the access of patient records by clinical faculty throughout the system of hospitals and outlying practices using a secure, Web-based system
- Allow patients to be followed throughout the inpatient and outpatient continuum of care
- Act as a common interface so that clinicians could examine problem lists, medication lists, laboratory data, clinical information, and reports with the same 'look' and navigation rules
- Incorporate problem and medication list information from the previous MMR, to retain all previous clinical information
- Link with legacy hospital systems such as patient registration, provider lists, and allergy systems, in real time and bi-directionally to the degree possible
- Build the system incrementally. For example, prompts and reminders would be added over time, as would linkages with the vendor systems such as Siemens computerized physician order entry (CPOE). This approach was very different from the "turnkey" approach attempted by other academic health systems.
- Ensure the use of coded information, such as International Classification of Disease (ICD-9-CM) codes for diagnoses and National Drug Codes (NDC) for medications, rather than free text, to enhance research and quality-of-care opportunities
- Use of a proprietary relational database, DB2, for data storage and production

The initial version of WebCIS became fully operational in April 2001, becoming the primary clinical record system used at UNC Hospitals. Small changes were incorporated over time. For example, in October of 2002, WebCIS 1.5 delivered telephone message services and a linkage to the Picture Archival System (PACS) for online display of medical imagery such as radiographs and computed axial tomography (CT) and magnetic resonance imaging (MRI) scans. A major overhaul of WebCIS occurred at the end of 2004, when version 2.0 fully replaced the outpatient MMR and incorporated the following functions:

- Direct entry of coded medications, allergies, and problem lists with ability to keep a dated, annotated record of active vs. inactive entries in these lists. The ability to print new and refill prescriptions was included in the medication module.
- Direct entry of coded vital signs and nursing notes at patient check-in the outpatient areas
- Incorporated automatic alerts to the treating provider for the following actions:
    - Specialty-specific reminders
    - Health maintenance, immunizations, and disease management tasks
    - Completed ancillary tests for current outpatients
    - Inpatient admissions
    - Inpatient deaths
- Expansion of Personal and Group patient lists to allow multiple personal lists and the ability to create or join new group lists
- Attending attestation fields and electronic signature of clinical notes for Medicare patients
- A script writer to generate prescriptions for outpatient medications

With each new version of WebCIS came increased functionality. The April 2005 revision produced version 2.5, which added a forms tool allowing direct data entry of inpatient notes (history and physical, progress, consults, procedure notes, and operative notes) and direct transmission of electronic prescriptions to area pharmacies. This tool pulls the relevant data from problems, medications, allergies, and the history into the note, thus eliminating double entry and streamlining documentation. The most important enhancement was the prescription-writing capability, which increased physician satisfaction and use of WebCIS dramatically.

The current WebCIS version includes information from these clinical and administrative areas:

- Patient appointment scheduling
- Patient demographic data
- Patient problem file
- Clinic visit data
- Hospital census information
- Medications prescribed
- Insurance provider
- Cardiology reports
- Pulmonary reports
- Peripheral vascular laboratory tests
- Laboratory results
- Pathology reports
- Radiology reports
- Respiratory therapy reports
- Electrocardiogram results
- Referring physician lists
- Transcribed notes
- Allergy lists
- Gastrointestinal procedures reports
- PACS medical imagery

Use of WebCIS for Clinical Research Projects

Before the DEcIDE project, WebCIS data were used for a UNCHCS diabetes management program under the direction of internists and pharmacists associated with the Department of Internal Medicine. This program was developed as a separate entity from

WebCIS, with the only linkage being a daily, automatic download of laboratory data for enrolled patients so that their diabetic control could be monitored. The current DEcIDE project is the first time that WebCIS has been used, in its entirety, for clinical research. Reporting the challenges encountered will be of assistance to future investigators.

As described above, WebCIS is maintained as a relational database that is composed of 47 separate data files (**Appendix B**). The medical record number is the primary linkage between the files, but patient name also appears in a few files. To deidentify the information, data in structured fields such as "name," "phone number," and "Social Security number (SSN)" were simply omitted from the data extraction. Other necessary fields such as "medical record number" were encrypted using a consistent encryption process. Fortunately, most of the WebCIS data used for this project consisted of structured data, which could be stripped of direct identifiers easily and accurately.

Because WebCIS had not previously been used for research, we could not anticipate all of the potential challenges we might face, overall and with respect to each individual data file. To gain an understanding of WebCIS and the data it held for diabetics, we asked the UNCHCS Information Technology (IT) team to extract all clinical data available for 5 diabetic patients in deidentified format. These data were downloaded from the "live" system as 41 (the 5 test patients did not have data for 6 of the 47 tables) separate text files. Although the IT group provided the data structure for reading the text files into SAS for data manipulation (see **Appendix B**), the SAS upload was cumbersome. The first WebCIS data extract contained null characters that were created when data from numeric fields containing integers were converted to character fields. The UNCHCS programmers minimized this problem in later data extracts, but some extraneous characters remained. Sheps programmers developed a program to identify and delete these extraneous characters from the files so that they could be more easily uploaded into SAS (**Appendix C**).

Reviewing the SAS data for the 5 test patients, we noted that 4 files appeared to be particularly useful for this study: the patient problems file (TPRBPAT.txt), drug file (TFDBPAT.txt), laboratory results file (TLABRSSC.txt), and visit transcription file (TTRNTEXT.txt). There were 2 very promising variables in the TPRBPAT.txt file: CPK_ICD9_CODE, or the ICD-9-CM code the clinician used to record the reason for the visit, and C_ONSET_DATE, the corresponding date of entry for the ICD-9-CM code. Although these variables might seem to indicate when a condition was first diagnosed, we quickly discovered that this assumption was true only for acute conditions such as MI and stroke. For chronic conditions such as diabetes, the clinician entered a new date each time care was provided; older dates were also retained in the files.

Given that the primary focus of the project was to compare the data from WebCIS and NC Medicaid as data resources for assessing current treatment patterns for newly diagnosed DMT2 patients, the validity of the medication data was critical. For WebCIS, the medication file (TFDBPAT.txt) contains the NDC, drug name, strength, units, quantity, American Hospital Formulary Service code, drug frequency code, the date the clinician prescribed the drug, and the date the drug was inactivated. Inactivation occurs when the clinician writes a new prescription for the same regimen of the same drug. If the clinician prescribes the drug only once with no refills (this may occur with antibiotics), then the inactivation date is blank.

As noted above, the ability to provide electronic prescriptions greatly aided drug prescribing and, with it, use of this WebCIS module, but its completeness was still a concern. We discuss how we addressed this issue in the text data-mining section of this report. We also

describe a small subproject undertaken to assess the agreement between the data from text mining and from the TFDBPAT.txt file.

The laboratory file (TLABRSSC.txt) was particularly useful, as it contained results from as far back as January 1994. The file contains information on the test, the result, the normal range for the test, and whether the result was outside of the normal range. The UNCHCS laboratories use their own coding system, not the industry standard, Logical Observation Identifier Names and Codes (LOINC®).[18] LOINC was developed in 1996, 2 years after the earliest UNC laboratory results were available. Thus issues of consistency arose during this project. For example, WebCIS test codes depend on where the patient's sample is drawn from within the UNCHCS system, to track use across clinical sites. Eight different codes were used to identify HbA1c level within the WebCIS laboratory files.

The transcription notes file (TTRNTEXT.txt) from outpatient visits is one of the most informative files in WebCIS, but data mining was required to realize its potential for research. The notes describe the salient aspects of the patient's medical care, including medications prescribed and being used, as well as historical data regarding disease onset. The transcribed notes are in the form of unstructured text derived from provider dictation, so they may contain confidential patient information. Special deidentification programs were developed and run on this file before it was used for research, based on review of the clinical data downloaded on the 5 test patients[a] (**Appendix D**).

Deidentification Procedure for the TTRNTEXT.txt File

The TTRNTEXT.txt data file contained long text fields that appeared to be "catch-all" fields for storing unstructured data from external systems. A large portion of UNC's medical school uses a voice-recognition system operated by an external vendor, but not all practices use the same external vendor. This led to variability in headers, formatting, etc. A single field in 1 record could contain a series of text strings formatted in a particular way, whereas the same field in a second record could contain different types of text data formatted completely differently. Further, the same field in a third record could contain a text translation of a voice memo. WebCIS staff confirmed that some specialty clinics use other systems in the clinic to collect information and then "push" the selected data into WebCIS. The UNCHCS also has speech-recognition software to convert doctors' dictated memos to text for inclusion in the medical record.

Storing these chunks of data in a single, long text field is a quick and easy way to keep the associated data with the appropriate patient, but this approach presents problems when trying to analyze individual elements within the chunks. Finding identifying data for removal was difficult. Some records contained attribute names that could be used for locating identifying data. Terms such as "Name:" or "Patient:" or "MRNO (medical record number):" sometimes preceded identifying data (e.g., "Patient: John Doe"). For some records, a text-filtering program could search for these markers and replace the text just after the markers. However, the markers were inconsistent across specialty clinic records, some records contained identifiers with no markers,

[a] As we had Institutional Review Board (IRB) approval and a data-use agreement for a limited dataset containing *deidentified* data, our access to identifiable data was considered an adverse event and was reported to the Public Health IRB.

and determining the number of characters to replace after the markers was even more problematic.

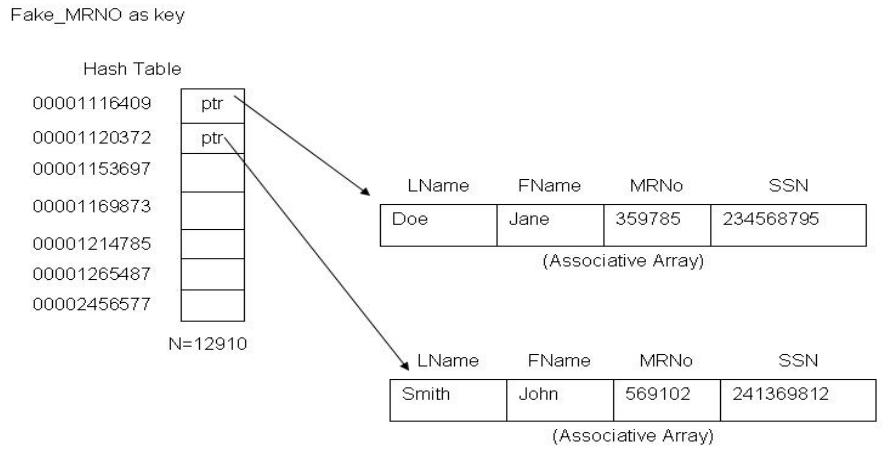
Spoken notes are free text with almost no data structure and few, if any, markers for identifying data. For example, a patient's name or medical record number can occur anywhere within a 2000-character text field. Moreover, physicians creating these notes refer to patients in varied ways: "Ms. Doe," "Jane," "Jane Doe," etc. Although "Jane" in 1 sentence may not be enough information to identify an individual, the combination of that with "Ms. Doe" in a later sentence would be. An effective deidentification program must be flexible enough to handle these variations and others that might emerge during the deidentification process. Given that 1 of the extracted files contained 2 million lines of text, however, it was not practical to explore a file of this size manually looking for all possible variations.

The realization that patterns in the text would not be sufficient to locate and remove enough direct identifiers led us back to UNCHCS IT to consider other ways to deidentify the extracted data. The group generated a dataset containing direct identifiers such as first name, last name, medical record number, address, phone number, SSN, etc. for each patient who had records in the extracted transcription file. This "key" dataset also contained a "fake medical record number" (fake_MRNo) that could serve as a linking field between the key dataset and the transcription file for each patient. Using the direct identifiers, we simplified the problem to replacing the occurrences of identifiers in the research dataset without having to locate them first. In other words, knowing the direct identifiers in advance eliminated the need to discover the identifiers using patterns in the text before replacing them.

Additional constraints further refined the requirements of the data-scrubbing program. Because the data extract was pulled directly from a live clinical information system, the extraction routines were required to run at night, fitting into the "mainframe" queue with other nightly jobs. External programs, such as the scrubbing program, were not permitted to run in UNCHCS' mainframe environment. In addition, to abide by HIPAA regulations and our data-use agreement, the data scrubbing had to be done at the IT site to prevent unauthorized disclosure of direct identifiers. The solution was to write the scrubbing program script in the Practical Extraction and Report Language (PERL) to run on a personal computer at the IT offices. The open-source licensing of PERL allowed a quick, easy installation at no direct cost, and its features provided the necessary file handling, regular expression, and data structure functionality.

The first step of the scrubbing program was to read all of the 12,910 records from the key dataset into a data structure for quick access. Each record contained a set of direct identifiers and a unique Fake_MRNo for a specific patient. Figure 1 shows a hash table of associative arrays that contain the first name (FName), last name (LName), medical (MRNo), and SSN represented by a pointer (ptr). The hash table stored these records in memory and provided very fast access to the individual records and to the direct identifiers in those records. The Fake_MRNo was used to link research dataset records to records in the key dataset. After the program was run, the output data set had the word "REMOVED" in place of the identifiable text.

Figure 1. Deidentification procedure using a hash table of associative arrays



EXAMPLE OF DE-IDENTIFIED TRANSCRIPTION FILE

00001116409 8/05/1998 UNC Hospitals PT REMOVED, REMOVED Chapel Hill, NC 27599 MR# REMOVED
DOS 08/05/98 TREATING PHYSICIAN XXXXX INTERNAL MEDICINE CLINIC - ESTABLISHED VISIT
HISTORY REASON FOR VISIT: Sinusitis/bronchitis. HISTORY OF PRESENT ILLNESS: I see Ms. REMOVED today on an added-on basis. She has a two-week history of a waxing and waning of her respiratory tract infection.

North Carolina Medicaid Data

In 2006, the NC Medicaid program provided care to more than 1.6 million individuals, including indigent children (36.7%), families with dependent children (29.2%), pregnant women (3.7%), the elderly (9.4%), blind persons (0.1%), and those otherwise disabled (16.1%).[19] About 61% of the NC Medicaid population in 2003 was female; 20% were age 0–4, 35% were age 5–20, and 32% were age 21–64. Medicaid eligibility in North Carolina is similar to other states for pregnant women and SSI disability, but eligibility based on financial need currently requires income of <45% of the Federal poverty level.[19] Despite the fact that UNCHCS is a public hospital, the populations served by UNCHCS and NC Medicaid may not be strictly comparable.

For this project, we had access to Medicaid inpatient, outpatient, and pharmacy data and a membership file that contained eligibility information and a linkage file to track those who lost and regained eligibility over time. The membership file has demographic data such as birth date, sex, and county code and information on enrollment periods (in months during the year). Inpatient data include the admitting and discharge dates and up to 9 ICD-9-CM diagnoses and 6 Current Procedural Terminology (CPT) codes per claim. The pharmacy data include the NDC code, the dispensing date, refill vs. new dispensing, the intended duration of the prescription (days' supply), and the quantity dispensed. We used NC Medicaid data from January 1, 2001 through December 31, 2004 to ensure that all claims would be complete and adjudicated for the research project.

Identification of Patients With Newly Diagnosed Type 2 Diabetes

This section describes the inclusion and exclusion criteria for identifying newly diagnosed DMT2 patients for both the WebCIS and NC Medicaid analyses. In the WebCIS section, we describe how we addressed several challenges using text data-mining approaches.

WebCIS

One of the most challenging tasks in using electronic databases of US patients (i.e., administrative claims or EMRs) is identifying when a chronic condition was first diagnosed. Each time a patient switches health plans or healthcare providers, which often occurs in the mobile US society, the longitudinal nature of their data is lost—no method exists for tracking care across insurers or healthcare systems. A further complication is referral patterns. As a tertiary care and academic medical center, UNCHCS often sees patients who have been referred to specialists from their local primary care physicians. Thus patients seen at UNCHCS may be seen only sporadically for evaluation and consultation.

These issues might have affected the ability to identify newly diagnosed patients with diabetes. We were particularly concerned about patients seeing only specialists at UNCHCS, with primary follow-up by their local practitioners. We knew we would not have complete medication and laboratory data on these patients, hindering the ability to identify when they were diagnosed with diabetes from the WebCIS files. Access to only 1 clinician's EMR data would not provide complete information on a patient, but it could be very valuable for research because it reflects actual care and should not be affected by payment source.

To address the potential limitations of the WebCIS data for conducting an observational study of diabetes, we developed inclusion and exclusion that differentiated patients seen regularly at UNC and followed by UNCHCS practitioners from those seen only episodically for emergency or referral care. Based on advice from our clinician researchers, we reasoned that patients who were seen at UNC several times a year and who underwent periodic HbA1c tests were likely to use UNCHCS for their primary care. In contrast, those with sporadic UNC visits and rare HbA1c tests were likely to be followed outside the UNCHCS system. Thus, a primary component of the inclusion criteria was the number of HbA1c tests patients had over time.

We obtained a data extract for all patients who had an HbA1c test on or after January 1, 2002 from UNCHCS. The original research plan was to identify patients who were newly diagnosed with diabetes as of January 1, 2003[b] and then compare their cardiovascular outcomes according to the initial antidiabetic medications they had received. However, given that many newly diagnosed patients are often treated by diet modification before receiving antidiabetic therapy, we were concerned that a 2-year follow-up (2003–2005) might not provide a sufficient number of adverse outcomes to observe longitudinal trends in antidiabetic medications or clinical care. Thus, we revised the initial research plan to define the base population as patients who had a first HbA1c test (as recorded in WebCIS) after January 1, 2001 and who had 2 or more HbA1c tests recorded in the laboratory file. We requested demographic, visit, medication,

[b] We used WebCIS data from January 1 to December 31, 2002 to ensure that we selected newly diagnosed patients.

electrocardiography, patient problems, vital status notes, and transcription files[c] for 12,424 patients with multiple HbA1c tests.

Based on an extensive review of the laboratory, patient visit, and patient problem files for these 12,424 patients, we developed the following exclusion and inclusion criteria for identifying newly diagnosed DMT2 patients who were regularly seen at UNC for their care:

Exclusion criteria

- Patients who had an ICD-9-CM diagnosis code of 250.xx in the patient problem file before January 1, 2001, and/or
- Medication prescribed for diabetes mellitus before January 1, 2001

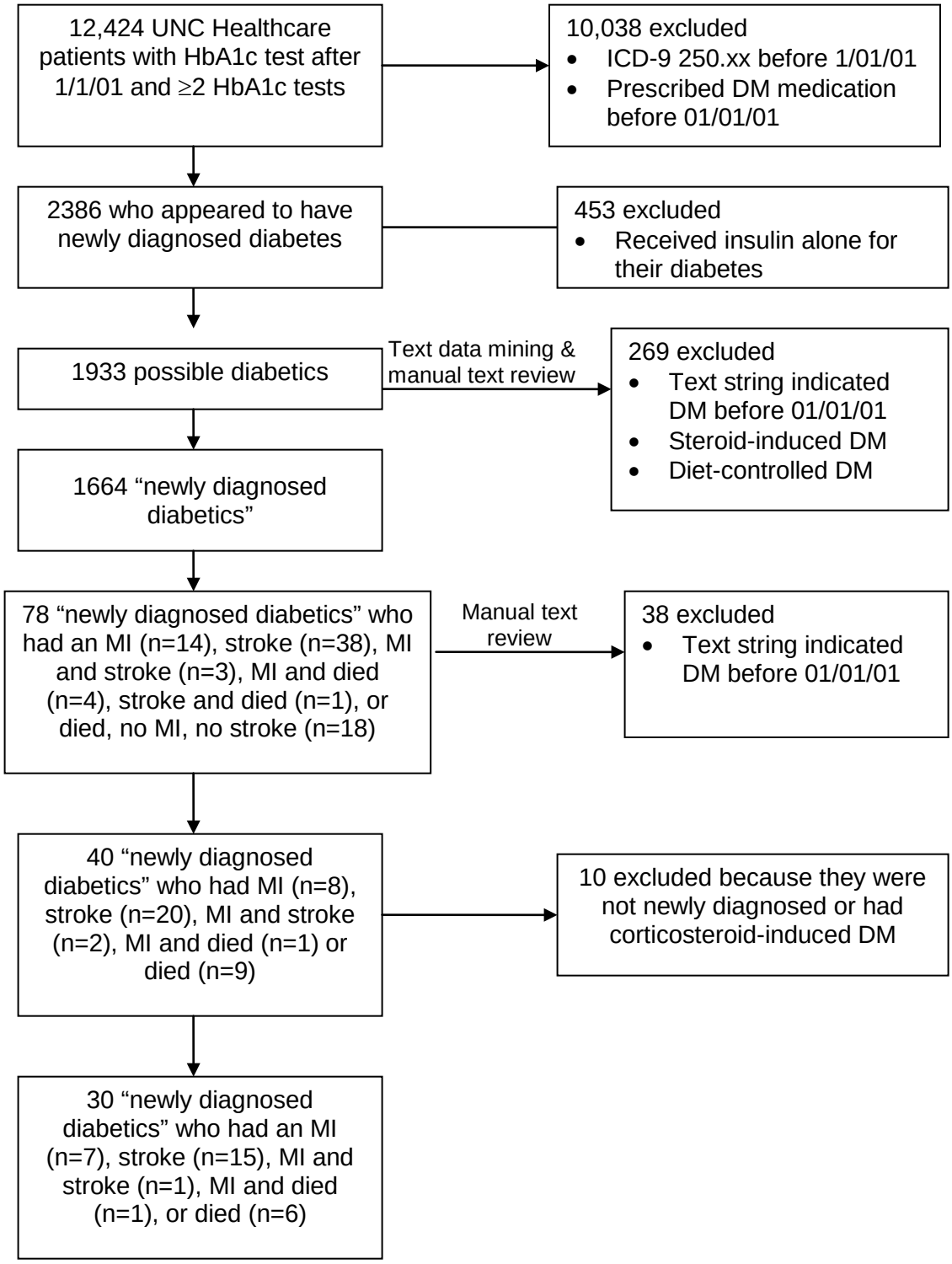
Inclusion criteria

- ≥2 HbA1c laboratory tests after January 1, 2001
- ≥2 outpatient visits in the 2-year period before the first HbA1c laboratory test at UNC
- Elevated HbA1c level after January 1, 2001
- Antidiabetic medication prescribed after January 1, 2001

Some of the inclusion criteria (≥2 HbA1c tests after January 1, 2001 and ≥2 outpatient visits) were developed to enhance capture of patients who sought care regularly from UNCHCS. Because the main purposes were to describe medication use in newly diagnosed diabetics and to evaluate long-term outcomes, we also required that patients be prescribed therapy for glycemic control after their first elevated HbA1c level. After applying these criteria, there were 2386 patients who appeared to have newly diagnosed diabetes. When we excluded patients who received insulin alone for their diabetes, 1933 diabetic patients remained. (We excluded patients who received insulin alone since we could not easily distinguish between patients with type 1 versus type 2 diabetes in this group and because we could not track dosing with any confidence.) Figure 2 provides an overview of the patient selection process.

[c] The transcription files were deidentified using the process described above, with the UNCHCS facility maintaining the WebCIS data before our receiving it for the project.

Figure 2. Patient identification process



We wished to further examine the ability to identify patients who were newly diagnosed after January 1, 2001 and who were not seen at UNCHCS as a referral or for emergency care. The goal was to develop an algorithm that would be highly specific for newly diagnosed diabetes. To that end, we decided to use text data-mining procedures on the transcription files to help with this determination. One investigator (CB) has been using text data mining to identify specific text in journal articles, as applied to systematic reviews and meta-analyses.

She applied her text-mining experience to the visit notes from the diabetic patients, providing a dataset that captured all occurrences of 3 text strings: diabetes (diabet*), insulin, and a diabetes medication (we provided her with a list of all the diabetes medications by brand and generic name), along with the patient anonymized ID number and transcription date. The text data mining procedure is described below.

After the text data-mining procedure was complete, we read the resulting text processed file into SAS and identified 677 patients who had a mention of any of the 3 text strings before January 1, 2001, nearly one-third of the newly diagnosed diabetics. As SAS could not be used for this determination, 5 of the investigators (SW, NM, MO, ZL, BS) reviewed the text strings from these 677 patients manually for indicators of prior diabetes or diabetes medications. These extensive validity checks were again part of the effort to develop a highly specific algorithm. We used the following criteria:

<u>Exclusion Criteria</u>

- Patients whose data indicated they had diabetes before January 1, 2001
- Patients with diabetes due to steroids
- Patients whose diabetes was diet-controlled

<u>Inclusion Criteria (these patients would remain in the cohort of newly diagnosed patients)</u>

- No mention of diabetes before January 1, 2001
- Transcription record stated that patient did not have diabetes
- Hyperinsulinemia with or without metformin treatment
- Borderline or questionable diabetes, i.e., firm diagnosis was not made
- Polycystic ovary syndrome
- Gestational diabetes

From this manual review, we excluded another 269 patients, leaving us with 1664 patients considered to have been newly diagnosed with diabetes after January 1, 2001. Note that we did not manually screen all of the data for these 1664 patients, as we had budgeted only a small amount of project time for this investigator (CB).

Text-Mining Procedure

The transcription data consisted of unformatted clinical records with personally identifying information removed. The format of these documents was inconsistent due to irregular spacing, punctuation, and division of records based on a fixed character limit. The text processing consisted of 4 runs. Each run was reviewed and the design revised for future runs. Text processing was done using GNU Awk, a UNIX scripting utility well suited for basic text processing. The platform used was Sun Solaris 10 on a Sparc Blade 1500 workstation.

The first run considered the full dataset for only 12,222 patients, as there were no transcription visit notes for the remaining 204 patients in the original diabetes cohort that had ≥2 HbA1c tests after January 1, 2001. Computer Program A used 3 kinds of triggers for extraction. The first trigger was a set of 8 drugs often used to treat diabetes: Humalog®, Humulin® Iletin®, Lantus®, Novolin®, Novolog®, Velosulin®, and Relion®. The second trigger was the word "insulin," and the third trigger was the string "diabet." The latter was used to ensure the detection of both "diabetes" and "diabetic." The output of this program comprised the patient ID, date of treatment, start and end character numbers, and the 50 characters before and after the matching term (see **Appendix E** for details).

Computer Program B was similar to Program A, but instead of returning a fixed number of characters surrounding the matching words, it returned a sentence. To achieve this, a sentence tokenizer was created that initially used punctuation and a list of exceptions, such as decimal point in numbers, period in some multiword abbreviations, etc., which was then replaced with a regular expression. The regular expression enabled the use of wildcards for letters, e.g., a period followed by any character or any digit. Records were sent first to the sentence tokenizer program and then to the extraction program used in the first run. The output of the second computer program was the patient ID, date of treatment, sentence ID, sentence, and 3 binary fields that indicated which trigger caused the record to be extracted, i.e., a drug name, "insulin," or "diabet."

Similar to Program B, Program C triggered output from a set of drug names, but its output was the actual drug and dosage indicated in the note rather than an entire sentence (see **Appendix E** for details).

In contrast to Programs A, B, and C, which all used drug names as triggers, Program D used numbers and Système International (SI) units to trigger the extraction of drug and dosage information from a note. The rationale for this change was 2-fold: 1) to ensure that drugs missing from the list were extracted, and 2) to ensure that this approach could be applied to conditions other than diabetes. Program D also identified the method of drug delivery when stated in the note. The output of Program D was the patient ID, visit date (separated by month, day, and year), and the drug name, dose, unit, delivery method, and frequency. Program D was tested using the standard metrics of *precision*, the number of correctly identified items (where an item was the drug name, amount, or method of delivery) divided by the number of extracted items; and *recall*, the number of correctly identified items divided by the number of items in the note. The training and test sets comprised 30 and 40 notes, respectively, each selected at random from the 1933 patients at high risk. The gold standard—i.e., the correct number of items—was identified manually from the each set of notes. Preliminary findings from these experiments were recently presented.[20]

Assessing the Completeness of the Medication Prescribing File (TFDBPAT.txt)

As discussed above, the completeness of the medication prescribing file was of concern given that the project's focus was to assess the use of antidiabetic medications for glycemic control. We therefore conducted an agreement study to compare the Medication text data with the structured TFDBPAT.txt data. The medication text file had several potential problems, including (1) physicians might not have entered all of the patient's medications into the visit notes, (2) the text-mining trigger (a number and an SI unit) may have missed medications that were dictated with the drug name alone, not the number and SI, and (3) misattribution of medications prescribed with medications discontinued. Visual review indicated that the

Medication text file contained many more drugs than appeared in the TFDBPAT.txt file, but we could not assess how many drug names and dates should have been in either data source, i.e., there was no gold standard for assessing sensitivity and specificity.

Our first agreement study focused on oral medications used for treating the 78 "newly diagnosed" DMT2 patients, i.e., those we thought were newly diagnosed before extensive reading of the visit notes. We assessed agreement for antidiabetic medications prescribed after 2002 and calculated agreement for the generic names of the sulfonylureas, biguanides, thiazolidinediones, and combinations of these drugs as well as for therapeutic class more generally (Agreement Study 1). We sought agreement for the name, entered date (when the drug was prescribed), and the window between the entered and the inactivation dates, when an inactivation date was available. Agreement was coded as 1 for "yes" and 0 for "no" to indicate whether the generic drug name/therapeutic class appeared in both the TFDBPAT.txt and Medication text files. We coded a match if there was agreement on either the brand or the generic name of the drug. To assess agreement for the WebCIS entered dates (and accompanying medication name), we coded whether there was a match for the visit date in the Medication text file (1 for "yes," 0 for "no") either exactly or within ± 7, 14, and 30 days. In addition, we evaluated whether dates from the Medication text data overlapped the entered and inactive dates in WebCIS, when the inactivation date was available.

We conducted a parallel agreement study (Agreement Study 2), in which we evaluated whether WebCIS's TFDBPAT.txt file could substantiate listings from the Medication text file. For this analysis, we considered each visit date from the Medication text file and sought agreement with the WebCIS time windows (date entered minus date inactivated) from the TFDBPAT.txt file.

Our agreement studies may be best illustrated with an example (see Tables 4 and 5 below for sample data). For Agreement Study 1, we see that Glucotrol appears in both the WebCIS and Medication text files, so the agreement for name is 100%. For the entered date in WebCIS, we see that there is an exact match for Glucotrol on 2/13/2004 and 2/24/2004. However, there is no exact match for the inactive date for the second Glucotrol prescription, 8/31/2005. The closest Medication text file date is 8/26/2005. The variable for an exact match would be coded as 0 and the variable for gap ± 7 days would be 1.

Table 4. Sample Web CIS text file

PATIENT ID	DRUG NAME	ENTERED DATE	INACTIVE DATE
W10	GLUCOTROL XL 5MG TABLET SA	2/13/2004	2/24/2004
W10	GLUCOTROL XL 5MG TABLET SA	2/24/2004	8/31/2005

Table 5. Sample medication text file

PATIENT ID	DRUG NAME	VISIT DATE
W10	GLUCOTROL	2/13/2004
W10	GLUCOTROL	2/24/2004
W10	GLUCOTROL	8/3/2004
W10	GLUCOTROL	1/24/2005
W10	GLUCOTROL	5/19/2005
W10	GLUCOTROL	8/26/2005

We also evaluated whether dates from the Medication text data overlapped the entered and inactive dates in WebCIS, when inactivation dates were available. For this example, the variable for coverage of the first WebCIS Glucotrol entry would be coded as 1 to indicate that there were Medication text file records that spanned the start and inactive dates of the WebCIS entry (from 2/13/2004 to 2/24/2004). The coverage variable would be coded 1 for the second WebCIS drug entry as well. Note that there were only 12 days of coverage for the first WebCIS drug entry and only 2 visit dates in the Medication text file that spanned this coverage. For the second WebCIS data entry, there were 552 days of coverage but 3 visit dates besides the entered and inactive dates—5 visit dates in total that fell within the coverage window. From this information, we calculated the following:

$$\text{Days coverage per visit (DCV)} = \frac{\text{Total number of days between WebCIS's entered and inactive dates for a particular drug}}{\text{Total number of visit dates from the Medication text file that fall within the WebCIS start and inactive date window for that drug}}$$

For the first WebCIS Glucotrol entry, the number of days between the entered and inactive dates was 12 and there were 2 patient visits, thus the DCV was 6, whereas the DCV for the second WebCIS Glucotrol entry was 110.4.

Using the above example to describe Agreement Study 2, we evaluated whether the 6 Glucotrol entries in the Medication text file would coincide with the 2 WebCIS data entries; we referred to this as coverage. Because all 6 dates are within the 2 WebCIS intervals, they would be classified as an exact match on coverage with the WebCIS TFDBPAT.txt file.

By looking at agreement in both directions, i.e., the overlap in dates between the 2 files, we could ascertain the potential completeness of each file for research purposes.

NC Medicaid

The extract of NC Medicaid data was done for diabetic patients aged 18 or older as of January 1, 2001 who had been eligible at any point from January 1, 2001 until December 31, 2004. The files excluded refugees, aliens, children on the State Children's Health Insurance Program (Health Choice), and those who were eligible for both Medicare and Medicaid. We identified 5624 patients who had a diagnosis of diabetes on or after January 1, 2002 who also 1) were eligible for coverage in 2001, 2) had no outpatient claims with a diabetes ICD-9-CM code before January 1, 2001, and 3) had no pharmacy claims for an antidiabetic drug before January 1, 2001. Because clinicians might have entered a diabetes diagnosis code so that glucose tests would be reimbursed, we also required that patients have >1 diabetes ICD-9-CM code. Thus we might have included patients who were not newly diagnosed, but we could not requisition medical records to confirm the diabetes diagnosis. Finally, we further restricted the NC Medicaid population to those who were also treated with an antidiabetic drug after January 1, 2001 (n=2794).

Identification of Deaths, Myocardial Infarction, and Stroke Among Diabetic Patients

Stroke and MI are serious macrovascular complications associated with diabetes and comorbid hypertension and dyslipidemia. We used these medium-term (months to a few years) cardiovascular outcomes to assess initial antidiabetic medications used by newly diagnosed DMT2 patients.

WebCIS

For the 1664 WebCIS patients who appeared to be newly diagnosed with diabetes, we used the TPRBPAT.txt file and the codes in **Appendix F** to determine whether: (1) any of them died (based on a WebCIS field, not death certificates), (2) developed an MI, or (3) had a stroke after their first elevated HbA1c level in WebCIS. In all, 78 patients had 1 or more of these events: 14 had an MI only, 38 had a stroke only, 18 died without having another complication, 4 had an MI and died; 1 had a stroke and died; and 3 had an MI and stroke (Figure 2).

For these 78 diabetics, we developed an Excel spreadsheet that incorporated information on:

- HbA1c test results
- Glucose tests: random and fasting
- Lipid levels: LDL, high-density lipoprotein (HDL), and total cholesterol (TC); triglycerides
- Blood pressure readings from ambulatory visits
- Body mass index (BMI)
- Visit dates
- Problems noted at each visit
- Medications from both the WebCIS TFDBPAT.txt file and the Medication text file: glycemic control agents, insulin, lipid-lowering drugs, antihypertensive drugs (see **Appendix G**).

Before 2 clinical investigators (TC, MO) reviewed the data in the Excel spreadsheet, the PI reviewed the problem file (TPRBPAT.txt) for additional information indicating that these 78 patients had not, in fact, been newly diagnosed with DMT2 or that their diabetes was corticosteroid-induced. In all, 38 patients were excluded after this review, leaving 40 patients for analysis. These 40 patients had the following events: 8 had MI only, 20 had stroke only, 9 died without another complication, 1 had an MI and died, none had a stroke and died, and 2 had an MI and stroke. To ensure that the clinicians analyzed only truly eligible patients, the PI and the project manager conducted a manual review of the text data-mining file to carefully screen for the timing of the DM diagnosis. Based on this additional review for the final WebCIS analysis, 10 additional patients were excluded because they were not newly diagnosed with DMT2. Thus the final analysis population included 30 patients: 6 who had MI only, 15 who had stroke only, 6 who died without another complication, 1 who had an MI and died, none who had a stroke and died, and 2 who had an MI and stroke.

The clinical investigators reviewed the data on these 30 patients to detect patterns suggesting worsening health in the 12 months preceding the MI, stroke, or death and reported their findings on the abstraction form (see **Appendix H**). The clinicians used evidence-based

clinical guidelines from 3 organizations[21-23] to assess the patients' clinical profiles. In particular, they focused on whether the DMT2 patients received medications that adequately controlled their diabetes and comorbid HT, DL, and hypertriglyceridemia (TG), if present. The clinicians based their opinions of disease control (DM, HT, DL, and TG) on clinical measurements such as HbA1c levels, blood pressures taken at each visit, and cholesterol tests. We had the information from the abstraction forms keyed with 100% verification. The WebCIS results we present below reflect analyses of the abstraction forms.

We assessed the factors listed above (HT, DL medications, and visits) in the 30 newly diagnosed DMT2 patients who had an MI, had a stroke, and/or died. We could not conduct a parallel analysis of the WebCIS DMT2 population who did not have these outcomes. To do this, we would have needed to perform validation reviews of ~1600 records, which was not possible given the current scope of the project.

NC Medicaid

We determined whether the 2794 patients had an MI and/or stroke. We did not evaluate death because this is not consistently captured in administrative claims files. Of the 2794 patients, 49 had an MI, 173 had a stroke, and 14 had both an MI and a stroke after the date of their first DM diagnosis reported. This left 2423 patients who did not have one of these outcomes.

We used the patterns discerned by the WebCIS clinical review to inform the Medicaid analyses. In particular, we determined which indicators of diabetes control could be detected only by reviewing clinical data such as laboratory test results and vital status measures (that were available in the WebCIS files) versus those that could be identified using administrative claims data such as available from NC Medicaid. The latter included:

- Whether patients had HT or DL identified before the first DMT2 indicator (diagnosis or medication)
- Medications taken by patients in the 12 months after DMT2 was identified
- For those who also had HT, whether they received an ACEI or ARBs within 12 months after DMT2 diagnosis
- Medications taken by patients in the 12 months before an outcome (MI, stroke, death)
- Number of patient visits in the 12 months before the outcome event

Based on this dichotomy, we developed the analysis plan for the NC Medicaid data to reflect these indicators.

Results

This section provides the findings from the literature scan and analyses of the WebCIS and NC Medicaid datasets addressing medication use in newly diagnosed DMT2 patients.

Literature Scan

Key Question 1: What are the primary DMT2 and cardiovascular outcomes and laboratory tests evaluated in comparative studies of antidiabetic therapy in newly diagnosed patients?

Eleven studies addressed this question. Six were observational studies, of which 4 used administrative claims data from Saskatchewan Health,[24-27] 1 used data from primary care practitioners in the UK,[28] and 1 was a reanalysis of data from 2 clinical trials.[29] The remainder were randomized controlled trials.[30-34]

The 4 observational studies using Saskatchewan Health data covered ~5 years of follow-up. The studies compared new users of diabetes medications who were dispensed sulfonylurea alone, metformin alone, or a combination of sulfonylurea and metformin. The first[26] compared the effects of therapy with sulfonylurea, metformin, or their combination on cardiovascular and all-cause mortality. In this study, metformin therapy and the combination regimen were superior to sulfonylurea alone. The adjusted odds ratio (OR) for cardiovascular-related mortality was 0.64 (95% CI, 0.49–0.84) for metformin and 0.64 (0.54–0.77) for combination therapy compared with sulfonylurea alone. The corresponding results for all-cause mortality were 0.60 (0.49–0.74) and 0.66 (0.58–0.75), respectively.

Eurich and colleagues conducted a similar study[25] but restricted their population to those who had a diagnosis of or hospitalization for heart failure. Metformin and combined metformin and sulfonylurea therapies were associated with lower hazard ratios (HR) for all-cause mortality at the end of the follow-up period (mean, 2.5 years; median, 2.1 years) compared with sulfonylurea monotherapy, in both the crude analysis (metformin HR, 0.63 [95% CI, 0.49–0.82]; combination therapy HR, 0.50 [0.43–0.58]) and after adjustment for age, sex, chronic disease score, and medications used to treat heart failure (metformin: HR, 0.70 [0.54–0.91]; combination therapy: HR, 0.61 [0.52–0.72]). For all-cause hospitalizations both at 1 year and at the end of follow-up, the crude analyses showed a lowered risk for metformin and combination therapies at 1 year [metformin: HR, 0.52 (0.35–0.76); combination therapy: HR, 0.41 (0.32–0.52)], but the adjusted estimates showed no significant difference between the therapies at 1 year [metformin: HR, 0.84 (0.67–1.04); combination therapy: HR, 0.92 (0.80–1.06)]; and at the end of follow-up (metformin: HR, 0.87 (0.73–1.05); combination therapy: HR, 0.93 (0.83-1.05)].

Johnson and colleagues conducted a second study in the Saskatchewan population looking at cardiovascular mortality, nonfatal cardiovascular hospitalizations, and the composite of these outcomes.[24] Metformin alone was associated with lower risk of cardiovascular morbidity or mortality compared with sulfonylurea in both the crude and adjusted analyses [HR, 0.67 (0.56–0.80) and HR, 0.81 (0.68–0.97), respectively]. Metformin was also superior to sulfonylurea for the endpoint of nonfatal cardiovascular hospitalizations in both crude and adjusted analyses [HR, 0.76 (0.61–0.93) and HR, 0.78 (0.63–0.97), respectively].

This same population was studied by Simpson and colleagues,[27] focusing on mortality from ischemic effects of poor adherence or underdosing of first-generation sulfonylureas, glyburide, and metformin. Poor adherence appeared to have a more negative effect on mortality among those taking first-generation sulfonylureas and glyburide versus metformin [HR for sulfonylurea, 2.20 (unadjusted), 2.45 (adjusted); HR for glyburide, 1.55 (unadjusted), 1.33 (adjusted), HR for metformin 1.10 (unadjusted), 0.98 (adjusted)]. It cannot be determined from these observational studies whether physician or patient factors (mild hepatic or renal dysfunction, BMI, or HbA1c level) might explain why patients were prescribed particular initial therapies for diabetes.

There were 2 additional observational studies, 1 that used EMR[28] and 1 that was a reanalysis of 4 randomized controlled studies.[29] Lusignan and colleagues used data from 142 general practices in the UK to identify patients with newly diagnosed diabetes in 1994. They collected information about the type of diabetes therapy and the proportion of patients meeting

HbA1c, BMI, blood pressure, and TC goals. They then reassessed these variables on the same patients in 2001. From 1994 to 2001, the use of short-acting sulfonylureas increased by 12%, and that of metformin, by 15%.[28] The analyses did not stratify by specific medication type, but HbA1c levels were not as well controlled in 2001 compared with 1997: 28.9% of diabetics had an HbA1c level <6.5% in 1997, but only 22.5% had such a level in 2001. This trend persisted after adjustment for age, sex, and practice. Blood pressures improved from 1994 to 2001, but only 22.5% of diabetics reached the target blood pressure of <140/80 mm Hg in 2001.

Rajagopalan et al.[29] reanalyzed data from 4 randomized controlled trials of pioglitazone versus other glycemic control medications (metformin, gliclazide) to determine the prevalence of metabolic syndrome at baseline and at Week 52. Metabolic syndrome wais defined as a combination of symptoms including impaired glucose tolerance, insulin resistance, or diabetes; BMI >30 kg/m^2; elevated blood pressure; microalbuminuria; and lipid abnormalities. The prevalence of metabolic syndrome decreased by 9.2% (95% CI, 6.5%–12%) in the pioglitazone group, 7.7% (3.5%–11.9%) in the metformin group, and 4.3% (0.4%–8.3%) in the gliclazide group, with the pioglitazone group showing greater improvements in HDL and triglyceride levels than the other 2 groups.

Four randomized, controlled trials of varying lengths—3 months,[31] 6 months,[30, 34] and 12 months[32, 33]—evaluated HbA1c levels and vital signs over time (Table 6). In all 4 studies, although HbA1c level decreased over the treatment period, there was little or no change in systolic or diastolic blood pressures.

Table 6. Randomized, controlled studies evaluating the effects of antidiabetic agents on HbA1c levels and blood pressure

	At Baseline	At End of Study
Nakamura (3 months)[31]	HbA1c Level P 7.7 ± 1.2 G 7.8 ± 1.1 V 7.6 ± 1.1 SBP (mm Hg) P 122 ± 17 G 122 ± 18 V 118 ± 16 DBP (mm Hg) P 74 ± 14 G 78 ± 14 V 78 ± 12	HbA1c Level P 6.8 ± 1.1* G 6.9 ± 1.2* V 6.8 ± 1.1* **p<0.05 vs. baseline* SBP (mm Hg) P 116 ± 15 G 124 ± 16 V 122 ± 18 DBP (mm Hg) P 72 ± 12 G 79 ± 12 V 80 ± 14
Hallsten (6 months)[30]	HbA1c Level (mean ± SE) R 6.8 ± 0.2 M 6.9 ± 0.2 P 6.3 ± 0.1 SBP (mm Hg) R 152 ± 5.0 M 145 ± 4.1 P 147.2 ± 3.2 DBP (mm Hg) R 90.5 ± 2.1 M 91.4 ± 2.5 P 85.1 ± 2.3	HbA1c Level (mean ± SE) R 6.5 ± 0.2* M 6.2 ± 0.2* P 6.1 ± 0.1 **p<0.05 vs. baseline; p=ns for R vs. M* SBP (mm Hg) R 149 ± 4.5 M 141.8 ± 4.0 P 144.4 ± 3.8 DBP (mm Hg) R 84.2 ± 2.4 M 85.5 ± 2.6 P 85.4 ± 2.7
Watanabe (6 months)[34]	HbA1c Level P 6.9 ± 0.2	HbA1c Level P 6.1 ± 0.33*

	G 7.2 ± 0.5	G 6.3 ± 0.4* **p<0.01 vs. baseline* Change in SBP (mm Hg) P ↓ 3 G ↓ 7.4, *p=ns* Change in DBP (mm Hg) P ↓ 11.6 G ↓ 0.9, *p=ns*
Tan (12 months)[32, 33]	HbA1c Level P 8.54 ± 0.9 G 8.45 ± 1 SBP (mm Hg): P 128.4 ± 14.6 G 127.8 ± 17.8 DBP (mm Hg) P 81.6 ± 9.6 G 80.4 ± 10.2	Mean Change in HbA1c Level at 52 Weeks P ↓ 0.78, *p<0.001* G ↓ 0.68, *p<0.001 (p=ns for P vs. G)* Mean Change in SBP at 52 Weeks P ↓ 3.5 G ↓ 1.4, *p=ns* Mean Change in DBP at 52 Weeks P ↓ 3.9 G ↓ 1.3, *p=0.028 in favor of P*
DBP = diastolic blood pressure; G = glimepiride (Tan[32, 33]) or glibenclamide (all others); M = metformin; P = pioglitazone; R = rosiglitazone; SBP = systolic blood pressure; V = voglibose.		

Summarizing the evidence for Key Question 1 suggests differences between glycemic control medications and their effect on cardiovascular risk factors. In particular, the Saskatchewan Health studies showed that metformin was associated with improved cardiovascular profiles compared with sulfonylurea. However, the Saskatchewan studies did not evaluate the effects of thiazolidinediones on cardiovascular morbidity and mortality and could not control for patient or physician factors in choice of therapy due to their observational study design. The randomized controlled studies were primarily focused on pioglitazone compared with other medications and their effect on surrogate markers of cardiovascular disease, such as blood pressure and BMI. However, all the pioglitazone studies were funded by the pharmaceutical industry so one cannot rule out publication bias as a possible explanation for these findings.[35] Understanding the differences among the sulfonylureas, biguanides, and thiazolidinediones for long-term treatment of DMT2 would be valuable for future research.

Key Question 2: Is there a difference in DMT2 outcomes such as HbA1c level, fasting plasma glucose level, or switching/augmentation of therapy depending on the initial medications DMT2 patients use?

There were 18 head-to-head trials,[14, 30-34, 36-47] of which 8 were open-label studies.[14, 31, 34, 36, 38, 39, 44, 46] Six of these 8 studies included a sulfonylurea treatment group,[14, 31, 34, 35, 44, 46] 5 had a pioglitazone arm,[14, 31, 34, 39, 44] and 5 had a metformin arm.[14, 36, 38, 44, 46] Repaglinide[38] was included in 1 study, and acarbose[40] in another. All of the open-label studies were conducted outside of the US: UK (n=1), Italy (n=2), Germany (n=1), Japan (n=3), and India (n=1). The length of the open-label trials varied from 3 months (n=2) to 6 months (n=2) to 1 year (n=4), and the number of subjects ranged from 30 to 265. Outcomes evaluated in nearly all 8 studies included HbA1c level, fasting plasma glucose (FPG), lipid parameters (TC, LDL, HDL, TG), blood pressure (BP), weight or BMI, and side effects. Some studies also included measures of insulin sensitivity and measures of cardiovascular risk such as lipoprotein (a), apolipoprotein A-I, apolipoprotein B, and plasminogen activator inhibitor.

All 8 open-label studies assessing HbA1c level and FPG showed decreases in these 2 variables between baseline and the end of follow-up, but only 2 studies showed a difference by treatment. Campbell and colleagues found that the HbA1c level was reduced by 25% at 52 weeks with metformin therapy, whereas glipizide therapy was associated with only a 17% reduction.[36] Pioglitazone showed a significant advantage over acarbose for HbA1c level in a 26-week study, reducing the mean HbA1c level from 8.98±1.20% to 7.82±1.95% compared with a reduction from 9.03±1.32% to 8.55±1.96% for acarbose ($p<0.001$).[40]

Three randomized, double-blind, noninferiority studies evaluated pioglitazone versus gliclazide (n=1270)[37] or versus metformin (n=205[43] and n=1199[45]). All 3 trials were conducted outside of the US. Two of the trials lasted 52 weeks,[37, 45] and the third was 32 weeks in duration.[43] Outcomes evaluated included: HbA1c level, FPG, lipid parameters, BP, body weight, liver enzymes, side effects, and tolerability. The investigators also measured fasting serum or plasma insulin levels and indicators of insulin sensitivity (HOMA-%S, QUICKI), proinsulin levels, B-cell function (HOMA-%B), and C-peptide. Reductions in HbA1c levels were similar among treatment groups at the study endpoints (declared noninferior). There was some variation in the effect on measures of insulin sensitivity and lipid parameters.

One trial included a 1-year extension of the follow-up period involving 98 of the original 206 study centers.[47] At 2 years, pioglitazone was superior to gliclazide (a sulfonylurea not available in the US) at maintaining HbA1c below either 7% (111/261, 42.5% for pioglitazone vs. 81/289, 28% for gliclazide; $p<0.001$) or below <8% (129/270, 47.8% versus 110/297, 37%, respectively, p=ns). FPG and measures of insulin sensitivity were also better in the pioglitazone group at 2 years (difference in FPG for pioglitazone minus gliclazide, –0.83 ± 0.22 mmol/L [95% CI, –1.26 to –0.39]); FSI, HOMA-%S, and HOMA-%B also showed improvement. However, the dropout/withdrawal rate in this trial was around 50% in both groups.

Six trials were randomized, double-blind comparisons.[30, 32, 33, 41, 42, 45, 48] Two studies compared pioglitazone and a sulfonylurea.[32, 33] The others compared metformin with rosiglitazone,[30] metformin with nateglinide,[41] glyburide with miglitol,[42] metformin with pioglitazone,[45] and glimepiride with repaglinide.[39, 48] The number of subjects in these trials ranged from 45 to 312. Three of the trials lasted 6 months; the remainder lasted 12 months. Only 1 trial included any US patients. The trials evaluated outcomes similar to those described for the noninferiority trials. None found a significant difference in control of HbA1c level at the study endpoint.

A small, prospective, observational study[49] used a database of general practitioners in Tayside, Scotland. The investigators evaluated prescribing patterns and HbA1c-level response to treatment in patients receiving a sulfonylurea versus metformin monotherapy. A subgroup analysis examined the effect of BMI on treatment response to sulfonylureas or metformin and found none.

The only study to evaluate secondary failure of antidiabetic monotherapy used the Saskatchewan Health data.[50] The investigators followed patients who had received monotherapy with a sulfonylurea or metformin for ≥2 years. After another 2 years of follow-up (4 years from start of therapy), the proportion of patients who reached secondary failure (adding or changing a drug) was 46.8% for the sulfonylurea group and 38% for the metformin group. Metformin monotherapy was associated with a delay in the onset of secondary failure: (unadjusted HR, 0.93 [95% CI, 0.85–1.02]; adjusted HR, 0.89 [0.82–0.98]), where adjustment was for age, sex, adherence, and chronic disease score. The adjusted Kaplan-Meier curves for time to secondary failure became significantly different at Year 4 (p=0.035) and continued to diverge throughout

follow-up. More diabetic patients who first received a sulfonylurea rather than metformin needed to begin taking combination therapy (39.8% vs. 29.6%, respectively; adjusted HR, 0.79 [0.71–0.87]) or insulin (9.1% vs. 5%, respectively; adjusted HR, 0.65 [0.51–0.82]).

The results of this literature search for both key questions indicated that few published studies of US patients had evaluated how well antidiabetic medications control diabetes in newly diagnosed DMT2 patients. In particular, few studies had compared the effects of metformin versus sulfonylureas on short-term indicators of diabetes control such as HbA1c level, FPG level, or the need for medication augmentation or change.[14, 36, 44, 46, 47]

We did not include the UKPDS in the evidence tables because subjects were randomized to receive metformin only if they were obese. The other comparator groups were sulfonylurea and insulin, which did not meet the inclusion criteria for study selection (insulin is not an oral therapy). Of the studies reviewed, investigators typically evaluated the following clinical outcomes: HbA1c level, FPG level, body weight, lipid profile, and blood pressure.

Results from the Electronic Medical Record (WebCIS) and Administrative Claims Analyses (NC Medicaid)

This section provides findings from the analyses of the WebCIS and NC Medicaid databases. For the WebCIS data, we provide the results for the informatics and clinical aims. For the informatics aim, we present 2 analyses: (1) results of text data mining with regard to precision and recall for both the training and test datasets and (2) results from the agreement studies. For the clinical aim—to assess medication use in newly diagnosed DMT2 patients—we provide results from the 30 newly diagnosed DMT2 patients with MI, stroke, and/or death. In particular, we assessed the presence of HT or DL at the time of DMT2 diagnosis and, if so, the treatment of these conditions. We also evaluated how well the medications controlled the patient's DMT2, HT, and/or DL. Finally, we used the clinical information and test results from WebCIS for HbA1c level; systolic and diastolic blood pressures; and HDL, LDL, and TC levels to determine whether the patients were being adequately treated for these conditions in the 12 months before the MI, stroke, and/or death.

We conducted a parallel analysis using the NC Medicaid dataset, with two differences: (1) we compared patients who did and did not have a cardiovascular outcome (MI, stroke, or both) and (2) we could not evaluate the control of DM, HT, and/or DL because clinical values or test results were not available in the administrative claims data. As with the WebCIS analyses, we assessed whether patients had HT or DL at the time they were first diagnosed with DMT2 and whether they were treated for these conditions. For the Medicaid patients who had an MI, a stroke, or both, we calculated the average number of outpatient visits they had in the 12 months before the outcome and whether they were taking medications to treat DM, HT, and/or DL if they had these conditions.

All of these analyses are descriptive, given the very small sample sizes.

WebCIS Analyses

In this section, we describe the results of the text-mining procedures, the agreement studies evaluating the completeness of the TFDBPAT.txt file, and the clinical findings.

Text-Mining Results for Computer Program D

Program D performed well on drug extraction for the training and test datasets (Table 7). An informal analysis showed that drugs paired with dosages were usually associated with current prescriptions rather than drugs used previously. **Appendix D** contains a more in-depth description of the text-mining results.

Table 7. Detailed training and test dataset results

	Training (n=30)		Test (n=40)	
	Precision	**Recall**	**Precision**	**Recall**
Drug	86.96%	82.76%	86.81%	82.78%
Amount	100.00%	91.39%	100.00%	96.08%
Method of delivery	100.00%	62.03%	100.00%	91.89%
Prescribed frequency	100.00%	41.73%	100.00%	48.15%
Average	**96.74%**	**69.48%**	**96.70%**	**79.72%**

Precision for drug amount was 100% for both training and test datasets, because the method uses an explicit pattern match on the unit of measurement and its quantifier. Recall addresses how well Program D correctly identified the item of interest (drug, amount, etc.) divided by the total number of the specific items in the note; this provides a measure of the utility of the program. Program D was able to correctly identify >80% of the drugs and 96% of the drug amounts listed in the visit note. Recall of prescribed frequency of medication use was low (41.73%/48.15%); however, this is on par with other text-mining methods. Overall, the recall results reflect variations in visit notes report frequency.

Agreement Study Results

This section describes findings on the agreement between the Medication text-mining file and the structured TFDBPAT.txt prescribing drug file for medication name, start date, and coverage date for each WebCIS medication record.

Agreement Study 1: Evaluation of Drug Name. There were 335 WebCIS diabetes medication prescription records for the 78 patients considered "newly diagnosed" who had an MI or stroke or who died. On average, each patient had 4.3 diabetes drug mentions in WebCIS. Table 8 provides results comparing the 335 specific drug names from WebCIS, by therapeutic class and generic name, with the Medication text file for each of the 78 patients.

The worst agreement (i.e., the fewest matches between the Medication text file and the WebCIS records) was for the sulfonylureas (85.7%), mainly because of poor agreement for glipizide (79.6%). Comparing across the generic drugs, for which there were substantial numbers of exposures to evaluate, agreement was >90% for glimepiride, glyburide, metformin, pioglitazone, and rosiglitazone.

Table 8. Agreement between WebCIS and medication text files on drug name

Drug Name	Total N	Number in Agreement	Percent Agreement
Overall	335	292	87.2
Therapeutic class*			
Biguanides	127	123	96.9
Sulfonylurea	140	120	85.7
Thiazolidinediones	63	61	96.8
Generic name			
Glimepiride	31	31	100.0

Drug Name	Total N	Number in Agreement	Percent Agreement
Glipizide	98	78	79.6
Glyburide	11	11	100.0
Metformin	127	123	96.9
Metformin & glyburide	4	2	50.0
Pioglitazone	26	26	100.0
Rosiglitazone	37	35	94.6
Rosiglitazone & metformin	1	1	100.0
* Excludes the combination antidiabetic agents (n=5).			

Agreement Study 1: Evaluation of Entered Dates. For assessing the agreement for the date the diabetes medications were prescribed, we compared the entered date from the WebCIS TFDBPAT.txt file with the patient visit date from the Medication text file. For this analysis, we focused on the WebCIS records for which there were both entered and inactive dates. We asked , "For the WebCIS drug listed and its corresponding entered date, was the same drug listed (by generic name) in the Medication text file with that same visit date?" We evaluated dates as an exact match as well as a match ±7, ±14, and ±30 days of the entered dates overall (Table 9), by therapeutic class (Table 10), and by generic name (Table 11).

There was a match within ± 30 days for about 50% of the WebCIS-entered dates based on the visit dates from the Medication text file (Table 9). Agreement was lowest for the sulfonylureas (42.9%), particularly, for glipizide (30.9%) (Table 10). Of the 3 types of sulfonylureas, the third-generation drug (glimepiride) had the best agreement with the Medication text file (Table 11).

Table 9. Agreement between WebCIS and medication text files on entered date

	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Entered date matches (N=239)	69	28.9	69	28.9
Additional matched on ± 7 days	15	6.3	84	35.1
Additional matched on ± 14 days	16	6.7	100	41.8
Additional matched on ± 30 days	19	7.9	119	49.8
Not matched within 30 days	120	50.2	239	100.0

Table 10. Agreement between WebCIS and medication text files on entered date, by therapeutic class

Entered date matches by therapeutic class (N=236*)	Total	Matched ± 30 days	Percent matched ± 30 days
Biguanides	89	48	53.9
Sulfonylurea	98	42	42.9
Thiazolidinediones	49	30	61.2
* Excludes the combination antidiabetic agents (n=3).			

Table 11. Agreement between WebCIS and medication text files on entered date, by generic drug name

Entered date matches by generic name (N=239)	Total	Matched ± 30 days	Percent matched ± 30 days
Glimepiride	23	15	65.2
Glipizide	68	21	30.9
Glyburide	7	4	57.1
Metformin	89	48	53.9
Metformin & glyburide	2	0	0.0
Pioglitazone	21	16	76.2

Entered date matches by generic name (N=239)	Total	Matched ± 30 days	Percent matched ± 30 days
Rosiglitazone	28	14	50.0
Rosiglitazone & metformin	1	1	100.0

Agreement Study 1: Evaluation of Days Coverage Per Visit. Of the 239 WebCIS medication records that had an inactive date, 153 records (64.0%) had patient visit dates from the Medication text files that fell on or within the timespan between the WebCIS entered and inactivation dates.

For the 153 records that had a visit date (any type of provider) within the timespan of interest, we calculated the DCV. The mean and median DCV were 114 days and 78 days, respectively (range, 1–825 days) (Figure 3). The distribution of DCV was positively skewed. Most patients appeared to see their clinicians fairly regularly, regardless of the type of diabetes medication.

Figure 3. Days coverage per visit (DCV)

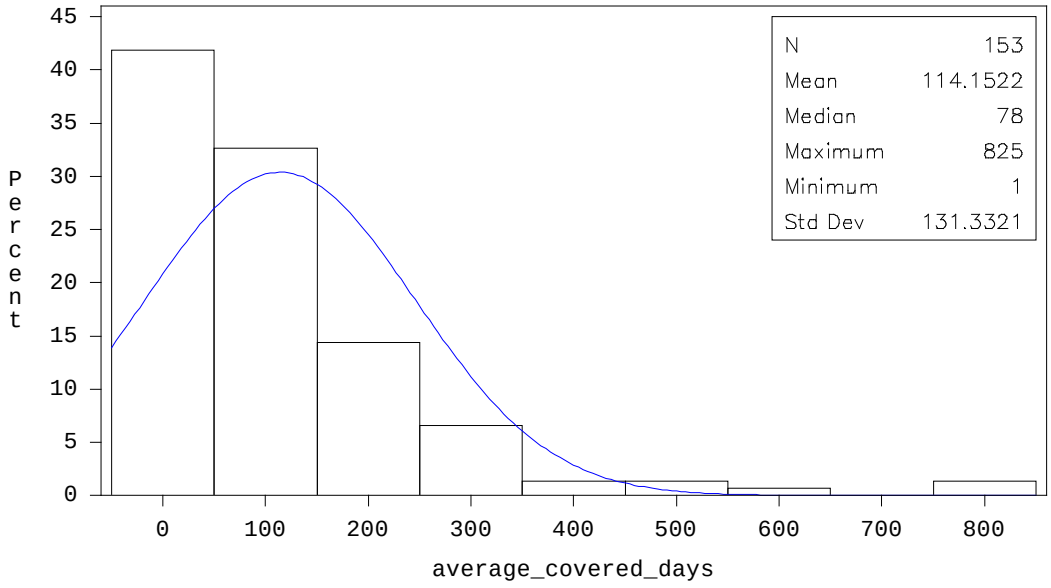


Table 12 shows the mean and median DCV for the 3 therapeutic classes and the 8 generic diabetes medications for the 239 WebCIS records with ≥1 match within the entered and inactivation timespan. The sulfonylureas had the largest mean and median DCV, both of which were driven by glipizide (mean and median DCV of 177.5 and 112.4, respectively). For the remaining therapeutic classes and generic diabetes medications, patients were being seen at UNC at median 2- to 3-month intervals (range, 43.7 for glimepiride to 84.3 for pioglitazone).

Table 12. Days coverage per visit (DCV), by therapeutic class and drug name

Name	N	Mean	Median
Therapeutic class			
Biguanides	63	84.4	52.5
Sulfonylurea	52	136	84.0
Thiazolidinediones	37	116.6	79.0

Name	N	Mean	Median
Generic drug names			
Glimepiride	16	88.6	43.7
Glipizide	30	177.5	112.4
Glyburide	6	77.4	59.5
Metformin	63	84.4	52.5
Metformin & glyburide*	—	—	—
Pioglitazone	20	118.3	84.3
Rosiglitazone	17	138.4	79.0
Rosiglitazone & metformin*	—	—	—
* Too infrequent to assess			

Agreement Study 2: Congruence between Dates from the Medication Text File and the WebCIS Drug File (TFDBPAT.txt). This analysis evaluated 714 diabetes drug listings for the 78 patients tagged by visit date from the Medication text file and compared the visit dates with the WebCIS intervals spanning the entered to inactivation dates (where inactivation date was available). For this analysis, we asked, "For each of the Medication text file dates, did it fall within a WebCIS interval that had both an entered and inactivation date?" We evaluated agreement as an exact match on generic drug name as well as a match ±7, ±14, and ±30 days of the entered and inactivated dates overall (Table 13), by therapeutic class ±30 days (Table 14), and by generic diabetes medication ±30 days (Table 15).

Table 13. Agreement between medication text file date and WebCIS interval dates

(N=714)	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Agreement for Medication text dates	364	51.0	364	51.0
Additional matched on ± 7 days	8	1.1	372	52.1
Additional matched on ± 14 days	7	1.0	379	53.1
Additional matched on ± 30 days	10	1.4	389	54.5
Not matched within 30 days	325	45.5	714	100.0

Table 14. Agreement between medication text file date and WebCIS interval dates, by therapeutic class

(N=714)	Total	Matched ± 30 days	Percent matched ± 30 days
Biguanides	300	168	56.0
Sulfonylurea	259	129	49.8
Thiazolidinediones	152	96	63.2
*Excludes the combination antidiabetic agents (n=3).			

Table 15. Agreement between medication text file date and WebCIS interval dates, by generic drug name

(n=714)	Total	Matched ± 30 days	Percent matched ± 30 days
Glimepiride	99	55	55.6
Glipizide	138	63	45.7
Glyburide	22	11	50.0
Metformin	300	153	51.0
Metformin & glyburide	1	0	0
Pioglitazone	64	52	81.3
Rosiglitazone	88	39	44.3
Rosiglitazone & Metformin	2	1	50.0

As discussed in the description of WebCIS development, there have been gradual WebCIS system improvements since the original version in 2001, with the most extensive modification occurring at the end of 2004. We repeated the analyses for Agreement Studies 1 and 2 using only data from 2005 to June 2006, with few or no changes in the results. Future analyses focusing on medication use in EMRs will require linkage with pharmacy dispensing records to ensure completeness of medication data.

Assessing Medication Use In Newly Diagnosed DMT2 Patients

This section provides results for the clinical aim of the study. We begin by providing a high-level view of the patient population seen by UNC clinicians. This is followed by analysis of the 23 patients who had an MI, a stroke, or both after their DMT2 diagnosis.

There were 175,094 patients seen at UNCHCS between January 1 and December 31, 2003. The demographics of the UNCHCS are provided in Tables 16 and 17 below, with Table 17 showing the comparison between the UNCHCS population in 2003 and the NC population in 2005 in terms of race, sex, and insurance type.[51]

Table 16. UNCHCS patient population in 2003, by age stratum

Age (years)	N	%
<20	41,713	23.8
20–29	23,646	13.5
30–39	26,696	15.3
40–49	26,829	15.3
50–59	23,794	13.6
60–69	15,793	9.0
≥70	16,623	9.5

Table 17. Demographic characteristics of the UNCHCS patient population in 2003 and the North Carolina population in 2005

	UNCHCS N	UNCHCS %	NC %
Race			
White	110,805	63.3	74.1
Black	37,955	21.7	21.8
Other	12,142	6.9	*
Hispanic	10,715	6.1	*
Asian	2,415	1.4	1.8
Native American	1,055	0.6	1.3
Sex			
Female	101,396	57.9	50.8
Male	73,698	42.1	49.2
Health Insurance			

Private	80,788	46.1	63.9
Medicare	37,797	21.6	14.7
Medicaid	33,569	19.2	12.8
Other	16,712	9.5	*
Military	6,228	3.6	5.3
* Unavailable			

From the population shown in Tables 16 and 17, we identified the patients who appeared to be newly diagnosed with DMT2 as of January 1, 2001 (n=1664). Because we could not determine which of the 1664 DM patients actually were newly diagnosed, we estimated the number of newly diagnosed patients by manually reviewing the transcription files for 200 of the 1664 patients. Based on this review, we estimated that ~1100 patients (65% of the 1664 cohort) were newly diagnosed with DMT2 in the 4.5-year period between January 1, 2001 and June 27, 2006. This equated to a diabetes incidence rate of 1.8/1000 persons. In 2004, the crude incidence rate for adults aged 18–79 was 7/1000,[52] indicating that the incidence rate in the UNC population was much lower than that for the nation as a whole. This finding may reflect the challenge of identifying newly diagnosed patients in the UNC population. Alternatively, the lower rate in the UNC clinic population may reflect a population different from that included in the National Health Interview Survey, from which the national incidence rates were derived.[52]

The analyses below are only for the 30 patients with verified, newly diagnosed DMT2 who had an MI or stroke or who died after DMT2 was diagnosed. Seven of these 30 patients died after their DM diagnosis, with all deaths unrelated to diabetes: 2 were due to lung cancer; 1, due to breast cancer; 1, due to HIV infection and liver disease; 2, due to surgical complications; and 1, due to multiple chronic conditions. Analyses are provided for the 23 patients who had an MI (n=7), stroke (n=15), or both (n=1). Because of cost and time constraints, we could not develop a text data-mining algorithm to differentiate between patients who were new to UNCHCS vs. new DMT2. Thus an analysis of patients without these 3 outcomes was not done.

Blood pressure measurements, weights, and lipid results were available for >90% of these 23 patients. At the time of DMT2 diagnosis, most patients (61%) already had both HT and DL, 6 (26%) had either HT or DL, and 2 of the stroke patients had DM only (Table 18). Table 19 shows the medications the patients were prescribed within 12 months after DMT2 diagnosis. In all, 3 of the MI patients (42%) and 2 of the stroke patients (13%) were not being treated with an antidiabetic drug. Only 42% of the patients with hypertension were prescribed an ACEI or ARB within 12 months after DMT2 diagnosis.

Table 18. Presence of HT and DL at time of DMT2 diagnosis among WebCIS patients who had MI or stroke

Outcome*	Comorbidity				
Frequency Row %	DM only	DM and HT	DM and DL	DM, HT, DL	Total
MI	0 0.00	0 0.00	2 28.57	5 71.43	7
Stroke	2 13.33	3 20.00	1 6.67	9 60.00	15
Total	2	3	3	14	22
*The 1 patient with an MI and stroke was excluded for classification purposes.					

Table 19. Drug classes prescribed within 12 months after DMT2 diagnosis in patients with HT, DL, or both

Outcome*	Therapeutic Class of Drugs Prescribed					
Frequency Row %	DM medication	DM and HT medication	DM and DL medication	HT and DL medications	DM, HT, DL medications	Total
MI	0 0.00	1 14.29	1 14.29	3 42.86	2 28.57	7
Stroke	6 40.00	2 13.33	1 6.67	2 13.33	4 26.67	15
Total	6 27.3	3 13.6	2 9.1	5 22.7	6 27.3	22
*The 1 patient with an MI and stroke was excluded for classification purposes.						

None of the MI or stroke patients was prescribed a thiazolidinedione as first-line therapy for DMT2. For the MI patients, 2 received a sulfonylurea as initial therapy for DMT2, 2 received metformin, and 1 received both. Of the 15 stroke patients, 13 were receiving antidiabetic treatment when the stroke occurred: 6 were taking sulfonylurea only; 4, metformin only; 2, sulfonylurea and metformin; and 1, metformin and insulin. In the 12 months before the first complication, only 50% of the MI patients had good control of their diabetes (HbA1c level ≤7%), as did 60% of the stroke patients (Table 20). Tables 21 and 22 show similar results for control of HT and DL, respectively. Only 20% of MI and stroke patients who had DM, HT, and DL had all 3 comorbidities controlled simultaneously, and in 12.5% of the patients, DM, HT, and DL were simultaneously uncontrolled (Table 23). In addition, DM was controlled, but not HT and DL, in about 20% of these patients.

Table 20. Outcomes by control of DMT2 during the 12 months before MI or stroke

Outcome	Diabetes well controlled?		
Frequency Row %	No	Yes	Total
MI	3 50.00	3 50.00	6
Stroke	6 40.00	9 60.00	15
Total	9	12	21
Frequency missing = 1. The 1 patient with MI and stroke was excluded from the analysis.			

Table 21. Outcomes by control of HT during the 12 months before MI or stroke

Outcome	Hypertension well controlled?		
Frequency Row %	No	Yes	Total
MI	3 42.86	4 57.14	7
Stroke	6 50.00	6 50.00	12
Total	9	10	19
Frequency missing = 3. The 1 patient with MI and stroke was excluded from the analysis.			

Table 22. Outcomes by control of DL during the 12 months before MI or stroke

Outcome	Hypertension Well Controlled?		Total
Frequency Row %	No	Yes	
MI	4 57.14	3 42.86	7
Stroke	5 45.45	6 54.55	11
Total	9	9	18
Frequency missing = 4. The 1 patient with MI and stroke was excluded from the analysis.			

Table 23. Control of DM, HT, and DL during the 12 months before MI or stroke

	Hypertension well controlled?	Cholesterol well controlled?		
	Frequency Row % Table %	No	Yes	Total
Diabetes NOT well controlled	No	2 66.67 12.5	1 33.33 6	3
	Yes	2 50.00 12.5	2 50.00 12.5	4
Diabetes well controlled	No	3 75.00 19	1 25.00 6	4
	Yes	2 40.00 12.5	3 60.00 19	5
Total		9	7	16
Frequency missing = 5. The 1 patient with MI and stroke was excluded from the analysis.				

The clinicians who reviewed the health information for the MI and stroke patients determined whether changes had been made to the patient's drug regimens to improve control of DM, HT, and/or DL. As shown in Table 24, more changes were made for DM control than for either HT or DL control, indicating that clinicians could have done more to improve HT and/or DL. Compared with soon after DMT2 diagnosis (Table 19), more patients were being treated for DM and HT in the 12 months before the outcome (Table 25); there was only a very slight increase in the number of patients being treated for all three comorbidities.

Table 24. Changes in drug therapy to improve control of DM, HT, and DL

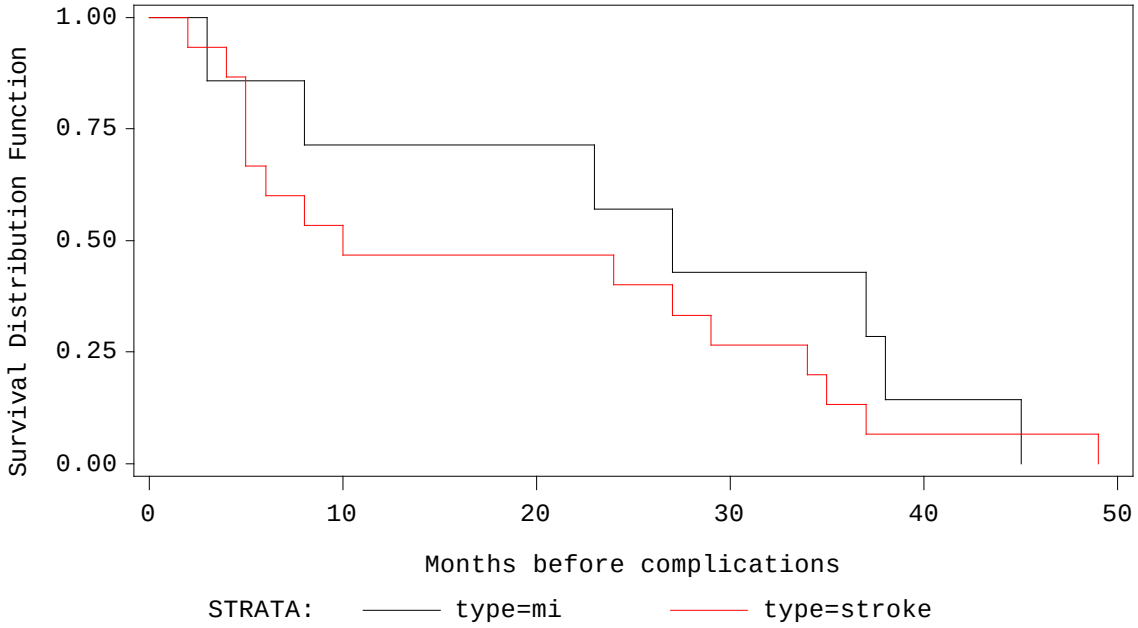
	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Diabetes regimen				
No change	9	39.13	9	39.13
Change	14	60.87	23	100.00
Hypertension regimen				
No change	11	57.89	11	57.89
Change	8	42.11	19	100.00
Dyslipidemia regimen				
No change	11	64.71	11	64.71
Change	6	35.29	17	100.00

Table 25. Drug classes prescribed within 12 months before MI or stroke

Outcome	Therapeutic Class of Drugs Prescribed					
Frequency Row %	DM medication	DM and HT medication	DM and DL medication	HT and DL medications	DM, HT, DL medications	Total
MI or Stroke	3 13.0	6 26.1	2 8.7	5 21.7	7 30.4	23

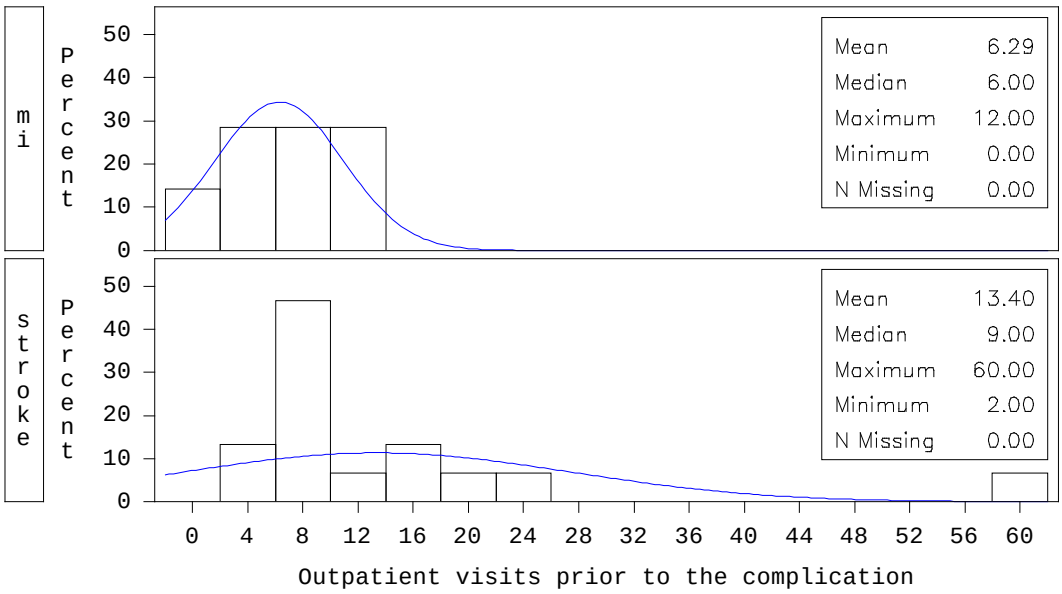
In the 7 patients who had an MI, the average time to this event was 25.6 months (median, 27 months; range, 3–45 months) (Figure 4). Thus, for most of these patients, significant time was available for risk factor modification before the negative outcome. For the 15 who had a stroke, the mean time to the event was 18.7 months (median, 10 months; range, 2–49 months) (Figure 4). These data suggest that stroke occurs more often than MI in this population and occurs sooner after DMT2 is diagnosed than does MI.

Figure 4. Time to MI or stroke



We hypothesized that patients who were at high risk of an MI or stroke might have more outpatient visits than those who had a lower risk of these outcomes. We therefore evaluated the number of visits that patients had had in the 12 months before an outcome in both the WebCIS and Medicaid files. Patients who had a stroke had visited their healthcare providers more often during the 12 months before the event than did patients who had had an MI (Figure 5). Again, there were substantial opportunities for intensification of their care before the event. Although the clinical review provided evidence of some dose titration, substantial opportunity for enhanced care remains. Some of these adverse outcomes may have been preventable.

Figure 5. Number of outpatient visits within 12 months before MI or stroke



Without an adequate comparison group (i.e., patients who did not die or have an MI or stroke after DMT2 diagnosis), we could not tell whether these trends explained MI and/or stroke occurrence. In all, 64% of the MI and stroke patients had HT and DL at the time of DMT2 diagnosis. Each of these conditions appeared to be well controlled in only half of the patients, and, among patients who had all 3 conditions, the conditions were well controlled in only 20% of cases. Clinicians appeared to be more responsive to changing DM medications than HT or DL medications. Stroke appeared to occur earlier and more often than did MI after DMT2 diagnosis, despite the fact that patients who had had a stroke had been seen much more frequently during the 12 months before the outcome than were those who had had an MI.

NC Medicaid Analyses

The NC Medicaid analysis mirrored that of the WebCIS analysis. For the 2794 patients with newly diagnosed DMT2, we determined the number (%) who had HT, DL, or both at the time of DMT2 diagnosis (Table 26). As would be expected, more of those who had an MI and/or stroke had HT and DL in addition to DMT2. Those who did not have an MI and/or stroke typically had DM alone or DM and HT.

Table 26. Prevalence of HT and/or DL at time of DMT2 diagnosis

Outcome	Comorbidity				
Frequency Row %	DM only	DM and HT	DM and DL	DM, HT, DL	Total*
MI	8 16.3	14 28.6	3 6.1	24 49.0	49
Stroke	32 18.5	80 46.2	8 4.6	53 30.6	173
MI or stroke	3 21.4	7 50.0	0 0.0	4 28.6	14
No MI or stroke	820 33.9	931 38.4	136 5.6	536 22.1	2423

Outcome	Comorbidity				
Frequency Row %	DM only	DM and HT	DM and DL	DM, HT, DL	Total*
Total	863 32.4	1032 38.8	147 5.5	617 23.2	2661

*Total sums to 2641 not 2794 because there were 17 MIs, 105 strokes, and 13 stroke & MIs prior to the first DMT2 diagnosis.

The medications dispensed within 12 months after DMT2 diagnosis are listed in Table 27. They reflect the comorbidities of the patients displayed in Table 26.

Table 27. Drug classes dispensed within 12 months after DMT2 diagnosis

Outcome	Therapeutic Class of Drugs Prescribed							
Frequency Row %	DM drug	DM and HT drugs	DM and DL drugs	HT and DL drugs	DM, HT, DL drugs	HT drugs	DL drugs	Total*
MI	2 4.1	14 28.6	0 0.0	2 4.1	30 61.2	1 2.0	0 0.0	49
Stroke	13 7.6	76 44.2	6 3.5	10 5.8	57 33.1	9 5.2	1 0.6	172
MI or stroke	1 7.1	4 28.6	0 0.0	0 0.0	9 64.3	0 0.0	0 0.0	14
No MI or stroke	533 22.7	925 39.3	131 5.6	45 1.9	625 26.6	87 3.7	5 0.2	2351
Total	549 21.2	1019 39.4	137 5.3	57 2.2	721 27.9	97 3.8	6 0.2	2586

* One stroke patient and 72 patients without an MI or stroke are missing because we there we no medication claims for these patients.

In comparison to the results from WebCIS, in which 42% of those with new-onset DMT2 and HT were prescribed an ACEI, an ARB, or both, 61.5% of NC Medicaid patients were dispensed these medications (Table 28).

Table 28. ACEIs, ARBs, or both dispensed within 12 months after DMT2 diagnosis for patients with both DMT2 and HT

Frequency Row %	ACEI	ARB	Both	Neither	Total
MI	20 46.5	7 16.3	5 11.6	11 25.6	43
Stroke	51 37.5	22 16.2	11 8.1	52 38.2	136
MI and Stroke	11 100.0	0 0.0	0 0.0	0 0.0	11
No MI or Stroke	617 39.5	237 15.1	97 6.2	612 39.2	1563
Total	699 39.9	266 15.2	113 6.4	675 38.5	1753

Figures 6–8 display the mean and median number of visits during the 12 months before MI and/or stroke. The mean number of pre-event visits did not differ substantially among the 3 outcomes, but there was a large difference between the mean (n=11.2) and median (n=5) number of visits among patients who had both an MI and a stroke. Most of the patients had few outpatient visits before the first outcome, although 1 patient had as many as 59 visits within 12 months before the complication, and several others had ~2 visits per month. Clinicians treating

NC Medicaid patients with DMT2 and comorbid HT and DL had sufficient opportunities to intervene to prevent an MI and/or stroke.

Figure 6. Number of physician or hospital outpatient visits within 12 months before MI

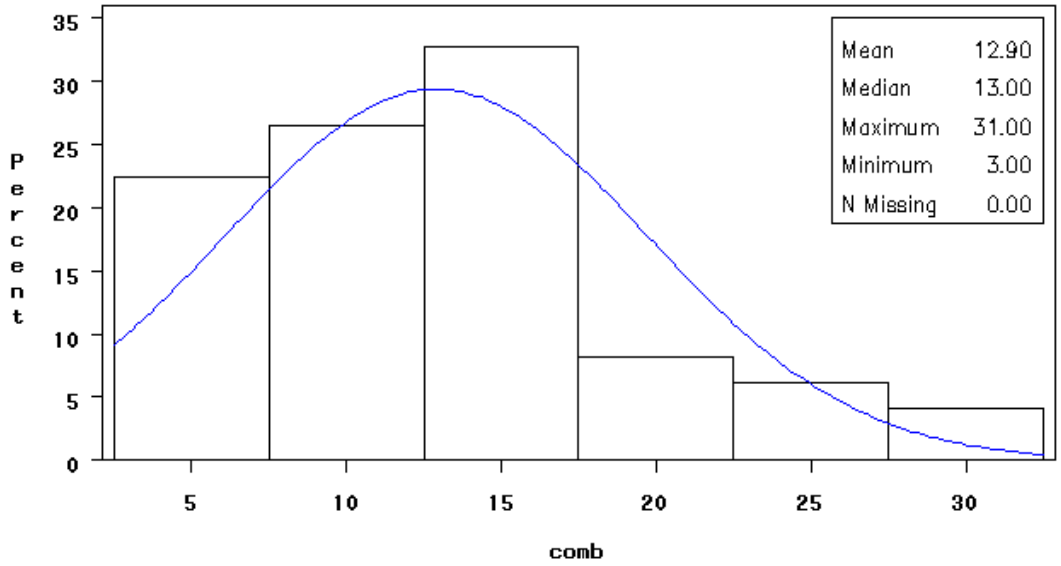


Figure 7. Number of physician or hospital outpatient visits within 12 months before stroke

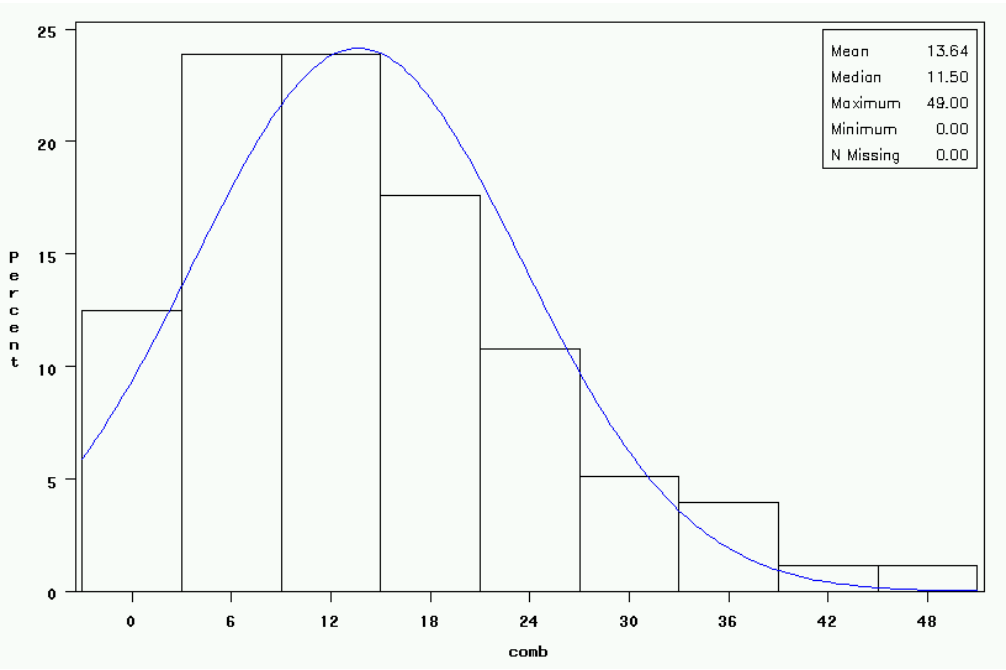
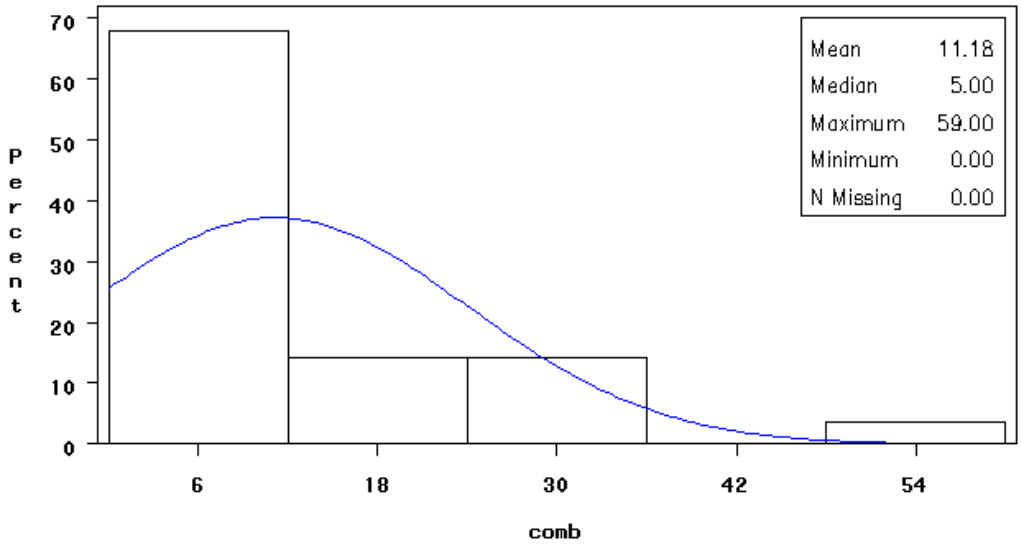


Figure 8. Number of physician or hospital outpatient visits within 12 months before MI and stroke



Discussion

The major findings from this study relate to the knowledge we gained and the processes we developed for using EMRs such as WebCIS for observational research. We also compared the utility of using EMR versus administrative claims data to study DM and comorbid HT and DL. The findings regarding the informatics goal of the project are summarized as bullet points below.

- EMRs provide a rich source of clinical data, but there are numerous challenges to its use and interpretation that depend on its penetration into clinical care.
- Text data that arise from transcription notes contain valuable information about the timing of diagnosis, over-the-counter medications, family history, and behavioral factors but require extensive computer processing to maximize their usefulness for research.
- The availability of medication information and laboratory data measuring glycemic control and lipids were critical for this project. The completeness of EMR data to be used for clinical research must be assessed because patients see many providers, have prescriptions filled at different pharmacies, and have laboratory tests performed in multiple locations.

Our clinical goal was to evaluate medication use in newly diagnosed DMT2 patients using EMR and administrative claims data. For this goal, we focused on patients who had died or had diabetic complications such as MI or stroke. We found that:

- Many DMT2 patients with adverse diabetic outcomes had coexisting dyslipidemia and/or hypertension.
- Given the frequency with which these patients sought care, there were ample opportunities to treat these comorbid conditions. More aggressive treatment might have reduced the rate of adverse outcomes.

The UNC DEcIDE project was focused on the overuse, underuse, and inappropriate use of medications to treat DMT2. We quickly realized that it would be very difficult to use administrative claims data to assess the adequacy of medication use, because a critical factor for such evaluations is the availability of clinical data such as HbA1c levels, systolic and diastolic blood pressure measurements, and lipid levels. We therefore turned to UNC's EMR, WebCIS, to evaluate medication use in newly diagnosed DMT2 patients, keeping in mind that medication recording in WebCIS might not be as complete as desired, given the stated goals.

We began the project with a literature scan to identify the clinical factors that often assessed in experimental and observational studies of glycemic control in newly diagnosed DMT2 patients. While identifying the literature for inclusion, we noted that there was very little information on initial treatment patterns in newly diagnosed patients with DMT2, i.e., do DMT2 patients receiving sulfonylureas do better or worse than those receiving metformin? We noted from our review that assessing the adequacy of glycemic control requires access to HbA1c test results. The project clinicians and the reviewed studies also stressed the importance of including measures of blood pressure and lipid levels to attain quality, evidence-based care of DMT2 patients. Thus, for patients with comorbid HT and/or DL, we also evaluated maintenance of systolic and diastolic blood pressures and appropriate lipid levels. We focused attention on WebCIS patients who had had an MI and/or stroke for these analyses. Unfortunately, because of resource constraints, we could not conduct a parallel investigation of newly diagnosed DMT2 patients in WebCIS who had not had these outcomes.

Our initial plan was to develop algorithms from the WebCIS analyses that could be applied to the NC Medicaid claims data. These algorithms would evaluate whether administrative claims data regarding diagnoses, pharmaceuticals, and procedures could be used to assess the adequacy of treating diabetes and its associated comorbidities. The WebCIS system has significant potential as a data source for clinical and health-services research. The system is comprehensive and spans all specialties. It represents an interesting example of a 'second-generation' EMR, and we need to learn more about how to conduct research from such clinical repositories. However, we had substantial challenges in taking the raw clinical data from the WebCIS system and making it into an easily analyzable research file. We especially had difficulty identifying cases of new-onset diabetes, as opposed to prevalent cases. The process required multiple steps with additional manual review at each step. In many cases, these were patients who were initially receiving only outpatient specialty care at UNC who, after a period of years, also began receiving more comprehensive care, including care for DMT2. This pattern would give the appearance of 'new-onset' diabetes, but in reality, the DM had been present for some time but treated elsewhere.

With extensive reanalysis and validation through manual review of clinic notes, we assembled a file of cases of new-onset DM with complications. However, we did not anticipate at the initiation and budgeting of this project the value that text mining would bring to the challenges we faced. Thus we did not have the financial resources to assemble a control group of patients with new-onset DM who had not had complications. Without the ability to compare findings in DMT2 patients with and without adverse outcomes, we could not develop algorithms for testing in the Medicaid dataset. However, we conducted similar analyses of WebCIS and NC Medicaid data, confirming that DMT2 patients with both HT and DL are at increased risk of MI and/or stroke.

With WebCIS, we had access to HbA1c values, blood pressure measurements, and lipid levels; thus clinicians could evaluate the adequacy of treatment for DM, HT, and DL among

patients with MI and/or stroke. Only 20% of the patients who had DM, HT, and DL were receiving adequate treatment for all 3 conditions. Clinicians appeared to change DM medications more often than those for HT or DL; they might have focused more on glucose control than on control of these important, and often difficult to manage, comorbidities. Further, patients were seen at UNC a great deal in the year before the complications. There were sufficient opportunities for care, such that patients could have received better management of their cardiovascular risk factors. Examining the reasons for this undertreatment in these complex patients is beyond the scope of this analysis, but they include 'clinical inertia' (the tendency of clinicians to continue past treatment) and clinical overload—clinicians may simply be overwhelmed by the number and complexity of acute and chronic problems faced by these patients.

Comparing the patients seen at UNCHCS with those covered by NC Medicaid might have been inappropriate, given the different populations covered by each system. A smaller percentage of NC Medicaid patients who had had an MI and/or stroke had comorbid HT and/or DL compared with WebCIS patients, possibly reflecting a greater burden of illness at the tertiary care center. A striking difference is the use of ACEI/ARBs in the hypertensive patients in these 2 populations: 42% in WebCIS patients versus 61.6% in Medicaid patients. The Medicaid patients thought to be newly diagnosed might actually have been prevalent cases not identified as such by the algorithm.

The ability to evaluate medication exposures in WebCIS was not ideal for this project. A fair number of UNCHCS practitioners did not use the drug entry fields before Version 2.0 of WebCIS, until they could use CPOE that would print out a prescription for them. The drug data residing in the WebCIS medication fields are therefore likely to have been incomplete. To address this issue, we explored the utility of the clinician's visit transcriptions for containing information on medication use. As described in the Methods section, text mining was used to identify drug names, doses, and regimen from transcriptions. An important constraint of the transcription data was the limited ability to identify when clinicians reported stopping the use of a drug, i.e., using a negation protocol as part of the text data-mining application.

This phenomenon of partial use of the EMR may be relatively common in practice. We attempted to examine how underuse of these systems might have affected the analysis by conducting 2 agreement studies. Neither data source, the WebCIS prescribing data or the Medication text file, contained more complete medication data. For WebCIS, the clinicians probably did not use the prescribing module regularly, given that the date (±30 days) from the Medication text file matched within an appropriate WebCIS interval only about half of the time (Agreement Study 2). The agreement differed by therapeutic class and type of medication, with the best agreement for pioglitazone (81.3%) and much lower for the remaining types of diabetes medications. Similar to the findings from Agreement Study 2, the best agreement between WebCIS entered date and the Medication text file was noted for pioglitazone, but most agreement by therapeutic class or generic type was in the 50% range. We attribute the mediocre agreement to the varied ways in which clinicians recorded medications within their transcription notations: some were very specific about their prescribed treatments (name, dose, frequency of use) and others provided only the drug name. The DCV analysis provided a measure of how often patients taking each therapeutic class or generic type of medication saw UNC care providers. The median DCVs clustered at 50 and 80 days, suggesting that patients were seeing their clinicians every 2 or 3 months.

Another potential reason for incomplete medication data, or incomplete clinical data overall, relates to the type of institution supplying the data. UNCHCS is an academic medical center, where patients treated in the community setting are often sent for expert consultation. A major challenge for this project was the ability to identify patients who were being cared for regularly by clinicians at UNCHCS versus those who sought care sporadically. We developed criteria to differentiate these 2 groups that may or may not be appropriate for other DMT2 studies or populations. UNC does not differ substantially from most other academic health centers; many patients receive care from both the academic health center and community practitioners. The EMR will provide only the portion of care the patient received at the academic center unless the health systems have been integrated. This is also the case with patients who receive care at Veterans Affairs hospitals. Staff model health maintenance organizations (HMOs), such as the Permanente Medical Group, will be somewhat less susceptible to this issue of 'leakage' of care outside the system being studied.

To enhance the use of the EMR medication list and prescription writing capability, systems must facilitate prescription writing compared with the conventional method of handing a prescription to the patient. New enhancements to WebCIS, such as the ability to e-prescribe to multiple area pharmacies, may lead to greater use of the medication list and prescription writing system. In the meantime, researchers working from such computerized medication lists should take into account the likely incomplete use of such systems. Of course, pharmacy claims databases are more complete than EMRs because they reflect prescriptions dispensed to the patient. Depending on the healthcare issue of concern, some data resources may be better than others. The challenge faced by researchers is having familiarity with multiple data sources to determine which is best for each project, given the nuances of each database.

By using both EMRs and administrative claims to conduct similar analyses, we could gauge the utility of both data sources. As always, there were advantages and disadvantages for each type of data. Whereas the EMR contained only a portion of the patient's medical history, the information was very rich for research purposes. Alternatively, when administrative claims were used for research, we had the entirety of the medical history while the patient was covered by the insurer, but the information indicated only the care that was received, not its results.

The fragmentation of the US healthcare system constitutes a major obstacle for researchers analyzing EMRs and administrative claims data. When patients change healthcare providers or insurers, the ability to bridge their medical information is extremely limited. As a result, researchers cannot follow patients over long periods. This is particularly important for addressing the long-term safety of medications, especially those that may be carcinogenic or cause other illnesses that take time to develop.

There is great hope that EMRs will lead to improved healthcare, more efficient provision of care, and reduced healthcare costs—that they will be the salvation of the US healthcare crisis. However, EMRs provide only a window of the care given to the patient by 1 clinical specialty or at a single clinical site. The UNC WebCIS system is unusual among academic medical centers, in that all clinical specialties use the same EMR and the system cuts across both inpatient and outpatient care. If we find ways to link EMRs with administrative claims using a unique medical code (preferred) or through probability matching (less optimal), the data systems will be much more robust. Given the technological advances that are occurring daily and the expertise that is being developed, the potential of EMRs for recording healthcare and for research is phenomenal. However, as this project indicates, there are many pitfalls that must be overcome. This will take

several years, skilled minds, standardization of terms, attention to detail, and a great deal of patience.

The preliminary text mining results suggest that automated information extraction from medical records can offer much with respect to detecting the early onset of diabetes. Identifying information from the medical record that other hospital information systems do not capture will clearly have the greatest impact; however, text mining can also identify data inconsistencies. As health care moves towards data mining methods that aggregate from multiple sources, the importance of data quality can not be overstated.

Although our preliminary results are encouraging, the scope of this project did not allow sufficient time to fully explore the degree to which text mining can contribute to early onset of diabetes. For example, the current technique identifies only the drug, the drug amount and the treatment schedule, but information such as adverse effects, negation and differentiating between current and historical treatment strategies would also aid in longitudinal studies of drug usage in a population. To help fill this void, Drs Blake and West wrote and were subsequently awarded a grant from the Computer Research Association's Committee on the Status of Women in Computing Research (CRAW). The new grant, which funds three female undergraduate students in the Schools of Information and Library Science and Public Health for a year, is a quintessential example of the multi-disciplinary team required to remove some of the technology limitations that we observed. Between this and the CRAW projects, Drs Blake and West hope to secure additional external funding that will be required to fully realize the potential of text mining in a public health setting.

Translation of Findings

Patients and Providers

In 2005, an estimated ~7% of the U.S. population, or 21 million Americans, had diabetes. The burden of this disease is extremely high in terms of the overall health of the nation as well as for healthcare costs. Strict control of glucose levels is critical for reducing diabetic complications. However, little is known about the comparative effects of initial antidiabetic medications on the likelihood of diabetic complications.

We identified patients who saw UNCHCS clinicians on a regular basis and were newly diagnosed with diabetes mellitus type 2 on or after January 1, 2001. These patients were followed for development of MI or stroke after the diabetes diagnosis (and who did not have 1 of these outcomes before diabetes diagnosis). Using UNC’s EMRs (WebCIS), clinicians evaluated whether glycemia, blood pressure, and lipids were well controlled before the development of these outcomes. Diabetes, hypertension, and dyslipidemia were well controlled with medication in only 20% of the patients. Substantial room for improvement exists for both better documentation and better care for the UNC patient population.

We conducted a similar evaluation using NC Medicaid data, looking at the proportion of newly diagnosed diabetes patients who had hypertension and dyslipidemia. We could not assess the adequacy of control in this group, however, because test results and vital sign measurements were not available.

Comparing the 2 data sources used for this study, WebCIS and NC Medicaid, we found advantages and disadvantages for each type of data. The ability to look at clinical factors such as laboratory test results and vital status in the data from UNC’s WebCIS was particularly valuable

for this project. However, it provides information only for the care given to patients at UNC. If a patient is seen by practitioners outside of UNC, this care is not captured by WebCIS.

Alternatively, NC Medicaid data does capture all of the medical care a patient receives that is reimbursable by Medicaid. If patients use over-the-counter medications, though, such medications will not be available for study. The major disadvantage in using NC Medicaid data to assess the adequacy of medication use in diabetics with and without comorbid hypertension and/or dyslipidemia is that test results such as HbA1c and cholesterol levels are not available; one knows only that the test was done. In addition, although the patients labeled as 'new-onset' diabetes in Medicaid had no claims for DM in the year before the observational period, we could not confirm this new-onset status with clinical data or notes as we could with the WebCIS patients. Therefore, the Medicaid data may have contained an unknown proportion of patients who had prevalent diabetes. Even with multiple analyses of WebCIS EMR data, we still had difficulty attaining >50% specificity for identification of new-onset DMT2. Future research should build on this finding and attempt to identify ways to develop more specific identification algorithms without the need for time-consuming review of clinical notes. Again, text mining may hold promise in this area.

The design of UNC's WebCIS began in the mid-1980s and has been improved incrementally since. A major redesign occurred in 2000, which greatly facilitated clinical practice throughout UNCHCS. The clinician/researchers involved in WebCIS's development had a view to the future and recognized the need for discrete data that could be used for research, in addition to full-text information that would allow more complete recording of clinical care for practitioners. The balance between discrete and full-text data makes use of WebCIS more efficient for both clinical care and research. UNC is now developing a full data warehouse, which will greatly facilitate research applications.

Policymakers

Our literature scan focused on publications that addressed the effectiveness of medications used by newly diagnosed diabetic patients. Most of the studies were randomized, controlled trials comparing thiazolidinediones (typically pioglitazone) with other oral antidiabetic medications. We did identify several observational studies that examined time to secondary failure of medications with regard to cardiovascular outcomes, but most of these studies used the Saskatchewan Health data. These are administrative claims data, so they typically provide a complete picture of patient care but do not allow assessment of clinical factors such as HbA1c level or blood pressure measurements.

Studies conducted by Nichols and colleagues[4] and Cook and colleagues[5] show the utility and value of EMRs for studying diabetes, an increasingly prevalent disease that puts a huge burden on the healthcare system. EMRs provide valuable clinical data that are not available from administrative claims, but both data sources are needed to obtain the full picture of a patient's clinical care. Thus a very critical need is to determine ways to link EMRs and administrative claims in the HIPAA environment, so that we can assemble a longitudinal record for each patient. Adequate funding and interdisciplinary teams are 2 requirements for addressing these rate-limiting steps. Realizing the potential of EMRs for assessing the quality and effectiveness of healthcare will be a challenge that takes time, effort, and funding. The challenges can be overcome to improve healthcare quality, enhance efficiency, and reduce costs.

References

1. National Institute of Diabetes and Digestive and Kidney Diseases. National Diabetes Statistics fact sheet: general information and national estimates on diabetes in the United States. 2005. Available at: http://diabetes.niddk.nih.gov/dm/pubs/statistics/index.htm#7. Accessed November 20, 2006.
2. North Carolina Rural Economic Development Center Inc. Rural Data Bank: Population in North Carolina. Available at: http://www.ncruralcenter.org/databank/trendpage_Population.asp. Accessed November 20, 2006.
3. Turner RC, Cull CA, Frighi V, et al. Glycemic control with diet, sulfonylurea, metformin, or insulin in patients with type 2 diabetes mellitus: progressive requirement for multiple therapies (UKPDS 49). UK Prospective Diabetes Study (UKPDS) Group. JAMA 1999 Jun 2;281(21):2005-2012.
4. Nichols GA, Alexander CM, Girman CJ, et al. Treatment escalation and rise in HbA1c following successful initial metformin therapy. Diabetes Care 2006 Mar;29(3):504-509.
5. Cook MN, Girman CJ, Stein PP, et al. Initial monotherapy with either metformin or sulphonylureas often fails to achieve or maintain current glycaemic goals in patients with Type 2 diabetes in UK primary care. Diabet Med 2007 Apr;24(4):350-358.
6. Selby JV, Peng T, Karter AJ, et al. High rates of co-occurrence of hypertension, elevated low-density lipoprotein cholesterol, and diabetes mellitus in a large managed care population. Am J Manag Care 2004 Feb;10(2 Pt 2):163-170.
7. Fuke D, Hunt J, Siemienczuk J, et al. Cholesterol management of patients with diabetes in a primary care practice-based research network. Am J Manag Care 2004 Feb;10(2 Pt 2):130-136.
8. Karter AJ, Moffet HH, Liu J, et al. Achieving good glycemic control: initiation of new antihyperglycemic therapies in patients with type 2 diabetes from the Kaiser Permanente Northern California Diabetes Registry. Am J Manag Care 2005 Apr;11(4):262-270.
9. Lesho EP, Myers CP, Ott M, et al. Do clinical practice guidelines improve processes or outcomes in primary care? Mil Med 2005 Mar;170(3):243-246.
10. Nau DP, Mallya U. Sex disparity in the management of dyslipidemia among patients with type 2 diabetes mellitus in a managed care organization. Am J Manag Care 2005 Feb;11(2):69-73.
11. Pladevall M, Williams LK, Potts LA, et al. Clinical outcomes and adherence to medications measured by claims data in patients with diabetes. Diabetes Care 2004 Dec;27(12):2800-2805.
12. Singh H, Kalavar J. Quality of care for hypertension and diabetes in federal- versus commercial-managed care organizations. Am J Med Qual 2004 Jan-Feb;19(1):19-24.
13. Timpe EM, Amarshi N, Reed PJ. Evaluation of angiotensin-converting enzyme inhibitor use in patients with type 2 diabetes in a state managed care plan. Am J Manag Care 2004 Feb;10(2 Pt 2):124-129.
14. Yamanouchi T, Sakai T, Igarashi K, et al. Comparison of metabolic effects of pioglitazone, metformin, and glimepiride over 1 year in Japanese patients with newly diagnosed Type 2 diabetes. Diabet Med 2005 Aug;22(8):980-985.
15. Bolen S, Feldman L, Vassy J, et al. Systematic Review: Comparative Effectiveness and Safety of Oral Medications for Type 2 Diabetes Mellitus. Ann Intern Med 2007 Jul 16.
16. Carey TS, Thomas D, Woolsey A, et al. Half a loaf is better than waiting for the bread truck. A computerized mini-medical record for outpatient care. Arch Intern Med 1992 Sep;152(9):1845-1849.

17. Hammond JE, Berger RG, Carey TS, et al. Making the transition from information systems of the 1970s to medical information systems of the 1990s: the role of the physician's workstation. J Med Syst 1991 Jun;15(3):257-267.

18. Regenstrief Institute Inc. Logical Observation Identifiers Names and Codes (LOINC®). Available at: http://www.regenstrief.org/medinformatics/loinc/. Accessed November 26, 2006.

19. North Carolina Department of Health and Human Services Division of Medical Assistance. Medicaid in North Carolina Annual Report: State Fiscal Year 2006. Available at: http://www.dhhs.state.nc.us/dma/2006report/2006report.pdf. Accessed September 1, 2007.

20. Kraus S, Blake C, West SL. Information Extraction from Medical Notes (P187). 12th World Congress on Health (Medical) Informatics; 2007; Brisbane, Australia: MedInfo; 2007. p. 1662-1664.

21. Standards of medical care in diabetes--2007. Diabetes Care 2007 Jan;30 Suppl 1:S4-S41.

22. Chobanian AV, Bakris GL, Black HR, et al. Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. Hypertension 2003 Dec;42(6):1206-1252.

23. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation 2002 Dec 17;106(25):3143-3421.

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

25. Eurich DT, Majumdar SR, McAlister FA, et al. Improved clinical outcomes associated with metformin in patients with diabetes and heart failure. Diabetes Care 2005 Oct;28(10):2345-2351.

26. Johnson JA, Majumdar SR, Simpson SH, et al. Decreased mortality associated with the use of metformin compared with sulfonylurea monotherapy in type 2 diabetes. Diabetes Care 2002 Dec;25(12):2244-2248.

27. Simpson SH, Majumdar SR, Tsuyuki RT, et al. Dose-response relation between sulfonylurea drugs and mortality in type 2 diabetes mellitus: a population-based cohort study. CMAJ 2006 Jan 17;174(2):169-174.

28. Lusignan S, Sismanidis C, Carey IM, et al. Trends in the prevalence and management of diagnosed type 2 diabetes 1994-2001 in England and Wales. BMC Fam Pract 2005 Mar 22;6(1):13.

29. Rajagopalan R, Iyer S and Khan M. Effect of pioglitazone on metabolic syndrome risk factors: results of double-blind, multicenter, randomized clinical trials. Curr Med Res Opin 2005 Jan;21(1):163-172.

30. Hallsten K, Virtanen KA, Lonnqvist F, et al. Rosiglitazone but not metformin enhances insulin- and exercise-stimulated skeletal muscle glucose uptake in patients with newly diagnosed type 2 diabetes. Diabetes 2002 Dec;51(12):3479-3485.

31. Nakamura T, Ushiyama C, Shimada N, et al. Comparative effects of pioglitazone, glibenclamide, and voglibose on urinary endothelin-1 and albumin excretion in diabetes patients. J Diabetes Complications 2000 Sep-Oct;14(5):250-254.

32. Tan M, Johns D, Gonzalez Galvez G, et al. Effects of pioglitazone and glimepiride on glycemic control and insulin sensitivity in Mexican patients with type 2 diabetes mellitus: A multicenter, randomized, double-blind, parallel-group trial. Clin Ther 2004 May;26(5):680-693.

33. Tan MH, Johns D, Strand J, et al. Sustained effects of pioglitazone vs. glibenclamide on insulin sensitivity, glycaemic control, and lipid profiles in patients with Type 2 diabetes. Diabet Med 2004 Aug;21(8):859-866.

34. Watanabe I, Tani S, Anazawa T, et al. Effect of pioglitazone on arteriosclerosis in comparison with that of glibenclamide. Diabetes Res Clin Pract 2005 May;68(2):104-110.

35. Lexchin J, Bero LA, Djulbegovic B, et al. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. BMJ 2003 May 31;326(7400):1167-1170.

36. Campbell IW, Menzies DG, Chalmers J, et al. One year comparative trial of metformin and glipizide in type 2 diabetes mellitus. Diabete Metab 1994 Jul-Aug;20(4):394-400.

37. Charbonnel BH, Matthews DR, Schernthaner G, et al. A long-term comparison of pioglitazone and gliclazide in patients with Type 2 diabetes mellitus: a randomized, double-blind, parallel-group comparison trial. Diabet Med 2005 Apr;22(4):399-405.

38. Derosa G, Mugellini A, Ciccarelli L, et al. Comparison of glycaemic control and cardiovascular risk profile in patients with type 2 diabetes during treatment with either repaglinide or metformin. Diabetes Res Clin Pract 2003 Jun;60(3):161-169.

39. Derosa G, Mugellini A, Ciccarelli L, et al. Comparison between repaglinide and glimepiride in patients with type 2 diabetes mellitus: a one-year, randomized, double-blind assessment of metabolic parameters and cardiovascular risk factors. Clin Ther 2003 Feb;25(2):472-484.

40. Goke B. Improved glycemic control and lipid profile in a randomized study of pioglitazone compared with acarbose in patients with type 2 diabetes mellitus. Treat Endocrinol 2002;1(5):329-336.

41. Horton ES, Foley JE, Shen SG, et al. Efficacy and tolerability of initial combination therapy with nateglinide and metformin in treatment-naive patients with type 2 diabetes. Curr Med Res Opin 2004 Jun;20(6):883-889.

42. Pagano G, Marena S, Corgiat-Mansin L, et al. Comparison of miglitol and glibenclamide in diet-treated type 2 diabetic patients. Diabete Metab 1995 Jun;21(3):162-167.

43. Pavo I, Jermendy G, Varkonyi TT, et al. Effect of pioglitazone compared with metformin on glycemic control and indicators of insulin sensitivity in recently diagnosed patients with type 2 diabetes. J Clin Endocrinol Metab 2003 Apr;88(4):1637-1645.

44. Ramachandran A, Snehalatha C, Salini J, et al. Use of glimepiride and insulin sensitizers in the treatment of type 2 diabetes--a study in Indians. J Assoc Physicians India 2004 Jun;52:459-463.

45. Schernthaner G, Matthews DR, Charbonnel B, et al. Efficacy and safety of pioglitazone versus metformin in patients with type 2 diabetes mellitus: a double-blind, randomized trial. J Clin Endocrinol Metab 2004 Dec;89(12):6068-6076.

46. Derosa G, Franzetti I, Gadaleta G, et al. Metabolic variations with oral antidiabetic drugs in patients with Type 2 diabetes: comparison between glimepiride and metformin. Diabetes Nutr Metab 2004 Jun;17(3):143-150.

47. Tan MH, Baksi A, Krahulec B, et al. Comparison of pioglitazone and gliclazide in sustaining glycemic control over 2 years in patients with type 2 diabetes. Diabetes Care 2005 Mar;28(3):544-550.

48. Horton ES, Clinkingbeard C, Gatlin M, et al. Nateglinide alone and in combination with metformin improves glycemic control by reducing mealtime glucose levels in type 2 diabetes. Diabetes Care 2000 Nov;23(11):1660-1665.

49. Donnelly LA, Doney AS, Hattersley AT, et al. The effect of obesity on glycaemic response to metformin or sulphonylureas in Type 2 diabetes. Diabet Med 2006 Feb;23(2):128-133.

50. Eurich DT, Simpson SH, Majumdar SR, et al. Secondary failure rates associated with metformin and sulfonylurea therapy for type 2 diabetes mellitus. Pharmacotherapy 2005 Jun;25(6):810-816.

51. U.S. Census Bureau. United States Census 2000. Available at http://www.census.gov/main/www/cen2000.html. Accessed September 1, 2007.

52. Centers for Disease Control and Prevention National Center for Chronic Disease Prevention and Health Promotion. National Diabetes Surveillance System Incidence of Diabetes in the Population Aged 18-79 Years. 2006. Available at: http://www.cdc.gov/diabetes/statistics/incidence/index.htm. Accessed September 1, 2007.

Appendixes

Appendix A. Glossary and Evidence Table

Glossary

Δ	change
↓	decrease
↑	increase
↔	no change
µg	micrograms
µg	micrograms per minute
%	percent
A	acarbose
ADR	adverse drug reaction
AER	albumin excretion rate
AIP	atherogenic index of plasma
Alk Phos	alkaline phosphatase
ALT	alanine aminotransferase
apo A-I	apolipoprotein A-I
apo B	apolipoprotein B
AST	aspartate aminotransferase
ba-PWV	brachial-ankle pulse wave velocity
BMI	body mass index
BP	blood pressure
BUN	blood urea nitrogen
C	combination therapy (see each trial for specific combination)
CAD	coronary artery disease
CDS	chronic disease score
CHF	congestive heart failure
CI	confidence interval
CNS	central nervous system
CrCL	creatinine clearance
CRP	C-reactive protein
DBP	diastolic blood pressure
DKA	diabetic ketoacidosis
DM	diabetes mellitus
ER	emergency room
ET-1	endothelian ET-1
FFA	free fatty acids
FPG	fasting plasma glucose
FSI	Fasting serum insulin
FSI	Fasting serum insulin
g	grams
g/L	grams per liter
G	second generation sulfonylurea such as glipizide, gliclazide, glimepiride, or glyburide

GGT	glutamyl transferase
GI	gastrointestinal
HbA1c	hemoglobin A1C
h	hour
h/CRP	high sensitivity C-reactive protein
Hcy	homocysteine
HDL	high-density lipoprotein cholesterol
HOMA-B	homeostasis model assessment of beta cell function
HOMA-BF	homeostasis model assessment beta cell function
HOMA-IR	homeostasis model assessment of insulin resistance
HOMA-S	homeostasis model assessment for insulin sensitivity
HR	hazard ratio
IRI	immunoreactive insulin
ITT	intention to treat analysis
kcal	kilocalorie
kg	kilograms
LDL	low-density lipoprotein cholesterol
LFTs	liver function tests
Lp (a)	lipoprotein (a)
M	metformin or miglitol
m^2	meters squared
mcU/mL	microunits per milliliter
mEq/L	milliequivalent per liter
mg	milligram
mg/dL	milligrams per deciliter
mmHg	millimiters of mercury
mmol/L	millimoles per liter
μmol/L	micromole per liter
μIU/L	micro-international units per liter
μU/mL	microunits/mL
mU/L	milliunits per liter
mU/L	milliunits per milliliter
N	nateglinide
n/a	not applicable
ng/mL	nanograms/mL
nmol/L	nanomoles per liter
NR	not reported
NS	not significant
OAM	oral antihyperglycemic medication
OR	odds ratio
P	pioglitazone or placebo
PAI-I	plasminogen activator inhibitor
PET	positron-emission tomography
pg/L	picograms per liter
pg/mL	picograms per milliliter
pmol/L	picomoles per liter

PPG	postprandial glucose
PPI	postprandial insulin
PSI	postprandial serum insulin
QUICKI	quantitative insulin sensitivity check index
R	repaglinide
RCT	randomized controlled trial
RR	risk ratio
S	sulfonylureas
SBP	systolic blood pressure
SCr	serum creatinine
SE	standard error
TC	total cholesterol
TG	triglycerides
U.K.	United Kingdom
U.S.	United States of America
UAE	urinary albumin excretion
U/L	units per liter
ULN	upper limit of normal
V	voglibose
y	years

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
Campbell IW, et al. Diabet Metab 1994;20:394-400.	Open-label, randomized study comparing: Metformin (M) up to 3000 mg/day vs. Glipizide (G) up to 30 mg/day	Diet-failed type 2 DM, age 40-69 y, most subjects were obese Sample size: 48 Mean age (y): M 57 ± 10, G 57 ± 9 Male: 33% Duration of diabetes: M: 2.3 ± 3.4 G: 2.8 ± 3.9	52 weeks	At baseline: M: 11.46 ± 1.92 G: 11.75 ± 2.11 At 52 weeks: M: 8.64 ± 1.21 (↓ 25%) G: 9.72 ± 1.91 (↓ 17%) Over the 52 week period, there was no significant difference between treatment groups at 24 or 36 weeks, however at 52 weeks, the difference was significant in favor of metformin.	FPG (mmol/L): At baseline: M: 11.5 ± 2.76 G: 12.22 ± 3.33 At 52 weeks: M: 7.11 ± 1.28 (↓ 36%) G 9.23 ± 3.69 (↓ 25%) Over the 52 week period, there was a significant difference in FPG at weeks 24, 36, and 52 weeks favoring metformin vs. glipizide. Body weight (kg): At baseline: M: 78.2 ± 15.7 G: 82.2 ± 16.8 Change at 52 weeks (kg): M: ↓ 2.0* G: ↑ 2.6 *Difference in weight became significant beginning at 4 weeks through 52 weeks in favor of	Albumin excretion rates (AER): Treatment with either agent produced a significant improvement in those subjects with elevated AER (> 20 μg/min) with 13/14 patients showing a fall in AER into the normal range. These changes corresponded to the improvement in glycemic control in these patients.	Small sample size Obese patients made up most of sample because study was conducted at a time when metformin was recommended as first-line therapy only in obese type 2 DM patients. Open label design	Metformin, in mainly obese type 2 DM patients, gave better glycemic control over a 52 week period than glipizide. Glipizide improved glycemic parameters in the early part of the study (through ~24 weeks), then the effect appeared to lessen. The effect of metformin was greater in the second 6 month period of the study.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					metformin. Lipid parameters: TC, TG, LDL, HDL, including HDL subfractions showed no significant change over the 52 weeks in either group. Side effects: No glipizide patients had any hypoglycemic episodes. Due to slow titration, no metformin patients reported serious GI side effects.			
Charbonnel BH, et al. Diabet Med 2005;22:399-405. Additional detail published in: **Belcher G**, et al. Diabet Med 2005;22:973-9. **Charbonnel B**, et al. Diabetologia 2005;48:553-60.	Randomized, double-blind, multicenter, non-inferiority study comparing: Pioglitazone (P) up to 45 mg/d vs. Gliclazide* (G) up to 160 mg/d *not marketed in the US	Non-US patients from 18 countries, aged 35-75 y, with type 2 DM inadequately treated with diet, and an HbA1c 7.5-11.0%. Patients who had previously used glucose-lowering pharmacotherapy at any time were excluded from the study. Sample size: 1270 total: P: 624, G 626	52 weeks of treatment consisting of 16-week forced titration period followed by a 36-week maintenance period at maximum tolerated dose	At baseline: P: 8.7 ± 1.0 G: 8.7 ± 1.1 At week 52: P: 7.2% G: 7.3% -0.08; 90% CI (-0.18-0.02); met requirements to declare P non-inferior to	FPG:* Mean change: P: -2.4 G: -2.0; p=0.002 Change as a % value: P: -17% G: -14%; p=NS Total cholesterol:	Fasting plasma insulin (pmol/L): P: -19 G: +15 p < 0.001 C-peptide (nmol/L), Mean change: P: -0.2 G: +0.3 p <0.001 in favor of P	Gliclazide is a second-generation sulfonylurea, however it is not commercial-ly available in the U.S.	Pioglitazone was equivalent to gliclazide in decreasing HbA1c at 52 weeks. There were some differences in lipid profiles and side effects. Pioglitazone

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
Ceriello A, et al. Diabetes Care 2005;28(2):266-72. **Erdmann E**, et al. Int J Cardiol 2006;107:147-153. Funding: Study sponsored by Takeda Europe Research and Development Centre and Eli Lilly and Company.		Mean age (y): P 56 ± 9.5, G 56 ± 9.6 Male: P 61.4%, G 61.7% Duration of diabetes (mean years): P: 2.8 ± 3.8 y, G: 3.0 ± 3.8 at start of study		G % of subjects reaching HbA1c < 7% after 1 y: P: 53% G: 47%; p = NS The response to pioglitazone was maintained at week 52 (Δ from nadir to endpoint < 0.2% compared with 0.4% in the gliclazide group).	P: +3% G: -5%; p<0.001 HDL*: P: + 20% G: + 6%; p< 0.001 Total cholesterol/HDL*: P: - 14% G: - 10%; p<0.001 LDL*: P: + 3% G: -5%; p<0.001 FFA*: P: -0.13 mmol G: -0.03 mmol; p<0.001 Side effects: Edema: P: 8.7% G: 4.5% Hypoglycemia: P: 3.5% G: 10.1% Abnormal LFTs: P: 3 patients (0.5%) G: 10 patients (1.6%) BP: NR	32,33 split pro-insulin levels (pmol/L) Mean change: P: -9 G: no change p < 0.001 in favor of P HOMA-%S At endpoint: P ↑, G ↓; p <0.001 in favor of P --QUICKI At endpoint: P improved more than G; p < 0.001 in favor of P HOMA-%B At endpoint: P small ↑ G ↑; p< 0.001 in favor of P		decreased FPG more than gliclazide. Pioglitazone had a slower onset, but efficacy was maintained at week 52. With gliclazide, there was a deterioration of HbA1c after the nadir point (0.4% change in mean HbA1c from nadir to endpoint vs. < 0.2% with P).

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					Weight: P: +2.8 kg G: +1.9 kg Albuminuria: % with (micro)-albuminuria: At baseline: P: 26.4% G: 22.3% % (micro)-albuminuria cases resolved at week 52: P: 9.9% M: 7.9% % new cases of (micro)-albuminuria at week 52: P: 5.5% M: 6.4% Liver Enzymes: AST (U/L), At baseline: P: 25 ± 12 G: 25 ± 12 AST, Mean change at week 52: P: - 6 ± 31 G: + 5 ± 38 ALT (U/L),At baseline:			

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					P: 32 ± 17 G: 32 ± 17 ALT, Mean change at week 52: P: - 21 ± 32 G + 9 ± 44 GGT (U/L), At baseline: P: 58 ± 5 G: 60 ± 6 GGT, Mean change at week 52: P: -24 ± 44 G: +9 ± 62 Alk Phos (U/L), At baseline: P: 70 ± 22 G: 80 ± 25 Mean change at 52 weeks: P: -12 ± 20 G: +3 ± 20 Bilirubin (µmol/L), At baseline: P: 9 ± 6 G: 10 ± 5 Biliruibin, Mean change at 52 weeks: P: -1 ± 37 G: +13 ± 42			

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					% with LFT values above ULN: At baseline: AST: P: 10.4 G: 11 ALT: P: 23.5 G: 24 GGT: P: 23.1 G: 20.1 Bilirubin: P: 3% G: 3.3% Alk phos: P: 3.6 G: 3.6 At last visit: AST: P: 5 G: 9.4 ALT: P: 8.1 G: 24.5 GGT: P: 12.8 G: 18.8 Bilirubin: P: 2.8 G: 2.9 Alk Phos: P: 2 G: 4.4 * significant difference vs. gliclazide			

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
Derosa, et al. Clin Ther 2003;25:472-84. Funding: not specified	Randomized, placebo-controlled, double blind, single-center study comparing: Repaglinide (R) 1mg/d vs. Glimepiride (G) 1mg/d titrated Mean final doses were: R 2.5 mg/d G 3 mg/d Diet was controlled	Single center in Italy, type 2 DM for at least 6 months receiving no antidiabetic medicine with a HbA1c > 7%, normotensive (< 130/85), no heart disease, normal renal function, LDL > 100 mg/dL, no lipid-lowering agents, no diuretics, beta-blockers, or thyroxine Sample size: 132 Mean age (y): R 56, G 54 Male: 50	4 week placebo washout, then 8 week titration period, then 12 month treatment period	At baseline: R: 8.0 ± 1.1 G 7.8 ± 1.2 At 12 months: R 6.8 ± 0.8* G 6.7 ± 0.9* *p<0.01 vs. baseline	FPG (mg/dL): At baseline: R 158 ± 22 G 164 ± 18 At 12 months: R 120 ± 24* G 125 ± 19* PPG (mg/dL): At baseline: R 194 ± 30 G 188 ± 32 At 12 months: R 148 ± 27* G 167 ± 28* (p<0.05 in favor of R) Lipid profile: There were no significant changes in TC, LDL, HDL, TG, apo A-I, or apo B after 6 or 12 months in either group. Blood pressure: No significant changes in SBP or DBP at 6 or 12 months in either group. Body weight (kg): At baseline: R 76.4 ± 5.2	FPI (µU/mL): At baseline: R 23.1 ± 3.1 G 23.9 ± 2,8 At 12 months: R 18.2 ± 2.9* G 21.4 ± 3.0 Cardiovascular Risk Factors: Lp(a) (mg/dL): At baseline: R 15.4 ± 7.2 G 17.4 ± 9.1 At 12 months: R 11.1 ± 6.8* G 10.5 ± 7.2* PAI-I (ng/mL): At baseline: R 39 ± 20 G 42 ± 17 At 12 months: R 31 ± 18* G 33 ± 15* --Homocysteine (µmol/L): At baseline: R 13.9 ± 3.3 G 14.2 ± 3.7 At 12 months: R 10.6 ± 2.9* G 8.5 ± 2.8* (p<0.05 in favor of G)		Repaglinide and glimepiride both improved glycemic control based on HbA1c, FPG, and PPG. PPG and FPI were significantly better with repaglinide vs. glimepiride. No differences in lipid profiles for either agent. Both significantly reduced Lp (a) [good effect]. Both decreased PAI-I and Hcy (cardiovascular risk factors).

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					G 77.1 ± 5.9 BMI (kg/m2) At baseline: R 26.1 ± 1.2 G 26.4 ± 1.0 *p<0.05 vs. baseline			
Derosa G, et al. Diabetes Research and Clinical Practice 2003;60:161-169. Funding: not specified	Open-label, randomized, single-center study comparing: Repaglinide (R) 2-4 mg/day vs. Metformin (M) 1500-2500 mg/day titrated Mean final doses were: R 3 ± 1 mg/day M 2000 ± 500 mg/day Diet was controlled (1400-1600 kcal/day; 55% carbohydrates, 25% proteins, 22% lipids [7% saturated], 105 mg cholesterol and 36 g fiber) and were asked to perform aerobic activity for at least 30 minutes 3-4 times per week.	Single outpatient clinic in Italy, type 2 DM for at least 6 months who had not previously received oral hypoglycemic agents, with a HbA1c > 7%, normotensive (< 130/85), no heart disease, normal renal function, LDL > 100 mg/dL. Lipid-lowering medications, if present, were stopped, and patients went through a 4 week placebo washout period. No patients were taking other medications that may have influenced cardiovascular risk parameters. Sample size: 112 Mean age (y): R 55 ± 10, M 52 ± 9 Male: 50 Duration of diabetes (y): R 4 ± 2, M 5 ± 2	4 week placebo-run in, followed by an 8 week dose titration period, followed by a 12 month treatment period Doses were titrated in order to achieve the following targets: FPG < 6.7 mmol/L 2 h PPG < 8.8 mmol/L	At baseline (end of washout period): R: 7.6 ± 0.9 M: 7.4 ± 0.9 At 12 months: R: ↓** M: ↓** ** p< 0.01 vs. baseline	FPG (mmol/L): At baseline: R: 8.5 ±1.33 M: 8.2 ± 1 At 12 months: R: 7.7 ± 1.22* M: 7.6 ± 0.94* % met FPG target < 6.7 mmol/L): R: 37/53 (69.8%) M: 42/49 (87.7%) PPG (mmol/L): At baseline: R: 10.2 ± 1.94 M: 10.8 ± 1.67 At 12 months: R: 8.6 ± 1.67*† M: 9.6 ± 1.55* Lipid parameters: TC (mmol/L): At baseline: R: 4.97 ± 0.98 M: 5.28 ± 0.80	FSI (pmol/L): At baseline: R: 190.3 ± 24.3 M: 179.2 ± 18.1 At 12 months: R: ↓* M: ↓*† Postprandial serum insulin (PSI): At baseline: R: 382 ± 59 M: 361.1 ± 63.9 At 12 months: R: ↔ M: ↓*† Lp(a)(µmol/L): At baseline: R: 0.84 ± 0.51 M: 0.74 ± 0.47 At 12 months: R: -0.18*† M: - 0.05 Apo A-I (g/L): At baseline: R: 1.02 ± 0.16	Open label design	The use of either repaglinide or metformin in therapy-naïve patients with type 2 DM is associated with a positive effect on both glycemic control and overall cardiovascular risk profile. Effects on FPG and HbA1c were similar. PPG was decreased more by repaglinide. Both medications reduced FSI from baseline. Only metformin reduced PSI.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					At 12 months: R: ↓ 0.34 M: ↓ 0.54* LDL(mmol/L): At baseline: R: 3.26 ± 0.65 M: 3.39 ± 0.67 At 12 months: R: ↓ 0.31 M: ↓ 0.39* HDL(mmol/L) At baseline: R: 1.09 ± 0.18 M: 1.19 ± 0.21 At 12 months: R: ↑ 0.08 M: ↓ 0.03 TG (mmol/L): At baseline: R: 1.76 ± 0.59 M: 1.98 ± 0.54 At 12 months: R: ↓ 0.18 M: ↓ 0.27* Mean values for BP (SBP and DBP), BMI, and weight did not change significantly from baseline to the end of	M: 1.19 ± 0.20 At 12 months: R: ↑ 0.15 M: ↓ 0.04 Apo B (g/L): At baseline: R: 1.26 ± 0.18 M: 1.35 ± 0.22 At 12 months: R: ↓ 0.17 M: ↓ 0.13 PAI-1 (ng/mL) At baseline: R: 43.6 ± 22.8 M: 40.2 ± 15.3 At 12 months: R: ↓ 9.4* M: ↓ 11.5** *p<0.05 vs. baseline † p <0.05 vs. M		

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					the 12 month treatment period. *p<0.05 vs. baseline † p <0.05 vs. M % met PPG target < 8.9 mmol/L): R: 38/49 (77.5%) M: 47/53 (88.6%) Side effects: Hypoglycemia R: 0 M: 0 Unsatisfactory glycemic control, causing patients to drop out of study: R: 3 M: 4 Nausea and diarrhea caused 2 patients in the metformin group to drop out of the study			
Derosa G, et al.	Randomized,	Italy, type 2 DM patients	8 week	At baseline:	FPG (mg/dL):	Lp(a) (mg/dL):	Not masked	Both G and M

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
Diab Nutr Metab 2004;17:143-150. Funding: not specified	controlled, multicenter, open-label study comparing: Glimepiride (G) 1-4 mg/day vs. Metformin (M) 1-3 g/day Mean final dose: G: 3 ± 1 mg/day M: 2.5 ± 0.5 g/day Patients were instructed to follow a 1400-1600 kcal/day diet [55% carb., 25% protein, 22% lipids] and to undertake physical aerobic activity (30 minutes at least 3-4 times week.	aged 46-67 y diagnosed in the past 6 months. Patients had to be normotensive (BP < 130/85), non-smokers, no coronary heart disease, and normal renal function (CrCL < 1.5 mg/dL). No. of subjects: 164 Male: 80/164 (%) Duration of diabetes: < 6 months	titration period, followed by a 12 month treatment period	G: 8.5 ± 1.2 M: 8.4 ± 1 At 12 months: G: 6.9 ± 0.7 p < 0.01 M: 7 ± 0.9 p < 0.01; p=NS for G vs. M Patients who met target < 7%: G: 61/73 (83.5%) M: 64/75 (85.3%)	Patients who met target < 120: G: 60/73 (82.1%) M: 65/75 (86.6%) At baseline: G: 165 ± 20 M: 174 ± 15 At 12 months: G: 123 ± 25 M: 125 ± 15; both p< 0.01 vs. baseline PPG (mg/dL): Patients who met target < 160: G: 58/73 (79.4%) M: 62/75 (82.6%) At baseline: G: 189 ± 33 M: 192 ± 28 At 12 months: G: 162 ± 26 p < 0.01 M: 166 ± 28 p < 0.01; p=NS for G vs. M LDL At baseline:	At baseline: G: 44.5 ± 21.3 M: 48.3 ± 16.8 At 12 months: G: 38.8 ± 13.2 p < 0.01 M: 43 ± 14.6 p < 0.05; p =NS for G vs. M PAI-1 At baseline: G: 38 ± 21 M: 41 ± 16 At 12 months: G: 30 ± 18, M: 31 ± 14, both p < 0.05 vs. baseline Homocysteine: At baseline: G: 13.4 ± 3.6 M: 14.4 ± 3.1 At 12 months: G: 9.7 ± 2.6 p < 0.01 M: 12.8 ± 2.9 p=NS; p < 0.05 in favor of G There was no statistically significant variation in Apo, A-I, Apo B, or	Run-in titration period XX efficacy trial	were well tolerated and improved glycemic control as evidenced by HbA1c, FPG, and PPG. M lowered FPI, PPI, and LDL. G and M lowered Lp (a). PAI-1, G lowered homocysteine Overall: Both agents improved glycemic control and cardiovascular risk factors.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					G 135 ± 20 M 144 ± 20 At 12 months: G 130 ± 15, p=NS M 130 ± 25, p < 0.05 vs baseline There was no statistically significant variation in BMI, SBP, DBP, TC, HDL, or TG. No patients in either group experienced signs or symptoms of mild or severe hypoglycemia.	fibrinogen.		
Donnelly LA, et al. Diabetic Medicine 2006;23:128-33. Funding: Authors are supported by a Wellcome Trust Clinical Training Fellowship and a Wellcome Trust Research Leave Fellowship.	Prospective, observational study comparing: Sulfonylurea (S) vs. Metformin (M)	Tayside, Scotland; all type 2 DM patients registered with Tayside General Practitioners [Part I] with an initial prescription for an oral hypoglycemic agent between 1/1/94 – 2/28/02. Part II evaluated response to treatment stratified by BMI in a subset of patients with both a baseline and treatment HbA1C. [I] Sample size: 5049 (S 3053, M 1996)	[Part II] Mean duration of diabetes: M: 2.5 ± 3.4 y S: 2.9 ± 4.3 y	[Part II] BMI did not affect HbA1c response to treatment with sulfonylureas. A decreased BMI was correlated with a greater response to metformin. However the effect was small and		Prescribing patterns during the period 1/1/94 and 2/28/02: Metformin prescribed in an increasing percentage of patients from 28% in 1994 to 56% in 2001. Subgroup Analysis: Prescribing	Large numbers of patients (~50%) excluded from treatment response analysis due to lack of either baseline HbA1c or post-treatment HbA1c (after	Glycemic response to sulfonylureas was not influenced by BMI. Slight negative correlation between BMI and glycemic response to metformin. A patient's BMI should not

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
		Mean age: n/a Males: n/a BMI: n/a 98.5% Caucasian [II] Sample size: 2064 patients (S 1083, M 981) Mean age: M 59.5 ± 11.7 y, S 66.5 ± 11.3 y Male: M 52%, S 56.2% BMI: M 33.1 ± 5.9 kg/m^2 S 27.7 ± 4.5 kg/m^2		statistically not significant: BMI: HbA1c reduction in metformin group was: In obese subgroup: –1.46 (95% CI 1.34-1.57) vs. In non-obese subgroup: -1.34 (95%CI 1.25-1.42); p=0.11		patterns compared with patient weight during the period 1/1/01 and 2/28/02 (S 2155 patients, M 1701 patients): <u>Obese</u>: M 62.1%, S 37.9% <u>Overweight (BMI > 25 kg/m2)</u>: M 33.6% S 64.4% <u>Normal (BMI < 25 kg/m2)</u>: M 13% S 87%	3-12 months of therapy). Number of patients with normal weight was small. Additional data comparing metformin and sulfonylurea in normal weight patients is needed to confirm results.	influence treatment choice when selecting between metformin and a sulfonylurea.
Eurich DT, et al. Diabetes Care 2005;28:2345-51. Funding: Alberta Heritage Foundation for Medical Research, Canadian Institutes for Health Research, Canadian Diabetes Association, the Heart and Stroke Foundation of Canada, and the Kidney Foundation of Canada.	Retrospective cohort study using administrative data from an electronic prescription database comparing: Sulfonylureas (S) vs. Metformin (M) vs. Combination of sulfonylurea + metformin (C)	Canada, 12,272 outpatients enrolled in health plan > 1 year, age > 29 y, new users of antidiabetic agents (first claim for oral antidiabetic agent with no claim for least 1 year prior), between 1/1/91 and 12/31/96. Patients had to have a hospital stay or physician service for heart failure between 12/1/91 and 12/31/99. Patients with heart failure in the 3 years before beginning an oral antidiabetic agent or those who ever had insulin were excluded.	Followed from index date to until death, termination of coverage, or 12/31/99. Mean duration of follow-up was 2.5 ± 2.0 years		Primary outcome was all-cause mortality, both at 1 year (i.e., short-term) and at the end of the follow-up period (i.e., long term). The sulfonylurea monotherapy group served as the reference group for all hazard ratio [HR (95%CI)] estimates. All-cause mortality at 1 year: S 200 deaths (26%) HR=1.0 (reference group) M 29 deaths (14%) HR=0.52 (0.35-0.76) unadjusted HR=0.66 (0.44-0.97) adjusted* C 97 deaths (11%) HR=0.41 (0.32-0.52) unadjusted		Observation-al study design; therefore, groups not randomized. Administra-tive database, therefore could not control for level of glycemic control, BMI, severity of heart failure, or other	In newly treated diabetic patients with heart failure, when compared with sulfonylurea therapy, --fewer deaths occurred with metformin alone (33%) or in combination with a sulfonylurea (31%) than with sulfonylurea

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
		Saskatchewan Health Database Sample size: 1,183 patients with heart failure who were also users of antidiabetic agents, including S 773 (42%), M 208 (11%), and C 852 (47%) Mean age (y): 72 ± 10.7 Male: 57% Race: not available Comorbidities: chronic disease score (CDS) median for S = 10, M = 11, C = 11; CDS mean for S = 10.7 ± 3.7, M 11.6 ± 3.6, C 11.7 ± 3.7			HR=0.54 (0.42-0.70) adjusted* All-cause mortality at the end of follow-up (mean 2.5 y, median 2.1 y): S 404 deaths (52%) HR=1.0 (reference group) M 69 deaths (33%) HR=0.63 (0.49-0.82) unadjusted HR=0.70 (0.54-0.91) adjusted* C 263 deaths (31%) HR=0.50 (0.43-0.58) unadjusted HR=0.61 (0.52-0.72) adjusted* Secondary outcomes were all-cause hospitalizations at 1 year and at the end of the follow-up period, as well as a composite outcome of all-cause hospitalizations or all-cause mortality. All-cause hospitalizations at 1 year: S 406 hospitalizations (53%) HR=1.0 (reference group) M 102 hospitalizations (49%) HR=0.52 (0.35-0.76) unadjusted HR=0.84 (0.67-1.04) adjusted* C 435 hospitalizations (51%) HR=0.41 (0.32-0.52) unadjusted HR=0.92 (0.80-1.06) adjusted* All-cause hospitalizations at the end of follow-up (mean 2.5 y, median 2.1 y): S 538 hospitalizations (70%) HR=1.0 (reference group) M 143 hospitalizations (69%) HR=0.87 (0.73-1.05) adjusted*		cardiovascu-lar risk factors. S group was slightly older, had fewer comorbidi-ties, and had fewer prescription claims for heart-failure related medications compared to the M and C groups. Database cannot distinguish between systolic and diastolic dysfunction heart failure, some patients labeled as having CHF may have had diastolic dysfunction. Residual potential for confounding by indication remains-clinicians	monotherapy (52%). --a reduction in deaths or hospitaliza-tions was also observed. --there was no difference in time to first hospitalization among the groups. --there were no hospitaliza-tions or deaths in any of the cohorts attributed to metabolic acidosis throughout the follow-up.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					C 632 hospitalizations (74%) HR=0.93 (0.83-1.05) adjusted* Composite endpoint (all-cause mortality or all-cause hospitalizations) at 1 year: S 480 events (63%) HR=1.0 (reference group) M 115 events (55%) HR=0.80 (0.65-0.98) unadjusted HR=0.79 (0.65-0.98) adjusted* C 480 events (56%) HR=0.82 (0.72-0.93) unadjusted HR=0.86 (0.75-0.98) adjusted* Composite endpoint (all-cause mortality or all-cause hospitalizations) at the end of follow-up (mean 2.5 y, median 2.1 y): S 658 events (85%) HR=1.0 (reference group) M 160 hospitalizations (77%) HR=0.84 (0.71-1.00) unadjusted HR=0.83 (0.70-0.99) adjusted* C 681 events (80%) HR=0.83 (0.75-0.93) unadjusted HR=0.86 (0.77-0.96) adjusted* *adjusted for age, sex, modified chronic disease score, therapies known to affect heart failure outcomes (i.e., ACE inhibitors, angiotensin II blockers, beta-blockers, antiplatelet agents, nitrates, lipid-lowering agents, antiarrhythmic agents, and spironolactone), and total physician visits before heart failure diagnosis.		may have prescribed agents alternative to metformin in frail patients, those with elevated creatinine, etc.	
Eurich, et al.	Retrospective,	Canada, outpatients	Followed from		Primary	Secondary	Administra-	Metformin may

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
Pharmacotherapy 2005;25(6):810-6. Funding: Supported in part by unrestricted grants from the Institute of Health Economics, Edmonton, Alberta, Canada; Eli Lilly, Canada; and the Alberta Heritage Foundation for Medical Research, Edmonton, Alberta, Canada.	cohort study using data from an electronic prescription database comparing: Sulfonylurea (S) vs. Metformin (M)	enrolled in health plan > 1 year, age > 29 y, new users of antidiabetic agents (first claim for oral antidiabetic agent with no claim for least 1 year prior), between 1/1/91 and 12/31/96. Patients had to continue on monotherapy x 2 years, Saskatchewan Health Sample size: 6,729 subjects (S 5,077, M 1,652) Mean age (y): 63.8 ± 12.7 Male: 56.3% Race: n/a Comorbidities: chronic disease score median = 8	index date to until death, termination of coverage, or 12/31/99. Mean duration of follow-up (y): S 5.83 ± 1.8 M 5.29 ± 1.6		endpoint: Time to secondary failure after 2 years of monotherapy. Failure defined as addition of a second agent or a switch to a different agent. At 2 years: Primary endpoint of secondary failure reached: S 2377/5077 (46.8%) vs. M 627/1652 (38%) Metformin monotherapy was associated with a delay in the onset of secondary failure: unadjusted HR = 0.93 (95% CI 0.85-1.02; p > 0.05) or adjusted HR = 0.89 (95%CI 0.82-0.98; p=0.025) after adjustment for	endpoints: Time to combination therapy: More patients receiving S than M started combination therapy: S 39.8% M 29.6%; (HR 0.79; p < 0.001) However, fewer patients in the S group had a switch in oral therapies: S 7.0% M 8.4%; (Adjusted HR = 1.43, p < 0.001) Time to addition of insulin: More patients in the S group switched to insulin: S 9.1% M 5%; adjusted HR 0.65; p < 0.001	tive claims data only (no clinical data such as glycemic control) No weight or BMI information, Limited applicability to practice since the study sample all were on an agent for 2+ years	delay onset of secondary failure compared to oral sulfonylurea in patients stable on agents for 2 years.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					age, sex, adherence and chronic disease score. On the adjusted Kaplan Meier curve for time to secondary failure, the lines become significantly different at year 4 (p = 0.035) and continue to diverge throughout follow-up.			
Goke B, et al. Treat Endocrinol 2002;1(5):329-336.	Open-label, randomized study comparing: Pioglitazone (P) 45 mg/day vs. Acarbose (A) up to 300 mg/day All patients were asked to follow a disease and body weight-oriented diet. Compliance was assessed by pill counts	Germany, 47 centers, patients with either newly diagnosed type 2 DM (~50%) or had previous treatment with oral antihyperglycemic agents, patients in the latter category had to stop the oral agent for at least 2 months prior to starting the study, patients also had to have an HbA1c between 7.5-11.5%, FPG≥ 140 mg/dL, and BMI between 25-43 kg/m^2. Sample size: P 129, A 136 Mean age (y): P 58.9 ± 9.1, A 58.8 ± 9.1 Male: P 53.5%, A 54.5% --% smokers: P 17%, A	1 week diet run-in period, followed by a 26 week treatment period At the end of 26 weeks, patients taking pioglitazone could continue to week 64 on pioglitazone monotherapy. Patients taking acarbose could start	All patients: At baseline: P 8.98 ± 1.2 A 9.03 ± 1.32 At 26 weeks: P 7.82 ± 1.95* A 8.55 ± 1.96 Treatment-naïve patients: At baseline: P 8.99 ± 1.26 A 8.85 ± 1.22 At 26 weeks: P 7.27 ± 1.75* A 7.94 ± 1.85 Previously-treated	Other parameters were reported, but the results for treatment-naïve patients were not reported separately.		Only 50% of the subjects in the study were treatment-naïve. The results for this subgroup were only reported for HbA1c. Open label study Significant crossover from acarbose to pioglitazone	Pioglitazone reduced HbA1c levels more than acarbose over 6 months in treatment-naïve type 2 DM patients.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
		19.1% Duration of diabetes (months): P 57 ± 55.4, A 59.1 ± 50.3 --Number (%) of treatment-naïve type 2 DM patients: P 69/129 (53.5%) A 71/136 (52.2%)	taking pioglitazone in addition to the acarbose for an additional 38 weeks (up to week 64).	patients: At baseline: P 8.98 ± 1.26 A 9.23 ± 1.40 At 26 weeks: P 8.46 ± 1.99** A 9.21 ± 1.88 Proportion of patients achieving HbA1c targets at week 26: HbA1c ≤ 6.5% P: 31%* A: 13% HbA1c ≤ 7.0% or a ↓ in HbA1c ≥ 0.6% or a ↓ in FPG ≥ 30 mg/dL P: 81%* A: 57% Comments: HbA1c values showed the most improvement in patients who were treatment-naïve. During the extension period to week 64, there was little further			(n=24)	

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
				decrease in HbA1c for patients continuing on pioglitazone (↓ 0.06 ± 0.69%). *p< 0.001 vs. acarbose **p< 0.009 vs. acarbose				
Hällsten K, et al. Diabetes 2002;51:3479-3485. Funding: Supported by grants from the Academy of Finland, the Novo Nordisk Foundation, the Finnish Diabetes Research Society, and GlaxoSmithKline. Additional detail published in: **Virtanen KA**, et al. Diabetes 2003;52:283-290. Funding: In addition to the above organizations, the Finnish Cultural Foundation, the Research and	Randomized, double-blind, placebo-controlled trial Rosiglitazone (R) up to 8 mg/day vs. Metformin (M) up to 2 g/day vs. Placebo	Finland; newly diagnosed or diet-controlled patients with type 2 DM and no diabetes complications, fasting plasma glucose between 6.1 and 11 mmol/L after run-in period, and BP < 160/100 mmHg, normal hepatic and renal function. -Sample size: 45 patients -Mean age: R 58.6 ± 2, M 57.8 ± 2.2, P 57.7 ± 1.9 --% Male: 28/41 --Race/Ethnicity: Finnish Duration of diabetes: n/a	4 week run-in period, on diet therapy, followed by 26 week treatment period	At baseline (mean ± SE): R 6.8 ± 0.2 M 6.9 ± 0.2 P 6.3 ± 0.1 At week 26 (mean ± SE): R 6.5 ± 0.2* M 6.2 ± 0.2* P 6.1 ± 0.1 * p < 0.05 vs. baseline; p = NS for R vs. M	FPG (mmol/L) at baseline: R 7.2 ± 0.3 M 8.0 ± 0.5 P 7.2 ± 0.3 At week 26: R 6.8 ± 0.3 M 6.8 ± 0.3* P 7.2 ± 0.3 *p < 0.001 vs. baseline; p = NS for R vs. M FFA (µmol/L) At baseline: R 595 ± 46 M 511 ± 63 P 607 ± 56 At week 26: R 512 ± 61 M 510 ± 47 P 519 ± 45 Mean body weight (kg): At baseline: R 83.7 ± 2.1	Patient compliance with treatment was > 95% (capsule count). FSI (mU/L) At baseline: R 8.6 ± 1.5 M 11.7 ± 2.1 P 10.1 ± 1.3 At week 26: R 6.6 ± 0.4 M 8.8 ± 1.1 P 9.6 ± 0.9 C-peptide (nmol/L) At baseline: R 0.78 ± 0.07 M 0.89 ± 0.10 P 0.86 ± 0.07 At week 26: R 0.58 ± 0.04 M 0.65 ± 0.07* P 0.71 ± 0.04* * P < 0.05 vs	Small study with primary objective to measure whole-body insulin sensitivity and skeletal muscle glucose uptake by measures not used in routine clinical practice and of uncertain clinical significance (PET, euglycemic clamp technique, cycle ergometer measurements, etc)	Glycemic control similarly improved with both rosiglitazone and metformin at 26 weeks. Whole-body insulin sensitivity and glucose uptake in skeletal muscle were unchanged by metformin, but improved by rosiglitazone (44% and 38%, respectively). In addition, rosiglitazone doubled the insulin-stimulated glucose uptake rate during physical

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
Science Foundation of Farmos, the Swedish Research Council, and the Swedish Diabetes Association also supported this study.					M 88.8 ± 3.0 P 88.3 ± 2.5 At week 26: R 84.3 ± 2.4 (↔) M 86.8 ± 2.9 (↓ 2 kg; p < 0.05 vs. baseline) P 88.4 ± 2.5 (↔) Blood pressure (mmHg): SBP At baseline: R 152 ± 5.0 M 145 ± 4.1 P 147.2 ± 3.2 At 26 weeks: R 149 ± 4.5 M 141.8 ± 4.0 P 144.4 ± 3.8 DBP: At baseline: R 90.5 ± 2.1 M 91.4 ± 2.5 P 85.1 ± 2.3 At 26 weeks: R 84.2 ± 2.4 M 85.5 ± 2.6 P 85.4 ± 2.7	baseline		exercise, while there was no change with metformin.
Horton E, et al. Curr Med Res Opinion 2004;20(6):883-89.	Randomized, double-blind, multicenter, placebo-controlled study comparing:	U.S. and U.K., patients aged ≥ 30 years with type 2 DM, BMI 20-35 kg/m2, HbA1c 6.8-11%, FPG ≤ 15 mmol/L, and not previously	24 weeks	At baseline: N 8.1 ± 0.1 M 8.3 ± 0.1 C 8.2 ± 0.1 P 8.2 ± 0.1	FPG (mmol/L) At baseline: N 10.2 ± 0.2 M 10.0 ± 0.2 C 10.3 ± 0.2	Liquid meal challenges: 30 minute post-challenge insulin increment	Subset analysis	Combination nateglinide/metformin is an effective and well-tolerated

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
Funding: Study supported by a grant from Novartis Pharmaceuticals.	Nateglinide (N) 120 mg tid before meals vs. Metformin (M) 500 mg tid with meals vs. Nateglinide + Metformin (C) vs. Placebo (P)	treated with an antidiabetic medication. This analysis is a subset of a larger RCT, the medication naïve patients were not randomized on this factor. Sample size: N 104, M 104, C 89, P 104 Mean age (y): N 57.9, M 55.4, C 57.7, P 59 Male: N 56.7%, M 67.3%, C 65.2%, P 64.4% Duration of diabetes (y): N 4.7 ± 0.6, M 3.7 ± 0.4, C 3.4 ± 0.4, P 4.2 ± 0.4 BMI (kg/m^2): N 29.9, M 29.9, C 30.6, P 29.5		Mean change at endpoint: N ↓ 0.8 ± 0.1* M ↓ 0.8 ± 0.1* C ↓ 1.6 ± 0.1* P ↑ 0.3 ± 0.1, p <0.05 vs. baseline *p <0.001 vs. both baseline and placebo % patients reaching HbA1c < 7.0%: N 34% M 41% C 70% P 17%	P 10.4 ± 0.2 Mean Δ at endpoint: N ↓ 1.1 ± 0.3 (↓ 11%)* M ↓ 1.2 ± 0.3 (↓ 12%)* C ↓ 2.3 ± 0.3, (↓22%)* P no change *p <0.001 vs. both baseline and placebo Mean Δ in body weight: N not reported M not reported C ↑ 0.2 ± 0.4 kg P ↓ 0.2 ± 0.4 kg % experiencing gastrointestinal side effects: N 16.3% M 27.9% C 27% P 14.4%	(pmol/L): At baseline: N 178 ± 14 M 149 ± 12 C 186 ± 21 P 160 ± 16 Mean change at endpoint: N ↑ 164 ± 26 M no change C ↑ 88 ± 32 P no change --2 hour post-challenge post-prandial glucose excursion (mmol/L): At baseline: N 2.5 ± 0.2 M 2.3 ± 0.2 C 2.6 ± 0.2 P 2.1 ± 0.2 Mean change at endpoint: N ↓ 1.9 ± 0.2* M ↓ 1.0 ± 0.2** C ↓ 2.3 ± 0.2* P ↓ 0.5 ± 0.2 *p <0.001 vs. both baseline and placebo **p <0.001 vs. baseline only		approach for treatment-naïve type 2 DM patients not controlled by diet and exercise. The agents have a similar effect on HgbA1c when used singly.
Johnson JA, et al. Diabetes Care 2002;25:2244-8. Funding: Institute	Retrospective cohort study using administrative data from an electronic prescription	Canada, 12,272 outpatients enrolled in health plan > 1 year, age > 29 y, new users of antidiabetic agents (first	Followed from index date to until death, termination of coverage, or		Number of deaths (all-cause mortality): S 750 (24.7%) M 159 (13.8%) C 635 (13.6%)		The sulfonylurea group was older, on average,	Metformin monotherapy or in combination with a

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
of Health Economics and Alberta Heritage Foundation for Medical Research See related publication	database comparing: Sulfonylureas (S) vs. Metformin (M) vs. Combination of sulfonylurea + metformin (C)	claim for oral antidiabetic agent with no claim for least 1 year prior), between 1/1/91 and 12/31/96. Patients had to receive the S, M, or C therapy for at least 1 year during the follow-up and not have received insulin to be included in the study; Saskatchewan Health Database Sample size: S 3,033, M 1,150, C 4,683 Mean age (y): 64.1 ± 13.0 Male: 55.9% Race: not available Comorbidities: chronic disease score (CDS) median = 8, CDS mean = 8.5 ± 4.1	12/31/99. Mean duration of follow-up (y): 5.1 ± 2.2		Crude OR (95% CI): S 1.0 (comparator group) M 0.49 (0.41-0.59) C 0.49 (0.43-0.54) Multivariate logistic regression analysis: Adjusted OR*: M 0.60 (0.49-0.74) C 0.66 (0.58-0.75) Multivariate survival analysis: Adjusted RR: M 0.78 (0.65-0.92) C 0.63 (0.57-0.71) _____ Number of cardiovascular-related deaths: S 351 (11.6%) M 80 (7%) C 299 (6.4%) Crude OR (95% CI): S 1.0 (comparator group) M 0.53 (0.41-0.68) C 0.48 (0.41-0.57) Multivariate logistic regression analysis Adjusted OR*: M 0.64 (0.49-0.84) C 0.64 (0.54-0.77) Multivariate survival analysis: Adjusted RR: M 0.84 (0.66-1.07) C 0.63 (0.54-0.74) _____ * adjusted for age, sex, CDS,		and included more men. The metformin group had fewer nitrate users. Patients not randomized. Administra-tive database, therefore could not control for level of glycemic control, BMI, or other cardiovascu-lar risk factors.	sulfonylurea, was associated with reduced all-cause and cardiovascular mortality compared with sulfonylurea monotherapy among new users of these agents.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					nitrate use			
Johnson JA, et al. Diabetes Med 2005;22:497-502. Funding: Institute of Health Economics and Alberta Heritage Foundation for Medical Research See related publication **Johnson JA**, et al. Diabetes Care 2002;25:2244-8.	Retrospective cohort study using administrative data from an electronic prescription database comparing: Sulfonylureas (S) vs. Metformin (M) vs. Combination of sulfonylurea + metformin (C)	Canada, 12,188 outpatients enrolled in health plan > 1 year, age > 29 y, new users of antidiabetic agents (first claim for oral antidiabetic agent with no claim for least 1 year prior), between 1/1/91 and 12/31/96. Patients had to receive the minimum recommended daily of doses of S, M, or C therapy for at least 6 months during the follow-up period and not have received insulin to be included in the study; Saskatchewan Health Database Sample size: 5,720, including S 2138, M 923, C 2641 (note: 1560 of these combination patients were later excluded because they used both drug concurrently for less than 50% of the follow-up period) Mean age (y): S 67.8 ± 12.4, M 64.3 ± 12.4, C 62 ± 12.3 Male: S 59%, M 52%, C 54% Race: not available Comorbidities: chronic disease score (CDS) median for S = 8, M = 8, C = 9; CDS mean for S = 8.4 ± 4.1, M 8.0 ± 3.9, C 8.9 ±	Followed from index date to until death, termination of coverage, or 12/31/99. Mean duration of follow-up (y): S 4.7 ± 1.9 M 4.6 ± 1.6 C 5.6 ± 1.9 Median duration of follow-up (y): S 4.5 M 4.3 C 5.5		There were 381 deaths from cardiovascular causes and 715 hospitalizations at least once for cardiovascular reasons. The composite rates of cardiovascular mortality and hospitalizations were as follows: S 54.2 events/1000 patient-years M 37.4 events/1000 patient-years C 43.6 events/1000 patient-years The cardiovascular mortality rates were as follows: S 25.5 deaths/1000 patient-years M 14.4 deaths/1000 patient-years C 90.2 deaths/1000 patients-years The non-fatal hospitalization rates were as follows: S 75.3 hospitalizations/1000 person-years M 53.7 hospitalizations/1000 person-years C 90.2 hospitalizations/1000 person-years _____ Primary outcome was the composite endpoint of fatal or non-fatal cardiovascular-related events [HR (95%CI)]: Unadjusted analyses: S 1.0 [comparator group] M 0.67 (0.56-0.80) C 0.80 (0.69-0.93) Adjusted analyses:* M 0.81 (0.68-0.97)		Patients not randomized. Administra-tive database, therefore could not control for level of glycemic control, BMI, or other cardiovascu-lar risk factors	When compared with sulfonylurea monotherapy, Metformin monotherapy was associated with a lower risk of cardiovascular-related morbidity and mortality. Combination metformin and sulfonylurea therapy was associated with a reduced risk of fatal cardiovascular events.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
		4.0			C 0.97 (0.84-1.13) Secondary outcomes: Cardiovascular mortality: Unadjusted analyses: M 0.58 (0.44-0.76) C 0.41 (0.31-0.53) Adjusted analyses:* M 0.75 (0.57-1.00) C 0.61 (0.46-0.80) _____ Nonfatal cardiovascular hospitalizations: Unadjusted analyses: M 0.76 (0.61-0.93) C 1.09 (0.92-1.29) Adjusted analyses: M 0.78 (0.63-0.97) C 1.05 (0.89-1.25 *adjusted for age, sex, CDS, nitrate use			
Lusignan S, et al. BMC Family Practice 2005;6(13) Available at www.biomedcentral.com/1471-2296/6/13 Funding: Study funded by a grant from Doctors' Independent Network (DIN) and	Retrospective cohort Computerized primary care database During the period between 1994-2001, the proportion of patients being treated with: diet alone: ↓ from	Computerized U.K. primary care database; selected cohort of 74 (out of 142) family practices considered to be good quality data providers with continuous recording from 1994 - 2001; algorithm used to identify within the database newly diagnosed type 2 DM patients		Between 1997 and 2001, % of patients meeting NICE targets for HbA1c of: < 6.5%: ↓ from 28.9% to 22.5% < 7.5% ↓ from 53.7% to 52.3%	Between 1994 and 2001, % of patient meeting NICE targets for: BMI < 25 kg/m^2 : fell from 27% to 19.4% BP < 140/80: increased to 22.5% BP < 160/100	Prevalence of type 2 DM is increasing: 1994: M 18/1000, F 16/1000 2001: M 27/1000, F 23/1000. Between 1994 and 2001, recording of (presented as	Author's comment on need to distinguish between changes in BMI as a result of DM therapies vs. the increasing BMI of the general population.	Prevalence of type 2 DM continues to increase. Measurements and electronic recording of data of value in the management of DM is increasing.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
supported by a Wellcome Trust Grant.	38.2% to 33%, any insulin ↑ from 13.2% to 15.1%, Oral agents only ↑ from 48.6% to 51.8%. Of oral agents, use of: ultra-long acting sulfonylureas fell to 0.1% Short-acting sulfonylureas ↑ from 23.3% to 35.1%. Metformin ↑ from 22.6% to 38.9%. Alpha-glucosidase inhibitors ↑ 1.7% to 2.4%.			(target of 6.5% vs. 7.5% dependent on CAD risk)	mmHg: *not reported* (target depends on CAD risk) TC < 5 mmol/L: increased to 46.2% Smoking reduction: *not reported*	practice medians): HbA1c ↑ from 34% to 74%. TC ↑ 17% to 61% BMI ↑ 43% to 55% BP ↑ 65% to 82%	Their study was unable to do this. Accuracy of records could affect these results.	Glycemic control has not improved. Author's did an analysis controlling for BMI and found that BMI accounted for little of the change in HbA1c levels over time.
Nakamura, et al. J Diabetes Complications 2000;14(5):250-4. Funding: not specified	Randomized trial comparing: Pioglitazone (P) 30 mg/d vs. Glibenclamide* (G) 5 mg/d vs. Voglibose (V) 0.6 mg/d *same as glyburide in US	Japan, normotensive patients with type 2 DM and microalbuminuria being treated by diet alone with fasting C-peptide > 0.33 mmol/L and HbA1c > 6.5%. 30 healthy controls were used for comparison. Sample size: 45 DM patients (15 each in P, G, V groups) and 30 healthy controls Mean age (y): P 60 ± 13 G 61 ± 10 M 56 ± 12 Male: 50% Comorbidities: normotensive patients <	3 months	At baseline: P 7.7 ± 1.2 G 7.8 ± 1.1 V 7.6 ± 1.1 At 3 months: P 6.8 ± 1.1* G 6.9 ± 1.2* V 6.8 ± 1.1* * p <0.05 vs. baseline	BP At baseline (mmHg): Systolic P 122 ± 17 G 122 ± 18 V 118 ± 16 Diastolic P 74 ± 14 G 78 ± 14 V 78 ± 12 At 3 months (mmHg): Systolic P 116 ± 15 G 124 ± 16 V 122 ± 18	--UAE At baseline (μg/min): DM: 156.2 ± 42.8 Controls: 8.2 ± 2.6 At 3 months: P decreased from 142.8 ± 42.4 to 48.4 ± 18.2, p < 0.01 G change NS V change NS --urinary ET-1 (ng/g UC): At baseline: DM: 8.7 ± 1.3	Small sample size The main outcome, urinary endothelin, is a biomarker for microalbu-min, which is a biomarker for renal insufficiency in DM.	At baseline, UAE and urinary ET-1 levels were higher in DM patients than in the 30 healthy controls, but plasma ET-1 levels did not differ. Only pioglitazone reduced UAE and urinary ET-1 levels in DM patients. HbA1c was

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
		140/90 mmHg Duration of diabetes (y): P 16 ± 4 G 14 ± 4 M 15 ± 5			Diastolic P 72 ± 12 G 79 ± 12 V 80 ± 14	Controls: 2.4 ± 0.2 At 3 months: P decreased from 8.6 ± 1.3 to 3.4 ± 0.5, p < 0.01 G change NS V change NS --plasma ET-1 (pg/mL): At baseline: DM: 1.3 ± 0.4 Controls: 1.0 ± 0.6 At 3 months: No significant change in any of the groups (P, G, or V)		reduced to the same degree in all 3 groups. SCr, BP, BUN were not affected by treatment.
Pagano G, et al. Diabete Metab 1995;21:162-7. Funding: Study supported by Bayer S.p.A and a grant from Consiglio Nazionale delle Ricerche	Randomized, double-blind, multicenter study comparing: Miglitol (M) 50 mg tid x 6 wks, then 100 mg tid x 18 wks vs. Glibenclamide* (G) 2.5 mg bid x 6 weeks, then 5 mg bid x 18 weeks All patients were assigned a diet consisting of 30	France, Type 2 DM patients from four outpatient clinics aged 40-70 y with a BMI < 30 kg/m^2, HbA1c 7-11%, no prior antidiabetic drugs, and SCr > 176.8 mmol/L. Sample size: M 49, G 47 Mean age (y): Male: M 33/49, G 24/47 Duration of diabetes (months): M 60 ± 6.9, G 84 ± 9.4; (p =0.04) BMI (kg/m^2): M 26.4 ± 0.4, G 26.7 ± 0.4	7 week placebo run-in period, followed by 24 weeks of treatment	At baseline: M 8.2 ± 0.2 G 7.8 ± 0.1 Mean change at 24 weeks: M ↓ 0.78 ± 0.21 G ↓ 1.18 ± 0.20 (p < 0.05 for both vs. baseline) # patients reaching HbA1c < 7% at 24 weeks:	FPG (mmol/L): At baseline: M 9.6 ± 0.3 G 9.1 ± 0.3 At 24 weeks: M 8.7 ± 0.3 G 8.0 ± 0.3 (p < 0.001 for both vs. baseline). No changes in body weight were noted in either group. No significant	--Glucose incremental area following a standard meal (mmol/L 180 minutes): At baseline: M 537 ± 44 G 620 ± 46 At 24 weeks: M 406 ± 40; p < 0.01 G 596 ± 41; p = NS -- Insulin incremental area	No placebo group, so it is hard to tell how much of the change is secular trend vs. drug effect While authors concluded that miglitol is "appropriate for initial application in diet-	Authors concluded that miglitol is appropriate for initial treatment of type 2 DM. The lower stimulation of insulin release after a meal with miglitol may be of interest for long-term treatment since hyperinsulinem

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
	kcal/kg of ideal body weight per day (60% carbohydrates, 25% lipids, 15% proteins, 35% dietary fibers. All patients maintained a light physical activity (1 h walking per day) *same as glyburide in US			M 26/49 (53%) G 31/47 (66%) # patients with HbA1c dropping by > 1%: M 25/49 (51%) G 31/47 (66%) Subgroup Analysis: There was a subgroup of non-responders in whom HbA1c did not vary: M 10 G 9 A comparison of patient demographics and baseline HbA1c could not differentiate these nonresponders from the responders.	variations in TC, HDL, TG in either group. Side effects: M 10 patients (8 flatulence, 2 diarrhea) G 10 patients (asthenia and sensation of hunger)	following a standard meal (pmol/L 180 min): At baseline: M 35,832 ± 3180 G 39,420 ± 3276 At 24 weeks: M 30,755 ± 3795; p < 0.05 G 41,747 ± 3908	resistant Type 2 diabetes" it has less effect than sulfonylurea	ia could play an important role in insulin resistance and diabetes complications.
Pavo I, et al. J Clin Endocrinology and Metab 2003;88(4):1637-45. Funding: not specified, Eli Lilly & Co employs	Randomized, double-blind, multicenter, noninferiority trial comparing: Pioglitazone (P) 30-45 mg once daily vs.	Hungary and Russia, recently diagnosed (< 12 months) type 2 DM patients at least 40 y with HbA1c of 7.5-11.0% and not receiving an oral antihyperglycemic medication (OAM).	3-5 week placebo lead-in period, then randomization occurred; followed by an 8 week titration period and 24 week	At baseline: P: 8.6% M: 8.6% At week 32, Δ from baseline was: P: ↓ 1.3% M: ↓ 1.5%	FPG (mmol/L) At baseline: P 11.8, M 12.4 Δ at 32 weeks: P↓3, p<0.0001 M ↓ 2.8, p<0.0001 (p=NS for P	FSI (pmol/L) At baseline: P 101.2 M 118.3 Δ at 32 weeks: P ↓ 22.7, p<0.0001 M ↓ 1.3, p=NS		Both pioglitazone and metformin are effective and safe first-line treatment options in recently diagnosed,

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
primary author, Eli Lilly appears to have sponsored the study	Metformin (M) 850 – 2550 mg/day Mean dose at end of 8 week titration period: P: 41.5 mg/d M: 2292 mg/d % reaching max dose during titration: P: 77% G: 73%	Sample size: 205 (P 105, M 100) Mean age (y): P 54.2 ± 9.1, M 55.8 ± 8.4 Male: P 43.8, M 56 Mean BMI (kg/m^2): P 31.3 ± 4.2, M 31.1 ± 4.4 Duration of DM (months): P 5.6 ± 3.8, M 6.3 ± 3.9	treatment period (32 weeks of treatment altogether)	P < 0.001 vs baseline for each treatment; P vs M met criteria for noninferiority	vs. M) --TG: both ↓ (P vs. M, p not given) --TC (mmol/L): P no change, M ↓ 0.37, p=0.002 (p =0.02 in favor of M) --HDL (mmol/L): P ↑ 0.22, p<0.0001 M ↑ 0.13, p<0.0001 (p=0.02 in favor of P) LDL (mmol/L): P ↑ 0.16, p=0.055 M ↓ 0.18, p=0.004 (p=0.003 in favor of M) LDL/ApoB ratio: P ↑ 0.25, p<0.0001 M no change (p<0.0001 in favor of P) Body Weight (kg) At baseline:	(p=0.003 in favor of P) Lp(a) [g/L]: P ↑ 0.02, p=0.003 M no change --HOMA-S ∆ at 32 weeks: P ↑ 14.9%; p=0.002) M ↓ 0.9%; p=0.87 p<0.005 in favor of P (increased insulin sensitivity) The correlation between changes in HbA1c and HOMA-S was not significant (r=0.08) for P.		OAM-naïve patients with type 2 DM. Pioglitazone and metformin are equally effective in lowering HbA1c and FPG, but improvements in indicators of insulin sensitivity (↑HOMA-S, ↓FSI) were more pronounced in patients on pioglitazone.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					P 86.1 M 88.8 ∆ at 32 weeks: P ↑ 0.7 ± 0.4, p=0.041 M ↓2.4 ± 0.4, p<0.0001 (p<0.001 in favor of M) Systolic BP (mmHg): At baseline: P 140.1 ± 15.4 M 142.6 ± 14.2 At 32 weeks: P ↓ 6.2 ± 1.2, p<0.0001 M ↓ 6.7 ± 1.2, p<0.0001 (P vs. M, p=NS) Diastolic BP (mm Hg): At baseline: P 87 ± 8.5 M 88 ± 8.2 At 32 weeks: P ↓ 3.9 ± 0.6, p<0.0001 M ↓ 3.9 ± 0.6, p<0.0001 (p=NS for P vs. M) --Heart Rate: not influenced			

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					by either P or M Liver enzymes: ALT (U/L): At baseline P 30.3 M 29.0 ∆ at 32 weeks: P ↓ 6.8 ± 1.6, p<0.0001 M ↑ 1.2 ± 1.6, p=NS (p=0.0002 in favor of P) AST (U/L): At baseline: P 24.2 M 22.6 ∆ at 32 weeks: P ↓ 2.2 ± 0.9, p=0.011 M ↑ 0.7 ± 0.9, p=NS (p=0.016 in favor of P) % reporting side effects: P 51.4% M 47% (p=NS for P vs. M) Number of patients halting treatment due to side effects: P 5			

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					M 9 % experiencing side effects of: lower-limb edema: P 12.4% M 4% (p=0.001) diarrhea: P 4%, M 12.4% (p=0.041)			
Rajagopalan R, et al. Curr Med Res Opin 2005; 21(1):163-72. Funding: Takeda Europe Research and Development Centre, Ltd.	Post-hoc analysis of a subset of patients from the Schernthaner, et al and Charbonnel et al, studies. Schernthaner et al, examined Pioglitazone (P) vs Metformin (M) Charbonnel et al, examined Pioglitazone (P) vs. Gliclazide (G) See entries for these studies for more detail	Subset of patients meeting WHO definition as having metabolic syndrome from Schernthaner et al, and Charbonnel et al, studies. Sample size: P 1221, M 597, G 626 Mean age (y): P 57 ± 9.4, M 57 ± 9.3, G 57 ± 9.6 Male: P 57%, M 58%, G 62% BMI (kg/m2): P 31 ± 5.5, M 31 ± 5.2, G 31 ± 5.1 Duration of diabetes (y): P 3.1 ± 4.04, M 3.0 ± 3.71, G 3.1 ± 3.75	52 weeks of treatment consisting of 16-week forced titration period followed by a 36-week maintenance period at maximum tolerated dose	At baseline: P 8.7 ± 1.0 M 8.7 ± 1.0 G 8.7 ± 1.1	Primary outcome was the presence of metabolic syndrome at baseline and at week 52. Proportion of patients with metabolic syndrome: At baseline: 72.1% overall Change from baseline (95%CI) at week 52: P ↓ 9.2% (6.5-12%) M ↓ 7.7% (3.5-11.9%) G ↓ 4.3% (0.4-8.3%)		Post-hoc analysis of two studies (Schernthaner, et al and Charbonnel, et al) combined data. Two of the component studies published as abstract only, so cannot examine full source data. Only intermediate outcomes examined (Rajagopa-lan).	Pioglitazone, metformin, and gliclazide all reduced the proportion of patients meeting the definition for metabolic syndrome from baseline. Most often the improvement was due to increases in HDL and decreases in TG. Pioglitazone improved both HDL and TG parameters more than metformin or gliclazide.
Ramachandran A, et al. J Assoc	Open-label, randomized study	India, new type 2 DM subjects age 30-60 years	12-14 weeks	At baseline C 7.5 ± 1.0	FPG (mg/dL): At baseline:	FSI (mU/mL) At baseline:	The authors note that	Glimepiride improved

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
Physicians India (JAPI) 2006;52:458-463. Funding: not specified Additional detail published in: **Ramachandran A**, et al. JAPI 2004;52:459-63.	Control (C) consisting of diet and exercise vs. Glimepiride (G) 1-2 mg/day vs. Metformin (M) 250-850 mg/day vs. Pioglitazone (P) 15-30 mg/day All groups were advised an appropriate diets with restricted calories, high carbohydrate (60%), and low fat (< 20%).	who had not received any antidiabetic treatment in the past and a BMI < 30 kg/m2. Those with a HbA1C < 8.5% were advised diet and exercise and made up the control group. Subjects with HbA1C between 8.5 and 11 were randomized to one of the treatment groups. Patients with an HbA1C > 11% and/or fasting plasma glucose >/= 200 mg/dL were excluded. Sample size: 97 subjects (C 20, G 25, M 24, P 28) Mean age (y): C 43.5 ± 8.7, G 45.3 ± 10.3, M 44.4 ± 10.6, P 45.1 ± 8.5 --		G 10.2 ± 2.2 M 9.6 ± 2.4 P 9.3 ± 1.8 At end of study: C 7.2 ± 1.1 G 7.7 ± 1.7* M 8.2 ± 2.5** P 6.7 ± 1.3* * p < 0.01 vs. baseline ** p < 0.05 vs. baseline	C 7.4 ± 1.3 G 10.7 ± 2.7 M 10.2 ± 3.3 P 9.3 ± 2.0 At end of study: C 7.1 ± 1.6 G 7.9 ± 2.6* M 8.6 ± 3.7 P 6.8 ± 1.4* 2 hr plasma glucose (mg/dL): At baseline: C 13.4 ± 1.7 G 18.6 ± 4.2 M 17.4 ± 4.2 P 15.9 ± 4.7 At end of study: C 11.4 ± 2.7* G 11.5 ± 4.4* M 12.4 ± 4.8 P 9.9 ±3.4* Lipids (mg/dL) TC At baseline: C 5.4 ± 1.1 G 5.3 ± 1.4 M 5.1 ± 0.95 P 5.8 ± 1.4 At end of study: C 5.4 ± 0.9 G 5.5 ± 1.7 M 4.7 ± 0.9**	C 20.1 ± 7.9 G 17.1 ± 7.2 M 19.9 ± 7.2 P 19.2 ± 10.2 At end of study: C 23.6 ± 12.7 G 20.8 ± 8.1** M 19.1 ± 15.6 P 15.5 ± 7.3** 2 h plasma insulin: At baseline: C 82.3 ± 47.7 G 44.7 ± 32.5 M 58.1 ± 30.2 P 68.2 ± 38.0 At end of study: C 128.1 ± 64.4* G 88.9 ± 58.8 M 80.8 ± 73.8 P 77.7 ± 33.6 HOMA-IR At baseline: C 6.5 ± 2.7 G 7.8 ± 2.9 M 9.4 ± 5.3 P 7.9 ± 4.7 At end of study C 7.0 ± 3.4 G 7.3 ± 3.7 M 7.1 ± 6.1** P 4.9 ± 3.0* --HOMA-IR: % of patients with > 10%	Indian patients with type 2 DM show several characteris-tic features such as high insulin resistance with low BMI and young age at diagnosis. The metformin dose used was much lower (250-1000 mg/day) than that commonly used in other studies (2550 mg/day). This lower dose range is used in Indian patients at least in part due to the characteris-tics described above. Study	insulin secretion including the early phase secretion and reduce plasma triglycerides. Metformin and pioglitazone had beneficial effects on lipid levels, improved insulin sensitivity, and also improved insulin secretion. In an open label study of limited generalizability using non-comparable dosage regimens, pioglitazone had somewhat better outcomes, but conclusions should be viewed with caution given the multiple limitations.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					P 5.3 ± 1.2* HDL At baseline: C 1.0 ± 0.2 G 0.95 ± 0.3 M 1.0 ± 0.2 P 0.98 ± 0.15 At end of study: C 1.1 ± 0.3 G 1.1 ± 0.2 M 1.1 ± 0.3 P 1.1 ± 0.2* TG At baseline: C 2.0 ± 1.1 G 2.2 ± 1.4 M 2.8 ± 2.5 P 2.9 ± 2.4 At end of study: C 2.1 ± 1.2 G 1.7 ± 0.9** M 2.5 ± 1.8 P 2.2 ± 1.4** * p < 0.01 vs. baseline ** p < 0.05 vs. baseline	improvement over baseline by 12 weeks: C 0 G 5 M 38* P 26 HOMA-BF At baseline: C 122.4 ± 81.6 G 58.6 ± 43.6 M 70.9 ± 42.9 P 73.5 ± 42.9 At end of study C 165.2 ± 142.1 G 121.0 ± 79.2* M 99.7 ± 72.9 P 99.8 ± 47.0* --HOMA-BF: % of patients with > 10% improvement over baseline by 12 weeks: C 20 G 50* M 10 P 26 Δ I/G At baseline: C 24.5 ± 23.7 G 10.6 ± 10.4 M 9.5 ± 8.3 P 8.7 ± 7.7 At end of study C 31.1 ± 27.2	evaluated patients over short period (12-14 weeks). Open-label design Baseline weight and BMI were highest in the C group and lowest in the G group at the start of the study. Limited external validity- pts with BMI > 30 excluded	

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
						G 33.5 ± 21.7* M 21.6 ± 26.8 P 16.6 ± 11.4* --Insulinogenic Index (ΔI/G): % of patients with > 10% improvement over baseline by 12 weeks: C 13 G 35** M 15 P 16 Fasting C-peptide (pmol/min) At baseline: C 0.57 ± 0.28 G 0.64 ± 0.38 M 0.59 ± 0.31 P 0.56 ± 0.16 At end of study C 0.70 ± 0.18** G 0.88 ± 0.29 M 0.75 ± 0.39 P 0.64 ± 0.23 2 h C-peptide: At baseline: C 1.4 ± 0.48 G 1.0 ± 0.41 M 1.2 ± 0.52 P 1.1 ± 0.39 At end of study C 1.4 ± 0.44 G 1.6 ± 0.40*		

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
						M 1.7 ± 0.58* P 1.6 ± 0.37* * p < 0.01 vs. baseline ** p < 0.05 vs. baseline		
Schernthaner G, et al. J Clin Endocrin Metab 2004;89:6068-76. Also reported in: **Belcher G**, et al. Diabetes Medicine 2005;22:973-9. **Ceriello A**, et al. Diabetes Care 2005;28(2):266-72. **Erdmann E**, et al. Int J Cardiol 2006;107:147-153. Funding: not specified	Randomized, double-blind, non-inferiority study comparing: Pioglitazone (P) up to 45 mg/day vs. Metformin (M) up to 2550 mg/day Mean dose at end of 12 week titration period: P: 43 mg/day M: 2124 mg/day % reaching max dose during titration: P: 85.9% M: 61.6% Antihypertensives and lipid-lowering agents were permitted.	12 European countries, type 2 DM ages 35-75 y inadequately treated with diet alone with an HbA1c of 7.5-11.0% Sample size: 1199 Mean age (y): P 57 ± 9.4, M 56 ± 9.3 % male: 55 --% taking lipid lowering meds: P 11, M 10 --% taking an ACE inhibitor: P 31, M 29 BMI (kg/m^2): P 31.2 ± 4.9, M 31.4 ± 5.2 Duration of DM (y): P 3.4 ± 4.3, M 3.1 ± 3.8	12 week forced titration period, followed by a 40 week treatment period (52 weeks of treatment altogether)	At baseline: P 8.69% M 8.68% At 52 weeks: P 7.28% (Δ ↓ 1.41%) M 7.18% (Δ ↓1.50%); met criteria for noninferiority Maximum response for both was at 32 weeks and was maintained thereafter to week 52	FPG (mg/dL): At baseline: P 205.4 M 203.6 At 52 weeks: P 160.4 M 164; p = 0.016 in favor of P --TC (mg/dL) Mean change: P ↑ 9.65 M ↓ 4.25 --HDL (mg/dL): Mean change: P ↑ 6.18 M ↑ 3.09; p=0.001 in favor of P LDL (mg/dL): Mean change: P ↑ 10.4 M ↓ 4.63; p=0.001 in favor of M --TG (mg/dL): Mean change: P ↓ 54 M ↓ 26.6;	FSI (pmol/L): Mean change P ↓ 14 M ↓ 2; p<0.001 in favor of P --C-peptide (ng/mL): At baseline: P ↓ 0.2 (↓ 5.3%) M ↑ 0.1 (↑ 2.7%), p=NS --32,33 split Proinsulin levels (μIU/mL): P ↓ 1.2 (↓30.2%) M ↓ 1.1 (↓34.2%), p=NS		Pioglitazone and metformin provide clinically equivalent and statistically noninferior control of HbA1c over a 1 year period. Effects on FPG, post-prandial glucose levels, insulin sensitivity, and lipid profile differ between the two agents. Significantly greater reductions in FPG and glucose excursions were seen with P. P decreased TG and increased HDL.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					p=0.001 in favor of P FFA (mg/dL): Mean change: P ↓ 3.39 M ↑ 1.13; p< 0.001 in favor of P? --TC/HDL: ↓8% in both groups BP: ↓ trend in both, but p=NS compared to baseline Body weight: Mean change: P ↑ 1.9 kg, M ↓ 2.5 kg Urinary albumin/crea-tine ratio: P ↓ 19% M ↑ 1% (p=0.002 in favor of P) % with (micro)albu-minuria: At baseline: P 26.4% M 23.1% % resolved at week 52: P 11.2%			M decreased LDL. Both decreased TC/HDL ratio to similar degree.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					M 7.4% % new cases of (micro)albu-minuria at week 52: P 5.0% M 7.6% Liver Enzymes: --AST (U/L) At baseline: P 25 ± 12 M 25 ± 12 Mean change at week 52: P ↓ 1 ± 51 M ↑ 0 ± 38 --ALT (U/L) At baseline: P 32 ± 17 M 32 ± 18 Mean change at week 52: P ↓ 14 ± 91 M ↓ 2 ± 50 --GGT (U/L) At baseline: P 58 ± 5 M 51 ± 3 Mean change at week 52: P ↓ 12 ± 84 M ↓ 8 ± 74			

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					--Alk Phos (U/L) At baseline: P 70 ± 22 M 68 ± 25 Mean change at 52 weeks: P ↓ 9 ± 22 M ↓ 11 ± 19 Bilirubin (µmol/L) At baseline: P 9 ± 6 M 10 ± 5 Mean change at 52 weeks: P ↑ 6 ± 53 M ↑ 5 ± 37 % with LFT values above ULN: At baseline: AST: P 10.4, M 12.9 ALT: P 23.5, M 24 GGT: P 23.1, M 23.2 Bilirubin: P 3%, M 2.3 Alk phos: P 3.6, M 3.7 At last visit: AST: P 5, M 9.4 ALT: P 8.1, M			

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					20 GGT: P 12.8, M 17.6 Bilirubin: P 2.8, M 2.7 Alk Phos: P 2, M 1.9 LFTs Hepatotoxicity: P 2, M 1; all recovered within 2 weeks of diagnosis of jaundice after study drug discontinued -Tolerability: % of patients halting treatment due to side effects: GI: P 1.5%, M 2.5% Edema: P 1.5%, M 0.3% CNS: P 1.7%, M 0.3% Abnormal LFTs: P 0%, M 0.3%n % judged as severe: P 4.9%, M 7.4% # of deaths (all judged not treatment-			

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					related): P 3, M 2			
Simpson SH, et al. CMAJ 2006;174(2):169-74. Funding: Institute of Health Economics and Alberta Heritage Foundation for Medical Research	Retrospective, inception cohort study using administrative data from an electronic prescription database comparing: Sulfonylureas-first generation (S^1) vs. Glyburide (G) vs. Metformin (M)	Canada; 12,272 outpatients enrolled in health plan > 1 year, age > 29 y, new users of antidiabetic agents (first claim for oral antidiabetic agent with no claim for least 1 year prior), between 1/1/91 and 12/31/96. Patients had to receive S, M, or G for at least 6 months and had not received insulin; Saskatchewan Health Database Sample size: 5795 patients, including S^1 120 (607 person-years of follow-up), G 4,138 (19,298 person-years of follow-up), M 1,537 (6,995 person-years of follow-up) Mean age (y): 66.3 ± 13.4 Male: 56.6% Race: n/a Comorbidities: chronic disease score median = 8	Followed from index date to until death, termination of coverage, or 12/31/99. Mean duration of follow-up (y): 4.6 ± 2.1		There were 1503 deaths during the study period, of which 372 (24.8% were attributable to an acute ischemic event. Mortality rates were as follows: S1 67.6 deaths/1000 person-years G 61.4 deaths/1000 person-years M 39.6 deaths/1000 person-years _____ Primary outcome was time from Rx of S, M, or G to death from any cause: Subjects with the lowest level of exposure to the drugs (either lower daily dose or poor adherence) served as the reference group for each hazard ratio (HR) reported.. HR were displayed graphically so no confidence intervals are provided. S^1 2.07 (unadjusted) 2.12 (adjusted*) G 1.32 (unadjusted) 1.29 (adjusted*) M 0.92 (unadjusted) 0.84 (adjusted*) Re-analysis using poor adherence subgroup as comparator for good adherence subgroup S^1 2.20 (unadjusted) 2.45 (adjusted*) G 1.55 (unadjusted)		Patients not randomized. Administrative database, therefore could not control for level of glycemic control, BMI, or other cardiovascular risk factors.	A greater risk of death was associated with higher daily doses of first-generation sulfonylureas (adjusted HR 2.1, 95% CI 1.0-4.7) and glyburide (HR 1.3, 95% CI 1.2-1.4), but not metformin (HR 0.8, 95% CI 0.7-1.1). Similar associations were observed for death caused by an acute ischemic event.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					1.33 (adjusted*) M 1.10 (unadjusted) 0.98 (adjusted*) _____ Secondary outcome was time to death attributable to an acute ischemic event: S[1] 1.55 (unadjusted) 1.21 (adjusted*) G 1.46 (unadjusted) 1.37 (adjusted*) M 1.22 (unadjusted) 1.10 (adjusted*) _____ * adjusted for age, sex, nitrate use, chronic disease score, number of physician visits, and hospital admissions NOTE: 95% CIs were not provided as HRs given in graphical format.			
Tan M, et al. Clin Ther 2004; 26(5):680-93. Funding: Eli Lilly & Co employs primary author, Eli Lilly sponsored the study	Randomized, double-blind, parallel group, multicenter study comparing: Pioglitazone (P) 15-45 mg/day vs. Glimepiride (G) 2-8 mg/day Mean final dose: P 37 mg/day G 6 mg/day Dose distribution: P 15 mg (14.4%), 30 mg (23.4%),	19 centers in Mexico, type 2 DM either OAM-naïve or currently receiving monotherapy, HbA1c of 7.5-11.0% (naïve-patients) or 7.5-9.5% (monotherapy patients) Sample size: 244 Mean age (y): P 55.1 ± 8, G 55.7 ± 9.3 Male: P 45, G 53 --Race: Hispanic 99%, White 1% Duration of diabetes (months): P 77.8 ± 79.2, G 81.2 ± 82.8 --% taking OAMs at	1-3 week washout, 12 week titration period, and 40 week treatment period (52 weeks of treatment altogether)	At baseline: P 8.54 ± 0.9 G 8.45 ± 1 Mean Δ at 52 weeks (ITT, n=208): P ↓ 0.78, p< 0.001 G ↓ 0.68, p<0.001 (p=NS for P vs. G) Mean Δ at 52 weeks (completers, n=174):	FPG (mmol/L) At baseline: P 9.1 ± 2.5 G 9.1 ± 2.68 Mean Δ at 52 weeks: P ↓ 0.6, p=NS G ↑ 0.6, p=NS (p=0.012 in favor of P) Body weight (kg): At baseline: P 74.2 ± 10.5 G 74.5 ± 10.8	FSI (pmol/L): P 107.6 ± 69.31 G 102.3 ± 54.72 Mean Δ at 52 weeks: P ↓ 21.1 G ↑ 15.1 (p<0.001 in favor of P) --HOMA-S (%) At baseline: P 71.7 ± 47.6 G 68.5 ± 37.9 Mean Δ at 52 weeks:	Dropout/with drawal: P 28% G 28%	HbA1c were similarly reduced by pioglitazone and glimepiride at 52 weeks. Pioglitazone increased measures of insulin sensitivity, while glimepiride decreased them at 52 weeks.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
	45 mg (62.2%) G 2 mg (25.5%), 4 mg (11.3%), 6 mg (25.5%), 8 mg (37.7%)	screening: P 76%, G 77.2%		P ↓ 1.27, p< 0.001 G ↓ 0.78, p < 0.001 (p=0.027 in favor of P) Pattern of HbA1c changes was different: P: unchanged at week 12, then declined to week 52. G: rapid decline to week 12, continued decline to week 20, then stable to week 52.	Mean Δ at 52 weeks: P ↑ 1.49 G ↑ 0.79 (p=NS) BMI (kg/m2): At baseline: P 29.3 ± 3.3 G 28.8 ± 3.2 BP (mmHg): At baseline: Systolic: P 128.4 ± 14.6 G 127.8 ± 17.8 Diastolic P 81.6 ± 9.6 G 80.4 ± 10.2 Mean Δ at 52 weeks: Systolic: P ↓ 3.5 G ↓ 1.4 (p=NS) Diastolic: P ↓ 3.9 G ↓ 1.3 (p=0.028, in favor of P) Lipid Profile: P: HDL: ↑0.07 LDL: ↑0.42 TC: ↑0.48 TG: no Δ apo B: no Δ	P ↑ 17.96, p<0.001 G ↓ 7.88, p=0.099 (p<0.001 in favor of P) --QUICKI P 0.312 ± 0.0283 G 0.311 ± 0.0220 Mean Δ at 52 weeks: P ↑ 0.013 G ↓ 0.007 (p < 0.001 in favor of P)		Pioglitazone caused more edema, while glimepiride had more hypoglycemic episodes.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					TC/HDL ratio: no Δ LDL/apo B ratio: ↑0.30 G: No signif Δ's in any lipid parameter --AST, ALT: No signif. Δ's in either group. --Side effects: Overall: P 86.8% G 76.4% Serious ADRs: P 8 patients G 5 patients Number discontinuing due to ADR: P 5 patients G 3 patients Peripheral edema: P 28.9%, G 13.8% (p=0.0005, in favor of G) Hypoglycemia P 15.7%, G 30.9% (p=0.024 in favor of P)			

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
Tan MH, et al. Diabet Med 2004;21:859-66. Funding: Eli Lilly & Co employs primary author, Eli Lilly sponsored the study	Randomized, multicenter study comparing: Pioglitazone (P) 30-45 mg/day vs. micronized glibenclamide*(G) 1.75-10.5 mg/day Doses were titrated to achieve FBG ≤ 7 mmol/L (126 mg/dL) and 1 h post-prandial blood glucose (BG) of ≤ 10 mmol/L (180 mg/dL). % reaching maximal dose during titration: P 75%, G 62% * Same as micronized glyburide in US.	Denmark, Finland, Norway, and Sweden, type 2 DM either OAM-naïve or currently receiving monotherapy, HbA1c of 7.5-11.0% (naïve-patients) or 7.5-9.5% (monotherapy patients), fasting serum C-peptide of 0.333 pmol/L (1.0 ng/mL) Sample size: 200 subjects (P 91, G 109) Mean age (y): P 60 ± 8.5, G 57.9 ± 9.2 Male: P 62, M 73 Race: 99% Caucasian BMI (kg/m^2): P 30.2 ± 5.6, G 29.6 ± 4.8 Duration of DM (months): P 57.1 ± 56.9, G 62.6 ± 56.1 --% taking lipid lowering meds: P 32, G 26 --% taking antihypertensive meds: P 50, G 43 --% taking OAMs at screening: P 70, G 69 Of these, sulfonylureas P 55%, G 58% metformin P 44%, G 42%	1-3 week washout, 12 week titration period, and 40 week treatment period (52 weeks of treatment altogether)	At baseline: P 8.4 ± 0.7 G 8.5 ± 0.8 ∆ at 52 weeks (ITT): P ↓ 0.5% G ↓ 0.4%; p=NS for P vs. G ∆ at 52 weeks (completers only): P ↓ 1.2% G ↓ 0.6%; p=0.001 in favor of P Author's notes: Both P and G showed maximal response after 36 weeks. HbA1c stayed relatively constant for the P group during the remainder of the trial, but rose toward baseline for the G group at weeks 44 and 52, suggesting possible diminished	FPG (mmol/L) Baseline: P 10.7 ± 2.0 G 10.6 ± 2.4 ITT: At 52 weeks: P 9.9 ± 0.4 G 10.9 ± 0.3 (p=NS for P vs G) Completers: At 52 weeks: p<0.001 in favor of P --TG: P ↓ more than G (p<0.019) --TC (mmol/L): Mean change: P ↑ 0.20 G ↑ 0.05; p=NS --HDL (mmol/L): Mean change: P ↑ 0.21 G ↑ 0.03; p<0.001 in favor of P LDL: both ↔ --TC/HDL: P ↓ more than G	FSI (pmol/L): Mean change +/- SE P ↓ 1.3 ± 7.3; p = NS G ↑23.8 ± 6.4; p < 0.001 ; p< 0.007 for P vs. G --HOMA-S (%): At baseline: P 84.8 ± 51.1 G 99.3 ± 64.8 Mean change at 52 weeks: P ↑ 17 G ↓ 13; p <0.001 in favor of P --QUICKI At baseline: P 0.312 ± 0.027 G 0.317 ± 0.028 Mean change at 52 weeks: P ↑ 0.011 G ↓ 0.007; p =0.007 in favor of P --AIP: P ↓ more than G (p < 0.001)	Primary endpoint was insulin sensitivity measures (HOMA-S and QUICKI), not HbA1c. Non-US subjects. Dropout/withdrawal: P 38% G 40%	At 52 weeks, pioglitazone shows improvement in insulin sensitivity measured by HOMA-S, QUICKI, and FSI compared to micronized glyburide. The ITT analysis showed both agents improved HbA1c equally at 52 weeks. Pioglitazone improved some components of lipid profiles, even though patients gained more weight than glyburide patients.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
				effectiveness over time of G.	(p<0.004) Systolic BP (mmHg): Mean change: P ↓ 5.0 ± 1.52 G ↓ 5.3 ± 1.34; p<0.001 vs baseline for both; p=NS for P vs G Diastolic BP (mmHg): Mean change: P ↓ 4.3 ± 0.82 G ↓ 2.5 ± 0.72; p<0.001 vs. baseline for both; p=NS for P vs G --ALT (U/L) At baseline: P 34.4 ± 15.4 G 37.6 ± 19.4 Mean change at 52 weeks: P ↓ 10.7 G ↑ 0.18; p < 0.001 in favor of P --AST: At baseline: P 25.8 ± 10.1 G 27.5 ± 11.4 Mean change at 52 weeks:			

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					P ↓ 4.4 G ↑ 0.93; p < 0.001 in favor of P Side effects: Serious events: incidence did not differ (n=8 vs 7). Peripheral edema: P26%, G 8% (p=0.001) Pedal edema: scores did not differ Number of patients with 1 or more episodes of hypoglycemia: P 4 (4%) G 32 (29%); p< 0.001 in favor of P --Weight (kg): At baseline: P 88.7 ± 17.4 G 89.1 ± 16 Mean change at 52 weeks: P ↑ 3 G ↑ 1.1; p < 0.002 in favor of G			
Tan MH, et al.	Randomized,	Patients who completed	Continued	Primary	FPG (mmol/L):	FSI (pmol/L)	Dropout/with	At 2 years,

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
Diabet Care 2005;28:544-50. Funding: Study sponsored by Takeda Europe Research and Development Centre and Eli Lilly and Company.	double-blind, parallel group study comparing: Pioglitazone (P) up to 45 mg/day vs. Gliclazide* (G) up to 160 mg/day *not marketed in the US	the 1 year parent study (Charbonnel, et al outlined above) at 98 of the 206 original study centers were invited to participate in this extension of the parent study. Sample size: P: 270 G: 297 Mean age (y): P: 57 ± 9.8 G: 56 ± 9.9 Male: P: 63.3% G: 61.3% Race: White: P: 93.7% G: 92.6% BMI (kg/m^2): P: 32 ± 6.4 G: 31 ± 5.6 Duration of diabetes (y): P: 2.7 ± 3.5 G: 2.9 ± 3.8	subgroup of patients from above study for an additional year (2 years altogether)	endpoint: HbA1c < 8% at 2 years: P: 129/270 (47.8%) G: 110/297 (37%) P vs. G HbA1c: -0.45 ± 0.11 (95% CI; -0.66, -0.23) At 2 years, a greater proportion of patients treated with P maintained a HbA1c < 8% than with G. The agents began to diverge at week 32, became significantly different at week 52. HbA1c < 7%: At 2 years: P: 111/261 (42.5%) G: 81/289 (28%) (p< 0.001 in favor of P)	At 2 years: P vs G: -0.83 ± 0.22 (95% CI -1.26, -0.39) Body weight (kg): At baseline: P: 91.7 ± 19.9 G: 89.2 ± 18.2 At 2 years: P: 95.6 ± 0.42 G: 93.4 ± 0.42 (p <0.001 in favor of G)	At 2 years: P vs G: -52.9 ± 9.9 (95% CI -72.6, -33.3) HOMA-S (%): At 2 years: P vs G: 36.2 ± .4.4% (95% CI 27.5, 45) HOMA-B(%): At 2 years: P vs. G: -9.1 ± 3.7% (95% CI -16.3, -1.82)	drawal: P: 45.6% G: 57.2%	pioglitazone is superior to gliclazide in sustaining glycemic control as measured by % patients with HbA1c < 8%. FPG, and measures of insulin sensitivity were also better in the pioglitazone group at 2 years.
Watanabe I, et al. Diabetes Res Clin Prac 2005;	Randomized study, unclear if open-label, comparing:	Japan, untreated type 2 DM patients with HbA1c of 6.5-8.0%;	6 months	At baseline: P 6.9 ± 0.2 G 7.2 ± 0.5	All are % change from baseline P vs.	All are % change from baseline:	Small sample size Open label?	Of the clinical outcome variables

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
68:104-10. Funding: not specified	Pioglitazone (P) 15 mg/day vs. Glibenclamide* (G) 1.25-2.5 mg/day *same as glyburide in US	30 subjects		At 6 months: P 6.1 ± 0.33* G 6.3 ± 0.4* *p < 0.01 vs. baseline	G at 6 months: FPG P ↓ 11.7 G ↓ 9.8; p=NS --TC: P ↑ 1.9 G ↓ 1.1; p=NS LDL: P ↑ 9 G ↑ 4.5; p=NS --HDL: P ↑ 7.5 ± 8.6% G ↓ 4 ± 16.5%; p=0.009 --TG: P ↓20. ± 27.6% G ↑ 14.4 ± 44.6%; p=0.027 BP (mmHg): Systolic BP: P ↓ 3 G ↓ 7.4; p=NS Diastolic BP: P ↓ 11.6 G ↓ 0.9, p=NS BMI: P ↑ 0.3 G ↓ 3; p=NS Side Effects: P: 2 stopped treatment due to edema	--IRI: P ↓ 26.7 ± 33.4 G ↑ 13.3 ± 41.2; p =0.0052 --HOMA-IR: P ↓ 37.2 ± 33.8 G ↑ 2.1 ± 39.4; p =0.014 --h/CRP: P ↑ 8.6 G ↑ 2.3; p =NS --ba-PWV: P ↓ 6.3 ± 5.6 G ↑ 0.8 ± 5.7, p = 0.004	Primary outcomes are intermediate markers	measured, HbA1c and FPG improved equally in both groups at 6 months. Pioglitazone improved HDL and TG values, compared to no change in the glyburide group. IRI, HOMA-IR, and ba-PWV improved in the pioglitazone group

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					G: 1 stopped treatment due to hypoglycemia			
Yamanouchi T, et al. Diabetic Medicine 2005;22:980-5 Funding: not specified	Randomized study comparing: After 3 month run-in period of diet, exercise, patients were randomized to pioglitazone [P](30-45 mg/d) vs. metformin [M] (750 mg/day) vs. glimepiride [G] (1-2 mg/d)	Japan, compare changes in major metabolites at 12 months in outpatients w/ DM who had never used oral hypoglycemic or lipid agents. Sample size: 114 Mean age (y): P 55.2, M 54.7, G 55.6 Male: 50% Race/ethnicity: 100% Japanese Co-morbidities: not given Duration of diabetes: "short"	12 months	At baseline: P 10.2 ± 0.8 M 9.9 ± 0.7 G 9.8 ± 0.7 At 12 months: P 7.9 ± 1.0 M 7.8 ± 1.0 G 7.7 ± 0.9 (p=NS for all comparisons)	FPG (mmol/L) At baseline: P 11.97 ± 1.90 M 11.82 ± 1.69 G 12.05 ± 1.64 At 12 months: P: 7.93 ± 2.25* M: 9.03 ± 2.01* G: 8.79 ± 1.78* *p significant vs. baseline Lipid profile: At 12 months: TC, HDL, and TG; all p= NS vs. baseline FFA (mEq/L) At baseline: P 542.2 ± 226.5 M 523.7 ± 263 G 518.6 ± 243.1 At 12 months: P 237.3 ± 139.1* M 455.8 ± 205.3 G 475.6 ± 200.6	At 12 months: 1,5-anhydrogluctiol and IRI; all p=NS vs. baseline Fasting plasma insulin (mcU/mL) At baseline: P 10.2 ± 4.1 M 9.8 ± 4.0 G 9.6 ± 4.4 At 12 months: P 8.0 ± 5.0 M 7.7 ± 3.3 G 7.3 ± 3.9	Japanese patients Drug doses may not have been comparable, as well as lack of dose titration in the study Small sample size Large standard deviations on some outcome measures	Pioglitazone, metformin, and glimepiride were equally effective at 12 months in reducing HbA1c in drug-naïve Japanese patients with type 2 DM.

Evidence Table								
Citation	**Study Design & Diabetes Drugs Evaluated**	**Study Population**	**Length or Range of Tx Duration**	**HbA1c (%)**	**Other clinical outcomes**	**Other research outcomes+**	**Limitations**	**Conclusions**
					* p < 0.01 vs. baseline At 12 months: BMI and BP; all p=NS vs. baseline			

Appendix B. Data Files Contained in WEBCIS

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
CPK_MRNO	TALRGPAT	CHAR	11	N	Patient Medical Record #	Allergy Patient Data
CPK_ALRG_ID	TALRGPAT	INTEGER	4	N	Sequence #	Allergy Patient Data
CFK_SRC_SYS_ID	TALRGPAT	CHAR	1	N	Source System Id	Allergy Patient Data
C_DATE	TALRGPAT	DATE	4	Y	Allergy Date	Allergy Patient Data
CFK_ALRG_CODE	TALRGPAT	CHAR	2	Y	Allergy Category Code	Allergy Patient Data
CFK_ALRG_TYPE_CD	TALRGPAT	CHAR	1	Y	Allergy Type Code	Allergy Patient Data
C_ALRG_DESC	TALRGPAT	CHAR	40	N	Allergy Description	Allergy Patient Data
C_REACTION	TALRGPAT	CHAR	15	Y	Allergy Reaction	Allergy Patient Data
C_ACTIVE_STS	TALRGPAT	CHAR	1	N	Allergy Active Status	Allergy Patient Data
C_SEVERITY	TALRGPAT	CHAR	8	Y	Allergy Severity	Allergy Patient Data
CFK_ENTERED_USERID	TALRGPAT	CHAR	6	N	Entering Userid	Allergy Patient Data
C_ENTERED_TS	TALRGPAT	TIMESTMP	10	N	Entered Timestamp	Allergy Patient Data
CFK_INACT_USERID	TALRGPAT	CHAR	6	Y	Inactivating Userid	Allergy Patient Data
C_INACTIVATED_TS	TALRGPAT	TIMESTMP	10	Y	Inactivated Timestamp	Allergy Patient Data
C_NOTE	TALRGPAT	VARCHAR	202	Y	Allergy Note	Allergy Patient Data
CPK_MRNO	TCARDPHY	CHAR	11	N	Patient Medical Record #	Cardiology Patient Data (Physician)

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_OBSERV_TIMESTAMP	TCARDPHY	TIMESTMP	10	N	Observation Timestamp	Cardiology Patient Data (Physician)
C_OBSERV_ID	TCARDPHY	CHAR	20	N	Observation Id	Cardiology Patient Data (Physician)
C_DISP	TCARDPHY	CHAR	2	N	Observation Disposition	Cardiology Patient Data (Physician)
C_PHYS_CARE_ID	TCARDPHY	CHAR	3	N	Physician Care Id	Cardiology Patient Data (Physician)
C_PHYNO	TCARDPHY	CHAR	5	N	Physician #	Cardiology Patient Data (Physician)
C_FAX_IND	TCARDPHY	CHAR	1	N	Fax Indicator	Cardiology Patient Data (Physician)
CPK_MRNO	TCARDTXT	CHAR	11	N	Patient Medical Record #	Cardiology Patient Data
C_OBSERV_TIMESTAMP	TCARDTXT	TIMESTMP	10	N	Observation Timestamp	Cardiology Patient Data
C_OBSERV_ID	TCARDTXT	CHAR	20	N	Observation Id	Cardiology Patient Data
C_DISP	TCARDTXT	CHAR	2	N	Observation Disposition	Cardiology Patient Data
C_SEQNUM	TCARDTXT	SMALLINT	2	N	Text Sequence #	Cardiology Patient Data
C_CARD_TEXT	TCARDTXT	VARCHAR	2000	N	Text	Cardiology Patient Data
CPK_MRNO	TECGLOCD	CHAR	11	N	Patient Medical Record #	EKG Patient Location Table
C_ECG_DATE	TECGLOCD	DATE	4	N	EKG Date	EKG Patient Location Table

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_ECG_TIME	TECGLOCD	CHAR	5	N	EKG Time	EKG Patient Location Table
C_MUSE_TIMESTAMP	TECGLOCD	TIMESTMP	10	N	EKG System Timestamp	EKG Patient Location Table
C_TECHNICIAN	TECGLOCD	CHAR	3	N	EKG Technician	EKG Patient Location Table
C_ECG_LOCATION	TECGLOCD	CHAR	10	N	EKG Location	EKG Patient Location Table
C_ECG_PERFORMED	TECGLOCD	CHAR	10	N	Type of EKG Performed	EKG Patient Location Table
C_INSERT_DATE	TECGLOCD	TIMESTMP	10	N	Record Insert Timestamp	EKG Patient Location Table
C_MEDICATION	TECGLOCD	CHAR	27	N	EKG Medication	EKG Patient Location Table
CFK_MRNO	TECGTEXT	CHAR	11	N	Patient Medical Record #	EKG Patient Text Table
C_ECG_DATE	TECGTEXT	DATE	4	N	EKG Date	EKG Patient Text Table
C_ECG_TIME	TECGTEXT	CHAR	5	N	EKG Time	EKG Patient Text Table
C_TEXT_SEQ	TECGTEXT	SMALLINT	2	N	Text Sequence #	EKG Patient Text Table
C_TEXT_REVISE_DATE	TECGTEXT	DATE	4	N	Text Revision Date	EKG Patient Text Table
C_INSERT_DATE	TECGTEXT	TIMESTMP	10	N	Record Insert Timestamp	EKG Patient Text Table
C_TEXT	TECGTEXT	VARCHAR	200	N	Text	EKG Patient Text Table
CFK_MRNO	TECGUITL	CHAR	11	N	Patient Medical Record #	EKG Patient Utility Table
C_ECG_DATE	TECGUITL	DATE	4	N	EKG Date	EKG Patient Utility Table

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_ECG_TIME	TECGUITL	CHAR	5	N	EKG Time	EKG Patient Utility Table
C_HEIGHT	TECGUITL	DECIMAL	2	N	Patient Height	EKG Patient Utility Table
C_WEIGHT	TECGUITL	DECIMAL	3	N	Patient Weight	EKG Patient Utility Table
C_VENT_RATE	TECGUITL	DECIMAL	3	N	Patient Ventracal Rate	EKG Patient Utility Table
C_PR_INTERVAL	TECGUITL	CHAR	3	N	Patient PR Interval	EKG Patient Utility Table
C_QRS_DURATION	TECGUITL	DECIMAL	3	N	Patient QRS Duration	EKG Patient Utility Table
C_QT_QTC	TECGUITL	CHAR	7	N	Patient QT_QTC Measurement	EKG Patient Utility Table
C_PRT_AXES	TECGUITL	CHAR	12	N	Patient PRT_AXES Measurement	EKG Patient Utility Table
C_INSERT_DATE	TECGUITL	TIMESTMP	10	N	Record Insert Timestamp	EKG Patient Utility Table
CPK_MRNO	TGIPRCDR	CHAR	11	N	Patient Medical Record #	GI Procedures Patient Data
C_PROCEDURE	TGIPRCDR	CHAR	21	N	GI Procedure Name	GI Procedures Patient Data
C_PROCEDURE_DATE	TGIPRCDR	DATE	4	N	GI Procedure Date	GI Procedures Patient Data
C_DOCUMENT_NUMBER	TGIPRCDR	CHAR	40	N	GI Procedure Document #	GI Procedures Patient Data
C_SEQ_NUM	TGIPRCDR	SMALLINT	2	N	Text Sequence #	GI Procedures Patient Data
C_UPDATE_TIMESTAMP	TGIPRCDR	TIMESTMP	10	N	Text Revision Date	GI Procedures Patient Data
C_INSERT_TIMESTAMP	TGIPRCDR	TIMESTMP	10	N	Record Insert Timestamp	GI Procedures Patient Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_TEXT	TGIPRCDR	VARCHAR	1896	N	Text	GI Procedures Patient Data
CPK_MRNO	TGIPRPHY	CHAR	11	N	Patient Medical Record #	GI Procedures Physician Data
C_PROCEDURE	TGIPRPHY	CHAR	21	N	GI Procedure Name	GI Procedures Physician Data
C_PROCEDURE_DATE	TGIPRPHY	DATE	4	N	GI Procedure Date	GI Procedures Physician Data
C_DOCUMENT_NUMBER	TGIPRPHY	CHAR	40	N	GI Procedure Document #	GI Procedures Physician Data
C_PHYS_CARE_ID	TGIPRPHY	CHAR	3	N	Physician Care Id	GI Procedures Physician Data
C_PHYNO	TGIPRPHY	CHAR	5	N	Physician #	GI Procedures Physician Data
C_FAX_ID	TGIPRPHY	CHAR	1	N	Fax Indicator	GI Procedures Physician Data
CPK_MRNO	TLABRQSC	CHAR	11	N	Patient Medical Record #	Lab Patient Result Request
C_TESTDATE	TLABRQSC	DATE	4	N	Lab Test Date	Lab Patient Result Request
C_SAMPLENO	TLABRQSC	CHAR	8	N	Lab Sample #	Lab Patient Result Request
C_REQ_TEST	TLABRQSC	CHAR	4	N	Lab Request Test #	Lab Patient Result Request
C_ORGANISM_ID	TLABRQSC	CHAR	2	N	Lab Organism ID	Lab Patient Result Request
C_REQ_STATUS	TLABRQSC	CHAR	1	N	Lab Request Status	Lab Patient Result Request
C_LOCATION_ID	TLABRQSC	CHAR	4	N	Lab Location ID	Lab Patient Result Request
C_SPECIMEN_SRC	TLABRQSC	CHAR	5	N	Lab Specimen Source	Lab Patient Result Request

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_SPECIMEN_SITE	TLABRQSC	CHAR	32	N	Lab Specimen Site	Lab Patient Result Request
C_TEST_TIME	TLABRQSC	CHAR	4	N	Lab Test Time	Lab Patient Result Request
C_ORDERNO	TLABRQSC	CHAR	11	N	Lab Order #	Lab Patient Result Request
C_ORDER_PHYS	TLABRQSC	CHAR	5	N	Lab Ordering Physician #	Lab Patient Result Request
C_ORDER_LOCATION	TLABRQSC	CHAR	5	N	Lab Ordering Location	Lab Patient Result Request
CFK_MRNO	TLABRSSC	CHAR	11	N	Patient Medical Record #	Lab Patient Result
C_TESTDATE	TLABRSSC	DATE	4	N	Lab Test Date	Lab Patient Result
C_SAMPLENO	TLABRSSC	CHAR	8	N	Lab Sample #	Lab Patient Result
C_REQ_TEST	TLABRSSC	CHAR	4	N	Lab Request #	Lab Patient Result
C_ORGANISM_ID	TLABRSSC	CHAR	2	N	Lab Organism ID	Lab Patient Result
C_LOCATION_ID	TLABRSSC	CHAR	4	N	Lab Location ID	Lab Patient Result
C_RES_TEST	TLABRSSC	CHAR	4	N	Lab Result Test #	Lab Patient Result
C_RESULT_SEQ	TLABRSSC	SMALLINT	2	N	Lab Result Seq #	Lab Patient Result
C_COMPLETE_DATE	TLABRSSC	TIMESTMP	10	N	Lab Completion Date	Lab Patient Result
C_TEST_RANGE	TLABRSSC	CHAR	20	N	Lab Test Range	Lab Patient Result
C_TEST_UNITS	TLABRSSC	CHAR	15	N	Lab Test Units	Lab Patient Result
C_ABNORMAL_IND	TLABRSSC	CHAR	2	N	Lab Abnormal Indicator	Lab Patient Result

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_MORETEXT	TLABRSSC	CHAR	1	N	Lab More Text Indicator	Lab Patient Result
C_RES_STATUS	TLABRSSC	CHAR	1	N	Lab Result Status	Lab Patient Result
C_RESULTS	TLABRSSC	VARCHAR	75	N	Lab Results	Lab Patient Result
CFK_MRNO	TLABTXSC	CHAR	11	N	Patient Medical Record #	Lab Patient Text
C_TESTDATE	TLABTXSC	DATE	4	N	Lab Test Date	Lab Patient Text
C_SAMPLENO	TLABTXSC	CHAR	8	N	Lab Sample #	Lab Patient Text
C_REQ_TEST	TLABTXSC	CHAR	4	N	Lab Request #	Lab Patient Text
C_ORGANISM_ID	TLABTXSC	CHAR	2	N	Lab Organism ID	Lab Patient Text
C_LOCATION_ID	TLABTXSC	CHAR	4	N	Lab Location ID	Lab Patient Text
C_RES_TEST	TLABTXSC	CHAR	4	N	Lab Result Test #	Lab Patient Text
C_RESULT_SEQ	TLABTXSC	SMALLINT	2	N	Lab Result Seq #	Lab Patient Text
C_TEXT_SEQ	TLABTXSC	SMALLINT	2	N	Lab Text Seq #	Lab Patient Text
C_TEXT	TLABTXSC	VARCHAR	78	N	Lab Text	Lab Patient Text
CFK_ACCOUNT_NUMBER	TPATINS	CHAR	12	N	Patient Visit Account #	Patient Visit Insurance Data
CFK_MRNO	TPATINS	CHAR	11	N	Patient Medical Record #	Patient Visit Insurance Data
CFK_CLINIC_CODE	TPATINS	CHAR	3	N	Clinic Code	Patient Visit Insurance Data
CFK_VISIT_DATE	TPATINS	TIMESTMP	10	N	Patient Visit Date	Patient Visit Insurance Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_SEQ_NO	TPATINS	SMALLINT	2	N	Sequence #	Patient Visit Insurance Data
C_INSERTION_DATE	TPATINS	TIMESTMP	10	N	Record insert timestamp	Patient Visit Insurance Data
CFK_INSUR_CODE	TPATINS	CHAR	5	N	Insurance Code	Patient Visit Insurance Data
CFK_ACCOUNT_NUMBER	TPATPHYS	CHAR	12	N	Patient Visit Account #	Patient Visit Physician Data
CFK_MRNO	TPATPHYS	CHAR	11	N	Patient Medical Record #	Patient Visit Physician Data
CFK_CLINIC_CODE	TPATPHYS	CHAR	3	N	Clinic Code	Patient Visit Physician Data
CFK_VISIT_DATE	TPATPHYS	TIMESTMP	10	N	Patient Visit Date	Patient Visit Physician Data
CFK_PHYS_CARE_ID	TPATPHYS	CHAR	3	N	Physician Care Id	Patient Visit Physician Data
CFK_PHYNO	TPATPHYS	CHAR	5	N	Physician #	Patient Visit Physician Data
C_INSERTION_DATE	TPATPHYS	TIMESTMP	10	N	Record Insert Timestamp	Patient Visit Physician Data
CPK_ACCOUNT_NUMBER	TPATVIST	CHAR	12	N	Patient Visit Account #	Patient Visit Data
C_MRNO	TPATVIST	CHAR	11	N	Patient Medical Record #	Patient Visit Data
CFK_CLINIC_CODE	TPATVIST	CHAR	3	N	Clinic Code	Patient Visit Data
C_VISIT_DATE	TPATVIST	TIMESTMP	10	N	Patient Visit Date	Patient Visit Data
CFK_PATIENT_STATUS	TPATVIST	CHAR	2	N	Patient Status	Patient Visit Data
CFK_HOSP_SVC	TPATVIST	CHAR	3	N	Hospital Service Code	Patient Visit Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
CFK_FINAN_CLASS	TPATVIST	CHAR	1	N	Financial Class	Patient Visit Data
C_ADMIT_DATE	TPATVIST	TIMESTMP	10	N	Admit Timestamp	Patient Visit Data
C_INSERTION_DATE	TPATVIST	TIMESTMP	10	N	Record Insert Timestamp	Patient Visit Data
CFK_PATIENT_TYPE	TPATVIST	CHAR	1	N	Patient Type	Patient Visit Data
C_CLERK_ID	TPATVIST	CHAR	11	N	Registration Clerk Id	Patient Visit Data
C_DEPT	TPATVIST	CHAR	3	N	Department	Patient Visit Data
C_DIVISION	TPATVIST	CHAR	4	N	Division	Patient Visit Data
C_END_TS	TPATVIST	TIMESTMP	10	Y	Visit End Timestamp	Patient Visit Data
C_VITAL_STATS_NOTE	TPATVIST	VARCHAR	1870	Y	Vital Stats Note	Patient Visit Data
CPK_ACCOUNT_NUMBER	TPATVISTEXT	CHAR	12	N	Patient Visit Account #	Patient Visit Extention
C_MRNO	TPATVISTEXT	CHAR	11	N	Patient Medical Record #	Patient Visit Extention
CFK_CLINIC_CODE	TPATVISTEXT	CHAR	3	N	Clinic Code	Patient Visit Extention
C_VISIT_DATE	TPATVISTEXT	TIMESTMP	10	N	Patient Visit Date	Patient Visit Extention
CFK_LAST_NSTATION	TPATVISTEXT	CHAR	4	Y	Patient Last Nurse Station	Patient Visit Extention
C_LAST_ROOM_BED	TPATVISTEXT	CHAR	6	Y	Patient Last Room/Bed	Patient Visit Extention
C_ADMIT_DIAGNOSIS	TPATVISTEXT	CHAR	30	Y	Patient Admit Diagnosis	Patient Visit Extention

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
CFK_DISPOSITION_CD	TPATVISTEXT	CHAR	3	Y	Patient Disposition	Patient Visit Extention
C_DISCHARGE_TS	TPATVISTEXT	TIMESTMP	10	Y	Patient Discharge Timestamp	Patient Visit Extention
CFK_VISIT_STS_CD	TPATVISTEXT	CHAR	1	Y	Patient Visit Status Code	Patient Visit Extention
CFK_MRNO	TPIDXADR	CHAR	11	N	Patient Medical Record #	Patient Demographic Address
C_PT_ADDRESS	TPIDXADR	CHAR	30	N	Patient Address	Patient Demographic Address
C_PT_CITY	TPIDXADR	CHAR	18	N	Patient City	Patient Demographic Address
C_PT_STATE	TPIDXADR	CHAR	2	N	Patient State	Patient Demographic Address
C_PT_ZIP	TPIDXADR	CHAR	9	N	Patient Zip	Patient Demographic Address
C_PT_PHONE	TPIDXADR	CHAR	10	N	Patient Home Phone	Patient Demographic Address
C_PT_PHONE_WK	TPIDXADR	CHAR	10	N	Patient Work Phone	Patient Demographic Address
C_DISTRICT_CODE	TPIDXADR	CHAR	3	N	Patient District Code	Patient Demographic Address
C_ORIGINAL_DATE	TPIDXADR	TIMESTMP	10	N	Original Record Timestamp	Patient Demographic Address
C_UPDATE_DATE	TPIDXADR	TIMESTMPH	10	N	Update record Timestamp	Patient Demographic Address
C_PT_ADDRESS2	TPIDXADR	CHAR	30	N	Patient Address Line 2	Patient Demographic Address
CPK_MRNO	TPIDXHIPAA	CHAR	11	N	Patient Medical Record #	Patient HIPAA
CFK_SOURCE	TPIDXHIPAA	CHAR	15	N	HIPAA Source System	Patient HIPAA

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
CFK_HIPAA_SIGN_CD	TPIDXHIPAA	INTEGER	4	N	HIPAA Sign Code	Patient HIPAA
C_HIPAA_SIGN_CD_DT	TPIDXHIPAA	DATE	4	N	HIPAA Sign Code Date	Patient HIPAA
C_NOTICE_VERSION	TPIDXHIPAA	CHAR	2	Y	HIPAA Version	Patient HIPAA
C_COMMENT	TPIDXHIPAA	CHAR	30	Y	HIPAA Comment	Patient HIPAA
C_TS	TPIDXHIPAA	TIMESTMP	10	N	Record Insert Timestamp	Patient HIPAA
CPK_MRNO	TPIDXHIPAACONF	CHAR	11	N	Patient Medical Record #	Patient HIPAA Confidentiality
CPK_HIPAA_CONF_CD	TPIDXHIPAACONF	INTEGER	4	N	HIPAA Conf Code	Patient HIPAA Confidentiality
C_OPT_OUT_DATE	TPIDXHIPAACONF	DATE	4	Y	HIPAA Op/Out Date	Patient HIPAA Confidentiality
CFK_SOURCE	TPIDXHIPAACONF	CHAR	15	N	HIPAA Source System	Patient HIPAA Confidentiality
C_TS	TPIDXHIPAACONF	TIMESTMP	10	N	Record Insert Timestamp	Patient HIPAA Confidentiality
CPK_MRNO	TPIDXHIPAAPERSON	CHAR	11	N	Patient Medical Record #	Patient HIPAA Person
CPK_PERSON	TPIDXHIPAAPERSON	CHAR	30	N	HIPAA Person Name	Patient HIPAA Person
C_OPT_OUT_DATE	TPIDXHIPAAPERSON	DATE	4	Y	HIPAA Op/Out Date	Patient HIPAA Person
CPK_SOURCE	TPIDXHIPAAPERSON	CHAR	15	N	HIPAA Source System	Patient HIPAA Person
C_TS	TPIDXHIPAAPERSON	TIMESTMP	10	N	Record Insert Timestamp	Patient HIPAA Person
CPK_MRNO	TPIDXINS	CHAR	11	N	Patient Medical Record #	Patient Demographic Insurance

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_PATIENT_NO	TPIDXINS	CHAR	12	N	Patient #	Patient Demographic Insurance
C_SEQ_NO	TPIDXINS	SMALLINT	2	N	Sequence #	Patient Demographic Insurance
C_INSERT_DATE	TPIDXINS	TIMESTMP	10	N	Record Insert Timestamp	Patient Demographic Insurance
CFK_INSUR_CODE	TPIDXINS	CHAR	5	N	Patient Insurance Code	Patient Demographic Insurance
CPK_MRNO	TPIDXMST	CHAR	11	N	Patient Medical Record #	Patient Demographic Master
C_PT_SS_NUM	TPIDXMST	CHAR	9	N	Patient Social Security #	Patient Demographic Master
C_MARITAL_STAT	TPIDXMST	CHAR	1	N	Patient Marital Status	Patient Demographic Master
C_RELIGION_CD	TPIDXMST	CHAR	3	N	Patient Religion Code	Patient Demographic Master
C_BIRTH_DATE	TPIDXMST	CHAR	10	N	Patient Birth Date	Patient Demographic Master
C_RACE	TPIDXMST	CHAR	5	N	Patient Race	Patient Demographic Master
C_SEX	TPIDXMST	CHAR	1	N	Patient Sex	Patient Demographic Master
C_PT_RECORD_KEY	TPIDXMST	CHAR	12	N	Patient Record Key	Patient Demographic Master
C_ORIGINAL_DATE	TPIDXMST	TIMESTMP	10	N	Original Record Timestamp	Patient Demographic Master
C_UPDATE_DATE	TPIDXMST	TIMESTMP	10	N	Update Record Timestamp	Patient Demographic Master
C_ADV_IND	TPIDXMST	CHAR	2	N	Patient Advance Directive Indicator	Patient Demographic Master

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_ADV_IND_DATE	TPIDXMST	DATE	4	N	Patient Advance Directive Date	Patient Demographic Master
C_DEATH_IND	TPIDXMST	CHAR	1	N	Patient Death Indicator	Patient Demographic Master
C_DIS_DEATH_DATE	TPIDXMST	DATE	4	N	Patient Discharge Death Date	Patient Demographic Master
C_CONTACT_CAUTION	TPIDXMST	CHAR	1	Y	Patient Contact Caution Indicator	Patient Demographic Master
CPK_LAST_NAME	TPIDXNME	CHAR	22	N	Patient Last Name	Patient Demographic Name
C_FIRST_NAME	TPIDXNME	CHAR	16	N	Patient First Name	Patient Demographic Name
C_MIDDLE_INIT	TPIDXNME	CHAR	16	N	Patient Middle Name	Patient Demographic Name
C_MRNO	TPIDXNME	CHAR	11	N	Patient Medical Record #	Patient Demographic Name
C_ORIGINAL_DATE	TPIDXNME	TIMESTMP	10	N	Original Record Timestamp	Patient Demographic Name
C_UPDATE_DATE	TPIDXNME	TIMESTMP	10	N	Update Record Timestamp	Patient Demographic Name
CPK_MRNO	TPRBPAT	CHAR	11	N	Patient Medical Record #	Patient Problem
CPK_PAT_PROBLEM_ID	TPRBPAT	INTEGER	4	N	Patient Problem Id	Patient Problem
CFK_DEPT	TPRBPAT	CHAR	3	Y	Department	Patient Problem
CFK_DIV	TPRBPAT	CHAR	4	Y	Division	Patient Problem
CPK_ICD9_CODE	TPRBPAT	CHAR	5	Y	ICD9 Code	Patient Problem
C_ONSET_DATE	TPRBPAT	DATE	4	Y	Onset Date	Patient Problem

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_ACTIVE_STS	TPRBPAT	CHAR	1	N	Active Status Flag	Patient Problem
C_DESC	TPRBPAT	CHAR	255	N	Problem Description	Patient Problem
C_ENTERED_TS	TPRBPAT	TIMESTMP	10	N	Entered Timestamp	Patient Problem
CFK_ENTERED_USERID	TPRBPAT	CHAR	6	Y	Entered WebCIS User Id	Patient Problem
C_UPDATED_TS	TPRBPAT	TIMESTMP	10	N	Updated Timestamp	Patient Problem
CFK_UPDATED_USERID	TPRBPAT	CHAR	6	Y	Update WebCIS User Id	Patient Problem
C_INACTIVE_DATE	TPRBPAT	DATE	4	Y	Inactivation Date	Patient Problem
C_NOTE	TPRBPAT	VARCHAR	800	Y	Notes	Patient Problem
CPK_MRNO	TPTHCASC	CHAR	11	N	Patient Medical Record #	Pathology Case Data
C_SPECDATE	TPTHCASC	DATE	4	N	Pathology Specimen Date	Pathology Case Data
C_CASENUM	TPTHCASC	CHAR	12	N	Pathology Case Number	Pathology Case Data
CFK_CASETYPE	TPTHCASC	CHAR	2	N	Pathology Case Type	Pathology Case Data
CFK_PHYNO	TPTHCASC	CHAR	5	N	Physician #	Pathology Case Data
C_LOCATION	TPTHCASC	CHAR	4	N	Pathology Location	Pathology Case Data
C_STATUS	TPTHCASC	CHAR	1	N	Pathology Status	Pathology Case Data
CFK_MRNO	TPTHTXSC	CHAR	11	N	Patient Medical Record #	Pathology Text Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_SPECDATE	TPTHTXSC	DATE	4	N	Pathology Specimen Date	Pathology Text Data
C_CASENUM	TPTHTXSC	CHAR	12	N	Pathology Case Number	Pathology Text Data
C_CASETYPE	TPTHTXSC	CHAR	2	N	Pathology Case Type	Pathology Text Data
C_TEXT_SEQ	TPTHTXSC	SMALLINT	2	N	Text Sequence #	Pathology Text Data
C_TEXT	TPTHTXSC	VARCHAR	162	N	Text	Pathology Text Data
CPK_MRNO	TPULDIAG	CHAR	11	N	Patient Medical Record #	Pulmomary Diagnosis Data
C_ORDER_TIMESTAMP	TPULDIAG	TIMESTMP	10	N	Pulmonary Order Timestamp	Pulmomary Diagnosis Data
C_ORDERNO	TPULDIAG	CHAR	22	N	Pulmonary Order #	Pulmomary Diagnosis Data
C_REPORT_TYPE	TPULDIAG	CHAR	3	N	Pulmonary Report Type	Pulmomary Diagnosis Data
C_RESULT_STATUS	TPULDIAG	CHAR	1	N	Pulmonary Result Status	Pulmomary Diagnosis Data
C_STATUS_SEQ	TPULDIAG	SMALLINT	2	N	Pulmonary Status Sequence	Pulmomary Diagnosis Data
C_AGE	TPULDIAG	SMALLINT	2	N	Patient Age	Pulmomary Diagnosis Data
C_AGE_KEYED	TPULDIAG	CHAR	8	N	Patient Age	Pulmomary Diagnosis Data
C_HEIGHT_IN	TPULDIAG	DECIMAL	2	N	Patient Height/in	Pulmomary Diagnosis Data
C_HEIGHT_CM	TPULDIAG	DECIMAL	3	N	Patient Height/cm	Pulmomary Diagnosis Data
C_ARMSPAN_IN	TPULDIAG	DECIMAL	4	N	Patient Armspan/in	Pulmomary Diagnosis Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_ARMSPAN_CM	TPULDIAG	DECIMAL	5	N	Patient Armspan/cm	Pulmomary Diagnosis Data
C_NORM_LB	TPULDIAG	DECIMAL	5	N	Normal pounds	Pulmomary Diagnosis Data
C_NORM_KG	TPULDIAG	DECIMAL	5	N	Normal KG	Pulmomary Diagnosis Data
C_WEIGHT_LB	TPULDIAG	DECIMAL	3	N	Patient Weight/lb	Pulmomary Diagnosis Data
C_WEIGHT_KG	TPULDIAG	DECIMAL	5	N	Patient Weight/kg	Pulmomary Diagnosis Data
C_BSA	TPULDIAG	DECIMAL	5	N	Patient BSA	Pulmomary Diagnosis Data
C_ROOM	TPULDIAG	CHAR	20	N	Patient Room	Pulmomary Diagnosis Data
C_TEMP	TPULDIAG	SMALLINT	2	N	Patient Temperature	Pulmomary Diagnosis Data
C_PBAR	TPULDIAG	SMALLINT	2	N	Patient PBAR	Pulmomary Diagnosis Data
C_PHYSICIAN	TPULDIAG	CHAR	20	N	Physician Name	Pulmomary Diagnosis Data
C_TECHNICIAN	TPULDIAG	CHAR	20	N	Technician Name	Pulmomary Diagnosis Data
C_DIAGNOSIS	TPULDIAG	CHAR	24	N	Diagnosis	Pulmomary Diagnosis Data
C_PF_REFERENCE	TPULDIAG	CHAR	20	N	PF Reference	Pulmomary Diagnosis Data
C_ARMSPAN_USED	TPULDIAG	CHAR	1	N	Patient Armspan Used	Pulmomary Diagnosis Data
C_PDFEV1	TPULDIAG	CHAR	6	N	PDFEV1	Pulmomary Diagnosis Data
C_PROTOCOL	TPULDIAG	CHAR	20	N	PROTOCOL	Pulmomary Diagnosis Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_DOSE	TPULDIAG	CHAR	11	N	Dose	Pulmomary Diagnosis Data
C_PDRAW	TPULDIAG	CHAR	6	N	PDRAW	Pulmomary Diagnosis Data
C_PCFEV1	TPULDIAG	CHAR	6	N	PCFEV1	Pulmomary Diagnosis Data
C_MET_REFERENCE	TPULDIAG	CHAR	20	N	MET REFERENCE	Pulmomary Diagnosis Data
C_EXERCISE_TIME	TPULDIAG	CHAR	4	N	EXERCISE TIME	Pulmomary Diagnosis Data
CFK_MRNO	TPULORDR	CHAR	11	N	Patient Medical Record #	Pulmonary Order Data
C_ORDER_TIMESTAMP	TPULORDR	TIMESTMP	10	N	Pulmonary Order Timestamp	Pulmonary Order Data
C_ORDERNO	TPULORDR	CHAR	22	N	Pulmonary Order #	Pulmonary Order Data
C_REPORT_TYPE	TPULORDR	CHAR	3	N	Pulmonary Report Type	Pulmonary Order Data
C_RESULT_STATUS	TPULORDR	CHAR	1	N	Pumomary Result Status	Pulmonary Order Data
C_STATUS_SEQ	TPULORDR	SMALLINT	2	N	Pulmonary Status Sequence #	Pulmonary Order Data
C_SEQ_NUM	TPULORDR	SMALLINT	2	N	Pulmonary Sequence #	Pulmonary Order Data
C_OBSERV_CODE	TPULORDR	CHAR	7	N	Pulmonary Observation Code	Pulmonary Order Data
C_OBSERV_PARM2	TPULORDR	CHAR	20	N	Pulmonary Observation Parm	Pulmonary Order Data
C_OBSERV_AMT_PRD	TPULORDR	CHAR	8	N	Pulmonary Observation PRD	Pulmonary Order Data
C_OBSERV_AMT_PRE	TPULORDR	CHAR	8	N	Pulmonary Observation PRE	Pulmonary Order Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_OBSERV_AMT_PPR	TPULORDR	CHAR	8	N	Pulmonary Observation PPR	Pulmonary Order Data
C_OBSERV_AMT_AFD	TPULORDR	CHAR	8	N	Pulmonary Observation AFD	Pulmonary Order Data
C_OBSERV_AMT_APP	TPULORDR	CHAR	8	N	Pulmonary Observation APP	Pulmonary Order Data
C_OBSERV_AMT_CMC	TPULORDR	CHAR	8	N	Pulmonary Observation CMC	Pulmonary Order Data
C_OBSERV_UNIT	TPULORDR	CHAR	10	N	Pulmonary Observation Unit	Pulmonary Order Data
CFK_MRNO	TPULTEXT	CHAR	11	N	Patient Medical Record #	Pulmonary Text Data
C_ORDER_TIMESTAMP	TPULTEXT	TIMESTMP	10	N	Pulmonary Order Timestamp	Pulmonary Text Data
C_ORDERNO	TPULTEXT	CHAR	22	N	Pulmonary Order #	Pulmonary Text Data
C_REPORT_TYPE	TPULTEXT	CHAR	3	N	Pulmonary Report Type	Pulmonary Text Data
C_RESULT_STATUS	TPULTEXT	CHAR	1	N	Pumomary Result Status	Pulmonary Text Data
C_STATUS_SEQ	TPULTEXT	SMALLINT	2	N	Pumomary Status Sequence #	Pulmonary Text Data
C_RESULT_SEQ	TPULTEXT	SMALLINT	2	N	Pumomary Result Sequence #	Pulmonary Text Data
C_RESULT_TEXT	TPULTEXT	VARCHAR	1970	N	Pumomary Result Text	Pulmonary Text Data
CPK_MRNO	TRDOORDR	CHAR	11	N	Patient Medical Record #	Radiology Order Data
C_ORDERNO	TRDOORDR	CHAR	8	N	Radiology Order #	Radiology Order Data
C_TESTDATE	TRDOORDR	DATE	4	N	Radiology Test Date	Radiology Order Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_DISPOSITION	TRDOORDR	CHAR	2	N	Radiology Test Disposition	Radiology Order Data
C_REASON	TRDOORDR	CHAR	76	N	Radiology Test Reason	Radiology Order Data
CFK_RAD_REFCODE	TRDOORDR	CHAR	5	N	Radiology Attending Physician #	Radiology Order Data
C_RES_REFCODE	TRDOORDR	CHAR	5	N	Radiology Referring Physician #	Radiology Order Data
C_ORDERING_PHYNO	TRDOORDR	CHAR	5	N	Radiology Ordering Physician #	Radiology Order Data
C_PHY_LNAME	TRDOORDR	CHAR	15	N	Radiologist Last Name	Radiology Order Data
C_PHY_FNAME	TRDOORDR	CHAR	15	N	Radiologist First Name	Radiology Order Data
C_TIMESTAMP	TRDOORDR	TIMESTMP	10	N	Record Insert Timestamp	Radiology Order Data
CFK_MRNO	TRDOTEST	CHAR	11	N	Patient Medical Record #	Radiology Test Data
C_ORDERNO	TRDOTEST	CHAR	8	N	Radiology Order #	Radiology Test Data
C_TESTDATE	TRDOTEST	DATE	4	N	Radiology Test Date	Radiology Test Data
C_DISPOSITION	TRDOTEST	CHAR	2	N	Radiology Test Disposition	Radiology Test Data
C_SEQNO	TRDOTEST	CHAR	3	N	Radiology Sequence #	Radiology Test Data
C_RAD_GROUPID	TRDOTEST	CHAR	16	N	Radiology Group Id	Radiology Test Data
CFK_PROC_NUM	TRDOTEST	CHAR	7	N	Radiology Procedure Id	Radiology Test Data
CFK_MRNO	TRDOTEXT	CHAR	11	N	Patient Medical Record #	Radiology Text Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_ORDERNO	TRDOTEXT	CHAR	8	N	Radiology Order #	Radiology Text Data
C_TESTDATE	TRDOTEXT	DATE	4	N	Radiology Test Date	Radiology Text Data
C_DISPOSITION	TRDOTEXT	CHAR	2	N	Radiology Test Disposition	Radiology Text Data
C_TEXT_DISPOSITION	TRDOTEXT	CHAR	3	N	Radiology Text Disposition	Radiology Text Data
C_RAD_GROUPID	TRDOTEXT	CHAR	16	N	Radiology Group Id	Radiology Text Data
C_TEXT_SEQ	TRDOTEXT	SMALLINT	2	N	Radiology Text Sequence #	Radiology Text Data
C_TEXT	TRDOTEXT	VARCHAR	1986	N	Text	Radiology Text Data
CPK_MRNO	TRESLDAT	CHAR	11	N	Patient Medical Record #	Epidemiology Result Data
C_ACCESS_NUM	TRESLDAT	CHAR	7	N	Epidemiology Accession #	Epidemiology Result Data
C_SEQ_NUM	TRESLDAT	SMALLINT	2	N	Epidemiology Sequence #	Epidemiology Result Data
CFK_RESULT_CD	TRESLDAT	CHAR	5	N	Epidemiology Result Code	Epidemiology Result Data
C_INSERT_DATE	TRESLDAT	TIMESTMP	10	N	Record Insert Timestamp	Epidemiology Result Data
CPK_MRNO	TRESORDR	CHAR	11	N	Patient Medical Record #	Ancillary Result Order Data
C_DATE_OF_SERVICE	TRESORDR	TIMESTMP	10	N	Result Date of Service	Ancillary Result Order Data
C_ORDER_NUMBER	TRESORDR	CHAR	13	N	Result Order #	Ancillary Result Order Data
CFK_SOURCE_ID	TRESORDR	CHAR	3	N	Result Source Id	Ancillary Result Order Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_RESULT_STATUS	TRESORDR	CHAR	3	N	Result Status	Ancillary Result Order Data
C_ACCESSION_NUMBER	TRESORDR	CHAR	16	Y	Result Accession #	Ancillary Result Order Data
C_ACCOUNT_NUMBER	TRESORDR	CHAR	12	Y	Result Account #	Ancillary Result Order Data
C_CLINIC_CODE	TRESORDR	CHAR	3	Y	Result Clinic Code	Ancillary Result Order Data
C_ORDER_PHYSICIAN	TRESORDR	CHAR	5	Y	Result Ordering Physician	Ancillary Result Order Data
C_INTERP_PHYSICIAN	TRESORDR	CHAR	5	Y	Result Interpreting Physician	Ancillary Result Order Data
C_PROCEDURE_ID	TRESORDR	CHAR	13	Y	Result Procedure Id	Ancillary Result Order Data
C_INSERT_TS	TRESORDR	TIMESTMP	10	Y	Record Insert Timestamp	Ancillary Result Order Data
C_ENTERING_USERID	TRESORDR	CHAR	8	N	Result Entering Userid	Ancillary Result Order Data
C_ENTERING_FNAME	TRESORDR	CHAR	30	N	Result Entering First Name	Ancillary Result Order Data
C_ENTERING_LNAME	TRESORDR	CHAR	30	N	Result Entering Last Name	Ancillary Result Order Data
C_REPORT_TYPE	TRESORDR	CHAR	1	Y	Result Report Type	Ancillary Result Order Data
CPK_MRNO	TRESPBAS	CHAR	11	N	Patient Medical Record #	Respiratory Base Data
C_BEGIN_DATE	TRESPBAS	DATE	4	N	Respiratory Order Begin Date	Respiratory Base Data
C_BEGIN_TIME	TRESPBAS	CHAR	6	N	Respiratory Order Begin Time	Respiratory Base Data
C_PLACER_ORDER	TRESPBAS	CHAR	30	N	Respiratory Placer Order Name	Respiratory Base Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_DICTATING_PERSON	TRESPBAS	CHAR	30	N	Respiratory Dictating Person	Respiratory Base Data
C_ACTIVITY	TRESPBAS	CHAR	30	N	Respiratory Activity	Respiratory Base Data
C_ORDER_PROVIDER	TRESPBAS	CHAR	30	N	Respiratory Ordering Provider	Respiratory Base Data
C_END_DATE	TRESPBAS	DATE	4	N	Respiratory End Date	Respiratory Base Data
C_END_TIME	TRESPBAS	CHAR	6	N	Respiratory End Time	Respiratory Base Data
CPK_MRNO	TRESPTXT	CHAR	11	N	Patient Medical Record #	Respiratory Text Data
C_BEGIN_DATE	TRESPTXT	DATE	4	N	Respiratory Order Begin Date	Respiratory Text Data
C_BEGIN_TIME	TRESPTXT	CHAR	6	N	Respiratory Order Begin Time	Respiratory Text Data
C_PLACER_ORDER	TRESPTXT	CHAR	30	N	Respiratory Placer Order Name	Respiratory Text Data
C_SEQNO	TRESPTXT	SMALLINT	2	N	Respiratory Order Sequence #	Respiratory Text Data
C_RESULT_ID	TRESPTXT	CHAR	40	N	Respiratory Order Result Id	Respiratory Text Data
C_UNITS	TRESPTXT	CHAR	20	N	Respiratory Units	Respiratory Text Data
C_RESULT	TRESPTXT	VARCHAR	250	N	Respiratory Results	Respiratory Text Data
CFK_MRNO	TRESTEXT	CHAR	11	N	Patient Medical Record #	Ancillary Result Text Data
C_DATE_OF_SERVICE	TRESTEXT	TIMESTMP	10	N	Result Date of Service	Ancillary Result Text Data
C_ORDER_NUMBER	TRESTEXT	CHAR	13	N	Result Order #	Ancillary Result Text Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
CFK_SOURCE_ID	TRESTEXT	CHAR	3	N	Result Source Id	Ancillary Result Text Data
C_RESULT_STATUS	TRESTEXT	CHAR	3	N	Result Status	Ancillary Result Text Data
C_SEQ_NUM	TRESTEXT	SMALLINT	2	N	Result Sequence #	Ancillary Result Text Data
C_TEXT	TRESTEXT	VARCHAR	1896	N	Result Text	Ancillary Result Text Data
CPK_MRNO	TSITEDAT	CHAR	11	N	Patient Medical Record #	Epidemiology Result Data
C_ACCESS_NUM	TSITEDAT	CHAR	7	N	Epidemiology Accession #	Epidemiology Result Data
CFK_SITE	TSITEDAT	CHAR	6	N	Epidemiology Site	Epidemiology Result Data
C_OLD_SITE_1	TSITEDAT	CHAR	3	N	Epidemiology Old Site 1	Epidemiology Result Data
C_OLD_SITE_2	TSITEDAT	CHAR	3	N	Epidemiology Old Site 2	Epidemiology Result Data
C_COMMENTS	TSITEDAT	CHAR	70	N	Epidemiology Comments	Epidemiology Result Data
C_INSERT_DATE	TSITEDAT	TIMESTMP	10	N	Record Insert Timestamp	Epidemiology Result Data
CFK_MRNO	TTRNDIAG	CHAR	11	N	Patient Medical Record #	Transcription Diagnosis Data
C_DISCHARG_DATE	TTRNDIAG	DATE	4	N	Patient Date of Service	Transcription Diagnosis Data
C_UDOCNUM	TTRNDIAG	CHAR	8	N	Transcription Document #	Transcription Diagnosis Data
C_DOC_TYPE	TTRNDIAG	CHAR	2	N	Transcription Doc Type	Transcription Diagnosis Data
C_DIAGTEXT	TTRNDIAG	VARCHAR	802	N	Transcription Diagnosis Text	Transcription Diagnosis Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
CPK_MRNO	TTRNSCRP	CHAR	11	N	Patient Medical Record #	Transcription Master Data
C_DISCHARG_DATE	TTRNSCRP	DATE	4	N	Patient Date of Service	Transcription Master Data
C_UDOCNUM	TTRNSCRP	CHAR	8	N	Transcription Document #	Transcription Master Data
C_DOC_TYPE	TTRNSCRP	CHAR	2	N	Transcription Doc Type	Transcription Master Data
C_ADMIT_DATE	TTRNSCRP	CHAR	10	N	Admit Date	Transcription Master Data
CFK_PHYNO	TTRNSCRP	CHAR	5	N	Attending Physician #	Transcription Master Data
C_NUMBER_SEGMENTS	TTRNSCRP	SMALLINT	2	N	Transcription Number of Segments	Transcription Master Data
CFK_MRNO	TTRNTEXT	CHAR	11	N	Patient Medical Record #	Transcription Text Data
C_DISCHARG_DATE	TTRNTEXT	DATE	4	N	Patient Date of Service	Transcription Text Data
C_UDOCNUM	TTRNTEXT	CHAR	8	N	Transcription Document #	Transcription Text Data
C_DOC_TYPE	TTRNTEXT	CHAR	2	N	Transcription Doc Type	Transcription Text Data
C_TEXT_SEQ	TTRNTEXT	SMALLINT	2	N	Transcription Sequence #	Transcription Text Data
C_TEXT	TTRNTEXT	VARCHAR	1896	N	Transcription Text	Transcription Text Data
CPK_MRNO	TTRPRDAT	CHAR	11	N	Patient Medical Record #	Epidemiology Treatment Data
C_ACCESS_NUM	TTRPRDAT	CHAR	7	N	Epidemiology Accession #	Epidemiology Treatment Data
C_SEQ_NUM	TTRPRDAT	SMALLINT	2	N	Sequence #	Epidemiology Treatment Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
CFK_TRPR_CD	TTRPRDAT	CHAR	3	N	Epidemiology Treatment Code	Epidemiology Treatment Data
C_INSERT_DATE	TTRPRDAT	TIMESTMP	10	N	Record Insert Timestamp	Epidemiology Treatment Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
CPK_ACCOUNT_NUMBER	TVISTNOTE	CHAR	12	N	Patient Account #	Patient Vital Stats Note Data
C_MRNO	TVISTNOTE	CHAR	11	N	Patient Medical Record #	Patient Vital Stats Note Data
CFK_CLINIC_CODE	TVISTNOTE	CHAR	3	N	Clinic Code	Patient Vital Stats Note Data
C_VISIT_DATE	TVISTNOTE	TIMESTMP	10	N	Patient Visit Timestamp	Patient Vital Stats Note Data
C_VISIT_NOTE_ID	TVISTNOTE	INTEGER	4	N	Patient Visit Note Id	Patient Vital Stats Note Data
C_ENTERED_TS	TVISTNOTE	TIMESTMP	10	N	Entered Timestamp	Patient Vital Stats Note Data
C_ENTERED_USERID	TVISTNOTE	CHAR	6	N	Entered Userid	Patient Vital Stats Note Data
C_VITAL_STATS_NOTE	TVISTNOTE	VARCHAR	1870	Y	Patient Vital Stats Note	Patient Vital Stats Note Data
CPK_ACCOUNT_NUMBER	TVITPAT	CHAR	12	N	Patient Account #	Patient Vital Stats Data
C_MRNO	TVITPAT	CHAR	11	N	Patient Medical Record #	Patient Vital Stats Data
CFK_CLINIC_CODE	TVITPAT	CHAR	3	N	Clinic Code	Patient Vital Stats Data
C_VISIT_DATE	TVITPAT	TIMESTMP	10	N	Patient Visit Timestamp	Patient Vital Stats Data
CPK_VITAL_STAT_ID	TVITPAT	INTEGER	4	N	Patient Vital Stats Id	Patient Vital Stats Data
CPK_SEQ_NBR	TVITPAT	INTEGER	4	N	Patient Vital Stats Sequence #	Patient Vital Stats Data
C_VALUE_DATE	TVITPAT	DATE	4	Y	Patient Vital Value Date	Patient Vital Stats Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_VALUE_NUMERIC	TVITPAT	DECIMAL	5	Y	Patient Vitals	Patient Vital Stats Data
CFK_UOM_ID	TVITPAT	SMALLINT	2	Y	Unit of measure id	Patient Vital Stats Data
C_VALUE_STRING	TVITPAT	CHAR	20	Y	Patient Vital Value	Patient Vital Stats Data
CFK_VITAL_SIGN_STS	TVITPAT	CHAR	1	N	Patient Vital Sign Status	Patient Vital Stats Data
CFK_ENTERED_USERID	TVITPAT	CHAR	6	N	Entered Timestamp	Patient Vital Stats Data
C_ENTERED_TS	TVITPAT	TIMESTMP	10	N	Entered Userid	Patient Vital Stats Data
C_NOTE	TVITPAT	VARCHAR	160	Y	Patient Vitals Note	Patient Vital Stats Data
CPK_MRNO	TFDBCLASSPAT	CHAR	11	N	Patient Medical Record #	Patient Drug Class Data
CPK_AHFS_CLASS	TFDBCLASSPAT	CHAR	6	N	Drug Class Code	Patient Drug Class Data
C_ACTIVE_STS	TFDBCLASSPAT	CHAR	1	N	Patient Active Status	Patient Drug Class Data
C_ENTERED_USERID	TFDBCLASSPAT	CHAR	6	N	Entered Userid	Patient Drug Class Data
C_ENTERED_TS	TFDBCLASSPAT	TIMESTMP	10	N	Entered Timestamp	Patient Drug Class Data
C_UPDATED_USERID	TFDBCLASSPAT	CHAR	6	Y	Updated Userid	Patient Drug Class Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_UPDATED_TS	TFDBCLASSPAT	TIMESTMP	10	Y	Updated Timestamp	Patient Drug Class Data
C_NOTE	TFDBCLASSPAT	VARCHAR	95	Y	Patient Drug Notes	Patient Drug Class Data
CPK_MRNO	TFDBPAT	CHAR	11	N	Patient Medical Record #	Patient Drug Data
CPK_PAT_DRUG_ID	TFDBPAT	INTEGER	4	N	Drug Id	Patient Drug Data
CFK_NDC	TFDBPAT	CHAR	11	Y	Drug NDC Code	Patient Drug Data
C_DRUG_NAME	TFDBPAT	CHAR	30	N	Drug Name	Patient Drug Data
C_DEA_IND	TFDBPAT	CHAR	1	Y	Dea Indicator	Patient Drug Data
C_STR	TFDBPAT	DECIMAL	11	Y	Drug Strength	Patient Drug Data
C_STR_UNITS	TFDBPAT	CHAR	10	Y	Drug Strength Units	Patient Drug Data
C_STR_VOL	TFDBPAT	DECIMAL	7	Y	Drug Strength Volumne	Patient Drug Data
C_STR_VOL_UNIT	TFDBPAT	CHAR	5	Y	Drug Strength Volumne Units	Patient Drug Data
C_GENERIC_CODE_NBR	TFDBPAT	CHAR	5	Y	Drug Generic Code	Patient Drug Data
C_DOSAGE	TFDBPAT	CHAR	20	Y	Drug Dosage	Patient Drug Data
C_DOSAGE_FORM	TFDBPAT	CHAR	10	Y	Drug Dosage Form	Patient Drug Data
C_DOSAGE_FORM_CD	TFDBPAT	CHAR	2	Y	Drug Dosage Form Code	Patient Drug Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_QUANTITY	TFDBPAT	CHAR	20	Y	Drug Quantity	Patient Drug Data
C_DAYS_SUPPLY	TFDBPAT	SMALLINT	2	Y	Drug Days	Patient Drug Data
C_AHFS_CLASS	TFDBPAT	CHAR	6	Y	Drug Classification	Patient Drug Data
C_SPEC_THER_CLASS	TFDBPAT	CHAR	6	Y	Drug Sepcialty Classification	Patient Drug Data
CFK_FREQ_CD	TFDBPAT	CHAR	10	Y	Drug Frequency Code	Patient Drug Data
C_ROUTE	TFDBPAT	CHAR	10	Y	Drug Route	Patient Drug Data
CFK_REFILL_CD	TFDBPAT	CHAR	3	Y	Drug Refill Code	Patient Drug Data
C_ENTERED_USERID	TFDBPAT	CHAR	6	N	Entered Userid	Patient Drug Data
C_ENTERED_TS	TFDBPAT	TIMESTMP	10	N	Entered Timestamp	Patient Drug Data
C_ACTIVE_STS	TFDBPAT	CHAR	1	Y	Active Status Id	Patient Drug Data
C_INACT_USERID	TFDBPAT	CHAR	6	Y	Inactive Userid	Patient Drug Data
C_INACT_TS	TFDBPAT	TIMESTMP	10	Y	Inactive Timestamp	Patient Drug Data
C_NOT_TOL_NOTE	TFDBPAT	CHAR	50	Y	Total Notes	Patient Drug Data
CFK_SUBSTITION_CD	TFDBPAT	CHAR	1	Y	Substitution Code	Patient Drug Data
C_NOTE	TFDBPAT	VARCHAR	220	Y	Drug Notes	Patient Drug Data
C_PHARMACY_NOTE	TFDBPAT	VARCHAR	155	Y	Pharmacy Notes	Patient Drug Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
CPK_MRNO	TIMMUPAT	CHAR	11	N	Patient Medical Record #	Patient Immunization Data
CPK_IMMU_ID	TIMMUPAT	INTEGER	4	N	Patient Immunization Id	Patient Immunization Data
CPK_SEQ_NO	TIMMUPAT	INTEGER	4	N	Sequence #	Patient Immunization Data
CFK_LOCATION_ID	TIMMUPAT	INTEGER	4	Y	Location	Patient Immunization Data
C_ORDERED_DATE	TIMMUPAT	DATE	4	Y	Order Date	Patient Immunization Data
C_LAST_DONE_DATE	TIMMUPAT	DATE	4	Y	Last Done Date	Patient Immunization Data
C_NOT_DONE_DATE	TIMMUPAT	DATE	4	Y	Not Done Date	Patient Immunization Data
CFK_NOT_DONE_ID	TIMMUPAT	SMALLINT	2	Y	Not Done Id	Patient Immunization Data
CFK_DISPLAY_UOM_ID	TIMMUPAT	SMALLINT	2	Y	UOM Id	Patient Immunization Data
C_CUSTOM_DUE_AGE	TIMMUPAT	INTEGER	4	Y	Custom Due Age	Patient Immunization Data
C_ENTERED_TS	TIMMUPAT	TIMESTMP	10	N	Entered Timestamp	Patient Immunization Data
C_UPDATED_TS	TIMMUPAT	TIMESTMP	10	N	Updated Timestamp	Patient Immunization Data
C_COMMENT	TIMMUPAT	VARCHAR	255	Y	Comments	Patient Immunization Data
CPK_MRNO	TINFECTC	CHAR	11	N	Patient Medical Record #	Epidemiology Infection Data
C_ACCESS_NUM	TINFECTC	CHAR	7	N	Epidemiology Accession #	Epidemiology Infection Data
C_INFECT_DATE	TINFECTC	DATE	4	N	Epidemiology Infection Date	Epidemiology Infection Data

Column Name	Table Name	Column Type	Column Length	Null Indicator	Column Description	Table Description
C_ADMIT_DATE	TINFECTC	DATE	4	N	Patient Admit Date	Epidemiology Infection Data
C_PREV_ADMIT_DATE	TINFECTC	DATE	4	N	Patient Previous Admit Date	Epidemiology Infection Data
C_NURSE_STATION	TINFECTC	CHAR	4	N	Nurse Station	Epidemiology Infection Data
C_ROOM_NUMBER	TINFECTC	CHAR	4	N	Room #	Epidemiology Infection Data
C_INFECT_SERVICE	TINFECTC	CHAR	4	N	Infected Service Code	Epidemiology Infection Data
C_INFECT_LOCATION	TINFECTC	CHAR	3	N	Infected Location	Epidemiology Infection Data
C_PHYNO	TINFECTC	CHAR	5	N	Physician #	Epidemiology Infection Data
C_OUTSIDE_XFER	TINFECTC	CHAR	1	N	Outside Transfer	Epidemiology Infection Data
C_INSERT_DATE	TINFECTC	TIMESTMP	10	N	Record Insert Timestamp	Epidemiology Infection Data

Appendix C. Program To Delete Extraneous Characters From Text Data Files

```
*****************************************************************
* BADCHAR.SAS
* HSRProj: Test for bad characters
*
* 4/12/06
*****************************************************************;
options pageno=1 ps=60;

*** read acceptable character set ***;
data charset;
  infile 'c:\nlm\charset.txt' lrecl=300;
  length goodchar1-goodchar91 $2;
  array gc $ goodchar1-goodchar91;
  input goodchar1-goodchar91;
run;

data hexfile;
infile 's:\sheps shared\s\schwartz_bob\diab_ttrndiag.txt' lrecl=1000
truncover
   delimiter='0D'x ignoredoseof;
length line $835;
input  line &;
len=length(line);
array c {835} $ c1-c835;
array h {835} $ h1-h835;
do i=1 to len;
  c{i}=substr(line,i,1);
  h{i}=put(c{i},$hex2.);
end;
run;
data newfile;
set hexfile;
if _n_=1 then set charset;
```

```
array h {835} $ h1-h835;
array c {835} $ c1-c835;
array goodchar {91} $ goodchar1-goodchar91;
do i=1 to len;
  badchar=1;
  do j=1 to 91;
    if h{i} = goodchar{j} then badchar=0;
  end;
  if badchar=1 then c{i}=' ';
end;
file 's:\sheps shared\s\schwartz_bob\diab_ttrndiag2.txt' lrecl=835;
put (c1-c835) ($char1.);
run;
```

Appendix D. Code for Deidentification Process for the TTRNTEXT.txt Data

```
#code to read in PHI from provided file then replace all of them in another file

$regex_PHI = '(^[a-zA-Z-]+)\s+([a-zA-Z-]+)\s([a-zA-Z])?\s+([a-zA-Z]+)\s+(\d{11})\s(\d{11})';

# define length of string at beginning to omit from replacing
#  - for ttrntext file
#$omitLength = 32;

#  - for ttrndiag file
$omitLength = 28;

# prompt user for files
print "\nPlease enter filename with PHI (i.e. tpidxnme.txt): ";
$phiFile = <STDIN>;

print "\nPlease enter file where PHI needs to be removed: ";
$inFile = <STDIN>;

print "\nPlease enter filename for the OUTPUT file: ";
$outFile = <STDIN>;

open (PHI_FileHdl,"$phiFile") or die "can not open PHI file\n";

$i=0;
while ($line=(<PHI_FileHdl>)) {

# for debugging
#  $line=(<PHI_FileHdl>);
#  print $line;

  if ($line =~/$regex_PHI/) {
     $aPHI{$5}{'lname'} = $1;
     $aPHI{$5}{'fname'} = $2;
     $aPHI{$5}{'realMRNO'} = $6;

 #   $aLastName[$i] = $1;
 #   $aFirstName[$i] = $2;
 #   $aMidIni[$i] = $3;
 #   $aMidName[$i] = $4;
 #   $aCodedMRNo[$i] = $5;
 #   $aRealMRNo[$i] = $6;

  }  # eo if match regex

  $i++;
}
```

```
$numEntries = $i;
print "Number of PHI Entries: $numEntries\n\n";

# #for debugging
#print "lname: " . $aPHI{$5}{'lname'} . "\n";
#print "fname: " . $aPHI{$5}{'fname'} . "\n";
#print "realMRNO: " . $aPHI{$5}{'realMRNO'} . "\n";

#use the following for debugging
#print "with coded lname: " . $aPHI{'00011739270'}{'lname'} . "\n";
#print "with coded fname: " . $aPHI{'00011739270'}{'fname'} . "\n";
#print "with coded mrno: " . $aPHI{'00011739270'}{'realMRNO'} . "\n";

#$k = <STDIN>;

close PHI_FileHdl;

# values are now in arrays - open target file and do substitutions
$strSub = "*REMOVED*";

open (IN_FileHdl,"$inFile") or die "Can not open input file\n";
binmode(IN_FileHdl);

open (OUT_FileHdl,">$outFile") or die "Can not Create the log text\n";
$|++;   # set buffer control to flush on each write

#open (LOG,">$logfile") or die "Can not Create the log text\n";

# preset coded - just for first pass
  $coded = "00012455856";

 $j=0;
while ($fullLine=(<IN_FileHdl>)) {

# for debugging/testing
# while ($j<100) {
#          $fullLine=(<IN_FileHdl>);
#   print "PRE-line\n";

#   for ($i=0;$i<$numEntries;$i++) {
#   for ($i=0;$i<20;$i++) {

    $headOfLine = "";
    if ($fullLine =~ /(^\d{11})/) {
            $coded = $1;
 #          print "coded: " . $coded . "\n";

     # account for omitting first omitLength of string
      $headOfLine = substr($fullLine,0,$omitLength);   #set aside for later
```

```
      $inLine = substr($fullLine,$omitLength);
          } else {
                    $inLine = $fullLine;    # lines that do not start with coded MRNOs
          }

# for debugging
# print "headOfLine: $headOfLine\n";
# print "inLine: $inLine\n";
#$k = <STDIN>;
#$coded = "00011111121";  #for testing only

   # substitute lastname in all UPPER CASE
    $strTarget = uc $aPHI{$coded}{'lname'};
    $inLine =~ s/$strTarget/$strSub/g;
 #   print "uc lastname: " . $strTarget . "\n";

   # substitute lastname with first letter Uppercase
    $strTarget = lc $strTarget;    #vanhorn
    $strTarget = ucfirst $strTarget;       #Vanhorn
    $inLine =~ s/$strTarget/$strSub/g;
 #   print "lc lastname: " . $strTarget . "\n";

  # - lastnames with 2 cap letters and perhaps broken into 2 words
    $iLastInd = (length($strTarget)-2);
    for ($ind=2;$ind<$iLastInd;$ind++) {
      $head = substr($strTarget,0,$ind);  #Van
      $tail = substr($strTarget,$ind);   #horn
      $tail = ucfirst $tail;          #Horn
      $strNew = $head . $tail;        #VanHorn
      $strWithSpace = $head . " " . $tail;   #Van Horn
      $strSpaceLCFirst = lc($head) . " " . $tail;
 #     print "cap $ind: $strNew\t $strWithSpace\t $strSpaceLCFirst\n";
      $inLine =~ s/$strNew/$strSub/g;
      $inLine =~ s/$strWithSpace/$strSub/g;
    }

    # handle hyphenated last names; uc OK otherwise ucfirst on both names
    #   strTarget is upFirst last name now
    if (($pos = index($strTarget,"-")) != -1) {
       $head = substr($strTarget,0,$pos);
       $tail = substr($strTarget,$pos+1);
       $hyphLName = $head . "-" . ucfirst($tail);
       $inLine =~ s/$hyphLName/$strSub/g;

       # also get first and second parts of hyphenated used alone
       $inLine =~ s/$head/$strSub/g;
       $tail = ucfirst($tail);
       $inLine =~ s/$tail/$strSub/g;

# for debugging
```

```
#       print "hyphLName: $hyphLName\n";
# $k = <STDIN>;
   }
# $k = <STDIN>;

   # substitute firstname in all UPPER CASE
    $strTarget = uc $aPHI{$coded}{'fname'};
    $inLine =~ s/$strTarget/$strSub/g;
  #  print "uc firstname: $strTarget \n";

   # substitute Firstname with first letter Uppercase
    $strTarget = lc $strTarget;
    $strTarget = ucfirst $strTarget;
    $inLine =~ s/$strTarget/$strSub/g;
  #  print "ucfirst firstname: $strTarget \n";

  # handle various MRNo possibilities
  # assume original MRNo is 11 digits: 4 leading 0's, 7 core digits
    $MRNo = substr($aPHI{$coded}{'realMRNO'},4);
  #  print "MRNo: $MRNo\n";
    $inLine =~ s/$MRNo/$strSub/g;

  # handle form ##-##-## OR ###-##-## OR #-##-##
  #    chop into form ###-##-## first then remove leading 0's
    $dashedMRNo = substr($MRNo,0,3) . "-" . substr($MRNo,3,2) . "-" . substr($MRNo,5);
 #   print "dashedMRNo: $dashedMRNo\n";
    $inLine =~ s/$dashedMRNo/$strSub/g;

# for debugging
#print "near end of replacements\n";
#$k = <STDIN>;

    if (index($dashedMRNo,"0") == 0) {
       $newDashed = substr($dashedMRNo,1);   # gives ##-##-##
#    print "newDashed: $newDashed\n";
       $inLine =~ s/$newDashed/$strSub/g;
    }

  # one more time to get #-##-##
    if (index($newDashed,"0") == 0) {
       $newDashed = substr($newDashed,1);   # gives #-##-##

#       print "newDashed2: $newDashed\n";
       $inLine =~ s/$newDashed/$strSub/g;
    }


  #handle leading 0's in MRNo - only up to 3 because of MRN's like 00000000011
  #   where "11" would then be a valid MRN and also a valid measurement in the text
    for ($char=0;$char<2;$char++) {
```

```
      if (index($MRNo,"0") == 0) {
           $MRNo = substr($MRNo,1);
#          print "MRNo 0's removed: $MRNo\n";
      }
     }
     $inLine =~ s/$MRNo/$strSub/g;

#    }   #eo for $i through numEntries

  $completeLine = $headOfLine . $inLine;

# for debugging/testing
#  print "fLine: $fullLine\n";
#  print "cLine: $completeLine\n";
#  $k=<STDIN>;

  print OUT_FileHdl $completeLine;
#  print "POST: " . $inLine;
#  print "\nPOST-line\n";

 $j++;
# print "\n" . $j . " - end of while\n";

#  show status on screen
  print $j ."\r";

}   # eo while

close IN_FileHdl;
close OUT_FileHdl;
```

Appendix E. Text Mining Programs A, B, C, and D

Run #1

Computer Program A

```
gawk '{
        drugs="(humalog|humulin|iletin|lantus|novolin|novolog|velosulin|relion|insulin)"
        lead=1
        x=1
        while (x != 0)
        {
                if (length($1) == 32) id=$1
                x = match(tolower(substr($0,lead)),drugs);
                if (x != 0)
                {
                        if (length($1) == 32) id=$1
                        print id "|" lead+x-2 "|" substr($0,lead+x-50,99);
                        lead=lead+x+99
                }
        }
}'
```

OUTPUT for Computer Program A

Output included the +/- 50 characters from the drug name identified

0001245585611/01/200101489174ln3|772| etiology. PAST MEDICAL HISTORY: Significant noninsulin-dependent diabetes mellitus for the last 8

0001245585606/26/200301219381lf3|1211|ntrolled. This also may be a sign of worsening insulin resistance from insulin shunting and decr

0001245589912/01/200301331560hp1|1697| (1) Lasix, 40 milligrams PO QD. (2) Lantus, 40 units subcu QD. (3) Protonix, 40 millig

0001245589912/01/200301331560hp4|836| milligrams PO QD and we will give sliding scale insulin as needed. We will check an Accu-Chek tid.

Run #2

The Sentence Tokenizer

```
tr -d "[]()" |\
gawk '{
exception=0
stopwords="^[a-z\.]\{2,3\}\."
        gsub(/\.  /,"\.\n")  # ".  " end of sentence -easy and common
        x=match(tolower($0),stopwords)
        if (!x) gsub(/\. /,"\.\n") # end of sentence with 1 space
        else
        {
                words=split($0,arr)
                while(w != words)
                {
                        printf "%s ",arr[w]
                        if(match(arr[w],/\.$/))
                        {
                                if (!match(tolower(arr[w]),stopwords)) printf "\n";
                        }
                        w++;
                } printf "%s\n",arr[w];
          exception=1
```

```
          w=0
        }
 # delete leading whitespace (spaces, tabs) from front of each line
 # aligns all text flush left
gsub(/^[ \t]+/, "")

 # delete trailing whitespace (spaces, tabs) from end of each line
gsub(/[ \t]+$/, "")

        if(!exception) print
        exception=0
}' | sed -e 's/^ //'
```

Computer Program B

```
gawk '{
drugs="(humalog|humulin|iletin|lantus|novolin|novlog|velosulin|relion|acetohexamide|chlorpropamid
e|glimepiride|glipizide|glyburide|tolazamide|tolbutamide|neteglinide|repaglinide|metformin|piogli
tazone|rosiglitazone|acarbose|miglitol|glipizide\+metformin|metformin\+glipizide|metformin\+glybu
ride|glyburide\+metformin|repaglinide\+metformin|metformin\+repaglinide|exenatide|pramlintide)"
    lead=1
    x=1
    if (length($1) == 32) {   id=$1   sentencenum=1       }
    if (match(tolower(substr($0,lead)),drugs) !=0)
    if (x !=0 ) drugmatch=1; else drugmatch=0
    y = match(tolower(substr($0,lead)),"insulin");
    if (y !=0 ) insulinmatch=1;  else insulinmatch=0
    z = match(tolower(substr($0,lead)),"diabetes");
    if (z !=0 ) sickmatch=1; else sickmatch=0
    if (x+y+z > 0) {
       if (length($1) == 32) id=$1
       patient=(substr(id,1,11))
       date=(substr(id,12,10))
       print patient "|" date "|" drugmatch
       print "|" insulinmatch "|" sickmatch "|" $0;
       sentencenum++
     }
}'
```

OUTPUT FOR PROGRAM B

0001111112101/03/2005E7062596db1|1|MEDICATIONS: Actos 45 mg, Altace 5 mg, atenolol 50 mg, calcium, CellCept, Centrum, clonidine 0.2 mg, Ecotrin 81 mg, fish oil 1000 mg t.i.d., folic acid, Glucotrol XL 20 mg, hydrochlorothiazide 25 mg, Klor-Con, Lantus 20 units, magnesium, Norvasc 5 mg, prednisone 2.5 mg every other day and Prograf 2 mg twice daily.

|0001111112101/05/2006E7830824n22 Lantus insulin 20 units q.h.s., Prograf 1 mg in the morning and 0.5 mg in the evening, glucosamine with chondroitin 2 tablets b.i.d. PHYSICAL EXAMINATION: VITAL SIGNS: Blood pressure is 160/75 with a temperature of 37.1 and pulse of 55.

0001111112101/08/2004E2783821n22|1|In order to potentially decrease her insulin resistance, we have discussed the possibility to taper her corticosteroids off over the course of the next 3 to 4 months.

0001111112101/15/200401018183da3|1|We will start Lantus 10 units to take at bedtime daily.

Run #3

Computer Program C

```
#data set to process
set original=trainingset

#filename to store temporary (unformatted) results
set result=result

#filename to store processed/formatted results
```

```
set finalout=finalout

#A recognized drug name for our purposes consists of a minimum 4 letter word with an optional
trailing parenthetical generic name.
set drug="([a-zA-Z\-]){4,}( \(([a-zA-Z\-]){1,}\))?"

#A number, which may have a decimal.  Allow worded numbers one-ten.
set number="((([0-9/]){1,}([\.-](([0-
9/]){1,}))?)|(one|two|three|four|five|six|seven|eight|nine|ten))"

#list of common s.i. units and volumetric units
set si_unit="(mg|ml|gram|grams|g|puff|tab|tablet)"

#Common delivery abbreviations (i.v., by mouth, rectal)
set delivery="(i\.?v\.?p?\|p\.?o\.?|p\.?r\.?|p\.?v\.?)"

#Latin frequencies
set frequency="([qbt]\.?[0-9a-z]\.?[dh]\.?)"

# The following patterns were used to develop and test during the training set
#pattern 1:  [drug] [number] [si unit] [delivery mode] [frequency]
#pattern 2:  [drug] [number] [si unit] [frequency]
#pattern 3:  [delivery mode] [drug]
#pattern 4:  [drug] [number] [si unit]
#pattern 5:  [drug] [number] [si unit] ([delivery or frequency] [frequency])
#pattern 6:  [drug] [number] [si unit] [delivery or frequency]

# The following patterns worked best and was used in the evaluation of method 3
#pattern 7:  [drug] [number] [si unit] ([delivery or frequency]) ([frequency])

set pattern7="$drug $number $si_unit ?($delivery|$frequency)? ?($frequency)?"
set numbereachline="gawk '{if (match($1,/:/)) {split($1,a,":"); id=a[1]; print $0} else print id
":" $0;'}"
echo Using "$pattern7"
egrep -ino "$pattern7" $original |gawk '{if (match($1,/:/)) {split($1,a,":"); id=a[1]; print $0}
else print id ":" $0;'} >>! $result

cat $result | ~/inls/istudy/code/schepps/mine/join.awk |tr -s ' ' \| |sort > finalout
rm $result
```

OUTPUT FOR SCRIPT C

The script used three flags to indicate the reason that the text had been selected. The first flag indicated that diabetes was in the sentence, the second indicated insulin and the last indicated that a drug name was present. These fields are not mutually exclusive, that is a sentence may have multiple flags set.
Eg.
|0|0|1| = Text was included because it contained a term starting with “diabet”
|0|1|0| = Text was included because it contained a term “insulin”
|1|0|0| = Text was included because it contained a known diabetes drug name

00011111121|04/08/1999|0|0|1|No history of diabetes, CVA.

00011111121|01/08/2004|0|1|0|In order to potentially decrease her insulin resistance, we have discussed the possibility to taper her corticosteroids off over the course of the next 3 to 4 months.

00011111121|02/05/2004|1|0|0|MEDICATIONS: Atenolol 50 mg b.i.d., clonidine 0.1 mg, folic acid, Allegra, prograf 1 mg 3 in the morning and 2 in the evening, CellCept 1000 mg b.i.d., prednisone 5 mg alternating with 2.5 mg, Altace 5 mg b.i.d., multivitamins, calcium, Ecotrin, magnesium oxide, Lipitor 10 mg, Lantus 14 units at bedtime, Glucotrol XL 10 mg every morning reduction to 5 mg was not made, Actos 45 mg every morning, fish oil, Fibercon, Norvasc 5 mg, hydrochlorothiazide 25 mg, Klor-Con and Zetia 10 mg.

Run #4

Computer Program D

```
#data set to process
set original=trainingset

#filename to store temporary (unformatted) results
set result=result
#filename to store processed/formatted results
set finalout=finalout

#A recognized drug name for our purposes consists of a minimum 4 letter word
with an optional trailing parenthetical ge-neric name.
set drug="([a-zA-Z\-]){4,}( \(([a-zA-Z\-]){1,}\))?"

#A number, which may have a decimal.  Allow worded numbers one-ten.
set
number="((([0-9/]){1,}([\.-]((([0-9/]){1,}))?)|(one|two|three|four|five|six
|seven|eight|nine|ten))"

#list of common s.i. units and volumetric units
set si_unit="(mg|ml|gram|grams|g|puff|tab|tablet)"

#Common delivery abbreviations (i.v., by mouth, rectal)
set delivery="(i\.?v\.?p?\|p\.?o\.?|p\.?r\.?|p\.?v\.?)"

#Latin frequencies
set frequency="([qbt]\.?[0-9a-z]\.?[dh]\.?)"

# The following patterns were used to develop and test during the training set
#pattern 1:  [drug] [number] [si unit] [delivery mode] [frequency]
#pattern 2:  [drug] [number] [si unit] [frequency]
#pattern 3:  [delivery mode] [drug]
#pattern 4:  [drug] [number] [si unit]
#pattern 5:  [drug] [number] [si unit] ([delivery or frequency] [frequency])
#pattern 6:  [drug] [number] [si unit] [delivery or frequency]

# The following patterns worked best and was used in the evaluation of method 3
#pattern 7:  [drug] [number] [si unit] ([delivery or frequency])
([frequency])

set pattern7="$drug $number $si_unit ?($delivery|$frequency)?
?($frequency)?"
set numbereachline="gawk '{if (match($1,/:/)) {split($1,a,":"); id=a[1];
print $0} else print id ":" $0;'}"
echo Using "$pattern7"
egrep -ino "$pattern7" $original |gawk '{if (match($1,/:/))
{split($1,a,":"); id=a[1]; print $0} else print id ":" $0;'} >>! $result

cat $result | ~/inls/istudy/code/schepps/mine/join.awk |tr -s ' ' \| |sort >
finalout
rm $result
```

OUTPUT FOR PROGRAM D

ID	MONTH	DAY	YEAR	DRUG	DOSAGE	DOSEUNIT	DELIVERY OR FREQ	FREQ
11657766	2	20	2006	Claritin	10	mg	p.r.n.	(null)
11657766	2	20	2006	Neurontin	300	mg	(null)	(null)
11657766	2	20	2006	Neurontin	300	mg	(null)	(null)
11657766	2	20	2006	Advair	100/50	mcg	(null)	(null)
11657766	2	20	2006	Claritin	10	mg	(null)	(null)
11657912	10	16	1996	Serevent	2	puffs	(null)	(null)
11657912	10	16	1996	Atrovent	2	puffs	(null)	(null)
11657912	5	15	1998	Theodur	200	mg	bid	(null)

Appendix F. Codes for Determination of Death, Myocardial Infarction, and Stroke

ICD-9 Codes

Coronary Heart Disease

ICD	PROCEDURE	DRUG	LAB
MI 410 Angina (413)	Coronary angioplasty (36.09, 00.66) Coronary artery bypass surgery (36.10,36.11,36.12,36.13,36.14,36.15,36.16,36.19)	• MONA(morphine, oxygen, nitro, aspirin) • Thrombolytic drugs(streptokinase, urokinase, alteplase, or reteplase) • Heparin alone • Aspirin	• Cardiac enzymes: -Creatine kinase -Troponin I or T -Lactate dehydrogenase isozymes specific for the heart

Ischemic stroke & hemorrhagic stroke

ICD	PROCEDURE	DRUG	LAB
431 Intracerebral hemorrhage 433 Occlusion and stenosis of precerebral arteries 434 Occlusion of cerebral arteries 435 Transient ischemia attack 436 Acute, but ill-defined, cerebrovascular disease		As ischemic stroke is due to a thrombus (blood clot) occluding a cerebral artery, a patient is given antiplatelet medication (aspirin, clopidogrel, dipyridamole), or anticoagulant medication (warfarin), dependent on the cause, when this type of stroke has been found. Hemorrhagic stroke must be ruled out with medical imaging, since this therapy would be harmful to patients with that type of stroke.	Blood tests (not specify)

Laboratory Codes

CPK_TEST_NUMBER	CPT CODE	C_TEST_NAME	HIGH VALUES
854	82553	CK-MB	>3.4 ng/mL
156	82553	CK-MB	>3.4 ng/mL
1755	84484,84512	TROPONIN I	>1.0 ng/mL

Appendix G. Complete Drug List

DRUGS FOR GLYCEMIC CONTROL	
PRECOSE 25, 50, 100	682092
METFORMIN HCL 500, 850, 1000	682092
GLUCOPHAGE 500, 850, 1000	682092
GLUCOPHAGE XR 500, 750	682092
AVANDAMET 1/500, 2/500, 4/500	682092
GLYSET 25, 50, 100	682092
AVANDIA 2, 4, 8	682092
STARLIX 60, 120	682092
PRANDIN 0.5, 1, 2	682092
ACTOS 15, 30, 45	682092
Note: did not include glucagon emergency kit or 1 mg vial	

DRUGS FOR GLYCEMIC CONTROL	
GLUCOVANCE 1.25/250, 2.5/500, 5/500	682020
METAGLIP 2.5/250, 2.5/500, 5/500	682020
GLYBURIDE MICRO 1.5, 3, 6	682020
GLYBURIDE 1.25, 2.5, 5	682020
GLYBURIDE POWDER	682020
TOLBUTAMIDE 500	682020
TOLAZAMIDE 100, 250, 500	682020
GLIPIZIDE 5, 10	682020
GLIPIZIDE POWDER	682020
CHLORPROPAMIDE 100, 250	682020
ACETOHEXAMIDE 250, 500	682020
TOLINASE 100, 250, 500	682020
ORINASE 500	682020
MICRONASE 1.25, 2.5, 5	682020
GLYNASE PRESTAB 1.5, 3, 6	682020
DIABETA 1.25, 2.5, 5	682020
AMARYL 1, 2, 4	682020
GLUCOTROL XL 2.5, 5, 10	682020
GLUCOTROL 5, 10	682020
DIABINESE 100, 250	682020

INSULIN FOR GLYCEMIC CONTROL	
HUMALOG 100U/ML VIAL	682008
HUMALOG MIX 75/25 VIAL, PEN	682008
HUMALOG 100U/ML PEN, CARTRIDGE	682008
HUMULIN R 100U/ML VIAL, CARTRIDGE ; 500U/ML VIAL	682008
HUMULIN N 100U/ML VIAL, CARTRIDGE	682008
HUMULIN L 100U/ML VIAL	682008
HUMULIN U 100U/ML VIAL	682008
HUMULIN 70/30 VIAL, CARTRIDGE, PEN	682008
HUMULIN 50/50 VIAL	682008
ILETIN I REGULAR 100U/ML	682008
ILETIN II PORK REG 100U/ML	682008
ILETIN I NPH 100U/ML VIAL	682008
ILETIN II PORK NPH 100U/ML	682008
ILETIN I LENTE 100U/ML VIAL	682008
ILETIN II PORK LEN 100U/ML	682008
LANTUS 100U/ML VIAL	682008
NOVOLIN 70/30 100U/ML VIAL, INNOLET, CARTRIDGE, 150U/1.5ML	682008
NOVOLIN R 100U/ML SYRINGE, VIAL, INNOLET, CARTRIDGE	682008

NOVOLIN N 100U/ML SYRINGE, VIAL, INNOLET, CARTRIDGE	682008
NOVOLIN L 100U/ML VIAL	682008
NOVOLOG 100U/ML CARTRIDGE, VIAL,	682008
NOVOLOG MIX 70/30 CARTRIDGE, VIAL, SYRINGE	682008
NOVOLOG FLEXPEN SYRINGE	682008
VELOSULIN HUMAN BR 100U VL	682008
RELION R 100U/ML VIAL	682008
RELION N 100U/ML VIAL	682008
RELION 70/30 100U/ML VIAL	682008
INSULIN R PURE/PORK U100 VL	682008
LIPID-LOWERING AGENTS	
LOVASTATIN 10, 20, 40	240608
MEVACOR 10, 20, 40	240608
PRAVACHOL 10, 20, 40, 80	240608
ZOCOR 5, 10, 20, 40, 80	240608
BAYCOL 0.2, 0.3, 0.4, 0.8	240608
LIPITOR 10, 20, 40, 80	240608
LESCOL 20, 40	240608
LESCOL XL 80	240608
ADVICOR 500/20, 750/20, 1000/20	240608
ALTOCOR 10, 20, 40, 60	240608
NIASPAN 500, 750, 1000	240600
GEMFIBROZIL 600 *(See note at bottom of table)*	240600
NIACOR 500	240600
ZETIA 10	240692
NIACIN CAPSULE SA 125, 250	240692
NICOTINIC ACID POWDER	240692
GEMFIBROZIL 600 *(See note at bottom of table)*	240606
GEMFIBROZIL POWDER	240606
ATROMID-S 500	240606
LOPID 600	240606
TRICOR 54, 67, 134, 160, 200	240606
FENOFIBRATE 67, 134, 200	240606
CLOFIBRATE 500	240606
LOFIBRA 67, 134, 200	240606
COLESTID GRANULES PACKET	240604
COLESTID GRANULES	240604
COLESTID FLAVORED GRANULES	240604
COLESTID 1GM TABLET	240604
QUESTRAN PACKET	240604
QUESTRAN LIGHT PACKET	240604
LOCHOLEST PACKET	240604
LOCHOLEST POWDER	240604
LOCHOLEST LIGHT PACKET	240604
LOCHOLEST LIGHT POWDER	240604
QUESTRAN POWDER	240604
QUESTRAN PACKET	240604
QUESTRAN LIGHT POWDER	240604
CHOLESTYRAMINE POWDER	240604
CHOLESTYRAMINE LIGHT POWDER	240604
CHOLESTYRAMINE RESIN POWDER	240604
CHOLESTYRAMINE PACKET	240604
CHOLESTYRAMINE LIGHT PACKET	240604
PREVALITE POWDER	240604

PREVALITE PACKET	240604
WELCHOL 625MG	240604
ACE-INHIBITORS AND ARBS	
ENALAPRIL MALEATE 2.5, 5, 10, 20	243204
CAPTOPRIL/HCTZ 25/15, 25/25, 50/15, 50/25	243204
LISINOPRIL-HCTZ 10/12.5, 20/12.5, 20/25	243204
CAPTOPRIL 12.5, 25, 50, 100	243204
LISINOPRIL 2.5, 5, 10, 20, 30, 40	243204
ENALAPRIL/HCTZ 5/12.5, 10/25	243204
LEXXEL 5-2.5, 5-5	243204
ZESTRIL 2.5, 5, 10, 20, 30, 40	243204
ALTACE **1.25, 2.5,** 5, 10	243204
ZESTORETIC 10/12.5, 20/12.5, 20/25	243204
CAPOTEN **12.5, 25,** 50	243204
ACCURETIC 10/12.5, 20/12.5, 20/25	243204
ACCUPRIL 5, 10, 20, 40	243204
MAVIK 1, 2, 4	243204
TARKA 2/180, 1/240, 2/240, 4/240	243204
LOTENSIN HCTZ 5/6.25, 10/12.5, 20/12.5, 20/25,	243204
LOTENSIN 5, 10, 20, 40	243204
MONOPRIL 10, 20, 40	243204
MONOPRIL HCTZ 10/12.5, 20/12.5	243204
UNIVASC 7.5, 15	243204
UNIRETIC 7.5/12.5, 15/12.5	243204
MOEXIPRIL 7.5, 15	243204
VASOTEC 2.5, 5, 10, 20	243204
PRINIVIL 2.5, 5, 10, 20	243204
CAPOZIDE 25/15, 50/25	243204
VASERETIC 5/12.5, 10/25	243204
PRINZIDE 10/12.5, 20/12.5, 25	243204
CAPTOPRIL POWDER	243204
** Note: Did not include injectable enalapril*	
DIOVAN HCTZ 80/12.5, 160/12.5, 160/25	243208
DIOVAN 40, 80, 160, 320	243208
HYZAAR 50/12.5, 100/25	243208
COZAAR 25, 50, 100	243208
TEVETEN TILTAB 400, 600	243208
AVAPRO 75, 150, 300	243208
AVALIDE 150/12.5, 300/12.5	243208
ATACAND 4, 8, 16, 32	243208
ATACAND HCTZ 16/12.5, 32/12.5	243208
MICARDIS 20, 40, 80	243208
MICARDIS HCTZ 40/12.5, 80/12.5	243208
TEVETEN HCTZ 600/12.5, 600/25	243208
BENICAR 5, 20, 40	243208

1. All drug entries above came from Drug Master File.A, with the exception of the drugs and/or strengths that are **in bold.** The drugs that are in bold were listed in Drug Master File.B, but not in A.
2. I excluded code 880800 (vitamins and minerals), which included niacinamide powder.
3. Gemfibrozil 600 mg was the only drug I came across that appears to be listed under two different codes.

Appendix H. Abstraction Form for WebCIS Data

Patient ID # (1-78)_____

1.	**Complication patient experienced (either MI, stoke OR death)** MI Stroke Death	**Yes**	**No**
2	**On what date did patient experience complication**? Date: |_____| -|_____|-|_____| Mon Day Year		
3	**On what date did patient have *first* HIGH HbA1c (test result > 6.0)?** Date: |_____| -|_____|-|_____| Mon Day Year		
4.	**How many years did patient have diabetes (use date of first HIGH HbA1c as the start of diabetes) before experiencing the complication?**	_____ # of months	
5.	**Looking back at the patient's history for the 12 months prior to the complication, was the diabetes well controlled? Note: HbA1c ≤ 7% is considered well controlled.**	Yes	No

6.	**Which DM therapies was the patient prescribed in the 12 months prior to the complication?**	**DM therapies** ✓ all that apply	**Date started**
6a	Sulfonylurea		|_| -|_| -|_| M -D -Yr
6b	Metformin		|_| -|_| -|_| M -D -Yr
6c	Thiazolidienedione		|_| -|_| -|_| M -D -Yr
6d	Insulin		|_| -|_| -|_| M -D -Yr
6e	Other (specify drug name)		|_| -|_| -|_| M -D -Yr

GO TO QUESTION 8 IF RESPONSE TO QUESTION 5 WAS "YES" (PATIENT'S DIABETES WAS CONTROLLED)

7	**If patient's diabetes was uncontrolled ("no" to Q5), when did the patient's diabetes become uncontrolled?** Date: |_____| -|_____|-|_____| Mon Day Year	

8a **Changes to the patient's diabetes medications in the 12 months prior to the complication: ☐ no changes**

Diabetes medication changes (describe changes)

____ Added a medication→ which medication added _____ and on what date |_| -|_| -|_|
M –D -Yr

____ Dropped a medication→ which medication dropped _____

____ Increased the dose of a medication→ which medication_____

____ Changed to a different medication class

8b **Outpatient visit frequency when patient's diabetes was controlled?**

____ avg # visits in 12 months when diabetes was controlled

Outpatient visit frequency in the year prior to the complication?

____ # visits in 12 months prior to complication

8c **Frequency of HbA1c testing 12 months prior to complication**

____ # HbA1c tests in 12 months prior to complication

8d. Other significant clinical changes in 12 months prior to complication?

		Yes	No or DK
9	**Does the patient have hypertension (or is the patient on antihypertensive medications)?**		If no, Q 14
10	**Did the patient have hypertension (or was the patient on antihypertensive medication) when he/she was diagnosed with diabetes?**		
11.	**Looking back at the patient's history for the 12 months prior to the complication, was the hypertension well controlled?**		

		HTN therapies ✓ all that apply	**Date started**
12.	**Which antihypertensive therapies was the patient prescribed in the 12 months prior to the complication?**		
12a	Diuretic		|_| -|_| -|_| M -D -Yr
12b	Beta blocker		|_| -|_| -|_| M -D -Yr
12c	ACE Inhibitor		|_| -|_| -|_| M -D -Yr
12d	Angiotensin Receptor Blocker (ARB)		|_| -|_| -|_| M -D -Yr
12e	Calcium channel blocker		|_| -|_| -|_| M -D -Yr
12f	Other (specify drug name)		|_| -|_| -|_|

M -D -Yr

13a	**Did the clinician make any changes to the hypertension medications in the 12 months prior to the complication?** ☐ **no changes** ____ **Added a medication→ which medication added** _____ **and on what date** |_| -|_| -|_| M –D -Yr ____ **Dropped a medication→ which medication dropped** _____ ____ **Increased the dose of a medication→ which medication**_____ ____ **Changed to a different medication class**	
13b	**Other changes in 12 months prior to complication for <u>hypertension control</u>?**	
14	**Does the patient have hypercholesterolemia and/or hypertriglyceridemia (or was the patient on lipid medications)?**	If no, Q 19
15	**Did the patient have hypercholesterolemia and/or hypertriglyceridemia (or was the patient on lipid medications) when he/she was diagnosed with diabetes?**	
16.	**Looking back at the patient's history for the 12 months prior to the complication, was the hypercholesterolemia and/or hypertriglyceridemia well controlled?**	

17.	**Which therapies was the patient prescribed in the <u>12 months prior to the complication</u> for hypercholesterolemia and/or hypertriglyceridemia?**	**Chol/TG therapies** ✓ all that apply	**Date started**
17a	Statin		| M -D -Yr
17b	Fibrate		| M -D -Yr
17c	Bile-acid resin		| M -D -Yr
17d	Niacin		| M -D -Yr
17e	Ezetimibe		| M -D -Yr
17f	Other (specify)		| M -D -Yr

18a **Did the clinician make any changes to the patient's medications for lipid or triglyceride control in the 12 months prior to the complication?** ☐ **no changes**

____ **Added a medication→ which medication added** _____ **and on what date** |_| -|_| -|_|
M –D -Yr

____ **Dropped a medication→ which medication dropped** _____

____ **Increased the dose of a medication→ which medication**_____

____ **Changed to a different medication class**

18b **Other changes in 12 months prior to complication for lipid or triglyceride control?**

19	**Was the patient obese, overweight or normal weight at the time he/she was diagnosed with diabetes?** BMI ≥ 36 BMI between 30 and 35) Overweight (BMI 25-29.9) Normal	Yes	No

20 **In your opinion as a clinician, what factor(s) led to this person's event (MI, stroke, death)?**

		Yes	**No**
21	First high HbA1C is first ever?		
22	DM diagnosed at time of complication?		
23	DM med at first visit		
24	DM Meds		
25	HTN Meds		
26	Lipid Meds		
27	Blood pressure recorded		
28	Weight recorded		
29	Lipid tests recorded		
30	Increased triglycerides only		
31	Gap in care? Length of gap_____months		