A: Specific Aims

Type 2 diabetes mellitus (T2DM) has doubled in the last 2 decades among Hispanic residents of the Lower Rio Grande Valley. In the absence of targeted intervention this epidemic is clearly set to expand within this currently relatively young population, as it both ages and undergoes rapid population growth. The increase is already being further aggravated by onset of T2DM at a younger age, including quite frequently children. The costly and debilitating consequences of complications of T2DM are being seen in larger numbers of patients, and earlier in life than hitherto. While susceptibility to type 2 diabetes mellitus has long been known to have a substantial genetic component, risk factors for T2DM and its many complications encompasses a range of behavioral and life style components which are at presently ill-defined. In 2004 we began recruiting a small cohort of randomly selected residents of Brownsville, Texas, examining genetic components, life style and socioeconomic factors associated with T2DM (MD000170 P20 National Center on Minority Health and Health Disparities). We propose in this study to expand this cohort to a sufficiently large population of participants who can be followed both for the development and evolution of their diabetes, as well as the complications arising as disease progresses. In particular, it is imperative that we expand our understanding beyond the descriptive social and behavioral components of disease and progression such as family history, ethnic background, age and obesity, to the actual biologic underpinnings of disease. It is also critical to assess the contribution of anxiety and depression, which is little studied in Mexican Americans, and which may be a major factor given the stresses within a poor population. The identification of precise risk factors for T2DM and its complications in a Mexican American cultural/gene interaction setting is urgently needed to allow us to move from general approaches to prevention of disease and complications in a Hispanic population.

STUDY HYPOTHESIS:

The prevalence of Type 2 diabetes (T2DM), already disproportionately high in the Hispanic population of the Lower Rio Grande Valley of South Texas, has doubled in the last 20 years. We hypothesize that the particular susceptibility to increase of T2DM among Hispanics is due to interactions of genes predisposing to development of T2DM with currently poorly defined cultural, socioeconomic and psychological factors.

SPECIFIC AIMS:

AIM 1 Expand our current 3000 phenotypically and genetically characterized cohort of Hispanics resident in Brownsville, Texas to 7000 participants. This cohort will continue be full characterized with respect to metabolic syndrome, T2DM, common complications such as premature cardiovascular disease (CVD) or renal failure.

AIM 2 Use genome scanning techniques and access to human gene expression maps to increase our knowledge of the genetic components of susceptibility to T2DM and its complications among Hispanics.

AIM 3 Expand the characterization of socioeconomic, behavioral and lifestyle factors in this cohort, particularly culturally appropriate measures of anxiety and depression, and determine their interaction with genetic susceptibility. Create a GIS linked database to utilize spatial analysis to examine the relationship between socioeconomic, behavioral and lifestyle factors with existing environmental factors (location of fast food, grocery stores).

B: BACKGROUND AND SIGNIFICANCE

Diabetes undeniably poses major health burdens for individuals, families, and populations. In the United States, some 15 million persons have diabetes. Half of these are likely to be undiagnosed and most (90 to 95%) have type 2 diabetes (Kenny SJ et al., 1995). Worldwide, diabetes affects more than 135 million people (King et al., 1998). Estimates in the United States of the economic impact of diabetes are as high as $92 billion or 1 in every 7 health-care dollars expended (Rubin et al., 1994). There is an upward trend in both frequency and costs so that for the foreseeable future the impact is only going to increase. The problem is especially acute in minority populations such as Mexican Americans and African Americans. Already, the Mexican American population bears a disproportionate burden of diabetes with age-specific prevalences that are three to five fold higher than the general population (Hanis et al., 1983). While diabetes unquestionably is a major burden to this population, the majority of the population has not yet reached the prime risk years. As this population ages, the burden will only increase and exceed the already stretched resources. As devastating as these numbers seem to be, they do not capture the real impact of diabetes in terms of its high morbidity and mortality (Hanis et al., 1983). Though we are beginning now to understand some of the metabolic pathways leading to pathology and the genetic predispositions to these, we have little or no understanding of their interaction with socioeconomic, cultural and psychological factors on which to base effective targeted intervention strategies.

Obesity and Diabetes Trends in the US: adults: Obesity is a major precipitating factor in development of the metabolic syndrome and T2DM (Reaven, 1988; Reaven, 1993; Lebovitz, 2003). By 2000, obesity resulting from poor diet and limited physical activity accounted for 18.1% of deaths in the US, a close second only to tobacco consumption, which it is shortly predicted to overtake(Mokdad et al., 2004). Age adjusted prevalence of obesity body mass index (BMI) $\geq$30 (calculated as weight in kilograms divided by the square of height in meters) in the United States has increased from 22.9% (NHANES III 1988-1994), to 30.5% in 1999-2000(Flegal et al., 1998), and extreme obesity (BMI$\geq$40) from 2.9% to 4.7%. These trends were observed in both sexes, all ages and all racial groups. In persons over the age of 40, the increase in prevalence of diabetes has risen from 8.9% (NHANES II 1976-1980) to 12.3% (NHANES III 1988-1994)(Harris et al., 1998). In Mexican Americans the risk of diabetes is almost twice that of non-Hispanic whites, and Mexican American men, in particular, have greatly increased rates of undiagnosed diabetes and impaired fasting glucose compared with other groups(Harris et al., 1998). Data from Texas, where there is a large Mexican American population, reveal that self reported obesity is high, 22.7% compared with 19.8% nationally, (BFRSS 2000) behind only Mississippi, Louisiana and West Virginia. Self-reported diabetes in Texas is 7.1% compared with 7.3% nationally(Harris et al., 1998; Mokdad et al., 2000b). Nationally age-adjusted prevalence of obesity in Mexican Americans increased in men from 23.9% to 28.9% (NHANES 1988-1994 and1999-2000) and in women from 35.3% to 39.7%(Flegal et al., 2002).

Nationally in the United States, fasting plasma diagnosed diabetes increased from 4.1% to 5.5% in men and 5.6% to 7.4% in women (NHANES 1988-1994 and 1999-2000)(Mokdad et al., 2000a). In persons over the age of 40, the increase in prevalence of diabetes has risen from 8.9% (NHANES II 1976-1980) to 12.3% (NHANES III 1988-1994) (Harris et al., 1998). It now appears that In Mexican Americans the risk of diabetes is almost twice that in non-Hispanic whites, and Mexican-American men, in particular, have greatly increased rates of undiagnosed diabetes and impaired fasting glucose compared with other ethnic groups (Harris et al., 1998). In Mexican Americans the percentage of diabetes in patients >20 years was 9.3% (NHANES 1988-1994)(Harris et al., 1998). In Texas the increase in fasting plasma glucose diagnosed diabetes in NHANES data (1988-1994 and 1999-2000) was from 4.8 to 6.6% (both sexes(Mokdad et al., 2000a), but self-reported diabetes (BFRSS) was 7.1%(Mokdad et al., 2001) On the South-Texas Mexico border age specific prevalence of self reported diabetes has also increased (see figure 1, preliminary data).

The Mexican American population of the Lower Rio Grande Valley: Brownsville is located along the US/Mexico border in south Texas in a region known as the Lower Rio Grande Valley. The city has a population of 139,722 (population density 671/km^2) living in 38,174 households with an average size of 3.62 persons. Forty-five percent of children in Brownsville live in poverty, the highest proportion of any city in the U.S. with a population over 100,000 (245th out of 245 cities >100,000). Thirty-two percent of families in Brownsville live below the poverty level, with a median household income of $24,468. Ninety-one percent of the population is Hispanic or Latino (74% Mexican American), and 8% are Anglo. Thirty-two percent are foreign born, and 98% of these

individuals were born in Latin America (primarily Mexico)(Federal Reserve Bank of Dallas, 2001). Rates of obesity and type 2 diabetes in the Lower Rio Grande Valley are the highest in Texas, which already has some of the highest rates in the United States(Office of Border Health and Texas A&M Public Policy Research Institute, 1 A.D.).

Genetics of T2DM and Mexican Americans: Population studies, pedigree investigations, animal models and molecular studies consistently implicate a substantial role of genes in determining the risk for type 2 diabetes (DeFronzo RA, 1997a; DeFronzo RA, 1997b). These studies establish that no simple genetic model adequately explains risk for diabetes; rather, there are likely to be multiple genes with small to modest effects that interact with environmental factors in producing susceptibility (Hanis CL, 1996). It is estimated that 31% of the contemporary Mexican American gene pool is Native American derived (Hanis et al., 1991). The observed 3.1 and 3.3 fold increased risks of diabetes in male and female siblings, respectively, of those with diabetes in the Mexican American population of Starr County are consistent with genetic variation at a number of loci having small to moderate effects (Hanis et al., 1996). This view of the genetics of diabetes further implies that we are looking for genes whose effects are neither necessary nor sufficient to produce diabetes. To date, no less than 9 complete genome searches have been or will soon be completed for type 2 diabetes or related traits. Of these, our own (Hanis et al., 1996) and those of Mahtani et al. (Mahtani et al., 1996), Hanson et al. (Hanson et al., 1998; Elbein et al., 1999), Elbein et al. (Elbein et al., 1999) and Hegele et al. (Hegele et al., 1999) have reported results for the disease diabetes. Increasingly, multipoint maps are becoming available and results are being posted on the World-Wide Web (e.g., results are being compiled by the International Type 2 Diabetes Linkage Analysis Consortium at http://www.sph.umich.edu/group/statgen/consortium/). While the picture is still evolving, these various studies have not identified a single major susceptibility locus that explains the majority of type 2 diabetes risk across population groups. This adds weight to the argument that there are likely several loci contributing to diabetes risk.

Mental Health in Hispanics: In the United States, mental health disorders have been shown to account for four of the five causes of premature death and disability in adult populations (National Institute of Mental Health, 2005). Research has shown that unadjusted rates of depression are more prevalent among minority populations, particularly Hispanics (Centers for Disease Control, 1998; Swenson et al., 2000; Dunlop et al., 2003);, In fact the World Health Organization predicts depression will be one of the leading causes of death and disability by the year 2020 (Motl et al., 2004).

In a national sample, prevalence rates of major depression were associated with sociodemographic, health and economic factors, such that upon adjustment Hispanic depression rates were similar to white rates. However , higher rates of depression were related to presence of serious diseases, functional limitations, absence of health insurance, and lifestyle / behavioral factors such as smoking and exercise patterns(Dunlop et al., 2003; Fisher et al., 2001). Another study of depressive symptoms in Hispanic and non-Hispanic white elderly found that after adjusting for sociodemographic factors and health, older Hispanic women with lower acculturation levels had increased risk for depression (Swenson et al., 2000).

Several studies have found that persons with diabetes have comorbid depression, anxiety or other mental disorders (Anderson et al., 2001; Grigsby et al., 2002; Black et al., 2003; Dinan et al., 2004; McVeigh KH et al., 2004). In fact, one study found that after controlling for sociodemographic factors (age, sex, education and acculturation and marital status) that the interaction of depression and diabetes not only predicted greater incidence of mortality, complications and disability, it also predicted earlier onset of adverse events (Black et al., 2003).

*Environmental / spatial distribution of diabetes risk factors and outcomes: O*ne of the modern technologies available to explore potential risk factors for T2DM and its complications in a Mexican American cultural/gene interaction setting is the use of geographic information system (GIS) and spatial analysis. This new technology provides an opportunity to examine spatial relationships between purely traditional statistical associations. As in most epidemiological studies, the questions that we seek to answer have inherently spatial components. We are anticipating that examining the spatial distribution of obesity measures and T2DM incidences may provide insight into their association.

A GIS is an automated system for the capture, storage, retrieval, display, and analysis of spatial data. Spatial data within a GIS are unique because their records are spatially referenced to their geographic positions. The

investigation of T2DM prevalence involves data concerning people, their health, their socio-economic characteristics, and the environments in which they live. Since all the elements of the data contain specific geographic locations, they can be spatially referenced into a GIS. With all the datasets organized as data layers in the GIS, further analysis of the layers can then be conducted within a common geographic framework. As an enabling technology, GIS can play an important role in providing an integrated environment, within which specialists from different disciplines can potentially work together to explore and integrate different methodological approaches in this project .

The increasing availability of public domain digital map data from many government agencies at different levels makes it less laborious to collect various data layers about the behavioral, socioeconomic, and environmental characteristics of the cohort. For this project, we plan to search and import digital map layers of road networks, hydrological drainage networks, digital elevation models, toxic release inventory (TRI) sites, air and water pollution monitoring stations, distributions of fast-food restaurants, census tract data, satellite images, and aerial photography into our spatial database. These imported data layers will most likely have varied datum and projection types. They will need to be reprojected into a common coordinate system to enable map overlay and other analyses within the GIS. Geoprocessing operations will also be performed to generate a variety of valuable secondary properties of the data layers, such as the slopes and aspects derived from digital elevation modeling (DEM), the distances to the nearest fast-food restaurants, parks, biking trails, exercise facilities, the air pollution concentration surfaces interpolated from monitoring stations, and the urbanization statuses, including density, interpreted from remote sensing imagery and YR 2000 Census information. This richer information will expand our abilities to explore and to discover potential contributors to T2DM in a wider space.

Maps have been traditionally used by epidemiologist for identifying the causes of illnesses of unknown origin (Clarke et al., 1996), with John Snow's classic maps of cholera cases in relation to the Broad Street pump being an early but excellent example. GIS technology has revolutionized the map-making process with powerful display tools that allow flexible comparisons of maps to explore associations between local environment and disease. However, pure spatial visualization can easily be misleading (and misused) without objective statistical tests. The use of GIS in this project will go “beyond the simple mapping” and will adopt an exploratory spatial data analysis (ESDA) approach. ESDA relies on spatial statistical tools to locate unusual data outliers, to identify possible clusters of high incidences of disease, to seek unique spatial patterns and associations, and to ultimately develop rigorous hypotheses in the absence of any known cause for the disease.

C: PRELIMINARY STUDIES/PROGRESS REPORT

We are in the process of expanding programs for diabetes research and prevention at the Brownsville Regional Campus of the University of Texas School of Public Health. Thus, this project builds on several currently funded projects. Common to all of these projects is the development of appropriate population for study of diabetes and its complications in Mexican Americans. Once developed (or once the infrastructure for development is in place), long-term research and intervention will be greatly facilitated.

The Starr County Studies: Starr County(98% Mexican-American); one of 14 Texas counties bordering with Mexico, is at the upper end of the Lower Rio Grande Valley, about 75 miles from Cameron County in which we find Brownsville. The population of both counties is predominantly Mexican American, with the same culture and origins, the only difference being that Brownsville is relatively more affluent, though still one of the poorest cities for its size in the United States. For more than 20 years, we have made a concerted effort to understand the distribution and burden that type 2 diabetes and its complications pose for the Mexican-American population of Starr County, Texas. A primary aim is to identify those genetic factors responsible for the three-fold increased risk of diabetes that this population bears. A series of genetic and epidemiologic investigations related to type 2 diabetes included the first demonstration of the elevated risk among Mexican Americans (Hanis et al., 1983); a fact that is now largely taken for granted. These studies resulted in the completion of the first genome scan to localize genes for type 2 diabetes (Hanis et al., 1996), and subsequently to the identification of a gene, *calpain 10*, that is the first gene identified as contributing to type 2 diabetes (Horikawa et al., 2000). This was the first identification of a gene for any common chronic disease using a paradigm of genetic linkage studies followed by fine scale mapping and gene identification. We participate in the single largest effort ever funded by the National Heart, Lung and Blood Institute to identify genes for hypertension.

The sample we are collecting in Starr County is the only Hispanic population represented in this large multi-centered effort. We are also conducting the single largest effort funded by the National Eye Institute to map genes for diabetic retinopathy.

Over the years, some 8,000 individuals have undergone from 1 to 11 physical evaluations. Over the next 5 years, we are funded in a variety of projects to perform approximately 6,000 additional examinations. We employ a staff of 10 full-time individuals in Rio Grande City who perform; electrocardiograms, echocardiography, stereoscopic fundus photography of the eye, oral glucose tolerance tests or intravenous glucose tolerance tests. Additionally, we perform approximately 4,500 annual glucose tests on walk-in individuals (something that we have done since 1988).

We have also engaged in a series of intervention studies to prevent disease or slow the development of the processes by which health gives way to disease. This has included state of the art interventions aimed at obesity and diabetes education. We are now in our 8th year of a diabetes education program that resulted in the development of a culturally sensitive program that we have demonstrated leads to significant improvement in diabetes control among Mexican Americans. We are now testing those elements of the program that are critical so that it can be implemented on a much broader scale.

NIDDM Genes in Blacks, Hispanics and Non-Hispanic Whites – DK47487 (PI – Hanis): The main purpose of this project currently in its 8th year is to localize and identify genes for type 2 diabetes using the original group of 346 affected sibling pairs on whom genetic markers have been typed across the genome; a representative sample of 1,004 individuals aged 15 to 74 with detailed phenotypic characterizations and 1,254 individuals distributed in 250 nuclear pedigrees. We continue to fine map regions of interest identified in the linkage scan and to sequence and genotype positional and candidate genes. Our approach is to resolve the contribution of variation in a candidate gene by typing a set of single nucleotide polymorphisms (SNPs) in the gene that captures the array of variation. In this way we can resolve the gene without being left with a conclusion that there may be variation in linkage disequilibrium with a typed marker. This project also involves recruiting and testing of a new cohort of 1,255 individuals aged 20 to 50 years. Each individual receives a complete examination and a full oral glucose tolerance test. A subset of 485 of these individuals will undergo a frequently sampled intravenous glucose tolerance test. This cohort is genotyped at select genes to test for potential physiologic effects of variation. Those undergoing the more detailed testing are invited according to their genotype so that we can utilize balanced statistical designs in order to maximize power for detecting effects.

Genes and High BP: Mexican American Field Center – HL54504 (PI – Hanis): This is a large multi-centered project to map and identify genes for hypertension. The project is now in its 8th year. In the first phase of the project, 1,805 Mexican Americans were recruited and examined. Recruitment required a pair of siblings with type 2 diabetes and then included all other siblings. Over 1,000 affected pairs were collected and genotyping recently completed for a genome scan. In the ongoing phase 2 of this project, all individuals are being reexamined to determine changes in critical risk factors. Additional phenotyping is also being conducted and includes echocardiography.

Genetics of Diabetic Retinopathy – EY12386 (PI – Hanis): The extensive sibling resource developed in the above project for whom genetic markers are already available made it possible to consider the genetics of diabetes complications. To this end, all diabetic individuals are being recruited for additional examinations that involve stereoscopic fundus photography for the classification of diabetic retinopathy. Phenotypic characterization has occurred and we are now determining the familial aggregation of diabetic retinopathy and will begin shortly a genome scan to localize genes for retinopathy. This project also involves a large laboratory commitment involving fine mapping, sequencing and genotyping of select candidate genes.

Education and Group Support for Diabetic Hispanics – DK48160 (PI – Brown): While we continue to conduct basic research aimed at understanding the pathways involved in diabetes and its complications, we also have sufficient information to begin interventions. We therefore developed a culturally sensitive diabetes education program that we have demonstrated to be effective in reducing glycosylated hemoglobin levels and increasing diabetes knowledge. In the renewal of this project (now in the 9th year) we are testing whether elements of the education program can be simplified so that it might be implemented on a larger community basis.

Increasing prevalence of diabetes and impaired glucose tolerance along the US/Mexico border (manuscript in preparation).

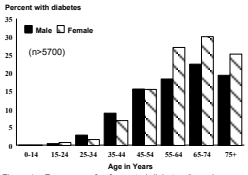

Figure 1. Frequency of sef-reported diabetes (based on census report) by gender in the Lower Rio Grande Valley, 2001-2004.

At our Brownsville site, we have since 2004 a growing program studying diabetes following the protocols established in Starr County. Frequency of diabetes in our population derived from our studies is shown in figure 2. These studies extend prospective, ongoing populations-based surveys of diabetes and the genetics of diabetes in Starr County in existence since 1981 (PI Dr. C. Hanis) The Brownsville diabetes studies form the diabetes research core of our EXPORT grant 5 P20 MD000170 funded by NIH NCMHD, entitled Creation of an Hispanic Health Research Center (HHRC) in the Lower Rio Grande Valley (PI Dr. McCormick). The diabetes core (PI Dr. Craig Hanis), aims to identify and phenotypically characterize a representative cohort of 800 Mexican Americans residing in Brownsville, Texas, with respect to T2DM status, permitting estimation of prevalence and establishing a cohort for genetic association testing (Bosque-Plata et al., 2004; Hanis et al., 1996; Hanis et al., 1983; Hanis et al., 1985; Hanis et al., 1991; Horikawa et al., 2000). Preliminary analysis of the first 250 Brownsville individuals screened in 2004 reveal the data in Table1.

Table 1: Prevalence of Previously Diagnosed and Newly Diagnosed T2DM (T2DM) and Impaired Fasting Glucose (IFG) among Mexican Americans in Brownsville, Texas, 2004 data.

Age Group	n	Previous T2DM (%)	Newly Diagnosed T2DM (%)	Any T2DM (%)	IFG (%)	T2DM or IFG (%)
35 – 44	107	2 (1.9%)	1 (0.1%)	3 (2.8%)	2 (1.9%)	5 (4.7%)
45 - 54	55	5 (9.1%)	6 (10.9%)	11 (20.0%)	7 (12.7%)	18 (32.7%)
55 – 64	57	11 (19.3%)	5 (8.8%)	16 (28.1%)	3 (5.3%)	19 (33.3%)
Totals	219	18 (8.2%)	12 (5.5%)	30 (13.7%)	12 (5.6%)	42 (19.2%)

Diabetes classification was based on the current recommendations of the National Diabetes Data Group(NDDG, 1997). Individuals previously told that they had diabetes, who were currently on glucose lowering medications represent the previously diagnosed group. Newly diagnosed T2DM was based on a fasting glucose of 126 mg/dl or greater. Impaired fasting glucose was based on a fasting glucose greater than or equal to 110 mg/dl, less than 126 mg/dl, and not otherwise classified as having diabetes. These data document a tremendous burden of T2DM in this population. Of particular concern are the numbers of newly diagnosed cases and the numbers with impaired fasting glucose. These speak to a much larger burden than would be anticipated otherwise. These prevalences are also higher than any that have been reported previously in the Valley (figure 1). (Hanis et al., 1983) Whether these numbers reflect a background rate of diabetes that is inherently higher than other areas of the Rio Grande Valley or whether this represents a quickening of the development of T2DM remains to be seen.

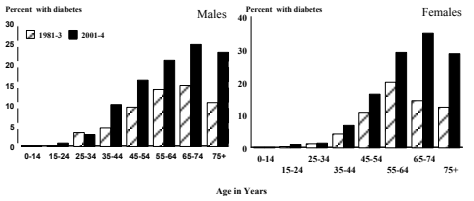

Figure 2. Comparison of frequency of diabetes in males and females

D: Research Design and Methods

Aim 1 – ***Expand our current 3000 phenotypically and genetically characterized cohort of Hispanics resident in the South Texas Border, to 7000 participants. This cohort will continue be fully characterized with respect to metabolic syndrome, T2DM, common complications such as premature cardiovascular disease (CVD) or renal failure.***

A major effort of this project is the examination of 7000 Mexican Americans representative of the adult population of South Texas border cities. Recruitment and examinations are patterned after those methods we have successfully used in Starr County over the past 2 decades. We will randomly select blocks according to dwelling unit density (i.e., number of units per block as opposed to numbers of individuals per dwelling unit). On selected blocks, the field team will enumerate all individuals to determine age, gender, ethnicity and diabetes status based on history. We will enroll individuals aged 18 years and over. These individuals will be invited for a physical examination at the Regional Campus Clinical Research Center. We will recruit and examine 5-10 individuals per week. Each examination takes an average of 2½ to 3 hours, and is patterned after that employed with the random cohort for the NIDDM genes project. Following obtaining informed consent, the examination will consist of:

Laboratory: Collection of fasting (overnight) blood specimens, total of 40 ml by venipuncture; two 10 ml EDTA vacutainers, one 10 ml Red top tube, one Pax Gene tube for RNA extraction, also two 2 ml EDTA and one 3.5 ml SST vacutainer tube. As cost effective and to gather more and fast data we are sending 7.5 ml of blood to the Community Reference Laboratory of Valley Baptist Medical Center to run CMP, Lipid panel, CBC-Diff, and A1C. Values of the already mentioned assays are going to be input into our database and use in analyses. Results of this test will also be given to participants to provide feedback. Following processing and aliquoting. Samples will be placed in a - 80° C freezer. We will also collect a urine specimen for determination of renal and liver function, acid-base balance, bacteriurea, and carbohydrate metabolism through the use of UA dipsticks remaining samples will be place in - 80° C freezer. The results of this test will be used to correlate with blood tests that are already being done. Sample aliquots (plasma, red cells, DNA, RNA and urine) will be maintained in Brownsville campus.

Anthropometric Examination: Weight, height, and body circumference measures are obtained on each individual. Weight and circumferences are each measured twice. For circumferences, we use two individuals - one to measure and one to insure that the tape is maintained parallel to the ground. Circumferences are measured 3 times. Additionally, we will determine body fat composition by bioimpedance.

Blood Pressure: Systolic and diastolic blood pressure will be determined following standard protocols. Participants will sit quietly for 5 minutes and then have 3 blood pressure determinations using a Welch Allyn vital signs machine in the dominant arm then another 3 measurements in the appositive arm.

Obstetrical Parameters: One page questionnaire to collect the total number of pregnancies, method of delivery, birth defects , gender, baby's weight at birth, mom's weight gain by pregnancy, complications during pregnancy, prenatal visits, and infant & childhood deaths, as well as mammogram and pap smear history and contraceptive method.

Medical and Medication Histories: Identifying and demographic history will be obtained. Smoking and alcohol usage will be determined, and fertility data obtained (if necessary). Diabetes histories and medication histories will also be completed. The medication history obtains information on both prescription and non-prescription medications (including folk remedies). The history will also review cardiovascular and cerebrovascular health. (Appendix 1)

Demographic data: Identification questions, gender, place of birth, parents and grandparents place of birth, and phone numbers, ethnicity, employment status, occupation, type of insurance if any, income per year, and number of years living in the county.

Family History data: Biological family history, parent's age, or age at death, medical history of parents, grandparents and siblings.

Diabetes History: Year or age at diagnosis, medication for diabetes, diabetic retinopathy if diagnosed & treatment, ketoacidosis & where diagnosed, kidney dialysis and treatments, history of sore, ulcers, blisters, and in what part of the body, complications due to diabetes, and if taking home remedies for the diabetes.

Smoking and Drinking History: Cigarettes and Alcohol consumption.

Electrocardiography: Patients will be invited to be examined by electrocardiography using standard equipment and techniques.

Exit Interview: All individuals are given an exit interview to review immediate findings. They are also provided with a written summary including their height, weight, blood pressure, glucose, CBC-Diff, A1C, cholesterol and triglyceride values. Values out of their expected range are flagged and referrals made as necessary. This interview is also used to address any questions or concerns that may be raised.

To facilitate classification of diabetes status, any individual with a fasting glucose level of 120 mg/dl or greater will be scheduled for a second fasting sample. While this procedure exceeds current recommendations for diabetes classification, it is comparable to what we are using in Starr County and will permit classification of diabetes based on the newer more relaxed criteria or the previous more restrictive criteria.

Aim 2 – ***Use genome scanning techniques and access to human gene expression maps to increase our knowledge of the genetic components of susceptibility to T2DM and its complications among Hispanics. Expand the characterization of socioeconomic, behavioral and lifestyle factors in this cohort, including culturally appropriate measures of anxiety and depression, and determine their interaction with other factors including genetic susceptibility. Using these analyses, we will develop targeted approaches to intervention for treatment and prevention.***

Table 2 . Genes/markers Currently being Analyzed

Locus or Marker	Details	Reference
Calpain 10 SNP43 and SNP56	2 SNPs (SNP43 and SNP56) based on linkage and association of SNP43 and some evidence of interaction of SNP56 with SNP43	Horikawa et al., 1999
Calpain 3	2 SNPs detectible following Nde or RsaI digestion in the 3' untranslated region based on evidence of linkage and association and interaction with Calpain 10	Ongoing work
PPP1R3	ARE2 interval alleles based on RNA-protein interactions among Pima Indians	Xia et al., 1999
CRP/APOA2	Chromosome 1, Utah Linkage, possible replication	Elbein et al., 1999
Sorbinol Dehydrogenase	Biological Candidate - Exon 9 pro357leu - preliminary retinopathy association; Assume 1 other polymorphism	Hanis Resequencing
Chromosome 7	D7S502 - Conditional analyses D7S1799 - Possible Pima linkage	Cox et al., 1999 Hanson et al., 1998
D11S4464	Possible linkage with diabetes and obesity among Pima	Hanson et al., 1998
D6S1009	Possible linkage with diabetes among Pima	Hanson et al., 1998
Chromosome 20	D20S197 -Diabetes linkage, nephropathy linkage D20S905 - Diabetes linkage D20S107	Ji et al., 1997 Ghosh et al., 1999 Ghosh et al., 1999
Aldose Reductase	Promotor C(-106)T Assume 1 other polymorphism	Kao et al., 1999 Hanis Resequencing
Sulfonylurea Receptor	Physiologic Candidate and Associations with diabetes - 2 polymorphisms (albeit they are not common)	Inoue et al., 1996
Other Calpains	Apparent role of the calpain gene family Assume 3 markers	Bell Resequencing Hanis Resequencing
Proopiomelanocortin	RsaI and 7566C/T Polymorphisms - linkage and association with leptin/obesity	Hixson et al., 1999
MTHFR	C677T - association with complications	Vaccaro et al., 1999
TNF-β	Association with complications	Hawrami et al., 1996

As described above, we are involved in a series of investigations to localize and identify genes for type 2 diabetes. We will expand these studies to include the Brownsville Cohort samples. We will also expand these studies to consider genetic variation at three additional genes. DNA sequencing will be performed in Houston, while genotyping will be done at both sites. This will lay the foundation for infrastructure and training opportunities at the Brownsville site. Samples collected in the Brownsville Cohort (n = 3000) will be analyzed in Brownsville, while our Starr County samples (focus will be on 1,000 representative samples) will be analyzed in Houston.

We are also involve in the genome sequencing and will transfer unidentified specimens to Washington University for analysis. The genomic characterization of our cohort population will be deidentified as in all our analyses of biological and other variables. The characterization will allow us to determine genetic determinants of disease and potential susceptibility to disease in Mexican Americans on which to base intervention strategies. This will provide us with the tools for a degree of personalized medicine at the population level. The range of diseAIM 1ases of interest include many serious complications of obesity and diabetes, both of which are highly prevalent in the cohort. These diseases include cardiovascular and renal disease as well as end-stage liver disease, eye diseases and other diseases associated with metabolic disorders. We hope also to be able to use this information for new drug discovery, as well as determine genetically governed responses to standard and new therapy for some of the diseases listed above.

The information in this study will not be linked to individuals in the cohort and will therefore not affect them directly in any way. However, the benefits to the community to which they belong will be considerable. This is one of the first attempts to determine genetic predisposition to disease in Mexican Americans, which have hitherto not been well studied, and will open new doors for prevention and intervention.

The genotyping effort is quite modest. We have considerable expertise genotyping on a large scale. Currently funded projects allow for some 500,000 locus tests. The choice of which markers to type is highly subjective. Initially we will type three markers in *calpain 10* that have been shown to be associated with diabetes among Mexican Americans in Starr County (SNPs 43, 19 and 63). The choice of the additional genes to be examined will be selected as we get closer in time. Genes that are currently being examined in funded projects are in the table 2.

To choose markers to be genotyped in the full data resources, we will sequence 20 individuals in both directions for a total of 40 reactions per individual per region. Assuming on average, that each gene will require sequencing of some 18 regions leads to 720 reactions per gene. Greatly facilitating this is that we are not sequencing de novo, but are re-sequencing regions for which the sequence has already been determined. Our use of a temperature gradient thermocycler has compressed the time required for optimization due to the large number of conditions that can be simultaneously tested in a single run. Sequencing proceeds using BigDye Terminator cycle sequencing according to manufacturers recommendations (ABI Prism Sequencing Kits, Foster City, California) and the ABI Prism 377 DNA Sequencer. Our experience is that the BigDye terminators are superior to other chemistries for detecting heterozygotes. This is critical because most variation detected will be in heterozygotes. Once the array of variation is determined, we will identify haplotype tags that capture the majority of genetic variation. These tag sites will be genotyped.

A variety of genotyping approaches are available. For polymorphisms altering restriction sites (or where a restriction site can be generated), amplified DNA will be digested with the appropriate enzyme following PCR. PCR amplifications will be carried out in 96-well plates with the aid of a Biomek 2000 workstation and multichannel pipettes. Oligonucleotide primers (20 to 24 bp) for PCR amplification will be purchased from Research Genetics, Inc (Huntsville, AL) or Gibco, Inc (Grand Island, NY). Each amplification series will contain a negative control (water) to ensure absence of contamination as well as a sample of known genotype. Digested PCR product will be electrophoresed on acrylamide or agarose gels depending on the expected size of the fragments. Ethidium bromide stained gels will be photographed under ultraviolet light and labeled with preprinted identification labels to allow independent genotyping by two technicians.

Not all polymorphisms will be detectible based on restriction digestion. We have the capability to type variation using direct sequencing. At the outset it should be pointed out that the methodology for typing single nucleotide polymorphisms is changing rapidly. For example, dense DNA arrays (a.k.a. chips) and mass spectroscopy show promise for single nucleotide polymorphism genotyping. However, their use is not yet routine and the sensitivity and specificity of competing techniques remains a major issue. Therefore, we propose to use the TaqMan assay for markers not detectible using restriction enzymes. This also corresponds to existing equipment in Brownsville (ABI7900 Prism).

The TaqMan assay uses fluorogenic probes in a 5= nuclease assay to identify differences in DNA sequence. For high through-put processing, we will use the laboratory=s ABI Prism 7700 Sequence Detection System (ABI 7900 in Brownsville). Briefly, allele-specific probes approximately 20-30 bp in length are labelled at the 5=

and 3= ends with fluorescent reporter and quencher dyes, respectively. These probes are blocked at the 3= end to prevent extension during PCR. The proximity of the reporter dye molecule to the quencher dye molecule masks the fluorescent activity of the reporter dye as long as the probe remains intact. During the annealing and extension phase of the PCR reaction, primers and probes bind to the DNA strand in a site-specific manner. As the *Taq* DNA polymerase extends the DNA strand from the primer, its 5= nuclease activity degrades the bound probe and releases the reporter dye, causing an increase in the fluorescence intensity of the reporter dye. Doubly-dye-labeled probes will be obtained from Perkin-Elmer (Norwalk) and primers will be obtained from Genosys (Houston). Fluorescent activity will be determined using an ABI 7700, which has the capability of reading and analyzing a PCR plate in less than one minute, within two hours after completion of PCR. Quantitation of fluorescence is made by comparing each sample=s fluorescent activity to that of a background dye (ROX) present in the reaction buffer, a blank standard containing no DNA, and samples of known genotype (when available).

As mentioned above cardio vascular disease is associated with metabolic disorder, thus CVD is a very costly disease both economically and in terms of human mobility and mortality. In an attempt to decrease the incidence of CVD, much attention has been focused on controlling or improving the lipid profile of patients who are at risk of disease.

Although CVD risk is heritable and there have been a number of successes localizing QTLs that influence disease risk, the genetic basis of this risk is still relatively unknown. Quantitative endophenotypes that are related to disease liability can offer more power for gene localization/identification than dichotomous disease status and thus serve as valuable phenotypes for disease gene identification. Evidence from epidemiological studies has shown that specific lipids and their constituent components are important factors in the development of CVD. The classical lipid traits (such as HDL-C, LDL-C, and triglycerides) that are most commonly examined in disease risk are complex molecules comprised of multiple lipid and protein components. The human lipidome contains many thousands of these individual lipid species. We hypothesize that these lipid components may represent unique biological measures that are closer to the direct action of genes and, hence, may lead to the more rapid discovery of genes causally involved in CVD risk.

We aimed at assessing the influence of genetic variation on the human lipidome and prediction of CVD in Mexican American individuals. We will perform lipid profiling and focused genotyping in an independent population of Mexican American individuals from the Cameron County Hispanic Cohort to replicate those variants identified in the San Antonio Family Study discovery cohort that appear to jointly influence lipidomic endophenotypes and CVD risk.

Targeted lipidomic analysis will be performed on all 2,500 CCHC subject, acquiring the same lipid profile, including additional untargeted lipid species identified as we acquired for the SAFS cohort. For replication, we will genotype all 2,500 CCHC individuals using the Illumina OmniExpressExome-8+ BeadChip customized with an additional 3,000 bead types.

AIM 3 Expand the characterization of socioeconomic, behavioral and lifestyle factors in this cohort, including culturally appropriate measures of anxiety and depression, and determine their interaction with genetic susceptibility. Use spatial analysis to examine the relationship between the above factors and environmental factors in a spatial analysis.

The proposed cohort study will assess the following outcomes and mediating variables associated with depression and diabetes. Proposed measures are based on a national cohort study of adults aged 54-65 who completed the Health and Retirement Survey including the World Health Organization's Composite International Diagnostic Interview- Short Form described by Dunlop, Song, Lyon et al, 2003. Other measures are based on items from the Behavioral Risk Factor Surveillance Survey. And still other measures are based on the National Health survey available at http://www.cdc.gov/nchs/nhis.htm.

Table 3. Proposed Socioeconomic, Behavioral and Lifestyle Measures Associated with Diabetes and Depression

Construct	Specific variables (as needed)	Instrumentation
Depression		Composite International Diagnostic Interview- Short Form
Sociodemographic characteristics	Race, age, gender, living arrangements, marital status, family structure	Behavioral Risk Factor Surveillance Survey Health and Retirement Survey National Health Survey
Health Needs	Physical health, health behavior, and functional limitations	Behavioral Risk Factor Surveillance Survey Health and Retirement Survey National Health Survey
Health Behaviors	Physical exercise, Current smoking, alcohol consumption	Behavioral Risk Factor Surveillance Survey
Economic Resources	Income, wealth, health insurance coverage, educational level, employment status	Behavioral Risk Factor Surveillance Survey Health and Retirement Survey National Health Survey
Acculturation	Use of Spanish and English language	Marin Scale of Acculturation

Statistical Analysis of mental health data

Analysis will be conducted using SAS version 8.2 (The SAS system for Windows.Windows version 5.1, 2001). SAS. The preliminary analysis of the cohort's baseline status will produce a profile of the study population with respect to demographic and baseline outcome characteristics. It will consist of descriptive statistics such as frequencies, means, cross-tabulations, and graphical representations of the data for demographic characteristics, depression and its other mediating variables. The primary analyses will examine the prevalence of depression and chi-square tests of subgroup patterns. Logistic regression will be done to test for interactions between mediating variables and the main outcomes. Exploratory analysis will be performed after all primary and secondary study questions have been addressed. The purpose is to identify relationships and patterns in the data that may not have been anticipated at the onset of the study. All results of this analysis will be labeled as exploratory in nature. Results will be used to suggest hypotheses for future studies.

Spatial Analysis of risks, lifestyle, cultural and other factors associated with diabetes and complications: The integration of relevant datasets into a specialized spatial database, in which the associations between the attribute information and their geographic references of the data are managed and maintained are key to the use of spatial analysis. We will geocode the data for this study employing address matching and/or GPS georeferencing techniques. Address matching allows data records with street addresses to be automatically placed within a map coordinate system as points based on a street network. The residential address of the cohort samples will be used to merge them with the remainder of the data within the GIS. Access to the Defense Department's global positioning system (GPS) provides additional capability to incorporate a sample into the spatial database, when street addresses can not be located within the road network. The location information of the samples acquired will then be linked to their attribute information, such as age, race, diagnostic class, and genome type.

The utilization of ESDA often involves the measurement of and testing for spatial autocorrelation, an embodiment of Tobler's First Law of Geography, which basically states that "everything is related to everything else, but near things are more related than distant things." Many global indices have been proposed to characterize spatial autocorrelation and to quantify the overall spatial pattern of a study area as being random, clustered, or dispersed. Popular global indices include Moran's (1948) I, Geary's (1954) C, and Getis-Ord's (1992) G statistics(Getis A, 1999; Gatrell A and Senior M, 1999; Cockings et al., 2004). The local level implementations of these indices were also developed to help pinpoint where the strongest clustering and dispersion are. These statistics are often based either on the covariance (Moran's I), or on the subtraction (Geary's c), or on the addition (Getis and Ord G) of a continuous variable between spatially associated samples. The limitation of these autocorrelation statistics is, however, that they can not be directly applied to binomial or categorical variables (such as the ones representing being diagnosed as diabetes or not and the genomic types in this project). Traditional spatial statistics will be ideal for continuous measures like the Body Mass Index and blood pressure. Another issue is that these measurements are not able to take the underlying population variation into consideration if quantifying spatial clusters or hotspots from point pattern data (such as the cohort sample used in this study). This is apparently a drawback because 50 cases of diabetes may be a small incidence rate within a large city, for example, but an epidemic if occurred within a small town.

We propose to use spatial kernel density estimation to approach these problems, wherein the distribution of discrete sample points representing, for example, diabetes incidences will be transformed into a continuous surface of disease risk. This avoids the widely criticized spatial aggregation of point pattern data into an arbitrary areal unit of artificial designation, and at same time preserves the individual nature of the point pattern approach(Sabel et al., 2000). Generally, the underlying population at risk is often not uniformly distributed and this necessitates the generation of a kernel density surface for a control sample. The ratio of the case to control kernel density surfaces can then be derived to estimate the relative risk at different geographic locations, with `peaks' and `troughs' in the data identifying high and low incidence "clusters" respectively. A Monte Carlo simulation with repeated random surfaces of case-control pairs can be used to highlight regions of significantly high or low relative risk.

ESDA is often performed with an intention to develop hypotheses to guide future confirmative research. Such hypotheses can be formulated by identifying spatial patterns of disease, and their precursors, and examining the proximities of the detected clusters to the distributions of potential risk factors in space. A causal study to confirm the association between spatial pattern of disease and potential socioeconomic and environmental factors is often conducted based on a regression type analysis. Traditional statistical models are based on an assumption of independent observations. However, the existence of spatial autocorrelation due to a spatial spillover effects or distance decline effects may violate the assumption. If the residuals of an ordinary least squares regression mode are tested to be significantly autocorrelated, then that regression is likely to be not properly specified. One or more new variables should be incorporated into the regression to account for spatial dependency (Getis 1999). This effort has led to the formulation of several spatial regression tools, such as spatial lag models and spatial error models. Unfortunately, as with the problems we encountered in the ESDA stage, these models are most designed to be applied to areal data with continuous variables. Logistic type of spatial regression models that can be used to deal with binomial and categorical variables have not been formed yet.

In this project, we propose to investigate the use of Geographically Weighted Regression (GWR) model to analyze the potential risk factors for diabetes and its complication. GWR is a new analytic tool that allows local spatial relationships to be measured and mapped using spatially varying regression parameters (Fortheringham AS et al., 2002). It does not only take the spatial dependence into consideration, but accounts for the spatial variance of this dependence, thus simultaneously addressing the problems of *spatial heterogeneity* and *spatial autocorrelation.* A binomial extension of the GWR model is also available to support local level logistic regression and will be employed to explore the spatial interaction of genomics with behavioral, socioeconomic and environmental factors in order to reach the most robust estimation on their impacts to T2DM and its complications.

STATISTICAL ANALYSIS

Data Management: All data are entered now using desktops and surface pros using screens and drop off menus for fast input and to prevent errors. Routines include range checking and field reconciliation at the time of entry. Reconciled files are then transferred to the Master data directory on the workstation. Copies are transferred to a writeable CD ROM to permit archival storage (one copy is made and stored off-site also). The genotype management routines were developed by us and include entry and reconciliation and permit automated exporting for linkage analysis on a variety of platforms. Increasingly, we use the ABI sequencers for routine genotyping. .The representative sample collected in Aim 1 will allow standard estimates of prevalence and standard errors of estimates. We will also employ logistic regression to test the impact of age, gender, body mass and distribution and related factors on predicting diabetes status. In addition to largely standard analyses, these samples will be used for testing associations with detected genetic variation.. Central to the cohort analysis is testing for association between type 2 diabetes and an identified variant. These analyses will include traditional contingency table analysis and transmission disequilibrium testing. For the contingency table analyses, a two by two table of allele counts in diabetics and non-diabetics will be constructed and tested using a standard χ^2 or Fisher=s Exact Test (Rosner 1995). We illustrate power based on SNP 43 at *calpain 10* whose frequency is estimated at 0.77. Assuming an allele frequency estimate of 5% in the random population, this sample will have 80% power to detect a frequency difference of 4.6% or more. If the frequency of an allele in the random sample is 40%, then we will have 80% power to detect a difference of 8.7% (i.e., 48.7% versus 40%). We will also employ transmission disequilibrium testing (Speilman et al., 1993). Because our hypothesis centers on association, we will choose 1 pair of affected siblings at random from each sibship (Monks et al., 1998).

Table 4. Time table for recruitment

2005/6	
Month	
May	Recruitment and Training
June	100
July	100
August	100
September	100
October	100
November	100
December	100
January	125
February	125

March	125
April	125
Totals	1200

E: Human Subjects Research

Risks to Human Subjects

This study will rely on material previously collected and now being collected in our ongoing studies of chronic disease among Mexican Americans in Starr County, Texas. To these resources we will add 1200 Mexican American individuals from Brownsville aged 35 to 54 representative of the population. All individuals will be Mexican American. Children are not included; the study is designed for equal representation of men and women though in practice our experience shows that differential participation leads to more women than men (estimated to be in the range of 60% women).

Sources of Research Material

This study will use the data files from the investigations described above. All previous data have been computerized. No identifiers are included in the files used for analysis. DNA has been extracted from buffy coats on all individuals and will be the raw source of genetic material. These samples are identified by ID number only. The collection of new data will involve protocols and safeguards currently in force for our existing approved and funded studies.

Potential Risks

For existing data resources, there are no substantive risks to participants other than confidentiality. The newly recruited (individuals 950 supported by this project) will have minor risks from the phlebotomy involved in blood collecting.

2. Adequacy of the Protection against Risk

Recruitment and Consent

All data collected and protocols (as well as ongoing protocols permitting the use and analysis of the data and samples) have been reviewed and approved by the institutional committee for the protection of human subjects (see attached consent forms and CPHS approval letter). Written consent was obtained from all previous participants for the examinations and use of biological materials obtained and will be obtained from all new participants. Written consent to disclose Protected Health Information (PHI) to the Rio Grande Valley Health Information Exchange (RGV HIE), the Department of Health and Human Services Texas Cancer Registry (TCR), and other specified health care entities will be provided by subjects during the consent process. This will be obtained from all participants. Participants will have the option to opt-out of the authorization by signing/revoking an opt-out form.

Compensation

A gift card from a local vendor in the amount of $50 will be given to each participant for their participation in our study after they have completed all tests.

Procedures for Protecting Against Risks

Risk to prior participants in this study is minimal. To protect confidentiality, all data is coded and identifiers removed. Original data forms are in locking file cabinets in rooms with limited access. The PI does have access to identifiers. Thus, there is permission for re-contacting individuals. We do anticipate this in the present study.

For new participants, we will employ the same safeguards and standards of confidentiality as before. Except as necessary for the conduct of the study, all identifiers will be eliminated from data files except for key files necessary for the conduct of the study and follow-up. These files will have extremely limited access. To minimize the risks of phlebotomy, only trained personnel will be used

Why Risks to Subjects are Reasonable

As described, the risk to subjects is minor. The benefits, however, are significant. The populations of South Texas border cities bear disproportionate burdens of diabetes and are among the most impoverished in the

United States. It is a young population that is only going to see this burden increase over the next few decades. For the first time, it appears that we have the ability to understand the molecular underpinnings of this disproportionate risk. Already, these studies are leading to novel metabolic pathways. This will lead to understanding the metabolic processes underlying type 2 diabetes and will lead to avenues for intervention that may prevent or slow the onslaught of diabetes and its complications.

Total enrollment

TABLE 5. TARGETED/PLANNED ENROLLMENT: 860 (Number of Subjects)			
Ethnic Category	Sex/Gender		
	Females	Males	Total
Hispanic or Latino	600	600	1200
Not Hispanic or Latino	25	25	50
Ethnic Category Total of All Subjects*	625	625	625
Racial Categories			
American Indian/Alaska Native	0	0	0
Asian	0	0	0
Native Hawaiian or Other Pacific Islander	0	0	0
Black or African American	13	13	26
White	444	444	888
Racial Categories: Total of All Subjects *	457	457	914

INCLUSION OF WOMEN, CHILDREN AND MINORITIES

The population to be studied is Mexican Americans, and all subjects will therefore belong to the Hispanic minority. In the survey, men and women will be selected randomly, and therefore have equal chances of inclusion and exclusion.

USE OF THE Clinical Research Center/Unit. The initial identification of subjects and interviews will be held in the community. Participants will then be invited either to present themselves at the CRC/CRU for a fasting blood sugar, and a full examination with laboratory and other measurements, or if they prefer, we can visit in the morning for these procedures. (See the description of research methods for details). The central location of the CRUs in the different research sites is convenient to most people's homes and workplaces, so more active participants may wish to visit the CRU for the initial and follow-up examinations. We will encourage all participants to visit the clinic because we can use the central location to reinforce the sense of community participation, and take the opportunity to expose our subjects to information on diabetes and other diseases affecting this population. We can also better ensure standardization of the examination at the clinic.

Any individual with a fasting glucose level of 120 mg/dl or greater used to be scheduled for a second fasting sample, but now with the A1c results it is easier to have a better idea of the glucose levels.

Data and Safety monitoring plan:

This is a non-interventional study of low risk therefore a formal data and safety monitoring board will not be needed. Dr. Joseph McCormick will serve as the medical monitor and record adverse events including those that are anticipated and unanticipated. He will determine the relationship if any with the study. These will be reported to the IRB and funding agency as appropriate. The Research subject advocate will monitor the frequency and severity of adverse events and will be available to research subjects for counseling and questions.

F: Vertebrate Animals

None used

G: Literature cited

Anderson,R.J., Freedland,K.E., Clouse,R.E., and Lustman,P.J. (2001). The prevalence of comorbid depression in adults with diabetes: a meta-analysis. Diabetes Care *24*, 1069-1078.

Black,S.A., Markides,K.S., and Ray,L.A. (2003). Depression predicts increased incidence of adverse health outcomes in older Mexican Americans with type 2 diabetes. Diabetes Care *26*, 2822-2828.

Bosque-Plata,L., Aguilar-Salinas,C.A., Tusie-Luna,M.T., Rami, rez-Jimenez,S., Rodri, guez-Torres,M., Auron-Gomez,M., Rami, rez,E., Velasco-Perez,M.L., Rami, rez-Silva,A., Gomez-Perez,F., Hanis,C.L., Tsuchiya,T., Yoshiuchi,I., Cox,N.J., and Bell,G.I. (2004). Association of the calpain-10 gene with type 2 diabetes mellitus in a Mexican population. Mol. Genet. Metab *81*, 122-126.

Centers for Disease Control (1998). Self-reported frequent mental distress among adults - United States, 1993 - 1996. MMWR *47*, 326-331.

Clarke,K.C., McLafferty,S.L., and Tempalski,B.J. (1996). On epidemiology and geographic information systems: a review and discussion of future directions. Emerg. Infect. Dis. *2*, 85-92.

Cockings,S., Dunn,C.E., Bhopal,R.S., and Walker,D.R. (2004). Users' perspectives on epidemiological, GIS and point pattern approaches to analysing environment and health data. Health Place. *10*, 169-182.

DeFronzo RA (1997a). Genetics of Diabetes - Part I.

DeFronzo RA (1997b). Genetics of Diabetes - Part II.

Dinan,T., Peveler,R., and Holt,R. (2004). Understanding schizophrenia and diabetes. Hosp. Med. *65*, 485-488.

Dunlop,D.D., Song,J., Lyons,J.S., Manheim,L.M., and Chang,R.W. (2003). Racial/ethnic differences in rates of depression among preretirement adults. Am. J. Public Health *93*, 1945-1952.

Elbein,S.C., Hoffman,M.D., Teng,K., Leppert,M.F., and Hasstedt,S.J. (1999). A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. Diabetes *48*, 1175-1182.

Federal Reserve Bank of Dallas. The Border Economy. 1-39. 10-6-2001. Federal Reserve Bank of Dallas. Ref Type: Report

Fisher,L., Chesla,C.A., Mullan,J.T., Skaff,M.M., and Kanter,R.A. (2001). Contributors to depression in Latino and European-American patients with type 2 diabetes. Diabetes Care *24*, 1751-1757.

Flegal,K.M., Carroll,M.D., Kuczmarski,R.J., and Johnson,C.L. (1998). Overweight and obesity in the United States: prevalence and trends, 1960-1994. Int. J. Obes. Relat Metab Disord. *22*, 39-47.

Flegal,K.M., Carroll,M.D., Ogden,C.L., and Johnson,C.L. (2002). Prevalence and trends in obesity among US adults, 1999-2000. JAMA *288*, 1723-1727.

Fortheringham AS, Brunsdon C, and Charlton M (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. (Chichester, England: John Wiley & Sons).

Gatrell A and Senior M (1999). Health and Health Care Applications. In Geographic Information Systems, Longley PA, Goodchild MF, Magiore DJ, and Rhind DW, eds. (New York: John Wiley), pp. 925-938.

Getis A (1999). Spatial Statistics. In Geographic Information Systems; Principles, Techniques, Management and Applications, Longley PA, Goodchild DJ, Maguire DJ, and Rhind DW, eds. (New York: John Wiley), pp. 239-251.

Grigsby,A.B., Anderson,R.J., Freedland,K.E., Clouse,R.E., and Lustman,P.J. (2002). Prevalence of anxiety in adults with diabetes: a systematic review. J. Psychosom. Res. *53*, 1053-1060.

Hanis CL (1996). Genetics of NIDDM among Mexican Americans. In Genetic Approaches to Noncommunicable Diseases, Berg K, Boulyjenkov V, and Christen Y, eds. (Berlin: Springer-Verlag), pp. 65-77.

Hanis,C.L., Boerwinkle,E., Chakraborty,R., Ellsworth,D.L., Concannon,P., Stirling,B., Morrison,V.A., Wapelhorst,B., Spielman,R.S., Gogolin-Ewens,K.J., Shepard,J.M., Williams,S.R., Risch,N., Hinds,D., Iwasaki,N., Ogata,M., Omori,Y., Petzold,C., Rietzch,H., Schroder,H.E., Schulze,J., Cox,N.J., Menzel,S., Boriraj,V.V., Chen,X., and . (1996). A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. Nat. Genet. *13*, 161-166.

Hanis,C.L., Ferrell,R.E., Barton,S.A., Aguilar,L., Garza-Ibarra,A., Tulloch,B.R., Garcia,C.A., and Schull,W.J. (1983). Diabetes among Mexican Americans in Starr County, Texas. Am. J. Epidemiol. *118*, 659-672.

Hanis,C.L., Ferrell,R.E., and Schull,W.J. (1985). Hypertension and sources of blood pressure variability among Mexican-Americans in Starr County, Texas. Int. J. Epidemiol. *14*, 231-238.

Hanis,C.L., Hewett-Emmett,D., Bertin,T.K., and Schull,W.J. (1991). Origins of U.S. Hispanics. Implications for diabetes. Diabetes Care *14*, 618-627.

Hanson,R.L., Ehm,M.G., Pettitt,D.J., Prochazka,M., Thompson,D.B., Timberlake,D., Foroud,T., Kobes,S., Baier,L., Burns,D.K., Almasy,L., Blangero,J., Garvey,W.T., Bennett,P.H., and Knowler,W.C. (1998). An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. Am. J. Hum. Genet. *63*, 1130-1138.

Harris,M.I., Flegal,K.M., Cowie,C.C., Eberhardt,M.S., Goldstein,D.E., Little,R.R., Wiedmeyer,H.M., and Byrd-Holt,D.D. (1998). Prevalence of diabetes, impaired fasting glucose, and impaired glucose tolerance in U.S. adults. The Third National Health and Nutrition Examination Survey, 1988-1994. Diabetes Care *21*, 518-524.

Hegele,R.A., Sun,F., Harris,S.B., Anderson,C., Hanley,A.J., and Zinman,B. (1999). Genome-wide scanning for type 2 diabetes susceptibility in Canadian Oji-Cree, using 190 microsatellite markers. J. Hum. Genet. *44*, 10-14.

Horikawa,Y., Oda,N., Cox,N.J., Li,X., Orho-Melander,M., Hara,M., Hinokio,Y., Lindner,T.H., Mashima,H., Schwarz,P.E., Bosque-Plata,L., Horikawa,Y., Oda,Y., Yoshiuchi,I., Colilla,S., Polonsky,K.S., Wei,S., Concannon,P., Iwasaki,N., Schulze,J., Baier,L.J., Bogardus,C., Groop,L., Boerwinkle,E., Hanis,C.L., and Bell,G.I. (2000). Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat. Genet. *26*, 163-175.

Kenny SJ, Aubert RE, and Geiss LS (1995). Prevalence and incidence of non-insulin-dependent diabetes mellitus. In Diabetes in America, Harris MI, Cowie CC, Stern MP, Boydo EJ, Reiber GE, and Bennett PH, eds. (Washington, DC: NIH), pp. 47-68.

King,H., Aubert,R.E., and Herman,W.H. (1998). Global burden of diabetes, 1995-2025: prevalence, numerical estimates, and projections. Diabetes Care *21*, 1414-1431.

Lebovitz,H.E. (2003). The relationship of obesity to the metabolic syndrome. Int. J. Clin. Pract. Suppl 18-27.

Mahtani,M.M., Widen,E., Lehto,M., Thomas,J., McCarthy,M., Brayer,J., Bryant,B., Chan,G., Daly,M., Forsblom,C., Kanninen,T., Kirby,A., Kruglyak,L., Munnelly,K., Parkkonen,M., Reeve-Daly,M.P., Weaver,A., Brettin,T., Duyk,G., Lander,E.S., and Groop,L.C. (1996). Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families. Nat. Genet. *14*, 90-94.

McVeigh KH, Mostahari F, and Thorpe LE (2004). Serious psychological distress among persons with diabetes. MMWR *53*, 1089-1092.

Mokdad,A.H., Bowman,B.A., Ford,E.S., Vinicor,F., Marks,J.S., and Koplan,J.P. (2001). The continuing epidemics of obesity and diabetes in the United States. JAMA *286*, 1195-1200.

Mokdad,A.H., Ford,E.S., Bowman,B.A., Nelson,D.E., Engelgau,M.M., Vinicor,F., and Marks,J.S. (2000a). Diabetes trends in the U.S.: 1990-1998. Diabetes Care *23*, 1278-1283.

Mokdad,A.H., Marks,J.S., Stroup,D.F., and Gerberding J.L. (2004). Actual Causes of Death in the United States, 2000. JAMA *291*, 1238-1245.

Mokdad,A.H., Serdula,M.K., Dietz,W.H., Bowman,B.A., Marks,J.S., and Koplan,J.P. (2000b). The continuing epidemic of obesity in the United States. JAMA *284*, 1650-1651.

Motl,R.W., Birnbaum,A.S., Kubik,M.Y., and Dishman,R.K. (2004). Naturally occurring changes in physical activity are inversely related to depressive symptoms during early adolescence. Psychosom. Med. *66*, 336-342.

National Institute of Mental Health, National Institutes of Health. Department of Health and Human Services FY 2005 budget report. 2005.
Ref Type: Report

NDDG (1997). Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care *20*, 1183-1197.

Office of Border Health and Texas A&M Public Policy Research Institute, Texas Department of Health. Survey of Health and Environmental Conditions in Texas Border Countries and *Colonias*. **79-10828**, 1-102. 1. **Austin, Texas**, **Texas Department of Health**.
Ref Type: Report

Reaven,G.M. (1988). Banting lecture 1988. Role of insulin resistance in human disease. Diabetes *37*, 1595-1607.

Reaven,G.M. (1993). Role of insulin resistance in human disease (syndrome X): an expanded definition. Annu. Rev. Med. *44:121-31.*, 121-131.

Rubin,R.J., Altman,W.M., and Mendelson,D.N. (1994). Health care expenditures for people with diabetes mellitus, 1992. J. Clin. Endocrinol. Metab *78*, 809A-809F.

Sabel,C.E., Gatrell,A.C., Loytonen,M., Maasilta,P., and Jokelainen,M. (2000). Modelling exposure opportunities: estimating relative risk for motor neurone disease in Finland. Soc. Sci. Med. *50*, 1121-1137.

Swenson,C.J., Baxter,J., Shetterly,S.M., Scarbro,S.L., and Hamman,R.F. (2000). Depressive symptoms in Hispanic and non-Hispanic White rural elderly: the San Luis Valley Health and Aging Study. Am. J. Epidemiol. *152*, 1048-1055.

The SAS system for Windows.Windows version 5.1. SAS®. [8.02]. 2001. Cary, North Carolina, Sas Institute Inc.