2019

Assessing Diabetes Risk Among Hispanic Populations: Development And Validation Of A Risk Score Using Readily Available Information

Lucia Juarez
University of Alabama at Birmingham

ASSESSING DIABETES RISK AMONG HISPANIC POPULATIONS: DEVELOPMENT AND VALIDATION OF A RISK SCORE USING READILY AVAILABLE INFORMATION

by

LUCIA D. JUAREZ

ANDREA CHERRINGTON AND ROBIN LANZI COMMITTEE CO-CHAIRS
JEFFREY GONZALEZ
GREGORY PAVELA
DOROTHY PEKMEZI

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2019

ASSESSING DIABETES RISK AMONG HISPANIC POPULATIONS: DEVELOPMENT AND VALIDATION OF A RISK SCORE USING READILY AVAILABLE INFORMATION

LUCIA D. JUAREZ

HEALTH EDUCATION AND HEALTH PROMOTION

ABSTRACT

Background: Diabetes prevalence is increasing worldwide affecting vulnerable populations, in particular Hispanics living in the United States (U.S.). The burden of diabetes and its complications can be reduced by identifying those at high risk and routing them to treatment and intensive behavioral interventions.

Aim: The aim of this study is to develop and validate a simple score to identify undiagnosed diabetes among Hispanics living in the U.S.

Methods: A representative sample of Hispanics living in the U.S. was used to develop and validate the score. Logistic regression was used to identify significant risk factors and model coefficients were used to assign points for the score. The risk score was validated using a split sample and comparing its performance to similar risk scores.

Results: Being of Mexican descent, male gender, older age, lower education, being born in the U.S., family history of diabetes, being overweight or obese, having had gestational diabetes and not complying with physical activity recommendations were risk factors included in the score. The model had an area under the ROC curve (AUC) of 0.76 in the development sample and of 0.77 in the validation sample. Other scores had AUCs between 0.69 and 0.75. The SOL risk score also performed well identifying prediabetes, with an AUC of 0.68. The risk score also demonstrated reasonable performance detecting dysglycemia, AUC of 77%, sensitivity of 74% and specificity of 70%.

Conclusion: This study presents the first diabetes risk score derived for Hispanics living in the U.S. It provides a simple and inexpensive tool to identify individuals at high risk of diabetes and prediabetes. Further work is still needed to validate this score in other Hispanic populations. Tools developed specifically for Hispanic populations may be more effective among Hispanics than those developed for the general population.

Keywords: Type 2 Diabetes Mellitus, Prediabetes, Dysglycemia, Risk Scores, Prediction Models, Hispanic, Latino

ACKNOWLEDGMENTS

I offer my sincere gratitude to each member of this dissertation committee for their support and guidance. I also thank Dr. Andres Azuero for his help and expertise in statistical methods and for being a constant source of support; Dr. April Carson for her valuable comments and insights and Ms. April Agne for her help reviewing thousands of references and for all her practical expertise. Lastly, I thank my family – Andrzej, Anna, and Izabella for being a constant source of encouragement, support and good cheer.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table *Page*

DIABETES RISK SCORES FOR HISPANICS LIVING IN THE UNITED STATES: A SYSTEMATIC REVIEW

A RISK SCORE FOR UNDIAGNOSED DIABETES FOR HISPANICS LIVING IN THE U.S.: THE HISPANIC COMMUNITY HEALTH STUDY/STUDY OF LATINOS (HCHS/SOL)

EVALUATING THE PERFORMANCE OF THE HCHS/SOL DIABETES RISK SCORE FOR IDENTIYING DYSGLYCEMIA AMONG HISPANICS IN THE U.S.: THE HISPANIC COMMUNITY HEALTH STUDY/STUDY OF LATINOS (HCHS/SOL)

LIST OF FIGURES

Figure *Page*

DIABETES RISK SCORES FOR HISPANICS LIVING IN THE UNITED STATES: A SYSTEMATIC REVIEW

A RISK SCORE FOR UNDIAGNOSED DIABETES FOR HISPANICS LIVING IN THE U.S.: THE HISPANIC COMMUNITY HEALTH STUDY/STUDY OF LATINOS (HCHS/SOL)

EVALUATING THE PERFORMANCE OF THE HCHS/SOL DIABETES RISK SCORE FOR IDENTIYING DYSGLYCEMIA AMONG HISPANICS IN THE U.S.: THE HISPANIC COMMUNITY HEALTH STUDY/STUDY OF LATINOS (HCHS/SOL)

INTRODUCTION

Diabetes mellitus or Type 2 diabetes is one of the most prevalent chronic diseases around the world. In 2015, 30.3 million Americans had diabetes and another 84.1 had prediabetes, which if not treated can develop into diabetes within five years (“Centers for Disease Control and Prevention (CDC)”, 2017). Diabetes can lead to vascular complications and death and it is not only the seventh leading cause of death in the United States (U.S.) but the second most expensive chronic disease according to the CDC. Increased screening, timely lifestyle and behavioral interventions and medical treatment can help manage hyperglycemia ("The Diabetes Prevention Program (DPP): description of lifestyle intervention," 2002; Gerstein et al., 2006; Knowler et al., 2002; Li et al., 2008). However, despite advances in prevention, control and treatment of diabetes over the past two decades, these benefits have not fully reached Hispanics or other minority populations in the U.S. (Aviles-Santa et al., 2016).

Hispanics represent 17.6% of the U.S. mainland population and by 2060, it is projected that nearly one in three individuals in the U.S. (29%) will be of Hispanic descent (Colby & Ortman, 2015). A comprehensive study that used data collected between 1988 to 2012 from the National Health and Nutrition Examination Surveys (NHANES) estimated that the prevalence of diabetes among Hispanics in the U.S. was 22.6%, two times the rate among non-Hispanic Whites (Menke, Casagrande, Geiss, & Cowie, 2015). Furthermore, among those individuals who had diabetes, 49% of Hispanics and 33.5% of

non-Hispanic Whites were undiagnosed, increasing their risk of complications (Menke et al., 2015).

Understanding diabetes risk factors among Hispanics is challenging on many levels. Hispanics are a heterogeneous group of peoples with varied heritages which is often not reflected in data collected in the U.S. This can lead to inconsistent use of the terms Hispanic or Latino when referring to ancestry group, place of birth, or immigrant generation (Krogstad & Lopez, 2014). The term Hispanic has also gone through several modifications in the U.S. Census since 1970 (Moy, 1977). Government and scientific reports in the U.S. tend to combine data from Hispanics of different backgrounds. Recent data have showed that diabetes prevalence varies by Hispanic ancestry group and that these differences may be attributed to sociodemographic and cultural characteristics within each group (Aviles-Santa et al., 2016; Katherine M Flegal et al., 1991; Schneiderman et al., 2014). In this study, the term Hispanic refers to people who identify their origins to Latin American countries and who are living in the U.S. (Krogstad & Lopez, 2014).

Since the 1960s, data have been collected to study chronic diseases in Hispanics populations in the U.S. (Cruz-Vidal, Costas, Garcia-Palmieri, Sorlie, & Hertzmark, 1979; Dawber et al., 1959). Examples of epidemiological studies that have collected data from Hispanics include the Laredo Project (Gaskill, Allen, Garza, Gonzales, & Waldrop, 1981), the San Antonio Heart Study (SAHS) (Gunby, 1980), the Hispanic Health and Nutrition Examination Survey (HHANES) (K. M. Flegal et al., 1991), the Behavior Risk Factor Surveillance Survey (BRFSS) (Stein, Lederman, & Shea, 1993), and the National

Health and Nutrition Examination Survey (NHANES) (Johnson, Dohrmann, Burt, & Mohadjer, 2014). However, comparisons across these studies are difficult due to the cross-sectional design of most studies and methodological variations in sample selection, diagnostic criteria and interpretation of the term Hispanic. The Hispanic Community Health Study/Study of Latinos (HCHS/SOL) studies chronic diseases among Hispanics living in four metropolitan areas in the U.S. It was established in 2006. The study identifies participants' ancestry from Central America, Cuba, Dominical Republic, Mexico, Puerto Rico, and South America (Sorlie et al., 2010).

Results from the HCHS/SOL baseline examination (2008-2011) have shown that diabetes prevalence varies among Hispanics of different heritage backgrounds. South Americans had the lowest prevalence of diabetes at 10.2% while Dominicans (18.1%), Mexicans (18.3%) and Puerto Ricans (18.3%) had the highest ones (Schneiderman et al., 2014). Age, gender, higher body mass index (BMI) are risk factors associated with higher prevalence of diabetes as well as having lived longer in the U.S. Combination of factors such as gender and education are also associated with diabetes, for example, diabetes prevalence was lower among women with education of high school and above than among women with lower education. Similarly, diabetes prevalence for men with an education of at least high school is lower than for men with education less than high school (Schneiderman et al., 2014). Other analyses show that among participants with diabetes, nearly four in ten meet one or more criteria for undiagnosed diabetes based on the American Diabetes Association (ADA) (Aviles-Santa et al., 2016). Furthermore, among those with diabetes 41.3% are unaware of their condition (Schneiderman et al., 2014).

In the U.S., universal screening for diabetes is recommended for high-risk adults. Different organizations publish screening recommendations based on risk factors associated with diabetes (American Diabetes Association, 2018; Group, 2010; Handelsman et al., 2011; Siu, 2015; Vijan, 2010). Although the risk factors listed in most recommendations are similar, including individual characteristics such as age, BMI, gender, being Hispanic or from any other minority group, having a family history of diabetes or gestational diabetes, hypertension and high cholesterol, the thresholds, ranges and combination of factors vary. Having multiple criteria can be confusing to public health providers. There is not enough information to evaluate the implementation of national screening recommendations in the U.S. (Abid, Ahmad, & Waheed, 2016). However, being a member of a minority population, having a low socioeconomic status and lacking health insurance are associated with a lower prevalence of laboratory testing to diagnose diabetes (Casagrande, Cowie, & Genuth, 2014). The lack of a confirmatory diagnosis delays treatment and increases the likelihood of complications in these populations.

The high prevalence of undiagnosed diabetes among Hispanics in the U.S. calls for effective screening tools that identify those more in need of treatment. Diabetes risk scores can be used as prediction models to help classify individuals according to their risk, along a continuous spectrum. Risk scores are built using risk factors associated with the disease in the populations in which they are developed (Royston, Moons, Altman, & Vergouwe, 2009). There are a number of risk scores developed and validated to identify undiagnosed diabetes in populations around the world and the U.S.(Brown, Critchley, Bogowicz, Mayige, & Unwin, 2012; Buijsse, Simmons, Griffin, & Schulze, 2011; Schwarz, Li, Lindstrom, & Tuomilehto, 2009; Thoopputra, Newby, Schneider, & Li,

2012; Witte, Shipley, Marmot, & Brunner, 2010). However, in order for a risk score to be reflective of and effective in a population, it needs to be validated in the population it will be applied (Royston et al., 2009). This dissertation was designed to address these issues.

The overall objective of this three paper format dissertation is to contribute to the understanding of diabetes prevention among Hispanics by (1) identifying screening tools that have been validated among Hispanic populations; (2) examining risk factors associated with diabetes among Hispanics living in the U.S., and (3) proposing a risk score that identifies undiagnosed diabetes using data from the HCHS/SOL study.

Specific Aims

As such, the specific aims for this dissertation are as follows:

Aim 1: Review the literature to investigate if there are risk scores to identify undiagnosed diabetes among Hispanic populations. The review will identify of risk scores for undiagnosed diabetes and will provide information on data availability for risk factors included in the risk scores and their performance to assess their applicability among Hispanics in the U.S.

Aim 2: Based on the results from the systematic review, develop and validate a risk score for undiagnosed diabetes among Hispanics in the U.S. We use data from the baseline examination from HCHS/SOL. We examine the association of risk factors associated with diabetes identified in the systematic review and explore the association of other factors relevant to Hispanics in the U.S. The performance of the newly developed risk score is evaluated and compared to that of other risk scores identified in the systematic review and to the ADA risk score, commonly applied to the general U.S. population.

Aim 3: Further assess the performance of the SOL risk score to identify diabetes and prediabetes and compare it to other screening criteria. Although we have seen significant and sustained improvements in glycemic control in the past two decades (Hoerger, Segel, Gregg, & Saaddine, 2008; Sperl-Hillen & O'Connor, 2005), it is clear that more attention needs to be geared towards early identification of individuals at risk and towards interventions that prevent and control diabetes. Whether national recommendations ever reach a common and practical agreement as to what constitutes a reasonable approach for screening for diabetes, practitioners will benefit from knowing how effective different criteria are identifying individuals who are at high risk of diabetes. The use of validated risk scores developed for specific populations such as Hispanics can help identify those who are at highest risk. Having a simple questionnaire that includes risk factors unique to Hispanics may facilitate the introduction to educational interventions.

The first paper in this dissertation addresses Aim 1. A systematic review of the literature that included over 16,000 references retrieved from four searching engines: PubMed, CINAHL, EMBASE and Cochrane Library including references through December 2016, without language restriction. The study helped identify risk factors associated with diabetes among Hispanics and highlighted the lack of research on diabetes prediction models developed for Hispanics, supporting the need to develop a risk score using recent data from Hispanics residing in the U.S. The second paper in this dissertation presents the development of a risk score for undiagnosed diabetes based on readily available information from the HCHS/SOL baseline cohort. The development of the SOL risk score considered risk factors not included in other scores such as education, ethnic background and being born in the U.S. The final risk score’s performance was superior to that

of the scores identified in the systematic review and marginally better than the ADA risk score's performance. The third paper further evaluates the performance of the SOL risk score to identify Hispanics with prediabetes or diabetes (referred to as dysglycemia). In this paper, the performance of the risk score identifying dysglycemia is compared to the performance of the ADA risk score. The performance of the risk scores was comparable when adjusting the optimal cutoff point of the ADA risk score recommended for the general population.

Successful implementation of diabetes prevention programs requires a comprehensive framework that includes sociodemographic, cultural and environmental factors. A diabetes risk score can help identify and stratify individuals by their risk of having the disease allowing the possibility to uncover individual characteristics and values that may impact the development of the disease or may pinpoint possible barriers to the successful completion of an intervention program. For example, intervention approaches for those whose risk score is high due to age, weight status and lack of physical activity will differ from those who are younger but may have an earlier onset of obesity.

This research represents the first attempt to develop a diabetes risk score based on readily available information from Hispanics living in the U.S. The simplicity of the risk score provides a viable alternative to screen for undiagnosed diabetes and prediabetes that can be easily implemented in clinics and communities. It is also the first time that the ADA risk score is validated on Hispanics populations in the U.S. The growing prevalence of diabetes and prediabetes among Hispanics indicates that this population has not fully benefitted from prevention intervention programs. The Diabetes Prevention Program (DPP) has shown that a behavioral intervention that targets loss of body weight of at least

7% and increased physical activity to at least 150 minutes per week can effectively reduce diabetes incidence (Knowler et al., 2002). The participation of Hispanics in DPP was 16% and the effectiveness of the treatment did not vary by race/ethnicity. Other studies modeled on the DPP that were conducted in diverse community settings (Ali, Echouffo-Tcheugui, & Williamson, 2012; Ockene et al., 2012) and pharmacological interventions to prevent or delay diabetes have achieved successful results among Hispanics (Boyko et al., 2010; Buchanan et al., 2002; O'Brien et al., 2017; Xiang et al., 2006). The question that remains is how to make these interventions more accessible, cost-effective and culturally tailored.

DIABETES RISK SCORES FOR HISPANICS LIVING IN THE UNITED STATES: A SYSTEMATIC REVIEW

by

LUCIA D. JUAREZ; JEFFREY S. GONZALEZ; APRIL A. AGNE; ANDRZEJ KULCZYCKI; GREGORY PAVELA; APRIL P. CARSON; JOHN P. SHELLEY; ANDREA L. CHERRINGTON

Diabetes Research and Clinical Practice
Juarez, L. D., Gonzalez, J. S., Agne, A. A., Kulczycki, A., Pavela, G., Carson, A. P., . . . Cherrington, A. L. (2018). Diabetes risk scores for Hispanics living in the United States: A systematic review. *Diabetes Research and Clinical Practice, 142*, 120-129. doi:https://doi.org/10.1016/j.diabres.2018.05.009

ABSTRACT

Aim: Despite the proliferation of risk scores, few have been validated in Hispanic populations. Undiagnosed diabetes is more prevalent among racial/ethnic minorities in the United States (U.S.). The aim of this study is to systematically review published studies that developed risk scores to identify undiagnosed Type 2 Diabetes Mellitus based on self-reported information that were validated for Hispanics in the U.S.

Methods: The search included PubMed, EMBASE, Cochrane and CINAHL from inception to 2016 without language restrictions. Risk scores whose main outcome was undiagnosed Type 2 diabetes reporting performance measures for Hispanics were included.

Results: We identified three studies that developed and validated risk scores for undiagnosed diabetes based on questionnaire data. Two studies were conducted in Latin America and one in the U.S. All three studies reported adequate performance (area under the receiving curve (AUC) range between 0.68 and 0.78). The study conducted in the U.S. reported a higher sensitivity of their risk score for Hispanics than whites. The limited number of studies, small size and heterogeneity of the combined cohorts provide limited evidence of the validity of risk scores for Hispanics.

Conclusions: Efforts to develop and validate risk prediction models in Hispanic populations in the U.S are needed, particularly given the diversity of this fast-growing population. Healthcare professionals providing should be aware of the limitations of applying risk scores developed for the general population on Hispanics.

Keywords: Type 2 Diabetes Mellitus, Risk Scores, Prediction Models, Hispanic

Background

The prevalence of diabetes is increasing worldwide [1-3]. Diabetes and its vascular complications are the seventh leading cause of disability worldwide and contribute to the deaths of two million adults per year [4]. Type 2 diabetes is the most common type affecting 90-95% of those with diabetes and may be asymptomatic for years [5]. Uncontrolled diabetes leads to microvascular (e.g., neuropathy, nephropathy, retinopathy) and macrovascular (e.g., myocardial infarction, stroke) complications and may increase mortality risk [6, 7]. Timely lifestyle interventions and clinical treatments can help manage hyperglycemia [8-11], reducing the risk of vascular complications. However, evidence-based interventions may not reach those with undiagnosed diabetes.

In the United States (U.S.), approximately 1 in 4 people with diabetes is undiagnosed and could be targeted for early intervention. Universal screening is recommended for high-risk adults but there is limited information on the implementation or effectiveness of national recommendations to identify those with diabetes [12]. Only half of those individuals meeting American Diabetes Association's (ADA) screening recommendation criteria report being screened, while the screening rate for those not meeting these criteria is 30% [13]. Actual diagnostic testing involving laboratory measures is less prevalent in racial/ethnic minorities, those with lower socioeconomic status, and those who lack health insurance [14]. In 2011-2012, 49% of Hispanics with diabetes living in the U.S. were undiagnosed, compared to 32.3% among non-Hispanic blacks and 33.5% among non-Hispanic whites [15]. There is also significant variability in diabetes prevalence within Hispanic groups, ranging from 10.2% among South Americans and 13.4%

among Cubans to 17.7% for Central Americans, 18.0% for Dominicans and Puerto Ricans, and 18.3% for Mexican-Americans [16]. In the Hispanic Community Health Study/Study of Latinos (SOL), one third of Hispanics with diabetes were not aware of having it, with Puerto Ricans being more likely know of their diabetes diagnosis (70%) [16]. Further efforts to identify Hispanics most at risk of having undiagnosed diabetes will help target appropriate interventions to improve health outcomes in this diverse and growing population.

Diabetes risk scores are prediction models that identify significant risk factors in the population in which they are developed. The accuracy of their predictions depends on the availability and completeness of demographic, anthropometric, clinical data, and diagnostic tests for the target population [21-24]. Limited data for Hispanic populations present multiple challenges. Electronic databases may not include detailed information on factors that increase risk of diabetes for Hispanics such as country of origin [16], length of residency in the U.S. [25, 26], stress [27] and depression [28, 29]. Questionnaire assessments may provide some of these variables but data on biomedical tests and clinical factors such as family history of diabetes, gestational diabetes or polycystic ovary syndrome depend on individual access to care. Another crucial aspect in the accuracy of risk scores is the selection of the cut-point used to define a positive test result. Lower cut-points increase the sensitivity of the model but decrease its specificity. For diabetes screening, moderate sensitivity (60%) but high specificity (90%), repeated every three years are recommended to balance disease detection avoiding false-positive results [30]. There are several risk scores to identify undiagnosed diabetes based on readily available information that have been developed and validated in the U.S. and

other countries [17-19, 31]. The purpose of this study is to identify diabetes risk scores for undiagnosed diabetes that have been developed and evaluated in Hispanic populations. We assess whether risk scores developed and validated in one cohort perform equally in other cohorts; we explore consideration of risk factors specific to Hispanics in the models, and examine methodological issues in the development, validation and comparison of diabetes risk scores in terms of their sensitivity and specificity.

2. Methods

2.1 Search strategy

We conducted a systematic literature search using PubMed, EMBASE, CINAHL and Cochrane Library from date of inception through December 31, 2016. The search strategy was based on type 2 diabetes, screening and the development and validation of a prediction tool: risk score/ assessment/algorithm/prediction/model. Subject specific and medical subject headings (MeSH) terms such as “Diabetes Mellitus Type 2”, “Prediabetes”, “screening”, “risk scores”, “algorithms” and other broad terms were included in the search. The detailed search strategy with all terms for the four databases is included in Appendix A. We also screened reference lists of previous systematic reviews of diabetes risk scores. As a validity check, 120 references from other reviews [17, 18, 31-34] were selected as an exemplary sample. A total of 16,249 references were retrieved from all four sources after removal of duplicates and the inclusion of manually searched articles. Our final sample included 103 of the 120 exemplary articles. Research library staff at our institution assisted with all electronic searches.

2.2 Inclusion criteria

The articles selected for this review met the following inclusion criteria: 1) were published in a peer reviewed journal, 2) used any study design with evidence of random selection of adult participants, 18 years or older, 3) were based on participants' data collected via questionnaire, 4) developed and validated a risk score to identify undiagnosed diabetes, 5) reported sensitivity, specificity, and area under the curve (AUC) as outcome measures specifically for Hispanics, and 6) the final instrument did not include genetic risk factors, or invasive laboratory measures. No language restriction was applied in the search.

2.3 Article selection

Three investigators (LJ, AA, JS) independently reviewed titles and abstracts. Each database had two independent reviewers. Discrepancies were resolved by a third investigator (AC) before continuing with full paper reviews. A data extraction form (Appendix B) was used for each eligible article. The initial examination of titles discarded the majority of articles because the main outcome was other than screening for Type 2 diabetes, or because the studies targeted specific populations, such as patients who had another disease, and not the general population. The reviewers examined 1,247 abstracts of which 43 were selected for full-paper evaluation. Eighteen articles studied risk factors for undiagnosed diabetes but did not develop a score or risk assessment tool; nine predicted development of diabetes but did not evaluate undiagnosed diabetes sepa-

rately, and two included invasive measures in their prediction algorithms. We also excluded six studies in which the main outcome combined undiagnosed diabetes with other forms of glucose intolerance. Only three of the remaining eight studies reported performance measures specific to Hispanics (Figure 1).

2.4 Assessment of study quality

In order to assess the quality of the risk score, we made sure the risk score was validated. Risk scores perform better for the populations they were developed in and their performance needs to be evaluated in a different population or setting [35]. Validation of a risk score can be internal, employing the part of the sample used to develop the score; temporal, using the same sample after a selected time period; or ideally, external, using a similar but not identical population [36]. In practice, simpler models that are easy to interpret and implement are preferred. When the purpose of the risk score is to identify individuals at high risk of diabetes for intervention, specific absolute risks are not necessary. However, accurate information of individual absolute risks based on modifiable risk factors will be useful to convey the benefit of the intervention to participants.

2.5 Performance of the risk scores

The predictive performance of a risk score can be evaluated using calibration and discrimination measures [35]. Calibration measures the agreement of the study's observed risk with the model's predicted risk. Discrimination (c-statistic) measures the

ability of the model to assign a higher predicted probability to those with the event compared with those without the event. The area under the receiver operating curve (AUC) may be used to assess risk score performance, such that a score of one indicates full concordance and a score of 0.50 chance agreement. Other test performance measures include sensitivity (true positive rate), specificity (true negative rate), positive predictive value or the probability that the disease is present when the test is positive, and the negative predictive value or the probability that the disease is not present when the test is negative [37].

3. Results

We identified three studies that met inclusion criteria [38-40]. All three studies developed and validated a diabetes risk score for the prediction of undiagnosed type 2 diabetes; two were conducted in Latin America (Peru and Brazil) [38, 40] and the third in the U.S.[39]. The total number of participants in the three cohorts was 7,466. Approximately 53% were Hispanics, the age ranged from 20-74 years of age. All three studies reported AUCs as a performance measure to assess discrimination. Only one study (Peru) [38] compared the performance of their newly developed score with other scores. General characteristics of all three studies are shown in Table 1.

3.1 Study characteristics

The main purpose of all three studies was to develop and validate a simple and inexpensive tool to identify individuals with undiagnosed diabetes. The Peru and U.S. studies [38, 39] used nationally representative data and the study from Brazil used two

urban populations for development and validation [40]. Diabetes was defined as fasting plasma glucose ≥126mg/dL for the Peruvian and Brazilian risk scores. The study from the U.S. [39] reported using the oral glucose tolerance test 2-h post challenge, following World Health Organization criteria [41]. Only participants who reported having no history of physician-diagnosed diabetes were included. The performance for all three scores was adequate, with the AUC ranging between 0.68 and 0.78. In the study from the U.S. the authors reported a sensitivity of 80% and specificity of 61% for its combined analysis of Hispanics, Blacks and Native Americans compared to a sensitivity of 78% and specificity of 65% among whites [39].

3.2 Study populations

Herman and colleagues [39] used data from the Second National Health and Nutrition Examination Survey (NHANES II, 1976-1980) that examined a nationally representative sample of the U.S. population. NHANES II classified individuals as Hispanics based on three questions: race (White, Black, Other); state or foreign country where the participant was born, and national origin or ancestry as reported by the participant (Central, South America; Chicano; Cuban; Mexican; Mexicano; Mexican-American; Puerto Rican; Other Spanish). All Hispanics were combined into a single group to develop the score.

The study from Peru [38] used data from a national population-based survey (2004-2005) designed to study chronic conditions. Peru is a middle-income country with a multiethnic society. Mestizos, a combination of Amerindian and European (mostly Spanish) ancestries, represent over half of the national population (59.5%); Amerindians

(mostly Quechua) represent 27.2%; Whites (4.9%) and Blacks and other (8.3%) [43]. The authors did not report data on ethnicity nor was it included as a risk factor to develop the score.

Pires de Sousa and colleagues' [40] study used data from two urban populations in Brazil collected between 1999 and 2005. Vitoria is a capital city with a population of 265,874 according to the 1996 Brazilian census. The racial makeup of its population is 52% white, 39% pardo (tri-racial heritage: European, Amerindian and West African), 7% black and 2% other [44]. The other urban location was Ouro Preto, a smaller city with a population of 37,603 and similar racial composition according to the 1996 Brazilian census. Data for the city of Vitoria were collected in 1999-2000 with a follow-up of the same cohort 5 years later. Dates were not provided for data collection in Ouro Preto. The authors included ethnicity as a risk factor when developing the diabetes risk score but it was not retained in the final model.

3.3 Risk factors considered in the risk scores

The sample size from all three studies was adequate to develop the models, based on the recommendation of a minimum of 10 events per variable to develop a predictive model [45]. All three scores tested demographic, behavioral and anthropometric risk factors. Age was included in all three scores. A complete list of risk factors considered and included in the scores is shown in Table 2.The risk factors considered by all three studies were similar. While the study from the U.S. used only self-reported information, the studies from Brazil and Peru included laboratory measurements that were not significant in their adjusted models and were not included in their final risk scores [38-40].

4. Discussion

This review identified three studies that developed and validated risk scores for undiagnosed diabetes in cohorts from Latin America and the U.S. The small number of studies; the limited number of participants, and the heterogeneity of the cohorts did not allow for a meta-analysis, providing limited evidence of the validity of undiagnosed diabetes risk scores among Hispanics. Nonetheless, findings from this study provide evidence for the need to develop and validate prediction risk models for Hispanics living in the U.S. who are at increased risk of diabetes.

The model developed by Herman and colleagues [39] has been used by the American Diabetes Association (ADA) to help identify those at risk of diabetes in the U.S.[46]. Although the performance of this model was adequate overall and among Hispanics, Native Americans and African-Americans, data used for this study dates from 1976-1980, before oversampling of Hispanics started with NHANES III [42]. We identified three other diabetes risk scores for undiagnosed diabetes that used more recent waves from NHANES [47-49]. These studies were not included in the review because the authors did not report performance measures of their final instruments for Hispanics. The Patient Self-Assessment Score developed by Bang and colleagues was developed using data from 1999-2004 and was validated using data from NHANES 2005-2006 and cohort data from two community studies [47]. Another study by He and colleagues used NHANES data from 2005-2006 for development and validation [49]. Although both studies included race as a risk factor when developing their models, race was not re-

tained in the final scores for undiagnosed diabetes. Both studies suggested that the performance of their scores needed to be evaluated for demographic subgroups. Cichosz and colleagues used NHANES data from 2005-2010 to develop an improved screening risk score based on extended predictive features [48]. Their final model included educational level and race/ethnicity, and performed better than the patient self-assessment developed by Bang [47]. However, its performance in population subgroups was not investigated.

In Latin America, Garcia-Alcala and colleagues [50] applied the Finish Diabetes Risk Score (FINDRISC) in a convenience sample to identify individuals with diabetes in Puebla, Mexico. Another study developed a score in rural Honduras using questionnaire data and point-of-care capillary glucose tests that were applied in clinics [51]. Recently, a Colombian diabetes risk score that combined undiagnosed diabetes and impaired glucose regulation as their main outcome was developed and validated using a sample of 2060 individuals in northern Colombia [52]. The relevance of these studies for application to specific national-origin Hispanic groups in the U.S. has not been explored.

The scarcity of research on diabetes risk prediction for Hispanics in Latin America and in the U.S. is a major limitation. Little is known about the recalibration of instruments already developed, particularly in national-origin subgroups. This suggests that the development of new tools in these large and growing populations who are at high risk for diabetes is warranted. Screening thresholds of common glycemic markers also need further investigation within Hispanic populations, particularly for diagnosing diabetes. A study conducted among low-income, elderly Mexicans showed disagreement between OGTT and HbA1c measurements, suggesting possible misclassification when

using HbA1c alone [53]. Other studies have confirmed significantly higher levels of HbA1c for Hispanics compared to non-Hispanic whites in the U.S. [54, 55].

Some of the limitations of this systematic review highlight the need for future research in Hispanic populations. Although this study used a comprehensive search without language restrictions, some articles may have been missed, particularly those recently published and others from Latin America not indexed by major scientific publication engines or those with inconsistent use of the term “Hispanic”. The heterogeneity among Hispanic populations is another limitation, detailed subgroup analyses of these populations were not available. Recent cohort studies that examine chronic diseases among Hispanics such as SOL in the U.S. [56] and others that combine populations from Latin America such as CESCAS I [57], INTERHEART [58] and ELSA in Brazil [59] will be important for development and validation of future risk scores for diabetes that account for subgroup variation and the inclusion of risk factors unique to Hispanics living in the U.S. and elsewehere.

It is well documented that on average Hispanics have better health upon arrival to the U.S. compared to their American counterparts [60, 61]. Although Hispanics tend to have better longevity than their socioeconomic status would predict [62], the overall health of Hispanics declines as they spend more time living in the U.S. leaning towards that of natives, or even worse [63, 64]. Migration and onward integration are major life experiences and present challenges such as discrimination, language proficiency, stress and depression. These factors may contribute to adverse health outcomes among Hispanics and their inclusion in screening tools deserves further investigation.

Diabetes risk scores are important to the prevention and timely management of a range of diseases and health complications [17-19]. However, none have been developed and validated explicitly for Hispanic populations living in the U.S. This systematic review of the literature on risk scores aimed to identify undiagnosed diabetes among Hispanic populations living in the U.S. has revealed that only a handful of studies published to date have reported performance measures specific for Hispanics. Further, only one such study developed a diabetes risk score that was validated for Hispanics in the U.S. and this used data from 1976-1980 with all Hispanic groups combined. This highlights the lack of evidence regarding the applicability of such tools for Hispanics living in the U.S. This review underscores the urgent need to develop and validate simple and inexpensive tools to identify undiagnosed diabetes for Hispanics in the U.S. who constitute a large, diverse and growing population at high risk for diabetes.

REFERENCES

1. Guariguata, L., et al., *Global estimates of diabetes prevalence for 2013 and projections for 2035.* Diabetes Res Clin Pract, 2014. **103**(2): p. 137-49.
2. NCD-RisC, *Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants.* The Lancet, 2016. **387**(10027): p. 1513-1530.
3. Shaw, J.E., R.A. Sicree, and P.Z. Zimmet, *Global estimates of the prevalence of diabetes for 2010 and 2030.* Diabetes Research and Clinical Practice, 2010. **87**(1): p. 4-14.
4. Global Burden of Disease, *Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013.* Lancet, 2015. **386**(9995): p. 743-800.
5. Beagley, J., et al., *Global estimates of undiagnosed diabetes in adults.* Diabetes Res Clin Pract, 2014. **103**(2): p. 150-60.
6. Fowler, M.J., *Microvascular and Macrovascular Complications of Diabetes.* Clinical Diabetes, 2011. **29**(3): p. 116-122.
7. Vinik, A. and M. Flemmer, *Diabetes and macrovascular disease.* Journal of Diabetes and its Complications, 2002. **16**(3): p. 235-245.
8. *The Diabetes Prevention Program (DPP): description of lifestyle intervention.* Diabetes Care, 2002. **25**(12): p. 2165-71.
9. Gerstein, H.C., et al., *Effect of rosiglitazone on the frequency of diabetes in patients with impaired glucose tolerance or impaired fasting glucose: a randomised controlled trial.* Lancet, 2006. **368**(9541): p. 1096-105.
10. Knowler, W.C., et al., *Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin.* N Engl J Med, 2002. **346**(6): p. 393-403.
11. Li, G., et al., *The long-term effect of lifestyle interventions to prevent diabetes in the China Da Qing Diabetes Prevention Study: a 20-year follow-up study.* Lancet, 2008. **371**(9626): p. 1783-9.
12. Abid, A., S. Ahmad, and A. Waheed, *Screening for Type II Diabetes Mellitus in the United States: The Present and the Future.* Clin Med Insights Endocrinol Diabetes, 2016. **9**: p. 19-22.
13. Kiefer, M.M., et al., *National patterns in diabetes screening: data from the National Health and Nutrition Examination Survey (NHANES) 2005-2012.* J Gen Intern Med, 2015. **30**(5): p. 612-8.
14. Casagrande, S.S., C.C. Cowie, and S.M. Genuth, *Self-reported prevalence of diabetes screening in the U.S., 2005-2010.* Am J Prev Med, 2014. **47**(6): p. 780-7.
15. Menke, A., et al., *Prevalence of and trends in diabetes among adults in the united states, 1988-2012.* JAMA, 2015. **314**(10): p. 1021-1029.
16. Schneiderman, N., et al., *Prevalence of diabetes among Hispanics/Latinos from diverse backgrounds: the Hispanic Community Health Study/Study of Latinos (HCHS/SOL).* Diabetes Care, 2014. **37**(8): p. 2233-9.
17. Buijsse, B., et al., *Risk assessment tools for identifying individuals at risk of developing type 2 diabetes.* Epidemiol Rev, 2011. **33**: p. 46-62.
18. Schwarz, P.E., et al., *Tools for predicting the risk of type 2 diabetes in daily*

practice. Horm Metab Res, 2009. **41**(2): p. 86-97.
19. Thoopputra, T., et al., *Survey of diabetes risk assessment tools: concepts, structure and performance.* Diabetes Metab Res Rev, 2012. **28**(6): p. 485-98.
20. Witte, D.R., et al., *Performance of existing risk scores in screening for undiagnosed diabetes: an external validation study.* Diabet Med, 2010. **27**(1): p. 46-53.
21. Colagiuri, S., et al., *Screening for type 2 diabetes and impaired glucose metabolism: the Australian experience.* Diabetes Care, 2004. **27**(2): p. 367-71.
22. Lindstrom, J. and J. Tuomilehto, *The diabetes risk score: a practical tool to predict type 2 diabetes risk.* Diabetes Care, 2003. **26**(3): p. 725-31.
23. Stern, M.P., K. Williams, and S.M. Haffner, *Identification of persons at high risk for type 2 diabetes mellitus: Do we need the oral glucose tolerance test?* Ann Intern Med, 2002. **136**.
24. Tabaei, B.P. and W.H. Herman, *A multivariate logistic regression equation to screen for diabetes: development and validation.* Diabetes Care, 2002. **25**.
25. Commodore-Mensah, Y., et al., *Length of Residence in the United States is Associated With a Higher Prevalence of Cardiometabolic Risk Factors in Immigrants: A Contemporary Analysis of the National Health Interview Survey.* J Am Heart Assoc, 2016. **5**(11).
26. Heiss, G., et al., *Prevalence of metabolic syndrome among Hispanics/Latinos of diverse background: the Hispanic Community Health Study/Study of Latinos.* Diabetes Care, 2014. **37**(8): p. 2391-9.
27. Silveira, M.L., et al., *Perceived psychosocial stress and glucose intolerance among pregnant Hispanic women.* Diabetes & Metabolism, 2014. **40**(6): p. 466-475.
28. Golden, S.H. and B. Mezuk, *The association of depressive symptoms with prediabetes versus diagnosed diabetes: is ignorance really bliss?* Phys Sportsmed, 2009. **37**(1): p. 143-5.
29. Kan, C., et al., *A systematic review and meta-analysis of the association between depression and insulin resistance.* Diabetes Care, 2013. **36**(2): p. 480-9.
30. Johnson, S.L., B.P. Tabaei, and W.H. Herman, *The efficacy and cost of alternative strategies for systematic screening for type 2 diabetes in the U.S. population 45-74 years of age.* Diabetes Care, 2005. **28**(2): p. 307-11.
31. Brown, N., et al., *Risk scores based on self-reported or available clinical data to detect undiagnosed type 2 diabetes: a systematic review.* Diabetes Res Clin Pract, 2012. **98**(3): p. 369-85.
32. Abbasi, A., et al., *Prediction models for risk of developing type 2 diabetes: systematic literature search and independent external validation study.* 2012.
33. Barber, S.R., et al., *Risk assessment tools for detecting those with pre-diabetes: a systematic review.* Diabetes Res Clin Pract, 2014. **105**(1): p. 1-13.
34. Collins, G.S., et al., *Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting.* BMC Medicine, 2011. **9**(1): p. 103.
35. Royston, P., et al., *Prognosis and prognostic research: Developing a prognostic model.* Bmj, 2009. **338**: p. b604.
36. Moons, K.G.M., et al., *Prognosis and prognostic research: application and*

impact of prognostic models in clinical practice. BMJ, 2009. **338**.
37. Harrell, F.E., K.L. Lee, and D.B. Mark, *Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors.* Stat Med, 1996. **15**.
38. Bernabe-Ortiz, A., et al., *Development and Validation of a Simple Risk Score for Undiagnosed Type 2 Diabetes in a Resource-Constrained Setting.* J Diabetes Res, 2016. **2016**: p. 8790235.
39. Herman, W.H., et al., *A new and simple questionnaire to identify people at increased risk for undiagnosed diabetes.* Diabetes Care, 1995. **18**(3): p. 382-7.
40. Pires de Sousa, A.G., et al., *Derivation and external validation of a simple prediction model for the diagnosis of type 2 diabetes mellitus in the Brazilian urban population.* Eur J Epidemiol, 2009. **24**.
41. Bartholomew, L., et al., eds. *Intervention Mapping: Designing theory and evidence-based health promotion programs*. 2001, Mayfield Publishing Company: mountain View, CA.
42. McDowell, A., et al., *Plan and operation of the Second National Health and Nutrition Examination Survey, 1976-1980.* Vital Health Stat 1, 1981(15): p. 1-144.
43. Instituto Nacional de Estadistica e Informatica, P. *PRINCIPALES INDICADORES DEMOGRÁFICOS, SOCIALES Y ECONÓMICOS A NIVEL DEPARTAMENTAL*. 2007 [cited 2017 June 17, 2017]; Available from: w w w . i n e i . g o b . p e.
44. IBGE. *Síntese de Indicadores Sociais 2000*. 2000 [cited 2017 June 17, 2017]; Available from: https://sidra.ibge.gov.br/Tabela/2094.
45. Peduzzi, P., et al., *A simulation study of the number of events per variable in logistic regression analysis.* J Clin Epidemiol, 1996. **49**.
46. Lindström, J. and J. Tuomilehto, *The diabetes risk score: a practical tool to predict type 2 diabetes risk.* Diabetes care, 2003. **26**(3): p. 725-731.
47. Bang, H., et al., *Development and validation of a patient self-assessment score for diabetes risk.* Ann Intern Med, 2009. **151**(11): p. 775-83.
48. Cichosz, S.L., et al., *Improved diabetes screening using an extended predictive feature search.* Diabetes Technol Ther, 2014. **16**(3): p. 166-71.
49. He, G., T. Sentell, and D. Schillinger, *A new public health tool for risk assessment of abnormal glucose levels.* Prev Chronic Dis, 2010. **7**(2): p. A34.
50. Garcia-Alcala, H., et al., *Frequency of diabetes, impaired fasting glucose, and glucose intolerance in high-risk groups identified by a FINDRISC survey in Puebla City, Mexico.* Diabetes Metab Syndr Obes, 2012. **5**: p. 403-6.
51. Milton, E.C., et al., *Validation of a type 2 diabetes screening tool in rural Honduras.* Diabetes Care, 2010. **33**(2): p. 275-7.
52. Barengo, N.C., et al., *A Colombian diabetes risk score for detecting undiagnosed diabetes and impaired glucose regulation.* Prim Care Diabetes, 2017. **11**(1): p. 86-93.
53. López López, R., et al., *Diabetic by HbA1c, Normal by OGTT: A Frequent Finding in the Mexico City Diabetes Study.* Journal of the Endocrine Society, 2017. **1**(10): p. 1247-1258.
54. Aviles-Santa, M.L., et al., *Differences in Hemoglobin A1c Between Hispanics/Latinos and Non-Hispanic Whites: An Analysis of the Hispanic*

Community Health Study/Study of Latinos and the 2007-2012 National Health and Nutrition Examination Survey. Diabetes Care, 2016. **39**(6): p. 1010-7.
55. Rivera-Hernandez, A., et al., *[Glycosylated hemoglobin A1c as a diagnostic test for diabetes mellitus in adolescents with overweight and obesity]*. Rev Med Inst Mex Seguro Soc, 2015. **53 Suppl 3**: p. S294-9.
56. Sorlie, P.D., et al., *Design and implementation of the Hispanic Community Health Study/Study of Latinos.* Ann Epidemiol, 2010. **20**(8): p. 629-41.
57. Rubinstein, A.L., et al., *Detection and follow-up of cardiovascular disease and risk factors in the Southern Cone of Latin America: the CESCAS I study.* BMJ Open, 2011. **1**(1): p. e000126.
58. Lanas, F., et al., *Risk factors for acute myocardial infarction in Latin America: the INTERHEART Latin American study.* Circulation, 2007. **115**(9): p. 1067-74.
59. Aquino, E.M., et al., *Brazilian Longitudinal Study of Adult Health (ELSA-Brasil): objectives and design.* Am J Epidemiol, 2012. **175**(4): p. 315-24.
60. Franzini, L., J.C. Ribble, and A.M. Keddie, *Understanding the Hispanic paradox.* Ethn Dis, 2001. **11**(3): p. 496-518.
61. Ruiz, J.M., P. Steffen, and T.B. Smith, *Hispanic mortality paradox: a systematic review and meta-analysis of the longitudinal literature.* Am J Public Health, 2013. **103**(3): p. e52-60.
62. Jasso, G. and D.S. Massey, *Immigrant health: selectivity and acculturation*. 2004, IFS Working Papers, Institute for Fiscal Studies (IFS).
63. Abraido-Lanza, A.F., M.T. Chao, and K.R. Florez, *Do healthy behaviors decline with greater acculturation? Implications for the Latino mortality paradox.* Soc Sci Med, 2005. **61**(6): p. 1243-55.
64. Oza-Frank, R. and S.A. Cunningham, *The weight of US residence among immigrants: a systematic review.* Obes Rev, 2010. **11**(4): p. 271-80.
65. Liberati, A., et al., *The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration.* Annals of internal medicine, 2009. **151**(4): p. W-65-W-94.

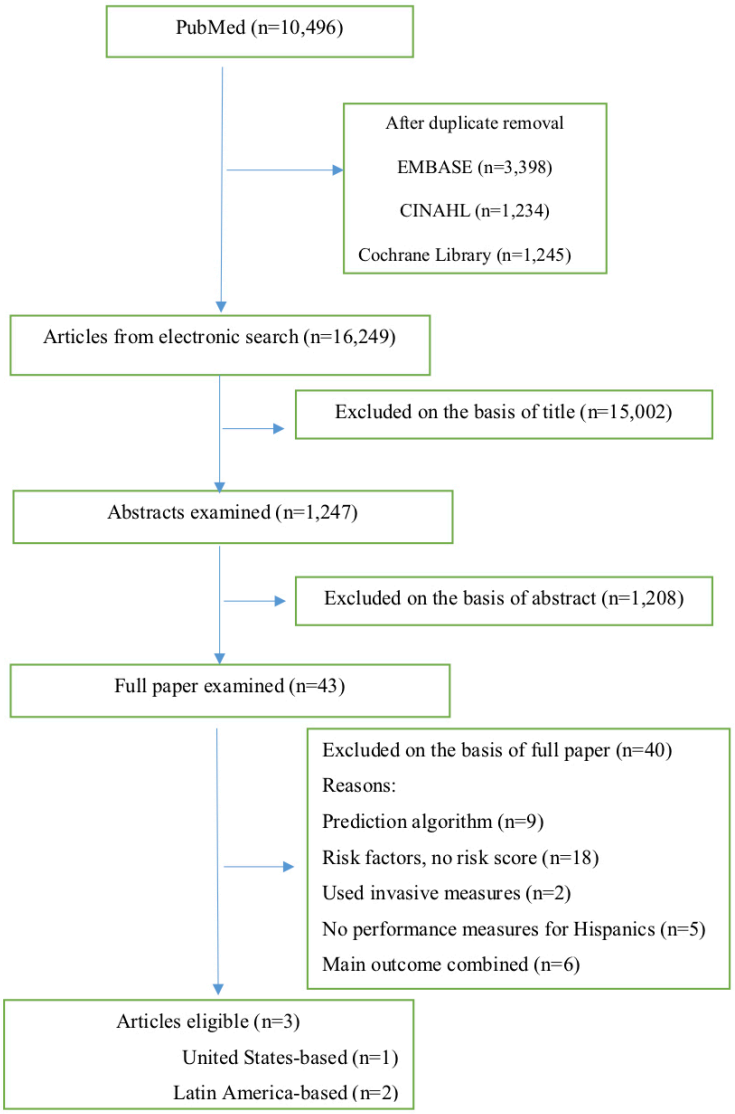

Figure 1. Flow diagram of articles selected for review of diabetes risk scores for Hispanic populations in the U.S. (Figure adapted from Liberati et al., 2009 [65])

Table 1. Diabetes risk score characteristics for included articles

	Peru [38]	Brazil [40]	USA [39]
Year published	2016	2009	1995
Year data development	2004-2005	1999-2000 c.	1976-1980
Year data validation	2010-2014	2005	1976-1980
Sample size	2472	1224	3770*
Target age group	≥ 35	> 35	20 - 74
Diabetes criteria	FPG ≥ 126 md/dL	FPG > 126 md/dL	OGTT (WHO)
# risk factors	3	3	6
Scoring criteria	0 -5	0 - 48	decision tree
AUC	0.68	0.77	0.78
Sensitivity	70	76	80
Specificity	59	67	61

* Total number of participants. The study did not report number of Hispanic participants. A similar study analyzing diabetes trends using NHANES II reported approximately 2.9% of Hispanics.

Table 2. Risk factors considered and odds ratio with 95% confidence interval for those included in risk scores

Risk factor	Peru [38] OR (95%CI)	Brazil [40] OR (95%CI)	USA* [39]
Demographic			
Age			
≥45 (versus <45 years)			●
≥55 (versus <55 years)	● 1.85 (1.30–2.63)		
45–54 versus 35–44 year		● 2.10 (1.15–3.84)	
55 or more versus 35–44 year		3.41 (1.89–6.15)	
Gender	❍	●	●
Ethnicity		❍	❍
Education	❍		❍
Behavioral			
Sendentarism/PA	❍	❍	●
Smoking	❍	❍	
Alcohol	❍		
General Health Assessment			❍
History of diabetes			❍
Family History Diabetes	● 2.34 (1.04–5.31)		●
Macrosomic infant			●
Anthropometric			
BMI	❍		
30 or more versus ≤30			●
25–29.9 versus ≤25		● 1.61 (0.90–2.87)	
30 or more versus ≤25		6.06 (3.49–10.52)	
Waist circumference			
90.0 to <99.9 cm (versus <90 cm)	● 2.09 (1.09–4.02)		
100+ cm (versus <90 cm)	4.07 (2.60–6.40)		
Waist to height ratio	❍		
Waist to hip ratio		❍	
Hypertension	❍	● 1.87 (1.18–2.97)	❍
Systolic blood pressure	❍	❍	
Diastolic blood pressure	❍	❍	
Total cholesterol	❍	❍	
HDL	❍	❍	
LDL		❍	
CT ≥ 240 versus <240		❍	
TG		❍	
Uric acid		❍	
Creatinine		❍	

● Included
❍ Considered
* Diabetes risk score used a classification tree. No odds ratios were reported.

APPENDIX A
SYSTEMATIC LITERATURE REVIEW WORKSHEET

by Murray Turner, University of Canberra, 2015. Designed to meet IOM Standard 3.4.1, and for use with the PRISMA 2009 Flow Diagram.[1]

SYSTEMATIC LITERATURE REVIEW WORKSHEET

by Murray Turner, University of Canberra, 2015. Designed to meet IOM Standard 3.4.1, and for use with the PRISMA 2009 Flow Diagram. [1]

Database searched	Date of Search	Search Terms	Filters / Limiters applied[2]
BIBLIOGRAPHIC DATABASES:			
PubMed	2/10/17	"Diabetes Mellitus, Type 2"[Mesh] OR NIDDM [tf] OR "Prediabetic State"[Mesh] OR prediabet* [tiab] OR pre-diabet* [tiab] OR ((unidentified [tiab] OR undiagnosed [tiab] OR unrecognized [tiab] OR unrecognised [tiab]) AND (diabetes [tiab] OR diabetic [tiab])) OR "Hyperglycemia"[Mesh] OR hyperglycemia [tiab] OR hyperglycaemia [tiab] OR "Glucose Tolerance Test"[Mesh] OR dysglycemia [tiab] OR dysglycaemia [tiab] OR "Diabetes Mellitus"[Mesh:NoExp] OR type-2-diabet* [tiab] OR type-ii-diabet* [tiab] OR adult-onset-diabet* [tiab] OR "Noninsulin Dependent Diabetes" [tiab] OR "Non-insulin Dependent Diabetes" [tiab] OR "impaired fasting glucose" [tiab] OR "impaired glucose tolerance" [tiab] OR "glucose intolerance" [tiab] AND assessment-tool* [tiab] OR assessment-score* [tiab] OR assessment-instrument* [tiab] OR assessment-model* [tiab] OR assessment-test* [tiab] OR Risk-score* [tiab] OR risk-scoring [tiab] OR clinical-scor* [tiab] OR clinical-model* [tiab] OR risk-tool* [tiab] OR risk-model* [tiab] OR risk-index [tiab] OR risk-indices [tiab] OR risk-algorithm* [tiab] OR risk-questionnaire* [tiab] OR risk-predict* [tiab] OR risk-equation* [tiab] OR risk-calculat* [tiab] OR "risk factor score" [tiab] OR "risk factor scores" [tiab] OR "risk factor scoring" [tiab] OR risk-factor-tool* [tiab] OR risk-factor-model* [tiab] OR risk-factor-algorithm* [tiab] OR risk-factor-questionnaire* [tiab] OR risk-factor-prediction* [tiab] OR risk-factor-equation* [tiab] OR risk-factor-calculat* [tiab] OR risk-factor-assessment* [tiab] OR "measuring risk" [tiab] OR "measure risk" [tiab] OR "calculating risk" [tiab] OR prediction-model* [tiab] OR model-predict* [tiab] OR predictive-model* [tiab] OR prediction-rule [tiab] OR screening-test* [tiab] OR screening-tool* [tiab] OR screening-instrument* [tiab] OR screening-model* [tiab] OR screening-score* [tiab] OR screening-questionnaire* [tiab] OR screening-protocol* [tiab] OR questionnaire score* [tiab] OR predictive-instrument* [tiab] OR prediction-instrument [tiab] OR predictive-equation* [tiab] OR prediction-equation* [tiab] OR prediction-tool* [tiab] OR prediction-rule* [tiab] OR predictive-rule* [tiab] OR predictive-tool* [tiab] OR prediction-scor* [tiab] OR predictive-scor* [tiab] OR predict-risk* [tiab] OR project-risk* [tiab] OR scoring-model* [tiab] OR score-model* [tiab] OR scoring-method* [tiab] OR scoring-tool* [tiab] OR scoring-questionnaire* [tiab] OR scoring-scheme* [tiab] OR scoring-instrument* [tiab] OR scoring-system* [tiab] OR classification-scheme* [tiab] OR numerical-scor* [tiab] OR recursive-partition* [tiab] OR (decision* [tiab] AND (model* [tiab] OR clinical* [tiab] OR "Logistic Models"[Mesh])) = 5997 OR "Diabetes Mellitus, Type 2"[Mesh] OR NIDDM [tf] OR "Prediabetic State"[Mesh] OR prediabet* [tiab] OR pre-diabet* [tiab] OR ((unidentified [tiab] OR undiagnosed [tiab] OR unrecognized [tiab] OR unrecognised [tiab])) AND (diabetes [tiab] OR diabetic [tiab])) OR "Hyperglycemia"[Mesh] OR hyperglycemia [tiab] OR hyperglycaemia [tiab] OR "Glucose Tolerance Test"[Mesh] OR dysglycemia [tiab] OR dysglycaemia [tiab] OR "Diabetes Mellitus"[Mesh:NoExp] OR type-2-diabet* [tiab] OR type-ii-diabet* [tiab] OR adult-onset-diabet* [tiab] OR "Noninsulin Dependent Diabetes" [tiab] OR "Non-insulin Dependent Diabetes" [tiab] OR "impaired fasting glucose" [tiab] OR "impaired glucose tolerance" [tiab] OR "glucose intolerance" [tiab] AND Mass Screening [mh] OR screen* [tiab] OR detect* [tf] OR predict* [tf] OR identif* [tf] OR multiphasic [tiab] AND score [tf] OR scores [tf] OR scoring [tf] OR tool [tf] OR tools [tf] OR instrument [tf] OR instruments [tf] OR model* [tf] OR equation [tf] OR equations [tf] OR "Surveys and Questionnaires"[Mesh:noexp] OR survey [tf] OR surveys [tf] OR questionnaire* [tf] OR Algorithms [mh] OR algorithm [tf] OR algorithms [tf] OR test [tf] OR tests [tf] OR measur* [tf] OR calculat* [tf] OR estimat* [tf] OR quantif* [tf] OR index [tf] OR instrumentation [sh] OR methods [sh] = 6233 Total = 10740	NOT ("animals"[MeSH Terms] NOT "humans"[MeSH Terms]) = 10543 refs AND 1950:2016 [dp] = 10496 refs
EMBASE	2/28/17	non insulin dependent diabetes mellitus'/exp OR NIDDM:ti OR 'impaired glucose tolerance'/exp OR prediabet*:ab,ti OR pre-diabet*:ab,ti OR ((Unidentified:ab,ti OR undiagnosed:ab,ti OR unrecognized:ab,ti OR unrecognised:ab,ti) AND (diabetes:ab,ti OR diabetic:ab,ti)) OR 'hyperglycemia'/exp OR hyperglycemia:ab,ti OR hyperglycaemia:ab,ti OR 'glucose tolerance test'/exp OR dysglycemia:ab,ti OR dysglycaemia:ab,ti OR 'diabetes mellitus'/de OR type-2-diabet*:ab,ti OR type-ii-diabet*:ab,ti OR adult-onset-diabet*:ab,ti OR "Noninsulin Dependent Diabetes":ab,ti OR "Non-insulin Dependent Diabetes":ab,ti OR "impaired fasting glucose":ab,ti OR "impaired glucose tolerance":ab,ti OR "glucose intolerance":ab,ti AND 'assessment tool*':ab,ti OR 'assessment score*':ab,ti OR 'assessment instrument*':ab,ti OR 'assessment model*':ab,ti OR 'assessment test*':ab,ti OR 'Risk score*':ab,ti OR 'risk scoring':ab,ti OR 'clinical scor*':ab,ti OR 'clinical model*':ab,ti OR 'risk tool*':ab,ti OR 'risk model*':ab,ti OR 'risk index':ab,ti OR 'risk indices':ab,ti OR 'risk algorithm*':ab,ti OR 'risk questionnaire*':ab,ti OR 'risk predict*':ab,ti OR 'risk equation*':ab,ti OR 'risk calculat*':ab,ti OR "risk factor score":ab,ti OR "risk factor scores":ab,ti OR "risk factor scoring":ab,ti OR 'risk factor tool*':ab,ti OR 'risk factor model*':ab,ti OR 'risk factor algorithm*':ab,ti OR 'risk factor questionnaire*':ab,ti OR 'risk factor prediction*':ab,ti OR 'risk factor equation*':ab,ti OR 'risk factor calculat*':ab,ti OR 'risk factor assessment*':ab,ti OR "measuring risk":ab,ti OR "measure risk":ab,ti OR "calculating risk":ab,ti OR 'prediction model*':ab,ti OR 'model predict*':ab,ti OR 'predictive model*':ab,ti OR 'prediction rule':ab,ti OR 'screening test*':ab,ti OR 'screening tool*':ab,ti OR 'screening instrument*':ab,ti OR 'screening model*':ab,ti OR 'screening score*':ab,ti OR 'screening questionnaire*':ab,ti OR 'screening protocol*':ab,ti OR 'questionnaire score*':ab,ti OR 'predictive-instrument*':ab,ti OR 'prediction instrument*':ab,ti OR 'predictive-equation*':ab,ti OR 'prediction equation*':ab,ti OR 'prediction tool*':ab,ti OR 'prediction rule*':ab,ti OR 'predictive rule*':ab,ti OR 'predictive tool*':ab,ti OR 'prediction scor*':ab,ti OR 'predictive scor*':ab,ti OR 'predict risk*':ab,ti OR 'project risk*':ab,ti OR 'scoring model*':ab,ti OR 'score model*':ab,ti OR 'scoring method*':ab,ti OR 'scoring tool*':ab,ti OR 'scoring questionnaire*':ab,ti OR 'scoring scheme*':ab,ti OR 'scoring instrument*':ab,ti OR 'scoring system*':ab,ti OR 'classification scheme*':ab,ti OR 'numerical scor*':ab,ti OR 'recursive partition*':ab,ti OR (decision* NEAR/5 (model* OR clinical*)):ab,ti = 16954 OR 'non insulin dependent diabetes mellitus'/exp OR NIDDM:ti OR 'impaired glucose tolerance'/exp OR prediabet*:ab,ti OR pre-diabet*:ab,ti OR ((Unidentified:ab,ti OR undiagnosed:ab,ti OR unrecognized:ab,ti OR unrecognised:ab,ti) AND (diabetes:ab,ti OR diabetic:ab,ti)) OR 'hyperglycemia'/exp OR hyperglycemia:ab,ti OR hyperglycaemia:ab,ti OR 'glucose tolerance test'/exp OR dysglycemia:ab,ti OR dysglycaemia:ab,ti OR 'diabetes mellitus'/de OR type-2-diabet*:ab,ti OR type-ii-diabet*:ab,ti OR adult-onset-diabet*:ab,ti OR "Noninsulin Dependent Diabetes":ab,ti OR "Non-insulin Dependent Diabetes":ab,ti OR "impaired fasting glucose":ab,ti OR "impaired glucose tolerance":ab,ti OR "glucose intolerance":ab,ti AND 'mass screening'/exp OR screen*:ab,ti OR detect*:ti OR predict*:ti OR identif*:ti OR multiphasi:ab,ti AND score:ti OR scores:ti OR scoring:ti OR tool:ti OR tools:ti OR instrument:ti OR instruments:ti OR model*:ti OR equation:ti OR equations:ti OR 'questionnaire'/exp OR survey:ti OR surveys:ti OR questionnaire*:ti OR 'algorithm'/exp OR algorithm:ti OR algorithms:ti OR test:ti OR tests:ti OR measur*:ti OR calculat*:ti OR estimat*:ti OR quantif*:ti OR index:ti = 10579 Total = 23967	NOT ['animals'/exp NOT 'humans'/exp] = 23640 AND [1950-2016]/py = 23427 AND [embase]/lim NOT [medline]/lim = 13021 NOT ([Conference Abstract]/lim OR [Conference Paper]/lim) = 1398

Record identification and Screening	**CINAHL**	2/13/17	MH Diabetes Mellitus, Type 2 OR TI NIDDM OR AB NIDDM OR MH Prediabetic State OR TI prediabet* OR AB prediabet* OR TI pre-diabet* OR AB pre-diabet* OR (JTI unidentified OR AB unidentified OR TI undiagnosed OR AB undiagnosed OR TI unrecognized OR AB unrecognized OR TI unrecognised OR AB unrecognised) AND (TI diabetes OR AB diabetes OR TI diabetic OR AB diabetic)) OR MH Hyperglycemia OR TI hyperglycemia OR AB hyperglycemia OR TI glycemia OR AB hyperglycaemia OR MH Glucose Tolerance Test OR TI dysglycemia OR AB dysglycemia OR TI dysglycaemia OR AB dysglycaemia OR MH Diabetes Mellitus OR TI type-2 diabet* OR AB type-2 diabet* OR TI type-ii diabet* OR AB type-ii diabet* OR TI adult-onset-diabet* OR AB adult-onset-diabet* OR TI "noninsulin dependent diabetes" OR AB "noninsulin dependent diabetes" OR TI "non-insulin dependent diabetes" OR AB "non-insulin dependent diabetes" OR TI "impaired fasting glucose" OR AB "impaired fasting glucose" OR TI "impaired glucose tolerance" OR AB "impaired glucose tolerance" OR TI "glucose intolerance" OR AB "glucose intolerance" OR MH Glucose Intolerance AND MH Health Screening OR TI mass-screening OR AB mass-screening OR TI screen* OR AB screen* OR TI detect* OR AB detect* OR MH Forecasting OR TI predict* OR AB predict* OR TI identif* OR AB identif* OR TI multiphasic OR AB multiphasic AND TI score OR TI scores OR TI scoring OR TI tool OR TI tools OR TI instrument OR TI instruments OR TI model* OR TI equation OR TI equations OR MH Surveys OR MH Questionnaires OR MH Structured Questionnaires OR TI survey OR TI surveys OR TI questionnaire OR TI questionnaires OR MH Algorithms OR TI algorithm OR TI algorithms OR TI test OR TI tests OR TI measur* OR TI calculat* OR TI estimat* OR quantif* OR TI index OR SH instrumentation OR SH methods = 4028 OR MH Diabetes Mellitus, Type 2 OR TI NIDDM OR AB NIDDM OR MH Prediabetic State OR TI prediabet* OR AB prediabet* OR TI pre-diabet* OR AB pre-diabet* OR (JTI unidentified OR AB unidentified OR TI undiagnosed OR AB undiagnosed OR TI unrecognized OR AB unrecognized OR TI unrecognised OR AB unrecognised) AND (TI diabetes OR AB diabetes OR TI diabetic OR AB diabetic)) OR MH Hyperglycemia OR TI hyperglycemia OR AB hyperglycemia OR TI hyperglycaemia OR AB hyperglycaemia OR MH Glucose Tolerance Test OR TI dysglycemia OR AB dysglycemia OR TI dysglycaemia OR AB dysglycaemia OR MH Diabetes Mellitus OR TI type-2 diabet* OR AB type-2 diabet* OR TI type-ii diabet* OR AB type-ii diabet* OR TI adult-onset-diabet* OR AB adult-onset-diabet* OR TI "noninsulin dependent diabetes" OR AB "noninsulin dependent diabetes" OR TI "non-insulin dependent diabetes" OR AB "non-insulin dependent diabetes" OR TI "impaired fasting glucose" OR AB "impaired fasting glucose" OR TI "impaired glucose tolerance" OR AB "impaired glucose tolerance" OR TI "glucose intolerance" OR AB "glucose intolerance" OR MH Glucose Intolerance AND MH Clinical Assessment Tools OR TI assessment-tool* OR AB assessment-tool* OR TI assessment-score* OR AB assessment-score* OR TI assessment-instrument* OR AB assessment-instrument* OR TI assessment-model* OR AB assessment-model* OR TI assessment-test* OR AB assessment-test* OR TI risk-score* OR AB risk-score* OR TI risk-scoring OR AB risk-scoring OR TI clinical-scor* OR AB clinical-scor* OR TI clinical-model* OR AB clinical-model* OR TI risk-tool* OR AB risk-tool* OR TI risk-model* OR AB risk-model* OR TI risk-index OR AB risk-index OR TI risk-indices OR AB risk-indices OR TI risk-algorithm* OR AB risk-algorithm* OR TI risk-questionnaire* OR AB risk-questionnaire* OR TI risk-predict* OR AB risk-predict* OR TI risk-equation* OR AB risk-equation* OR TI risk-calculat* OR AB risk-calculat* OR TI "risk factor score" OR AB "risk factor score" OR TI "risk factor scores" OR AB "risk factor scores" OR TI "risk factor scoring" OR AB "risk factor scoring" OR TI risk-factor-tool* OR AB risk-factor-tool* OR TI risk-factor-model* OR AB risk-factor-model* OR TI risk-factor-algorithm* OR AB risk-factor-algorithm* OR TI risk-factor-questionnaire* OR AB risk-factor-questionnaire* OR TI risk-factor-prediction* OR AB risk-factor-prediction* OR TI risk-factor-equation* OR AB risk-factor-equation* OR TI risk-factor-calculat* OR AB risk-factor-calculat* OR TI risk-factor-assessment* OR AB risk-factor-assessment* OR TI "measuring risk" OR AB "measuring risk" OR TI "measure risk" OR AB "measure risk" OR TI "calculating risk" OR AB "calculating risk" OR TI prediction-model* OR AB prediction-model* OR TI model-predict* OR AB model-predict* OR TI predictive-model* OR AB predictive-model* OR TI prediction-rule OR AB prediction-rule OR TI screening-test* OR AB screening-test* OR TI screening-tool* OR AB screening-tool* OR TI screening-instrument* OR AB screening-instrument* OR TI screening-model* OR AB screening-model* OR TI screening-score* OR AB screening-score* OR TI screening-questionnaire* OR AB screening-questionnaire* OR TI screening-protocol* OR AB screening-protocol* OR TI questionnaire-score* OR AB questionnaire-score* OR TI predictive-instrument* OR AB predictive-instrument* OR TI prediction-instrument* OR AB prediction-instrument* OR TI predictive-equation* OR AB predictive-equation* OR TI prediction-equation* OR AB prediction-equation* OR TI prediction-tool* OR AB prediction-tool* OR TI prediction-rule* OR AB prediction-rule* OR TI predictive-rule* OR AB predictive-rule* OR TI predictive-tool* OR AB predictive-tool* OR TI prediction-scor* OR AB prediction-scor* OR TI predictive-scor* OR AB predictive-scor* OR TI predict-risk* OR AB predict-risk* OR TI project-risk* OR AB project-risk* OR TI scoring-model* OR AB scoring-model* OR TI score-model* OR AB score-model* OR TI scoring-method* OR AB scoring-method* OR TI scoring-tool* OR AB scoring-tool* OR TI scoring-questionnaire* OR AB scoring-questionnaire* OR TI scoring-scheme * OR AB scoring-scheme * OR TI scoring-instrument * OR AB scoring-instrument * OR TI scoring-system* OR AB scoring-system* OR TI classification-scheme* OR AB classification-scheme* OR TI numerical-scor* OR AB numerical-scor* OR TI recursive-partition* OR AB recursive-partition* OR (JTI decision* OR AB decision*) AND (TI model* OR AB model* OR TI clinical* OR AB clinical* OR MH	Exclude PubMed =1234
	Cochrane	3/3/17	ID Search Hits #1 (screen or screening or screened or detect or detected or detecting or detection or predict or predicted or prediction or predicting or identify or identifies or identified or identification or risk or risks or assessment) near/3 (score or scores or scoring or tool or tools or instrument or instruments or model or models or modeling or modelling or equation or equations or survey or surveys or questionnaire or algorithm or algorithms or test or tests or measure or measures or measured or measuring or calculate or calculated or calculating or estimate or estimating or estimated or estimation or quantify or quantified or quantification or index or indices or instrumentation or rule):ti (Word variations have been searched) 3675 #2 (screen or screening or screened or detect or detected or detecting or detection or predict or predicted or prediction or predicting or identify or identifies or identified or identification or risk or risks or assessment) near/3 (score or scores or scoring or tool or tools or instrument or instruments or model or models or modeling or modelling or equation or equations or survey or surveys or questionnaire or algorithm or algorithms or test or tests or measure or measures or measured or measuring or calculate or calculated or calculating or estimate or estimating or estimated or estimation or quantify or quantified or quantification or index or indices or instrumentation or rule):ab (Word variations have been searched) 29590 #3 #1 or #2 31538 #4 type 2 diabetes or type ii diabetes:ti (Word variations have been searched) 10755 #5 type 2 diabetes or type ii diabetes:ab (Word variations have been searched) 16368 #6 #4 or #5 19105 #7 "type 2 diabetes" or "type ii diabetes":ti (Word variations have been searched) 10418 #8 "type 2 diabetes" or "type ii diabetes":ab (Word variations have been searched) 13044 #9 #7 or #8 15846 #10 #3 and #9 1124 #11 recursive partition:ti,ab,kw (Word variations have been searched) 139 #12 #9 and #11 0 #13 risk near/3 predict*:ti,ab,kw (Word variations have been searched) 2451 #14 #9 and #13 116 #15 (score or scoring) near/3 (model* or tool* or questionnaire* or instrument*):ti,ab,kw (Word variations have been searched) 4012 #16 #9 and #15 71 #17 #10 or #14 or #16 1245 refs	
	GRIY LITERATURE DATABASES:			
	OTHER SOURCES (Eg. handsearching, personal			

Include Eligibility	**RECORDS AFTER DUPLICATES REMOVED:**[4]	
	STUDIES AFTER FULL-TEXT ASSESSED FOR ELIGIBILITY:[5]	
	TOTAL STUDIES INCLUDED IN QUALITATIVE SYNTHESIS:	
	TOTAL STUDIES INCLUDED IN META-ANALYSIS:[6]	

NOTES:

1 - PRISMA Statement and 2009 Flow Diagram available at: http://www.prisma-statement.org/statement.htm
IOM Standard available at: http://iom.nationalacademies.org/Reports/2011/Finding-What-Works-in-Health-Care-Standards-for-Systematic-Reviews/Standards.aspx

2 - Filters / Limiters: Provide justification for applying filters / limiters to the search.

3 - Screening search results: Initial screening of records based title and abstract information. Keep a log of records excluded.

4 - Duplicate record removal: easily achieved by citation management software such as EndNote.

5 - Exclusion of full-text articles: Final exclusion of records based on pre-defined exclusion/inclusion criteria applied to full-text. Add to log of records excluded.

6 - Optional step: Only required if including a synthesis of quantitative data.

APPENDIX B

DATA EXTRACTION FORM

General information

Title of paper
First author
Name of risk score
Journal
Date published
Notes:

Sample information

Name of study which the data is from
Year
Country/countries
% of Hispanics
Primary reason for cohort
Sampling frame (inclusion/exclusion criteria and key characteristics)
Sample size
Outcome & definition used (if appropriate)
Outcome Incidence rate and (number of events) (if appropriate)
Notes:

Model information

Method for developing risk score
Number of variables considered for inclusion in the model
Variables included in final model
Variables considered but not included in final model
How was model/variables selected?
Treatment of continuous data
Treatment of missing data
Method of choosing cut-off (decided in advance?)
Internal sensitivity/specificity at recommended cut-off (if reported)
Internal PPV/NPV at recommended cut-off (if reported)
% Needing further testing (if reported)
Area under ROC
Any internal validation of the model
Calibration
Way in which the risk score can be completed (e.g. self-assessment, MD assessment, etc.)?
Notes:

External validation

Was an external validation carried out by the author in the same paper?
Year
Country/countries
Primary reason for cohort

Sampling frame
Sample size & number of events
Was the same outcome definition used?
Sensitivity/specificity at recommended cut-off (if reported)
PPV/NPV at recommended cut-off (if reported)
Area under ROC
Calibration
Any other external assessments of the model using this dataset
Notes:

Author's assessment
Intended use of risk score (who will use score and mechanism of implementation)
Recommended action should be taken by those who score above the cut-off
Strengths of score
Weaknesses of score
Notes:

Adapted from [33].

A RISK SCORE FOR UNDIAGNOSED DIABETES FOR HISPANICS LIVING IN THE U.S.: THE HISPANIC COMMUNITY HEALTH STUDY/STUDY OF LATINOS (HCHS/SOL)

by

LUCIA D. JUAREZ; ANDRES AZUERO; JEFFREY S. GONZALEZ; GREGORY A. TALAVERA; MATTHEW J. O'BRIEN; ANDREA L. CHERRINGTON; NEIL SCHNEIDERMAN

In preparation for Diabetes Care

Format adapted for dissertation

ABSTRACT

Objective: Hispanics in the United States (U.S.) are at a higher risk of undiagnosed diabetes. We developed and validated a risk score for undiagnosed diabetes based on self-reported information using data from Hispanics living in the U.S. and compared its performance to that of risk scores previously developed.

Research and Design Methods: We used a split-sample from the baseline visit of the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) to develop (n=7,765) and validate (n=4,688) the risk score. Multivariate logistic regression model coefficients were used to assign scores to each variable category. The SOL risk score was defined as the sum of these individual scores.

Results: Being of Mexican descent, male gender, older age, lower education, being born in the U.S., having a family history of diabetes, being overweight or obese, having had gestational diabetes and not complying with physical activity recommendations were associated with higher risk of having undiagnosed diabetes. Prevalence of undiagnosed diabetes was 7.9%. The score value ≥ 11 had sensitivity of 78%, specificity of 60% and positive predictive value of 15%. The Area under the Receiver Operating Curve (AUC) for the SOL score was 77%. Comparison scores had AUCs between 69% and 75%, sensitivities between 37% and 84% and specificities between 46% and 80%.

Conclusions: The SOL score includes education, ethnic background and being born in the U.S. as risk factors. Its performance was superior to that of other available scores and comparable to the performance of the ADA risk score. Further evaluation of the SOL risk score's performance is needed.

Keywords: Undiagnosed diabetes, Type 2 Diabetes Mellitus, Risk Scores, Prediction Models, Hispanic, Latino

INTRODUCTION

The prevalence of Type 2 diabetes continues to increase worldwide. Diabetes may be asymptomatic for years, which may lead to serious complications. In the United States (U.S.), nearly one in four people who have diabetes are undiagnosed. Many of them remain undiagnosed until they present disease-related complications. Recommendations to identify adults at high risk of diabetes are set in the U.S. However, there is not enough information to assess their implementation or effectiveness (1). There is evidence that clinical testing to diagnose diabetes involving laboratory measures is less prevalent in minority populations and among less privileged individuals with low socioeconomic status or no health insurance (2). Based on nationally representative data from the National Health and Nutrition Examination Survey (NHANES), in 2011-2012, almost half (49%) of Hispanics who had diabetes in the U.S. were undiagnosed. A significantly higher rate than among non-Hispanic blacks (32.3%) or non-Hispanic whites (33.5%) (3). Risk scores are simple prediction models that can identify and classify individuals according to their risk for undiagnosed diabetes. Risk scores that are validated in the populations they will be applied to in practice are effective screening tools. A recent systematic review identified few tools that have been developed for use among Hispanics in the U.S. and none that had been tested using recent data or that take into account the heterogeneity among Hispanic populations in the U.S. (4).

The aim of this study is to develop and validate a risk score to identify Hispanics with undiagnosed diabetes using data from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) which includes variables such as Hispanic background, whether

they were born in the U.S. and years living in the U.S. each of which has been associated with diabetes prevalence (5-7).

RESEARCH DESIGN AND METHODS

The HCHS/SOL (SOL) is a prospective, multicenter community-based study of Hispanic adults in the U.S. The original baseline cohort includes 16,415 men and women 18-74 years of age between 2008-2011 from four cities: Bronx, NY; Chicago, IL; Miami, FL; and San Diego, CA. The study design and protocol of the study have been published elsewhere (8). Study measures included 2-hour oral glucose tolerance test (OGTT), fasting plasma glucose (FPG) and Glycosylated hemoglobin (A1C). Participants were required to fast for at least 8 h prior to the visit, consuming only water and necessary medications. Venous blood specimens were collected, processed, and frozen on site toward the beginning of the visit and also 2 h after a 75 g glucose load. Plasma glucose was assessed using a hexokinase enzymatic method (Roche Diagnostics Corporation, Indianapolis, IN). Glycosylated hemoglobin (A1C) was measured in EDTA whole blood using a Tosoh G7 automated high-performance liquid chromatography analyzer (Tosoh Bioscience Inc., San Francisco, CA). Personal and family history of medical diagnoses, including diabetes, was obtained via a self-report interview administered by certified assessors (7).

The outcome variable used in the analysis is presence of newly diagnosed diabetes defined as either FPG ≥126 mg/dL (7 mmol/L), a 2-h post load glucose level (2-h OGTT) ≥200 mg/dL (11.2 mmol/L) or A1C level ≥6.5% (48 mmol/mol). Participants who reported having been diagnosed with diabetes by a doctor were excluded from the analysis.

The risk factors considered in the algorithm were selected from a pool of studies that developed similar risk scores (4). We used gender, age, Hispanic background heritage group, years in the U.S., education, family history of diabetes, smoking, alcohol use, waist circumference, BMI, gestational diabetes, self-reported high cholesterol, self-reported hypertension, and whether the individual followed physical activity recommendations (according to 2008 activity level guidelines). Physical activity is an established risk factor for diabetes (12). The HCHS/SOL study collected data from the Global Physical Activity Questionnaire (GPAQ) to estimate the number of active minutes of each individual. In practice, other risk scores like the ADA risk score use a simple question asking the individual whether he or she is physically active. Although, diet-related data were available and despite the fact that diet quality has been associated with the risk of type 2 diabetes (10; 11), we excluded this factor because assessing diet quality would require a more intricate data collection.

The risk score was derived from a development sample (62% of the original sample) using logistic regression. The random split of 62%-38% was based on the probability of each unique observation being included in a resample of equal size to the original sample size. On average, a random sample will include 62% of the original observations while the rest, 38% will be missing from that sample (13). Each potential risk factor was assessed in bivariate logistic models using undiagnosed diabetes as a dependent variable. Risk factors with a p-value of 0.10 or less were included in multivariable logistic models using manual stepwise elimination with a significance level of 5%. This approach was chosen because it allowed for the incorporation of sampling weights and the development of a clinically applicable risk score. Interaction terms were

tested in pairs systematically before obtaining the final model. Decisions on whether to keep or remove a variable or interaction were based on changes in the log likelihood ratio. The effect of gestational diabetes was examined by running models stratified by gender and by adding an interaction term with gender and gestational diabetes to the final models including all participants. The effect of being female and having had gestational diabetes was similar in all models confirming the effect was not exaggerated. Participants with missing outcome variables were excluded from the analysis. Only complete records were included in the analysis. All continuous variables were assessed for linearity using basic regression diagnostic procedures and standardized deviance residual plots, and non-linear relationships were explored.

The apparent discrimination of the model was measured using the Area under the Receiver Operating Curve (AUC). We used internal validation relying on a split-sample approach with 38% of the original sample randomly selected for validation. We ensured that the distribution of sampling weights was similar in both model development and validation samples. We applied the final risk score to participants from the validation sample; the outcome was compared to the true outcome to generate a receiver operating characteristic curve, and the AUC from the validation data set was compared to the AUC from the development data set. We also compared it to similar risk scores previously developed and validated in Hispanic populations in Peru (14) and Brazil (15), and two risk scores developed for the U.S. general population. The Centers for Disease Control and Prevention (CDC) prediabetes screening test and the American Diabetes Association (ADA) Type 2 diabetes risk test have been validated in the U.S. general population (16).

The choice of these scores was based on a systematic review aimed to identify risk scores applicable to Hispanic populations (4).

The point scoring system was based on the logistic regression coefficients of the predictors in the model using methodology by Sullivan in the Framingham Heart Study (17). Briefly, age was categorized using the age group 18-39 years as a referent risk factor profile. We computed the difference from the base in regression units for each category of every variable setting a constant coefficient corresponding to one point set for 5 years increase in age. Points were assigned such that a higher point total conveyed more risk. The cut-off point for the developed risk score was selected using the Youden index that maximizes sensitivity and specificity. All analyses were conducted in SAS V9.4.

RESULTS

The characteristics of the participants in the development sample (n=7,765) are detailed in Table 1. The overall age-adjusted prevalence of undiagnosed diabetes was 7.9% (661/7,765 unweighted). Higher prevalence was observed among those of Mexican or mixed background, participants 50 years and older, those not born in the U.S., with education of high school or less, not meeting exercise guidelines, having a family history of diabetes and with highest body mass index (BMI) status.

Model development

The final model included gender, age, education, family history of diabetes, having had gestational diabetes, BMI, not meeting with physical activity guidelines, having Mexican background and being born outside the U.S. Points assigned to each risk factor are shown in Table 2. The model performance for undiagnosed diabetes in the development sample measured by the AUC was 77% (95% CI: 75.8-79.1). After assigning points

to create the score, the AUC was comparable, 76% (95% CI: 74.4-77.8). Figure 1 illustrates the comparison between the ROC curves of the original model and the derived risk score in the development sample. The derived risk score had a maximum value of 30 points. The optimal cutoff point to identify undiagnosed diabetes was 11 according to the Youden index (Figure 1). Increasing age and BMI were associated with undiagnosed diabetes. Having a history of gestational diabetes contributed 6 points and not being U.S. born contributed 2 points to the score.

Model validation

In the validation sample (n=4,688), there were 378 individuals with undiagnosed diabetes. Table 3 shows the performance of the HCHS/SOL score compared to other scores in the validation sample. In terms of the ability to discriminate cases from non-cases measured by the AUC, the HCHS/SOL was superior to all other scores with an AUC of 77% followed closely by the ADA score at 75%. The AUC is the only general index of the accuracy of a screening measure that is independent of the cutoff point selected. At the suggested cutoff points, the sensitivity and specificity of the scores varied broadly (37-84%). The Peruvian risk score had the highest sensitivity (84%) but the lowest specificity (46%).

The HCHS/SOL score correctly identified 81% or 307 of the 378 undiagnosed cases and 60% of those without the disease. Moreover, 70% among the false positives identified by the HCHS/SOL risk score (1,186 of 1,716) had impaired glucose tolerance based on the American Diabetes Association definition. In comparison, the ADA risk score identified 74% of the 378 undiagnosed cases; 66% of those without diabetes and

71% among the ADA socre's false positives (1,043 of 1,473) had impaired glucose tolerance.

The performance of the risk scores by ethnic backgrounds yielded similar results. The AUCs were higher for the HCHS/SOL and ADA risk scores (Table 4). The Peruvian risk score had the lowest AUC (65%) in the South American subgroup. By gender, the performance of the HCHS/SOL and ADA scores was slightly better for females than for males (HCHS/SOL 78% females, 75% males and ADA 77% females, 73% males). Overall, the performance of the risk scores by subgroups was similar to that in the total validation sample. However, by age groups the HCHS/SOL risk score had an AUC of 76% among those 40 to 49 years of age, 70% among those younger than 40 and 64% for those 50 or above.

The number of variables used by the risk scores varied between three and nine. The Peruvian and Brazilian risk scores used only three variables however, implementation of either risk score may not be straightforward in the general population. The Peruvian risk score requires waist measurements which are seldom accurate without trained personnel and the Brazilian risk score defined hypertension as mean systolic blood pressure (SBP) above 140 mmHg or diastolic blood pressure (DBP) above 90 mmHg, or use of antihypertensive drugs. The HCHS/SOL risk score uses nine variables and like the CDC and ADA scores can be implemented using a self-assessment questionnaire.

CONCLUSIONS

The aim of this study was to develop and validate a diabetes risk score using data from Hispanics in the U.S. that can be applied as an initial screening tool to identify indi-

viduals at high risk of diabetes. The HCHS/SOL risk score identified 81% of the individuals with undiagnosed diabetes. Furthermore, 70% among the false positives had prediabetes highlighting the potential of the risk score to identify individuals at high risk of developing diabetes who would benefit from prevention programs. The new risk score is based on nine risk factors that can be collected as a self-assessment questionnaire making it an alternative to existing tools for the prediction of diabetes. Moreover, because this risk score was developed using data from Hispanics of diverse backgrounds living in the U.S., it may be tested in other Hispanic populations.

Risk factors included in this score have been widely used in other scores with the exception of ethnic background and being born in the U.S. Waist circumference was not significant in the final HCHS/SOL model. There is evidence that suggests waist circumference may be a better predictor than BMI for diabetes (18; 19). However, the International Diabetes Federation (IDF) recommendations do not include measurements for Hispanic populations due to lack of specific data (20). The HCHS/SOL model development tested waist circumference as a continuous and categorical variable following definitions from the IDF and the one used by the Peru risk score. The predictive value of waist circumference should be tested in other Hispanic populations. Hypertension and high cholesterol were tested as self-report only and combined with clinical measures. Although hypertension was significant when using the clinical measurements, we decided that to keep only self-reported measures that are easily implemented in a non-clinical setting.

The HCHS/SOL risk score is a feasible approach for diabetes screening. In populations where the prevalence of the disease is low, the performance of the screening

test may be enhanced by targeting individuals with a higher probability of having the disease. The use of risk scores combined with other demographic characteristics will enhance the identification of individuals with the disease. Application of the score in other population will help calibrate score ranges to help identify those at highest risk. Studies have shown that individuals that are labeled high risk by risk scores tend to have a higher risk of cardiovascular diseases and mortality making it worthwhile to intervene in this group (21-23).

This study is unique in that it used data from Hispanics living in the U.S. to develop and validate a risk score for undiagnosed diabetes. Although the sample includes Hispanics of different backgrounds across the country, the sample was based on four metropolitan areas which may not represent Hispanics living in rural areas who are more difficult to reach. However, data from the HCHS/SOL allowed to test the significance of Hispanic heritage background as a risk factor of undiagnosed diabetes. More research is needed to evaluate existing scores and adapting them to local settings and populations for Hispanics in the U.S. and to investigate the impact and cost-effectiveness of using risk scores as a public health tool and in clinical practice towards prevention of diabetes and its complications. In comparing the performance of the SOL risk score to the ADA and CDC risk scores, this study validates the effectiveness of the ADA and CDC scores for identifying diabetes in a sample of Hispanics living in the U.S.

As the burden of diabetes increases, risk scores offer an alternative for earlier identification of individuals at high risk of diabetes and its complications. This study used data from Hispanics in the U.S. to develop and validate a risk score for undiagnosed diabetes. The resulting score includes nine variables that can be obtained using a simple

questionnaire. The performance of the risk score was comparable to that of the ADA risk score. Application of the score in other populations will help refine the score and test its performance compared to other available scores. Diabetes risk scores allow for secondary prevention, such as early treatment of hypertension or high lipid profile that will prevent future complications. They can also help identify those at high risk of prediabetes providing a chance to route these individuals to prevention programs.

Limitations of this study include the use of a cross-sectional sample. Although the sample represents the most recent and comprehensive data available for Hispanics in the U.S., it may not be fully representative of all Hispanics living in the U.S. The use of clinical measurements at a single point in time may overestimate the prevalence of diabetes. The ADA recommends repeating positive tests to discard false positives, so having only one measurement may include false positives in the prevalence. In addition, the data for validation came from the same study. More studies are needed to study the applicability and validity of this score in practice. Strengths of the study include recent data, availability of all variables to compute the risk scores considered and adequate sample size to develop and validate the new risk score.

REFERENCES

1. Abid A, Ahmad S, Waheed A: Screening for Type II Diabetes Mellitus in the United States: The Present and the Future. Clinical Medicine Insights Endocrinology and Diabetes 2016;9:19-22

2. Casagrande SS, Cowie CC, Genuth SM: Self-reported prevalence of diabetes screening in the U.S., 2005-2010. Am J Prev Med 2014;47:780-787

3. Menke A, Casagrande S, Geiss L, Cowie CC: Prevalence of and Trends in Diabetes Among Adults in the United States, 1988-2012. JAMA 2015;314:1021-1029

4. Juarez LD, Gonzalez JS, Agne AA, Kulczycki A, Pavela G, Carson AP, Shelley JP, Cherrington AL: Diabetes risk scores for Hispanics living in the United States: A systematic review. Diabetes Res Clin Pract 2018;142:120-129

5. Kandula NR, Diez-Roux AV, Chan C, Daviglus ML, Jackson SA, Ni H, Schreiner PJ: Association of Acculturation Levels and Prevalence of Diabetes in the Multi-Ethnic Study of Atherosclerosis (MESA). Diabetes Care 2008;31:1621-1628

6. Pérez-Escamilla R, Putnik P: The Role of Acculturation in Nutrition, Lifestyle, and Incidence of Type 2 Diabetes among Latinos. The Journal of Nutrition 2007;137:860-870

7. Schneiderman N, Llabre M, Cowie CC, Barnhart J, Carnethon M, Gallo LC, Giachello AL, Heiss G, Kaplan RC, LaVange LM, Teng Y, Villa-Caballero L, Aviles-Santa ML: Prevalence of diabetes among Hispanics/Latinos from diverse backgrounds: the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). Diabetes Care 2014;37:2233-2239

8. Sorlie PD, Avilés-Santa LM, Wassertheil-Smoller S, Kaplan RC, Daviglus ML, Giachello AL, Schneiderman N, Raij L, Talavera G, Allison M, Lavange L, Chambless

LE, Heiss G: Design and implementation of the Hispanic Community Health Study/Study of Latinos. Ann Epidemiol 2010;20:629-641

9. Daviglus ML, Talavera GA, Aviles-Santa ML, Allison M, Cai J, Criqui MH, Gellman M, Giachello AL, Gouskova N, Kaplan RC, LaVange L, Penedo F, Perreira K, Pirzada A, Schneiderman N, Wassertheil-Smoller S, Sorlie PD, Stamler J: Prevalence of major cardiovascular risk factors and cardiovascular diseases among Hispanic/Latino individuals of diverse backgrounds in the United States. JAMA 2012;308:1775-1784

10. de Koning L, Chiuve SE, Fung TT, Willett WC, Rimm EB, Hu FB: Diet-quality scores and the risk of type 2 diabetes in men. Diabetes Care 2011;34:1150-1156

11. Jannasch F, Kroger J, Schulze MB: Dietary Patterns and Type 2 Diabetes: A Systematic Literature Review and Meta-Analysis of Prospective Studies. J Nutr 2017;147:1174-1182

12. Ghaderpanahi M, Fakhrzadeh H, Sharifi F, Badamchizade Z, Mirarefin M, Ebrahim RP, Ghotbi S, Nouri M, Larijani B: Association of physical activity with risk of type 2 diabetes. Iran J Public Health 2011;40:86-93

13. Chernick MR, LaBudde RA: *An introduction to bootstrap methods with applications to R*. John Wiley & Sons, 2014

14. Bernabe-Ortiz A, Smeeth L, Gilman RH, Sanchez-Abanto JR, Checkley W, Miranda JJ, Study Group CC: Development and Validation of a Simple Risk Score for Undiagnosed Type 2 Diabetes in a Resource-Constrained Setting. Journal of diabetes research 2016;2016:8790235

15. Pires de Sousa AG, Pereira AC, Marquezine GF, Marques do Nascimento-Neto R, Freitas SN, Nicolato RLdC, Machado-Coelho GL, Rodrigues SL, Mill JG, Krieger JE:

Derivation and external validation of a simple prediction model for the diagnosis of type 2 diabetes mellitus in the Brazilian urban population. Eur J Epidemiol 2009;24

16. Poltavskiy E, Kim DJ, Bang H: Comparison of screening scores for diabetes and prediabetes. Diabetes Res Clin Pract 2016;118:146-153

17. Sullivan LM, Massaro JM, D'Agostino RB: Presentation of multivariate data for clinical use: the Framingham Study risk score functions. Stat Med 2004;23:1631-1660

18. Chan DC, Watts GF, Barrett PH, Burke V: Waist circumference, waist-to-hip ratio and body mass index as predictors of adipose tissue compartments in men. QJM 2003;96:441-447

19. Langenberg C, Sharp SJ, Schulze MB, Rolandsson O, Overvad K, Forouhi NG, Spranger J, Drogan D, Huerta JM, Arriola L, de Lauzon-Guillan B, Tormo MJ, Ardanaz E, Balkau B, Beulens JW, Boeing H, Bueno-de-Mesquita HB, Clavel-Chapelon F, Crowe FL, Franks PW, Gonzalez CA, Grioni S, Halkjaer J, Hallmans G, Kaaks R, Kerrison ND, Key TJ, Khaw KT, Mattiello A, Nilsson P, Norat T, Palla L, Palli D, Panico S, Quiros JR, Romaguera D, Romieu I, Sacerdote C, Sanchez MJ, Slimani N, Sluijs I, Spijkerman AM, Teucher B, Tjonneland A, Tumino R, van der AD, van der Schouw YT, Feskens EJ, Riboli E, Wareham NJ: Long-term risk of incident type 2 diabetes and measures of overall and regional obesity: the EPIC-InterAct case-cohort study. PLoS Med 2012;9:e1001230

20. Alberti KG, Zimmet P, Shaw J: The metabolic syndrome--a new worldwide definition. Lancet 2005;366:1059-1062

21. Janssen PG, Gorter KJ, Stolk RP, Akarsubasi M, Rutten GE: Three years follow-up of screen-detected diabetic and non-diabetic subjects: who is better off? The ADDITION

Netherlands study. BMC Fam Pract 2008;9:67

22. Sandbaek A, Griffin SJ, Rutten G, Davies M, Stolk R, Khunti K, Borch-Johnsen K, Wareham NJ, Lauritzen T: Stepwise screening for diabetes identifies people with high but modifiable coronary heart disease risk. The ADDITION study. Diabetologia 2008;51:1127-1134

23. Spijkerman A, Griffin S, Dekker J, Nijpels G, Wareham NJ: What is the risk of mortality for people who are screen positive in a diabetes screening programme but who do not have diabetes on biochemical testing? Diabetes screening programmes from a public health perspective. J Med Screen 2002;9:187-190

Table 1. Prevalence of undiagnosed diabetes by participant characteristics, development sample

	Men and women			Women			Men		
		Undiagnosed			Undiagnosed			Undiagnosed	
	n	Prevalence (95% CI)		n	Prevalence (95% CI)		n	Prevalence (95% CI)	
Overall	7765	7.85	(6.99, 8.71)	4603	8.08	(6.92, 9.24)	3162	7.56	(6.32, 8.80)
Hispanic background									
Dominican	689	7.84	(5.24, 10.44)	461	6.79	(3.61, 9.96)	228	7.70	(4.12, 11.28)
Central American	803	7.51	(4.73, 10.28)	486	8.74	(5.09, 12.38)	317	5.19	(2.81, 7.57)
Cuban	1071	7.22	(5.70, 8.75)	549	7.66	(5.39, 9.93)	522	6.60	(4.29, 8.90)
Mexican	3183	8.51	(6.90, 10.13)	1961	8.90	(6.84, 10.96)	1222	7.76	(5.92, 9.61)
Puerto Rican	1190	7.39	(5.61, 9.16)	680	7.68	(5.49, 9.88)	510	7.33	(4.58, 10.07)
South American	571	6.56	(4.17, 8.94)	329	7.33	(4.18, 10.48)	242	4.44	(1.44, 7.43)
Mixed/other	258	9.72	(4.81, 14.64)	137			121	13.81	(5.93, 21.69)
Age group, years									
18–29	1373	0.95	(0.46, 1.43)	764	0.53	(0.08, 0.98)	609	1.39	(0.50, 2.27)
30–39	1271	2.96	(1.78, 4.14)	734	3.39	(1.55, 5.23)	537	2.54	(1.03, 4.05)
40–49	2158	7.47	(5.94, 9.00)	1291	7.76	(5.62, 9.90)	867	7.18	(5.16, 9.20)
50–59	1919	10.47	(8.86, 12.09)	1191	11.69	(9.44, 13.93)	728	8.98	(6.75, 11.21)
60–69	893	17.85	(14.06, 21.64)	532	16.89	(11.85, 21.93)	361	18.97	(13.68, 24.25)
70–74	151	25.92	(16.06, 35.79)	91	27.87	(14.52, 41.21)	60	23.15	(9.02, 37.29)
US born	1454	3.60	(2.39, 4.81)	827	2.83	(1.57, 4.10)	627	4.22	(2.22, 6.22)
Non US born	6311	8.15	(7.24, 9.07)	3776	8.43	(7.19, 9.67)	2535	7.79	(6.49, 9.09)
Education up to HS	4738	8.90	(7.79, 10.00)	2749	9.79	(8.32, 11.25)	1989	7.78	(6.15, 9.42)
Education above HS	3027	6.24	(5.01, 7.47)	1854	5.43	(3.96, 6.90)	1173	7.19	(5.28, 9.10)
No GPAQ activity guidelines	2566	9.96	(8.49, 11.43)	1847	10.16	(8.36, 11.96)	719	9.77	(7.30, 12.24)
GPAQ activity guidelines	5199	6.58	(5.61, 7.55)	2756	6.40	(5.03, 7.76)	2443	6.82	(5.45, 8.19)
Family history of diabetes	3275	9.39	(7.97, 10.80)	2085	9.69	(7.81, 11.57)	1190	8.81	(6.71, 10.92)
No family history of diabetes	4490	6.74	(5.75, 7.72)	2518	6.73	(5.35, 8.11)	1972	6.78	(5.32, 8.23)
Gestational diabetes				177	26.81	(21.57, 32.04)			
Body Mass Index (BMI)									
BMI < 25	1630	3.44	(2.25, 4.64)	982	3.63	(1.75, 5.52)	648	3.38	(1.70, 5.07)
Overweight (25.0-29.9))	3021	7.02	(5.76, 8.28)	1639	7.00	(5.17, 8.82)	1382	7.02	(5.26, 8.78)
Obese I (30.0-34.9)	1939	9.23	(7.56, 10.90)	1147	10.28	(7.98, 12.58)	792	8.10	(5.88, 10.31)
Obese II (35+)	1175	13.42	(10.67, 16.17)	835	11.65	(8.64, 14.66)	340	18.20	(13.13, 23.27)

Values except for sample size are weighted for study design and nonresponse, and age is standardized to Census 2010 U.S. population.

Table 2. Logistic regression estimates for risk factors associated with undiagnosed diabetes and score points

Variables	β	OR	95% CI	p-value	Points
Age (continuous)	0.068	1.07	(1.06, 1.08)	<.0001	
18-39*					0
40-49					3
50-59					5
60-69					7
70-74					9
Gender (being male)	0.401	1.49	(1.15, 1.94)	0.0028	1
Education: Up to HS	0.315	1.37	(1.06, 1.78)	0.0176	1
Family Hx	0.363	1.44	(1.14, 1.82)	0.0024	1
BMI (continuous)	0.081	1.08	(1.07, 1.11)	<.0001	
14-25*					0
25-30					2
30-40					4
40-70					9
Gest. Diab.	2.166	8.72	(5.13, 14.81)	<.0001	6
Not meeting PA guidelines	0.424	1.53	(1.20, 1.95)	0.0006	1
Mexican	0.319	1.38	(1.08, 1.75)	0.0088	1
Not US born	0.595	1.81	(1.24, 2.64)	0.002	2

c-statistic = 0.761, cutoff point=11 (sensitiviy=78%, specificity 60%)

Table 3. Performance of different diabetes risk scores compared to HCHS/SOL risk score using validation sample from HCHS/SOL.

Risk Score	# of variables	AUC	Sensitivity	Specificity	PPV	NPV	Accuracy
HCHS/SOL (≥11)	9	76.7%	81.2%	60.2%	15.2%	97.3%	61.9%
Peru (≥2)	3	70.4%	84.4%	46.4%	12.1%	97.1%	49.4%
Brazil (≥18)	3	71.5%	73.5%	52.5%	12.0%	95.8%	54.2%
ADA (≥5)	7	75.0%	73.5%	65.8%	15.9%	96.6%	66.4%
CDC (≥10)	6	68.7%	37.1%	79.9%	14.0%	93.6%	76.4%

Cutoff points are in parentheses.

Table 4. Performance of HCHS/SOL risk score by background of participant using validation sample

Background	AUC	Sensitivity	Specificity	PPV	NPV	Accuracy
Dominican	75.4%	77.1%	65.1%	17.1%	96.8%	66.1%
Central American	80.1%	79.5%	67.9%	18.7%	97.3%	68.9%
Cuban	71.8%	78.4%	57.8%	13.5%	97.0%	59.4%
Mexican	76.7%	82.6%	55.8%	14.8%	97.2%	58.0%
Puerto Rican	77.4%	83.3%	59.6%	15.3%	97.6%	61.5%
South American	75.5%	77.8%	66.8%	12.0%	98.1%	67.4%
Other	87.1%	88.9%	77.7%	21.6%	99.0%	78.4%

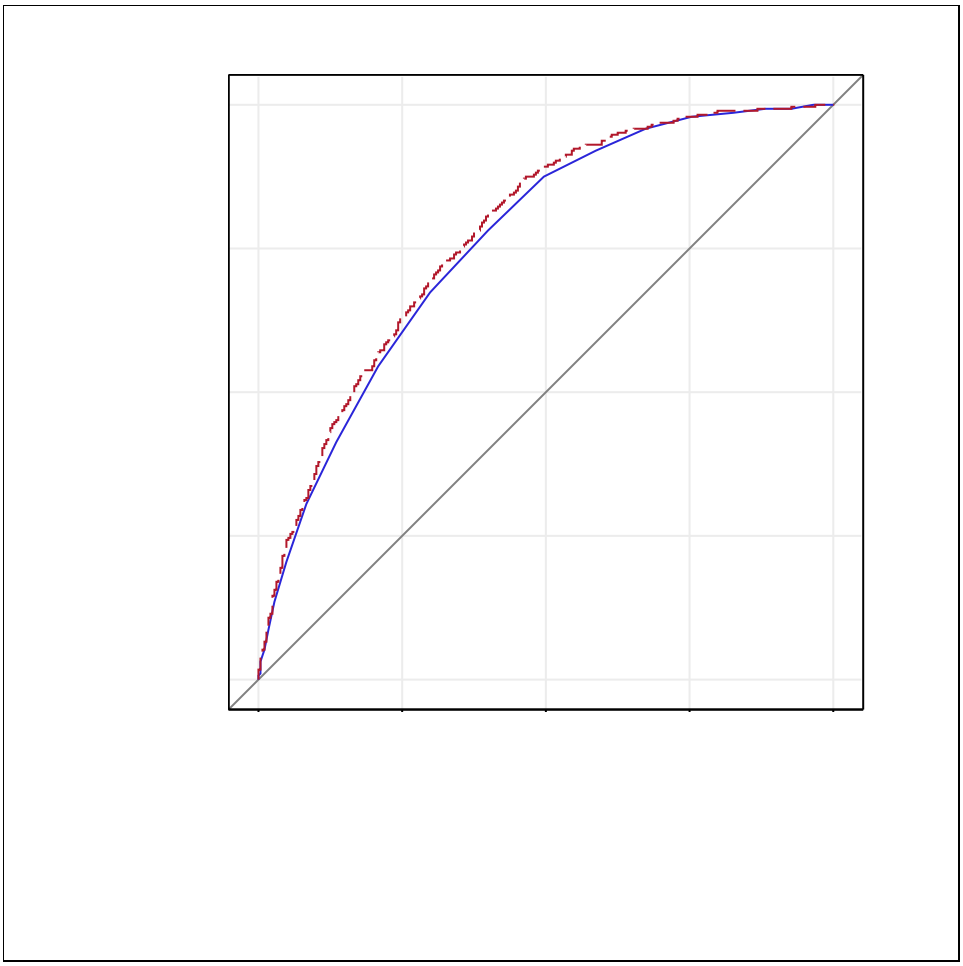

Figure 1. Area under the curve for risk factors logistic regression and HCHS/SOL risk score.

Supplemental table. Risk factors considered in each risk score

Risk factor	Peru	Brazil	ADA	CDC	SOL
Demographic					
Age	●	●	●	●	●
Gender	❍	❍	●	❍	●
Ethnicity		❍	❍	❍	●
Education	❍	❍	❍	❍	●
Born in U.S.	❍	❍	❍	❍	●
Behavioral	❍	❍	❍	❍	❍
Sendentarism/PA	❍	❍	●	●	●
Smoking	❍	❍	❍	❍	❍
Alcohol	❍	❍	❍	❍	❍
General Health Assessment		❍	❍	❍	❍
History of diabetes	❍	❍	❍	❍	❍
Family History Diabetes	●	❍	●	●	●
Gestational diabetes	❍	❍	●	●	●
Anthropometric	❍	❍	❍	❍	❍
BMI	❍	●	●	●	●
Waist circumference	●	❍	❍	❍	❍
Hypertension	❍	●	●	❍	❍
High cholesterol	❍	❍	❍	❍	❍

❍ Considered

● Included

EVALUATING THE PERFORMANCE OF THE HCHS/SOL DIABETES RISK SCORE FOR IDENTIYING DYSGLYCEMIA AMONG HISPANICS IN THE U.S.: THE HISPANIC COMMUNITY HEALTH STUDY/STUDY OF LATINOS (HCHS/SOL)

by

LUCIA D. JUAREZ; ANDRES AZUERO; JEFFREY S. GONZALEZ; GREGORY A. TALAVERA; ANDREA L. CHERRINGTON; NEIL SCHNEIDERMAN

In preparation for *Diabetes Care*

Format adapted for dissertation

ABSTRACT

Objective: We evaluate the performance of the SOL risk score for undiagnosed diabetes as a screening tool to detect diabetes and prediabetes (collectively known as dysglycemia) comparing it to the American Diabetes Association risk score for the same purpose.

Research and Design Methods: Data from the baseline visit of the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). We computed the c-statistic or area under the receiver curve (AUC) for both scores and identified optimal cutoff values to identify dysglycemia using the Youden index. We compared test performance characteristics, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for different thresholds of diabetes risk for each risk score.

Results: The SOL and ADA risk scores demonstrated reasonable performance detecting dysglycemia with AUCs of 77% in this population. The average ADA and SOL risk scores for this were 4.0 for ADA and 9.1 for SOL. The optimal cutoff point for detecting dysglycemia in the SOL risk score was 9 and 4 for the ADA risk score. At these cutoff points, the performance of the scores was comparable. For SOL AUC, sensitivity and specificity were 77.7%, 74% and 70%; for ADA, 77.1%, 74% and 69% respectively.

Conclusions: The SOL risk score performance identifying dysglycemia was comparable to the performance of the ADA risk score, developed for the general population. Future research should evaluate the performance of screening tools in Hispanic populations. Tools developed specifically for Hispanic populations may be more effective among Hispanics than those developed for the general population.

Keywords: Undiagnosed diabetes, Type 2 Diabetes Mellitus, Prediabetes, Dysglycemia, Risk Scores, Prediction Models, Hispanic, Latino

INTRODUCTION

Untreated diabetes is a leading cause of adverse health complications, which may result in cardiovascular mortality and morbidity. Diabetes and prediabetes (collectively known as dysglycemia) affect nearly half of the U.S. adult population, with higher rates among Hispanics and other racial/ethnic minorities (1). Diabetes and prediabetes may be asymptomatic for years and their diagnosis is based on laboratory tests. Current national recommendations stipulate that, in the absence of clear symptoms, the diagnosis should be confirmed by repeated testing (2). However, the prevalence of clinical laboratory testing to diagnose diabetes or prediabetes is lower among Hispanics and members of other racial/ethnic minorities; and among individuals with low socioeconomic status or those without health insurance (3). Studies based on national data report that nearly half of Hispanics with diabetes are undiagnosed (4) and that one-third of Hispanics with diabetes are not aware of having it (5). This is problematic because timely behavioral interventions and clinical treatment can help manage diabetes and prediabetes reducing the risk of complications (6-8). Accordingly, identifying individuals at high risk could help prevent new-onset diabetes and reduce the burden of diabetes-related complications.

As the prevalence of diabetes increases, it is imperative to find screening strategies that are cost-effective, accurate and easy to implement. Risk scores can help identify those at risk of diabetes and provide a continuous spectrum of risk that can be used to classify individuals to decide appropriate mode, the urgency of care or type of intervention needed. Many diabetes and prediabetes risk scores have been developed in different populations (9-14). However, risk scores need to be validated in the population that they will be applied to as their performance depends on the prevalence of the disease and the

characteristics of individuals in each population (15). Only a handful of risk scores have been developed and validated in Hispanic populations (16).

The American Diabetes Association (ADA) offers a diabetes risk self-assessment questionnaire on-line. This risk score assigns points to seven individual characteristics: age, gender, history of gestational diabetes, family history of diabetes, high blood pressure, physical inactivity, and body mass index (BMI). A total score of five or above indicates an increased risk of prediabetes or diabetes (2). A similar risk score based on data from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) was developed to assess risk of undiagnosed diabetes among Hispanics living in the U.S. The SOL risk score is based on nine individual characteristics: gender, age, Mexican background, education, being born in the U.S., family history of diabetes, gestational diabetes and physical inactivity. A score of 11 or more points indicates a higher risk of having undiagnosed diabetes (Juarez et al, unpublished). The ADA risk score helps identify increased risk for dysglycemia. The SOL score was developed to identify undiagnosed diabetes. The performance of the SOL score was adequate with an AUC of 77%. The score correctly identified 81% of the participants who had undiagnosed diabetes. In addition, among the risk score's false positives, close to 70% had prediabetes (Juarez et al, unpublished).

This study examines the effectiveness of the SOL risk score for detecting dysglycemia and compares its performance to that of the ADA risk score developed for the general population. An ideal screening tool should accurately identify the highest number of individuals with dysglycemia (area under the receiving curve (AUC) of at least 75%) using the minimum number of screening tests (high sensitivity and specificity).

RESEARCH DESIGN AND METHODS

The HCHS/SOL (SOL) is a prospective, multicenter community-based study of Hispanic adults in the U.S. The original baseline cohort includes 16,415 men and women 18-74 years of age between 2008-2011 from four communities: Bronx, NY; Chicago, IL; Miami, FL; and San Diego, CA. The study design and protocol of the study have been published elsewhere (17). Study measures included 2-hour oral glucose tolerance test (OGTT), fasting plasma glucose (FPG) and Glycosylated hemoglobin (A1C). Participants were required to fast for at least 8 h prior to the visit, consuming only water and necessary medications. Venous blood specimens were collected, processed, and frozen on site toward the beginning of the visit and also 2 h after a 75 g glucose load. Plasma glucose was assessed using a hexokinase enzymatic method (Roche Diagnostics Corporation, Indianapolis, IN). Glycosylated hemoglobin (A1C) was measured in EDTA whole blood using a Tosoh G7 automated high-performance liquid chromatography analyzer (Tosoh Bioscience Inc., San Francisco, CA). Personal and family history of medical diagnoses, including diabetes, was obtained via a self-report interview administered by certified assessors (5).

The SOL risk score (Juarez et al, unpublished) was derived from a development sample (62% of the original SOL baseline sample) using logistic regression. The random split of 62%-38% was based on the probability of each unique observation being included in a resample of equal size to the original sample size. On average, a random sample would include 62% of the original observations while the rest, 38% would be missing from that sample (18). We used the validation pool (38% of the original SOL baseline sample) to compare the performance of the ADA risk score based on the self-assessment

questionnaire and the SOL risk score. This validation pool included participants who may have been diagnosed diabetic before baseline who were not included for the development of the SOL risk score. The final sample size was 5,431.

Data for the study were collected from standardized questionnaires. The information was self-reported and included variables such as participant's gender, age, Hispanic heritage background; previous diagnosis of diabetes, hypertension, gestational diabetes for women; family history of diabetes, education, nativity (born in one of the 50 U.S. states, District of Columbia or elsewhere), activity level and activity level for work and leisure time. The body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared from objective height and weight measures.

Dysglycemia was defined using ADA criteria (19) as either FPG ≥100 mg/dL, a 2-hour post load glucose level (2-h OGTT) ≥140 mg/dL, A1C level ≥5.7% or documented use of hypoglycemic agents (scanned medications). ADA and SOL risk scores were calculated for each participant following the ADA and SOL scoring points shown in Table 1. Participants with missing data in any of the characteristics required to compute either risk score were excluded from the analysis.

Summary statistics were used to describe demographic characteristics, diabetes risk factors, ADA and SOL risk scores and A1C for all participants according to glycemic status. Chi-squared and t-tests were used to examine the association between participants' characteristics and glycemic status.

We computed the c-statistic or area under the receiver curve (AUC) for both scores for the total population and by Hispanic heritage background. We analyzed the receiver operating characteristic (ROC) curve to identify the most appropriate cutoff values

to identify dysglycemia. The optimal cutoff point for each score was identified using the Youden index, which maximizes sensitivity and specificity. Table 3 shows test performance characteristics, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for different thresholds of diabetes risk for each risk score. The threshold values were within two points of the optimal cutoff value for each score, two to six for the ADA score and seven to eleven for the SOL risk score. All analyses were weighted to adjust for sampling probability and nonresponse to account for the complex sample design of the data. Statistical analyses were performed in SAS version 9.4 (SAS Institute).

RESULTS

Table 2 shows participants' characteristics, diabetes risk factors, average ADA and SOL scores and average A1C by glycemic status. In this population, over half (53%) had dysglycemia, the average age of participants was 41.7 years, the average BMI was 29.4 and eight out of ten participants were not born in the U.S. The prevalence of diabetes risk factors in this population was high. Four in ten participants had a family history of diabetes, eight in ten were overweight or obese and three in ten did not meet physical activity guidelines. The average ADA risk score for the whole population was 4 and 9.1 for the SOL risk score. The average A1C was 5.8%. These three measures were significantly higher among the participants who had dysglycemia compared to those with normal glycemic status. The diabetes risk factors considered in both scores were significantly higher among those with dysglycemia than those with normal glycemic status, with the exception of sex. Participants with dysglycemia were older, not born in the U.S., had lower education and had higher BMI.

The apparent accuracy of both models, measured by the AUC, was adequate and comparable (77.1% for ADA and 77.7% for SOL). The performance of both scores was similar for all Hispanic backgrounds (Figure 1). Based on the Youden index, the optimal cutoff points were four for the ADA score and nine for the SOL score. At the optimal cut-off points, the sensitivity, specificity, PPV, and NPV were 74%, 70%, 73% and 71% for the SOL score and 74%, 69%, 72% and 71% for the ADA score (Table 3). There were incremental decreases in sensitivity and NPV from two points below the optimal cutoff point to two points above it with incremental increases of sensitivity and PPV. Changes in performance measures were steeper in the ADA score because the score ranged from 0 to 10, while the SOL score ranged from 0 to 30. At the suggested cutoff point of five, for the general population, the sensitivity of the ADA score decreased from 74% to 56%, while the specificity increased from 69% to 82%.

CONCLUSIONS

This study demonstrates the effectiveness of the SOL and ADA risk scores to identify dysglycemia in a sample of Hispanics living in the U.S. The performance of the SOL risk score, originally developed to detect undiagnosed diabetes among Hispanics living in the U.S., had an AUC of 77.7% detecting dysglycemia and sensitivity of 74% at the cutoff point of nine. The performance of the SOL risk score was comparable to that of the ADA score with a cutoff point of four for the same outcome. The ADA risk score was originally developed to detect undiagnosed diabetes and later shown to be effective to detect dysglycemia at a cutoff point of four for the entire U.S. population (20). All performance measures for both risk scores were comparable to those obtained in nationally representative samples for the general population (21). We believe that the performance of

both risk scores was comparable due to the similarity of categories used of age and BMI. The wider range of scores derived from the SOL risk score provides more options to group individuals in different risk categories. It remains to investigate whether the items specific to Hispanic populations may contribute to a more effective tool to identify dysglycemia among Hispanics in other settings.

Participants with dysglycemia in this sample had significantly higher ADA and SOL risk scores than those with normal glucose levels. All risk factors were significantly higher among those with dysglycemia, with the exception of sex. More than half (53%) of the individuals in these populations had dysglycemia. Their average ADA and SOL scores were 4 and 9.1, respectively. Both average scores coincide with the cutoff values to detect dysglycemia. Clinical trials have shown that individuals at high risk of diabetes may cut their risk by more than half when intervened to follow intensive life-style modification programs (7; 22; 23). The ADA and SOL risk scores can aid health practitioners decide which individuals should undergo further testing. Early diagnosis is essential to reduce the burden of diabetes and its complications.

Screening recommendations for blood glucose testing are not widely followed in the U.S. resulting in about 30% of people with the disease being undiagnosed (1). Despite current recommendations to screen Hispanics at earlier ages or lower BMIs (24), widespread blood testing for diabetes are not the norm and may not be the most efficient way to identify individuals at high risk of diabetes in large communities with limited resources.

Our goal was to show that the SOL risk score can be used to detect individuals at high risk of prediabetes and undiagnosed diabetes using different cutoff values, 9 for prediabetes and 11 for undiagnosed diabetes. The SOL risk score can be used in a great variety of community and clinical settings via a simple pencil-and-paper questionnaire, a clinical encounter, or via the internet. The instrument has good feasibility properties, including nine items that can be answered in minimal time providing a score easy to calculate. The risk score provides a continuous spectrum of risk that can be used to create categories that reflect the urgency and availability of resources in specific settings. The SOL screening score also has great educational potential, as it will highlight important risk factors that need to be modified in high-risk individuals. As a recruitment tool, the SOL risk score can be used when targeting high-risk populations with limited resources. The score may be used to stratify the population to help prioritize those at highest risk, increasing specificity to save on potential tests.

Some limitations to the use of this risk score are that some of the variables in the score may not be available for all individuals in all settings. However, the majority of the items will be available. In addition, increasing use of risk scores may encourage clinical settings to collect the data needed. Further validation in other settings and populations is still needed to ascertain the effectiveness of the SOL risk score. Although data from the HCHS/SOL study provides the most up to date and comprehensive data on Hispanics living in the U.S., it only included data from four communities which may not represent Hispanics living in rural areas or more isolated settings. Despite diabetes being more

common and the general population being more aware of the risks associated with diabetes and its complications, more education is needed, especially in immigrant populations that face numerous challenges.

We used a combination of blood tests to identify dysglycemia in this study. Research has shown that different tests and definitions result in an overlapping majority of cases identified, however, each test may identify a small proportion of cases that other tests do not (25-27). By using all tests available in this population, we may be identifying a larger pool of participants which will impact estimates of sensitivity, specificity, PPV, and NPV.

In conclusion, the SOL risk score performed well and comparable to the ADA score. Its performance was robust and may be used for undiagnosed diabetes and dysglycemia. Using this score may help with active identification of cases with prediabetes, which deserves attention from healthcare providers. Identifying individuals with prediabetes gives health providers the chance to delay or even reverse diabetes. Self-assessments help individuals decide whether they should seek medical care and may lead to more informed interactions with their health providers. The risk score could also be applied to health care databases to identify high-risk cases and verify that a blood test and adequate care have been prescribed.

REFERENCES

1. Control CfD, Prevention: National diabetes statistics report, 2017. Atlanta, GA: Centers for Disease Control and Prevention 2017;

2. American Diabetes Association: 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2018. Diabetes Care 2018;41:S13-s27

3. Casagrande SS, Cowie CC, Genuth SM: Self-reported prevalence of diabetes screening in the U.S., 2005-2010. Am J Prev Med 2014;47:780-787

4. Menke A, Casagrande S, Geiss L, Cowie CC: Prevalence of and Trends in Diabetes Among Adults in the United States, 1988-2012. JAMA 2015;314:1021-1029

5. Schneiderman N, Llabre M, Cowie CC, Barnhart J, Carnethon M, Gallo LC, Giachello AL, Heiss G, Kaplan RC, LaVange LM, Teng Y, Villa-Caballero L, Aviles-Santa ML: Prevalence of diabetes among Hispanics/Latinos from diverse backgrounds: the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). Diabetes Care 2014;37:2233-2239

6. Gerstein HC, Yusuf S, Bosch J, Pogue J, Sheridan P, Dinccag N, Hanefeld M, Hoogwerf B, Laakso M, Mohan V, Shaw J, Zinman B, Holman RR: Effect of rosiglitazone on the frequency of diabetes in patients with impaired glucose tolerance or impaired fasting glucose: a randomised controlled trial. Lancet 2006;368:1096-1105

7. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Nathan DM: Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002;346:393-403

8. Li G, Zhang P, Wang J, Gregg EW, Yang W, Gong Q, Li H, Li H, Jiang Y, An Y, Shuai Y, Zhang B, Zhang J, Thompson TJ, Gerzoff RB, Roglic G, Hu Y, Bennett PH: The long-

term effect of lifestyle interventions to prevent diabetes in the China Da Qing Diabetes Prevention Study: a 20-year follow-up study. Lancet 2008;371:1783-1789

9. Abbasi A, Peelen LM, Corpeleijn E, van der Schouw YT, Stolk RP, Spijkerman AM, Moons KG, Navis G, Bakker SJ, Beulens JW: Prediction models for risk of developing type 2 diabetes: systematic literature search and independent external validation study. 2012;

10. Brown N, Critchley J, Bogowicz P, Mayige M, Unwin N: Risk scores based on self-reported or available clinical data to detect undiagnosed type 2 diabetes: a systematic review. Diabetes Res Clin Pract 2012;98:369-385

11. Buijsse B, Simmons RK, Griffin SJ, Schulze MB: Risk assessment tools for identifying individuals at risk of developing type 2 diabetes. Epidemiol Rev 2011;33:46-62

12. Schwarz PE, Li J, Lindstrom J, Tuomilehto J: Tools for predicting the risk of type 2 diabetes in daily practice. Horm Metab Res 2009;41:86-97

13. Thoopputra T, Newby D, Schneider J, Li SC: Survey of diabetes risk assessment tools: concepts, structure and performance. Diabetes Metab Res Rev 2012;28:485-498

14. Witte DR, Shipley MJ, Marmot MG, Brunner EJ: Performance of existing risk scores in screening for undiagnosed diabetes: an external validation study. Diabet Med 2010;27:46-53

15. Royston P, Moons KG, Altman DG, Vergouwe Y: Prognosis and prognostic research: Developing a prognostic model. BMJ 2009;338:b604

16. Juarez LD, Gonzalez JS, Agne AA, Kulczycki A, Pavela G, Carson AP, Shelley JP, Cherrington AL: Diabetes risk scores for Hispanics living in the United States: A systematic review. Diabetes Res Clin Pract 2018;142:120-129

17. Sorlie PD, Avilés-Santa LM, Wassertheil-Smoller S, Kaplan RC, Daviglus ML, Giachello AL, Schneiderman N, Raij L, Talavera G, Allison M, Lavange L, Chambless LE, Heiss G: Design and implementation of the Hispanic Community Health Study/Study of Latinos. Ann Epidemiol 2010;20:629-641

18. Chernick MR, LaBudde RA: *An introduction to bootstrap methods with applications to R*. John Wiley & Sons, 2014

19. American Diabetes Association: 2. Classification and diagnosis of diabetes: standards of medical care in diabetes—2018. Diabetes Care 2018;41:S13-S27

20. Bang H, Edwards AM, Bomback AS, Ballantyne CM, Brillon D, Callahan MA, Teutsch SM, Mushlin AI, Kern LM: Development and validation of a patient self-assessment score for diabetes risk. Ann Intern Med 2009;151:775-783

21. Poltavskiy E, Kim DJ, Bang H: Comparison of screening scores for diabetes and prediabetes. Diabetes Res Clin Pract 2016;118:146-153

22. Schmidt MI, Duncan BB, Bang H, Pankow JS, Ballantyne CM, Golden SH, Folsom AR, Chambless LE: Identifying individuals at high risk for diabetes: The Atherosclerosis Risk in Communities study. Diabetes Care 2005;28:2013-2018

23. Tuomilehto J, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, Ilanne-Parikka P, Keinanen-Kiukaanniemi S, Laakso M, Louheranta A, Rastas M, Salminen V, Uusitupa M: Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med 2001;344:1343-1350

24. Siu AL: Screening for Abnormal Blood Glucose and Type 2 Diabetes Mellitus: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med 2015;163:861-868

25. Aviles-Santa ML, Hsu LL, Arredondo M, Menke A, Werner E, Thyagarajan B, Heiss G, Teng Y, Schneiderman N, Giachello AL, Gallo LC, Talavera GA, Cowie CC: Differences in Hemoglobin A1c Between Hispanics/Latinos and Non-Hispanic Whites: An Analysis of the Hispanic Community Health Study/Study of Latinos and the 2007-2012 National Health and Nutrition Examination Survey. Diabetes Care 2016;39:1010-1017

26. Avilés-Santa ML, Pérez CM, Schneiderman N, Savage PJ, Kaplan RC, Teng Y, Suárez EL, Cai J, Giachello AL, Talavera GA: Detecting prediabetes among Hispanics/Latinos from diverse heritage groups: does the test matter? Findings from the Hispanic Community Health Study/Study of Latinos. Prev Med 2017;95:110-118

27. Cowie CC, Rust KF, Byrd-Holt DD, Gregg EW, Ford ES, Geiss LS, Bainbridge KE, Fradkin JE: Prevalence of diabetes and high risk for diabetes using A1C criteria in the U.S. population in 1988-2006. Diabetes Care 2010;33:562-568

Table 1. Point assignment for ADA and SOL risk scores

	Points	
Characteristic	**SOL**	**ADA**
Age		
18-39*	0	0
40-49	3	1
50-59	5	2
60-69	7	3
70-74	9	3
Gender (being male)	1	1
Education: High School or less	1	NA
Ever been diagnosed with hypertension	NA	1
Family history of diabetes	1	1
BMI		
14-25*	0	0
25-30	2	1
30-40	4	2
40-70	9	3
If female, history of gestational diabetes	6	1
Not physically active	1	1
Mexican	1	NA
Not US born	2	NA

Table 2. Study Characteristics (standard error) of participants with Dysglycemia, HCHS/SOL Visit 1

Characteristic	Overall	Normal	Dysglycemia	p-value
N	5431	2220	3211	
Female	51.4 (1.06)	53.0 (1.48)	49.9 (1.50)	0.1378
Hispanic background				0.0218
Dominican	8.4 (12.42)	12.3 (1.55)	8.7 (1.01)	
Central American	6.0 (8.43)	7.3 (0.92)	7.1 (0.59)	
Cuban	15.4 (22.12)	18.5 (1.94)	19.0 (1.87)	
Mexican	37.1 (44.20)	41.1 (2.21)	40.2 (2.07)	
Puerto Rican	12.5 (16.29)	12.4 (1.21)	16.2 (1.18)	
South American	4.0 (5.47)	5.2 (0.56)	4.3 (0.44)	
Mixed/other	2.9 (4.80)	3.2 (0.49)	4.4 (0.80)	
Age group, years				<.0001
18–39	46.5 (1.11)	66.7 (1.42)	28.3 (1.56)	
40–49	22.6 (0.78)	19.5 (1.12)	25.5 (1.19)	
50–59	17.3 (0.71)	9.8 (0.87)	24.1 (1.11)	
60–69	10.2 (0.65)	3.6 (0.47)	16.2 (1.12)	
70–74	3.3 (0.36)	0.4 (0.14)	5.9 (0.66)	
Non US born	79.3 (1.02)	73.7 (1.60)	84.3 (1.18)	<.0001
Education HS or less	59.9 (1.18)	55.7 (1.67)	63.7 (1.39)	<.0001
Does not meet GPAQ activity guidelines	34.2 (0.98)	29.7 (1.45)	38.3 (1.30)	<.0001
Family history of diabetes	40.7 (1.13)	32.7 (1.46)	48.1 (1.47)	<.0001
Gestational diabetes*	21.4 (3.03)	80.2 (7.78)	18.1 (3.05)	0.0009
History of hypertension	26.8 (0.97)	14.1 (1.21)	38.3 (1.38)	<.0001
BMI Status				<.0001
25-30	21.9 (0.86)	33.6 (1.47)	11.3 (0.81)	
30-40	38.0 (0.99)	36.6 (1.43)	39.3 (1.50)	
40-70	40.0 (1.04)	29.8 (1.60)	49.3 (1.48)	
ADA risk score MEAN (SD)	4.0 (0.05)	2.9 (0.07)	5.0 (0.06)	<.0001
SOL risk score MEAN (SD)	9.1 (0.09)	7.1 (0.10)	11.0 (0.10)	<.0001
A1c(%)	5.8 (0.02)	5.2 (0.01)	6.2 (0.04)	<.0001

*591 women who had been told they had diabetes were asked whether that happened during pregnancy only.

Table 3. Performance of ADA and SOL Risk Scores in detecting Dysglycemia Among HCHS/SOL participants at different cutoff values

	Sensitivity, %		Specificity, %		PPV, %		NPV, %	
	SOL	**ADA**	**SOL**	**ADA**	**SOL**	**ADA**	**SOL**	**ADA**
-2 points	88.4	96.4	47.5	24.0	65.0	58.3	78.8	85.6
-1 point	81.0	88.0	58.7	47.3	68.4	64.8	73.7	78.1
optimal cutoff*	73.9	74.0	70.1	68.8	73.2	72.4	70.9	70.6
+1 point	64.7	55.9	77.3	81.7	75.8	77.2	66.5	62.7
+2 points	56.5	40.2	82.9	89.7	78.5	81.2	63.3	57.6

AUC for total population: ADA score 77.1%, AUC for SOL score 77.7%

* Optimal cutoff for ADA score was 4 and for SOL was 9.

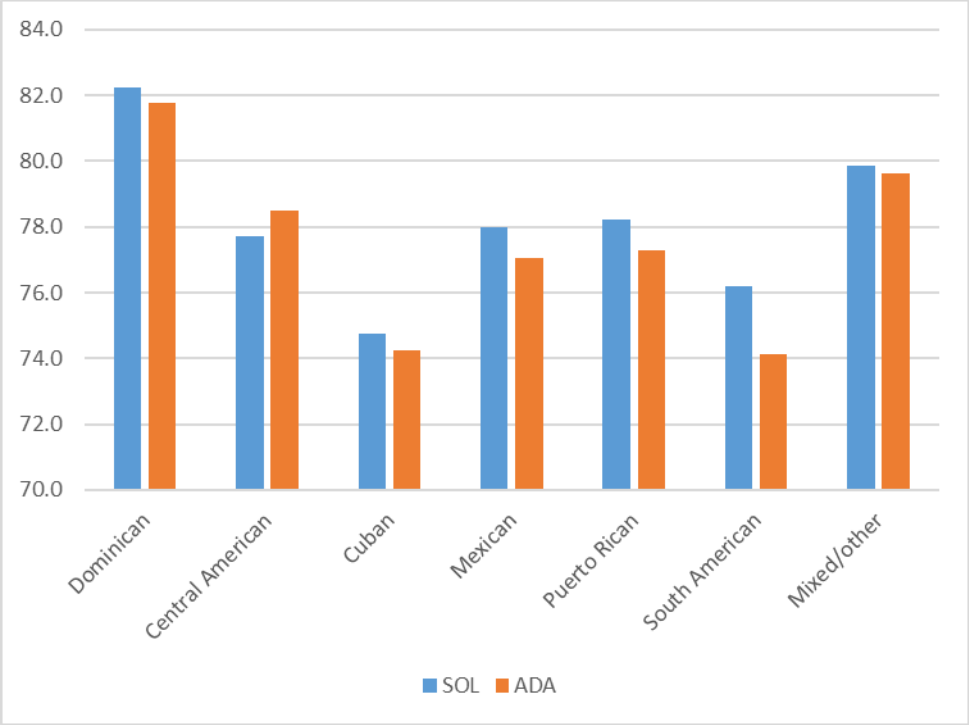

Figure 1. Area Under the Curve (AUC) by Hispanic background for ADA and SOL risk scores

CONCLUSIONS AND FUTURE DIRECTIONS

The three manuscripts comprised in this dissertation contribute to our understanding of diabetes among Hispanics living in the U.S. It is the first attempt to develop and validate a risk score for Hispanics in the U.S. who are at risk of undiagnosed diabetes. Hispanics are a heterogeneous group of peoples with diverse cultures, ancestries and sociodemographic characteristics. The prevalence of diabetes varies among Hispanic groups and nearly 40% of all Hispanics are undiagnosed. Finding simple screening tools that help categorize individuals at high risk of diabetes can facilitate routing those most in need to treatment and behavioral interventions. Risk scores can also be used as an educational tool to raise awareness of the individual's relative risk of having a disease and the risk factors associated with the disease. Combining the use of risk scores with the implementation of healthy eating and physical activity guidelines presents an opportunity to develop strategies that take individual's relative risk into account and has the potential to help route them to an appropriate intervention plan to help prevent the disease.

Multi-step screening approaches have been effective in diagnosing diabetes (Khunti et al., 2015). Cost-effectiveness in diagnosing diabetes increases when a non-invasive risk assessment tool is applied prior to a blood test (Health & Excellence, 2012; Khunti et al., 2012). A successful risk classification strategy is one that can stratify individuals into clinically relevant risk categories and whose performance is between 70% and 80% (Buijsse et al., 2011). The performance of the SOL risk score was 76% and the score was able to classify individuals at risk into prediabetes and diabetes with a performance of 77%, comparable to the ADA risk score.

As the prevalence of diabetes increases, the key to control and prevention of the disease will be to ensure that screening influences the course of diabetes and its complications. Screening for any disease is justified when early detection and treatment translate into a tacit benefit compared to its natural clinical presentation (Siu, 2015). For diabetes, this benefit is proven when high-risk individuals undergo intensive lifestyle interventions (Gillies et al., 2008; Health & Excellence, 2012). As the 2015 USPSTF national recommendations include being Hispanic as a risk factor in their clinical considerations, this would translate into universal screening for Hispanics up to age 70. However, universal screening may not be cost-effective, especially when there are ways to identify and target individuals at high-risk (Wareham & Griffin, 2001). Optimal screening strategies should be based on costs, availability of services, the effectiveness of the screening tools available and the type of diagnostic test. Evidence of the benefit of adequate management of individuals with complications and those at high-risk is increasing ("The Diabetes Prevention Program (DPP): description of lifestyle intervention," 2002; Gillies et al., 2008; Holman, Paul, Bethel, Matthews, & Neil, 2008; Khunti & Davies, 2012; Khunti et al., 2012; Lindström et al., 2006).

As more data on Hispanics become available, studies like the ones presented in this dissertation will improve our understanding of the burden of diabetes in Hispanic populations. The first manuscript in this dissertation provided important insights on the issues related to the scarcity of research focusing on diabetes risk prediction models for Hispanics. This scarcity of published research prevented the implementation of a meta-analysis that compared prediction models. Further, the definition of the term Hispanic

was inconsistent among studies. Studies that included data on Hispanics presented analyses that combined data from Hispanics and the general population, or combined data from Hispanics of diverse backgrounds into a single category. From the three studies identified in the systematic review, the study from the U.S. used data from 1976-1980 (Herman, Smith, Thompson, Engelgau, & Aubert, 1995). The other two studies were from Brazil (Pires de Sousa et al., 2009) and Peru (Bernabe-Ortiz et al., 2016). Hispanics living in the U.S. have unique characteristics that are influenced by migration and onward integration and challenges such as discrimination, language proficiency, and stress. Results from this systematic review and the increasing prevalence of diabetes among Hispanics motivated the development and validation of a diabetes risk score for Hispanics that was based on readily available information.

Building on the findings from the systematic review, the second manuscript used data from the baseline cohort of HCHS/SOL to develop and validate a risk score for undiagnosed diabetes. HCHS/SOL is a prospective study. It collects data from Hispanics living in the U.S. to describe the prevalence of chronic diseases and identify their risk factors. The cohort of participants included over 16,000 participants from diverse Hispanic backgrounds living in San Diego, Chicago, Miami and New York (Sorlie et al., 2010). The development of a risk score for Hispanics living in the U.S. is significant because the prevalence of diabetes, particularly undiagnosed diabetes is high and because this growing population already represents close to 20% of the total U.S. population. The development of the risk score included known diabetes risk factors used in other risk scores and additional factors collected by HCHS/SOL known to be associated with diabetes among Hispanics. These additional risk factors included Hispanic background, being born in the

U.S. and years living in the U.S. Potential risk factors were limited to those that a person could answer in a brief questionnaire and did not require clinical measures. For example, although several other risk scores included hypertension and high cholesterol clinical measurements, we opted for self-reported answers that would result in a simpler tool easy to implement in clinical and non-clinical settings.

The SOL risk score included nine risk factors: being of Mexican descent, male gender, older age, lower education, being born in the U.S., having a family history of diabetes, being overweight or obese, having had gestational diabetes and not complying with physical activity recommendations. The risk score's performance was superior to the performance of the other two risk scores developed for Hispanics (Bernabe-Ortiz et al., 2016; Pires de Sousa et al., 2009) and marginally better to that of the ADA self-assessment risk score. Results from this study showed that the risk score could also be used to identify prediabetes.

The third paper further investigated the performance of the SOL risk score in identifying Hispanics with diabetes and prediabetes (jointly known as dysglycemia). The performance of the SOL risk score to identify dysglycemia was comparable to that of the ADA risk score. On average, the risk scores increased with severity of the disease which allows investigators to stratify individuals and prioritize those who need urgent attention.

In summary, risk scores are a feasible alternative to identify and classify Hispanics at risk of diabetes and prediabetes. They provide information on the risk factors in the population under study and can be used as an educational tool to promote awareness of how these risk factors affect an individual's risk of diabetes in the Hispanic population.

Given the increasing prevalence of diabetes, preventive efforts need more intensive approaches that can be guided by risk stratification. Based on our findings, intensive behavioral programs can be customized for those at higher risk according to their different characteristics and risks factors. Ultimately, disease prevention requires an understanding of the population at risk within a comprehensive framework. Combining multidimensional and transdisciplinary strategies can result in viable alternatives that translate into clinical care to Hispanics at risk of diabetes.

Particularly noteworthy is that the development and validation of the diabetes risk score for Hispanics in the U.S. were conducted with limited resource settings in mind. The resulting score includes nine risk factors that can be easily obtained from individuals. The SOL risk score may be helpful in primary health care where diagnosis services may not be available. It can be used to identify individuals at high-risk who can be referred to further testing or educational programs. The study's main limitation is the use of one database for development and validation. The score was developed to identify undiagnosed diabetes and was proven to be effective in identifying dysglycemia. This is the first diabetes risk score for Hispanics living in the U.S. Although the performance of the score was comparable to that of the ADA score for the general population. Having a score for Hispanics has the potential to draw attention to diabetes screening in a population where prevalence remains high. The use of this score with a comprehensive approach that effectively identifies Hispanics at high-risk of diabetes routes them to testing, treatment or intensive lifestyle interventions has the potential to improve health outcomes in this population.

Limitations to manuscripts two and three include the use of baseline data and the need to split the sample to develop and validate the risk score. Future steps include validation of the risk score in subsequent waves of HCHS/SOL data and in other Hispanic populations and settings. This will be possible as more data become available in Latin America and the U.S. Additionally, although the performance of the risk score was adequate, it will need to be validated in other settings and populations to determine whether the same factors hold true.

The steady increase of diabetes rates among Hispanics indicates the pressing need to orchestrate preventive efforts that take into consideration the causes andcircumstances of the burden of disease specific to these populations. Multidisciplinary approaches that are culturally tailored can help create programs such as the DPP that translate into improved quality of life and better health outcomes. Many challenges such as limited human and financial resources and lack of public awareness need to be addressed. This dissertation adds to the understanding of diabetes among Hispanics in the U.S. The use of risk scores in clinical and research programs can help not only identifying individuals at risk but promoting communication between clinic and research staff and the individuals at risk. An open dialogue of cultural, socioeconomic and environmental circumstances between staff and individuals at risk will help design and implement health programs that are coherent and sustainable.

LIST OF GENERAL REFERENCES

Abid, A., Ahmad, S., & Waheed, A. (2016). Screening for Type II Diabetes Mellitus in the United States: The Present and the Future. *Clin Med Insights Endocrinol Diabetes, 9*, 19-22. doi:10.4137/cmed.s38247

Ali, M. K., Echouffo-Tcheugui, J., & Williamson, D. F. (2012). How effective were lifestyle interventions in real-world settings that were modeled on the Diabetes Prevention Program? *Health Affairs, 31*(1), 67-75. doi:10.1377/hlthaff.2011.1009

American Diabetes Association. (2018). 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2018. *Diabetes Care, 41*(Suppl 1), S13-s27. doi:10.2337/dc18-S002

Aviles-Santa, M. L., Schneiderman, N., Savage, P. J., Kaplan, R. C., Teng, Y., Perez, C. M., . . . Cowie, C. C. (2016). IDENTIFYING PROBABLE DIABETES MELLITUS AMONG HISPANICS/LATINOS FROM FOUR U.S. CITIES: FINDINGS FROM THE HISPANIC COMMUNITY HEALTH STUDY/STUDY OF LATINOS. *Endocrine Practice*. doi:10.4158/ep151144.or

Bernabe-Ortiz, A., Smeeth, L., Gilman, R. H., Sanchez-Abanto, J. R., Checkley, W., Miranda, J. J., & Study Group, C. C. (2016). Development and Validation of a Simple Risk Score for Undiagnosed Type 2 Diabetes in a Resource-Constrained Setting. *J Diabetes Res, 2016*, 8790235. doi:10.1155/2016/8790235

Boyko, E. J., Gerstein, H. C., Mohan, V., Yusuf, S., Sheridan, P., Anand, S., & Shaw, J. E. (2010). Effects of ethnicity on diabetes incidence and prevention: results of the Diabetes REduction Assessment with ramipril and rosiglitazone Medication (DREAM) trial. *Diabetic Medicine, 27*(11), 1226-1232. doi:10.1111/j.1464-5491.2010.03064.x

Brown, N., Critchley, J., Bogowicz, P., Mayige, M., & Unwin, N. (2012). Risk scores based on self-reported or available clinical data to detect undiagnosed type 2 diabetes: a systematic review. *Diabetes Research and Clinical Practice, 98*(3), 369-385. doi:10.1016/j.diabres.2012.09.005

Buchanan, T. A., Xiang, A. H., Peters, R. K., Kjos, S. L., Marroquin, A., Goico, J., . . . Azen, S. P. (2002). Preservation of pancreatic beta-cell function and prevention of type 2 diabetes by pharmacological treatment of insulin resistance in high-risk hispanic women. *Diabetes, 51*(9), 2796-2803.

Buijsse, B., Simmons, R. K., Griffin, S. J., & Schulze, M. B. (2011). Risk assessment tools for identifying individuals at risk of developing type 2 diabetes *Epidemiologic Reviews, 33*, 46-62. doi:10.1093/epirev/mxq019

Casagrande, S. S., Cowie, C. C., & Genuth, S. M. (2014). Self-reported prevalence of diabetes screening in the U.S., 2005-2010. *American Journal of Preventive*

Medicine, 47(6), 780-787. doi:10.1016/j.amepre.2014.07.039

Centers for Disease Control and Prevention. (2017). *Fast Facts and Fact Sheets: Current Cigarette Smoking among Adults in the United States*. Retrieved from https://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/

Colby, S. L., & Ortman, J. M. (2015). Projections of the size and composition of the US population: 2014 to 2060, Current Population Reports, P25-1143, US Census Bureau, Washington, DC.

Cruz-Vidal, M., Costas, R., Jr., Garcia-Palmieri, M. R., Sorlie, P. D., & Hertzmark, E. (1979). Factors related to diabetes mellitus in Puerto Rican men. *Diabetes, 28*(4), 300-307.

Dawber, T. R., Kannel, W. B., Revotskie, N., Stokes, J., 3rd, Kagan, A., & Gordon, T. (1959). Some factors associated with the development of coronary heart disease: six years' follow-up experience in the Framingham study. *American Journal of Public Health and the Nations Health, 49*, 1349-1356.

The Diabetes Prevention Program (DPP): description of lifestyle intervention. (2002). *Diabetes Care, 25*(12), 2165-2171.

Flegal, K. M., Ezzati, T. M., Harris, M. I., Haynes, M. G., Juarez, R. Z., Knowler, W. C., . . . Stern, M. P. (1991). Prevalence of diabetes in Mexican Americans, Cubans, and Puerto Ricans from the Hispanic health and nutrition examination survey, 1982–1984. *Diabetes Care, 14*(7), 628-638.

Flegal, K. M., Ezzati, T. M., Harris, M. I., Haynes, S. G., Juarez, R. Z., Knowler, W. C., . . . Stern, M. P. (1991). Prevalence of diabetes in Mexican Americans, Cubans, and Puerto Ricans from the Hispanic Health and Nutrition Examination Survey, 1982-1984. *Diabetes Care, 14*(7), 628-638.

Gaskill, S. P., Allen, C. R., Garza, V., Gonzales, J. L., & Waldrop, R. H. (1981). Cardiovascular risk factors in Mexican Americans in Laredo, Texas. I. Prevalence of overweight and diabetes and distributions of serum lipids. *American Journal of Epidemiology, 113*(5), 546-555.

Gerstein, H. C., Yusuf, S., Bosch, J., Pogue, J., Sheridan, P., Dinccag, N., . . . Holman, R. R. (2006). Effect of rosiglitazone on the frequency of diabetes in patients with impaired glucose tolerance or impaired fasting glucose: a randomised controlled trial. *Lancet, 368*(9541), 1096-1105. doi:10.1016/s0140-6736(06)69420-8

Gillies, C. L., Lambert, P. C., Abrams, K. R., Sutton, A. J., Cooper, N. J., Hsu, R. T., . . . Khunti, K. (2008). Different strategies for screening and prevention of type 2 diabetes in adults: cost effectiveness analysis. *BMJ, 336*(7654), 1180-1185. doi:10.1136/bmj.39545.585289.25

Group, M. o. D. M. U. W. (2010). VA/DoD clinical practice guideline for the management of diabetes mellitus. Version 4.0. *Update August*.

Gunby, P. (1980). San Antonio heart study compares ethnic groups. *JAMA, 244*(3), 225.

Handelsman, Y., Mechanick, J., Blonde, L., Grunberger, G., Bloomgarden, Z., Bray, G., . . . Wyne, K. (2011). American Association of Clinical Endocrinologists Medical Guidelines for Clinical Practice for Developing a Diabetes Mellitus Comprehensive Care Plan. *Endocrine Practice, 17*(Supplement 2), 1-53. doi:10.4158/ep.17.s2.1

Health, N. I. f., & Excellence, C. (2012). Preventing type 2 diabetes: risk identification and interventions for individuals at high risk. *NICE Guidelines PH38*.

Herman, W. H., Smith, P. J., Thompson, T. J., Engelgau, M. M., & Aubert, R. E. (1995). A new and simple questionnaire to identify people at increased risk for undiagnosed diabetes. *Diabetes Care, 18*(3), 382-387.

Hoerger, T. J., Segel, J. E., Gregg, E. W., & Saaddine, J. B. (2008). Is glycemic control improving in U.S. adults? *Diabetes Care, 31*(1), 81-86. doi:10.2337/dc07-1572

Holman, R. R., Paul, S. K., Bethel, M. A., Matthews, D. R., & Neil, H. A. W. (2008). 10-year follow-up of intensive glucose control in type 2 diabetes. *New England Journal of Medicine, 359*(15), 1577-1589.

Johnson, C. L., Dohrmann, S. M., Burt, V., & Mohadjer, L. K. (2014). National health and nutrition examination survey: sample design, 2011–2014.

Khunti, K., & Davies, M. (2012). Should we screen for type 2 diabetes: Yes. *BMJ, 345*, e4514.

Khunti, K., Gillies, C. L., Taub, N. A., Mostafa, S. A., Hiles, S. L., Abrams, K. R., & Davies, M. J. (2012). A comparison of cost per case detected of screening strategies for Type 2 diabetes and impaired glucose regulation: modelling study. *Diabetes Research and Clinical Practice, 97*(3), 505-513.

Khunti, K., Mani, H., Achana, F., Cooper, N., Gray, L. J., & Davies, M. J. (2015). Systematic Review and Meta-Analysis of Response Rates and Diagnostic Yield of Screening for Type 2 Diabetes and Those at High Risk of Diabetes. *PloS One, 10*(9), e0135702. doi:10.1371/journal.pone.0135702

Knowler, W. C., Barrett-Connor, E., Fowler, S. E., Hamman, R. F., Lachin, J. M., Walker, E. A., & Nathan, D. M. (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England Journal of Medicine, 346*(6), 393-403. doi:10.1056/NEJMoa012512

Krogstad, J. M., & Lopez, M. H. (2014). Hispanic nativity shift. *Washington, DC: The*.

Li, G., Zhang, P., Wang, J., Gregg, E. W., Yang, W., Gong, Q., . . . Bennett, P. H. (2008). The long-term effect of lifestyle interventions to prevent diabetes in the China Da Qing Diabetes Prevention Study: a 20-year follow-up study. *Lancet, 371*(9626), 1783-1789. doi:10.1016/s0140-6736(08)60766-7

Lindström, J., Peltonen, M., Eriksson, J. G., Louheranta, A., Fogelholm, M., Uusitupa, M., & Tuomilehto, J. (2006). High-fibre, low-fat diet predicts long-term weight loss and decreased type 2 diabetes risk: the Finnish Diabetes Prevention Study. *Diabetologia, 49*(5), 912-920.

Menke, A., Casagrande, S., Geiss, L., & Cowie, C. C. (2015). Prevalence of and Trends in Diabetes Among Adults in the United States, 1988-2012. *JAMA, 314*(10), 1021-1029. doi:10.1001/jama.2015.10029

Moy, C. S. (1977). Determining ethnic origin in an interview survey. Problems and recommendations. *Public Health Reports, 92*(5), 414-420.

O'Brien, M. J., Perez, A., Scanlan, A. B., Alos, V. A., Whitaker, R. C., Foster, G. D., . . . Homko, C. (2017). PREVENT-DM Comparative Effectiveness Trial of Lifestyle Intervention and Metformin. *American Journal of Preventive Medicine, 52*(6), 788-797. doi:10.1016/j.amepre.2017.01.008

Ockene, I. S., Tellez, T. L., Rosal, M. C., Reed, G. W., Mordes, J., Merriam, P. A., . . . Ma, Y. (2012). Outcomes of a Latino community-based intervention for the prevention of diabetes: the Lawrence Latino Diabetes Prevention Project. *American Journal of Public Health, 102*(2), 336-342. doi:10.2105/ajph.2011.300357

Pires de Sousa, A. G., Pereira, A. C., Marquezine, G. F., Marques do Nascimento-Neto, R., Freitas, S. N., Nicolato, R. L. d. C., . . . Krieger, J. E. (2009). Derivation and external validation of a simple prediction model for the diagnosis of type 2 diabetes mellitus in the Brazilian urban population. *European Journal of Epidemiology, 24*. doi:10.1007/s10654-009-9314-2

Royston, P., Moons, K. G., Altman, D. G., & Vergouwe, Y. (2009). Prognosis and prognostic research: Developing a prognostic model. *BMJ, 338*, b604. doi:10.1136/bmj.b604

Schneiderman, N., Llabre, M., Cowie, C. C., Barnhart, J., Carnethon, M., Gallo, L. C., . . . Aviles-Santa, M. L. (2014). Prevalence of diabetes among Hispanics/Latinos from diverse backgrounds: the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). *Diabetes Care, 37*(8), 2233-2239. doi:10.2337/dc13-2939

Schwarz, P. E., Li, J., Lindstrom, J., & Tuomilehto, J. (2009). Tools for predicting the risk of type 2 diabetes in daily practice. *Hormone and Metabolic Research, 41*(2), 86-97. doi:10.1055/s-0028-1087203

Siu, A. L. (2015). Screening for Abnormal Blood Glucose and Type 2 Diabetes Mellitus: U.S. Preventive Services Task Force Recommendation Statement. *Annals of Internal Medicine, 163*(11), 861-868. doi:10.7326/m15-2345

Sorlie, P. D., Aviles-Santa, L. M., Wassertheil-Smoller, S., Kaplan, R. C., Daviglus, M. L., Giachello, A. L., . . . Heiss, G. (2010). Design and implementation of the Hispanic Community Health Study/Study of Latinos. *Annals of Epidemiology,*

20(8), 629-641. doi:10.1016/j.annepidem.2010.03.015

Sperl-Hillen, J. M., & O'Connor, P. J. (2005). Factors driving diabetes care improvement in a large medical group: ten years of progress. *American Journal of Managed Care, 11*(5 Suppl), S177-185.

Stein, A. D., Lederman, R. I., & Shea, S. (1993). The Behavioral Risk Factor Surveillance System questionnaire: its reliability in a statewide sample. *American Journal of Public Health, 83*(12), 1768-1772.

Thoopputra, T., Newby, D., Schneider, J., & Li, S. C. (2012). Survey of diabetes risk assessment tools: concepts, structure and performance. *Diabetes/Metabolism Research and Reviews, 28*(6), 485-498. doi:10.1002/dmrr.2296

Vijan, S. (2010). Type 2 diabetes. *Annals of Internal Medicine, 152*(5), ITC3-1-1.

Wareham, N. J., & Griffin, S. J. (2001). Should we screen for type 2 diabetes? Evaluation against National Screening Committee criteria. *BMJ, 322*(7292), 986-988.

Witte, D. R., Shipley, M. J., Marmot, M. G., & Brunner, E. J. (2010). Performance of existing risk scores in screening for undiagnosed diabetes: an external validation study. *Diabetic Medicine, 27*(1), 46-53. doi:10.1111/j.1464-5491.2009.02891.x

Xiang, A. H., Peters, R. K., Kjos, S. L., Marroquin, A., Goico, J., Ochoa, C., . . . Buchanan, T. A. (2006). Effect of pioglitazone on pancreatic beta-cell function and diabetes risk in Hispanic women with prior gestational diabetes. *Diabetes, 55*(2), 517-522.

THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Institutional Review Board for Human Use

Exemption Designation
Identification and Certification of Research
Projects Involving Human Subjects

UAB's Institutional Review Boards for Human Use (IRBs) have an approved Federalwide Assurance with the Office for Human Research Protections (OHRP). The Assurance number is FWA00005960 and it expires on November 8, 2021. The UAB IRBs are also in compliance with 21 CFR Parts 50 and 56.

Principal Investigator:	JUAREZ, LUCIA D
Co-Investigator(s):	
Protocol Number:	**E161207005**
Protocol Title:	*Assessing Diabetes Risk Among Hispanic Populations: Development and Validation of A Risk Score Using Readily Avaliable Information*

The above project was reviewed on 12/14/16. The review was conducted in accordance with UAB's Assurance of Compliance approved by the Department of Health and Human Services. This project qualifies as an exemption as defined in 45CFR46.101(b), paragraph 4.

This project received EXEMPT review.

Date IRB Designation Issued: 12/14/16

Cari Oliver, CIP
Assistant Director, Office of the
Institutional Review Board for Human
Use (IRB)

Investigators please note:

Any modifications in the study methodology, protocol and/or consent form/information sheet must be submitted for review to the IRB prior to implementation.

470 Administration Building
701 20th Street South
205.934.3789
Fax 205.934.1301
irb@uab.edu

The University of
Alabama at Birmingham
Mailing Address:
AB 470
1720 2ND AVE S
BIRMINGHAM AL 35294-0104