Biobanking in Latinos: current status, principles for conduct, and contribution of a new biobank, El Banco por Salud, designed to improve the health of Latino patients of Mexican ancestry with type 2 diabetes

Oscar D Parra,[1] Lindsay N Kohler ![ORCID],[2,3] Lori Landes,[4,5] Alexis A Soto,[1] Diana Garcia,[1] Jacqueline Mullins,[4] Patty Molina,[6] Eladio Pereira,[6] Douglas J Spegman,[4] Lisa Soltani,[4] Lawrence J Mandarino[1]

For numbered affiliations see end of article.

Correspondence to
Dr Lawrence J Mandarino;
mandarino@arizona.edu

ABSTRACT

Underserved Latino communities experience a greater burden of type 2 diabetes mellitus (T2DM) than the general population. Predictors of glycemic control are likely to include both biological/genetic and social determinants of health (SDOH). A variety of approaches have been used with cohorts of Latino patients to study aspects of this health disparity, and those are reviewed briefly here. Such projects range from cohorts that are studies for a primary purpose, for example, to discover genetic variation associated with T2DM or to examine a particular aspect of SDOH that might be involved. Other studies have been conducted more as infrastructure that is broadly based in order to provide a resource that can be used by many investigators to address a variety of questions. From our experience and those of others, we propose a set of principles to ensure that needs of the community are identified and taken into account during the conduct of these studies. As an example of the implementation of these principles, we also describe a new biobank *El Banco por Salud* (El Banco), which was designed to improve access to studies designed to improve glycemic control and health in Latinos in partnership with Federally Qualified Health Centers in Arizona.

INTRODUCTION

The rise of precision medicine has made it abundantly clear that genetic variation is often an important factor in treatment of disease.[1] A number of precision medicine and biobanking initiatives, including the All of Us project,[2] have recognized that patients from all communities and ethnicities need to be engaged in these efforts.[3] Historically, there have been barriers to engaging diverse communities, including lack of access to healthcare and distrust among community members.[4 5] Among the communities in

the USA that experience healthcare disparities[6] and fewer opportunities to participate in biomedical research are groups that identify as Hispanic or Latino.[7] Within that broad category there also is wide variation, including communities as diverse biologically as Latinos derived from Caribbean populations and communities of predominantly Mexican ancestry.[8] These communities differ both genetically and culturally.[9] Many Latino communities often experience disparities in health, healthcare, income and education.[10] In general, Latino populations experience a greater burden of chronic diseases such as type 2 diabetes mellitus (T2DM) and obesity. However, complex diseases such as T2DM are phenotypes that have a biological (genetic) and environmental basis. Here, we define environmental as behavioral, social and cultural factors that interact with a person's biology. Especially in the case of underserved communities, social factors that contribute to disease have been termed social determinants of health (SDOH). Any biobank designed to create infrastructure to implement precision medicine therefore needs to include biological and genetic variables, and information regarding SDOH on the patient and community level.

This review was undertaken to (1) describe the status of medium to large scale studies of genetic basis of T2DM in the context of biobanking in Latino communities of Mexican ancestry, (2) identify gaps in these projects, and (3) propose a set of principles for developing biobanks for Latinos of Mexican ancestry that are community-based.

This is not intended to be an exhaustive review. Rather we summarize the status of the field, using brief examples and introduce a new biobank, El Banco por Salud, that attempts to address identified gaps through adhering to a set of principles for conducting such studies that we propose here. El Banco specifically addresses T2DM, obesity, and related diseases, which are among the most critical public health problems in communities of Latinos of Mexican ancestry living in the USA.[11 12]

Current status of biobanking in Latinos of Mexican ancestry; emphasis on T2DM

Broadly speaking, a biobank is any collection of biospecimens and data that are used for biomedical research.[13] Most biobanks derive from cohort-type studies that may be cross-sectional or longitudinal and usually are driven by the purpose of studying a particular disease, like T2DM, that is particularly important for the community from which participants in the biobank are derived. Such 'biobanks' are usually designed for the exclusive research use of the investigators carrying out these projects and their direct collaborators. Examples of such projects, which we call purpose-driven biobanks, are provided in table 1. This list is not intended to be comprehensive but rather provide some examples of these biobanks. The projects cited in table 1 focus specifically on the Latino community whose members primarily are derived from the peoples of Mexico. Other biobanks not listed here may focus on Latinos of Central American or Caribbean ancestry or from the diverse peoples of South America. The studies in table 1 focus on projects designed to answer questions regarding the epidemiology and genetic origin of T2DM and related conditions. With respect to T2DM and metabolic disease, such studies have yielded valuable information regarding the biological and genetic pathogenesis of disease, providing important clues regarding the genes involved in the development of T2DM in Latinos of Mexican ancestry. However, such studies have some serious limitations by their nature. First, access to samples or data may be restricted, either because of complexities of collaborations or the way consent was obtained that can limit the usefulness of biospecimens for research outside the specific, original study question. Second, because the study was done with a specific purpose in mind, there may not be sufficient ancillary data to allow other investigators to answer different questions. Third, there may be limited data bearing on the SDOH. Fourth, these studies were not designed to provide direct interventions or to improve the health of the participants. Finally, the participants in these studies were unlikely to have provided consent for research access to their medical record.

More recently, biobanking in underserved communities in general and in Latinos specifically has become more directly connected to the community and conducted in a manner that allows multiple investigators to access data and biospecimens to answer a multitude of questions rather than having one primary question

('Resource-focused biobanks, table 1). In other words, these more modern biobanks have become resources to connect investigators to the community through an access process mediated by patients, clinics, and academic investigators that is designed to maximize benefit to patients/participants in the community.

By far the largest such project is the 1 million patient, National Institutes of Health-funded, All of Us Research Program, which includes a large proportion of Latino participants.[14] Because this project connects data on SDOH and biospecimens with the electronic health record (EHR) in a large number of patients, it will be invaluable for developing precision treatments for a wide variety of common and less common diseases in a manner that accounts for patient engagement. However, because of the vast expanse of this project and the diverse communities included therein, it is more difficult, although not impossible, to use this resource on a local community level to design precision interventions for specific patients in a particular local community. Particularly with Latino patients, local community context is important both culturally and genetically, especially with respect to T2DM. Even more recently, the Mil Familias study is a cohort study of patients with T2DM and their families that is designed to accrue data on biology related to diabetes and SDOH, specifically to determine the 'real-world' burden of T2DM in Latinos, primarily of Mexican ancestry.[15] This study addresses some of the gaps and concerns present in other Latino biobanks. For example, engagement with the community was addressed by recruitment from and engagement with Federally Qualified Health Centers (FQHCs), other Hispanic Serving Institutions, and the use of 'Especialistas', similar to the role of a promotora. However, this cohort was still established as a resource for a single group of investigators and a single main purpose. As far as is evident, engagement of patients and patient-serving institutions in having input into design or approval of ancillary studies is not included in the overall design of this study. Moreover, the multiple recruitment approaches and multiple clinical referrals likely limit the utility of EHR data. We had previously engaged the Latino community of Maricopa County, Arizona (Phoenix and environs) in the Arizona Insulin Resistance Registry, a project designed and conducted from approximately 2007–2009 to recruit and characterize Latinos in a manner that would allow the use of biospecimens and data in future studies and also increase the opportunities for participants to engage in biomedical research.[16] The Latino community in Maricopa County consists primarily of individuals with ancestry connected to Mexico and the Southwest USA. Although this project was successful in fulfilling those aspects of its purpose, it was conducted before Medicaid expansion with the Affordable Care Act so many biobank participants did not have a regular primary care provider or stable care at a single clinic. Therefore, those patients also were not connected to an ongoing EHR, limiting utilization and definition of phenotypes to those that

Table 1 Purpose and resource-based studies of risk for T2DM and related conditions in Latinos of Mexican ancestry

Purpose-driven biobanks

Name	Location	Purpose/Study type
Laredo Project[22]	Laredo, Texas	Epidemiology of T2DM and cardiovascular risk cohort.
Starr County Study[23]	Starr County, Texas	Epidemiology and genetics of T2DM cohort.
Insulin Resistance and Atherosclerosis Study[24]	San Antonio, Texas	Epidemiology and characterization of insulin resistance and dyslipidemia cohort.
Genetics of Non-insulin-dependent Diabetes[25]	Multicenter (San Antonio, Texas)	Identification of genes causing T2DM. Multiple ethnic group cohort.
San Antonio Family Heart Study[26]	San Antonio, Texas	Identification of genes causing heart disease and related conditions cohort.
Multi-Ethnic Study of Atherosclerosis[27]	Multicenter	Study of genetics, prevalence, and progression of cardiovascular disease in a prospective multiethnic cohort
San Antonio Family Diabetes/Gallbladder Study[28]	San Antonio, Texas	Identification of genes causing T2DM and related diseases cohort.
Veterans Administration Genetic Epidemiology Study[29]	San Antonio, Texas	Identification of genes causing T2DM and related diseases cohort.
Brownsville study[30]	Brownsville, Texas	Identification of biological and social risk factors for T2DM cohort.

Resource-focused biobanks

Name	Location	Purpose/Resources
Arizona Insulin Resistance Registry[16]	Maricopa County, Arizona	Research registry of individuals of Mexican ancestry at risk of T2DM and cardiometabolic risk willing to participate in future studies (data, plasma, serum, DNA, immortalized white blood cells).
All of US Research Program[14]	Entire US population	Diverse US cohort of >1 million persons to accelerate biomedical research and improve health by connecting biological (genetic), EHR, and SDOH data. Ongoing genetic, epidemiological, and EHR cohort.
Million Vet Program[31]	Population of US Veterans	Diverse ongoing cohort of 1 million patients receiving care at Veterans Administration healthcare facilities by connecting EHR to genetic data.
Sangre por Salud biobank[17]	Maricopa County, Arizona	To expand precision medicine research to include under-represented Latino patients. Biobank embedded in a FQHC that provides infrastructure for precision medicine implementation as well as genetic and clinical information (connected to EHR).
Mil Familias[15]	Santa Barbara, California and environs	To establish an observational cohort of 1000 Latino families with at least one family member with T2DM to understand the impact of genetics, biology, behavior, psychology, and society/environment.

Representative examples of purpose-driven studies of T2DM and related conditions such as cardiovascular disease and dyslipidemia among Latinos of Mexican ancestry are listed under the heading 'Purpose-driven biobanks'. Examples of projects that are designed to provide more comprehensive biobank data and samples to support many investigators are listed under the heading 'Resource-focused biobanks'. These lists are not intended to be comprehensive.

EHR, electronic health record; FQHC, Federally Qualified Health Center; SDOH, social determinants of health; T2DM, type 2 diabetes mellitus.

Table 2 Principles of biobanking in underserved Latino communities

Principle	Description
1. Community engagement	A Latino biobank created by an academic institution should involve engagement on multiple levels with participants, community representatives, and clinical partners to ensure projects using biobank data and patients provide benefit to the community.
2. Cultural appropriateness	A Latino biobank should be culturally appropriate in the context of the nature of the local Latino community and involve research staff who are bilingual and are culturally embedded in the local community.
3. Community partnership	A Latino biobank undertaken by an academic institution should be a partnership with a limited number of community clinics that primarily serve the community. Local Federally Qualified Health Centers are good examples of such clinics.
4. Broadest impact	Data and biospecimens should be collected in a manner consistent with consent for use in future studies to allow access to as many approved investigators as possible to broaden the impact of the biobank to the utmost extent.
5. Joint governance	Clinic staff and academic investigators should have joint and equal input into the conduct of the biobank and ancillary studies.

were study specific. This also limited the ability to recontact participants for further studies because contact information changes frequently in individuals unable to afford continuity of cell phone number or who may change addresses frequently.

Based on our experience and the limitations of other projects, we developed a set of principles for community engaged biobanking in Latinos. These principles optimally would apply to the study of all complex diseases that are phenotypes, that is, that occur as a result of the interaction of the environment (eg, SDOH, built environment, etc) with genetics. These principles are given in table 2.

With these principles in mind, we engaged a, Mountain Park Health Center (MPHC) in Maricopa County for the purpose of developing a Latino biobank, known as Sangre por Salud (SPS), consisting of patients connected to the MPHC EHR who were consented for use of biospecimens and data in future studies and phenotyped with laboratory studies, anthropometrics, personal and family health histories, and questionnaire data using patient-reported measures of health, diet, behavioral wellness, and potential environmental exposures.[17] The purposes of the SPS biobank were to improve early T2DM detection at MPHC and create infrastructure that would facilitate precision medicine research in a Latino patient population. The project was successful in both instances, revealing the extent of prediabetes and diabetes in the MPHC patient population[17] and allowing patients to participate in Electronic Medical Records and Genomics (eMERGE) III, providing experience with return of genetic results in underinsured patients from a Latino community with health disparities.[18] Patients who were without a diagnosis of T2DM, cancer, or other chronic disease were recruited. This biobank, which is in ongoing use by the Mayo Clinic, has been especially important for providing experience regarding the implementation of precision, genomic medicine in a community clinic setting.[18] This biobank also proved useful to MPHC by

revealing that nearly 50% of non-diabetic patients met criteria for prediabetes, and that measurement of hemoglobin A1c (HbA1c) should be performed at every visit. However, since this biobank, by design, recruited patients who did not have a prior diagnosis of T2DM, its utility for studies involving patients with diabetes or providing tailored, targeted interventions in these patients is highly limited.

Latino communities are at higher risk of obesity and T2DM.[19] Moreover, in diverse communities with healthcare disparities, glycemic control may be worse than in communities that have ready access to healthcare, prescription drug programs, diabetes technology, and diabetes self-management education leading to an excess burden of complications of diabetes.[9] It is important to identify the determinants of poor glycemic control in these patients to target appropriate variables that would have the greatest impact. There is evidence that health insurance status, diabetes care utilization, patient-self management, language, acculturation, and social support influence glycemic control specifically in Latino populations.[20] Because of the lack of Latino biobanks focused on T2DM, since predictors of glycemic control are likely to include both biological/genetic and SDOH, and because these factors may be specific to ethnicity and community, we partnered with El Rio Community Health Center in Tucson, Arizona, and Mariposa Community Health Center in Nogales, Arizona to develop and conduct *El Banco por Salud* (El Banco). The purposes of El Banco are to (1) provide infrastructure that will aid researchers and FQHC partners in identifying clinical, genetic, and sociocultural determinants of glycemic control in Latino patients with T2DM, (2) create more opportunities for Latino patients to participate in biomedical research pertinent to their community, and (3) to create research opportunities for studies that are designed to understand the biological and social mechanisms underlying poor glycemic control in Latino patients. This biobank is a natural extension of our previous efforts, using the same

Figure 1 El Banco por Salud enrollment site locations.

principles of engagement of a community health center, but focused on T2DM, obesity, and other conditions related to insulin resistance and metabolic syndrome. Here, we report the design of the study and describe the demographic, clinical, and sociocultural characteristics of first 1111 patients enrolled El Banco.

Establishing community health clinic partnerships

To develop El Banco, partnerships were developed between the UA Center for Disparities in Diabetes, Obesity, and Metabolism (CDDOM) and FQHCs El Rio Community Health Center (El Rio) in Tucson, Arizona and Mariposa Community Health Center (Mariposa) serving Nogales, Arizona and Santa Cruz County (figure 1). These partnerships were established after extensive discussions over a period of 9 months. Considerations included focus of the biobank and the relationship to precision medicine, patient eligibility, data collection, patient privacy, access to electronic health records, operations, and governance. Integral to these partnerships would be a joint academic/community health center governance model, where El Rio and Mariposa leadership would participate equally with UA CDDOM staff on an executive leadership board. The function of this board is to oversee operations and use of data, biospecimens, and access to patients for recruitment into future studies. The inclusion of two community health centers was purposeful. El Rio is located in a predominantly urban area and is about a 1 hour drive from the US-Mexico border, whereas Mariposa is in a micropolitan area directly adjacent to the international border, which could provide information regarding location as a factor in health disparities. Additionally, there is a great deal of daily border crossing among residents of Nogales, Arizona and Nogales, Sonora. The closer

proximity of patients from Mariposa with Mexico may produce interesting and important cultural differences. Another potentially important factor are the potential differences in clinical care and diabetes education/lifestyle programs between the two separate institutions. Finally, we deemed it to be important to cover a larger geographical area than the Tucson metropolitan area.

El Banco activities at El Rio and Mariposa were funded by UA Health Sciences through subcontracts with FQHCs employing study coordinators at each site for recruitment and study activities. These El Rio and Mariposa study staff, along with UA personnel, were designated as members of the core research team responsible for all recruitment and enrollment efforts. It was jointly agreed on that the focus of this biobank would be self-identified Latino patients diagnosed with T2DM or prediabetes and their family. The rationale for focusing on T2DM within Latino communities was based on (1) the underrepresentation of Latino patients in clinical research and biobank studies, (2) the largest minority demographic in Arizona is Latino/Hispanic (8), (3) most of the demographic served by our FQHC partners are Latinos, and (4) T2DM comprises a large proportion of healthcare delivered by the FQHC partners. It was reasoned during these deliberations that such a design would create a biobank of consented patients who could participate in studies designed to make use of a family centered community biobank to address the causes for disparities in glycemic control in FQHC Latino patients and lead to pragmatic trials to improve glycemic control in Latino patients with poorly controlled T2DM. The principles laid out above were the foundation on which El Banco was built.

To collect data and biospecimens at El Rio and Mariposa in the most consistent manner, the study designs for each site were similar while allowing for practical differences related to space and staff preferences, consistent with the principle of community engagement. A principal investigator for each site was selected from clinical leadership. Funding for CDDOM through subcontracts supports FQHC-employed bilingual clinical research coordinators responsible for recruitment and enrollment. Bilingual/bicultural CDDOM research staff are responsible for onsite data collection. Biospecimen samples and blood samples for research use are collected by onsite phlebotomy laboratory. Clinical laboratory measurements were analyzed by clinical laboratories each clinic normally uses.

Recruitment and enrollment strategies that minimized clinical workflow disruptions and were synchronous with clinic operations and best practices were employed at each site (figure 2). Study eligible patients (probands) were prescreened by their primary care physicians and clinic staff through their EHR. Eligibility criteria for patients include self-reported Latino ethnicity, age 18–75 years, and a HbA1c of 5.7 or greater. Family members (defined as siblings, parents, children, aunts/uncles, cousins, and significantly close friends proband considers kin/family) were recruited simultaneously. Patients and
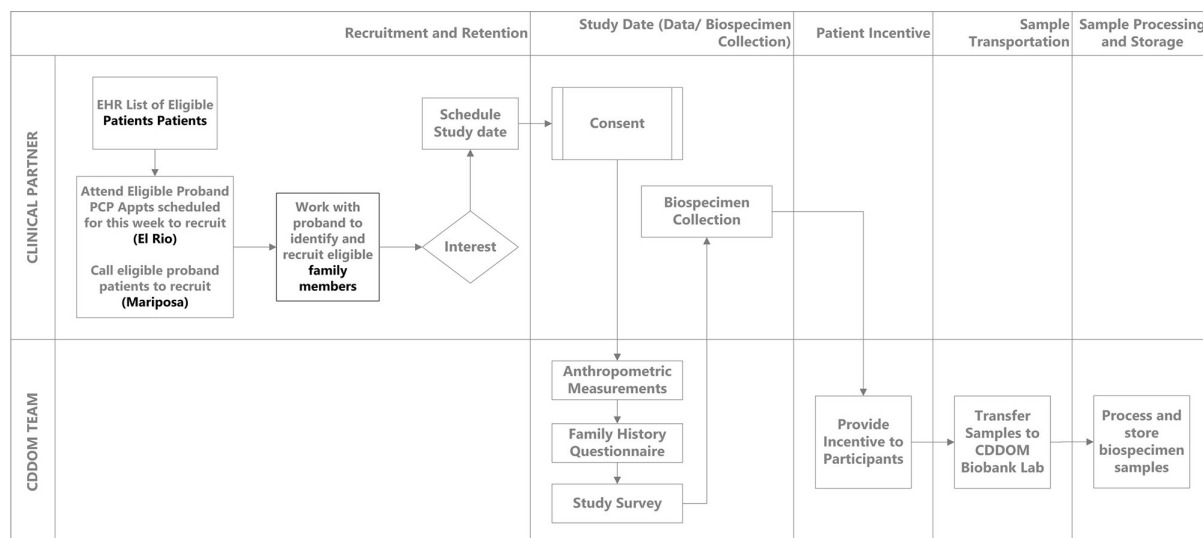
Figure 2 UA Center for Disparities in Diabetes, Obesity, and Metabolism (CDDOM) El Banco por Salud study work flow. EHR, electronic health record.

family members are linked through a unique identifier for the family unit. Exclusion criteria includes (1) history of cancer, excluding non-melanoma skin cancers, in the past 3 years (remission >3 years); (2) currently pregnant, delivered a baby within the last 12 months, or currently breast feeding; and (3) feel they are unable to refrain from smoking for 1.5 hours. Eligible patients were contacted either by 'cold calls' or in-person at regular clinic appointments. For Mariposa, recruitment of patients identified as eligible was primarily accomplished via cold calls. The Mariposa clinical research coordinators contacted eligible patients to schedule an in-person appointment for data collection to complete the study at their Nogales, Arizona clinic. The clinical research coordinators worked with the proband to identify and recruit eligible family members to participate in the study and accompany the proband. Transportation to and from the clinic was provided if necessary.

Impact of COVID-19: due to the COVID-19 pandemic, a virtual component was implemented to allow for physical distancing practices. For El Rio, the virtual component was introduced to both recruitment and study visits. For Mariposa, the virtual component was only introduced to the study visits since Mariposa's primary recruitment occurred via phone calls. A virtual telehealth component was introduced to reduce the required in-person time component from 2 hours to half an hour. The in-person component included the blood draw, saliva sample, and the anthropometric measures, while the remaining data collection was conducted virtually. Along with the changes to the study design and workflow, new protocols and procedures were introduced including additional personal protective equipment, sanitation practices, and a COVID-19 prescreening to assure the safety of both the research staff and study participants. In-person enrollment resumed in January 2021.

All participants were scheduled for the study visit at their usual clinic after an overnight fast. Written consent was obtained from participants in their preferred language by a bilingual research member. Participants were asked to complete a comprehensive survey offered in English or Spanish covering general health and personal health perceptions, detailed health history, health behaviors, demographics, and quality of life. Anthropometric measures including height, weight, waist circumference, seated blood pressure, and pulse were collected along with a blood draw for both the clinical panels including glucose, cholesterol, and complete blood count with differentials and the samples for banking. The study was designed to minimize human error in relation to data transfer from paper to electronic by eliminating the need for printed data collection tools, such as questionnaires and surveys. All data collection tools and eventually the consent form, were designed as electronic forms directly embedded into our REDCap Database. As the participant completed the study forms, the data were directly stored into the study database, eliminating the need for data transfers and data entry.

A goal of El Banco is to promote research across the university that is inclusive of the Latino population. To accomplish this goal, El Banco was structured as a centralized resource accessible (on review and approval) to investigators interested in Latino populations and the robust data and biospecimen bank El Banco offers. To oversee the review and approval process of each request for El Banco data and samples, the CDDOM Biobank Executive Joint Governance Committee was created consisting of the CDDOM Leadership/El Banco Leadership, and the FQHC Leadership and El Banco research staff.

A data access request process was created for investigators to provide a summary of their proposed study. The request includes logistics (sample size, data and/or biospecimens required, funding, and timeline) and

Table 3 El Banco por Salud participant characteristics by enrollment site (November 2021)

	El Rio (n=867)	Mariposa (n=244)	Overall (n=1111)	P value
Proband	454 (52.4%)	86 (35.2%)	540 (48.6%)	<0.001
Female	567 (65.4%)	174 (71.3%)	741 (66.7%)	0.0834
Age, years	52.6 (14.6)	49.5 (16.0)	51.9 (15.0)	0.0045
Body mass index, kg/m^2	32.2 (6.8)	31.6 (6.8)	32.1 (6.8)	0.1803
Waist measurements, inches	41.2 (6.1)	40.1 (6.5)	41.0 (6.2)	0.0151
Systolic blood pressure	131.8 (24.6)	128.5 (20.1)	131.1 (23.7)	0.0575
Diastolic blood pressure	79.0 (8.2)	78.1 (8.6)	78.8 (8.3)	0.1077
Fasting plasma glucose, mmol/L	8.4 (4.3)	7.8 (4.4)	8.3 (4.3)	0.0657
Hemoglobin A1c, %	7.8 (2.3)	7.5 (2.3)	7.7 (2.3)	0.095
Fasting plasma insulin, uIU/mL	16.5 (23.8)	12.3 (11.8)	15.5 (21.5)	0.0075
HOMA-IR	6.5 (15.2)	4.7 (8.0)	6.1 (13.8)	0.0682
Cholesterol, mg/dL	173.1 (44.8)	179.5 (44.1)	174.5 (44.7)	0.0503
Triglyceride, mg/dL	187.1 (189.9)	168.4 (115.3)	183.0 (176.4)	0.145
Self-reported general health	3.5 (1.0)	3.3 (1.2)	3.4 (1.1)	0.0095
1 Excellent	37 (4.3%)	18 (7.4%)	55 (5.0%)	0.0043
2 Very good	90 (10.4%)	43 (17.6%)	133 (12.0%)	
3 Good	305 (35.3%)	77 (31.6%)	382 (34.5%)	
4 Fair	290 (33.6%)	66 (27.0%)	356 (32.2%)	
5 Poor	141 (16.3%)	40 (16.4%)	181 (16.4%)	

HOMA-IR, homeostasis model assessment-insulin resistance.

a description of how their study aligns with the overall mission of CDDOM, contributes to the field of science, benefits each FQHC, and mostly importantly, how the proposed study impacts and benefits the study and clinic patients. The joint executive committee then reviews the data access request and if deemed appropriate and in alignment with the mission of the biobank, the request is approved, and the investigator can proceed with human subjects' determination and application (if required).

Overall recruitment began in January 2018 and continued through March of 2020. Recruitment was paused at that time due to the COVID-19 pandemic. Enrollment visits began again in January 2021 and are ongoing. For El Rio, a total of 2456 prescreened eligible probands have been contacted. Of these, 1452 probands and family members agreed to be scheduled for a study visit. As of November 1, 2021, 1114 patients have been consented, 961 were enrolled, and 867 have laboratory values available from El Rio. For Mariposa, a total of 1078 prescreened eligible probands had been contacted. Of these, 412 probands and referred family members agreed to be scheduled for a study appointment. Currently, 312 participants have consented, 290 were enrolled with 244 participants with available lab values from Mariposa. Participant characteristics are shown in table 3 by enrollment site.

Overall, statistically significant differences were seen between enrollment sites for participant type, age, waist circumference, fasting plasma insulin, and general health. Average body mass index was obese at 32.2 and 31.6 kg/m^2 while waist circumference was 41.2 and 40.1 inches for El Rio and Mariposa, respectively. Fasting plasma glucose averaged 8.4 and 7.8 mmol/L at El Rio and Mariposa, respectively. HbA1c levels were similar between sites with a mean 7.8% for El Rio and 7.5% for Mariposa. Fasting insulin levels were also higher for El Rio than Mariposa (16.6±23.9 vs 12.3±11.8; p=0.0071). Homeostasis model assessment-insulin resistance differences did not reach statistical significance between sites (p=0.0682) with El Rio averaging 6.5 (15.2) and Mariposa 4.7 (8.0). Sites differed in self-reported general health on a scale of excellent to poor ranging from 1 to 5 with El Rio averaging 3.5 and Mariposa 3.3 (p=0.0095).

Cardiometabolic risk factor frequency as defined by the Adult Treatment Panel III[21] is shown in figure 3. Abdominal obesity as measured by waist circumference was the most prevalent risk factor followed by hyperglycemia (fasting plasma glucose >5.55 mmol/L). Nearly half of all participants exhibited high triglycerides while 60.4% showed low high-density lipoprotein (HDL)-cholesterol. Low HDL-cholesterol (p=0.029) and fasting plasma glucose (p<0.001) were statistically significantly different by enrollment site with El Rio having a greater percentage of patients experiencing each risk factor.

By design, El Banco is an infrastructure project to provide data, specimens, and patients to simplify
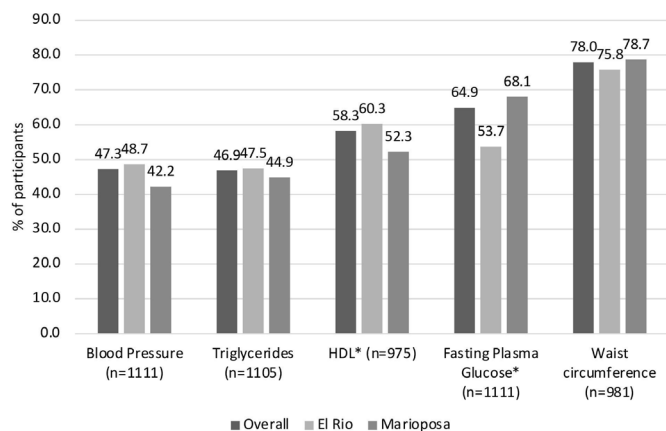
Figure 3 Cardiometabolic risk factors in El Banco por Salud. *Statistically significant (p<0.05) difference between enrolment site for overall participants assessed by χ^2. HDL, high-density lipoprotein.

collaborations between investigators and the community through a trusted community health center intermediary. In the first 4 years, El Banco has generated several diverse projects ranging from ancillary studies to patient recall to secondary data analysis. Most of the ongoing studies include students, trainees, and/or are led by junior faculty from departments including health promotion sciences, epidemiology, nutritional sciences, and medicine (endocrinology, nephrology, and cardiology).

DISCUSSION

This review provides a brief summary of the status of Latino biobanking in the USA, particularly with respect to patients with T2DM or at high risk of developing diabetes. The various approaches that have been taken to accomplish this are described, and gaps in these studies identified. Based on our experiences in creating three separate Latino biobanks and the experiences of others, we set forth and propose a set of principles by which the underserved Latino community should be engaged for conducting these projects. Finally, as an example of the implementation of these principles, we describe a new Latino biobank, El Banco por Salud, that focuses on Latino patients with T2DM and their family members to provide infrastructure for connecting patients in the community with academic investigators through the trusted intermediaries of Federally Qualified Community Health Centers where they receive primary care. The value of these principles is demonstrated by the ongoing and successful recruitment of patients, the strong and continuing relationships between an academic institution and FQHCs, and the utilization of biospecimens and patient data for studies conducted by academic investigators.

The key to establishing and maintaining this successful biobank is the collaborative model partnering community health centers with an academic institution focusing on true joint governance, teamwork, and transparency to advance knowledge and improve diabetic healthcare.

Author affiliations
[1]Department of Medicine, University of Arizona Medical Center—University Campus, Tucson, Arizona, USA
[2]Health Promotion Sciences, The University of Arizona Mel and Enid Zuckerman College of Public Health, Tucson, Arizona, USA
[3]Mel and Edith Zuckerman College of Public Health, University of Arizona Medical Center—University Campus, Tucson, Arizona, USA
[4]El Rio Community Health Centers, Tucson, Arizona, USA
[5]Department of Family Medicine and Population Health, Virginia Commonwealth University Medical Center, Richmond, Virginia, USA
[6]Mariposa Community Health Centers, Nogales, Arizona, USA

ORCID iD
Lindsay N Kohler http://orcid.org/0000-0002-0550-9282

REFERENCES

1 Jameson JL, Longo DL. Precision medicine--personalized, problematic, and promising. *N Engl J Med* 2015;372:2229–34.
2 The All of Us Research Program Investigators. The "All of Us" Research Program. *New England Journal of Medicine* 2019;381:668–76.
3 Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature* 2016;538:161–4.
4 Arevalo M, Heredia NI, Krasny S, *et al*. Mexican-American perspectives on participation in clinical trials: a qualitative study. *Contemp Clin Trials Commun* 2016;4:52–7.
5 Heredia NI, Krasny S, Strong LL, *et al*. Community perceptions of biobanking participation: a qualitative study among Mexican-Americans in three Texas cities. *Public Health Genomics* 2017;20:46–57.
6 In:National Academies of Sciences E, Medicine, Health, et al.. *Communities in action: pathways to health equity*. Washington (DC: National Academies Press (US), 2017.

7 Ceballos RM, Knerr S, Scott MA, *et al*. Latino beliefs about biomedical research participation: a qualitative study on the U.S.-Mexico border. *J Empir Res Hum Res Ethics* 2014;9:10–21.

8 Noe-Bustamante L. Key facts about us Hispanics and their diverse heritage: Pew research center, 2019. Available: https://pewrsr.ch/31rC4Kx [Accessed 26 Apr 2021].

9 Aguayo-Mazzucato C, Diaque P, Hernandez S, *et al*. Understanding the growing epidemic of type 2 diabetes in the Hispanic population living in the United States. *Diabetes Metab Res Rev* 2019;35:e3097.

10 Vega WA, Rodriguez MA, Gruskin E. Health disparities in the Latino population. *Epidemiol Rev* 2009;31:99–112.

11 Garcia L, Gold EB, Wang L, *et al*. The relation of acculturation to overweight, obesity, pre-diabetes and diabetes among U.S. Mexican-American women and men. *Ethn Dis* 2012;22:58–64.

12 Centers for disease control and prevention. Morbidity and mortality Weekly report — vital signs: leading causes of death, prevalence of diseases and risk factors, and use of health services among Hispanics in the United States — 2009–2013. Available: https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6417a5.htm?s_cid=mm6417a5_w

13 Paskal W, Paskal AM, Dębski T, *et al*. Aspects of modern biobank activity - comprehensive review. *Pathol Oncol Res* 2018;24:771–85.

14 All of Us Research Program Investigators, Denny JC, Rutter JL, *et al*. The "All of Us" Research Program. *N Engl J Med* 2019;381:668–76.

15 Morales J, Glantz N, Larez A, *et al*. Understanding the impact of five major determinants of health (genetics, biology, behavior, psychology, society/environment) on type 2 diabetes in U.S. Hispanic/Latino families: Mil Familias - a cohort study. *BMC Endocr Disord* 2020;20:4.

16 Shaibi GQ, Coletta DK, Vital V, *et al*. The design and conduct of a community-based registry and biorepository: a focus on cardiometabolic health in Latinos. *Clin Transl Sci* 2013;6:429–34.

17 Shaibi G, Singh D, De Filippis E, *et al*. The Sangre POR Salud Biobank: facilitating genetic research in an underrepresented Latino community. *Public Health Genomics* 2016;19:229–38.

18 Shaibi GQ, Kullo IJ, Singh DP, *et al*. Developing a process for returning medically actionable genomic variants to Latino patients in a federally qualified health center. *Public Health Genomics* 2018;21:77–84.

19 U.S. Department of Health and Human Services. National diabetes statistics report. In: *Prevention CfDCa, ED*. Atlanta, GA: centers for disease control and prevention, 2020.

20 Marquez I, Calman N, Crump C. A framework for addressing diabetes-related disparities in US Latino populations. *J Community Health* 2019;44:412–22.

21 Grundy SM, Cleeman JI, Daniels SR, *et al*. Diagnosis and management of the metabolic syndrome: an American heart Association/National heart, lung, and blood Institute scientific statement. *Circulation* 2005;112:2735–52.

22 Stern MP, Gaskill SP, Allen CR, *et al*. Cardiovascular risk factors in Mexican Americans in Laredo, Texas. II. prevalence and control of hypertension. *Am J Epidemiol* 1981;113:556–62.

23 Hanis CL, Ferrell RE, Barton SA, *et al*. Diabetes among Mexican Americans in Starr County, Texas. *Am J Epidemiol* 1983;118:659–72.

24 Wagenknecht LE, Mayer EJ, Rewers M, *et al*. The insulin resistance atherosclerosis study (IRAS) objectives, design, and recruitment results. *Ann Epidemiol* 1995;5:464–72.

25 Raffel LJ, Robbins DC, Norris JM, *et al*. The GENNID study. A resource for mapping the genes that cause NIDDM. *Diabetes Care* 1996;19:864–72.

26 Mitchell BD, Kammerer CM, Blangero J, *et al*. Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. The San Antonio family heart study. *Circulation* 1996;94:2159–70.

27 Bild DE, Bluemke DA, Burke GL, *et al*. Multi-Ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* 2002;156:871–81.

28 Hunt KJ, Lehman DM, Arya R, *et al*. Genome-Wide linkage analyses of type 2 diabetes in Mexican Americans: the San Antonio family Diabetes/Gallbladder study. *Diabetes* 2005;54:2655–62.

29 Abdul-Ghani MA, Tripathy D, DeFronzo RA. Contributions of beta-cell dysfunction and insulin resistance to the pathogenesis of impaired glucose tolerance and impaired fasting glucose. *Diabetes Care* 2006;29:1130–9.

30 Fisher-Hoch SP, Vatcheva KP, Rahbar MH, *et al*. Undiagnosed diabetes and pre-diabetes in health disparities. *PLoS One* 2015;10:e0133135.

31 Gaziano JM, Concato J, Brophy M, *et al*. Million veteran program: a mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol* 2016;70:214–23.