Published in final edited form as:

Public Health Genomics. 2016 ; 19(4): 229–238. doi:10.1159/000447347.

The Sangre Por Salud Biobank: Facilitating Genetic Research in an Underrepresented Latino Community

Gabriel Shaibi^a, Davinder Singh^b, Eleanna De Filippis^c, Valentina Hernandez^b, Bill Rosenfeld^b, Essen Otu^b, Gregorio Montes de Oca^b, Sharon Levey^d, Carmen Radecki Breitkopf^e, Richard Sharp^f, Janet Olson^e, James Cerhan^e, Stephen Thibodeau^g, Erin Winkler^d, and Lawrence Mandarin^h

^aCenter for Health Promotion and Disease Prevention, College of Nursing and Health Innovation, Arizona State University, Phoenix, Ariz

^bMountain Park Health Center, Phoenix, Ariz

^cDivision of Endocrinology, Department of Internal Medicine, Mayo Clinic Arizona, Scottsdale, Ariz

^dCenter for Individualized Medicine, Mayo Clinic Arizona, Scottsdale, Ariz

^eDepartment of Health Sciences Research, Mayo Clinic Rochester, Rochester, Minn

^fBiomedical Ethics Research Program, Mayo Clinic Rochester, Rochester, Minn

^gDepartment of Laboratory Medicine and Pathology, Mayo Clinic Rochester, Rochester, Minn

^hCenter for Disparities in Diabetes, Obesity, and Metabolism, Division of Endocrinology, Department of Medicine, University of Arizona, Tucson, Ariz., USA

Abstract

Background/Aims—The Sangre Por Salud (Blood for Health; SPS) Biobank was created for the purpose of expanding precision medicine research to include underrepresented Latino patients. It is the result of a unique collaboration between Mayo Clinic and Mountain Park Health Center, a federally qualified community health center in Phoenix, Arizona. This report describes the rationale, development, implementation, and characteristics of the SPS Biobank.

Methods—Latino adults (ages 18–85 years) who were active patients within Mountain Park Health Center's internal medicine practice in Phoenix, Ariz., and had no history of diabetes were eligible. Participants provided a personal and family history of chronic disease, completed a sociodemographic, psychosocial, and behavioral questionnaire, underwent a comprehensive cardiometabolic risk assessment (anthropometrics, blood pressure and labs), and provided blood samples for banking. Laboratory results of cardiometabolic testing were returned to the participants and their providers through the electronic health record.

Gabriel Q. Shaibi, PhD, Center for Health Promotion and Disease Prevention, College of Nursing and Health Innovation, Arizona State University, 500 N. 3rd Street, Phoenix, AZ 85004 (USA), gshaibi@asu.edu.

Disclosure Statement

The authors have nothing to disclose.

Results—During the first 2 years of recruitment into the SPS Biobank, 2,335 patients were approached and 1,432 (61.3%) consented to participate; 1,354 (94.5%) ultimately completed all requisite questionnaires and medical evaluations. The cohort is primarily Spanish-speaking (72.9%), female (73.3%), with a mean age of 41.3 ± 12.5 years. Most participants were born outside of the US (77.9%) and do not have health insurance (77.5%). The prevalence of overweight (35.5%) and obesity (45.0%) was high, as was previously unidentified prediabetes (55.9%), type 2 diabetes (7.4%), prehypertension (46.8%), and hypertension (16.2%). The majority of participants rated their health as good to excellent (72.1%) and, as a whole, described their overall quality of life as high (7.9/10).

Conclusion—Collaborative efforts such as the SPS Biobank are critical for ensuring that underrepresented minority populations are included in precision medicine initiatives and biomedical research that seeks to improve human health and reduce the burdens of disease.

Keywords

Biobank; Bioethics; Community-based research; Hispanic populations; Latino; Health disparities

Introduction

Biomedical research has entered a remarkable age of innovation with the emergence of high-speed, low-cost methods for sequencing human DNA. This new capacity is spurring a variety of initiatives in precision medicine focused on identifying early biomarkers of disease and designing treatment plans based on the unique biological features of individual patients [1]. Alongside these remarkable advances is the unfortunate reality that many historically underserved communities continue to struggle with an ever-widening health disparities gap, particularly with respect to advanced medical technologies [2]. If precision medicine is to realize its full potential as a tool for improving human health and reducing the burdens of disease, it will be critical to reach individuals with limited financial resources and other historically underrepresented populations that have difficulty accessing healthcare services.

Ensuring biobanks produce information that is relevant to a diverse population must be coupled with approaches that facilitate integration of precision medicine into diverse clinical practice settings [3]. As the practice of precision medicine becomes more common, it will likely extend beyond academic medical institutions and into community health centers that serve low-income and/or uninsured patients. These community health centers become increasingly important with the Affordable Care Act as they emphasize coordinated primary and preventive care for vulnerable and underserved populations through a comprehensive patient-centered medical home [4].

Potential barriers to ensuring sociodemographic diversity in biobanks are similar to documented challenges for enrolling diverse and vulnerable populations into clinical research studies. These obstacles include disparities in access, health literacy, trust, transparency, communication, and mutual respect [5]. Clinical researchers often collaborate with community agencies and stakeholders to overcome these difficulties in enrolling diverse populations into clinical research studies [6]. Although this model may work well for

recruitment into clinical trials, biobanks are somewhat unique because linking samples with the participant's medical record, preferably within their medical home where they receive regular care, is essential [7]. Therefore, new models of academic-community collaboration to support precision medicine through biobanking may be necessary. Ideally, these new models would include a collaborative governance structure, clearly defined protocols for handling return of results, support for ongoing contact with patients, and a platform for regular engagement and dialogue with the larger community.

The leadership from the Mayo Clinic Biobank and the Center for Individualized Medicine recognized these issues and the need for new biobanking models that can move the field of precision medicine forward. In this report, we describe the process by which the Mayo Clinic undertook a deliberate campaign to extend biobanking to an underrepresented Latino community through collaborating with Mountain Park Health Center (MPHC), a Federally Qualified Health Center that provides primary medical care and behavioral health services to low-income patients in Phoenix, Arizona. The end result is the Sangre Por Salud (Blood for Health; SPS) Biobank that is focused on recruiting and enrolling adult Latino patients from MPHC with no history of diabetes who consent to (1) specimen banking; (2) standardized cardiometabolic disease risk phenotyping; (3) completing questionnaires assessing sociodemographic characteristics, personal and family health history, lifestyle, and other behaviors (e.g. cancer prevention) and exposures (e.g. environmental toxins); (4) accessing data from their medical records, and (5) re-contacting for future studies. With this context, the aims of this paper are to (1) present select sociodemographic and clinical characteristics of the participants enrolled during the first 2 years of the SPS Biobank, and (2) identify practices and procedures that have contributed to the overall success of this unique collaboration.

Methods

Background and History

The Mayo Clinic Center for Individualized Medicine is committed to providing infrastructure for biobanking. The first biobank, the Mayo Clinic Biobank, leverages patient, physician, research, and health systems resources to support a wide array of health-related research studies and has been described in more detail elsewhere [8]. The data in the Mayo Clinic Biobank are rich in terms of health conditions and deep in terms of patient records, as more than half of the participants have more than a 15-year history of medical record data to link with results of studies using their biospecimens or other data. Despite this strength, a limitation of the Mayo Clinic Biobank is the lack of diversity in terms of race/ethnicity (~95% report European ancestry) and indicators of socioeconomic status (e.g. >80% have at least an associates degree). Recognizing the need to enhance diversity within the Mayo Clinic Biobank in order to produce more generalizable information that could be applied to various practice settings, the Center for Individualized Medicine at the Mayo Clinic looked beyond their patient-care network to establish a new model for biobanking. This model capitalized on the successful experience and infrastructure in order to extend precision medicine research to an underrepresented Latino patient population served by a Federally Qualified Community Health Center.

Establishing the Model

Prior to developing the SPS Biobank, an established partnership was in place between Mayo Clinic and MPHC; this partnership included research as well as patient care. A master agreement between the two institutions was in place and allowed for MPHC to rely on the Mayo Clinic Institutional Review Board (IRB) as the IRB of record. MPHC leadership and the research team recognized the importance of comprehensively integrating the SPS Biobank into the MPHC practice so that a separate agreement was needed to support this collaborative endeavor. In order to further embed the SPS Biobank within the practice and ensure its success, it was decided that (1) MPHC staff would become members of the research 'core team'; (2) only active MPHC patients would be enrolled into the SPS Biobank, and (3) the project needed to bring direct benefit to MPHC and their patients. To accomplish this, sample collection for banking would occur in the context of a routine clinical blood draw ordered by MPHC providers. SPS Biobank also provided funding for selected clinical labs [complete blood count with differential, complete metabolic panel, lipid panel, hemoglobin A1c (HbA1c)] ordered during this blood draw that were not covered by the patient's insurance. These labs were decided upon collaboratively by the SPS research team and MPHC staff to provide a risk profile for cardiovascular disease and type 2 diabetes, the major health issues for this population. This was also perceived as a benefit by MPHC leadership, as their providers are often faced with the challenge of ordering medical tests that patients cannot afford. As a consequence, those orders often go unfilled and the tests are not performed. By providing funding for these tests, the SPS Biobank would gain the benefit that all participants enrolled in the biobank would receive a set of tests that would form the basis for a standardized phenotype (described below). This was intended to extend and complement the Electronic Health Record. The rationale for the focus on Latino patients was based on three points: (1) the widespread underrepresentation of Latinos in clinical research and biobank studies; (2) the predominance of Latino patients served by MPHC (>70% of active patients), and (3) the extensive experience of the research team in working with Latino participants in clinical studies.

Once the conceptual framework for integrating the SPS Biobank into MPHC's clinical flow was established, attention was turned to developing the protocol and obtaining IRB approval. The protocol needed to take into account operational logistics for (1) handling research specimens that needed to be transported by courier ~ 30 miles each study day for processing in the Mayo Clinic Biospecimen Accessioning and Processing Core laboratory; (2) transferring clinical data from the MPHC electronic health record to a research database housed within Mayo Clinic, and (3) training and certifying MPHC personnel in the responsible conduct of research. As part of the protocol, a detailed health history form from the Mayo Clinic Biobank questionnaire was adapted and modified to include pertinent constructs and measures that are culturally and environmentally congruent with the local Latino population of Phoenix, Ariz. Adaptation of the self-administered questionnaire and selection of appropriate measures was conducted collaboratively between the institutions, with particular consideration given to the issue of available and psychometrically sound Spanish-language translations. In addition, a comprehensive family history questionnaire was developed that specifically targets common, chronic health conditions in the population, type 2 diabetes, cardiovascular disease, and cancer. It was decided that this particular

instrument would be interviewer administered in order to develop the most accurate family pedigrees as possible. Overall quality of life (QoL) rating was obtained by a single question, 'How would you describe your overall quality of life?', rating 0 ('as bad as it can be') to 10 ('as good as it can be') [9].

Recruitment and Enrollment

In order to facilitate enrollment and minimize any negative impact on clinical flow, the SPS Biobank research team and MPHIC providers held several discussions to identify a recruitment strategy that would work efficiently within the MPHIC internal medicine practice. The mechanism that was deemed to work most efficiently was for biobank staff to screen daily patient logs for general inclusion/exclusion criteria (e.g., age, ethnicity, health status) and alert MPHIC providers and medical assistants that a patient may be eligible. During a visit, the MPHIC medical providers offered potentially eligible patients a very brief overview of the biobank and asked if they wanted to learn more. Patients who expressed interest in hearing more were then referred to SPS Biobank research staff. Eligibility criteria included self-reported Latino ethnicity, age 18–85 years, and classification as an active MPHIC patient with a recorded visit within 6 months of biobank enrollment. Patients were excluded if they (1) had a confirmed diagnosis of type 2 diabetes; (2) had a history of cancer in the past 3 years (excluding non-melanoma skin cancers), patients in remission >3 years were eligible; (3) were pregnant, had delivered a child in the last 12 months, or were breastfeeding within the last 3 months, or (4) were unable to abstain from smoking during the duration of the study visit (~ 4 h). Patients who met these criteria and expressed interest were referred to a clinical research coordinator for a full explanation of study procedures and, if interested in participating, were scheduled for administration of informed consent. Informed consent was obtained by bilingual/bicultural research coordinators for the collection of metabolic, anthropometric, sociodemographic, and family history data, as well as 50 ml of blood for biobanking, access to their medical record, the ability to recontact in the future. Coordinators were trained by an experienced researcher with a long track record of recruiting and enrolling Latino participants into clinical research studies. Participants were provided the consent to read over, and then the study purpose, procedures, risks, benefits, and alternative approaches were explained in detail by the coordinator in the participants' preferred language. Any questions were answered prior to signing the consent. All documents were available in English and Spanish with translations approved by MPHIC staff prior to submission to the IRB in order to ensure attention to the local Spanish dialect. All procedures were approved by the Mayo Clinic IRB, and a Certificate of Confidentiality from NIH was secured.

Data Collection

Participants were scheduled for the study visit in the morning after an overnight fast. Height, weight, waist circumference, seated blood pressure, pulse, and temperature (to rule out acute illness) were measured. A fasting blood draw was collected for complete blood count with differential, complete metabolic panel, lipid panel (triglycerides, total, LDL, HDL cholesterol), HbA1c, and glucose. Samples for banking were collected simultaneously with clinical samples. After fasting samples were drawn, participants consumed 75 g of glucose in solution for the oral glucose tolerance test (OGTT). An OGTT is considered the gold-

standard for diagnosing type 2 diabetes and may provide additional risk stratification above and beyond fasting glucose and HbA1c [10]. Given the disproportionate rates of diabetes in the Latino community and the overrepresentation of diabetic patients in the MPHHC practice, the OGTT was perceived as an added value by MPHHC staff and would allow participants sufficient time to complete the family history and questionnaire. Laboratory analysis for clinical testing was performed by the commercial laboratory used by MPHHC to facilitate integration of results into the electronic medical record and ensure appropriate communication and follow-up with a provider in a timely manner.

Access to the SPS Biobank Data and Specimens

Investigator access to samples (DNA, RNA, serum, plasma) or data from the SPS Biobank or re-contact participants follows a 5-step process (fig. 1). Step 1 is a written access request submitted to the SPS biobank. This step may include input from the research coordinators and/or staff from MPHHC. Step 2 is a review by the SPS Biobank Core Team (MPHHC leadership and staff, along with investigators and Mayo Clinic administrators) to determine whether the request is congruent with MPHHC's mission. MPHHC's mission includes working with the communities it serves to sustain and improve health by providing affordable primary care. This review allows MPHHC the opportunity to clarify what impact the request may have on their practice, patient populations, and/or reputation and occurs prior to reviews by the Mayo Clinic Biobank Access Committee or the IRB. Importantly, this step allows MPHHC to identify requests that might be perceived to result in harm to the MPHHC practice or community trust and have input on modifying such requests in a way that would mitigate risks. Investigators are provided with MPHHC's mission, encouraged to visit the MPHHC website, and asked to describe any benefits their project may have to MPHHC and/or their patients. These benefits can include immediate, long-term, or potential benefits. In addition to gaining access to specimens and data from the SPS Biobank, this process may encourage investigators to work with MPHHC on collaborative research and clinical projects that will further benefit the practice and patient population. Requests that present potentially challenging issues could be taken to a Community Advisory Board (CAB; described below) for feedback, discussion, and advice. Step 3 is a 'screening meeting' to collect study specific details and to ensure a complete understanding of the logistics, timeline and study costs. Step 4 is a review by the Mayo Clinic Biobank Access Committee comprised of researchers, providers, and staff from both institutions. If approved by the Access Committee, the project is submitted to the Mayo Clinic IRB for final approval (step 5). The above process mirrors the process used by the Mayo Clinic Biobank, with the additional step described in step 2 for MPHHC 'mission fit' review. MPHHC is involved in all 5 steps of this process.

Community Advisory Board

In parallel with establishing an access process, both the SPS research team and MPHHC agreed that community input would be essential for transparency, broader engagement, shared decision making for complex ethical issues, and overall governance of the biobank. To facilitate these goals, a CAB was established through MPHHC. A bilingual/bicultural CAB coordinator was hired by MPHHC and charged with recruitment of members that would be comprised of MPHHC patients, SPS Biobank participants, community members, and other stakeholders. CAB meetings have been held every other month at MPHHC and include

general research education, updates on the progress of the SPS Biobank, presentations by researchers with active SPS Biobank protocols, and opportunities for the CAB to provide feedback and advice. Case studies and real life examples are used to stimulate discussion among CAB members. CAB meetings are conducted in Spanish since many members are monolingual Spanish-speaking, with simultaneous translation provided for attendees and non-Spanish speaking presenters. Although the CAB can play a role in providing input on perceived benefits, risks, and the impact of research on the community, they serve as an advisory, not approval, board.

Statistical Analysis

Descriptive data are presented as percentages for categorical data or means $\pm$ standard deviations for continuous data. Sex differences were assessed by χ^2 and independent sample t tests or Wilcoxon rank sum. To describe the diversity between the SPS Biobank and the Mayo Clinic Biobank, select demographics variables (age, sex, weight status, education level, self-rated health, and QoL) were compared between the two biobank populations. To compare SPS with the Mayo Clinic Biobank, categorical variables were compared by χ^2 , while continuous variables were evaluated by independent sample t tests or Wilcoxon rank sum.

Results

During the first 2 years of enrollment, 2,335 MPHC patients met general inclusion/exclusion criteria as determined from a review their medical record. These individuals were subsequently approached and 1,432 (61.3%) consented to participate. Of those who consented, 1,354 (94.5%) were phenotyped for cardiometabolic risk, provided samples for banking, and completed the family health history as well as the sociodemographic questionnaires. One participant withdrew consent within 1 week of completing the study visit, and these data are not included in this report. Men and women in the SPS Biobank were similar in terms of BMI, age, education levels, self-rated health, and QoL (table 1). Overall, the SPS Biobank cohort is primarily Spanish-speaking (72.9%), female (73.3%), with a mean age of 41.3 ± 12.5 years. Most of the participants were born outside of the US (77.9%) and did not have health insurance (77.5%) despite varying employment status (table 2). The prevalence of overweight and obesity was high (35.5 and 45.0%, respectively) as were the percentages of patients with previously unidentified prediabetes or type 2 diabetes (55.9 and 7.4%, respectively) and prehypertension or hypertension (46.8 and 16.2%, respectively). The majority of participants rated their health as good to excellent (72.1%) and, as a whole, described their overall QoL near the positive end on the Likert scale (mean QoL 7.9/10). These findings were corroborated in the relatively low prevalence of the top 10 health conditions reported in the cohort (table 3) while family histories suggest that the risk for future chronic disease was likely high (table 4).

Comparison of select sociodemographic and phenotypic data between the SPS Biobank and the Mayo Clinic Biobank patients is presented in table 5. The SPS cohort was younger ($p < 0.0001$), more likely to be female ($p < 0.0001$), had a lower proportion of lean and higher proportion of obese patients ($p < 0.0001$), completed less formal education ($p < 0.0001$), and

reported lower ratings of health ($p < 0.0001$). The cohorts were nearly identical in terms of their reported QoL ($p > 0.05$).

Discussion

The SPS Biobank is an infrastructure project that was established to support precision medicine research among Latinos by (1) linking Mayo Clinic Biobank resources with a Latino-serving Federally Qualified Community Health Center, and (2) fostering research opportunities for Latino patients. The initial 2-year period supports the ability to identify, recruit, and enroll a patient population that is historically underrepresented in the biobanking and genomic literature [11] and is disproportionately impacted by a number of health conditions that have a genetic component [12]. As such, the SPS Biobank could serve as a model to facilitate precision medicine and genomic research in a highly vulnerable and understudied population.

Data from the SPS Biobank highlight significant differences between SPS participants and those in the Mayo Clinic Biobank. One of the more striking differences between the two populations is the considerably younger age of the SPS cohort. This age difference mirrors population demographics across the US, where Latinos are the youngest population subgroup and non-Hispanic whites are the oldest [13]. The demographic landscape of the US is changing, and Latinos accounted for over 50% of the US growth rate between 2000 and 2010 [14]. Further, new immigrants and their children have been estimated to account for over 80% of US population growth between 2005 and 2050 [15]. This changing demographic landscape will likely lead to shifts in attitudes and access to health care, the types and rates of disease, as well as the genomic diversity of the US population. In anticipation of these changes, precision medicine initiatives that include sociodemographic and genomic diversity may be better able to address the health disparities gap. In order to support efforts to diversify biobanks by recruiting and enrolling vulnerable and underserved populations, we have found that a successful strategy is embed in the local health care center, hire bilingual and bicultural research staff who are familiar with the local community (important for building trust in order to capture sensitive sociodemographic data in a meaningful manner), and address an actionable health need directly relevant to the community as part of the enrollment process (here, cardiometabolic and diabetes testing).

In addition to age, income and education play a major role in population-level health [16]. Approximately 50% of the Mayo Clinic Biobank cohort completed college or graduate school compared to 6% in the SPS cohort. Educational attainment is a strong protective factor against mortality in the US [17] and, in particular, among highly educated individuals [18]. While it would seem that more educated populations are universally healthier, less educated Latinos experience a longer life expectancy than non-Hispanic whites [19]. This so-called ‘Latino Paradox’ may offer insight into resiliency factors that could be applied to other populations to promote health and prevent disease. Although we did not directly compare income between the two cohorts, using insurance coverage as a surrogate would suggest that the SPS Biobank cohort with 77.5% uninsured patients is at an income disadvantage compared to the predominantly insured Mayo Clinic Biobank cohort. This income discrepancy is particularly relevant in the context of increasing costs for medical

testing and treatment. As advances in medical technologies lead to better diagnostic procedures and individualized care, finding mechanisms to apply precision medicine to low income populations will require innovative programs that incentivize cost-effective approaches to ensure that all populations benefit from biomedical advances.

Despite sociodemographic differences between SPS and Mayo Clinic Biobank participants, it is noteworthy that, as a whole, SPS participants reported self-ratings of health and QoL similar to participants in the Mayo Clinic Biobank. Self-rated health is predictive of future health outcomes [20] independent of racial and ethnic background [21]. However, the predictive value of self-rated health for future health trajectories may be partially mediated by genetic factors [22]. Therefore, integrating self-ratings of health with genomic data may strengthen the predictive value of these seemingly disparate measures. With this context, biobanks that integrate various types of data (e.g., genomic, family history, behavior, sociodemographic, and perceived health, etc.) may be best situated to support prevention, early identification, and treatment models of precision medicine.

The Mayo Clinic and MPHC share a common philosophy for coordinated, patient-centered health care and have a long history of partnership. The common philosophy and track record were essential for building trust within MPHC's staff and patient population who have little exposure to academic biomedical research opportunities. Additional considerations that may further support collaborative efforts include, the integration of MPHC staff into the research team to provide input on the design, implementation, and ongoing refinement of the project, developing mutually identified goals, coordinated planning, frequent and ongoing dialogue with staff and providers not involved in the biobank, shared governance structure, and broad community input. This model of collaboration may serve as an ideal framework to leverage existing biobanking expertise and samples from large academic medical institutions to smaller institutions and community health practices that serve diverse and underrepresented populations. Collaboration may be a more cost-effective mechanism for enriching relatively homogeneous biobanks with genomic and sociodemographic diversity. To date, genomic studies have not resulted in convincing evidence that genetic differences explain racial disparities in chronic disease [23]. There are several hypotheses as to why this is the case, but a glaring limitation across these studies is the considerably larger sample size of cohorts enriched with European ancestry compared to other racial and ethnic groups [24]. Moreover, disease and disease risk phenotypes in many cases are likely to represent an interaction between genomic diversity and environmental differences. In the case of communities with health disparities such as the community from which SPS Biobank patients are drawn, contextual factors including education, health literacy, income level, and cultural beliefs are also important. Therefore, prioritizing not only genomic but also sociodemographic and cultural diversity within existing, large-scale biobanks will provide an opportunity to leverage infrastructure and capitalize on current and planned high-throughput sequencing activities that may result in more clinically relevant findings.

Another important lesson from the first 2 years of experience with the SPS collection is the importance of promoting community engagement. To sustain successful biobanks over time, both institutional and community support is critical. The complexity of biobanking and the implementation of precision medicine may be poorly understood and can create concerns,

particularly among members of historically underrepresented communities [25]. Both the Mayo Clinic collection and the SPS biobank have sought to address these challenges through the creation of CABs. This priority has proven to be an essential aspect of the SPS Biobank and has contributed to increased awareness and support for biobanking research in the local communities served by MPHIC. The CAB has provided advice on topics ranging from re-contacting previous participants to policies regarding withdrawal of consent. The CAB has also hosted Mayo Clinic researchers who have presented results of their work involving biobank samples, which has led to discussion of how their work may be expanded to have a greater impact in the Latino community. Although these activities have a regional focus, the hope is to share these experiences with others, highlighting potential opportunities to expand this model in other communities [26]. Efforts are also underway to coordinate CAB activities across regional biobank sites, to foster a broader sense of community, share perspectives across communities, and promote dialogue among multiple community advisors.

Strengths of the SPS Biobank include (1) the focus on an underrepresented population in biomedical research; (2) leveraging the experience and infrastructure of the Mayo Clinic Biobank; (3) embedding the biobank in a federally qualified community health center, and (4) nurturing a robust local CAB. Despite these strengths, there are limitations that are worthy of comment. First, the SPS Biobank is not intended to be representative of all Latinos nor is it representative of the local community. By design, we recruited otherwise healthy, non-pregnant Latino patients, but the MPHIC practice is more diverse, with a large proportion of diabetic patients as well as pregnant and postpartum women. In addition, we excluded patients below age 18 despite the fact that children make up approximately 45% of the MPHIC patient population. Second, the overall sample size is relatively small compared to other diverse biobanks, but the size and pace of enrollment will increase over time. Third, despite being embedded within the patient's medical home, the richness of preexisting health information included within the medical record is likely limited as the MPHIC practice only recently transitioned to an electronic medical record system. Fourth, we have not included a formal process for evaluating the collaborative project from a broader institutional or community perspective. Therefore, the ability to objectively identify key aspects for success or opportunities in need of further refinement is limited to speculation. Lastly, the purely descriptive nature of the dataset that is restricted to a subset of select clinical and sociodemographic variables is limiting. Further, since genomic sequencing is not a standard procedure within the SPS Biobank, we have included the three most prevalent and preventable diseases in the population. We acknowledge that there are limitations in the accuracy of family history data that could be minimized by enrolling family members. As future studies utilize the infrastructure created by the SPS Biobank and employ additional methodologies, interventions, or follow-up data points, these elements will be integrated to provide a more comprehensive picture of the cohort and the individual patients enrolled.

Conclusion

The SPS Biobank was created to promote diversity among participants in precision medicine research and provide opportunities for low-income Latino patients to participate in biomedical research. Efforts to foster transdisciplinary research projects and expand the SPS

Biobank are underway and will likely yield unique opportunities to address health disparities in Latino communities. As these efforts materialize, the long-term success of the SPS Biobank will be fully appreciated by advancements in the field and measurable improvements in health outcomes. Although the success of the infrastructure will be measured in the future, we believe the lessons learned in developing the SPS Biobank offer implications for future work to build upon. These implications include the importance of follow-up collections, opportunities to recruit family members for an enhanced family history with more complete pedigrees, enhanced phenotyping beyond what may be in the medical record, initiation of genomic sequencing, the assessment of health and genomic literacy, the collection of relevant psychosocial and behavioral measures, and a formal evaluation of the process by which participants are enrolled and institutions collaborate to diversify precision medicine efforts.

Acknowledgments

The SPS Biobank is made possible by research support from the Mayo Clinic Center for Individualized Medicine. We are indebted to the patients who contributed samples, the MPHIC providers and staff who support this initiative, the SPS CAB for their commitment and dedication to their community, and the staff of the Biospecimen Accessioning and Processing Core. We thank M'hamed (Hamy) Temkit for his help with data analysis, Jeannette Mendez, Monica Garcia, Cynthia Ramirez, and Elsa Palomar, for their help with study coordination, recruitment, and enrolling patients, and Guadalupe Estrada for her help with establishing the CAB.

References

1. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015; 372:793–795. [PubMed: 25635347]
2. Dankwa-Mullan I, Rhee KB, Stoff DM, Pohlhaus JR, Sy FS, Stinson N Jr, Ruffin J. Moving toward paradigm-shifting research in health disparities through translational, transformational, and transdisciplinary approaches. *Am J Public Health*. 2010; 100(suppl 1):S19–S24. [PubMed: 20147662]
3. Haga SB. Impact of limited population diversity of genome-wide association studies. *Genet Med*. 2010; 12:81–84. [PubMed: 20057316]
4. Kulesher RR. Health reform's impact on federally qualified community health centers: the unintended consequence of increased Medicaid enrollment on the primary care medical home. *Health Care Manag*. 2013; 32:99–106.
5. George S, Duran N, Norris K. A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. *Am J Public Health*. 2014; 104:e16–e31.
6. De las Nueces D, Hacker K, DiGirolamo A, Hicks LS. A systematic review of community-based participatory research to enhance clinical trials in racial and ethnic minority groups. *Health Serv Res*. 2012; 47:1363–1386. [PubMed: 22353031]
7. Belmont J, McGuire AL. The futility of genomic counseling: essential role of electronic health records. *Genome Med*. 2009; 1:48. [PubMed: 19439060]
8. Olson JE, Ryu E, Johnson KJ, Koenig BA, Maschke KJ, Morrisette JA, Liebow M, Takahashi PY, Fredericksen ZS, Sharma RG, Anderson KS, Hathcock MA, Carnahan JA, Pathak J, Lindor NM, Beebe TJ, Thibodeau SN, Cerhan JR. The Mayo Clinic Biobank: a building block for individualized medicine. *Mayo Clin Proc*. 2013; 88:952–962. [PubMed: 24001487]
9. Coates A, Dillenbeck CF, McNeil DR, Kaye SB, Sims K, Fox RM, Woods RL, Milton GW, Solomon J, Tattersall MH. On the receiving end—II. Linear analogue self-assessment (LASA) in evaluation of aspects of the quality of life of cancer patients receiving therapy. *Eur J Cancer Clin Oncol*. 1983; 19:1633–1637. [PubMed: 6315445]

10. Sacks DB. A1c versus glucose testing: a comparison. *Diabetes Care*. 2011; 34:518–523. [PubMed: 21270207]
11. Duran DG. Lack of Hispanics' involvement in research – is it Hispanics or scientists? *Community Genet*. 1998; 1:183–189. [PubMed: 11657308]
12. Vega WA, Rodriguez MA, Gruskin E. Health disparities in the Latino population. *Epidemiol Rev*. 2009; 31:99–112. [PubMed: 19713270]
13. US Census Bureau PD. Annual estimates of the resident population by sex, age, race, and Hispanic origin for the United States and States: April 1, 2010 to July 1, 2014. 2015
14. Passel, JS., Cohn, D., Lopez, MH. Hispanics Account for More Than Half of Nation's Growth in Past Decade. Washington: Pew Research Center; 2011.
15. Passel, JS., Cohn, D. US Population Projections: 2005–2050. Washington: Pew Research Center; 2008.
16. Backlund E, Sorlie PD, Johnson NJ. A comparison of the relationships of education and income with mortality: The National Longitudinal Mortality Study. *Soc Sci Med*. 1999; 49:1373–1384. [PubMed: 10509827]
17. Krueger PM, Tran MK, Hummer RA, Chang VW. Mortality attributable to low levels of education in the United States. *PLoS One*. 2015; 10:e0131809. [PubMed: 26153885]
18. Hayward MD, Hummer RA, Sasson I. Trends and group differences in the association between educational attainment and US. Adult mortality: implications for understanding education's causal influence. *Soc Sci Med*. 2015; 127:8–18. [PubMed: 25440841]
19. Ruiz JM, Steffen P, Smith TB. Hispanic mortality paradox: a systematic review and meta-analysis of the longitudinal literature. *Am J Public Health*. 2013; 103:e52–e60.
20. DeSalvo KB, Bloser N, Reynolds K, He J, Muntner P. Mortality prediction with a single general self-rated health question. A meta-analysis. *J Gen Intern Med*. 2006; 21:267–275. [PubMed: 16336622]
21. Lima-Costa MF, Macinko J, Mambrini JV, Cesar CC, Peixoto SV, Magalhaes WC, Horta BL, Barreto M, Castro-Costa E, Firmo JO, Proietti FA, Leal TP, Rodrigues MR, Pereira A, Tarazona-Santos E. Genomic ancestry, self-rated health and its association with mortality in an admixed population: 10 year follow-up of the Bambui-Epigen (Brazil) cohort study of ageing. *PLoS One*. 2015; 10:e0144456. [PubMed: 26680774]
22. Harris SE, Hagenaars SP, Davies G, Hill WD, Liewald DC, Ritchie SJ, Marioni RE, Sudlow CL, Wardlaw JM, McIntosh AM, Gale CR, Deary IJ. Molecular genetic contributions to self-rated health. *bioRxiv*.
23. Kaufman JS, Dolman L, Rushani D, Cooper RS. The contribution of genomic research to explaining racial disparities in cardiovascular disease: a systematic review. *Am J Epidemiol*. 2015; 181:464–472. [PubMed: 25731887]
24. Haiman CA, Stram DO. Exploring genetic susceptibility to cancer in diverse populations. *Curr Opin Genet Dev*. 2010; 20:330–335. [PubMed: 20359883]
25. Burgess MM. From 'trust us' to participatory governance: deliberative publics and science policy. *Public Underst Sci*. 2014; 23:48–52. [PubMed: 24434712]
26. Allyse MA, McCormick JB, Sharp RR. Prudentia populo: involving the community in biobank governance. *Am J Bioethics*. 2015; 15:1–3.

Fig. 1.

Approval process for accessing data, samples, and re-contact of participants in the SPS Biobank. Mission fit: the SPS Core team determines whether the access request is congruent with the mission of MPHC, and this provides MPHC the opportunity to clarify what impact the request may have on their practice, patient population, and/or reputation. CAB input: CAB input may be sought if broader community input regarding the science, ethics, or operations is warranted before initiating the project.

Table 1

Characteristics of SPS Biobank participants by gender

	Female (n = 992)	Male (n = 362)	Total (n = 1,354)	p value
BMI	30.0±6.6	29.3±5.4	29.8±6.3	0.207
Age, years (range = 18–85 years)	41.3±12.2	41.6±13.5	41.4±12.6	0.935
School years completed				0.945
Missing	107	28	135	
Never attended school or only attended kindergarten	25 (2.8)	12 (3.6)	37 (3.0)	
Grades 1 through 8 (elementary)	263 (29.7)	105 (31.4)	368 (30.2)	
Grades 9 through 11 (some high school)	227 (25.6)	83 (24.9)	310 (25.4)	
Grade 12 or GED (high school graduate)	182 (20.6)	68 (20.4)	250 (20.5)	
College 1 year to 3 years (some college)	115 (13.0)	42 (12.6)	157 (12.9)	
College 4 years or more (college graduate)	39 (4.4)	12 (3.6)	51 (4.2)	
Graduate or professional school	20 (2.3)	5 (1.5)	25 (2.1)	
Other	14 (1.6)	7 (2.1)	21 (1.7)	
Self-rated general health (1 = excellent, 5 = poor)	3.0±0.9	2.9±0.9	3.0±0.9	0.572
Overall QoL (0 = as bad as it can be, 10 = as good as it can be)	7.9±1.8	7.8±1.7	7.9±1.7	0.113

Data are presented as mean ± SD or n (%).

Table 2

Selected sociodemographic characteristics

	Total (n = 1,354)
Born in the US	
Missing	157
No	932 (77.9%)
Yes	265 (22.1%)
Years living in the US	
Missing	422
0 – 5	21 (2.3%)
5 – 10	115 (12.3%)
10 – 15	248 (26.6%)
>15	548 (58.8%)
Consented language	
English	366 (27.1%)
Spanish	986 (72.9%)
Language(s) spoken at home	
Missing	144
Only Spanish	548 (45.3%)
More Spanish than English	360 (29.8%)
Both equally	138 (11.4%)
More English than Spanish	104 (8.6%)
Only English	60 (5.0%)
Employment status	
Missing	246
Working full time for pay (35 or more hours a week)	415 (37.5%)
Working part-time for pay	201 (18.1%)
Not working for pay at present	492 (44.4%)
Insurance coverage	
Missing	76
Uninsured	990 (77.5%)
Medicaid	123 (9.6%)
Medicare	31 (2.4%)
Commercial	134 (10.5%)

Table 3

Frequency of top 10 self-reported health conditions

Rank	Self-reported health condition	n	%
1	Depression	210	15.5
2	High cholesterol (hyperlipidemia)	206	15.2
3	Abnormal distance vision	203	15.0
4	High blood pressure (hypertension)	197	14.5
5	Anxiety	161	11.9
6	Migraine headaches	133	9.8
7	Asthma	101	7.5
8	Arthritis (osteoarthritis)	88	6.5
9	Acid reflux and gastroesophageal reflux	84	6.2
10	Arthritis (rheumatoid)	76	5.6

Table 4

Family history of chronic disease

<i>Cancer</i>	
Maternal	
No	1,232 (91.0%)
Yes	122 (9.0%)
Paternal	
No	1,253 (92.5%)
Yes	101 (7.5%)
Sibling	
No	1,225 (90.5%)
Yes	129 (9.5%)
<i>Diabetes</i>	
Maternal	
No	1,007 (74.4%)
Yes	347 (25.6%)
Paternal	
No	1,110 (82.0%)
Yes	244 (18.0%)
Sibling	
No	1,042 (77.0%)
Yes	312 (23.0%)
<i>Cardiovascular disease</i>	
Maternal	
No	954 (70.5%)
Yes	400 (29.5%)
Paternal	
No	1,093 (80.7%)
Yes	261 (19.3%)
Sibling	
No	1,076 (79.5%)
Yes	278 (20.5%)

Table 5

Comparison of selected characteristics between SPS and Mayo Clinic Biobank

	SPS Biobank (n = 1,354)	Mayo Clinic Biobank (n = 50,231)
Gender		
Female	992 (73.3%)	29,496 (58.7%)
Male	362 (26.7%)	20,735 (41.3%)
Age category		
18 – 44 years	805 (59.5%)	7,977 (15.9%)
45 – 54 years	306 (22.6%)	7,851 (15.6%)
55 – 64 years	127 (9.4%)	11,869 (23.6%)
65 years	116 (8.6%)	22,534 (44.9%)
Weight status		
Missing	2	13,908
Normal weight (BMI < 25)	263 (19.5%)	12,680 (34.9%)
Overweight (BMI 25 – 30)	480 (35.5%)	12,817 (35.3%)
Obese (BMI ≥ 30)	609 (45.0%)	10,826 (29.8%)
School years completed		
Missing	136	242
8th grade or less	414 (33.1%)	255 (0.5%)
Grades 9 through 11 (some high school)	310 (25.5%)	673 (1.3%)
Grade 12 or GED (high school graduate)	250 (20.5%)	7,432 (14.9%)
College 1 year to 3 years (some college)	157 (12.9%)	15,972 (32.0%)
College 4 years or more (college graduate)	51 (4.2%)	12,539 (25.1%)
Graduate or professional school	25 (2.1%)	12,535 (25.1%)
Other	21 (1.7%)	583 (1.2%)
Self-rated general health		
Missing	112	174
Excellent	80 (6.4%)	6,776 (13.5%)
Very good	241 (19.4%)	20,099 (40.2%)
Good	575 (46.3%)	16,632 (33.2%)
Fair	318 (25.6%)	5,330 (10.6%)
Poor	28 (2.3%)	1,220 (2.4%)
Overall QoL (mean ± SD)	7.9 ± 1.7	7.8 ± 1.8