Building national patient registries in Mexico: insights from the MexOMICS Consortium

Paula Reyes-Pérez[1†], Ana Laura Hernández-Ledesma[1†],
Talía V. Román-López[2†], Brisa García-Vilchis[3],
Diego Ramírez-González[2], Alejandra Lázaro-Figueroa[3],
Domingo Martinez[1,4,5], Victor Flores-Ocampo[1],
Ian M. Espinosa-Méndez[2], Lizbet Tinajero-Nieto[6],
Angélica Peña-Ayala[6,7], Eugenia Morelos-Figaredo[8],
Carlos M. Guerra-Galicia[9], Estefania Torres-Valdez[10],
María Vanessa Gordillo-Huerta[11], Nadia A Gandarilla-Martínez[12],
Karla Salinas-Barboza[13], Guillermo Félix-Rodríguez[14],
Gabriel Frontana-Vázquez[10], Yamil Matuk-Pérez[15],
Ingrid Estrada-Bellmann[16], Deshiré Alpizar-Rodríguez[17],
Mayela Rodríguez-Violante[18], Miguel E. Rentería[19,20]*,
Alejandra E. Ruíz-Contreras[3]*, Sarael Alcauter[2]* and
Alejandra Medina-Rivera[1]*

[1]Laboratorio Internacional de Investigación Sobre el Genoma Humano, Universidad Nacional
Autónoma de México, Santiago de Querétaro, Mexico, [2]Departamento de Neurobiología Conductual y
Cognitiva, Instituto de Neurobiología, Universidad Nacional Autónoma de México, Santiago de
Querétaro, Mexico, [3]Laboratorio de Neurogenómica Cognitiva, Unidad de Investigación de
Psicobiología y Neurociencias, Coordinación de Psicobiología y Neurociencias, Facultad de Psicología,
Universidad Nacional Autónoma de México, Ciudad de México, Mexico, [4]Unidad de Genómica
Avanzada, Langebio, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional,
Irapuato, Mexico, [5]Escuela Nacional de Estudios Superiores, Unidad Juriquilla, Universidad Nacional
Autónoma de México, Santiago de Querétaro, Mexico, [6]Hospital General Regional No. 1, Instituto
Mexicano del Seguro Social, Querétaro, Santiago de Querétaro, Mexico, [7]Instituto Nacional de
Rehabilitación "Luis Guillermo Ibarra Ibarra", Ciudad de México, Mexico, [8]Hospital Regional, Instituto de
Seguridad y Servicios Sociales de los Trabajadores del Estado, Morelia, Mexico, [9]Neurociencias PRISMA
AC, San Luis Potosí, Mexico, [10]Hospital General Regional No. 2, Instituto Mexicano del Seguro Social, El
Marqués, Mexico, [11]Hospital General Querétaro, Instituto de Seguridad y Servicios Sociales de los
Trabajadores del Estado, Santiago de Querétaro, Mexico, [12]Departament of Neurology, Neurological
Center, ABC Medical Center, México City, México, [13]Hospital General de México, Ciudad de México,
Mexico, [14]Hospital Star Médica Querétaro, Santiago de Querétaro, Mexico, [15]Facultad de Medicina,
Universidad Autónoma de Querétaro. Unidad de Neurociencias, Hospital Angeles Centro Sur, Santiago
de Querétaro, Mexico, [16]Movement Disorders Clinic, Neurology Division, Internal Medicine
Department, University Hospital "Dr. José E. González", Universidad Autónoma de Nuevo León,
Monterrey, Mexico, [17]Research Unit, Colegio Mexicano de Reumatología, Ciudad de México, Mexico,
[18]Laboratorio Clínico de Enfermedades Neurodegenerativas, Instituto Nacional de Neurología y
Neurocirugía Manuel Velasco Suárez, Mexico City, Mexico, [19]Mental Health and Neuroscience Program,
QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia, [20]School of Biomedical Sciences,
Faculty of Medicine, The University of Queensland, Brisbane, QLD, Australia

Objective: To introduce MexOMICS, a Mexican Consortium focused on
establishing electronic databases to collect, cross-reference, and share health-
related and omics data on the Mexican population.

Methods: Since 2019, the MexOMICS Consortium has established three
electronic-based registries: the Mexican Twin Registry (TwinsMX), Mexican
Lupus Registry (LupusRGMX), and the Mexican Parkinson's Research Network
(MEX-PD), designed and implemented using the Research Electronic Data
Capture web-based application. Participants were enrolled through voluntary

participation and on-site engagement with medical specialists. We also acquired DNA samples and Magnetic Resonance Imaging scans in subsets of participants.

Results: The registries have successfully enrolled a large number of participants from a variety of regions within Mexico: TwinsMX ($n = 2,915$), LupusRGMX ($n = 1,761$) and MEX-PD ($n = 750$). In addition to sociodemographic, psychosocial, and clinical data, MexOMICS has collected DNA samples to study the genetic biomarkers across the three registries. Cognitive function has been assessed with the Montreal Cognitive Assessment in a subset of 376 MEX-PD participants. Furthermore, a subset of 267 twins have participated in cognitive evaluations with the Creyos platform and in MRI sessions acquiring structural, functional, and spectroscopy brain imaging; comparable evaluations are planned for LupusRGMX and MEX-PD.

Conclusions: The MexOMICS registries offer a valuable repository of information concerning the potential interplay of genetic and environmental factors in health conditions among the Mexican population.

1 Introduction

1.1 Brief history of patient registries

Patient registries are organized systems that collect, store, and analyze information about individuals from a specific population or living with a particular health condition (1). They are used for clinical research, recruitment into clinical trials, and improving patient care (2). When coupled with genomic information, they provide further opportunities for genomic medicine and research (3).

The National Leprosy Registry of Norway, established in 1856, was presumably the first national patient registry, implemented as a response to the high prevalence of leprosy at the time and included a collaborative effort between health officers and ministers of the church, which resulted in the policy of isolation that ultimately led to the disappearance of leprosy in Norway (4).

Since then, several registries have been set up around the globe with various aims. As expected, the two leading causes of death, cancer and heart disease, have led the efforts through the past century. There are at least 343 cancer registries in 65 countries (5), while 155 Cardiac disease registries in 49 countries have enrolled 73.1 million patients (6). One of the greatest examples is the US National Cardiovascular Data Registry (NCDR), which comprises ten registries and has more than 60 million records now entering the qualification of Big Data (7). In Latin America, population-based cancer registries, arguably the most extensively documented among patient registries, cover 20% of the population, and it is estimated that only 7% contain high-quality clinical information (8).

Although patient registries usually aim to estimate the prevalence of a disease or trait within a population and establish what demographics and environmental factors affect risk, they don't always aim to estimate how much genetic or environmental factors explain a trait. Historically, twin registries have been frequently used to improve our comprehension of human health and the genetic and environmental basis of complex diseases (9). The classical twin design consists of acquiring phenotypic data of monozygotic (MZ) and dizygotic (DZ) twins and measuring how correlated a trait is within pairs of both types. The heritability of a trait, the percentage of variance explained by genetic factors, can be estimated through the difference between the correlation of MZ twins and DZ twins, and the remaining variance in the trait is attributed to environmental factors (10). However, other family research designs can be used to estimate heritability. The so-called extended twin design acquires data from other family members of the twin participants, potentially improving the accuracy of heritability estimates (11). Genome-wide association studies (GWAS) can be used to estimate SNP-based heritability (12). Nevertheless, the costs of expanding twin studies or performing a GWAS are prohibitive for Lower and Middle-Income Countries that systematically lack funding for scientific research (13), making twin studies a more viable option.

Twin studies serve not only to estimate the heritability of a trait but also to assess how two different traits share genetic and environmental factors, elucidating their variance. Furthermore, they can assist in establishing causal relationships between associated traits (14). Consequently, insights gained from twin registries have deepened our understanding of genotype-environment interactions and their role in shaping the mechanisms underlying individual differences and their impact on health and disease conditions (9).

Despite these contributions, it is important to note that results from twin studies may not directly apply to populations with different genetic and environmental backgrounds. Latin America allegedly has some of the smallest sample sizes for twin registries, with around 5,000 individuals, Brazil has the most participants ($n = 4,826$), followed by Mexico (9).

The often mandatory implementation of patient records and twin registries across some countries has promoted relevant progress (15). Nevertheless, a considerable fraction of the world's population remains uncovered, particularly in African and Asian countries (16), due to the struggle to collect and integrate data to establish and maintain national registries (17).

1.2 Advancements and challenges in Mexican health registries

Although healthcare in Mexico has undoubtedly improved its coverage and scope over the last two decades (18), according to the 2020 census, about 33 million people in the country remain with no health services insurance (19), and 48.49% have no adequate access to health services (18).

Relevant efforts have been undertaken to establish registries in Mexico aimed at understanding the leading causes of death, which include heart disease, diabetes, and malignant tumors (20). These initiatives include the Mexican Registry of Atrial Fibrillation (ReMeFa) (21), the Mexican Registry of Pediatric Cardiac Surgery (22), and the National Registry of Cardiac Rehabilitation Programs (RENAPREC), alongside others focused on atrial fibrillation (23) and diabetes (24).

In some cases, efforts must be reinforced; for instance, the National Heart Failure Registry Program (25) was initiated in 2002 to generalize the characterization of patients with this condition; once the aim was achieved in 2004, the project was no longer pursued.

Although the government established a federal norm (NOM-024-SSA3-2012) in 2012 to regulate Electronic Health Record Information Systems (26), the diversity of electronic clinical records and database storage systems greatly hinders data sharing and interaction among health institutions (27). In addition, these registries are mainly survey-based and lack biobanking or genetic analyses, unlike their homologs in the European population, which limits their scope (28–30).

Biobanking encompasses the collection, preservation, analysis, and integration of biological information from biological samples along with personal and health-associated data (31). A few projects have been initiated in Mexico to advance biobanking efforts.

A partnership between the Carlos Slim Foundation and the Broad Institute launched the Slim Initiative in Genomic Medicine for the Americas (SIGMA), focusing on the study of type 2 diabetes (T2D), which has collected over 3,800 T2D cases and 4,072 healthy controls, of which 1,522 cases and 1,546 controls have been integrated into the 52K T2D exome sequencing study (32). The Metabolic Analysis in an Indigenous Sample (MAIS) cohort, which evaluated genetic contribution to traits associated with metabolic syndrome, has recovered samples of over 2,500 individuals from 60 different ethnic groups of Mexico (33).

The Mexico City Prospective Study (MCPS) and the Mexican Biobank (MXB) represent Mexico's most ambitious biobanking efforts. The MCPS aims to assess associations between risk factors and common causes of death in Mexico, including heart disease, stroke, diabetes, and alcoholic liver disease (34), and has generated genotyping and exome-sequencing for over 140,000 and whole genome sequencing of 9,950 individuals; this study is focused only on inhabitants from two districts (Coyoacán and Iztapalapa) in Mexico City (35).

Furthermore, MXB has analyzed over 6,000 DNA samples from over 800 rural and urban locations across Mexico and will serve as a reference for demographic and evolutionary studies. MXB includes anthropometric, disease, and lifestyle variables and biochemical traits derived from blood samples (36).

These two projects have generated invaluable genetic resources from Mexican individuals. However, they are constrained by limitations in the collected data and ethical restrictions when re-contacting participants. Specifically, there is minimal or no information regarding family history, mental and reproductive health, cognition, and chronic, degenerative, and immune diseases.

The MexOMICS Consortium was established to design and implement the required infrastructure for consolidating electronic-based databases and biobanks. This infrastructure is intended to collect, compare, cross-reference, and share valuable clinical, genetic, sociodemographic, and psychosocial information about Mexican individuals with various healthcare-related conditions. The objective is to facilitate a comprehensive characterization of our population.

The consortium is currently focused on three traits in the admixed Mexican population: (a) twins and members of multiple births (TwinsMX) (37), (b) people with Systemic Lupus Erythematosus (LupusRGMX) (38), and (c) patients with Parkinson Disease (MEX-PD) (39). There is a lack of other nationwide registries that include genetic data and publicly available information characterizing Mexican people with these traits. The developed databases offer valuable clinical and sociodemographic insights and include genetic profiling, biobanking, and the acquisition of Magnetic Resonance Imaging (MRI). These comprehensive efforts aim to deepen our understanding of genetic-environmental interactions and their impact on the phenotypic traits within the Mexican population, shedding light on potential implications for health and disease. In the present work, we will discuss the design, implementation, operational issues, and the current state and future directions of these three registries.

2 Methods

2.1 Ethics and data protection

The Research Ethics Committee of the Instituto de Neurobiología at the Universidad Nacional Autónoma de México (UNAM) approved all three registries. TwinsMX (40) was the first registry, established in 2019, followed by LupusRGMX and MEX-PD in 2021. Information is anonymized and stored at the Laboratorio Nacional de Visualización Científica Avanzada at UNAM. Participants provided informed consent before registration, and a privacy statement was given to them per the Federal Law on the Protection of Personal Data Held by Private Parties.

2.2 Platform and surveys design and implementation

Each registry follows specific guidelines to obtain necessary data. For such purposes, surveys are available on the web-based application Research Electronic Data Capture (REDCap), which allows secure data capture and storage (41, 42). Instruments integrated in the three registries are summarized in Table 1 to provide a general picture of the information being collected.

TABLE 1 Instruments employed and information required in twinsMX, LupusRGMX, and MEX-PD registries. The variables and instruments shared by two or more registries are in bold letters.

| Section | Description | | |
	TwinsMX	LupusRGMX	MEX-PD
Sociodemographic variables	Full name Date of birth Age Contact information Occupation Last educational degree (participants and parents) State of residency Family's income in minimum wages	Full name Date of birth Age Contact information Occupation Last educational degree State of residency Number of habitants in home Socioeconomic level index rule of the Mexican Association of Market and Opinion Intelligence Agencies (AMAI) (43).	Full name Date of birth Age Contact information Occupation Last educational degree State of residency Ethnicity Ancestry Genealogy
Medical history	Personal and family history of cardiovascular, metabolic, endocrine, neoplastic, respiratory, psychiatric, neurologic, urinary, reproductive, gastrointestinal, rheumatologic, developmental, pregnancy, dermatologic, and ophthalmologic disorders, surgical history and medication use	Personal and family history of cardiovascular, metabolic, endocrine, neoplastic, respiratory, psychiatric, neurologic, urinary, reproductive, gastrointestinal, rheumatologic, developmental, pregnancy, dermatologic, and ophthalmologic disorders, surgical history, and medication use. Lupus-associated symptoms and clinical manifestations based on the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) (44) and the American College of Rheumatology (ACR) criteria (45) Treatment	Clinical site and neurologist where the participant is registered. Motor and non-motor symptoms Medication Specific scales: Unified Parkinson's Disease Rating Scale (MDS-UPDRS) (46) and the Hoehn & Yahr scale (47)
Mental health	State-Trait Depression Inventory (ST-DEP) (48) State-Trait Anxiety Inventory (STAI) (49) Paranoid ideation Psychoticism and others (50) Self-confidence Social competence Family support Ability of people to organize themselves in difficult times (51)	State-Trait Depression Inventory (ST-DEP) (48) State-Trait Anxiety Inventory (STAI) (49) Symptom Checklist, Revised (SCL-90-R) (52) Coping Strategies Inventory (CSI) (53) Mexican Resilience Scale (RESI-m) (51)	State-Trait Depression Inventory (ST-DEP) (48) State-Trait Anxiety Inventory (STAI) (49) Symptom Checklist, Revised (SCL-90-R) (52) Parkinson Anxiety Scale (54)
Cognitive function	Brief IQ test (Shipley-2) (55), Synchronization-continuation task (56) and Speech-to-Speech Synchronization task (57) Spatial n-back (0,2-b) task for assessing working memory during fMRI (58) Creyos' computerized cognitive battery of 12 tasks (59)	Montreal Cognitive Assessment (MoCA) soon to be applied (60)	Montreal Cognitive Assessment (MoCA) (60) Creyos' computerized cognitive battery, eight selected tasks (59)
Environmental factors	Parental Bonding Questionnaire (61) Childhood Trauma Questionnaire (62).		Questionnaires about pesticides, heavy metals, water, air or soil contaminants, exposure, and head trauma were also included.
Reproductive health	Menstruation Pregnancy Breastfeeding Use of hormonal treatments		PD women-specific questionnaire (63)
Lifestyle and quality of life	Pittsburgh Sleep Quality Index (64) World Health Organization Quality of Life (WHOQoL-bref) (65). 36-Item Short Form Health Survey (SF-36) (66)	Pittsburgh Sleep Quality Index (64) World Health Organization Quality of Life (WHOQoL-bref) (65). 36-Item Short Form Health Survey (SF-36) (66) Lupus Quality of Life (LupusQoL) (67)	Not applied.
Personality traits	Big Five Personality inventory (68) Behavioral Activation System and Behavioral Inhibition System (69)	Big Five Personality inventory (68)	Not applied.
Diet and eating behaviors	Food intake Food-related behaviors Traits suggestive of eating disorders (70)		Not applied.
Physical activity	Global Physical Activity Questionnaire (GPAQ) (71)		Not applied.
Substance use and abuse	Frequency and amount of: Tobacco (72), Alcohol Use Disorders Identification Test (AUDIT) (73), Cannabis Use Disorders Identification Test (CUDIT)	Not applied.	Frequency and amount of: Tobacco Alcohol Caffeine

(Continued)

TABLE 1 Continued

| Section | Description | | |
	TwinsMX	LupusRGMX	MEX-PD
	(74), Consumption of other drugs Substance use disorder identification.		
Cognitive reserve	Not applied		The Cognitive Reserve Index Questionnaire (75)
Zygosity	Evaluation of the similitude of phenotypic physical traits and the frequency that their relatives confound twins. This questionnaire identifies if twins are identical (monozygotic) or nonidentical (dizygotic) with an accuracy of 98% (76)	Not applied.	
Laterality	Edinburgh Handedness Inventory (77)	Not applied.	

All three projects generally share an initial structure, with questionnaires focused on identifying socio-demographic traits and characterizing medical history and mental health. The shared technology information infrastructure of the registries has allowed us to design, implement, and validate surveys in one of the registries to further establish them in the other registries. In addition to these shared questionnaires, each registry includes surveys focused on bringing light on specific traits associated with the studied sample, such as zygosity in TwinsMX or cognitive impairment in MEX-PD.

Although the selection of each instrument responds to particular research questions, one primary guideline is the ability to produce harmonized measures to link and compare the data across registries and their counterparts in other populations (78). For instance, TwinsMX was built to be comparable to the Australian (79) and UK (80) Twins Registries. At the same time, MEX-PD is akin to the newly established Australian Parkinson's Genetics Study (81) and the Latin American Research Consortium on the Genetics of Parkinson's Disease (82).

2.3 Project identity

The term "-omics" encloses the acquisition, analysis, and interpretation of large datasets specific to various biological systems. These approaches, including genomics, proteomics, and metabolomics, allow for in-depth characterization of biological phenomena within specific contexts, thus helping to unravel the complexity of a given biological system or condition (83). In this context, the omics in our Consortium name, "MexOMICS" reflects our aim to generate and integrate omic data, particularly genomic and transcriptomic data, from Mexican individuals. Our goal is to offer a thorough comprehension of specific traits within our population by assessing clinical data and integrating the influences of genetic and environmental factors.

The logo of "MexOMICS" is accompanied by four entangled DNA strands, each representing the conjunction of projects (TwinsMX, LupusRGMX, and MEX-PD) in a Mexican-styled design that resembles a prehispanic symbol known as ollin, which represents "movement" (Figure 1A).

For the Mexican Twin Registry (Registro Mexicano de Gemelos in Spanish), we selected the acronym TwinsMX; "Twins" to highlight the multiple births and "MX" for Mexican. The logo is centered in this acronym and includes "Registro Mexicano de Gemelos" in the bottom part. The "I" and "X" letters include two dots on the top, resembling a pair of heads, representing the pairs of twins (Figure 1B). Color palette includes bright green and a shade of pink known as Mexican pink in our culture; both colors are widely used on traditional artisanal crafts and therefore represent the vibrancy of our culture, as well as the spirit, charisma, liveliness, and vitality of our people. TwinsMX's website (https://twinsmxofficial.unam.mx/) includes the "Sign up" section, along with an "About us" section focused on the academic profiles of the TwinsMX team members, frequently asked questions, information about the projects well as an interactive section that allows the people to see how is the registry advancing through time. It also includes a "Blog" section that shares science-based information of interest in a non-technical language. TwinsMX was launched in May 2019 and represents the first platform implemented by our consortium.

For the Mexican Lupus Registry's acronym, the word Lupus was selected as the central element, followed by the letters "RG" from ReGistry and "MX" from MeXican. The registry's logo includes the acronym along with "Registro Mexicano de Lupus" in the bottom part and integrates a butterfly, which resembles the malar rash, one of the most common manifestations of SLE, along with a meaning of hope; the purple color palette was selected because of its broad utilization in awareness campaigns as it symbolizes a middle point between the intensity and motivation of the red and the calmness and stability of the blue (Figure 1C). The website (https://lupusrgmx.liigh.unam.mx/) was created considering the same color palette. It includes a "Who are we" section, which includes academic profiles of the research team members, a "Communication" part with the structure of an informative blog, a "Lives" section that shows the Facebook Live events hosted by members of the project, a "Contact" section to provide a communicative platform with the team, and the "Sign up" button to access to the registry. LupusRGMX was successfully launched in May 2021, was socialized through a live Facebook session, and has worked since then using these platforms.

The Mexican Parkinson's Research Network, Red Mexicana de Investigación en Párkinson in Spanish, is represented with the acronym "MEX-PD", which includes "MEX" for Mexican and

FIGURE 1
Logos designed for (A) MexOMICS, (B) twinsMX, (C) LupusRGMX and (D) MEX-PD.

"PD" for Parkinson Disease (Figure 1D). The logo includes a tulip silhouette, a flower that has been broadly used to raise awareness about Parkinson Disease since a horticulturist living with PD developed a new variant of tulip and named it after Dr. James Parkinson, the first physician to describe and document the disease. MEX-PD website (https://mexpd.liigh.unam.mx/) includes a "Frequently asked questions" section, an "About us" section that includes the academic information about the team, and a "Pre-registry" option. MEX-PD was launched in July 2021.

2.4 Inclusion and exclusion criteria

For the TwinsMX cohort, Mexican multiple-birth individuals of all ages are welcome to participate in the online survey. Minors must obtain consent from a parent or legal guardian and are provided with age-appropriate questionnaires for minors, parents can assist pre-pubescent children in completing them. For the MRI study, only twin participants between 18 and 60 years of age with no major neurological disease (e.g., moderate or severe traumatic brain injury, neurodegenerative disease, tumor, stroke, and epilepsy), intellectual disability (an estimated IQ below 70), and who can read are included.

In LupusRGMX, Mexican individuals aged over 18 years with SLE diagnosis were included. SLE participants must have a confirmed diagnosis and fulfill ≥4 American College of Rheumatology (ACR) criteria (Hochberg) at the moment of diagnosis (45). Current or past corticoids and immunosuppressants usage were also used to confirm SLE diagnosis. Disease activity, additional treatments, and comorbidities were not considered as exclusion criteria.

MEX-PD participants were required to have been born in Mexico and/or have spent most of their lives in the country and must be at least 45 years old. PD diagnosis must be confirmed by neurologists specialized in movement disorders, based on the UK Brain Bank Criteria.

LupusRGMX and MEX-PD included non-affected controls who fulfilled the same age criteria and were not consanguineous

with participants with SLE and PD, respectively. Additionally, people with another autoimmune disease were excluded from the LupusRGMX control cohort, whereas individuals with a diagnosis of neurodegenerative diseases were excluded from the MEX-PD control cohort.

2.5 Volunteers recruitment

As one of the main objectives of the registries is to include as many Mexicans as possible across the country, it became evident that a decentralized approach was needed. Given the varied nature of each registry, specific procedures for participant recruitment have been followed: for TwinsMX, the enrollment is made through social media campaigns and on-site event invitations. Given the complexity of PD diagnosis, the recruitment in MEX-PD requires the clinical expertise of neurologists; a team of neurologists from different sites across Mexico has been assembled, and volunteers are being identified and registered through their practice. For LupusRGMX, the recruitment relies on a hybrid strategy: A team of rheumatologists actively register patients in their practice through social media and on-site events where volunteers are invited to self-register.

2.6 Genetic data acquisition

The collection of biological samples through buccal swabs for DNA acquisition is being performed at multiple sites for the three cohorts, allowing for the representation of diverse genetic backgrounds. Additionally, in LupusRGMX, DNA has also been isolated from blood samples.

For TwinsMX, genotypes were generated using the Illumina Infinium Global Diversity Array-8 v1.0 array, which comprises a 1,825,277 SNPs global backbone designed

for cross-population imputation coverage of the genome [https://www.illumina.com/products].

In LupusRGMX, whole genome sequencing was performed by Novogene Inc. in Sacramento, CA, United States (https://www.novogene.com/). The genome libraries were prepared (350 bp), and sequencing was performed using the Illumina NovaSeq X Plus Series® platform with a paired-end 150 bp sequencing strategy, with 90G raw data per sample and a coverage depth of 4X. Simultaneously, 50 samples have been collected to perform single-cell RNA-seq of peripheral blood mononuclear cells.

For MEX-PD genotyping is being performed through the support and collaboration with the Global Parkinson's Genetics Program (GP2) (84) and the Latin American Research Consortium on the Genetics of Parkinson's Disease (LARGE-PD) (82). The genotyping platform is the Illumina NeuroBooster array (NBA), containing 1,914,934 variants and specifically designed to screen for genetic variation in neurological disorders across diverse populations (85).

3 Results

3.1 Progress at a glance

As of June 2023, our registries included 2,915 volunteers for TwinsMX (April 28th, 2023), 1,761 participants for LupusRGMX (June 21st, 2023), and 750 for MEX-PD (June 29th, 2023) (Table 2). Additionally, for LupusRGMX, 153 participants were registered as controls, whereas MEX-PD had 397 volunteers registered as non-affected controls. Female participation represents 71.9% of TwinsMX, 94.3% in LupusRGMX, and 59.2% in MEX-PD, whereas 28.1%, 6.7% and 40.8% of participants were male in each study.

TwinsMX covers an age range from 0 to 80 years, with the majority between 20 and 40 y.o. (Figure 2A), and a mean age of 29.2 ± 11.2 years at the moment of enrollment. Age distribution

among monozygotic and dizygotic participants is shown in Figure 2, and a similar pattern can be observed in both groups.

In the case of LupusRGMX, the age range is between 18 and 79 years (Figure 2B), with the majority of the cohort aged between 20 and 45 and a mean age of diagnosis of 28.3 ± 10.9 years. Healthy controls' age distribution is also exhibited, with an age range from 10 to 60 years and the majority of volunteers being in their early 20s.

For MEX-PD, participants' ages range from 40 to 90 years (Figure 2C), with the majority being between 60 and 80 and a mean age of 62.17 ± 10.6 years. Control participants exhibited a similar age range, with the majority being between 40 and 60 y.o.

As mentioned, genetic analyses and other clinical measurements can help enrich the information collected through the registries, giving insights into the biological conditions of our participants. In this line, our team has been recovering DNA samples from volunteers of the three projects. As of June 2023, TwinsMX had collected 232 swab samples, of which 47 have already been genotyped. In LupusRGMX 109 samples have been recovered to perform whole genome sequences, of which 71 have been successfully sequenced: 61 from volunteers with SLE and ten healthy controls. In the case of MEX-PD, 495 samples have been obtained to perform whole genome genotyping.

As the nervous system is compromised in Parkinson's Disease, cognitive functions are evaluated with the Montreal Cognitive Assessment (MoCA). As of June of 2023, 126 controls and 250 patients have been evaluated. Due to the high prevalence and severity of neurologically related manifestations in SLE (also known as neuro-lupus), LupusRGMX will apply 90 MoCA evaluations by 2024. Additionally, structural, functional, and spectroscopy brain images have been acquired through MRI sessions of a subset of 267 TwinsMX volunteers (June 2023), although 500 participants are expected to be collected by the Spring of 2024. MEX-PD and LupusRGMX will perform at least 90 MRI evaluations each by 2024.

Regarding the geographical coverage of the three projects, we have a substantial number of volunteers from the center of Mexico, probably due to the location of the leading laboratories (Mexico City and Queretaro). In TwinsMX (Figure 2A), the majority of registers are concentrated in Mexico City ($n = 1,176$), Mexico State ($n = 295$), Veracruz ($n = 149$), Jalisco ($n = 111$), Puebla ($n = 105$), and Guanajuato ($n = 101$); with less participation among people from the north-western and south regions. For LupusRGMX (Figure 2B), we can observe higher participation among people living in Mexico City ($n = 255$), Mexico State ($n = 145$), and Queretaro ($n = 119$), whereas the west and northern west regions are the least represented. In the case of MEX-PD (Figure 2C) most of the participants come from Mexico City ($n = 197$), Queretaro ($n = 109$), Morelos ($n = 75$), Michoacan ($n = 72$), San Luis Potosí ($n = 71$), Mexico State ($n = 62$), and Nuevo Leon ($n = 42$); the states with no volunteers registered are shown in light gray.

3.2 Challenges and limitations

Implementing a National Registry always faces challenges, especially at the beginning. For instance, TwinsMX relies entirely on the self-registration of volunteers who are reached by our

TABLE 2 Sociodemographic characteristics of volunteers across registries.

Information	TwinsMX ($n = 2,915^*$)	LupusRGMX ($n = 1,761$)	MEX-PD ($n = 750$)
Gender (n, %)			
Female	2,096 (71.9%)	1,662 (94.3%)	444 (59.2%)
Male	819 (28.1%)	99 (6.7%)	306 (40.8%)
Age at enrollment (mean, SD)	29.2 ± 11.2	37.1 ± 10.9	62.17 ± 10.6
Time with the diagnosis (mean, SD)	NA	8.5 ± 7.8	6.7 ± 5.7
Age at diagnosis (mean, SD)	NA	28.3 ± 10.9	59.7 ± 11.8
Genetic samples collected (n, %)	232 (7.9%)*	108 (6.1%)	495 (66%)
MRI sampled (n)	267 (9.1%)	To start in Q2 2024	To start in Q2 2024
Cognitive function evaluations	267 (9.1%)**	To start in Q2 2024	376 (50.1%)

*Individual subjects.
**Subjects in MRI study.
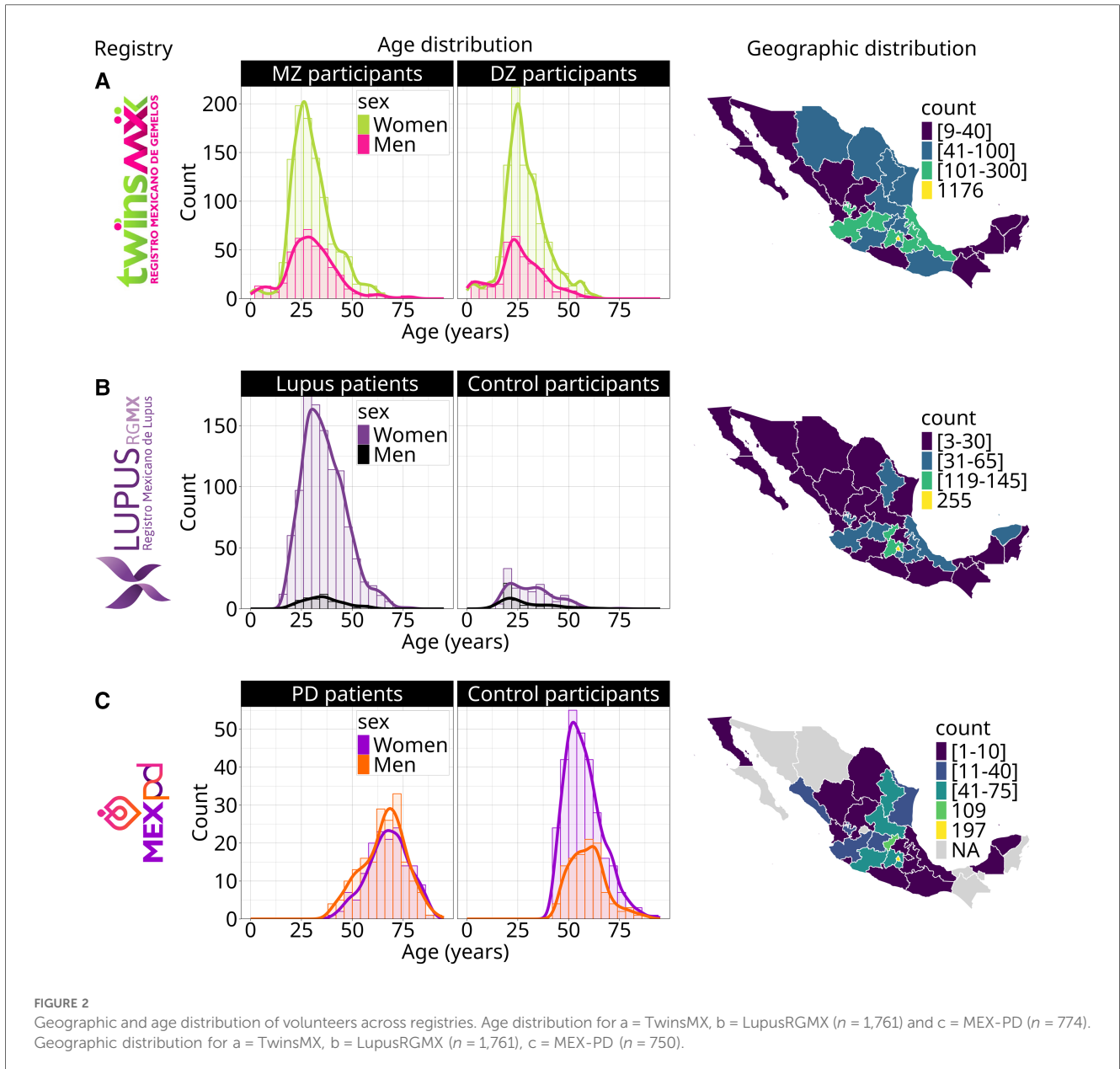
FIGURE 2
Geographic and age distribution of volunteers across registries. Age distribution for a = TwinsMX, b = LupusRGMX (n = 1,761) and c = MEX-PD (n = 774). Geographic distribution for a = TwinsMX, b = LupusRGMX (n = 1,761), c = MEX-PD (n = 750).

invitations through social media or events since there is no database available to allow us to identify and contact potential participants, which limits the scope of the project. Besides, it is necessary that all the siblings of multiple births complete the surveys. Still, in most cases, only one of the twins is interested or motivated to finish the questionnaires, which leads to incomplete data. Another limitation is that in Mexico, there is a word used explicitly for fraternal twins, "cuate" (/'kwate/), thus often, they do not identify themselves as twins, making them think that they are excluded from the registry. More vigorous and targeted media campaigns are being implemented to address this issue.

LupusRGMX mainly works on a voluntary self-registration basis of people reached by our communication campaigns. By far, the main obstacle we have faced has been reaching Mexican individuals with SLE with the interest and disposition to start and complete the surveys: this is reflected in the fact that

from the 1761 started enrollments, only 1,271 have completed the clinical questionnaire. Different strategies have been undertaken to increase survey completion, such as spacing out survey invitations and motivating participants through group dynamics, contests, and offering giveaways. In addition, since the beginning, LupusRGMX has tried to approach rheumatologists through the Mexican College of Rheumatology to establish collaborations in which the rheumatologist gives a formal clinical evaluation at the beginning of the registry and helps their patients through the initial clinical manifestations questionnaires, improving the reliability of our data. As of June 2023, only 147 participants have been registered through this option. With the knowledge that self-reported data can be threatened by self-reporting bias and thus considered unreliable (86), strategies are being undertaken to assess them; we are currently working on comparing the outcomes from self-reported

instruments with those provided by patients' medical practitioners to evaluate reliability.

Similarly, as MEX-PD patients must be registered through a Neurologist with expertise to perform a clinical evaluation, reaching people with access to this level of medical care has been challenging. The relative shortage of Neurologists with valid certification in Mexico (87), along with their substantial workloads, driven by the necessity to attend to a large number of patients, limits the scope of the registry. Nonetheless, 774 participants have been enrolled so far. Furthermore, as aforementioned, Neurologists in Mexico usually have a heavy workload, making the survey's completion daunting. To ease this task, a trained assistant team helps patients complete particular surveys (i.e., environmental exposure). Moreover, we are currently working on transferring suitable instruments for participants to self-complete.

Likewise, a shared challenge between the registries is acquiring data from more diverse individuals within Mexico. Reviewing the current progress of the cohorts, it became evident that most enrollments are representative of the central states of Mexico and from specific regions where we have collaborating associations. For this reason, we are working on specifically targeting the underrepresented regions of the country, taking advantage of online tools, using social media, and establishing new collaborations. However, Magnetic Resonance Imaging, performed in suitable TwinsMX participants and blood samples for LupusRGMX, demands on-site presence of participants and is only carried out in Queretaro, limiting the participation of people from northern or southern states in the country. These techniques' complexity and required infrastructure prevent its mobility to other Mexican states.

Finally, because national registries are rare in Mexico, constant invitation reminders via traditional and social media are required to boost participation continuously. As of 2020, 72.0% aged six or older in Mexico residents are internet users (88). Therefore recruitment through social media has shown to be a very effective tool, with similar models successfully implemented in other registries (89, 90). As of June 2023, our registries have implemented an integrative approach through social media, turning them into reliable platforms for the exposition and discussion of topics associated with each community.

3.3 Sharing is caring: community building

In Mexico, according to the General Science, Technology and Innovation State Inform 2018, over 69.8% of the population perceive that the knowledge generated through scientific and technological research positively impacts our economy, and up to 84.5% think that scientific research plays a vital role in the technological development of the country. Still, 45.7% of the population perceive the scientific community as dangerous due to their knowledge and the potential misuse of their discoveries by third parties (91). In this sense, as a consortium, we believe that the participants in the three registries are not subjects of study but active citizens with a major role in biomedical research. As such, it is imperative to provide informed consent, maintain transparent procedures, disclose progress reports, and open communication channels where participants can provide feedback.

As twins, Lupus, and Parkinson's communities in Mexico were already active, one of the first steps in building the three registries was approaching community leaders, civil organizations, and foundations, which have been critical in reaching participants and gaining credibility in their communities. TwinsMX, the first of this consortium, received support from the Multiple Birth Association (Asociación de Nacimientos Múltiples, A.C.). TwinsMX now holds a significant community, with a substantial presence in social media with over 3.2 K followers on Facebook (https://www.facebook.com/TwinsMXofficial) and 748 on Instagram (https://www.instagram.com/twinsmxofficial). The TwinsMX's team constantly works on generating informative content (https://twinsmxofficial.unam.mx/blog/) and fosters participation in community activities.

Furthermore, LupusRGMX receives direct support and active participation from representatives of the biggest communities of people with Lupus in the country: LUPUS MX (https://www.facebook.com/LUPUSMXOficial), Fundación Proayuda Lupus Morelos A.C. (https://www.facebook.com/LUPUSMORELOS), the Centro de Estudios Transdisciplinarios Athié-Calleja por los Derechos de las Personas con Lupus A.C. (https://www.facebook.com/Cetlu), and Despertar de la Mariposa A.C. (https://www.facebook.com/DespertarLupus); who have reviewed the surveys, and expressed their opinions and needs. We have also approached rheumatologists and the Colegio Mexicano de Reumatología to establish collaborations to help provide a more integrated evaluation of SLE in our cohort. The formation of this community has helped us identify necessities and areas of opportunity for clinicians and researchers to generate knowledge with future applications and impact on the lives of people with lupus. Through contact with experts on topics of interest such as pediatric SLE, grief associated with a chronic disease diagnosis, and reproductive health, we have established monthly open talks through Facebook Live sessions (https://www.facebook.com/lupusrgmx). Besides these talks, the team of LupusRGMX has been working on providing other sources of reliable information (https://lupusrgmx.liigh.unam.mx/comunicacion.html) and actively participating in activities for lupus awareness in Mexico.

Similarly, MEX-PD has sought support from the Red Mexicana de Asociaciones de Párkinson (https://www.facebook.com/RedMXdeAsociacionesdeParkinson) and different communities of people with Parkinson's and their loved ones, including Parkinson Laredo (https://www.facebook.com/groups/768321520680203) and Asociación Mexiquense de Parkinson IAP (https://www.facebook.com/parkinsonmexiquense).

3.4 Key findings from ongoing research

Deeper analyses of the sociodemographic and clinical characteristics of each registry will help answer each project's

specific questions; these analyses are and will continue to be presented in independent articles. However, we summarize here the main findings to exemplify the potential of these resources.

The TwinsMX registry aims to understand the interplay of genetic and environmental factors influencing the most prevalent diseases, subclinical symptoms, and cognitive aspects of the Mexican population. The first steps of this project have been previously described (37). The latest results include a study on myopia and astigmatism, revealing notably high heritability (over 80%) estimated for both conditions, accompanied by a substantial cross-trait genetic correlation (92). These findings are consistent with prior research conducted in other populations, suggesting that both traits are influenced by a shared set of genes. One of the further goals of this project is to explore associations of genetic variants with other publicly relevant health conditions, as well as neuropsychiatric measures such as depression and anxiety symptoms. For instance, our latest published advance for this project (92) revealed a high prevalence of gastritis, colitis, overweight, allergic rhinitis, and gastroesophageal reflux, in addition to myopia and astigmatism. On the other hand, anxiety, depression, insomnia, migraine, attentional deficit/ hyperactivity disorder, and obsessive-compulsive disorder were the mental health conditions with the highest occurrence. Additionally, we have reported our progress in collecting MRI data, including high resolution brain structural images, resting state functional MRI, magnetic resonance spectroscopy and task-related functional MRI (40). Altogether, this project is expected to aid on the characterization of the health profiles of our population, and the relative influence of genetic and environmental factors on a great variety of health, cognitive and brain phenotypes in the Mexican population.

Hernandez-Ledesma et al. (38) used LupusRGMX data to evaluate quality of life disparities, finding a lower quality of life in people with lupus than healthy participants, and that the quality of life was predicted by socioeconomic status, delay in diagnosis, and corticosteroid consumption, highlighting that lower socioeconomic status leads to a lower quality of life. Pedraza-Meza et al. (93) analyzed how social, clinical, psychological, and demographic variables affect social and temporal decision-making in people with lupus, reporting the importance of age and hostility to predict social decisions, as well as anxiety and obsessive-compulsiveness to predict temporal decisions, whereas clinical factors, i.e., being in remission and taking glucocorticoids, predict both kinds of decisions.

Regarding MEX-PD, Lázaro-Figueroa et al. (39) described a general analysis of the cohort demographic and clinical data. The age of onset of Parkinson's Disease (PD) was found to be younger than the reported for other populations as 19.8% of patients with PD were identified as early onset (age <50 years). Contrary to what was previously reported in other populations, the MEX-PD cohort did not show an association between head injury and PD. The more frequent initial motor symptom that MEX-PD patients reported was tremor in the upper limbs, and the more frequent treatment was levodopa. These preliminary results give new information about the characteristics of PD in the Mexican population.

3.5 Perspectives

From a general perspective, TwinsMX, LupusRGMX, and MEX-PD will allow us to characterize the genetic and environmental contribution of different traits and diseases with data collected in the Mexican population, which is heavily underrepresented in epidemiological and genetic studies.

Projects such as the MXB, the MAIS, and the MCPS studies represent a breakthrough in the knowledge of genetic diversity in Mexico and have added invaluable data regarding the genetic structure of the Mexican population. However, there remains an unmet need to expand our understanding of the impact of genetic variants, environmental factors, and their interactions on health-related traits. Overall, this characterization will provide medically relevant information to design, execute, and monitor clinical trials, and improve the time and accuracy of diagnosis, treatments, and therapies.

The collaborative design implemented in MexOMICS' registries, which contemplates recontacting based on email and integration of new questionnaires, alongside genetic data acquisition, facilitates the conduct of longitudinal studies. Furthermore, the shared questionnaires and variables recorded among the three registries enables data comparison between studies (38).

On the other hand, the registries within MexOMICS are also expanding to include non-invasive techniques to visualize and analyze brain structure and function. Multimodal neuroimaging will aid in the characterization of brain-health associations and their variability associated to genes and environment in our cohorts.

In particular, for TwinsMX, we expect to start to collect DNA from direct relatives of twins individuals (i.e., parents and not-twins siblings) to build a biobank. In the future, it is expected that we will acquire data through electroencephalography (EEG) to further investigate associations with cognitive performance. Meanwhile, for LupusRGMX MRI and cognitive function evaluations will be performed in at least 90 participants to explore the underlying functional and structural brain substrate. Finally, MEX-PD will also acquire imaging data through MRI, employing state-of-the-art techniques such as resting-state functional imaging, high-resolution multicontrast structural acquisitions, diffusion, and neuromelanin-weighted imaging. By also exploring non-affected controls, the main aim will be to characterize the functional and structural brain substrate of the clinical and cognitive deficits identified in our cohort.

4 Conclusions

Patient registries are fundamental tools in public health research, providing comprehensive population-based data on diseases' etiology and progression. Establishing registries in Mexico, including those initiated at the Universidad Nacional

Autónoma de México through the MexOMICS Consortium, has enriched the health research landscape with epidemiological and genetic information. These registries enable researchers to explore the complex interplay between genetic factors and environmental influences in traits in the general population by studying twin participants and those involved in complex diseases like systemic lupus erythematosus and Parkinson's disease. However, limitations, such as inconsistent data entry and patient participation, together with potential biases due to voluntary participation, must be addressed to ensure that the creation of these databases can serve as robust foundations for subsequent studies and investigations. Future directions involve refinement and expansion, optimizing experimental design, data collection methodologies, and fostering interdisciplinary collaboration.

The MexOMICS Consortium's registries represent a valuable resource for addressing questions associated with health and disease in the Mexican people.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics on Research Committee of the Institute of Neurobiology at Universidad Nacional Autónoma de México. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

PR-P: Data curation, Investigation, Writing – original draft, Writing – review & editing, Methodology. AH-L: Data curation, Investigation, Writing – original draft, Writing – review & editing, Methodology. TR-L: Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing. BG-V: Data curation, Methodology, Writing – original draft, Writing – review & editing. DR-G: Data curation, Methodology, Writing – review & editing. AL-F: Data curation, Methodology, Writing – review & editing. DM: Data curation, Methodology, Writing – original draft. VF-O: Methodology, Writing – review & editing. IE-M: Methodology, Writing – review & editing. LT-N: Investigation, Writing – review & editing. AP-A: Investigation, Writing – review & editing. EM-F: Investigation, Writing – review & editing. CG-G: Investigation, Writing – review & editing. ET-V: Investigation, Writing – review & editing. MG-H: Investigation, Writing – review & editing. NG-M: Investigation, Writing – review & editing. KS-B: Investigation, Writing – review & editing. GF-R: Investigation, Writing – review & editing. GF-V: Investigation, Writing –

review & editing. YM-P: Investigation, Writing – review & editing. IE-B: Investigation, Writing – review & editing. DA-R: Writing – review & editing, Investigation, Methodology. MR-V: Investigation, Writing – review & editing. MR: Conceptualization, Investigation, Methodology, Writing – review & editing. AR-C: Conceptualization, Investigation, Methodology, Supervision, Writing – review & editing. SA: Conceptualization, Investigation, Methodology, Supervision, Writing – review & editing. AM-R: Conceptualization, Investigation, Methodology, Supervision, Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Gliklich RE, Dreyer NA, Leavy MB, editors. Patient registries. *Registries for Evaluating Patient Outcomes: A User's Guide [e-book]*. Rockville, MD: Agency for Healthcare Research and Quality (US) (2014). Available online at: https://www.ncbi.nlm.nih.gov/books/NBK208643/ (cited January 16, 2023).

2. Kodra Y, de la Paz M P, Coi A, Santoro M, Bianchi F, Ahmed F, et al. Data quality in rare diseases registries. In: Posada de la Paz M, Taruscio D, Groft SC, editors. *Rare Diseases Epidemiology: Update and Overview*. Cham: Springer International Publishing (2017). p. 149–64.

3. Thompson R, Johnston L, Taruscio D, Monaco L, Béroud C, Gut IG, et al. RD-connect: an integrated platform connecting databases, registries, biobanks and clinical bioinformatics for rare disease research. *J Gen Intern Med*. (2014) 29(Suppl 3):S780–7. doi: 10.1007/s11606-014-2908-8

4. Irgens LM. The origin of registry-based medical research and care. *Acta Neurol Scand Suppl*. (2012) 195:4–6. doi: 10.1111/ane.12021

5. Merabishvili V. Review cancer incidence in five continents Vol. XI / ed. By F. Bray, M. Colombet, L. Mery, M. Piñeros, A. Znaor, R. Zanetti and J. Ferlay. IARC Sci. Publ. № 166, Lyon, France, 1545 p. *Issues Oncol*. (2021) 67:855. doi: 10.37469/0507-3758-2021-67-6-855

6. Dawson LP, Biswas S, Lefkovits J, Stub D, Burchill L, Evans SM, et al. Characteristics and quality of national cardiac registries: a systematic review. *Circ Cardiovasc Qual Outcomes*. (2021) 14(9):e007963. doi: 10.1161/CIRCOUTCOMES.121.007963

7. Weintraub WS. Role of big data in cardiovascular research. *J Am Heart Assoc*. (2019) 8(14):e012791. doi: 10.1161/JAHA.119.012791

8. Piñeros M, Abriata MG, Mery L, Bray F. Cancer registration for cancer control in Latin America: a status and progress report. *Rev Panam Salud Publica*. (2017) 41:e2. doi: 10.26633/RPSP.2017.221

9. Hur YM, Bogl LH, Ordoñana JR, Taylor J, Hart SA, Tuvblad C, et al. Twin family registries worldwide: an important resource for scientific research. *Twin Res Hum Genet*. (2019) 22(6):427–37. doi: 10.1017/thg.2019.121

10. Willoughby EA, Polderman TJC, Boutwell BB. Behavioural genetics methods. *Nat Rev Methods Primers*. (2023) 3(1):1–16. doi: 10.1038/s43586-022-00191-x

11. Keller MC, Medland SE, Duncan LE. Are extended twin family designs worth the trouble? A comparison of the bias, precision, and accuracy of parameters estimated in four twin family models. *Behav Genet*. (2010) 40(3):377–93. doi: 10.1007/s10519-009-9320-x

12. Tang M, Wang T, Zhang X. A review of SNP heritability estimation methods. *Brief Bioinform*. (2022) 23(3):1–15. doi: 10.1093/bib/bbac067

13. George SHL, Medina-Rivera A, Idaghdour Y, Lappalainen T, Gallego Romero I. Increasing diversity of functional genetics studies to advance biological discovery and human health. *Am J Hum Genet*. (2023) 110(12):1996–2002. doi: 10.1016/j.ajhg.2023.10.012

14. van Dongen J, Slagboom PE, Draisma HHM, Martin NG, Boomsma DI. The continuing value of twin studies in the omics era. *Nat Rev Genet*. (2012) 13(9):640–53. doi: 10.1038/nrg3243

15. Bakken IJ, Ariansen AMS, Knudsen GP, Johansen KI, Vollset SE. The Norwegian patient registry and the Norwegian registry for primary health care: research potential of two nationwide health-care registries. *Scand J Public Health*. (2020) 48(1):49–55. doi: 10.1177/1403494819859737

16. Grundmann E, Pedersen E, editors. Cancer registries. *Recent Results in Cancer Research*. Heidelberg: Springer-Verlag (1975). p. 215.

17. Morsy A, Lim TO, Varatharajan S, Lim JY. *National registries in developing countries: understanding construction challenges and implementation steps*. In: *2010 5th Cairo International Biomedical Engineering Conference*. (2010). p. 75–8.

18. Gutiérrez JP, García-Saisó S, Dolci GF, Hernández Ávila M. Effective access to health care in Mexico. *BMC Health Serv Res*. (2014) 14:186. doi: 10.1186/1472-6963-14-186

19. Instituto Nacional de Estadística y Geografía (INEGI). Instituto Nacional de Estadística y Geografía. Demografía y Sociedad (2021). Available online at: https://www.inegi.org.mx/temas/derechohabiencia/ (cited 2022 8).

20. Instituto Nacional de Salud Pública (INSP). INSP. De qué mueren los mexicanos (2020). Available online at: https://www.insp.mx/avisos/5111-dia-muertos-mexicanos.html (cited January 27, 2023).

21. Lara-Vaca S, Cordero-Cabra A, Martinez-Flores E, Iturralde-Torres P. Registro Mexicano de fibrilación auricular (ReMeFa). *Gac Med Mex*. (2014) 150(s1):48–59. Available online at: https://www.anmm.org.mx/GMM/2014/s1/GMM_150_2014_S1_048-059.pdf

22. Cervantes-Salazar J, Calderon-Colmenero J, Ramirez-Marroquin S, Palacios-Macedo A, Cerdan AB, Alarcon AV, et al. El registro mexicano de cirugía Cardiaca pediátrica. Primer informe. *Evid Med Invest Salud.* (2014) 7(2):56–62. Available online at: https://www.medigraphic.com/pdfs/evidencia/eo-2014/eo142b.pdf

23. González-Hermosillo JA, Márquez MF, Ocampo-Peña S, en nombre del Comité Científico del Registro CARMEN-AF. Design of an atrial fibrillation and embolic risk registry in Mexico: cARMEN-AF. *Arch Cardiol Mex.* 2017;87(1):5–12. doi: 10.1016/j.acmx.2016.11.003

24. Faradji RN, Valenzuela-Lara M, Vidrio-Velázquez M, Yepez-Rodríguez AE, González-Galvez G, Sainz de la Maza-Viadero ME, et al. RENACED-DT1: a national type 1 diabetes registry initiative in Mexico. *Salud Publica Mex.* (2020) 62 (3):232–4. doi: 10.21149/11185

25. Arturo OT, Lilia CM, Sergio FS, Arturo OS. Programa nacional de registro de insuficiencia Cardiaca. *Medicina Interna de México.* (2004) 20(4):243–52. Available online at: https://www.imbiomed.com.mx/articulo.php?id=25925 (cited 2023 Jan 27).

26. de Salud S. Diario Oficial de la Federación. NORMA Oficial Mexicana NOM-024-SSA3-2012, Sistemas de información de registro electrónico para la salud. Intercambio de información en salud (2012). Available online at: http://dof.gob.mx/nota_detalle.php?codigo=5280847&fecha=30/11/2012 (cited March 8, 2022).

27. Morales-Rocha VM. Technical reference framework for the interoperability of electronic medical record: Conceptual model. *Revista Mexicana de Ingenieria Biomedica* (2020). Available online at: http://www.rmib.mx/index.php/rmib/article/view/886

28. Gordon H, Trier Moller F, Andersen V, Harbord M. Heritability in inflammatory bowel disease: from the first twin study to genome-wide association studies. *Inflamm Bowel Dis.* (2015) 21(6):1428–34. doi: 10.1097/MIB.0000000000000393

29. Kho AN, Hayes MG, Rasmussen-Torvik L, Pacheco JA, Thompson WK, Armstrong LL, et al. Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study. *J Am Med Inform Assoc.* (2012) 19(2):212–8. doi: 10.1136/amiajnl-2011-000439

30. arcOGEN Consortium; arcOGEN Collaborators; Zeggini E, Panoutsopoulou K, Southam L, Rayner NW, et al. Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. *Lancet.* (2012) 380(9844):815–23. doi: 10.1016/S0140-6736(12)60681-3

31. Annaratone L, De Palma G, Bonizzi G, Sapino A, Botti G, Berrino E, et al. Basic principles of biobanking: from biological samples to precision medicine for patients. *Virchows Arch.* (2021) 479(2):233–46. doi: 10.1007/s00428-021-03151-0

32. dbGaP Study. Available online at: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001388.v2.p1 (cited March 27, 2024).

33. Mendoza-Caamal EC, Barajas-Olmos F, García-Ortiz H, Cicerón-Arellano I, Martínez-Hernández A, Córdova EJ, et al. Metabolic syndrome in indigenous communities in Mexico: a descriptive and cross-sectional study. *BMC Public Health.* (2020) 20(1):339. doi: 10.1186/s12889-020-8378-5

34. Tapia-Conyer R, Kuri-Morales P, Alegre-Díaz J, Whitlock G, Emberson J, Clark S, et al. Cohort profile: the Mexico city prospective study. *Int J Epidemiol.* (2006) 35 (2):243–9. doi: 10.1093/ije/dyl042

35. Ziyatdinov A, Torres J, Alegre-Díaz J, Backman J, Mbatchou J, Turner M, et al. Genotyping, sequencing and analysis of 140,000 adults from Mexico city. *Nature.* (2023) 622(7984):784–93. doi: 10.1038/s41586-023-06595-3

36. Sohail M, Palma-Martínez MJ, Chong AY, Quinto-Cortés CD, Barberena-Jonas C, Medina-Muñoz SG, et al. Mexican Biobank advances population and medical genomics of diverse ancestries. *Nature.* (2023) 622(7984):775–83. doi: 10.1038/s41586-023-06560-0

37. Leon-Apodaca AV, Chiu-Han E, Ortega-Mora I, Román-López TV, Caballero-Sánchez U, Aldana-Assad O, et al. TwinsMX: uncovering the basis of health and disease in the Mexican population. *Twin Res Hum Genet.* (2019) 22(6):611–6. doi: 10.1017/thg.2019.112

38. Ana Laura H-L, Domingo M, Elizabeth F-B, Román-López TV, Karen N-R, Sandra Valentina VDV, et al. Lupus RGMX: Social and Clinical Characteristics and their Contribution to Quality of Life in a Mexican Cohort with SLE. *medRxiv.* (2023). doi: 10.1101/2023.02.23.23286331v1

39. Lázaro-Figueroa A, Reyes-Pérez P, Morelos-Figaredo E, Guerra-Galicia CM, Estrada-Bellmann I, Salinas-Barboza K, et al. MEX-PD: a national network for the epidemiological & genetic research of Parkinson's disease. *medRxiv.* (2023):2023.08.28.23294700. doi: 10.1101/2023.08.28.23294700v1 (cited August 31, 2023).

40. García-Vilchis B, Román-López TV, Ramírez-González D, López-Camaño XJ, Murillo-Lechuga V, Díaz X, et al. TwinsMX: exploring the genetic and environmental influences on health traits in the Mexican population. *Twin Res Hum Genet.* (2024):1–12. doi: 10.1017/thg.2024.18

41. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform.* (2019) 95:103208. doi: 10.1016/j.jbi.2019.103208

42. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* (2009) 42(2):377–81. doi: 10.1016/j.jbi.2008.08.010

43. Asociación Mexicana de Agencias de Investigación de Mercados y Opinión Pública (AMAI). AMAI. NSE—Nivel Socioeconómico (2020). Available online at: https://www.amai.org/NSE/index.php?queVeo=NSE2020 (cited July 7, 2023).

44. Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH. Derivation of the SLEDAI. A disease activity index for lupus patients. The committee on prognosis studies in SLE. *Arthritis Rheum.* (1992) 35(6):630–40. doi: 10.1002/art.1780350606

45. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. *Ann Rheum Dis.* (2019) 78 (9):1151–9. doi: 10.1136/annrheumdis-2018-214819

46. Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord.* (2008) 23(15):2129–70. doi: 10.1002/mds.22340

47. Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. *Neurology.* (1998) 50(2):318. and 16 pages following. doi: 10.1212/WNL.50.2.318

48. Martín-Carbonell M, Riquelme-Marín A, Ortigosa-Quiles JM, Meda RM, Córdova AM, Enríquez-Santos JA, et al. Utilidad del inventario de depresión rasgo-estado (IDERE) para el diagnóstico clínico. Un estudio con muestras de cuatro países iberoamericanos. *Anal Psicol.* (2013) 28(3):763–71. doi: 10.6018/analesps.28.3.133711

49. Spielberger CD. Manual for the state-trait anxiety inventory (self-evaluation questionnaire) (1970). Available online at: https://cir.nii.ac.jp/crid/1370285712575158016 (cited July 12, 2023).

50. Garcia B. 19 escalas de evaluacion psicologica. (2017). Available online at: https://www.academia.edu/31038102/19_ESCALAS_DE_EVALUACION_PSICOLOGICA (cited August 28, 2023).

51. Palomar Lever J, Gómez Valdez NE. Desarrollo de una escala de medición de la resiliencia con mexicanos (RESI-M). *Interdisci Archaeol.* (2010) 27(1):7–22. Available online at: https://www.redalyc.org/pdf/180/18014748002.pdf

52. Oddy M. The use of psychological testing for treatment planning and outcomes assessment (2nd edn) edited by Mark E. Maruish London: Lawrence Erlbaum. 1999. 1507 pp. £200.00 (hb). ISBN 0 80582761 7. *Br J Psychiatry.* (2001) 178:93–4. doi: 10.1192/bjp.178.1.93-a

53. Tobin DL, Holroyd KA, Reynolds RV, Wigal JK. The hierarchical factor structure of the coping strategies inventory. *Cognit Ther Res.* (1989) 13(4):343–61. doi: 10.1007/BF01173478

54. Leentjens AFG, Dujardin K, Pontone GM, Starkstein SE, Weintraub D, Martinez-Martin P. The Parkinson anxiety scale (PAS): development and validation of a new anxiety scale. *Mov Disord.* (2014) 29(8):1035–43. doi: 10.1002/mds.25919

55. Shipley WC, Gruber CP, Martin TA, Klein AM. Shipley-2. Escala breve de inteligencia. Manual Moderno. (2009).

56. Sánchez-Moncada I, Concha L, Merchant H. The timing network is engaged in the practice of internally driven tapping independently of the learning transfer from perceptual to motor timing. *bioRxiv.* (2020):2020.12.17.423301. doi: 10.1101/2020.12.17.423301v2

57. Lizcano-Cortés F, Gómez-Varela I, Mares C, Wallisch P, Orpella J, Poeppel D, et al. Speech-to-speech synchronization protocol to classify human participants as high or low auditory-motor synchronizers. *STAR Protoc.* (2022) 3(2):101248. doi: 10.1016/j.xpro.2022.101248

58. Ruiz-Contreras AE, Soria-Rodríguez G, Almeida-Rosas GA, García-Vaca PA, Delgado-Herrera M, Méndez-Díaz M, et al. Low diversity and low frequency of participation in leisure activities compromise working memory efficiency in young adults. *Acta Psychol.* (2012) 139(1):91–6. doi: 10.1016/j.actpsy.2011.10.011

59. Hampshire A, Highfield RR, Parkin BL, Owen AM. Fractionating human intelligence. *Neuron.* (2012) 76(6):1225–37. doi: 10.1016/j.neuron.2012.06.022

60. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* (2005) 53(4):695–9. doi: 10.1111/j.1532-5415.2005.53221.x

61. Parker G, Tupling H, Brown LB. A parental bonding instrument. *Br J Med Psychol.* (1979) 52(1):1–10. doi: 10.1111/j.2044-8341.1979.tb02487.x

62. Bernstein DP, Fink L, Handelsman L, Foote J, Lovejoy M, Wenzel K, et al. Initial reliability and validity of a new retrospective measure of child abuse and neglect. *Am J Psychiatry.* (1994) 151(8):1132–6. doi: 10.1176/ajp.151.8.1132

63. Rao SC, Pattipati S, Salim A, Gutierrez N, Naito A, Ghosh K, et al. Promoting gender inclusivity: a questionnaire for women's health factors in Parkinson's disease. *Mov Disord.* (2022) 37(5):1112–3. doi: 10.1002/mds.29020

64. Jimenez-Genchi A, Monteverde-Maldonado E, Nenclares-Portocarrero A, de la Vega-Pacheco GEA. Confiabilidad y análisis factorial de la versión en español del índice de calidad de sueño de Pittsburgh en pacientes psiquiátricos. *Gac Med Mex.* (2008) 144(6):491–6. Available online at: https://www.anmm.org.mx/GMM/2008/n6/27_vol_144_n6.pdf

65. World Health Organization (WHO). The world health organization quality of life assessment (WHOQOL): position paper from the world health organization. *Soc Sci Med*. (1995) 41(10):1403–9. doi: 10.1016/0277-9536(95)00112-K

66. Zúniga MA, Carrillo-Jiménez GT, Fos PJ, Gandek B, Medina-Moreno MR. Evaluation of health status using survey SF-36: preliminary results in Mexico. *Salud Publica Mex*. (1999) 41(2):110–8. doi: 10.1590/S0036-36341999000200005

67. McElhone K, Abbott J, Shelmerdine J, Bruce IN, Ahmad Y, Gordon C, et al. Development and validation of a disease-specific health-related quality of life measure, the LupusQol, for adults with systemic lupus erythematosus. *Arthritis Rheum*. (2007) 57(6):972–9. doi: 10.1002/art.22881

68. Boyle GJ, Matthews G, Saklofske DH. Personality theories and models: an overview. In: *The SAGE handbook of personality theory and assessment. Vol. 1. Personality Theories and Models*. Los Angeles: Sage Publications (2008). p. 1–29.

69. Muris P, Meesters C, de Kanter E, Timmerman PE. Behavioural inhibition and behavioural activation system scales for children: relationships with eysenck's personality traits and psychopathological symptoms. *Pers Individ Dif*. (2005) 38 (4):831–41. doi: 10.1016/j.paid.2004.06.007

70. Fadanelli B, del Pilar M. Actitud hacia la ingesta y peso corporal. *Acta Médica Grupo Angeles*. (2014) 12(4):181–8. Available online at: https://www.medigraphic.com/pdfs/actmed/am-2014/am144b.pdf

71. World Health Organization (WHO). Global Physical Activity Questionnaire (GPAQ) (2002). Available online at: https://www.who.int/publications/m/item/global-physical-activity-questionnaire (cited August 28, 2023).

72. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The fagerström test for nicotine dependence: a revision of the fagerström tolerance questionnaire. *Br J Addict*. (1991) 86(9):1119–27. doi: 10.1111/j.1360-0443.1991.tb01879.x

73. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the alcohol use disorders identification test (AUDIT): wHO collaborative project on early detection of persons with harmful alcohol consumption–II. *Addiction*. (1993) 88(6):791–804. doi: 10.1111/j.1360-0443.1993.tb02093.x

74. Adamson SJ, Sellman JD. A prototype screening instrument for cannabis use disorder: the Cannabis use disorders identification test (CUDIT) in an alcohol-dependent clinical sample. *Drug Alcohol Rev*. (2003) 22(3):309–15. doi: 10.1080/0959523031000154454

75. Nucci M, Mapelli D, Mondini S. Cognitive reserve Index questionnaire (CRIq): a new instrument for measuring cognitive reserve. *Aging Clin Exp Res*. (2012) 24 (3):218–26. doi: 10.1007/BF03654795

76. Ordoñana JR, Bartels M, Boomsma DI, Cella D, Mosing M, Oliveira JR, et al. Biological pathways and genetic mechanisms involved in social functioning. *Qual Life Res*. (2013) 22(6):1189–200. doi: 10.1007/s11136-012-0277-5

77. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*. (1971) 9(1):97–113. doi: 10.1016/0028-3932(71)90067-4

78. Gliklich RE, Leavy MB, Cosgrove L, Simon GE, Gaynes BN, Peterson LE, et al. Harmonized outcome measures for use in depression patient registries and clinical practice. *Ann Intern Med*. (2020) 172(12):803–9. doi: 10.7326/M19-3818

79. Hopper JL, Foley DL, White PA, Pollaers V. Australian twin registry: 30 years of progress. *Twin Res Hum Genet*. (2013) 16(1):34–42. doi: 10.1017/thg.2012.121

80. Moayyeri A, Hammond CJ, Hart DJ, Spector TD. The UK adult twin registry (TwinsUK resource). *Twin Res Hum Genet*. (2013) 16(1):144–9. doi: 10.1017/thg.2012.89

81. Bivol S, Mellick GD, Gratten J, Parker R, Mulcahy A, Mosley PE, et al. Australian Parkinson's genetics study. *BMJ Open*. (2022) 12(2):e052032. doi: 10.1136/bmjopen-2021-052032

82. Zabetian CP, Mata IF. Latin American Research Consortium on the genetics of PD (LARGE-PD). LARGE-PD: examining the genetics of Parkinson's disease in Latin America. *Mov Disord*. (2017) 32(9):1330–1. doi: 10.1002/mds.27081

83. Vailati-Riboni M, Palombo V, Loor JJ. What are omics sciences? In: Ametaj BN, editors. *Periparturient Diseases of Dairy Cows: A Systems Biology Approach*. Cham: Springer International Publishing (2017). p. 1–7.

84. GP2. The global Parkinson's genetics program. *Mov Disord*. (2021) 36 (4):842–51. doi: 10.1002/mds.28494

85. Bandres-Ciga S, Faghri F, Majounie E, Koretsky MJ, Kim J, Levine KS, et al. Neurobooster array: a genome-wide genotyping platform to study neurological disorders across diverse populations. *medRxiv* [preprint]. (2023). Available online at: https://www.medrxiv.org/content/10.1101/2023.11.06.23298176v1 (Accessed June 12, 2023).

86. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. *J Multidiscip Healthc*. (2016) 9:211–7. doi: 10.2147/JMDH.S104807

87. Heinze G, Canchola VHO, Miranda GB, Fuentes NAB, Sánchez DPG. Los médicos especialistas en México. *Gac Med Mex*. (2018) 154(3):342–51. doi: 10.24875/GMM.18003770

88. Instituto Nacional de Estadística y Geografía (INEGI). Instituto Nacional de Estadística y Geografía (2022). Demografía y Sociedad. Available online at: https://www.inegi.org.mx/temas/ticshogares/ (cited March 8, 2022).

89. Zwan MD, van der Flier WM, Cleutjens S, Schouten TC, Vermunt L, Jutten RJ, et al. Dutch brain research registry for study participant recruitment: design and first results. *Alzheimers Dement*. (2021) 7(1):e12132. doi: 10.1002/trc2.12132

90. Obrutu O, Maughan J, Tjoe B, Herscovici R, Moy P, Rojas N, et al. Smidt Heart Institute Takotsubo Registry—study design and baseline characteristics. *Amer Heart J Plus Cardiol Res Pract*. (2022) 13:100086. doi: 10.1016/j.ahjo.2022.100086

91. Conacyt CN de C y. T. Informe general del estado de la Ciencia, la Tecnología y la Innovación. Conacyt Ciudad de México (2018).

92. Román-López TV, García-Vilchis B, Murillo-Lechuga V, Chiu-Han E, López-Camaño X, Aldana-Assad O, et al. Estimating the genetic contribution to astigmatism and myopia in the Mexican population. *Twin Res Hum Genet*. (2023) 16:1–9. doi: 10.1017/thg.2023.41

93. Pedraza-Meza ML, Hernández-Ledesma AL, Ruiz-Contreras AE, Medina-Rivera A, Martínez D. Clinical, psychosocial and demographic factors affect decisions in SLE people. *medRxiv*. (2024):2024.03.25.24304643. doi: 10.1101/2024.03.25.24304643v1.abstract