biobank.cy: The Biobank of Cyprus past, present and future

Eleni M. Loizidou[1], Maria Kyratzi[1], Maria A. Tsiarli[1,2], Andrea C. Kakouri[1], Georgia Charalambidou[1], Stella Antoniou[1], Stylianos Pieri[1], Panagiota Veloudi[1], Michaela Th. Mayrhofer[3], Andrea Wutte[3], Lukasz Kozera[3], Jens Habermann[3], Heimo Muller[4], Kurt Zatloukal[4], Karine Sargsyan[4], Alexandros Michaelides[5], Maria Papaioannou[6], Christos Schizas[6], Apostolos Malatras[1], Gregory Papagregoriou[1], Constantinos Deltas[1,7*]

[1]biobank.cy, Center of Excellence in Biobanking and Biomedical Research, University of Cyprus, Nicosia, Cyprus.

[2] Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, USA

[3]Biobanking and BioMolecular Resources Research Infrastructure-European Research Infrastructure Consortium (BBMRI-ERIC) Graz, Austria

[4]Diagnostic & Research Institute of Pathology, Medical University of Graz, Graz, Austria

[5]RTD TALOS LTD

[6]Cyprus National e-Health Authority

[7]School of Medicine, University of Cyprus, Nicosia, Cyprus

*Corresponding author:

Prof. Constantinos Deltas

University of Cyprus, School of Medicine

1, University Avenue, 2109 Nicosia

E-mail: Deltas@ucy.ac.cy

Tel: +357 22892882

Abstract

Background: The Cyprus Biobank collects biosamples and medical, biological, and lifestyle information with the aim of reaching 16,500 Cypriot volunteers aged ≥18-years, by year 2026. Volunteers are both from the general population and from disease cohorts of focused research projects, who amongst others will contribute to canvas the architecture of the Cyprus human genome and study the healthy and morbid anatomy of Cypriots, a population with a unique historical and epidemiological background. The Cyprus Biobank is a research infrastructure pillar of the biobank.cy Center of Excellence in Biobanking and Biomedical Research.

Methods: Within 3-years (November 2019-October 2022), 1348 participants of the general population were enrolled in the Cyprus Biobank pilot study. Extensive information was collected from each participant at enrolment, including biochemistry, complete blood count, physiological, anthropometric, socio-demographic, diet, and lifestyle characteristics. Prevalent health conditions along with medication use and family history were recorded, including 58 biomarkers based on blood and urine samples. With a systematic recruitment campaign, the Biobank is continuously increasing the number of individuals in the general population cohort and is developing separate disease cohorts of the Cypriot population.

Results: The pilot study enrolled 1348 participants (579 men and 769 women), aged between 18-85 years (median 48-years). The enrollment takes 40 minutes on average, including the collection of biological samples and phenotypic information. More than half (55%) of the pilot participants are educated to college level or above. Statistically significant differences were found between men and women regarding their education level ($p < 0.001$), marital status ($p = 0.01$) and employment status ($p < 0.001$) but not their age ($p = 0.29$). The most prevalent medical conditions recorded within the studied population are hypertension (17.2%), osteoporosis (6.9%) and diabetes (6.0%).

Conclusions: The Cyprus biobanking pilot study has successfully collected extensive baseline information from enrolled participants. The Biobank will comprise a rich data resource used to examine the major risk factors leading to public health burdens and develop strategies for disease prevention. In the same context,

the first 1000 whole exomes, as the Phase-1 of the Cyprus Human Genome Project, CYPROME, have been completed and are being analysed.

Keywords: Biobank, Cohort Study, Public Health, Population Health, Non-communicable diseases

Introduction

Cyprus is the third largest island in the Mediterranean Sea and has been at the crossroad of many civilizations over the centuries. Its prestigious strategic location within the southeast corner of the Mediterranean basin, bridging the Middle East and Europe, always attracted foreign rulers that sought to control the area and thereby extend the trade pathways. The island of Cyprus was first inhabited during the aceramic neolithic period by peoples from the Middle East and Asia Minor, more than 10,000 years ago, according to archeological evidence and genetic phylogeographic investigations[1]. Over the years it was influenced, or was occupied by, several rulers (Table 1), the most recent occupations of which were the Ottoman Empire between the years of 1571-1878, after which it was taken over by the British, 1878-1960. After a liberation struggle, 1955-1959, the Republic of Cyprus was first established as an independent country in 1960, at which time it joined membership to the United Nations. In 2004, Cyprus joined the European Union and in 2008 adopted the Euro currency. The island's recent and mainly older tumultuous history, including major and minor migration waves, has had its contribution to the Cypriot gene pool, as evidenced by several genetic studies performed over the past 30 years[2].

Table 1. Brief report of visitors-conquerors and key events in the history of Cyprus
• Proto-Cypriots, settlers from Asia Minor, Middle and Near East, more than 12,000 years ago • Mycenaeans (1400–1300 BC), Achaeans (1230–1190 and 1150–1050 BC) • Phoenicians, 950–850 BC • Assyrians, 709–669 BC • Egyptians, 565–546 BC • Persians (fights with Cypriot kings Onisilos and Evagoras, Athenian Engagement and Persian Victory), 546–332 BC • Alexander the Great – Expulsion of the Persians, 332–323 BC • Ptolemies of Egypt, successors of Alexander the Great • Roman Empire, 58 BC • Byzantine Empire, 330 AD • Arab raids, 647–965 AD (Relocation to Artaki, 690 AD) • Crusaders, Richard Lionheart & Templars, 1191–1192 AD • Lusignan French, 1192–1474 AD • Venetians, 1474–1571 AD • Ottoman Empire, 1571–1878 AD • British Colonialism, 1878–1960 AD • Independence and accession of the new state to the British Commonwealth and the United Nations, 1960 AD • Accession to the European Union, 2004 • 1974–present, emigration to and from Cyprus, waves of foreign workers, whether citizens or non-citizens of the European Union

In the past 50 years, the Cypriot population has experienced major economic and lifestyle changes, due to major historical events and the economic crisis that has impacted especially the Eastern European countries in the past 10 years. According to a 2021 census, the population of Cyprus was estimated at 74% Greek Cypriots, including three small religious minorities of the Maronites, Armenians, and Latins, as well as Turkish Cypriots [3],[4]. Since the Turkish invasion in 1974, the population of Cyprus in the European non-occupied part of the island has increased from 648,585 to 923,272 individuals [4], not necessarily of Greek or Turkish origin. Despite the recent COVID-19 pandemic, life expectancy has only slightly changed between 2015 and 2021 (81.7 vs 81.2 years on average, respectively).

Chronic non-communicable diseases (NCDs), such as cardiovascular disease (CVD), diabetes, thyroid disorders, and cancer, pose a substantial health burden, significantly impacting morbidity and mortality rates. According to a 2018 report by the World Health Organization (WHO), NCDs accounted for over 90% of morbidity in Cyprus [5]. This prevalence of NCDs can be attributed to factors including lifestyle, dietary habits, and tobacco use. Additionally, genetic factors and gene-environment interactions also contribute

87 significantly to NCD development. Regional variation in NCDs has also been reported, suggesting the
88 involvement of non-genetic factors that have yet to be identified. However, the presence of potential
89 limitations in sample collection, such as the lack of representative samples and under-representation of
90 specific regions may contribute to the fragmented understanding of NCD landscape in Cyprus [5].

91 Although most Cypriots declare a healthy status, two out of five suffer from a chronic disease, with CVD,
92 cancer, and diabetes being the most prevalent conditions [6]. While CVD rates have been steadily decreasing
93 since 2015, cancer rates remain at the same levels, albeit lower compared to the rest of the EU. On the
94 contrary, tobacco use rates in Cyprus are amongst the highest in the EU, and together with unhealthy dietary
95 habits, they comprise two of the most frequent mortality risk factors in Cyprus (19% and 14%, respectively)
96 [6]. Additionally, while obesity rates in Cyprus are comparable to the average rates of other EU countries,
97 childhood obesity is rapidly increasing and poses a significant public health issue for Cyprus.

98 To gain insights into the etiology and risk factors associated with NCDs and address limitations including the
99 sampling of data from a diverse and representative population, the establishment of a Cypriot Biobank was
100 imperative. While some causes of these diseases are already known based on studies conducted in other
101 populations and regions [7, 8] current knowledge is insufficient to explain the substantial geographic
102 differences in disease rates observed globally and within Cyprus itself [9]. The development of a Cypriot
103 Biobank will enable large-scale studies of both genetic and environmental contributions, as well as their
104 interactions, in the development of chronic diseases. Furthermore, it will address the limitations of previous
105 studies, including inadequate sample sizes, lack of blood and/or urine samples, inconsistencies in data
106 collection and processing, limited data collection on risk exposures (e.g. lifestyle, diet, disease comorbidities)
107 and outcomes, and restricted study populations.

108 The biobank.cy Center of Excellence (CoE) in Biobanking and Biomedical Research was established in 2019 at
109 the University of Cyprus (UCY), after funding through the Horizon 2020 Widening Participation/Teaming
110 program (Grant Agreement no. 857122). The program is a partnership of the UCY with two Advanced
111 Partners, the European Biobanking consortium BBMRI-ERIC [10] and the Medical University of Graz [11] , and
112 a local Partner, the small to medium-sized enterprise TALOS Rtd LTD [12] . Through this partnership,

biobank.cy has successfully received funding to advance the country's medical research infrastructure to match European standards, with two major objectives: a) Upgrade the first Biobank of the country and complete the Cyprus Human Genome Project; b) build the capacity and upgrade its research infrastructure with the aim to prepare Cyprus for the next generation of biomedical research. For the history, biobank.cy is the result of upgrading the Molecular Medicine Research Center (MMRC) which was first established in 2010 with another strategic funding of the Regional Development Fund of the European Union and the Republic of Cyprus through the Cyprus Research Promotion Foundation, as part of the Smart Specialization Strategy of the country. In its current status, the Biobank is one of five Pillars of the biobank.cy CoE, the other four being, the MMRC as the research arm of the CoE, and the Molecular Diagnostics, the Education and Innovation pillars.

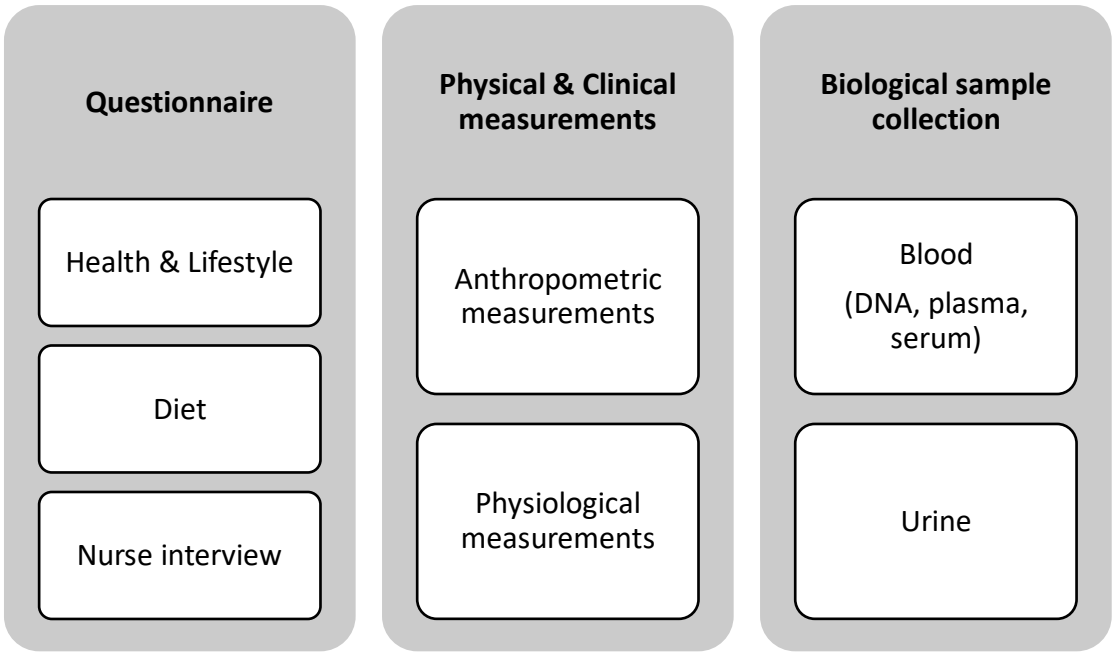
Towards the public health benefit and to support the development of the Cyprus Human Genome Project, the Government of Cyprus, and the University of Cyprus co-funded biobank.cy. As a first milestone, up to 16,500 permanent Cypriot residents will be invited to partake in the biobank cohort, voluntarily, until 2026. Data collection and storage includes baseline clinical, biochemical, and behavioral traits as well as blood and urine samples.

We herein provide comprehensive information on the operation of the first Biobank in Cyprus, sharing data on the demographics and health archives of participants. Some associations are also attempted with the understanding and anticipation that deeper and more useful results will be derived as the Biobank is enriched over the next years.

Methods

The biobank.cy, as an EU-funded CoE in Research and Innovation, hosts and maintains the Cyprus Biobank, which is not based in a hospital setting but enrolls volunteers who respond to open invitations while also collaborating with public and private clinics and ad hoc volunteer enrolment centers around the island. The Biobank obtained first approval by the Cyprus National Bioethics Committee in November 2011 (approval code: ΕΕΒΚ/ΕΠ/2011/04).

139 It gathers information through a structured and comprehensive questionnaire, clinical phenotyping, and
140 biological samples collected from Cypriot residents of the general population, aged ≥ 18 years old. Population
141 data of the pilot study are collected and stored using REDCap® [13, 14] and OpenSpecimen® [15] tools
142 respectively, to enable data security and tracking as well as to eliminate collection bias. Specifically, REDCap®
143 (Research Electronic Data Capture) is a secure, web-based application designed for data collection and
144 management in research studies, used to amass the following measurements: personal and family history as
145 well as clinical (complete blood count, biochemical) and biometric (blood pressure, height, weight).
146 OpenSpecimen® is used as the Biobank's Information Management System, by organizing information
147 related to the collection, processing, storage, and distribution of samples. Figure 1 presents a summary of
148 the full primary data collected at the Biobank.



149

150 **Figure 1: Measurement and samples collected at baseline assessment**

151 **Participant/volunteer enrolment**

152 Since the very first enrollment on the 1st of November 2019, participant awareness and engagement in the
153 project were achieved through media and the press (television, radio, social media and newspapers, leaflets
154 respectively) as well as social events which are organized by the Biobank personnel. Additionally, friends and
155 family referrals were frequent throughout the 3-year period. Interested individuals book an appointment

either through the biobank.cy website [16] or directly through a phone call. At baseline, participants are invited to the central collection point of the Biobank facility at the Shacolas Educational Center for Clinical Medicine, or elsewhere during outreach activities, where they undergo a 3-stage interview (25 min on average) and physical and clinical measurements (15 min on average) (Figure 2; Table 2). The Biobank participation process commences with trained research nurses providing an initial briefing to the participants. During this session, participants are fully informed about the research project's purpose, their rights, the procedures for participation and withdrawal, while they may ask questions about the project, prior to signing a broad consent form. The consent form is approved by the Cyprus National Bioethics Committee (approval code: ΕΕΒΚ/ΕΠ/2020/04), ensuring compliance with ethical requirements. A copy of the signed consent form is given to the participant.

Upon completion of the informed consent process, participants are requested to fill out a questionnaire providing essential information about their personal and family medical history. This information is securely collected in dedicated REDCap® projects.

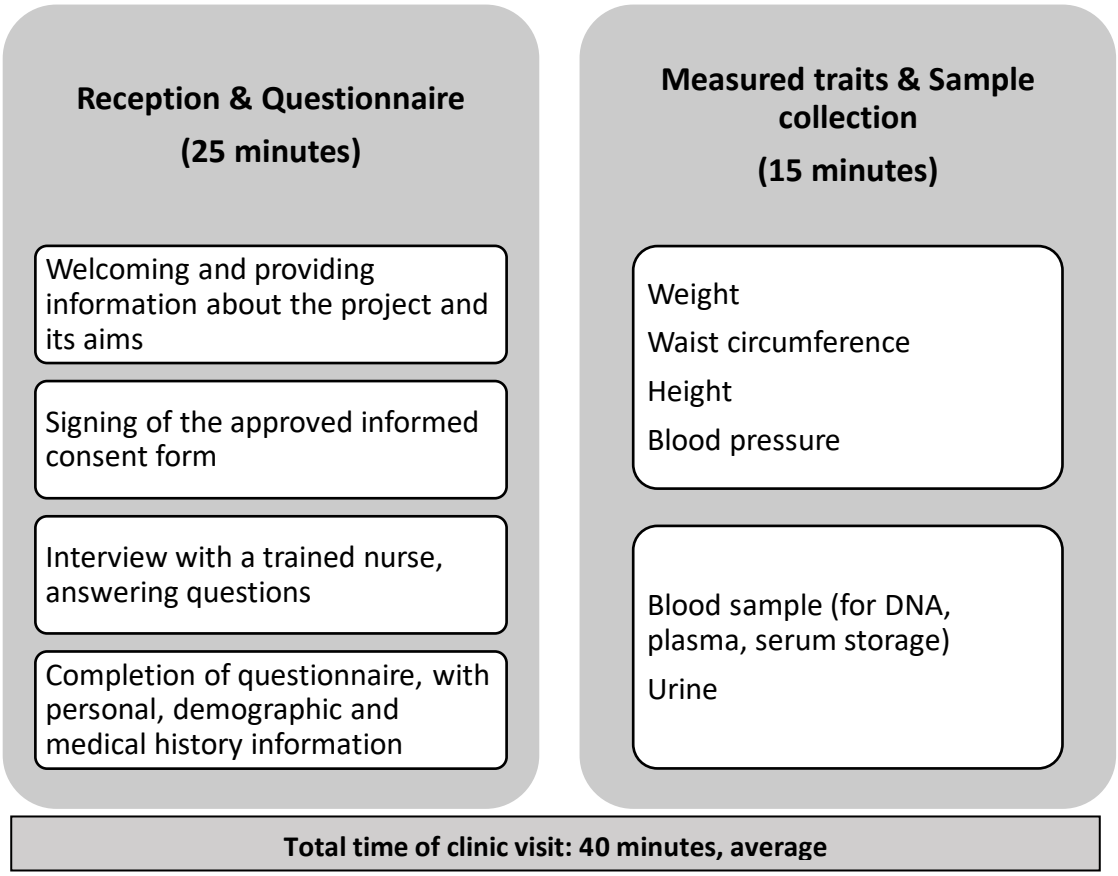


Figure 2: Baseline visit summary

Health and lifestyle questionnaire administered by a trained research nurse

In addition to data of personal character, the health and lifestyle section of the questionnaire includes questions on socio-demographic factors, including the origin of both parents, medical conditions along with medication used, family history and quality-of-life (QoL) questions, including smoking behavior, dietary habits, alcohol consumption, intensity of physical activity [17] and sleeping patterns [18], education level and employment status. As the questionnaire is part of the registration and clinic visit, a Biobank-trained research nurse is always available to answer any questions of the participant.

Importantly, questions on dietary assessment are directly linked to the Mediterranean score (MedDietScore) [19] which indicates the level of the participants' adherence to the traditional Mediterranean diet. This specific type of diet has been widely used for the past ~20 years as it has been proven to significantly reduce total mortality [20], thus having a great impact on clinical and public health. More specifically, a higher MedDiet score showed strong evidence against incidents of cardiovascular disease, diabetes, hypertension, obesity, and Alzheimer's disease among others [21]. Participants are specifically asked to report the consumption frequency of whole grain foo/d (bread, pasta, rice etc), potatoes, fruit and juice, vegetables and salads, legumes, fish and soup, red meat and its products/derivatives, poultry, full-fat dairy, and olive oil.

Table 2: Instruments used at the Biobank for data measurements

Measurement category	**Instrument**
Questionnaires	
Health & Lifestyle	Questions on socio-demographic factors, medical conditions, medication use, family history, smoking behavior, alcohol consumption, physical activity levels and sleeping patterns
Nurse Interview	Additional questions on QoL assessment (smoking, alcohol consumption, and physical activity frequency), family status, doctor-diagnosed conditions, education, employment status and dietary patterns
Physical and clinical measurements	
Anthropometric	Height, weight, and waist circumference measurements*
Physiological	Blood pressure using a digital sphygmomanometer

Biological samples	
Biological samples collected	Whole peripheral blood, urine

195 **Waist circumference measures available for 960 individuals*

196 **Physical-clinical measurements and biological material collected**

197 At baseline, anthropometric and physiological traits were measured for each Biobank participant along with
198 blood and urine sample collection. Anthropometric and physiological traits comprised body weight/height
199 and diastolic/systolic blood pressure respectively. Blood pressure was measured once, in a seated position.
200 A blood pressure measurement was repeated if the first measurement returned a value of >160mmHg for
201 systolic blood pressure or <80 mmHg for diastolic blood pressure.

202 At baseline enrolment visit, the participants provided peripheral blood for the isolation of DNA, plasma, and
203 serum, which are stored at -80° C. Urine was also collected and aliquots were stored likewise. Specifically, we
204 collected: 1x10ml EDTA for DNA extraction, 1x4ml EDTA for FBC, 1x4ml EDTA for plasma, 1x8.5ml serum,
205 1x3.5ml serum (prompt test), and 1x100ml urine. Aliquoted and stored are 4 aliquots of 50µl DNA at 50ng/µl,
206 plus a concentrated DNA stock, 4 aliquots of 2ml serum, 2 aliquots of 1ml plasma, and 5 aliquots of 2ml urine.

207 Fifty-eight clinical biomarkers are measured by a private diagnostic provider using a proportion of the stored
208 blood (Table 3).

Table 3: Clinical biomarkers measured at biobank.cy (N = 58)	
Biomarker group	**Measure**
Blood biochemistry	
Lipid profile	Total cholesterol
	Low density lipoprotein
	High density lipoprotein
	Triglycerides
Diabetes related tests	Glucose
	Glycated Haemoglobin A1c %
Bone and joint markers	Uric acid
Complete Blood Count	
White blood cells	White blood cells

Differential white cell count	Lymphocytes
	Lymphocytes %
	Eosonophils
	Eosonophils %
	Monocytes
	Monocytes %
	Basophils
	Basophils %
	Neutrophils
	Neutrophils %
	LUC
	LUC %
Red blood cells	Haemoglobin
	Red blood cells
	Haematocrit
	Mean Corpuscular Hemoglobin (MCH)
	Mean corpuscular hemoglobin concentration (MCHC)
	Mean corpuscular volume (MCV)
	Red cell distribution width coefficient of variation (RDW CV)
	Haemoglobin Distribution Width (HDW)
Platelets	Platelets
	Platelet Distribution Width (PDW)
	Mean Platelet Volume (MPV)
	Plateletcrit (PCT)
Urine Biochemical and Microscopy	
Biochemistry	Protein
	Glucose
	Ketones
	Bilirubin
	Urobilinogen
	Nitrates
	Erythrocytes/Haemoglobin (Ery/µL)
	Leucocytes (µ/L)
Macroscopy	Appearance

	Colour
	Reaction (pH)
	Specific gravity
Microscopy	WBC (/hpf)
	RBC (/hpf)
	Epithelial cells
	Casts
	Casts Hyaline
	Cast Granular
	Crystals
	Microorganisms
	Yeasts
Urine Biochemistry	
Microalbumin	Albumin
	Microalbumin
	Urine Albumin/Creatinine ratio
Protein/Creatinine Urine Ratio	Proteins
	Creatinine
	Urine Protein/Creatinine ratio

Results

Socio-demographic characteristics

The participants of the Biobank pilot study comprised both men (43%) and women (57%) spanning an age range from 18-85 years, with a median age of 48 years (Fig. 3). More than half of men (55.8%) and women (54.6%) were educated to a university degree level and above, while the most frequent employment status for both men and women was paid employee (62% and 63% respectively) (Table 4). The most common status of the enrolled individuals was either single or married (20.9% and 72.8% respectively).

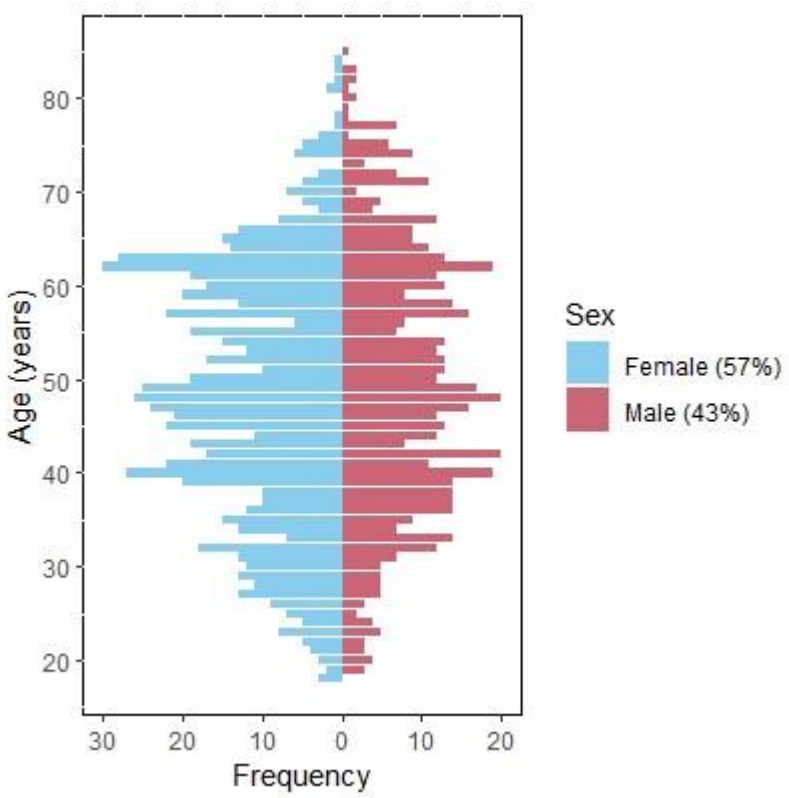


Figure 3: Age distribution of men and women enrolled in the Biobank pilot study

Table 4: Socio-demographic characteristics of the enrolled participants in the Biobank pilot study (October 2019 – October 2022)

Characteristic	**Men**	**Women**	
	N (%)	**N (%)**	***P*-value**
	579 (43)	769 (57)	
Age group (years)			0.29
< 25	22 (3.8)	30 (3.9)	
25-34	65 (11.2)	116 (15.1)	
35-44	135 (23.3)	163 (21.2)	
45-54	141 (24.4)	192 (25.0)	
≥ 55	216 (37.3)	268 (34.9)	
Marital status			0.01
Single	128 (22.1)	154 (20.0)	
Married	429 (74.1)	552 (71.8)	
Divorced	18 (3.1)	44 (5.7)	
Widowed	4 (0.7)	18 (2.3)	
Civil partnership	0 (0)	1 (0.1)	
Education			< 0.001
Lower than primary school	1 (0.1)	3 (0.4)	
Primary school	17 (2.9)	34 (4.4)	

Secondary school	161 (27.8)	167 (21.7)	
Technical/professional school	38 (6.6)	21 (2.7)	
College diploma (degree not equivalent to a BSc/BA)	39 (6.7)	124 (16.1)	
Bachelor's degree	200 (34.5)	261 (33.9)	
Master's degree	91 (15.7)	141 (18.3)	
PhD degree	32 (5.5)	18 (2.3)	
Employment status			< 0.001
Paid employee	358 (61.8)	486 (63.2)	
Self-employed	50 (8.6)	44 (5.7)	
Housewife	0 (0)	41 (5.3)	
Retired	87 (15.0)	76 (9.9)	
Unemployed	12 (2.1)	27 (3.5)	
Other	33 (5.7)	33 (4.3)	
Not reported	39 (6.7)	62 (8.1)	

227 *P*-values for the comparison between sex groups were calculated using Chi-squared and Fisher's exact tests

228 Medical history

229 Based on the NHS and Heart UK guidelines (https://www.nhs.uk/conditions/high-cholesterol/cholesterol-
230 levels/) (Table 4), almost half of the participants had high total (N = 681) and LDL (N = 630) cholesterol levels
231 (151/681 and 150/630 post-fasting respectively), while 15.8% (N = 213, 31/213 post-fasting) had low HDL
232 levels. Additionally, 19% (N = 257, 34/257 post-fasting) of individuals had high levels of triglycerides. The
233 majority of these participants did not report any medication use, despite the abnormal features in their blood
234 lipid profile (Table 5). The results of our pilot study are in line with recent findings regarding the high
235 prevalence of elevated blood lipid levels in the European population [22-24].

236 Although the vast majority of the participants self-reported as healthy, a variable proportion of them had
237 underlying past or current conditions, with the most prevalent being hypertension (17.2%), osteoporosis
238 (6.9%) and diabetes (6.0%) (Fig. 4). Additional conditions recorded include cardiovascular disease (myocardial
239 infarction and stroke), rheumatoid arthritis, osteoarthritis, asthma, kidney disease, cancer and depression
240 (Fig.4). Most individuals with a previously diagnosed condition used medication (Table 5).

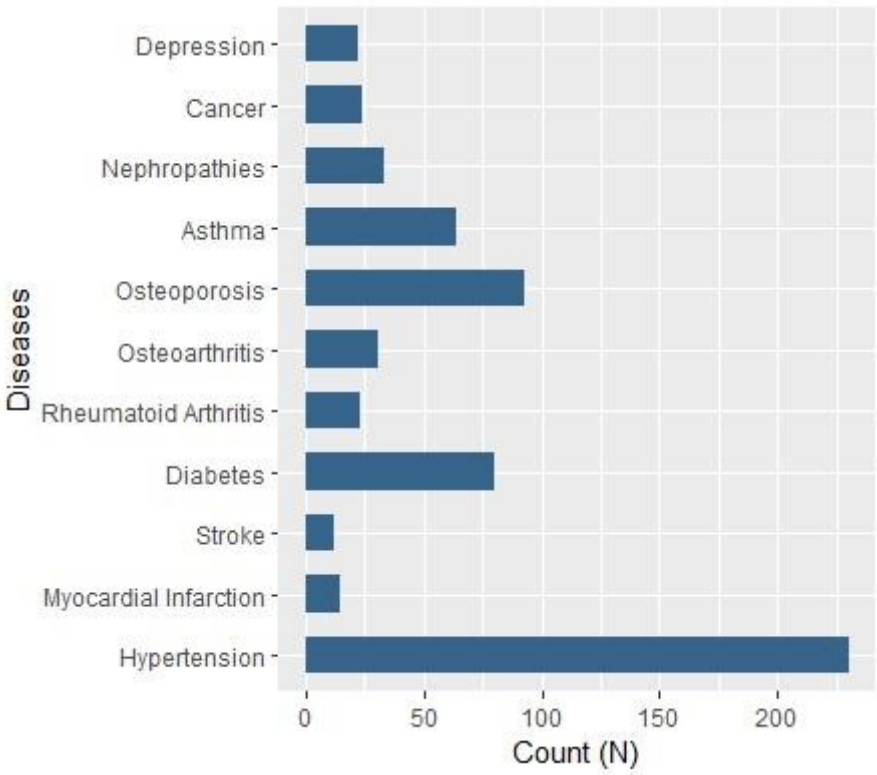


Figure 4: Diseases recorded among the participants of the pilot study

Table 5: Medical history and medication use (N = 1348)

Disease		**N (%)**	**Medication use (%)**
Lipid profile*	TC > 5 mmol/L	681 (51)	102 (15)
	LDL-C > 3 mmol/L	630 (47)	80 (12.7)
	HDL-C < 1 mmol/L (men) < 1.2 mmol/L (women)	 117 (8.7) 96 (7.1)	 24 (20.5) 16 (16.7)
	TG > 1.7 mmol/L	275 (20.4)	85 (30.9)
Hypertension‡		230 (17.2)	219 (16.3)
Myocardial Infarction‡		15 (1.1)	13 (1.0)
Stroke‡		12 (0.9)	10 (0.7)
Diabetes‡		80 (6.0)	78 (5.8)
Rheumatoid Arthritis‡		23 (1.7)	21 (1.6)
Osteoarthritis‡		31 (2.3)	29 (2.2)
Osteoporosis‡		93 (6.9)	86 (6.4)
Asthma‡		64 (4.8)	45 (3.4)
Kidney conditions‡		33 (2.5)	27 (2.0)
Cancer‡		24 (1.8)	7 (0.5)
Depression‡		22 (1.6)	17 (1.3)

*Abnormal lipid levels reported

‡ Doctor-diagnosed conditions

TC = Total cholesterol; LDL-C = LDL cholesterol, HDL-C = HDL cholesterol, TG = Triglycerides

Discussion

The CY-Biobank

The establishment of the first Biobank in Cyprus, in 2011, and its enhancement in 2019 as part of the biobank.cy Center of Excellence, provides an opportunity for ground-breaking research both nationally and internationally, through a rich collection of high-quality phenotypic data that will unveil the causes of common and rare diseases affecting the Cypriot population. The data are retrieved from organized questionnaires, the collection of blood and urine samples and detailed clinical measurements. The Biobank aims to reach a total sample size of ~ 16,500 individuals by 2026 and intends to follow a sub-group of volunteers in the long-term to record disease progression, based on criteria to be determined. After informed consent from the participants, the data are stored for research purposes with the goal to canvas the architecture of the Cyprus human genome. biobank.cy is equipped with clinic, laboratory and liquid sample storage facilities and offices for non-clinical research staff.

biobank.cy was founded and launched as a national research infrastructure in 2019, just before the COVID-19 pandemic started. Despite the challenges and the relatively short period since its establishment, it already has a high impact on the health of the Cypriot population. During the pandemic, in addition to the original scope, biobank.cy collected samples from COVID-19 cases to study the production of circulating immunoglobulin class G (IgG) antibodies against SARS-CoV-2 in the serum of convalescent individuals [25]. Individuals who participated in the project received the results from their sample collection two weeks later, thus gaining the opportunity to discuss potential abnormal ranges with their personal doctors. The whole exome sequencing (WES) of 1500 convalescent participants in the project is pending.

According to the data collected in this pilot study, volunteers diagnosed with hypertension represent 17.2%, diabetics represent 6%, patients with osteoporosis 6.9%, and patients with asthma 4.8% of the cohort presented. With the understanding that these were not targeted patient groups, it is obvious that with the projected increase in participants, the Cypriot population will soon be represented in greater depth and lead to safer conclusions, with its separate disease cohorts available for further research.

Future plans

Cyprus Human Genome Project

biobank.cy was formed with the vision of being the island's central node for archiving high-quality biomaterial and associated clinical records, with state-of-art research infrastructure facilities utilized for precision medicine. Towards this goal, biobank.cy has launched the largest research study to ever run in Cyprus, with the precise aim of understanding and treating diseases by characterizing the Cypriot genome. Specifically, biobank.cy has initiated the Cyprus Human Genome Project (CHGP). The CHGP is a first in its kind of scale and prospect study in Cyprus, with the aim of implementing whole genome sequencing (WGS) of at least 500 family trios, which will be the blueprint for the first Cypriot reference genome. Concurrently, additional volunteers of the general population, including chronic disease patients, are selected for WES based on additional ongoing separately funded research projects. Currently, more than 1000 Cypriot WES have been completed, while additional samples are being processed to reach a projected number of 5,000 exomes in the next 3 years.

The establishment of the Biobank was a landmark development that enabled the materialization of the CHGP as a significant milestone and a turning point for the health of the Cypriot population and precision medicine (Figure 5). The CHGP will provide unprecedented insights in terms of the pathogenic variant load of Cypriots and disease prevalence in the Cypriot population, while it will be of use to many genetic projects across all disciplines and non-communicable diseases investigations. Furthermore, the CHGP will provide the backbone to address specific diseases on the island by documenting their prevalence and revealing putative environmental factors that affect them. Moreover, biobank.cy with its increasing potential as a central medical research hub, will provide an unsurpassable opportunity for Cypriot researchers and medical professionals to expand their collaborations nationally and internationally while it will offer opportunities to Cypriot patients to have access to clinical trials and treatments not currently available in Cyprus.

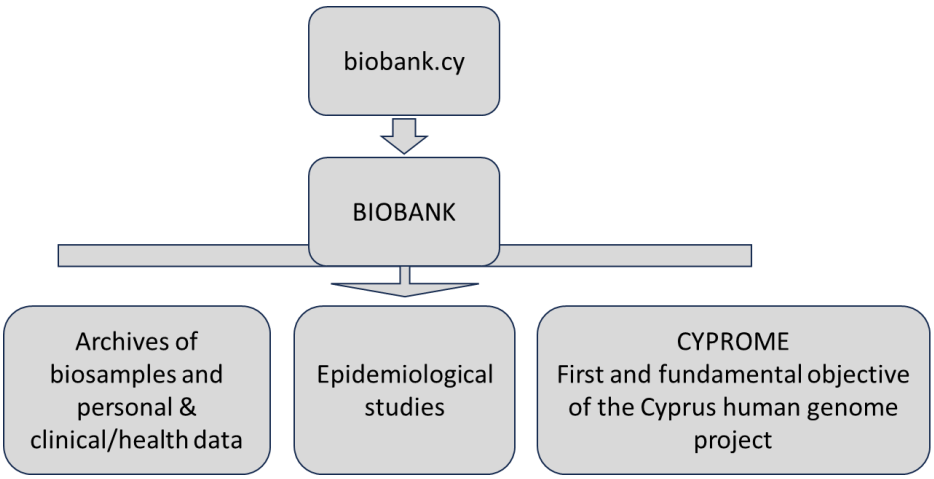


Figure 5: A snapshot of the Biobank activities enabling pioneering work towards precision medicine

Over the past 30+ years, genetic studies in the Cypriot population unraveled many interesting findings concerning mainly rare diseases, including kidney conditions, neurological conditions, endocrinological conditions and others [2]. Founder mutations were discovered for several diseases while a Cypriot consensus whole genome sequence (including the non-coding part) is still lacking, but individual volunteers and family trios are being enrolled to this effect. The CYPROME project is promising to identify more DNA variants of pathogenic potential and enable comparative studies vis-à-vis the minor allele frequencies among the neighboring populations, thus leading to a better understanding of our genetic architecture. Preliminary data suggest that a substantial number of DNA variants have impressively lower or higher frequencies, compared to other European or Asiatic databases, for which detailed analysis is pending.

With regards to banking itself, a major effort is to increase the number of participants to cover the population to a great depth, thereby empowering relevant epidemiological and genetics studies on all pathologies and enabling twin and population screening studies.

biobank.cy, since its inception, actively engages in Cyprus's eHealth ecosystem with a focus on becoming a fully operational eHealth biobank. Aligned with Cyprus's 2019 eHealth law, it emphasizes citizen-centricity and interoperability. The core infrastructure includes a data warehouse (DWH) serving as the nerve center for diverse health data, integrating clinical, samples, laboratory, genomic, and imaging data. The DWH is designed for integration with Cyprus's national EHR system, using the HL7 Fast Healthcare Interoperability Resources (FHIR) standard aiming to actively contribute to primary and secondary health data use. With a

327 vision of becoming a key player in the Cyprus National and European Health Data Space, biobank.cy designs
328 its infrastructure to become an early adopter of initiatives like International and European Patient Summary
329 while it already achieved integration with the BBMRI-ERIC Locator system using HL7 FHIR.

330 In conclusion, as with any other Biobank, the Biobank of Cyprus is creating the prospects for next-generation
331 biomedical research. It is empowering investigations that include epidemiological and genetic determinants
332 with the aim to identify and evaluate the contribution of factors involved in major public health burdens
333 caused by primarily not communicable diseases.

334 **Acknowledgements**

335 This work was funded by the *CY*-Biobank, which is an EU Horizon 2020 Research and Innovation Programme
336 under Grant Agreement no. 857122, the Republic of Cyprus, and the University of Cyprus. The authors thank
337 the research nurses, Biobank technicians, and administrative personnel who contributed to making this
338 operation possible. The Biobank Directorship thanks all participants who voluntarily enrolled in the Biobank
339 and made this momentum task a reality for the health benefit of all.

340 **Author contributions**

341 Material preparation, participant recruitment, collection of biological material and data collection were
342 performed with variable contribution of all authors. Statistical analyses were performed by EML. The first
343 draft of the manuscript was written by EML, CD, GP, AM, MK, ACK, MT, GC and SA and all authors commented
344 on the manuscript. All authors read and approved the final manuscript.

345 **Funding**

346 The biobank.cy pilot study was supported by *CY*-Biobank, an EU Horizon 2020 Research and Innovation
347 Programme under Grant Agreement no. 857122.

348 **Availability of data and materials**

349 The datasets used and/or analysed during the current study are available from the corresponding author,
350 upon a reasonable request.

DECLARATIONS

Ethics approval and consent to participate

Ethical approval was granted from the Cyprus National Bioethics Committee, code ΕΕΒΚ/ΕΠ/2020/04.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Voskarides, K., et al., *Y-chromosome phylogeographic analysis of the Greek-Cypriot population reveals elements consistent with Neolithic and Bronze Age settlements.* Investig Genet, 2016. **7**: p. 1.
2. Deltas, C., *The genetic heritage of Cypriots through special topics of genetics*. 2014, Athens: BHTA Medical Arts. 224.
3. Cyprus, M.o.F.A.-R.o., *Population Census.* 2021.
4. Cyprus, S.S., *Census of population 2021.* 2021.
5. Charalampous, P., et al., *Burden of non-communicable diseases in Cyprus, 1990-2017: findings from the Global Burden of Disease 2017 study.* Arch Public Health, 2021. **79**(1): p. 138.
6. OECD, O.f.E.C.-o.a.D. *State Of Health In The EU* 2021; Available from: https://health.ec.europa.eu/system/files/2021-12/2021_chp_cyprus_cypriot.pdf.
7. Charalampous, P., et al., *Burden of non-communicable disease studies in Europe: a systematic review of data sources and methodological choices.* Eur J Public Health, 2022. **32**(2): p. 289-296.
8. Bai, J., et al., *Global Epidemiological Patterns in the Burden of Main Non-Communicable Diseases, 1990-2019: Relationships With Socio-Demographic Index.* Int J Public Health, 2023. **68**: p. 1605502.
9. Kyprianidou, M., et al., *Prevalence and sociodemographic determinants of selected noncommunicable diseases among the general population of Cyprus: A large population-based cross-sectional study.* Chronic Dis Transl Med, 2023. **9**(3): p. 222-237.
10. BBMRI-ERIC, *BBMRI-ERIC.* 2023.
11. Graz, M.U., *Medical University Graz.*
12. Talos, R., *RTD Talos.* 2023.
13. Harris, P.A., et al., *Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support.* J Biomed Inform, 2009. **42**(2): p. 377-81.
14. Harris, P.A., et al., *The REDCap consortium: Building an international community of software platform partners.* J Biomed Inform, 2019. **95**: p. 103208.
15. Specimen, O., *Open Specimen.* 2023.
16. biobank.cy, *The CY-Biobank.* 2023.
17. Craig, C.L., et al., *International physical activity questionnaire: 12-country reliability and validity.* Med Sci Sports Exerc, 2003. **35**(8): p. 1381-95.
18. Roth, T., et al., *A new questionnaire to detect sleep disorders.* Sleep Med, 2002. **3**(2): p. 99-108.

390 19. Panagiotakos, D.B., et al., *MedDietScore: a computer program that evaluates the adherence to the*
391 *Mediterranean dietary pattern and its relation to cardiovascular disease risk.* Comput Methods
392 Programs Biomed, 2006. **83**(1): p. 73-7.
393 20. Trichopoulou, A., et al., *Adherence to a Mediterranean diet and survival in a Greek population.* N
394 Engl J Med, 2003. **348**(26): p. 2599-608.
395 21. Guasch-Ferre, M. and W.C. Willett, *The Mediterranean diet and health: a comprehensive overview.*
396 J Intern Med, 2021. **290**(3): p. 549-566.
397 22. Timmis, A., et al., *European Society of Cardiology: Cardiovascular Disease Statistics 2017.* Eur Heart
398 J, 2018. **39**(7): p. 508-579.
399 23. Laufs, U., et al., *Why is hypercholesterolaemia so prevalent? A view from evolutionary medicine.* Eur
400 Heart J, 2019. **40**(33): p. 2825-2830.
401 24. Steinhagen-Thiessen, E., et al., *Dyslipidemia in primary care--prevalence, recognition, treatment*
402 *and control: data from the German Metabolic and Cardiovascular Risk Project (GEMCAS).*
403 Cardiovasc Diabetol, 2008. **7**: p. 31.
404 25. Mamais, I., et al., *Circulating IgG Levels in SARS-CoV-2 Convalescent Individuals in Cyprus.* J Clin
405 Med, 2021. **10**(24).

406

407