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

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

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

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

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

[4]RTD TALOS LTD

[5]Cyprus National e-Health Authority

[6]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 marital status ($p = 0.01$), education level ($p < 0.001$), 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%). 1000 whole exomes, as the Phase-1 of the Cyprus Human Genome Project, CYPROME, has been completed.

Conclusions: The Cyprus biobanking pilot study has successfully collected extensive baseline information from enrolled participants. With expanding beyond the pilot study, the Biobank will comprise a rich data

resource that will be used to examine the major risk factors leading to public health burden and to develop strategies for disease prevention.

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 phylogeography 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 (http://www.mfa.gov.cy, https://www.cystat.gov.cy/). 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 (https://www.cystat.gov.cy/), 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 [3]. This prevalence of NCDs can be attributed to factors including lifestyle, dietary habits,

87 and tobacco use. Additionally, genetic factors and gene-environment interactions also contribute
88 significantly to NCD development. Regional variation in NCDs has also been reported, suggesting the
89 involvement of non-genetic factors that have yet to be identified. However, the presence of potential
90 limitations in sample collection, such as the lack of representative samples and under-representation of
91 specific regions may contribute to the fragmented understanding of NCD landscape in Cyprus [3].

92 Although most Cypriots declare a healthy status, two out of five suffer from a chronic disease, with CVD,
93 cancer, and diabetes being the most prevalent conditions [4]. While CVD rates have been steadily decreasing
94 since 2015, cancer rates remain at the same levels, albeit lower compared to the rest of the EU. On the
95 contrary, tobacco use rates in Cyprus are amongst the highest in the EU, and together with unhealthy dietary
96 habits, they comprise two of the most frequent mortality risk factors in Cyprus (19% and 14%, respectively)
97 [4]. Additionally, while obesity rates in Cyprus are comparable to the average rates of other EU countries,
98 childhood obesity is rapidly increasing and poses a significant public health issue for Cyprus.

99 To gain insights into the etiology and risk factors associated with NCDs and address limitations including the
100 sampling of data from a diverse and representative population, the establishment of a Cypriot Biobank was
101 imperative. While some causes of these diseases are already known based on studies conducted in other
102 populations and regions [5, 6] current knowledge is insufficient to explain the substantial geographic
103 differences in disease rates observed globally and within Cyprus itself [7]. The development of a Cypriot
104 Biobank will enable large-scale studies of both genetic and environmental contributions, as well as their
105 interactions, in the development of chronic diseases. Furthermore, it will address the limitations of previous
106 studies, including inadequate sample sizes, lack of blood and/or urine samples, inconsistencies in data
107 collection and processing, limited data collection on risk exposures (e.g. lifestyle, diet, disease comorbidities)
108 and outcomes, and restricted study populations.

109 The biobank.cy Center of Excellence (CoE) in Biobanking and Biomedical Research was established in 2019 at
110 the University of Cyprus (UCY), after funding through the Horizon 2020 Widening Participation/Teaming
111 program (Grant Agreement no. 857122). The program is a partnership of the UCY with two Advanced
112 Partners, the European Biobanking consortium BBMRI-ERIC (https://www.bbmri-eric.eu/) and the Medical

University of Graz (https://www.medunigraz.at/en/), and a local Partner, the small to medium-sized enterprise TALOS Rtd LTD (http://www.talos-rtd.com/). 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.

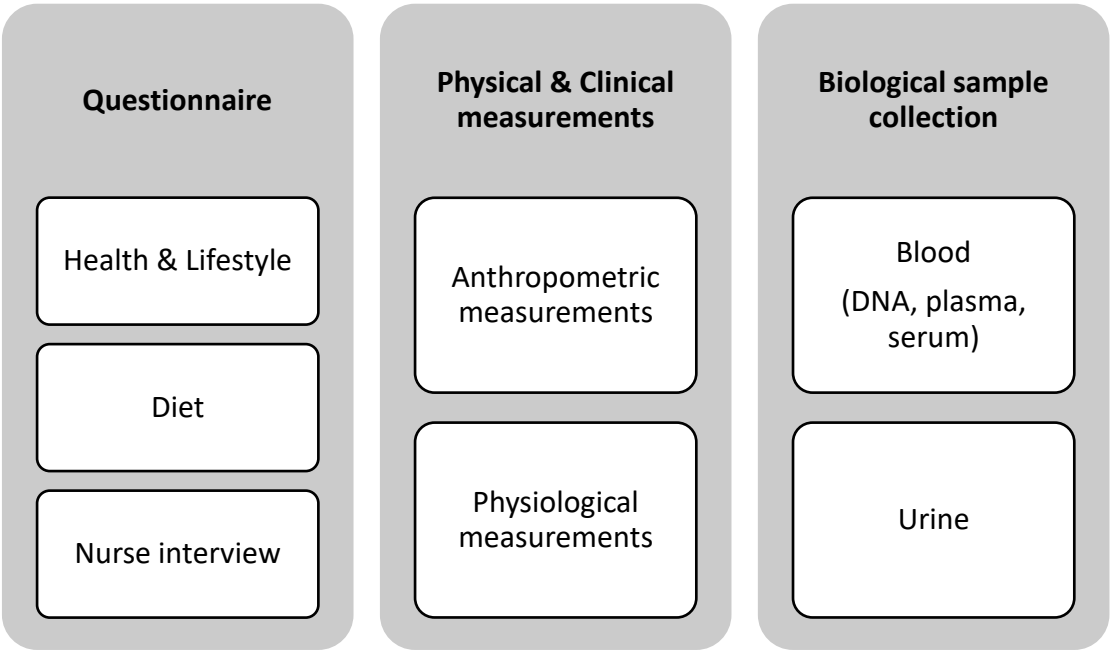
We herein provide comprehensive information on the operation of the first Biobank in Cyprus, sharing data on the demographics and health archives of participants. Some associations are also attempted with the understanding and anticipation that deeper and more useful results will be derived as the Biobank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

138 Biobank obtained first approval by the Cyprus National Bioethics Committee in November 2011 (approval
139 code: ΕΕΒΚ/ΕΠ/2011/04).

140 It gathers information through a structured and comprehensive questionnaire, clinical phenotyping, and
141 biological samples collected from Cypriot residents of the general population, aged ≥ 18 years old. Population
142 data of the pilot study are collected and stored using REDCap® [8, 9] and OpenSpecimen®
143 (https://www.openspecimen.org/) tools respectively, to enable data security and tracking as well as to
144 eliminate collection bias. Specifically, REDCap® (Research Electronic Data Capture) is a secure, web-based
145 application designed for data collection and management in research studies, used to amass the following
146 measurements: personal and family history as well as clinical (complete blood count, biochemical) and
147 biometric (blood pressure, height, weight). OpenSpecimen® is used as the Biobank's Information
148 Management System, by organizing information related to the collection, processing, storage, and
149 distribution of samples. Figure 1 presents a summary of the full primary data collected at the Biobank.



150

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

152 **Participant/volunteer enrolment**

153 Since the very first enrollment on the 1st of November 2019, participant awareness and engagement in the
154 project were achieved through media and the press (television, radio, social media and newspapers, leaflets

155 respectively) as well as social events which are organized by the Biobank personnel. Additionally, friends and
156 family referrals were frequent throughout the 3-year period. Interested individuals book an appointment
157 either through the biobank.cy website (https://biobank.cy/) or directly through a phone call. At baseline,
158 participants are invited to the central collection point of the Biobank facility at the Shacolas Educational
159 Center for Clinical Medicine, or elsewhere during outreach activities, where they undergo a 3-stage interview
160 (25 min on average) and physical and clinical measurements (15 min on average) (Figure 2; Table 2). The
161 Biobank participation process commences with trained research nurses providing an initial briefing to the
162 participants. During this session, participants are fully informed about the research project's purpose, their
163 rights, the procedures for participation and withdrawal, while they may ask questions about the project, prior
164 to signing a broad consent form. The consent form is approved by the Cyprus National Bioethics Committee
165 (approval code: ΕΕΒΚ/ΕΠ/2020/04), ensuring compliance with ethical requirements. A copy of the signed
166 consent form is given to the participant.

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

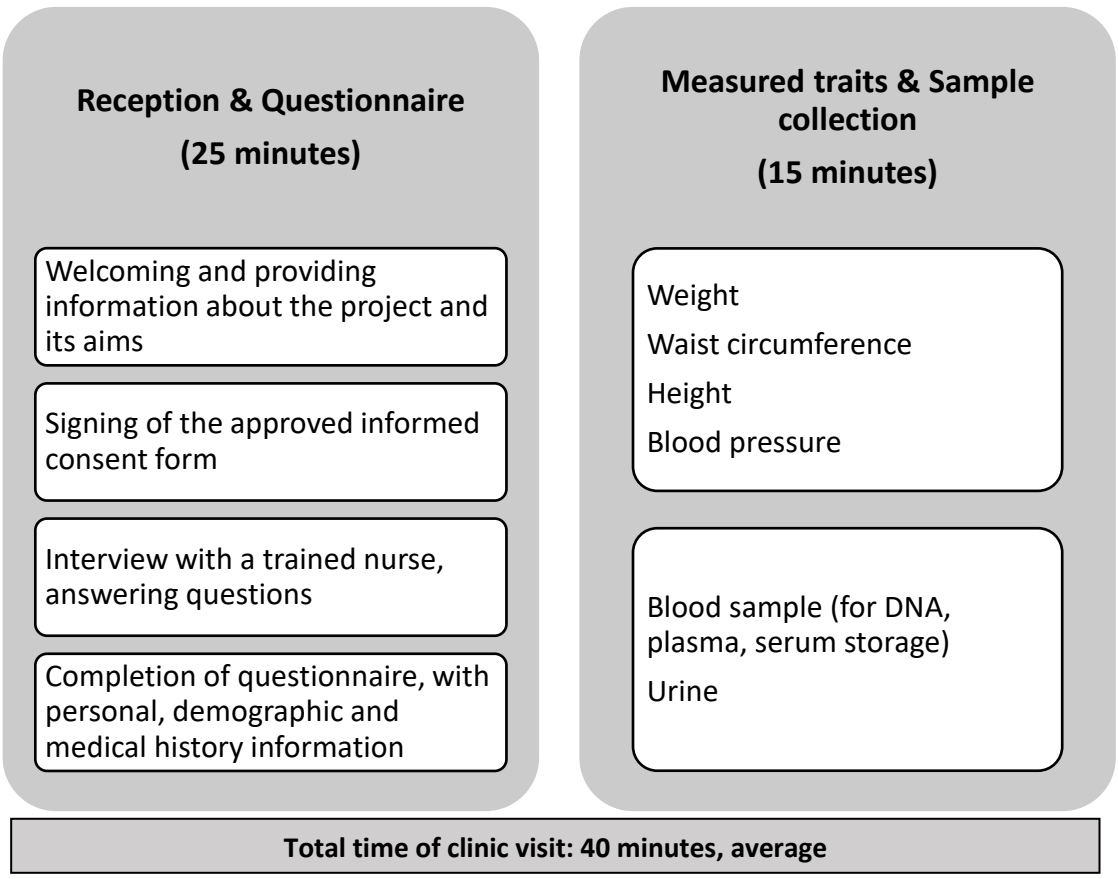


Figure 2: Baseline visit summary

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 [10] and sleeping patterns [11], 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) [12] 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 [13], 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,

193 obesity, and Alzheimer's disease among others [14]. Participants are specifically asked to report the
194 consumption frequency of whole grain foo/d (bread, pasta, rice etc), potatoes, fruit and juice, vegetables and
195 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

196 **Waist circumference measures available for 960 individuals*

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

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

203 At baseline enrolment visit, the participants provided peripheral blood for the isolation of DNA, plasma, and
204 serum, which are stored at -80° C. Urine was also collected and aliquots were stored likewise. Specifically, we
205 collected: 1x10ml EDTA for DNA extraction, 1x4ml EDTA for FBC, 1x4ml EDTA for plasma, 1x8.5ml serum,

206 1x3.5ml serum (prompt test), and 1x100ml urine. Aliquoted and stored are 4 aliquots of 50µl DNA at 50ng/µl,

207 plus a concentrated DNA stock, 4 aliquots of 2ml serum, 2 aliquots of 1ml plasma, and 5 aliquots of 2ml urine.

208 Fifty-eight clinical biomarkers are measured by a private diagnostic provider using a proportion of the stored

209 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).

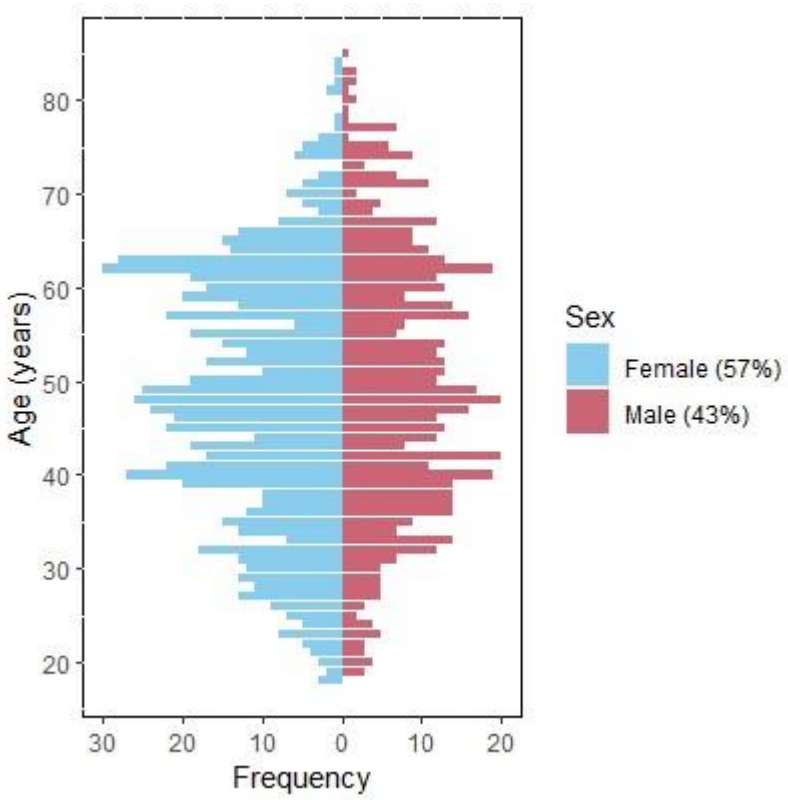


Figure 3: Age distribution of men and women enrolled in the Biobank pilot study

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)	

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

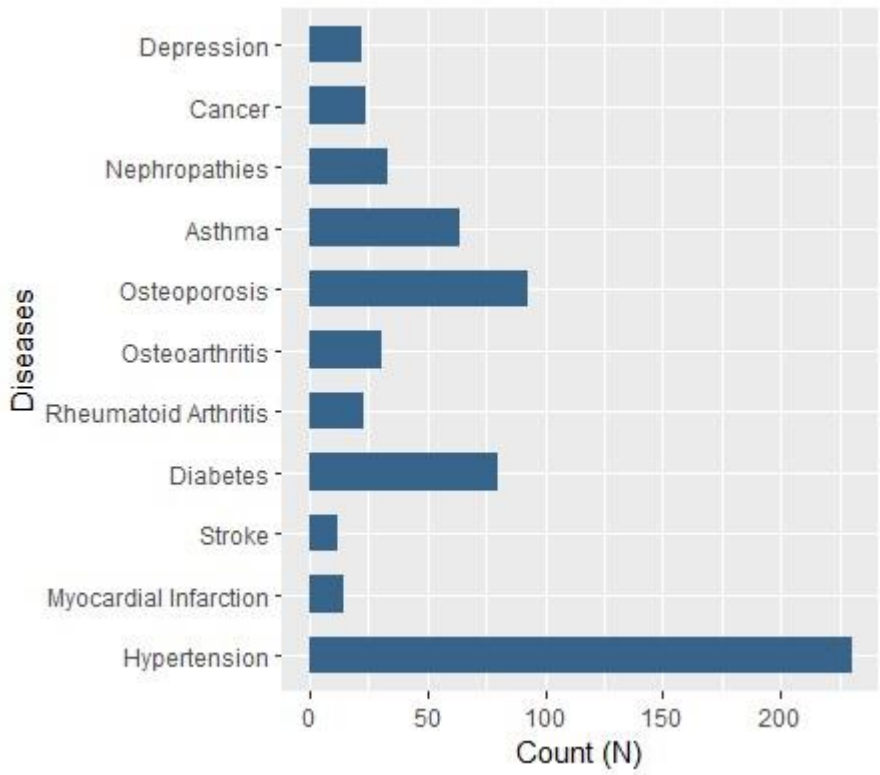
231 Medical history

232 Based on the NHS and Heart UK guidelines (https://www.nhs.uk/conditions/high-cholesterol/cholesterol-
233 levels/) (Table 4), almost half of the participants had high total (N = 681) and LDL (N = 630) cholesterol levels
234 (151/681 and 150/630 post-fasting respectively), while 15.8% (N = 213, 31/213 post-fasting) had low HDL
235 levels. Additionally, 19% (N = 257, 34/257 post-fasting) of individuals had high levels of triglycerides. The
236 majority of these participants did not report any medication use, despite the abnormal features in their blood
237 lipid profile (Table 5). The results of our pilot study are in line with recent findings regarding the high
238 prevalence of elevated blood lipid levels in the European population [15-17].

239 Although the vast majority of the participants self-reported as healthy, a variable proportion of them had
240 underlying past or current conditions, with the most prevalent being hypertension (17.2%), osteoporosis
241 (6.9%) and diabetes (6.0%) (Fig. 4). Additional conditions recorded include cardiovascular disease (myocardial
242 infarction and stroke), rheumatoid arthritis, osteoarthritis, asthma, kidney disease, cancer and depression
243 (Fig.4). Most individuals with a previously diagnosed condition used medication (Table 5).

244

245
246
247
248
249
250
251
252
253



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

Table 5: Medical history and medication use (N = 1348)				
Disease		**N (%)**	**Medication use (%)**	**None reported (N)**
Lipid profile*	TC > 5 mmol/L	681 (51)	102 (15)	579
	LDL-C > 3 mmol/L	630 (47)	80 (12.7)	550
	HDL-C < 1 mmol/L (men) < 1.2 mmol/L (women)	 117 (8.7) 96 (7.1)	 24 (20.5) 16 (16.7)	 93 80
	TG > 1.7 mmol/L	275 (20.4)	85 (30.9)	190
Hypertension‡		230 (17.2)	219 (16.3)	7
Myocardial Infarction‡		15 (1.1)	13 (1.0)	8
Stroke‡		12 (0.9)	10 (0.7)	7
Diabetes‡		80 (6.0)	78 (5.8)	9
Rheumatoid Arthritis‡		23 (1.7)	21 (1.6)	9
Osteoarthritis‡		31 (2.3)	29 (2.2)	8
Osteoporosis‡		93 (6.9)	86 (6.4)	9
Asthma‡		64 (4.8)	45 (3.4)	9
Kidney conditions‡		33 (2.5)	27 (2.0)	6
Cancer‡		24 (1.8)	7 (0.5)	8
Depression‡		22 (1.6)	17 (1.3)	7

*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

269 architecture of the Cyprus human genome. biobank.cy is equipped with clinic, laboratory and liquid sample
270 storage facilities and offices for non-clinical research staff.

271 biobank.cy was founded and launched as a national research infrastructure in 2019, just before the COVID-
272 19 pandemic started. Despite the challenges and the relatively short period since its establishment, it already
273 has a high impact on the health of the Cypriot population. During the pandemic, in addition to the original
274 scope, biobank.cy collected samples from COVID-19 cases to study the production of circulating
275 immunoglobulin class G (IgG) antibodies against SARS-CoV-2 in the serum of convalescent individuals [18].
276 Individuals who participated in the project received the results from their sample collection two weeks later,
277 thus gaining the opportunity to discuss potential abnormal ranges with their personal doctors. The whole
278 exome sequencing (WES) of 1500 convalescent participants in the project is pending.

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

284 **Future plans**

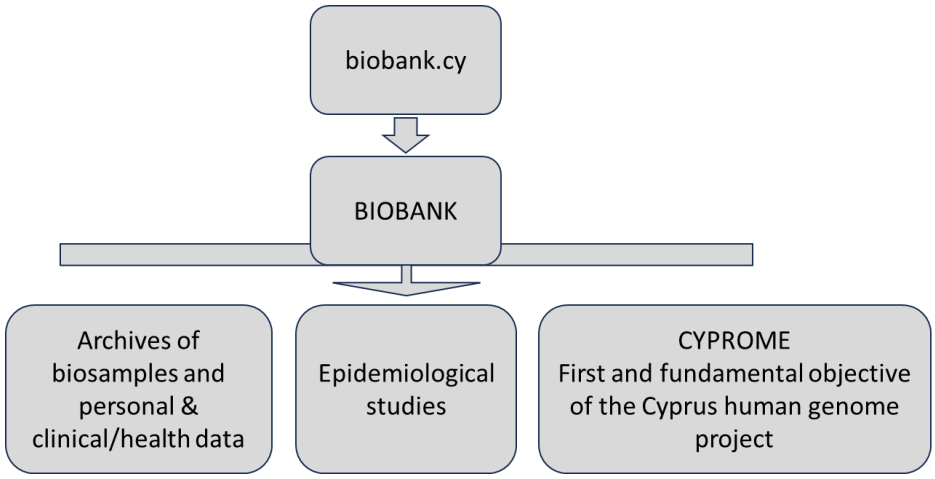
285 **Cyprus Human Genome Project**

286 biobank.cy was formed with the vision of being the island's central node for archiving high-quality biomaterial
287 and associated clinical records, with state-of-art research infrastructure facilities utilized for precision
288 medicine. Towards this goal, biobank.cy has launched the largest research study to ever run in Cyprus, with
289 the precise aim of understanding and treating diseases by characterizing the Cypriot genome. Specifically,
290 biobank.cy has initiated the Cyprus Human Genome Project (CHGP). The CHGP is a first in its kind of scale
291 and prospect study in Cyprus, with the aim of implementing whole genome sequencing (WGS) of at least 500
292 family trios, which will be the blueprint for the first Cypriot reference genome. Concurrently, additional
293 volunteers of the general population, including chronic disease patients, are selected for WES based on

294 additional ongoing separately funded research projects. Currently, more than 1000 Cypriot WES have been
295 completed, while additional samples are being processed to reach a projected number of 5,000 exomes in
296 the next 3 years.

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

307
308
309
310
311
312



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

314 Over the past 30+ years, genetic studies in the Cypriot population unraveled many interesting findings
315 concerning mainly rare diseases, including kidney conditions, neurological conditions, endocrinological and
316 others [2]. Founder mutations were discovered for several diseases while a Cypriot consensus whole genome
317 sequence (including the non-coding part) is still lacking, but individual volunteers and family trios are being
318 enrolled to this effect. The CYPROME project is promising to identify more DNA variants of pathogenic

319 potential and enable comparative studies vis-à-vis the minor allele frequencies among the neighboring
320 populations, thus leading to a better understanding of our genetic architecture. Preliminary data suggest that
321 a substantial number of DNA variants have impressively lower or higher frequencies, compared to other
322 European or Asiatic databases, for which detailed analysis is pending.

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

326 It is a solid objective that our Biobank will be the first to be connected to the Cyprus electronic health record
327 system, hence broadening, and strengthening its anticipated potential for contributing to the next generation
328 of research, encompassing powerful tools and a wealth of information for the first time.

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

333 **Acknowledgements**

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

339 **Author contributions**

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

Funding

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

Availability of data and materials

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

DECLARATIONS

Ethics approval and consent to participate

Ethical approval was granted from the Cyprus National Bioethics Committee, code EEBK/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 (Greek).
3. 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.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.

9. Harris, P.A., et al., *The REDCap consortium: Building an international community of software platform partners.* J Biomed Inform, 2019. **95**: p. 103208.
10. Craig, C.L., et al., *International physical activity questionnaire: 12-country reliability and validity.* Med Sci Sports Exerc, 2003. **35**(8): p. 1381-95.
11. Roth, T., et al., *A new questionnaire to detect sleep disorders.* Sleep Med, 2002. **3**(2): p. 99-108.
12. Panagiotakos, D.B., et al., *MedDietScore: a computer program that evaluates the adherence to the Mediterranean dietary pattern and its relation to cardiovascular disease risk.* Comput Methods Programs Biomed, 2006. **83**(1): p. 73-7.
13. Trichopoulou, A., et al., *Adherence to a Mediterranean diet and survival in a Greek population.* N Engl J Med, 2003. **348**(26): p. 2599-608.
14. Guasch-Ferre, M. and W.C. Willett, *The Mediterranean diet and health: a comprehensive overview.* J Intern Med, 2021. **290**(3): p. 549-566.
15. Timmis, A., et al., *European Society of Cardiology: Cardiovascular Disease Statistics 2017.* Eur Heart J, 2018. **39**(7): p. 508-579.
16. Laufs, U., et al., *Why is hypercholesterolaemia so prevalent? A view from evolutionary medicine.* Eur Heart J, 2019. **40**(33): p. 2825-2830.
17. Steinhagen-Thiessen, E., et al., *Dyslipidemia in primary care--prevalence, recognition, treatment and control: data from the German Metabolic and Cardiovascular Risk Project (GEMCAS).* Cardiovasc Diabetol, 2008. **7**: p. 31.
18. Mamais, I., et al., *Circulating IgG Levels in SARS-CoV-2 Convalescent Individuals in Cyprus.* J Clin Med, 2021. **10**(24).