Paracelsus 10,000: An Observational Cohort Study About the Health Status of the Population of Salzburg, Austria. Rationale, Objectives and Study Design

Vanessa Freya, Patrick Langthalerab, c, Emanuel Raphaelisb, Susanne Ring-Dimitriud, Ludmilla Kedeno, Elmar Aignerb, Jennifer Martinezb, Isabella Gostnerb, Arne C. Bathkeb, cf, Paul Sunglerg, Michael Studnickah, Elisabeth Haschke-Becher, Eugen Trinkak, Bernhard Iglsederl, Bernhard Paulwebera

aDepartment of Neurology, Christian Doppler University Hospital, Paracelsus Medical University and Centre for Cognitive Neuroscience, Affiliated member of the European Reference Network Epicare, Salzburg, Austria, bDepartment of Artificial Intelligence and Human Interfaces, Paris Lodron University of Salzburg, Salzburg, Austria, cTeam Biostatistics and Big Medical Data, IDA Lab Salzburg, Paracelsus Medical University Salzburg, Salzburg, Austria, dDepartment of Sport and Exercise Science, Paris Lodron University of Salzburg, Salzburg, Austria, eDepartment of Internal Medicine, St. Johanns University Hospital, Private Medical University of Salzburg, Salzburg, Austria, fDepartment of Statistics, University of Kentucky, Lexington, KY, USA, gDepartment of Surgery I, St. Johanns University Hospital, Paracelsus Medical University, Salzburg, Austria, hDepartment of Pulmonary Medicine, St. Johanns University Hospital, Paracelsus Medical University, Salzburg, Austria, iDepartment of Laboratory Medicine, University Hospital, Paracelsus Private Medical University, Salzburg, Austria, jDepartment of Public Health, Health Services Research and Health Technology Assessment, UMIT – University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria, kNeuroscience Institute, Christian Doppler University Hospital, Paracelsus Medical University and Centre for Cognitive Neuroscience, Salzburg, Austria, lDepartment of Geriatric Medicine, Christian Doppler University Hospital, Paracelsus Medical University, Salzburg, Austria

Key Words
Paracelsus 10,000 • Study design • Epidemiological study • Salzburg • Austria

Abstract:
Background/Aims: Paracelsus 10,000 is an observational cohort study with the objective to investigate the health status of the population in and around the city of Salzburg. The focus lies on common non-communicable diseases, mainly cardiovascular, cerebrovascular, and metabolic diseases, and their risk factors. The aim of the study is to establish a large
database enabling to study associations between life style, genetic factors and development of common diseases. Extensive phenotyping and biobanking constitute a basis for preventive measures, multi-omics and precision medicine. **Methods:** Inhabitants of the city of Salzburg and the surrounding areas were randomly selected and underwent a detailed investigation: The program included medical examinations, as well as assessments on mental disorders and life style including nutrition and physical activity. From all study participants biological samples were stored in a biobank at -80°C enabling future investigations of biomarkers and utilization of the whole spectrum of multi-omics. **Results:** Between the years 2013 and 2020, 10,044 participants, 5,176 women and 4,868 men aged between 40 and 77 years, completed a baseline assessment. Special emphasis was put on participants between 50 and 59 years old, representing an age group especially prone to changes in health status. In this group, 2,606 participants underwent an extended investigation program. **Conclusion:** The Paracelsus 10,000 study has the potential to gain new insight into the role of associations between genetic predisposition and life style factors for disease development in the Salzburg population. This report describes rationale, objectives and design of the study and provides insight into the main characteristics of the study cohort.

**Introduction**

Non-communicable conditions, including hypertension, diabetes, obesity and atherosclerosis are causing numerous deaths and loss of life years in good health, worldwide. The Global Burden of Diseases Study signals an increase in death by neurological disorders, including stroke, dementia, and epilepsy [1, 2]. Socioeconomic parameters, including education, income, and wealth, are emerging as fundamental causes of these health outcomes [3]. It is important to understand if and how vascular risk factors, local health care policy, genetic, or environmental factors might be leveraged to intervene.

Epidemiological studies at the local level can be informative about health exposures and outcomes. The Copenhagen City Heart Study, launched in 1975, follows a cohort of 20,000 for cardiovascular (CVD) risk factors and markers of aging [4]. The KORA (Cooperative Health Research in the Region Augsburg) research platform, with a cohort of 18,000 participants in Augsburg, Germany, focuses on CVD, environment and stress exposures [5]. The SHIP study (Study of Health in Pomerania) investigates risk factors in two independent cohorts in northeast Germany [6]. Epidemiological research in the Salzburg population is missing so far, disclosing a gap in this region.

According to a survey performed in 2010, 98% of interviewed Austrians regard hypertension as a serious health problem (i.e. they agree on an association of hypertension and heart attack) [7]. However, only 15% identified themselves as hypertensive [7], whereas studies showed that 38.1% of all subjects in Austria have a blood pressure higher than 140/90 mm Hg [8]. Numbers for the city of Salzburg and the surrounding areas are missing, uncovering a gap in local research on this field. The number of ischemic strokes is forecasted to about double by 2075 [9] and there is a rising incidence of stroke in young adults in European countries, as well as in developing countries. European data on trends in stroke incidences are conflicting, with a raised incidence in Sweden and a decrease in the United Kingdom, Italy, Russia and Estonia [10-14]. The geographical variations could be due to differences in local health care policy, genetic, meteorological and environmental factors or differences in distribution of vascular risk factors [15]. Taking this into account, it is of high interest to investigate the health status of populations within a defined geographical area.

Data of the Paracelsus 10,000 study (P10) could support the county wide health initiative “Salzburger Gesundheitsziele” that focuses amongst others on cardio-vascular, cerebrovascular and metabolic diseases, exercise, and nutrition for people at the age 60+ [16].
In our study, special emphasis was put on assessment of cerebrovascular, metabolic, and cardiovascular diseases, along with socioeconomic background and genetic risk factors, focusing on the balance of environment and genetics. Extensive phenotyping of study participants and the collection of a large number of biological samples in a biobank constitute the basis for the use of cutting-edge technologies, such as multi-omics and precision medicine, aiming at a better understanding of the natural history of diseases of interest.

This report describes the Paracelsus 10,000 study design, methods, and cohort characteristics at baseline. To highlight potential application for Paracelsus 10,000 data, the following paragraphs explain the structure of the basic (BP) and extended (EP) investigation program and check for sample generalizability and variability in exposures and outcomes assessed.

**Materials and Methods**

The Paracelsus 10,000 study is a population based, observational cohort study and so far consists of two parts: 1) First phase, cohort recruitment and baseline assessments; 2) Second phase, follow-up assessments and detection of clinical events.

As the first follow-up phase is still ongoing, this manuscript focuses only on the first phase of the study. In the years 2012 and 2013, a formative evaluation of required resources was completed and the project structure was established. Recruitment of participants started in April 2013 and was finished in March 2020, while examinations were performed throughout the whole study period from Monday to Friday (except during school holidays). For details of daily practice, see chapter 3.1 Investigations. Trained staff with a background in medicine, biology, or sports science conducted data collection. Facilities for the study were provided by the Uniklinikum Salzburg (Landeskrankenhaus and Christian Doppler Klinik), as sponsor of the Paracelsus 10,000 study. Facilities on site include examination rooms, office space, computer and waiting rooms for the participants, storage rooms, and biobank facilities.

Recruitment of the study population

People living in the city of Salzburg and surrounding areas were randomly selected from the local population registry. Within the predefined groups (40-49 years, 50-59 years, 60-69 years, women, men) individuals received a computer-based randomly distributed number. These numbers were sorted in ascending order and data of the first 10,000 people were extracted. This was repeated several times during the first study period. We aimed to recruit a balanced sample of men and women, with oversampling of individuals 50-59 years. We invited the selected people by mail; however, a certain number of invitations did not reach the intended recipients due to changed address or death. Because of data security guidelines and an unmanageable effort, it was not possible to conduct statistics about unanswered invitations. Due to the long duration of the first study phase, some individuals who were invited exceeded their 70th birthday at the time of the baseline examination.

In total, we examined 10,044 participants who were stratified into two sub-cohorts: 1) participants aged 40-77 years who received the basic examination program (basic program, BP) and 2) participants aged 50-59 years who received the basic program and additional examinations (extended program, EP). We chose this age group for the extended program and invited more individuals of this age group as we assume that the health status undergoes substantial changes in this period of life. The subjects receiving the extended program were randomly chosen from the pool of 50-59 years old participants. For response rate and distribution of age, gender, and program see Fig. 1 and Fig. 2b, for the list of basic and extended examination, see Table 1. In the end, 323 participants older than 69 years were excluded from the calculations due to better comparability with data of Statistics Austria.

In the beginning of the study, only participants living in the city of Salzburg were invited, of whom everyone received the EP. With increasing numbers of participants per day (Fig. 2a), the majority received only the basic program due to staff and room limitations (Fig. 2d). In the last third of the study period, participants from the areas surrounding Salzburg were invited. Hence, participants recruited between the second half of 2017 and 2020 came from less urbanized areas (Fig. 2c).
Fig. 1. The Figure shows the response rate of invited people and the distribution of the two sub-cohorts including sex and age. *participants approached the study team proactively and were included if they lived in Salzburg or the surrounding areas and were aged between 40 and 69.

Invited
n = 56,595
40-49 years: 18.359
50-59 years: 23.336
60-69 years: 14.900

Response Rate:
40-49 years: 14.3%
50-59 years: 18.3%
60-69 years: 21.2%

*Non-invited & participating: n = 286

Invited & participating: n = 9,758

Participants:

Invited & participating: n = 9,758

Basic Program
n = 7,438
w = 3902 m = 3536
age Ø 55.9

Extended Program
n = 2,606
w = 1274 m = 1332
age Ø 55.3

Fig. 2. Distribution of number (a) and age (b) of participants, as well as the degree of urbanization (c) and examination program (d) over time.
Participation was voluntary and without financial reward, while participants benefited from a preventive medical check-up. All participants signed an informed consent, and the study protocol was approved by the ethics committee of the State of Salzburg (415-E/1521/3-2012).

Even though single investigations had to be omitted in case of physical inability, there were no predefined exclusion criteria for the invited participants.

**Follow Up**

The first follow up phase of the study was planned in 2019, started in autumn 2020 and is scheduled to be finished in 2026. All participants are invited again, approximately 5-7 years after their first visit. The main investigations contain the basic program, a fundus photography, and three additional measurements: circumference of waist, thigh and upper arm. Participants who received the extended program in the first recruitment phase, additionally undergo magnetoencephalography (MEG), brain magnet resonance imaging (MRI), and cognitive testing. The aim is to capture clinical events like acute coronary syndrome or stroke via interviews, questionnaires, and review of hospital records. To this end, separate projects and applications will be designed. Further follow up studies will be conducted in 6 years intervals; ideally, the cohort will be followed until the end of the life of each individual participant. The collected pool of data will present the health status of our cohort at different points from their middle to higher ages. Information about lifestyle and medical events during this time will enable us to find relevant associations and novel risk factors for e.g. cardiovascular and cerebrovascular diseases.

The COVID-19 pandemic had great impact on the progress of the study. Within the 2 years since the follow up started, the study paused during a period of 17 months. This was necessary to reduce the risk of exposing patients in the hospital to the virus. In addition, relevant structural adaptations in the outpatient clinic impeded the progression of the study. Hence, the number of participants investigated in the follow up so far is lower than planned.

**Investigations**

Participants were directed to fast 10h before the visit and to avoid exercise, as well as visiting a sauna the day before. Fig. 3 shows the timeline of daily practice with 8-10 participants each day.

All examinations of an individual participant were carried out on a single day. Investigations that lasted longer than the visit (24h Blood pressure, movisens etc.) started on the day of the visit. In addition to a complete blood count and urinalysis, the following examinations were included in the study program:

**Anthropometry and body composition**
- Body height
- Body weight
- Abdominal circumference
- Body composition: Multiple frequency bioelectric impedance analysis was performed while participants adopted a supine position with electrodes placed on the dorsum of the right hand and foot. (Body Impedance Analyzer™, Nutri Plus Data Input, Germany).
- Dexa-Scan X-ray densitometry for measuring bone mineral density and body composition (Hologic Discovery A™, Hologic APEX, USA).

**Laboratory parameters**

Blood was drawn in sitting position between 7.30 and 10.30 a.m. after at least 10 hours of fasting. The whole blood was collected in special Vacutainer tubes (Greiner Bio-One™) for preparation of different types of biological materials. The analysis was conducted in a laboratory that is certified according to ISO 9001:2015 [17] and voluntarily follows the requirements of ISO 15189:2012 [18].

- Blood count:
  - Red blood cells (RBC), white blood cells (WBC), hemoglobin, hematocrit, mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), platelets, reticulocytes measured with Sysmex XN-1000™ (Sysmex Austria GmbH).
**Coagulation:**
Thrombin time (TT), partial thromboplastin time (PTT), prothrombin time (PT)/international normalized ratio (INR), fibrinogen, antithrombin III measured with BCS XP™ (Siemens Healthcare Diagnostics GmbH).

**Blood Sediment:**
Red blood cells (RBC), white blood cells (WBC) measured with Sysmex XN-1000™ (Sysmex Austria GmbH)

**Lipid metabolism:**
Total cholesterol, Low-density lipoprotein (LDL)-cholesterol, High-density lipoprotein (HDL)-cholesterol, triglycerides, apolipoprotein-B, apolipoprotein-AI, lipoprotein (a) measured with Cobas 6000™ (Roche Diagnostics GmbH).

**Glucose metabolism:**
Fasting blood glucose measured with Cobas 6000 (Roche Diagnostics GmbH), HbA1c measured with Cobas Integra 400 Plus Analyzer™ (Roche Diagnostics GmbH) and fasting insulin.

**Renal function:**
Serum creatinine, glomerular filtration rate (eGFR), serum urea, serum electrolytes (Na, K, Cl, Ca, Mg, P), serum uric acid, urine chemistry (albumin, creatinine, albumin/creatinine ratio, total protein, glucose), urine sediment (specific weight, pH value, red blood cells (RBC), white blood cells (WBC), nitrite, total protein, glucose, ketone bodies, urobilinogen, bilirubin), electrolytes (Na, K, Cl, Ca, Mg, P) in urine measured with Cobas 6000™ (Roche Diagnostics GmbH) while urine sediment was measured with Cobas U 601™ (Roche Diagnostics GmbH).

**Liver function:**
Aspartate transaminase (AST), alanine aminotransferase (ALT), gamma-glutamyltransferase (gGT), alkaline phosphatase (AP), cholinesterase (CHE) measured with Cobas 6000™ (Roche Diagnostics GmbH).

**Pancreas function:**
P-amylase, lipase measured with Cobas 6000™ (Roche Diagnostics GmbH).

**Iron parameters:**
Iron, transferrin, transferrin saturation, ferritin measured with Cobas 6000™ (Roche Diagnostics GmbH).

**high sensitivity c-reactive protein (hsCRP), thyroid-stimulating hormone (TSH), prostate specific antigen (PSA, for men), total serum protein measured with Cobas 6000™ (Roche Diagnostics GmbH).**

---

**Fig. 3.** Timeline of daily practice. DEXA= Dual Energy X-ray Absorptiometry, CT= Computed Tomography, US= Ultrasound, IPAQ= International Physical Activity Questionnaire. *All investigations except for DEXA Scan, coronary CT, exercise testing, liver US, 7 days nutrition protocol, IPAQ, accelerometer, 24h Blood pressure. **Time might be variable, up to the next day. This depended on participants’ personal reasons.
Cardiovascular parameters

- Blood pressure was measured bilaterally in a sitting position. The measurement was repeated three times on each side with a previous resting period of each 60s (Watch BP office AFIB™ from the company Microlife AG Swiss Corporation, Switzerland).
- Electrocardiography (ECG) 12-lead, in a supine position (Schiller Cardiovit AT-102 Plus 12-Kanal-EKG™, Germany).
- Ankle-brachial pressure index was measured three times in a supine position (Boso ABI-system 100 PWV™, Bosch + Sohn, Germany).
- 24 h blood pressure: A portable device was taken home by participants and measured blood pressure, pulse and pulse wave velocity in 15-minute intervals from 8 a.m. to 10 p.m. and in 30-minute intervals during night (Mobile-O-Graph PWA™, Industrielle Entwicklung Medizintechnik und Vertriebsgesellschaft mbH (I.E.M.), Germany).
- Coronary Artery Calcium Score (Agatston-Score) using Computed Tomography (Siemens Somatom Definition AS+™, Erlangen, Germany) [19].
- Ultrasound examination of the carotid artery bilaterally in a supine position (Panasonic Healthcare Diagnostics US). Intima media thickness and total plaque area were measured.

Spirometry, measurement of liver stiffness, oral glucose tolerance test

- Pulmonary function test was carried out, determining forced vital capacity and forced expiratory volume over 1s using a portable spirometer (Easy OneTM Spirometer™, NDD Medical Technologies, Zurich, Switzerland). Exclusion criteria were recent operations, cardiovascular and acute pulmonary diseases [20].
- Liver stiffness measurement and semi-quantitative estimation of fat content via continuation attenuated parameter using Fibroscan 530 Compact™ (Echosens Paris, France).
- An oral glucose tolerance test (OGTT) was performed following the standard protocol of the American Diabetes Association. Before and at 30min, 60min, 90min and 120min after oral intake of 75g glucose venous blood was drawn for measurement of plasma glucose and insulin levels.

Medical history

Trained staff conducted a face-to-face interview recorded as audio file that included medical history, family history, medication and the Salzburg Dementia Prediction Test [21]. Further inventories were merged to one computer-based questionnaire that was answered discretely by participants either in the study rooms (possibly with assistance by the staff) or at home. Participants who were not familiar with computer work or who needed translation answered the questionnaire manually and the study staff digitized the data.

The questionnaire included:

- ODQ-D (Oswestry Low Back Pain Disability) [22]
- WOMAC (Western Ontario and McMaster Osteoarthritis Index) about coxarthrosis [23]
- Screening questionnaire of the Austrian Social Insurance [24].
- BDI II (Beck Depression Inventory) [25].
- Barrat Impulsiveness Scale (BIS-15) [26; 27].

The face-to-face interview also included questions on major depressive disorder, bulimia and anorexia nervosa, generalized anxiety disorder, manic and hypomanic episodes, panic disorder, agoraphobia, social phobia, obsessive-compulsive disorder, posttraumatic stress disorder, substance and alcohol abuse.

Assessment of lifestyle factors

Further inventories on lifestyle that were answered discretely by participants were included:

- AUDIT (Alcohol Use Disorders Identification Test) [28]
- Inventory on diet according to the EPIC study (European Prospective Investigation into Cancer and Nutrition) [29]. The German version of the EPIC questionnaire was kindly provided by the DIFE (Deutsches Institut für Ernährungsforschung, Potsdam, Germany).
- Quality of life Assessment (QoL) in F 12 short form including twelve questions about quality of life influenced by health [30, 31].
- A seven-day diet survey including a program that calculates nutritional value (DGE–PC professional, version 5.1) by the German Nutrition Society [32].
- ESS (Epworth-Sleepiness-Scale) [33]
Additionally, questions about origin and family background, religion, employment, education, domestic circumstances, smoking habits, menopause, Mediterranean diet, allergies and sexuality were included. Further, questions were added on usage of mobile phone, sensitivity towards chemical smells, electromagnetic pollution, air and noise pollution.

**Health related physical activity**
- The objective measurement of physical activity was calculated from a triaxial accelerometer (3D-Activity Sensor Move II™, movisens GmbH, Karlsruhe, Germany), which collected raw data at a sampling rate of 64 Hz. The accelerometer was placed on the right hip during seven consecutive days (24h) in a typical working week, but was removed for activities involving water (i.e. showering, swimming).
- International Physical Activity Questionnaire (IPAQ) was completed after accelerometer recording to display the last seven days [34, 35]. We extended the inventory slightly by adding questions on electrically assisted bicycle and the times in standing or sitting position, as well as sleeping times.
- Questionnaire about regular exercise in different phases of life were recorded in the self-assessed inventory.

**Physical Fitness Testing**
- Participants completed an incremental cardiopulmonary exercise test (CPET) on a stationary bicycle ergometer (ergo select 200™, ergo line GmbH, Blitz, Germany) to measure cardio-respiratory fitness as peak oxygen uptake and aerobic capacity as ventilatory threshold with a gas-exchange analyzer (facemask, Hans Rudolph, Kansas, USA; volume sensor Triple-V™; semipermeable sampling line Twin Tube; Master Screen CPX-Jaeger™, VIASYS Healthcare – Respiratory Technologies, LabManager V.5.32.0.; all Jaeger™, Höchberg, Germany). Starting with 0.9 Watt per kg body mass the load (Watt) was increased every minute with 10 Watt until volitional exhaustion was reached [36, 37]. At the end of each step heart rate (12-lead ECG) and blood pressure (Riva-Rocci method) were measured to control for exhaustion. Subjects with a manifest cardiovascular disease, anemia and/or musculoskeletal disorders were excluded from testing.
- Six-meter gait speed test (6 GST): maximum and habitual walking speed was recorded via a wireless photoelectric beamer over a distance of 10 m, where the walking speed in m/s over 6 m was recorded (TC™, Brower Timing System, Draper, USA).
- Isometric grip strength of the right and left hand was measured in a sitting position via a hydraulic hand dynamometer by recording the best of three repetitions for right and left hand grip strength in kg (5030J1™, Jamar, Patterson Medical, Dallas, USA).

We used the internet based software EvaSys™ to collect parts of the data. By means of individual codes, the study staff or the participants themselves could fill in data into pre-programmed masks. Even though the program is internet based, it was incorporated into the intern hospital IT system, protected by the associated firewall. Devices were not changed during the first study phase.

**Biobank sample collection**
Blood was drawn as explained in 2.2. Laboratory Parameters. Blood (EDTA, citrate and heparin), fasting serum and spontaneous urine were centrifuged at 2.000 g for 15 minutes in a refrigerated centrifuge (CF5804R™, Eppendorf). The supernatant was aliquoted into 2D Data-Matrix coded Screw Cap tubes, that were placed in barcoded Lobarack-96™ from Micronic™ for longtime storage at -80°C. The tubes were closed with pierceable TPE Capcluster using electric Capper (Micronic™). The barcodes on the tubes and racks were scanned using Traxer Code Reader (Micronic™). Information about the type of biological material and location in the freezers was managed using Track-it sample management (Micronic™) and LabCollector (AgileBio™) softwares.

We collected 24 aliquots of fasting serum, 24 aliquots of fasting plasma EDTA, 8 aliquots of fasting citrate plasma, four aliquots of fasting heparin plasma, 8 aliquots of urine, 12 ml whole EDTA blood, two aliquots ofuffy coat and approximately 50g of stool from each participant. All biological materials are stored at -80°C in Ultra-Low Temperature Freezers from Thermo Fischer Scientific™.

Up to now, no genetic analyses have been conducted.
Quality control

During the course of the study, several measures of quality control were implemented. In the first phase of the study, all data manually entered in EvaSys was checked for accuracy. With rising numbers of participants and data sets, only randomly selected data of single participants were checked in detail. The first major quality control was conducted analytically after recruitment of 6,000 subjects. Here, data was checked for plausibility via determination of threshold values. If data exceeded defined thresholds, the raw data were consulted again. Additionally, duplicate observations were deleted and missing values were completed if possible. Of all the values (blood pressure and anthropometry parameters) of the 307 participants that were randomly checked for mistakes, 0.54 % was found to be faulty. Details about cut-off values can be found in the supplementary material (for all supplementary material see www.ppexmed.com). In the next and for now last phase of quality control, the statistical software R™ [38] was used to check for internally inconsistent answers (e.g. medication without known diseases or differences in self-report and interview). In case of doubt, single values were excluded from analysis.

Cohort Characteristics

Study population demographics and risk factors. To get an overview of characteristics of our study cohort, we conducted side by side comparisons of the P10 cohort with national official statistics for the population of Austria and the State of Salzburg, ages 40 to 69 years [39]. Supplementary Fig. S1-S10 compare the cohorts with respect to sex, age, migration background, educational attainment, type of employment, degree of urbanization, smoking habits, and physical activity. For this, the statistical software R™ (Version 4.0.2) [38] was used, including the package tidyverse [40].

Demographic characteristics of the cohort. The gender distribution in the study cohort shows a balance between women (51.6%) and men (48.4%), similar to the population of the State of Salzburg and of Austria as a whole. The study cohort differs in age distribution from these two reference populations: Due to the study design, participants between 50 and 59 are overrepresented (Supplementary Fig. S1). The Salzburg and Austrian population are much more likely to live in rurally urbanized regions, while the largest part of our cohort lives in densely and intermediately urbanized regions (Supplementary Fig. S2). Compared to the Salzburg and Austrian population, our cohort is composed of slightly more individuals with Austrian citizenship and no migration background (Supplementary Fig. S3). The study cohort includes more individuals with a higher educational level (Supplementary Fig. S5) and a higher number of “white-collar” workers (Supplementary Fig. S7) in the study cohort.

Due to this divergence, the variables education, employment, and migration background were additionally analyzed with regard to three urbanization degrees (Supplementary Fig. S4, S6, S8). There was only a difference in the dense and intermediate urbanization degree regarding migration. The study included less participants with first generation migration background than the reference group.

Life style: Exercise and Smoking. There is a smaller proportion of people who smoke and a higher proportion of ex-smokers in the study cohort compared to the Austrian and Salzburg population (Supplementary Fig. S9). Additionally, the study participants do less moderate and intense exercise in minutes per week than the reference populations (Supplementary Fig. S10).

Results

Health parameters captured

Although, the number of valid observations varied by type of assessment (Table 1), a wide variety of data were successfully collected on thousands of study participants. Overall, at their baseline assessment, 31.3% of the full study population aged 40-69 years was insulin resistant based on Homeostasis Model Assessment for Insulin Resistance (HOMA-IR) over 2.5; 58% of the study population was overweight or obese, and 8.6% had a moderate to very high risk for chronic kidney disease. A total of 20.8% of the cohort took antihypertensive medication, resulting in an average blood pressure of 138/84 in this group, while 24.4% had a blood pressure >140/90mmHg without taking antihypertensive medication (mean 148/93 mmHg).
Table 1. List of examinations additional to a complete blood count and urinalysis, the respective program and sample size. * includes: questions on major depressive disorder, bulimia and anorexia nervosa, generalized anxiety disorder, manic and hypomanic episodes, panic disorder, agoraphobia, social phobia, obsessive-compulsive disorder, posttraumatic stress disorder, substance and alcohol abuse, ** regular exercise in different phases of life recorded in the self-assessed inventory. *** includes: questions on origin and family background, religion, employment, education, domestic circumstances, smoking habits, menopause, Mediterranean diet, allergies and sexuality were included. Further, questions were added on usage of mobile phone, sensitivity towards chemical smells, electromagnetic pollution, air and noise pollution. BP: Basic program, EP: Extended program, OGTT: Oral glucose tolerance test, ODQ-D: Oswestry Low Back Pain Disability, WOMAC: Western Ontario and McMaster Osteoarthritis Index, US: Ultrasound, BDI: Beck Depression Inventory, BIS: Barrat Impulsiveness Scale, IPAQ: International Physical Activity Questionnaire, CPET: Cardiopulmonary exercise testing, AUDIT: Alcohol Use Disorders Identification Test, EPIC: European Prospective Investigation into Cancer and Nutrition, QoL: Quality of Life, ESS: Epworth-Sleepiness-Scale.

Data on anthropometry, lipid and glucose metabolism, kidney function, blood pressure and sociodemographic data of the sub-cohort (receiving the extended program) and the sub-cohort of participants aged 50-59 years who received the basic program, split up in women and men were compared (Table 2).

We decided to test the success of the randomization procedure of the two groups further for some important variables. To this end we calculated 95% confidence intervals for the difference in means between the BP and EP groups from a t-test. We then chose cutoff values such that we considered a randomization procedure successful if it produces samples with means within one cutoff value of each other. Cutoff values were 0.5 kg/m^2 for Body Mass Index (BMI), 1 mmHg for Blood Pressure, 2 ml/min/1.73m for epidermal Growth Factor Receptor (eGFR) and 0.1 (mg*µU)/(405*dL*mL) for HOMA-IR. We deemed randomization successful if the confidence interval was completely contained within the interval (-cutoff, cutoff), not successful if the confidence interval was completely outside this interval and use the term „inconclusive“ if the confidence interval crossed the boundaries of the cutoff interval. We thus demonstrated successful randomization for BMI in men and eGFR in men and women. Randomization was not successful regarding systolic blood pressure in women and diastolic blood pressure in both men and women. All other comparisons were inconclusive.
Table 2. Data on anthropometry, lipid and glucose metabolism, kidney function, blood pressure and sociodemographic data of the sub-cohort (receiving the extended program) and the group of participants aged 50-59 years who received the basic program, split up in women and men. BP: Basic program, EP: Extended program, BMI: Body Mass Index, LDL: Low-density Lipid, HDL: High-density Lipid, HbA1c: Hemoglobin A1c, HOMA-IR: Homeostasis Model Assessment for Insulin Resistance, eGFR: epidermal Growth Factor Receptor, SD: Standard Deviation. Values were calculated from data including participants between 50 and 59 years. For further details on age and gender distribution of the two groups, see Fig. 1.

<table>
<thead>
<tr>
<th>Examinations</th>
<th>Exam Type</th>
<th>Women (Mean, SD, N)</th>
<th>Men (Mean, SD, N)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anthropometric Data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>BP</td>
<td>55 (3), 1153</td>
<td>55 (3), 1006</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>55 (3), 1022</td>
<td>55 (3), 1075</td>
</tr>
<tr>
<td>Abdominal Circumference (cm)</td>
<td>BP</td>
<td>88 (13), 1153</td>
<td>99 (12), 1003</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>88 (13), 1017</td>
<td>98 (11), 1067</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>BP</td>
<td>26 (5), 1153</td>
<td>27 (4), 1005</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>26 (5), 1020</td>
<td>27 (4), 1072</td>
</tr>
<tr>
<td><strong>Lipid Metabolism</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Cholesterol (mg/dL)</td>
<td>BP</td>
<td>219 (36), 1147</td>
<td>210 (38), 1001</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>219 (38), 998</td>
<td>210 (39), 1056</td>
</tr>
<tr>
<td>LDL-Cholesterol (mg/dL)</td>
<td>BP</td>
<td>145 (35), 1144</td>
<td>146 (37), 997</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>142 (37), 993</td>
<td>144 (36), 1046</td>
</tr>
<tr>
<td>HDL-Cholesterol (mg/dL)</td>
<td>BP</td>
<td>72 (18), 1144</td>
<td>55 (15), 997</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>72 (18), 993</td>
<td>55 (15), 1046</td>
</tr>
<tr>
<td>Triglycerides (mg/dL)</td>
<td>BP</td>
<td>102 (57), 1147</td>
<td>133 (92), 1001</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>102 (60), 998</td>
<td>135 (85), 1057</td>
</tr>
<tr>
<td>Lipoprotein(a) (mg/dL)</td>
<td>BP</td>
<td>52 (76), 1015</td>
<td>49 (71), 853</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>53 (75), 899</td>
<td>46 (67), 918</td>
</tr>
<tr>
<td>Apolipoprotein B (mg/dL)</td>
<td>BP</td>
<td>108 (24), 1147</td>
<td>115 (27), 1000</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>105 (25), 999</td>
<td>111 (26), 1058</td>
</tr>
<tr>
<td><strong>Glucose Metabolism</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>BP</td>
<td>5.5 (0.4), 1034</td>
<td>5.5 (0.6), 985</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>5.5 (0.4), 962</td>
<td>5.5 (0.5), 1044</td>
</tr>
<tr>
<td>Glucose (mg/dL)</td>
<td>BP</td>
<td>92 (12), 1147</td>
<td>99 (21), 1001</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>91 (14), 999</td>
<td>98 (17), 1060</td>
</tr>
<tr>
<td>Insulin (mIU/L)</td>
<td>BP</td>
<td>9.1 (8.1), 1107</td>
<td>11.2 (9.4), 951</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>8.7 (5.7), 844</td>
<td>11.5 (9.0), 919</td>
</tr>
<tr>
<td>HOMA-IR</td>
<td>BP</td>
<td>2.20 (1.28), 1107</td>
<td>2.87 (2.91), 958</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>2.07 (1.86), 844</td>
<td>2.93 (3.00), 918</td>
</tr>
<tr>
<td><strong>Kidney Function</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creatinine (mg/dL)</td>
<td>BP</td>
<td>0.76 (0.11), 1147</td>
<td>0.95 (0.14), 1002</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>0.76 (0.13), 998</td>
<td>0.95 (0.16), 1066</td>
</tr>
<tr>
<td>eGFR (mL/min/1.73m²)</td>
<td>BP</td>
<td>82 (16), 1147</td>
<td>85 (19), 999</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>82 (15), 996</td>
<td>84 (14), 1058</td>
</tr>
<tr>
<td><strong>Blood Pressure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic (mmHg)</td>
<td>BP</td>
<td>126 (17), 1152</td>
<td>132 (15), 1004</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>123 (15), 1016</td>
<td>130 (14), 1069</td>
</tr>
<tr>
<td>Diastolic (mmHg)</td>
<td>BP</td>
<td>80 (10), 1152</td>
<td>86 (10), 1004</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>77 (10), 1016</td>
<td>83 (9), 1069</td>
</tr>
<tr>
<td><strong>Sociodemographic Data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest Education (ISCED) 0-2/3-5/6-8/other</td>
<td>BP</td>
<td>6.7/67.4/24.0/1.9, 1081</td>
<td>5.0/62.2/32.0/0.8, 946</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>6.2/67.4/22.2/2.2, 990</td>
<td>5.2/58.6/35.1/1.1, 1032</td>
</tr>
<tr>
<td>Urbanisation dense/intermediate/rural</td>
<td>BP</td>
<td>48.9/45.7/5.3, 1142</td>
<td>43.8/50.7/5.5, 998</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>63.3/31.5/5.2, 1009</td>
<td>60.5/34.2/5.3, 1053</td>
</tr>
<tr>
<td>Migration Background yes/no</td>
<td>BP</td>
<td>9.2/90.8, 1051</td>
<td>10.0/90.0, 908</td>
</tr>
<tr>
<td></td>
<td>EP</td>
<td>9.0/91.0, 941</td>
<td>9.6/90.4, 999</td>
</tr>
</tbody>
</table>

The distribution of socioeconomic data varies slightly between the groups, especially regarding urbanization. Participants who received the extended program show a higher degree of urbanization than the other group. For percentiles of distribution in entire cohort, see supplementary (Table S3).
Discussion

During the course of the Paracelsus 10,000 study, 10,044 participants were examined. We collected data on physical and mental health, including fitness parameters, health related behavioral variables and socioeconomic parameters with a focus on risk factors for cardiovascular and cerebrovascular diseases. A biobank with biological samples from study participants was established for future analysis of biomarkers such as genomics, epigenomics, proteomics, metabolomics and metagenomics.

In order to get some insight into our study cohort and to assess differences between the study cohort and the general population in Salzburg and Austria, we compared characteristics of our study participants with data acquired by the national statistics institute Statistik Austria [39]. While there is a rather representative distribution of gender, our study population has a higher proportion of participants aged 50 to 59 years. We intentionally selected participants in this age range, as we believe that the health status undergoes substantial changes in this period of life. There is evidence that lifestyle changes in this age span can prevent early onset of diseases and premature death [41]. In Austria, healthy life years above the age of 50 (m=14.53 years, w=15.66 years) are below average compared with the other 24 EU countries in the year 2005 (e.g. UK m=19.74, w=20.78; Italy m=20.63, w=20.86; Poland m=16.48, w=20.16; Denmark m=23.64, w=24.12) [42]. From the years 2005 to 2010, healthy life years are even diverging further between the EU countries, suggesting that not all countries are ageing in good health with implications for health and social care provision [43]. Hence, observing the health status in the age group 50-59 is of special interest and relevant for disease prevention strategies and early intervention. To achieve the goal of successful aging, interventions should aim to prolong the period in good health.

The comparison between the sub-cohort receiving the extended program and the group of participants aged 50-59 years who received only the basic program shows a similar distribution in metabolic parameters (except for blood pressure, see Table 3). This accounts for a successful randomization of participants in this age group allowing us to draw conclusions about the main cohort in this age group. The distribution of socioeconomic data varies slightly between the groups, especially regarding urbanization. Participants who received the extended program show a higher urbanization than the other group. This is due to the mentioned changes over time in recruiting: In the early phase of the study less participants per day were invited, hence proportionally more participants were randomly selected for the extended program. Inhabitants of the surrounding areas of Salzburg were recruited in the later phase of the study, resulting in a lower number of extensively investigated participants of the intermediately urbanized regions.

Coming back to the general age distribution, weighting is needed in order to generalize the study cohort to the Salzburg or Austrian populations. Our study cohort is less engaged in physical activity compared to the Salzburg or Austrian populations, with respect to intense and moderate exercise (including fast or uphill walking and walking while carrying items). This difference may be explained, at least in part, by differences in the data collection. Compared to the Salzburg and Austrian population, individuals without migration background are slightly overrepresented in our study cohort, going along with a higher level of education and a greater proportion of so called white-collar and a lower proportion of blue-collar workers, while there are no differences in the group of self-employed individuals. It has to be mentioned that the questionnaires

Table 3. Confidence intervals (95%) and p-values for the difference in means between the BP and EP groups: BMI: Body Mass Index, eGFR: epidermal Growth Factor Receptor; HOMA-IR: Homeostasis Model Assessment for Insulin Resistance, from a t-test split up in men and women for the parameters (BMI), systolic and diastolic blood pressure, eGFR and HOMA-IR.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Men (95% CI, p-value)</th>
<th>Women (95% CI, p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI</td>
<td>(-0.33, 0.39) p=0.6737</td>
<td>(-0.80, 0.605) p=0.0816</td>
</tr>
<tr>
<td>Systolic Blood Pressure</td>
<td>(0.81, 3.21) p=0.0012</td>
<td>(1.70, 4.41) p=0.0008</td>
</tr>
<tr>
<td>Diastolic Blood Pressure</td>
<td>(2.67, 3.67) p&lt;0.0001</td>
<td>(3.99, 4.66) p&lt;0.0001</td>
</tr>
<tr>
<td>eGFR</td>
<td>(-0.89, 2.00) p=0.4524</td>
<td>(-1.92, 0.70) p=0.3599</td>
</tr>
<tr>
<td>HOMA-IR</td>
<td>(-0.32, 0.21) p=0.6899</td>
<td>(-0.1, 0.36) p=0.2723</td>
</tr>
</tbody>
</table>
on lifestyle were issued in German (official and national language in Austria) and not all non-native speakers could be assisted in completing the inventories. Hence, subjects with migration background might have been underrepresented in the dataset acquired by the inventories due to the fact that they were not able to complete it reliably because of language barriers. This might have influenced other variables as well.

As most of our participants were recruited from the city of Salzburg, participants living in rural areas are underrepresented in the study cohort as compared to the State of Salzburg and Austria. Yet, according to our data, this does not explain the shifted distribution of employment type and education, as white-collar workers and people with higher education are overrepresented in all three urbanization groups. Hence, the discussed characteristics of our cohort seem to be stable throughout urbanization degrees. The fraction of the cohort living in the countryside, potentially discloses interesting associations between urbanization and mental as well as physical health, lifestyle, and living situation. The baseline data on health parameters show a broad variability from physiological to pathological values, enabling us to observe incidences and risk factors, as well as progression throughout the life span via data collected in the follow up period. Adding genetic analyses will increase our ability to answer important research questions, raising relevance and impact of the Paracelsus 10,000 study. The data pool will allow researchers of different fields to analyze associations between e.g. nutrition and the biome, metabolic incidents like a fatty liver and BMI, mental health and lifestyle factors, or the distribution of risk groups in the cohort.

The recruitment of all participants was completed before the first policies regarding the Covid-19 pandemic were implemented in Austria. Hence, collected data such as lifestyle, mental health etc. were not influenced by consequences of the pandemic.

Besides the known limitations of single investigation methods, there are some further weaknesses to discuss regarding the study design. Even though the invited people were selected randomly, a certain bias in the structure of the study population is inevitable. Approximately one sixth of the invited people actually participated in the study, probably excluding to some extent individuals with a lower degree of health consciousness. Additionally, people suffering from severe diseases might be underrepresented due to logistic reasons and physiological or psychological barriers.

With the provided resources, it was not possible for us to focus on the question to what extent the study cohort is representative of the Austrian population. This would require more detailed reference data and further statistical analysis, which are beyond the scope of this manuscript. Yet, answering the question of representativeness of the cohort might be an interesting future project.

Results of the quality control revealed only a small number of missing values and incorrect data, emphasizing the high quality of the study implementation and validity of the collected data. This might be due to a high level of standardization of investigations, conscientious training of staff, and background knowledge from different scientific fields including sports science, biology, psychology, and statistics.

Conclusion

In sum, the P10 project successfully collected data on a wide range of health conditions, socio-economic, and lifestyle factors for people aged 40-69 years. The large P10 dataset will enable high quality epidemiological analyses focused on determinants of health in mid-life.

The P10 is particularly well suited to address gaps in knowledge about incidence and progression of a wide variety of chronic health conditions like insulin resistance, obesity, hypertension or chronic kidney diseases. It is of high relevance to further develop and adapt preventive and therapeutic strategies in health care including lifestyle modification. Results of the Paracelsus 10,000 study might contribute to achieving these goals by providing detailed information about the health status of 10,044 Salzburg inhabitants in an epidemiological setting and deepened understanding of the association between exposures and diseases.
The follow up period of the study has already started and will give valuable insight into the health development of our cohort during the years to come.

**Abbreviations**

6GST (Six-meter gait speed test); AUDIT (Alcohol Use Disorders Identification Test); BDI (Beck Depression Inventory); BMI (Body Mass Index); CPET (cardiopulmonary exercise testing); CVD (Cardiovascular Disease); DALYs (Disability adjusted Life Years); ECG (Electrocardiography); eGFR (epidermal Growth Factor Receptor); EPIC (European Prospective Investigation into Cancer and Nutrition); ESS (Epworth-Sleepiness-Scale); HDL (High-density Lipid); HOMA-IR (Homeostasis Model Assessment for Insulin Resistance); IPAQ (International Physical Activity Questionnaire); ISCED (International Standard Classification of Education); KORA (Cooperative Health Research in the Region Augsburg); LDL (Low-density Lipid); MEG (magnetoencephalography); MRI (magnet resonance imaging); ODQ-D (Oswestry Low Back Pain Disability); OGTT (oral glucose tolerance test); P10 (Paracelsus 10,000); QoL (Quality of Life); SHIP (Study of Health in Pomerania); US (Ultra sound); WOMAC (Western Ontario and McMaster Osteoarthritis Index);

**Acknowledgements**

The authors thank all contributors of the study, particularly: Stefanie Drach, Johann Sellner, Slaven Pikija, Igor Kedenko, Josef Lang, Alexandra Jarosch, Markus Andress, Katja Vierling, Martin Pühringer, Michael Jordan, Marc Müller, James Anthony and all staff who supported the study during the course of the seven years. We specially express our gratitude to the late Prof. Gunther Ladurner and Priv.-Doz. Dr. Reinhold Fartacek who conceptualized and started the preparation of this project.

**Author Contributions**

Study initiation: Paulweber, Iglseder, Trinka, Haschke-Becher, Studnicka, Sungler.
Data collection: Frey, Raphaelis, Langthaler, Gostner, Martinz, Kedenko.
Conceptualization of manuscript: Frey, Raphaelis, Langthaler, Ring-Dimitriou, Paulweber, Iglseder, Trinka.
Writing: Frey, Raphaelis, Langthaler.
Revision: Ring-Dimitriou, Paulweber, Iglseder, Trinka, Haschke-Becher, Studnicka, Aigner, Kedenko, Bathke.

**Funding**

The study was financially supported by the Land Salzburg, the Salzburger Landeskliniken (SALK, the Health Care Provider of the University Hospitals Landeskrankenhaus and Christian Doppler Klinik), the Paracelsus Medical University Salzburg, the Austrian health insurance (Österreichische Gesundheitskasse-ÖGK) and by unrestricted grants from Bayer, AstraZeneca, Sanofi-Aventis, Boehringer-Ingelheim.

**Statement of Ethics**

All participants signed an informed consent and the study protocol was approved by the ethics committee of the country of Salzburg (415-E/1521/3-2012).
Disclosure Statement

The authors declare that no conflict of interests exists.

References


30 McHorney CA, Ware J, Raczek AE: The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31:247-263.


32 DGE. Nährwertberechnungsprogramm der DGE. Deutsche Gesellschaft für Ernährung e. V. (DGE), Bonn. URL: https://www.dgexpert.de/startseite/.


37 Puhringer M, Ring-Dimitriou S, Stoggl T, Iglseder B, Paulweber B: Differences in the point of optimal ventilatory efficiency and the anaerobic threshold in untrained adults aged 50 to 60 years. Respir Physiol Neurobiol 2020;282.


