מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Consortium Partners:

1. Prof. Eran Segal group Weizmann Institute of Science, Israel
2. Pheno.AI company, SME, Israel

Key Words:
Personalized Medicine, Personalized Nutrition, Large Cohort Study, Deep Phenotyping, Multi-Omics, Microbiome, machine learning, AI, Infrastructure for data accessibility.

Prof. Eran Segal group, Weizmann Institute of Science:

Prof. Eran Segal

Prof. Eran Segal's group is developing computational models for Personalized Medicine. We are a multi-disciplinary team of data scientists, biologists, experimental scientists, medical doctors, and nutritionists working in the area of Computational and Systems biology. The group's research focuses on Genetics, Microbiome, Proteins, Immune System, Gene Regulation and Nutrition, and their effect on health and disease. We aim to develop personalized medicine using machine learning, and analysis of heterogeneous high-throughput omics and clinical data.

The group has constructed a large (>20,000 participants), deep-phenotype, multi-omics cohort of healthy individuals, 40-70 years old (https://www.project10k.org.il/en).
The 10K cohort includes many common disease states such as hypertension, prediabetes and hyperlipidemia, migraine, gynecological, gastrointestinal disorders and many others. Additional data comes from condition specific cohorts including coronary heart disease cohort, active and recovered cancers (pancreas, ovary, breast colorectal), inflammatory bowel disease (IBD), atopic dermatitis, gynecological conditions (endometriosis and polycystic ovary syndrome) and more.

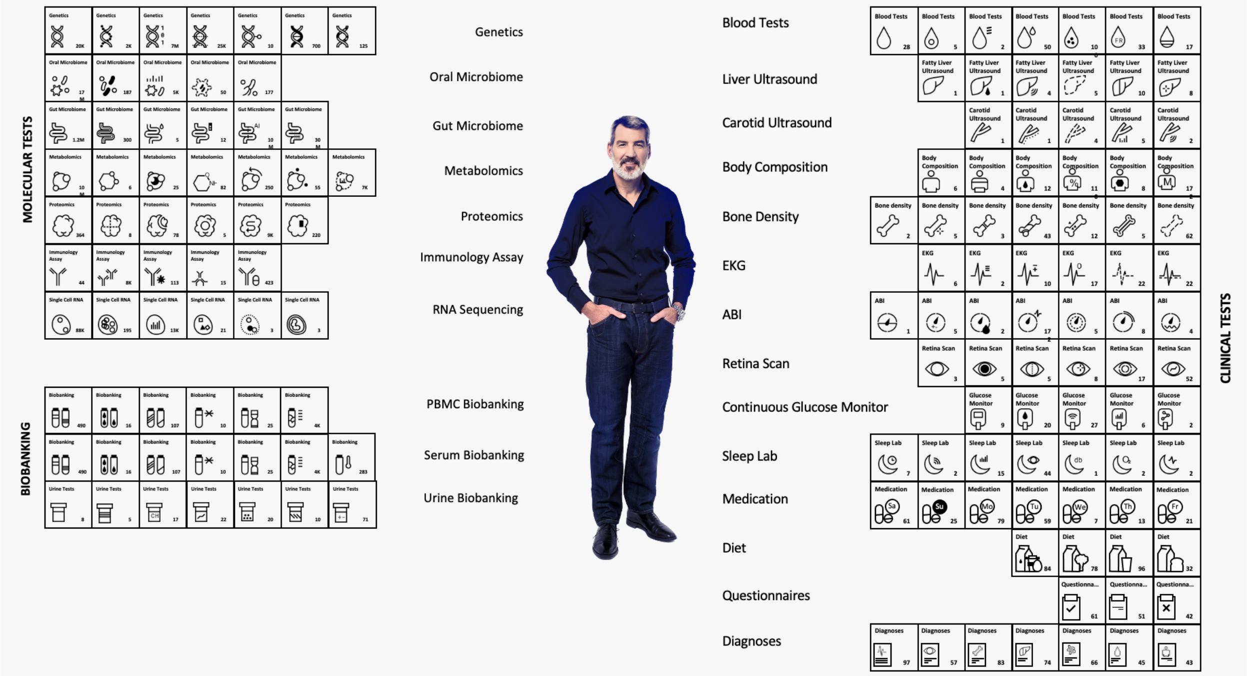


Figure 1: 10K Project Deep-Phenotype Cohort

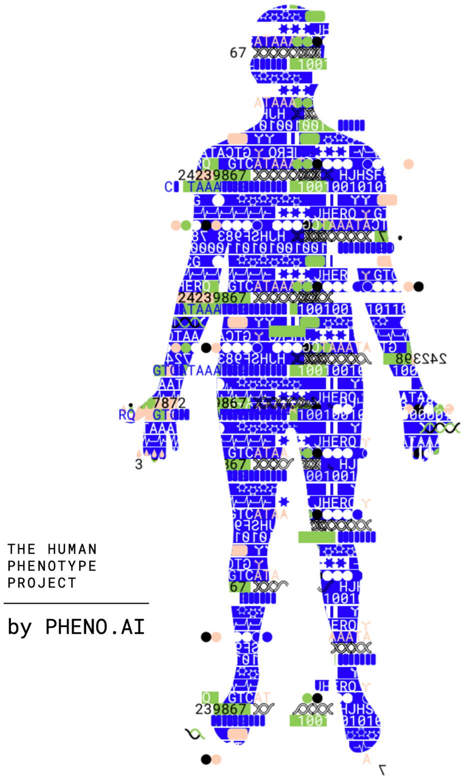
Data sources and group collaborations include:

1. Project 10K, a large, deeply phenotyped cohort study with multiple follow-up points including baseline data of >10,000 participants 40-70 years old with over 2,500 that have already completed a 2 years follow-up. Participants will be followed longitudinally for the next 25 years.
2. Experience working on multiple research protocols on the UKBB cohort databases and constructing disease risk and outcome models.
3. Research on Kupat Holim Clalit insured members data, the largest HMO in Israel with data from >50% of the Israeli population, with experience in constructing AI-based outcome models.
4. Multiple cohort studies in cooperation with hospitals in Israel.
5. Collaborations on microbiome and immunology profiling research with the Lifelines Biobank for health research.

The Segal Group has the expertise and developed high throughput omic tests capabilities including **Human microbiome**, **Metabolomics (lipidomics)**, **Innate and adaptive immune responses, Human Genetic profiling, Transcriptomics** as well as multiple phenotypic tests data collected as part of the cohort (picture above). A great research benefit could result from having multiple types of omic data measured on the same cohort and in multiple reference cohorts and combining those with diet, lifestyle, medical history, and other measures.

The Segal group has extensive experience in analyzing heterogeneous types of biological data, including DNA sequence, RNA structure, genetic, epigenetic, gene expression data, and analysis of microbiome data in numerous research projects. Our work spans a wide variety of advanced statistical methods, probabilistic models (e.g., Bayesian networks), thermodynamic models, and machine learning methods for gene regulation, human genetics, microbiome, and metabolomics. We also devised novel methods, including a new method for extracting information about growth rates of bacteria from metagenomic samples; novel methods for the combined multi-omics analysis of microbiome and host transcriptomics and metabolomics; algorithms for identifying sub-genomic variation in gut bacterial genomes and linking them to host phenotypes, and a new approaches to personalized medicine based on matching healthy controls to single patients and using transfer learning from external cohorts to find insights.

Pheno.AI Small-Medium Enterprise (SME) Collaborator

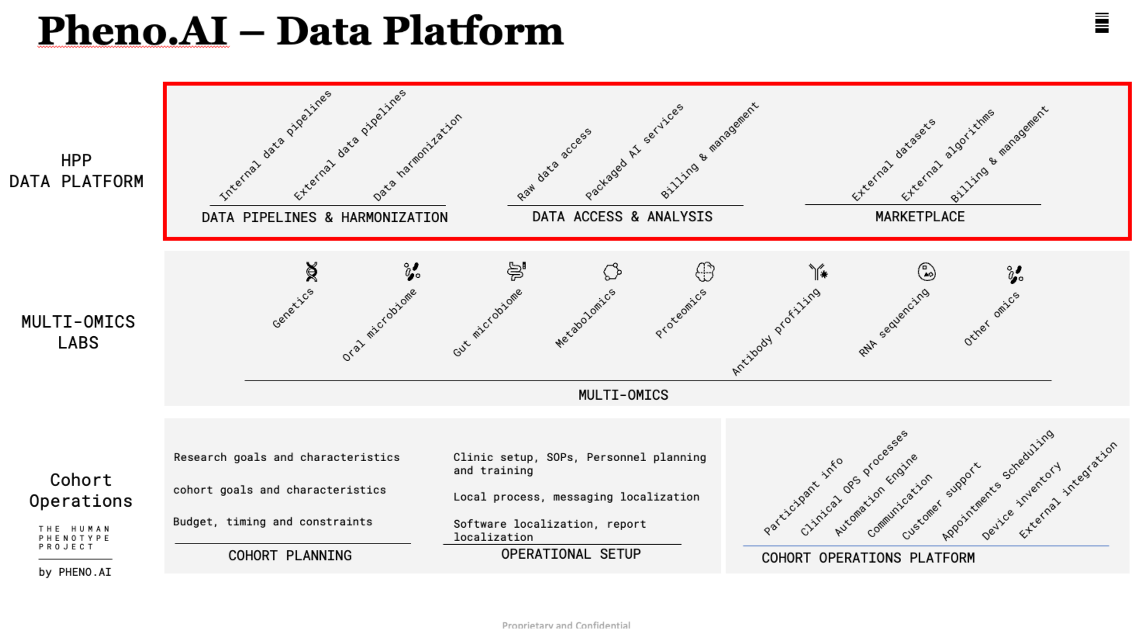


Pheno.AI (https://www.pheno.ai/home) is developing the infrastructure for the setup and operations of large-scale, inter-connected global cohorts and cloud-based systems for study-data accessibility in a secure and protected environment. The infrastructure includes data collection systems, data pipelines, multi-omic labs, and advanced analysis options and services for academic and commercial researchers (Figure 2 below).

Pheno.AI leads the global expansion of **The Human Phenotype Project** - a landmark effort to empower researchers to make the next breakthroughs in human health. The Human Phenotype Project's proprietary datasets are already yielding insights that are advancing science and driving innovation across the pharmaceutical, healthcare, and life-sciences industries.

The company developed proprietary algorithms & methodologies to collect and analyze omic, medical, and cohort studies that are managed under the infrastructure of the Human Phenotype Project Platform which includes the Weizmann Institute 10K Project.

The company business includes a Data-as-a-Service (DaaS) layer on top of deeply phenotyped cohorts. This opens up the platform to researchers who wish to create their own datasets & build their own algorithms. Using the DaaS platform, researchers will be able to exchange algorithms & datasets to collaborate on new discoveries.



Pheno.AI – Data Platform

HPP
DATA PLATFORM

Internal data pipelines
External data pipelines
Data harmonization
DATA PIPELINES & HARMONIZATION

Raw data access
Packaged AI services
Billing & management
DATA ACCESS & ANALYSIS

External datasets
External algorithms
Billing & management
MARKETPLACE

MULTI-OMICS
LABS

Genetics
Oral microbiome
Gut microbiome
Metabolomics
Proteomics
Antibody profiling
RNA sequencing
Other omics
MULTI-OMICS

Cohort
Operations

THE HUMAN
PHENOTYPE
PROJECT
by PHENO.AI

Research goals and characteristics
cohort goals and characteristics
Budget, timing and constraints
COHORT PLANNING

Clinic setup, SOPs, Personnel planning and training
Local process, messaging localization
Software localization, report localization
OPERATIONAL SETUP

Participant info
Clinical OPS processes
Automation Engine
Communication
Customer support
Appointments Scheduling
Device inventory
External integration
COHORT OPERATIONS PLATFORM

Proprietary and Confidential

Figure 2: Pheno.AI platform layers.

Contact Information:

Yotam Reisner MD, PhD
Medical Advisor 10K Project
E: yotam.reisner@weizmann.ac.il
M: +972-54-3123643

HORIZON-HLTH-2024-DISEASE-03-14-two-stage.

"Tackling high-burden for patients, under-researched medical conditions"

Key Words: multi-omics, deep phenotyping cohort, machine learning, microbiome, personalized medicine, personalized nutrition, data accessibility infrastructure, migraine, endometriosis

Prof. Segal Group, Weizmann Institute of Science

The Segal group's research covers data and technologies that provide deep insights into different disease stages. The 10K cohort includes many common disease states such as hypertension, prediabetes and hyperlipidemia, migraine, gynecological, gastrointestinal disorders and many others. Additional data comes from specific condition cohorts (e.g. coronary heart disease cohort, inflammatory bowel disease (IBD), atopic dermatitis, gynecological conditions (endometriosis and polycystic ovary syndrome) and more.

The group can have major contributions to a consortium in having varied ***data sources access*** (including genetic profiling), ***multi-omics methods and high throughput know-hows*** with special expertise in microbiome analysis, and ***data analysis computational capabilities*** that can assist in proposing novel treatment or interventions or technologies for uptake into health systems, especially discovery of new biomarkers and targets for intervention in multiple diseases and medical conditions, identify modifiable factors for prevention and improvement of treatment efficacy.

Pheno.AI company SME

The Pheno.AI contributions: provision of infrastructure for the operations and collaborative analysis of variable datasets including large, multi-omic data sets. Analysis can be performed in secured data rooms, compliant with the highest security standards, enabling within team collaborations. The infrastructure may also enable external entities (academia, industry, pharma) access to analysis of the data, providing new insights. Pheno.AI large Data Science department can perform complex machine learning analysis and building of AI models for cancer prediction.

In addition, Pheno.AI provides the infrastructure for end-to-end operations of longitudinal, deep-phenotype, multi-omics cohorts (healthy and disease) if initiation of such a cohort is planned.

Contact Information:
Yotam Reisner MD, PhD
Medical Advisor 10K Project
E: yotam.reisner@weizmann.ac.il
M: +972-54-3123643

מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

HORIZON-HLTH-2024-STAYHLTH-01-05-two-stage

“Personalized prevention of noncommunicable diseases - addressing areas of unmet needs using multiple data sources”.

Key Words: Deep phenotyping cohort, multi-omics, microbiome, sleep monitoring, nutrition, personalized medicine, personalized nutrition, Clalit HMO data, data accessibility infrastructure, machine learning

Prof. Segal Group, Weizmann Institute of Science
The group can have major contributions to a consortium in varied **data sources access** including the datasets specified on the call like multi-omics data (microbiome, metabolomics, genetic sequencing), sleep monitoring cohort, expertise in food and nutrition logging combined with continuous glucose monitoring. The group has high **data analysis** and **computational capabilities** that can assist in discovery of new biomarkers and vast experience in development of personalized medicine and diet approaches including microbiome based recommended diet algorithms proven effective in prediabetes (Ben-Yacov et al. Diabetes Care 2021;44:1–12).

The 10K Project led by the group includes many common disease states such as hypertension, prediabetes and hyperlipidemia, migraine, gynecological, gastrointestinal disorders and many others NCDs. Additional data comes from specific condition cohorts (e.g. coronary heart disease cohort, inflammatory bowel disease (IBD), atopic dermatitis, gynecological conditions (endometriosis and polycystic ovary syndrome) and more. The 10K Project data together with external datasets (UKBB, Clalit HMO) will enable large datasets for building models for disease risk, course and prevention.

Pheno.AI company SME
An infrastructure for the operations and collaborative analysis of variable datasets including large, multi-omic data sets. Analysis can be performed in secured data rooms, compliant with the highest security standards, enabling within team collaborations. Pheno.AI large Data Science department can perform complex machine learning analysis and building of AI models for cancer prediction.
In addition, Pheno.AI provides the infrastructure for end-to-end operations of longitudinal, deep-phenotype, multi-omics cohorts (healthy and disease) if initiation of such a cohort is planned.

Contact Information:
Yotam Reisner MD, PhD
Medical Advisor 10K Project
E: yotam.reisner@weizmann.ac.il
M: +972-54-3123643

מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

HORIZON-MISS-2023-CANCER-01-01 Call:

“Addressing poorly-understood tumor-host interactions to enhance immune system-centered treatment and care interventions in childhood, adolescent, adult and elderly cancer patients.”

Key Words: Cancer cohort, deep phenotyping healthy cohorts, microbiome, machine learning, multi-omics, infrastructure for data access and collaboration.

Prof. Segal Group, Weizmann Institute of Science
The Segal group’s research covers multiple data and technologies to enable deep insight into tumor-host interactions at different stages of the disease. Data includes a large, deeply phenotyped, *healthy controls* cohort data (10K Project) and samples from *cancer diagnosed patients*, including Pancreas, Breast and Ovarian Cancers. We also established a BRCA-positive cohort (males and females).
The group can have major contributions to a consortium in having varied ***data sources access*** (including genetic profiling), ***multi-omics methods and high throughput know-hows*** with special expertise in microbiome analysis, and ***data analysis computational capabilities*** that can assist in proposing novel treatment or interventions or technologies for uptake into health systems, especially discovery of new cancer biomarkers, targets for intervention.

Pheno.AI company SME
The Pheno.AI contributions: provision of infrastructure for the operations and collaborative analysis of variable datasets including large, multi-omic data sets. Analysis can be performed in secured data rooms, compliant with the highest security standards, enabling within team collaborations. Pheno.AI large Data Science department can perform complex machine learning analysis and building of AI models for cancer prediction.
In addition, Pheno.AI provides the infrastructure for end-to-end operations of longitudinal, deep-phenotype, multi-omics cohorts (healthy and disease) if initiation of such a cohort is planned.

Contact Information:
Yotam Reisner MD, PhD
Medical Advisor 10K Project
E: yotam.reisner@weizmann.ac.il
M: +972-54-3123643