
بسم الله الرحمن الرحیم

ژمیران

ژنوم مرجع ایرانیان

وبینار

کاربرد هوش مصنوعی در پیشبرد پزشکی شخصی

نقش هوش مصنوعی در پیشبرد اهداف پزشکی شخصی
دکتر حسین لنجانیان
متخصص بیوانفورماتیک

پزشکی شخصی و راه های اجرایی شدن آن در ایران
دکتر مریم دانشپور
متخصص ژنتیک پزشکی

پنجشنبه ۴ خرداد ۱۴۰۲ (۱۲-۱۰)
ادرس ثبت نام: ircme.ir
لینک ورود (Adobe Connect): lms1.sbmu.ac.ir/emrc
لینک ضبط شده وبینار: emrc.sbmu.ac.ir
امتیاز باز آموزی: ۲/۵
شناسه برنامه: ۱۹۵۹۴۷

GENOSCOPE DIAGNOSTIC
ژن کاوش آزما

Outline

- Definition
- Theory
- Big data in healthcare
- Example: Pharmacogenetic, Monogenic, Multifactorial
- Medical coding
- PM Initiation
- Gemiran (Tehran Cardiometabolic Genetic Study)

What is precision medicine?

Quantile-dependent expressivity

Quantile-specific heritability

- It appears that environmental factors sort of set the groundwork in which your genes start to have an effect
- If your surroundings predispose you to drinking more coffee – like your coworkers or spouse drink a lot, or you live in an area with a lot of cafes – then the genes you possess that predispose you to like coffee will have a bigger impact. These two effects are synergistic

This approach will allow doctors and researchers to **predict more accurately** which treatment and prevention strategies for a particular disease will work in **which groups of people**.

Pharmacogenetic

In precision medicine

Pharmacogenomics is a part of precision medicine

- Pharmacogenomics is the study of how genes affect a person's **response to particular drugs**.
- This relatively new field combines
  - Pharmacology (the science of drugs)
  - Genomics (the study of genes and their functions)
- To develop effective, safe medications and doses that are tailored to variations in a person's genes.

Example of a pharmacogenetic investigation

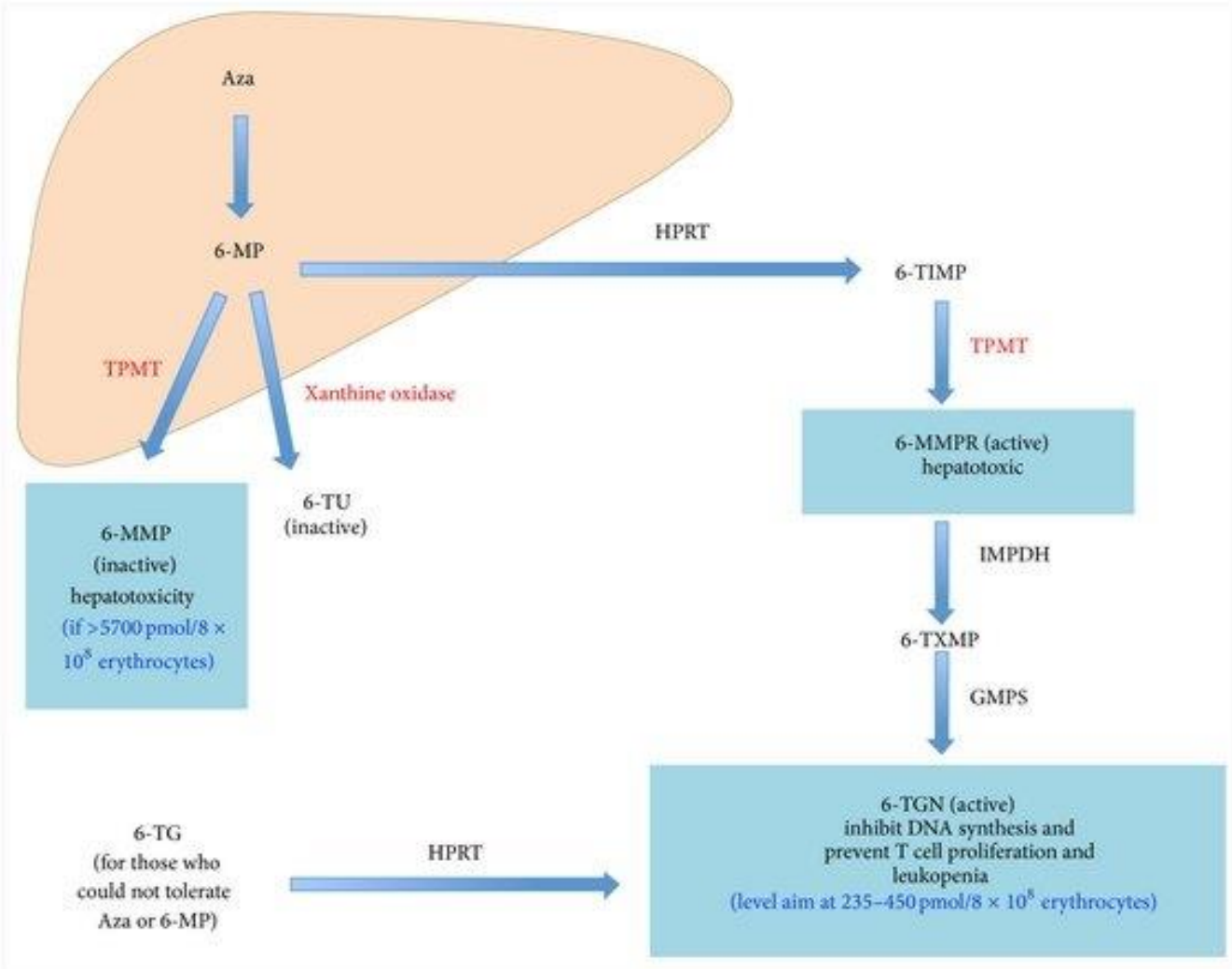
A 19-year-old patient diagnosed at the age of 8 years with common variable immunodeficiency (CVI) with mainly gastrointestinal manifestations.

He has been diagnosed with ulcerative jejunitis receiving different lines of treatment (mesalazine and corticotherapy) without achieving a correct control of the disease. Faced with this situation, the responsible physician decided to initiate treatment with azathioprine (AZT) as the next therapeutic step.

AZT is a purine analog that interferes with protein and nucleic acid synthesis. It is an immunomodulatory drug that exerts its action intracellularly, and its effect is delayed, the maximum effect being achieved 10-12 weeks after the start of treatment.

10-12 weeks after the start of treatment.

The effect of the treatment will depend on the balance between different enzymatic reactions related to the metabolism of Thiopurines: the inactivating ones (catalyzed by Thiopurine methyltransferase -TPMT- and also Xanthine oxidase), and those producing active metabolites (mainly Hypoxanthine ribosyltransferase). This balance will control the intracellular levels of 6-thioguanine; the antimetabolite that is incorporated in DNA synthesis, inhibits lymphocyte proliferation and, in excess, causes toxicity.



Thiopurin methyltransferase is encoded by the TPMT gene which is a polymorphic gene. The enzymatic activity of TPMT (phenotype) is conditioned by single nucleotide polymorphisms (SNPs). More than 20 allelic variants affecting enzyme activity have been described. The wild type genotype of TPMT (*1) is observed in approximately 90%. The most frequent mutant allelic variants in the population are *2, *3A *3B and *3C which result from the combination of two SNPs and may result in the substitution of one or two amino acids in the translated protein, leading to loss of enzyme activity, increased levels of 6-thioguanine and increased risk of myelosuppression.

NUDT15 polymorphisms			
Exon	Allele	SNP	Amino acid change
1	NUDT 15*5	c. 52 G > A	Val 18 Il
1	NUDT 15*2	c. 36_37 ins GGAGTC	Val 18_alV 19 ins GlyVal
3	NUDT 15*3	c. 415 C > T	Arg 139 Cys
3	NUDT 15*4	c. 416 C > T	Arg 139 His

The NUDT15 gene encoding for the enzyme nudix hydrolase 15, which is also involved in the metabolism of thiopurines. It is also a polymorphic gene. The wild type genotype of NUDT15 (*1) is observed in approximately 90%. Four allelic variants affecting enzyme activity have been described. The most relevant polymorphism in Caucasian population is NUDT15*3.

Depending on the combinations of the genotypes of these polymorphisms, different phenotypes can be defined for each individual:

Phenotype	TPMT	NUDT15
Normal metabolizer (NM)	'1 / '1	'1 / '1
Intermediate metabolizer (IM)	'1 / '2, '1 / '3A, '1 / '3B, '1 / '3C, '1 / '4	'1 / '3
Poor metabolizer (PM)	'2 / '2, '2 / '3A, '2 / '3B, '2 / '3C, '2 / '4, '3A / '3A '3A / '3B '3A/'3C, '3A/ '4, '3B / '3B, 3B / '3C, '3B / 4, '3C / '3C, '3C / '4, '4 / '4	'3 / '3

The metabolizing phenotype:

01

The metabolizing phenotype of the patient is associated with the risk of suffering adverse effects,

02

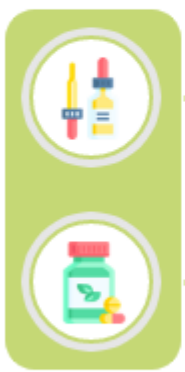
So depending on the phenotype of each patient it is possible to recommend using other treatments or dose reductions of 90 % in PM phenotype (2 deficient alleles) and 50 % in IM phenotype (one deficient allele).

It was decided to perform a pharmacogenetic analysis for the TPMT (rs2842934, rs2842934, rs1800460, rs1800584 and rs1142345) and NUDT15 (rs116855232, rs147390019, rs554405994 and rs186364861) polymorphisms by sequencing technique after signing the informed consent.

The results of the study showed that the patient did not carry any mutations in the series of genes analysed.

Gen	Polymorphism	Resulta
TPMT	rs 1800462 G>C	G/G (WT)
TPMT	rs 1800460 G>A	G/G (WT)
TPMT	rs 1142345 A>G	A/A (WT)
TPMT	rs 180058 ª G>A	G/G (WT)
NUDT15	Rs 116855232 C>T	G/G (WT)
NUDT15	Rs 147390019 G>T	G/G (WT)
NUDT15	Rs 554405994 (GGA/GTCG)> (GGAGTCG)	(GGA/GTCG) (WT)
NUDT15	Rs 186364861 G>A	G/G (WT)

Based on the results of the pharmacogenetic test, treatment with azathioprine at full doses was started.

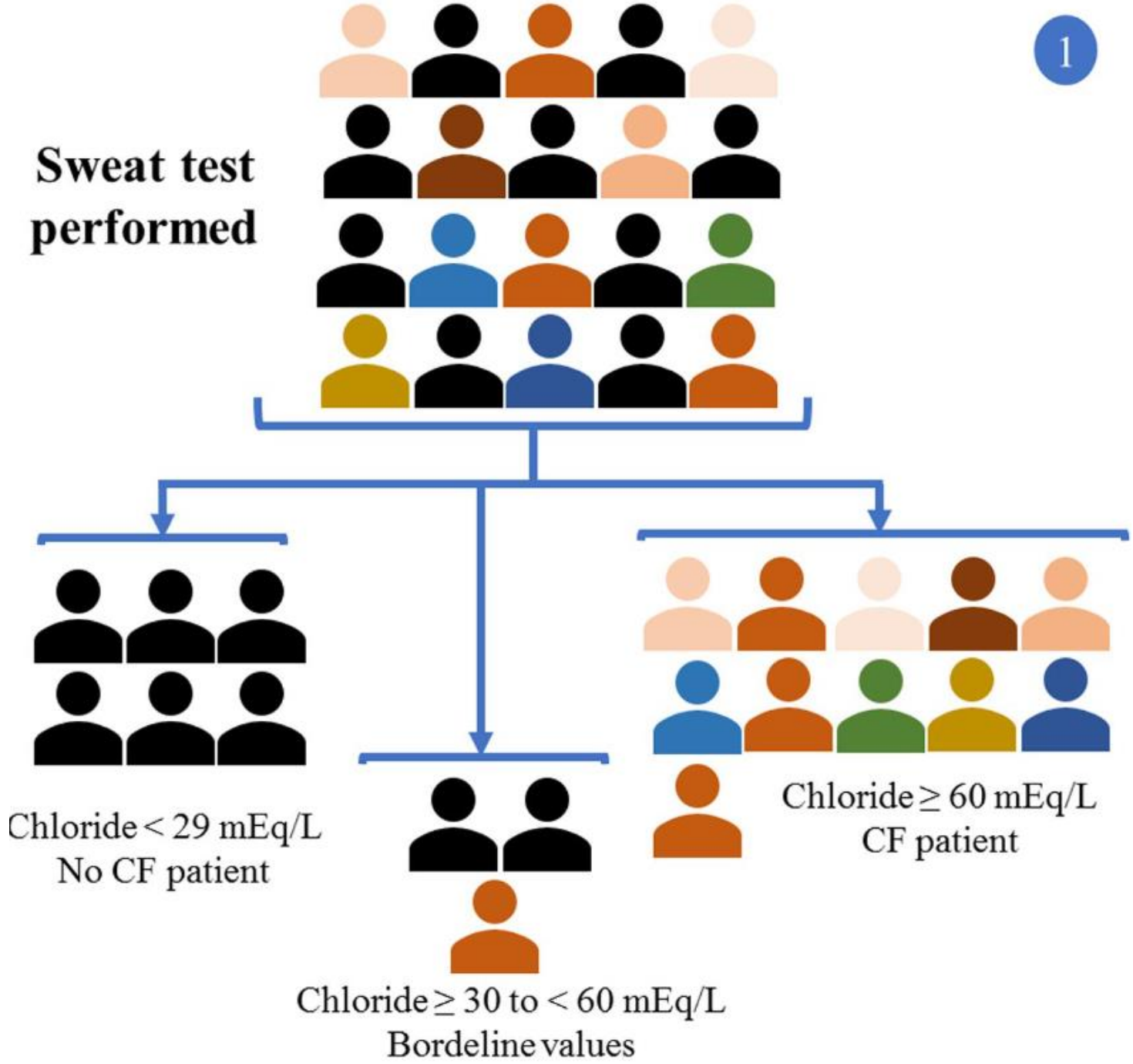


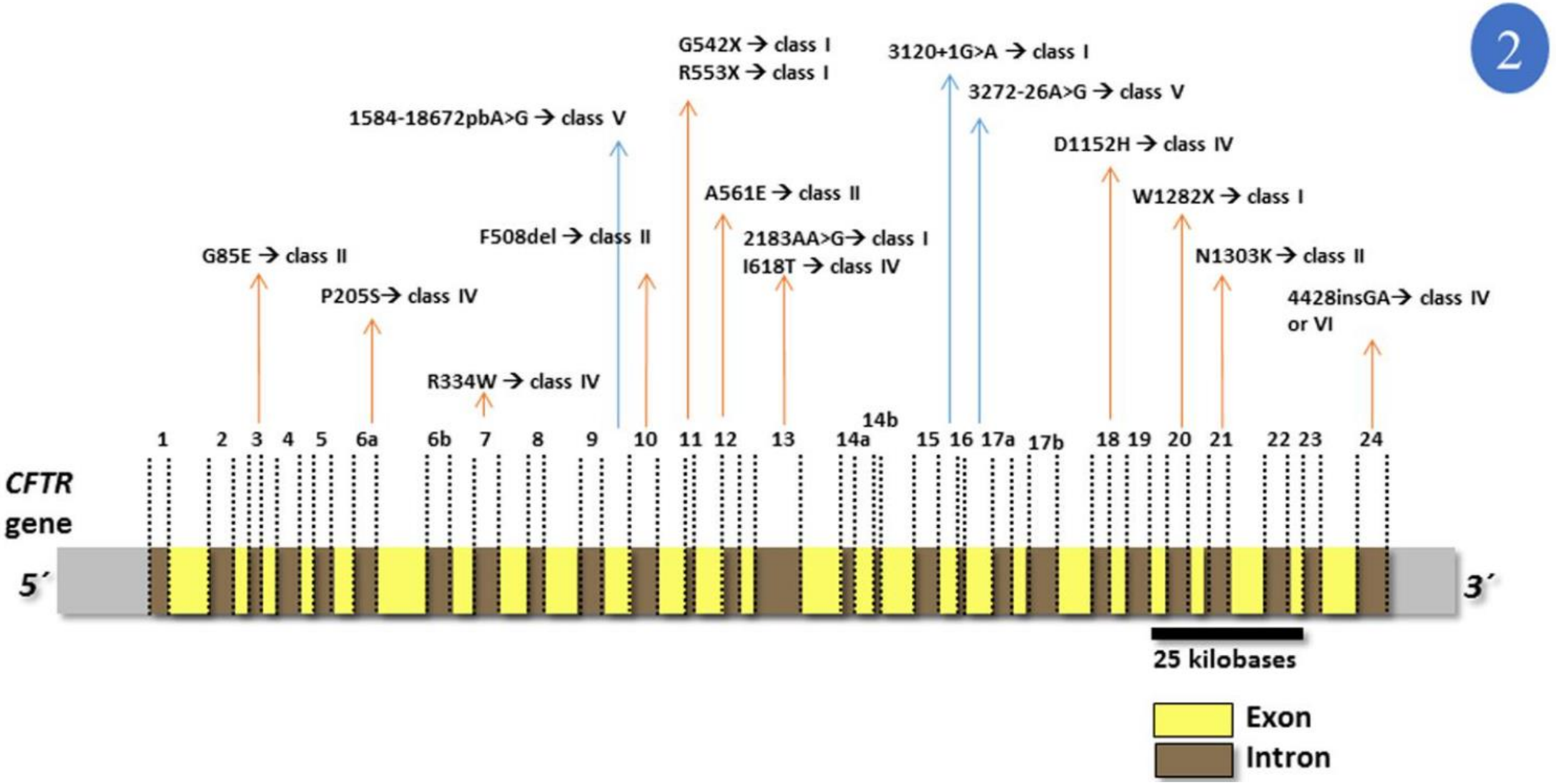
After six months of treatment, adequate control of inflammatory bowel disease was achieved, with no alteration of neutrophils or other toxicity associated with azathioprine treatment.

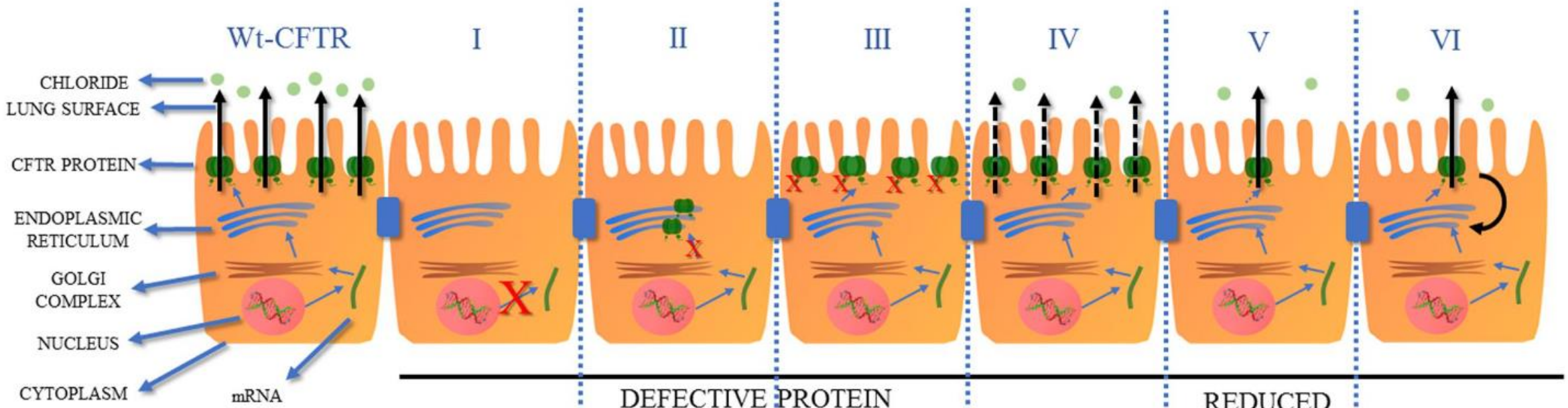
Currently the patient continues with his usual treatment.

Monogenic disorders

In Precision Medicine


1
Sweat test performed
Chloride < 29 mEq/L
No CF patient
Chloride ≥ 30 to < 60 mEq/L
Bordeline values
Chloride ≥ 60 mEq/L
CF patient


2
G542X → class I
R553X → class I
3120+1G>A → class I
3272-26A>G → class V
1584-18672pbA>G → class V
D1152H → class IV
A561E → class II
W1282X → class I
F508del → class II
2183AA>G→ class I
I618T → class IV
N1303K → class II
G85E → class II
P205S→ class IV
4428insGA→ class IV
or VI
R334W → class IV
1 2 3 4 5 6a 6b 7 8 9 10 11 12 13 14a 14b 15 16 17a 17b 18 19 20 21 22 23 24
CFTR gene
5´
3´
25 kilobases
Exon
Intron

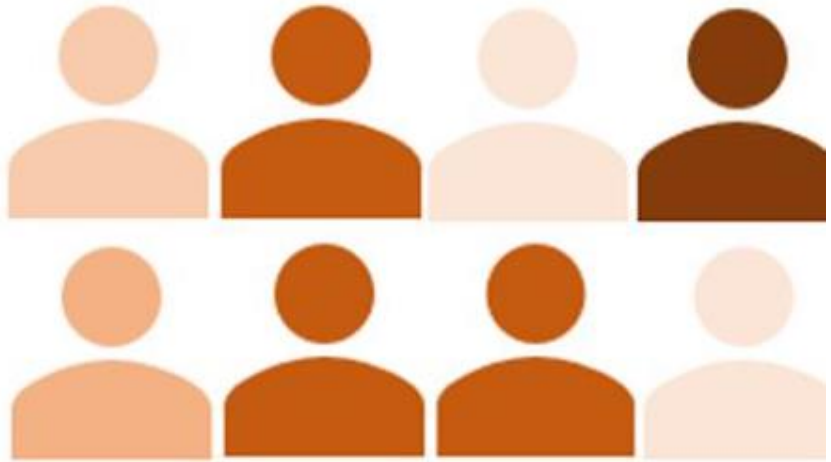


	IA	IB					
CFTR defect	No mRNA	No protein	No traffic	Impaired gating	Decreased conductance	Less protein	Less stable
Mutation examples	Dele2,3(21 kb), 1717-1G→A	Gly542X, Trp1282X	F508del, Asn1303Lys, Ala561Glu	Gly551Asp, Ser549Arg, Gly1349Asp	Arg117His, Arg334Trp, Ala455Glu	3272-26A→G, 3849+10 kg C→T	c.120del23, rPhe580del
Corrective therapy	Unrescuable	Rescue synthesis	Rescue traffic	Restore channel activity	Restore channel activity	Correct splicing	Promote stability
Drugs (approved)	Bypass therapies (no)	Read-through compounds (no)	Correctors (yes)	Potentiators (yes)	Potentiators (no)	Antisense oligonucleotides, correctors, potentiators? (no)	Stabilizers (no)
Clinical features (global aspect)	**MORE SEVERE DISEASE**				**LESS SEVERE DISEASE**		

Numerous mutation classes can be identified and different therapeutic strategies can be used

Other mutations

The same treatment based on personalized medicine can be applied



F508del/F508del mutation (Class II)

Multifactorial disorders

In Precision Medicine

Diabetes in Precision Medicine era

Radical changes in our ability to characterize and understand human biological variation through:

1) Assessment of the genetic and metabolic state
2) Leveraging data to inform disease categories
3) Science-guided preventive and treatment decisions tailored to specific pathological conditions.

Coupling these with detailed (digital)information about:

- lifestyle
- environment

present opportunities to optimize diabetes medicine.

Diabetes in Precision Medicine era

Our modern capacity to comprehensively interrogate diverse axes of biology may radically transform the practice of medicine. These axes include:

1) Developmental/ metabolic context
2) Genomic variation
3) Chromatin signals, that mark genes as active or repressed in tissues
4) Expressed transcripts
5) Biomarkers of disease
6) Increased knowledge of life style/environmental risk factors.
7) Parallel advances in computational power and analytical methods

Diabetes in Precision Medicine era

There are, however, several reasons for **hope**:

1) Diabetes caused by single gene defects can be characterized and targeted therapies are particularly effective.
2) Islet autoantibody biomarkers and genomic risk have clarified autoimmune diabetes from other forms of the disease, thereby facilitating immune intervention trials and preonset monitoring to reduce risk of severe complications and aiding in detection of environmental triggers.
3) Multiple biomarkers and genetic variants have been shown to alter risk of T2D.
4) T2D has been shown to be a complex combination of multiple conditions and processes, defined by process specific subgroups in which individuals with extreme burdens of risk in particular pathways reside and for whom a specific therapeutic approach may be optimal.
5) The tools, resources, and data now exist to determine the biological and lifestyle/environmental predictors of drug response.

Diabetes in Precision Medicine era

- Precision diabetes medicine refers to an approach to optimize the
  1) Diagnosis
  2) Prediction
  3) Prevention
  4) Treatment

by integrating multidimensional data, accounting for individual differences.

- The major distinction from standard medical approaches is the use of complex data to characterize the individual's health status, predisposition, prognosis, and likely treatment response.
- Precision medicine also focuses on identifying patients who, despite a diagnosis, do not require treatment (or require less than might conventionally be prescribed).

The pathophysiology of diabetes is complex with multiple causes, phenotypes, trajectories and consequences

- Theoretically, there can be many diabetes subphenotypes characterized by different combinations of
  - Molecular features
  - Pathophysiological processes
  - Risk factors
  - Complications
  - Comorbidities.
- These phenotypes can be altered by
  - Self-management
  - Quality of care
  - Drug treatments
  - All of which can influence clinical outcomes.

Type 1 diabetes

- Autoimmune destruction of the pancreatic islet, due to interaction between genetic susceptibility, perturbed immunology and environmental factors
  - INS (insulin)
  - PTPN22 (protein tyrosine phosphatase, nonreceptor type 22)
  - IL2RA (interleukin-2 receptor subunit alpha)
  - IFIH1 (interferon induced with helicase C domain 1)
  - CTLA4 (cytotoxic T-lymphocyte associated protein 4) loci
  - TCF7-P19T (transcription factor 7-P19T)

Type 2 diabetes

Classical examples of polygenic and complex diseases as a result of interactions between multiple genetic and environmental factors 113 loci have been associated with type 2 diabetes

The majority of these SNPs are associated with

- ✓Islet development and glucose sensing
- ✓Insulin synthesis
- ✓Secretion
- ✓Signaling or resistance

Whereas others are associated with metabolic traits, such as obesity, which frequently coexists with diabetes

Gestational diabetes mellitus (GDM)

Diabetogenic condition characterized by multiple hormonal changes with increased insulin resistance

Maturity-onset diabetes of the young (MODY)

- 1–2% of all cases of diabetes
- Autosomal dominant disorders characterized by nonketotic and/or non-acute presentation, typical of type 2 diabetes, but occurring at a younger age, usually before the age of 25 years
- Because of their rapid failure with oral drugs and/or young onset of presentation, MODY patients can be misdiagnosed as type 1 diabetes
- Alternatively, because of their low risk of ketosis, they might simply be classified as having type 2 diabetes

- 14 genetic subtypes of MODY each with distinct clinical characteristics and responsible genes
- In these young individuals with familial early-onset diabetes with or without typical features, genetic testing is required to increase the precision of diagnosis, which has implications on treatment selection and family screening

Subtype	Gene	Location	Etiology	Features
MODY 1(82)	*HNF4α*	20q13.12	Insulin secretion defect	Progressive hyperglycemia
MODY 2(75)	Glucokinase	7p13	Glucose sensing and insulin secretion defect	Early onset; mild hyperglycemia, minor microvascular disease
MODY 3(83)	*HNF1α*	12q24.31	Insulin secretion defect	Progressive hyperglycemia, sensitive to SU
MODY 4(84)	*PDX1/IPF1*	13q12.2	Insulin secretion defect	Early onset
MODY 5(85)	*HNF1β*	17q12	Insulin secretion defect	Variable age at onset, range infancy to adult; progressive hyperglycemia, renal cysts; renal failure, require insulin treatment
MODY 6(86)	*NeuroD1*	2q31.3	Insulin secretion defect	Early onset
MODY 7(87)	*KLF11*	2p25.1	Insulin secretion defect	Very rare
MODY 8(88)	*CEL*	9q34.13	β-cell defect	Endocrine and exocrine pancreatic insufficiency
MODY 9(89)	*PAX4*	7q32.1	Little data	Very rare
MODY 10(90)	*INS*	11p15.5	Insulin secretion defect	Diagnosed in patients aged in their 20s to 30s. Can cause neonatal diabetes, antibody negative type 1 diabetes, and MODY
MODY 11(91)	*BLK*	8p23.1	Defect in insulin synthesis and secretion	Onset often before age 25 years; some patients require insulin for treatment
MODY 12(92)	*ABCC8*	11p15.1	Little data	Frequent cause of neonatal diabetes, but can rarely cause MODY
MODY 13(93)	*KCNJ11*	11p15.1	Insulin secretion defect	Sulfonylurea therapy effective
MODY 14(94)	*APPL1*	3p14.3	Defect in insulin signaling pathway	With elevated FBG and HbA1C and onset between 30s and 50s

EXETER DIABETES

The Exeter Diabetes App provides information on diagnosing and treating subtypes of diabetes

MODY Calculator

Tests for Diabetes Subtypes

Type 1/Type 2
Diabetes Classification

Treatment decisions
in Type 2 diabetes

Information about the app/calculators

Guidelines For Genetic Testing In MODY

Introduction:

The purpose of these criteria is to ensure that NHS resources are used effectively to diagnose monogenic diabetes. These new NHS England National Genomic Test Directory Testing Criteria for Rare and Inherited Disease (R141 Monogenic diabetes, R142 Glucokinase-related fasting hyperglycaemia and R143 Neonatal Diabetes) were published in October 2021 (https://www.england.nhs.uk/wp-content/uploads/2018/08/rare-and-inherited-disease-eligibility-criteria-v3-2.pdf.pdf). The criteria has been set to keep the positive rate of tests performed at 25% overall as it has been over the past decade of testing for monogenic diabetes. They will be reviewed and revised should there be any significant change to this pick-up rate.

The criteria apply to the proband (i.e. the first member of a family with diabetes to be tested). Once a genetic diagnosis of monogenic diabetes has been confirmed in the proband, other family members will be eligible for testing of the familial variant.

Genetic testing for monogenic diabetes (R141 and R143) will only be performed on patients confirmed to have diabetes by laboratory blood glucose or HbA1c according to the WHO definition unless they meet criteria for Glucokinase related fasting hyperglycaemia (R142).

Testing indications:

- Diabetes – Maturity-onset diabetes of the young (R141)
- Glucokinase related fasting hyperglycaemia (R142)
- Neonatal diabetes (R143)
- Diabetes and non-autoimmune extra pancreatic features (R141)
- Diabetes with severe insulin resistance (R141)

Diabetes – Maturity-onset Diabetes Of The Young (R141)

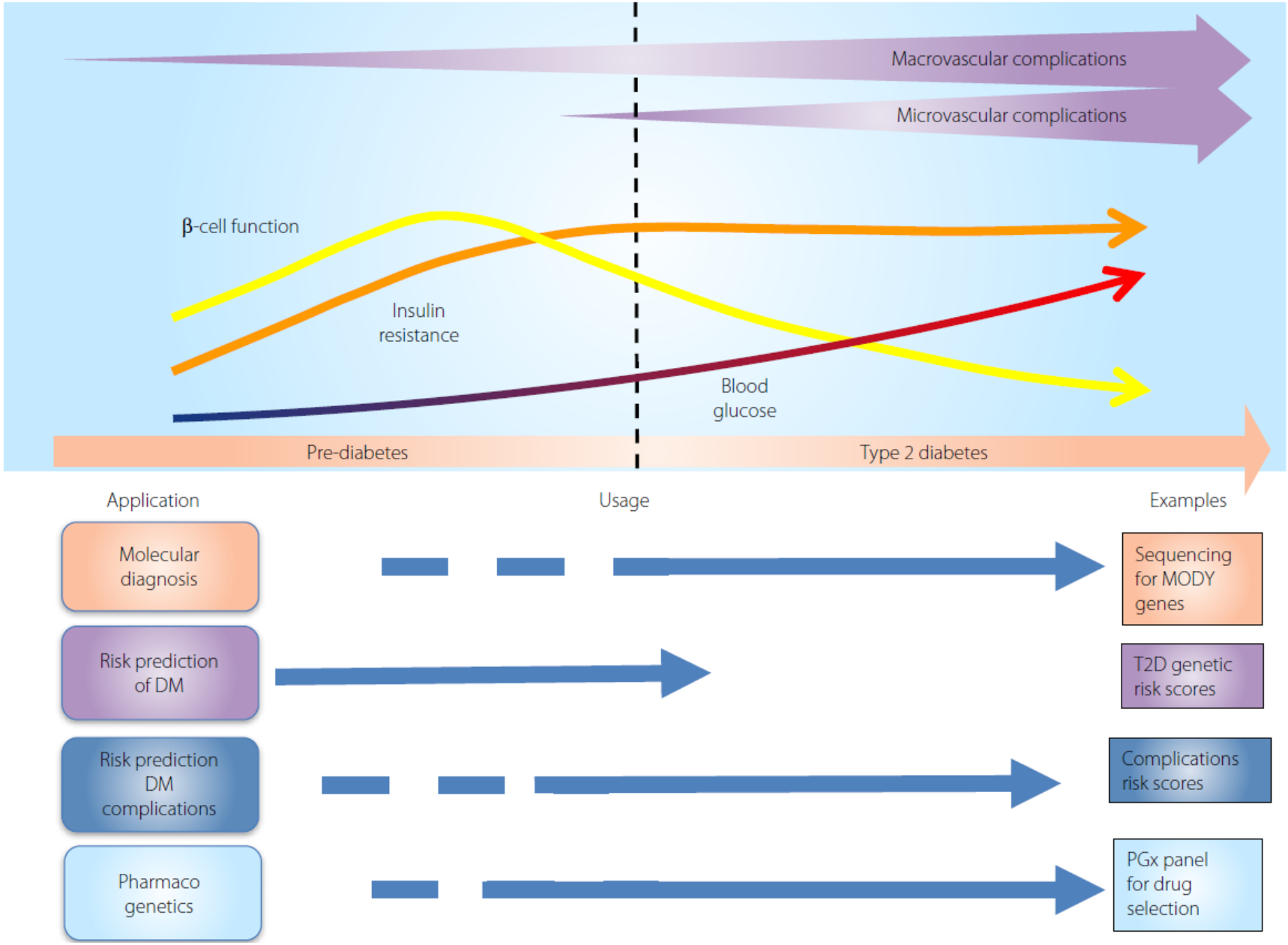
Glucokinase-related Fasting Hyperglycaemia (R142)

Neonatal Diabetes (R143)

Syndromic Diabetes (R141): Patients With Diabetes AND Non-autoimmune Extra-pancreatic Features

Neonatal diabetes

- Several national registers show that the incidence of neonatal diabetes, diagnosed before the age of 6 months, is approximately one in 100,000 live births.
- Amongst the reported mutations associated with neonatal diabetes nearly half were as a result of mutations in the genes encoding the KATP (ABCC8 and KCNJ11) with good response to SU


Macrovascular complications
Microvascular complications
β-cell function
Insulin resistance
Blood glucose
Pre-diabetes
Type 2 diabetes
Application
Usage
Examples
Molecular diagnosis
Risk prediction of DM
Risk prediction DM complications
Pharmaco genetics
Sequencing for MODY genes
T2D genetic risk scores
Complications risk scores
PGx panel for drug selection

New disease taxonomy construction and validation

- These behavioral and social determinants and other exogenous factors can now be tracked and measured by wearables and a range of medical devices.
- These factors account for about 60% of our determinants of health (behavioral, socio-economical, physiological, and psychological data), our genes account for about 30%, and last our actual medical history accounts for a mere 10%.
- Over the course of our lifetimes, we will each generate the equivalent of over **300 million books of personal and health-related data** that could unlock insights to a longer and healthier life.

Phenomenon of big data (five Vs)

1. **Volume:** Vast amount of complex and heterogenous data, which makes data sets too large to store and analyze using traditional database technology.
2. **Velocity:** Speed at which new data are generated and moves around.
3. **Variety:** Different types of structured, semistructured, and unstructured data, such as social media conversations and voice recordings.
4. **Veracity:** Certainty, accuracy, relevance, and predictive value of the data.
5. **Value:** Conversion of data into business insights.

AI and spectrum of health care (5 Ps)

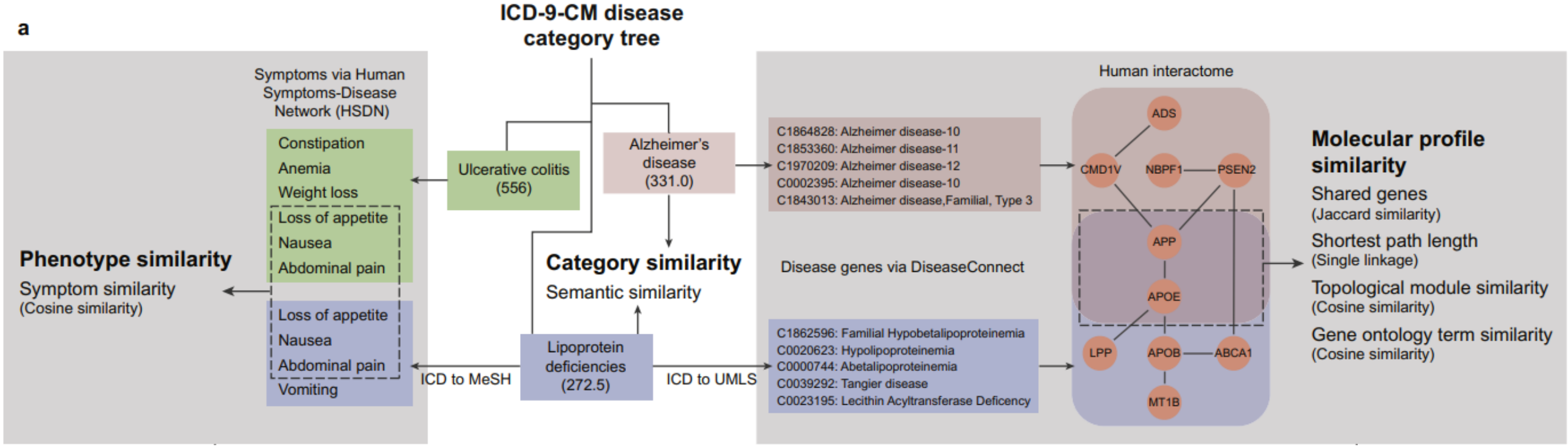
1. **Payer**
2. **Provider**
3. **Policy maker/government**
4. **Patients**
5. **Product manufacturers**

Reliable identification of medical coding errors and incorrect claims positively impacts payers, providers, and governments by saving inordinate amounts of money, time, and efforts

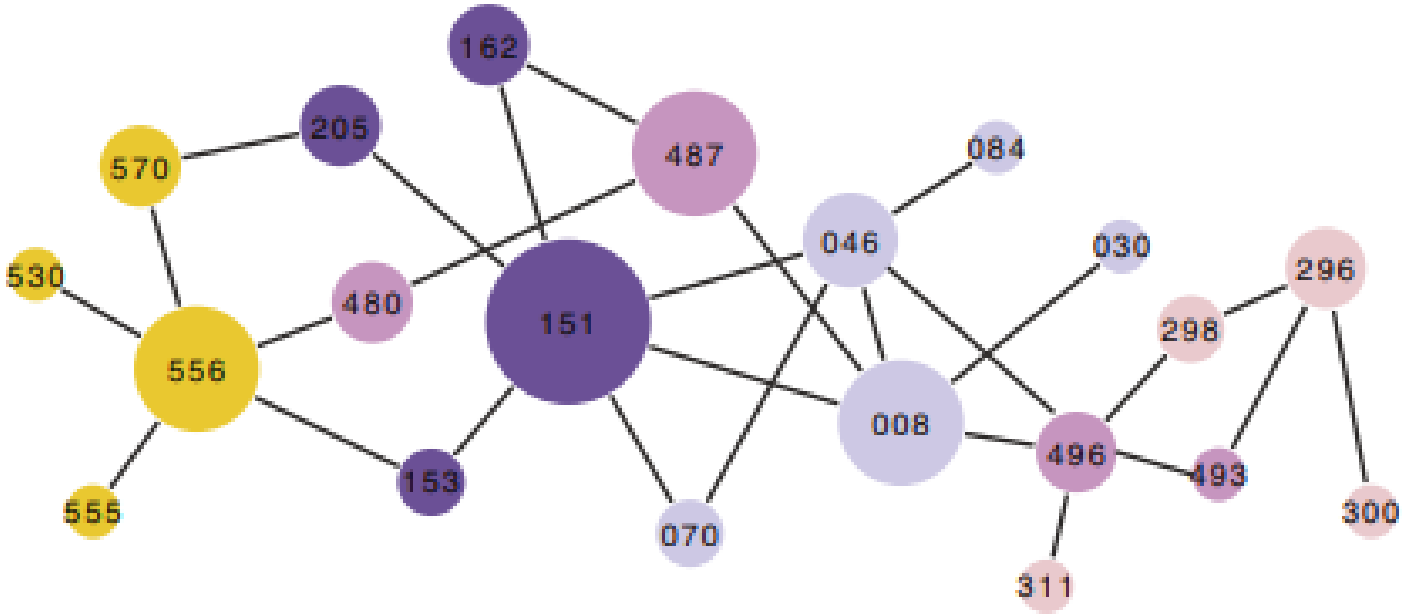
Current classification system for diseases

- International Classification of Diseases (ICD)
  - Clinical features
    - Symptoms
    - Examination of diseased tissues and cells
- Is it insufficient!
  - Lacks the depth required for precision medicine
    - Its rigid hierarchical structure and does not take into account the rapidly expanding **molecular insights** of disease phenotypes.
    - Many diseases have high **genetic heterogeneity or manifestation diversity**, which makes it difficult to tailor treatment to a patient's pathophysiology.
    - Additionally, **disease comorbidities, temporal disease trajectories, and various molecular** relationships between disease-associated cellular components and their connections in the interactome demonstrate the vague boundary between different diseases in current disease taxonomies.
- Therefore, efforts to reclassify diseases based on molecular insights have increased in the past decade to meet the needs of precision medicine.

Similarity calculation between the disease pairs in ICD taxonomy



Module or community annotations of disease association network by chapters in ICD

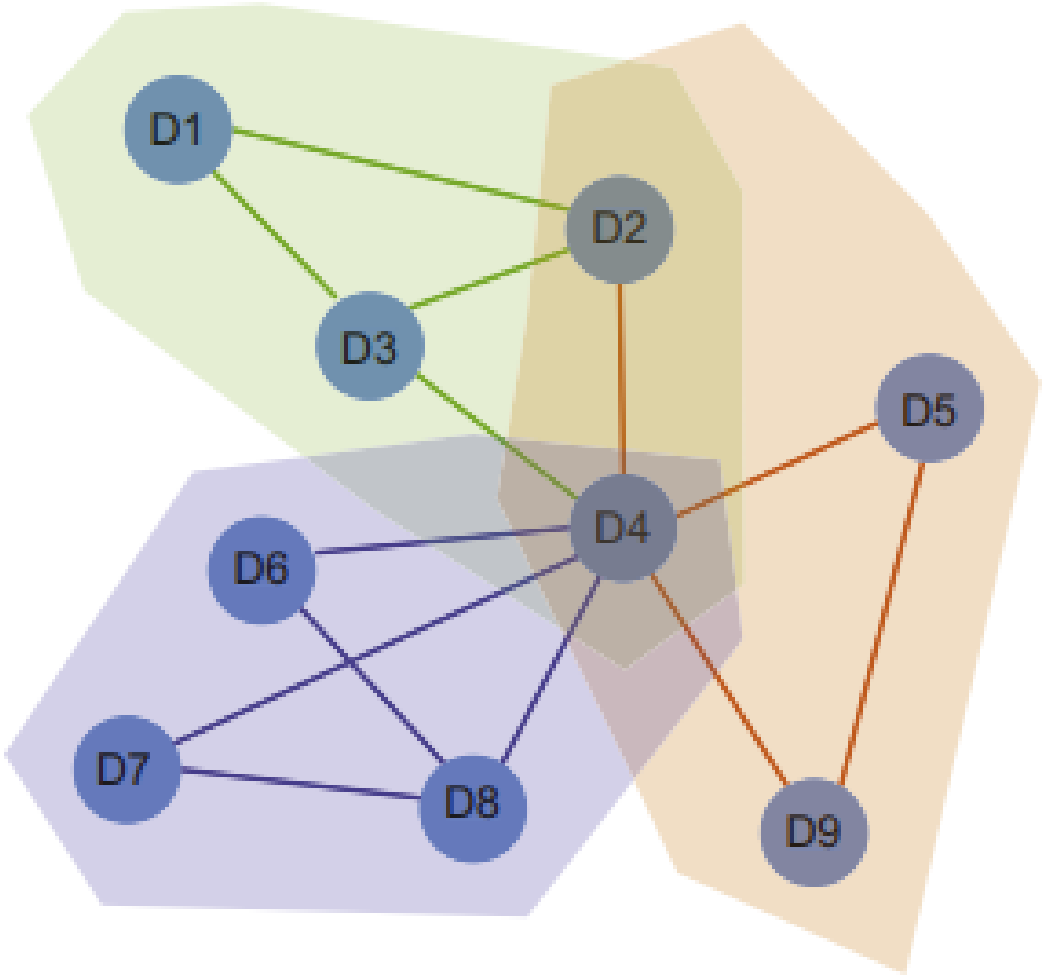


Disease association network
(with phenotype or molecular profile similarity)

Module annotations

- ICD Chapter 1 (Infectious diseases)
- ICD Chapter 2 (Neoplasms)
- ICD Chapter 5 (Mental disorders)
- ICD Chapter 8 (Respiratory diseases)
- ICD Chapter 9 (Digestive diseases)

Construction of integrated disease network (IDN) and generation of NCD

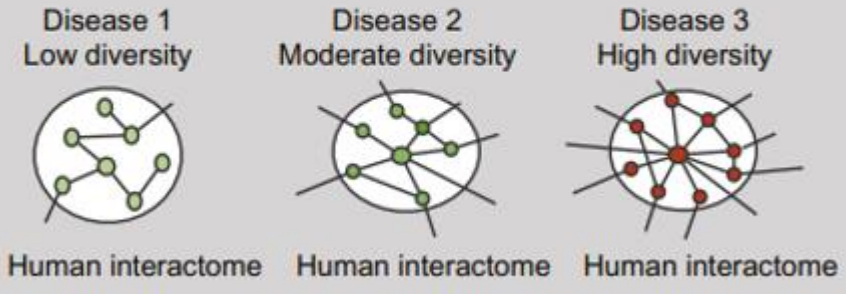


Integrated disease network (IDN) to generate overlapping new disease categories (NCD)

- Shared protein-protein interaction modules
- Shared genes
- Shared symptoms

Quality evaluation and validation of ICD and NCD

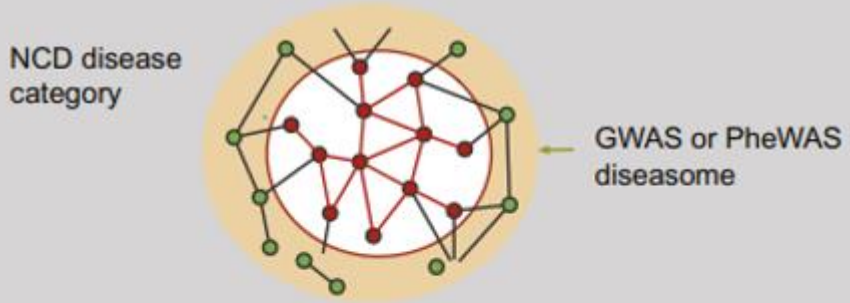
Disease and category diversity to show the molecular specificity of disease phenotypes



Quality evaluation and validation

Network modularity to evaluate the association density of disease phenotypes in disease categories

Disease overlap with GWAS and PheWAS data to validate the robustness of NCD



Precision Medicine Initiative

What is the Precision Medicine Initiative?

- The Precision Medicine Initiative is a long-term research endeavor, involving the National Institutes of Health (NIH) and multiple other research centers, which aims to understand

how a person's genetics, environment, and lifestyle can help determine the best approach to prevent or treat disease

- Goals
  - Short-term
    - Expanding precision medicine in the area of cancer research. Researchers at the National Cancer Institute (NCI) hope to use an increased knowledge of the genetics and biology of cancer to find new, more effective treatments for various forms of this disease.
  - Long-term
    - bringing precision medicine to all areas of health and healthcare on a large scale.

What are some potential benefits of precision medicine and the Precision Medicine Initiative?

- To this end, the NIH has launched a study, known as the **All of Us Research Program**, which involves a group (cohort) of at least 1 million volunteers from around the United States.
- Participants are providing genetic data, biological samples, and other information about their health.
- To encourage open data sharing, participants can access their health information, as well as research that uses their data, during the study.
- Researchers can use these data to study a large range of diseases, with the goals of better predicting disease risk, understanding how diseases occur, and finding improved diagnosis and treatment strategies.
- Precision medicine holds promise for improving many aspects of health and healthcare. Some of these benefits will be apparent soon, as the All of Us Research Program continues and new tools and approaches for managing data are developed.
- Other benefits will result from long-term research in precision medicine and may not be realized for years.

U.S. Department of Health & Human Services | National Institutes of Health

NIH National Institutes of Health
All of Us Research Program

Log In | JOIN NOW | Search

About | Get Involved | Funding and Program Partners | Protecting Data and Privacy | News and Events

Questions about COVID-19?
VISIT cdc.gov/coronavirus | Español

Learn how the *All of Us* Research Program is addressing COVID-19.

The future of health begins with you.

The *All of Us* Research Program is inviting one million people across the U.S. to help build one of the most diverse health databases in history. We welcome participants from all backgrounds. Researchers will use the data to learn how our biology, lifestyle, and environment affect health. This may one day help them find ways to treat and prevent disease.

JOIN NOW!

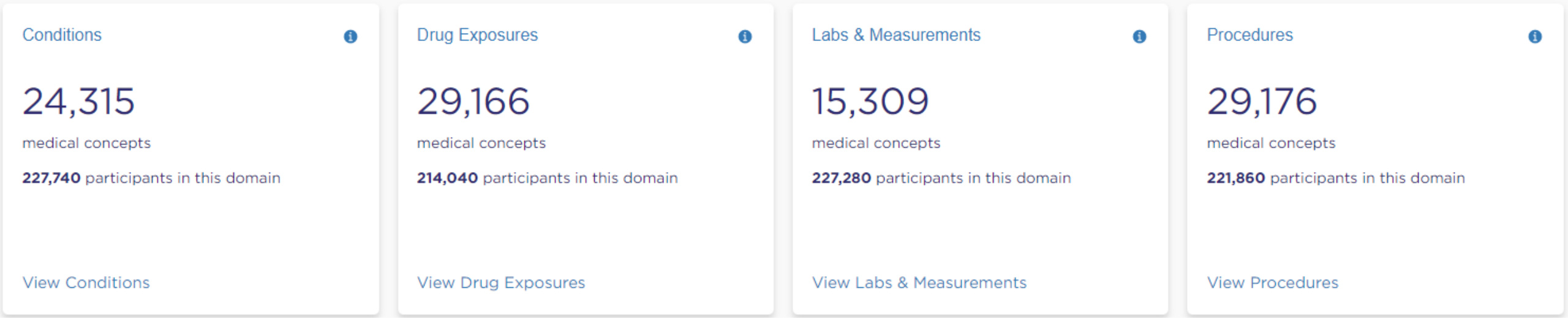
519,000+
Participants

362,000+
Participants who have completed initial steps of the program

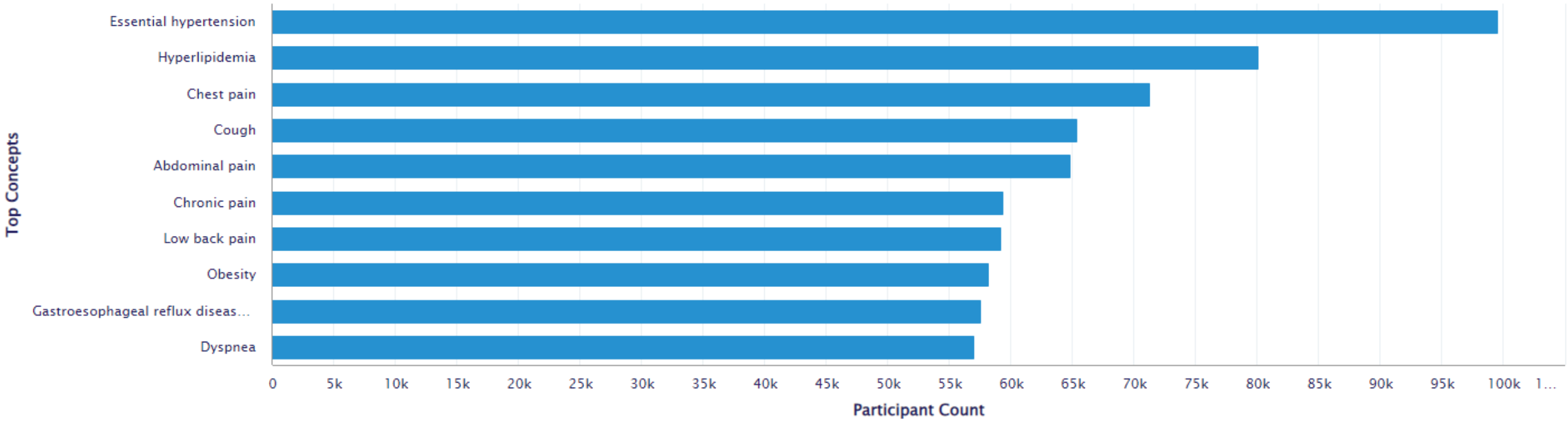
314,000+
Electronic Health Records

383,000+
Biosamples

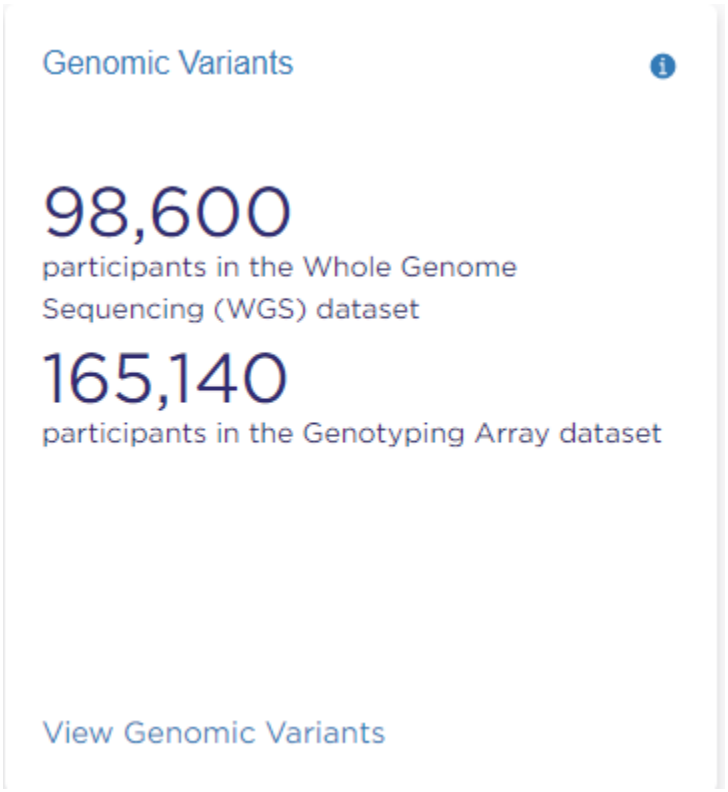
EHR domains


Conditions
24,315
medical concepts
227,740 participants in this domain
View Conditions
Drug Exposures
29,166
medical concepts
214,040 participants in this domain
View Drug Exposures
Labs & Measurements
15,309
medical concepts
227,280 participants in this domain
View Labs & Measurements
Procedures
29,176
medical concepts
221,860 participants in this domain
View Procedures

Top 10 Conditions by Descending Participant Counts



Genomics


Genomic Variants
98,600
participants in the Whole Genome Sequencing (WGS) dataset
165,140
participants in the Genotyping Array dataset
View Genomic Variants

Genomic Variants

Variant Search | Participant Demographics

Use the Variant Search to explore allele frequencies for a gene or genomic region. Drill down into specific variants to view select annotations and genetic ancestry associations.

BRCA2

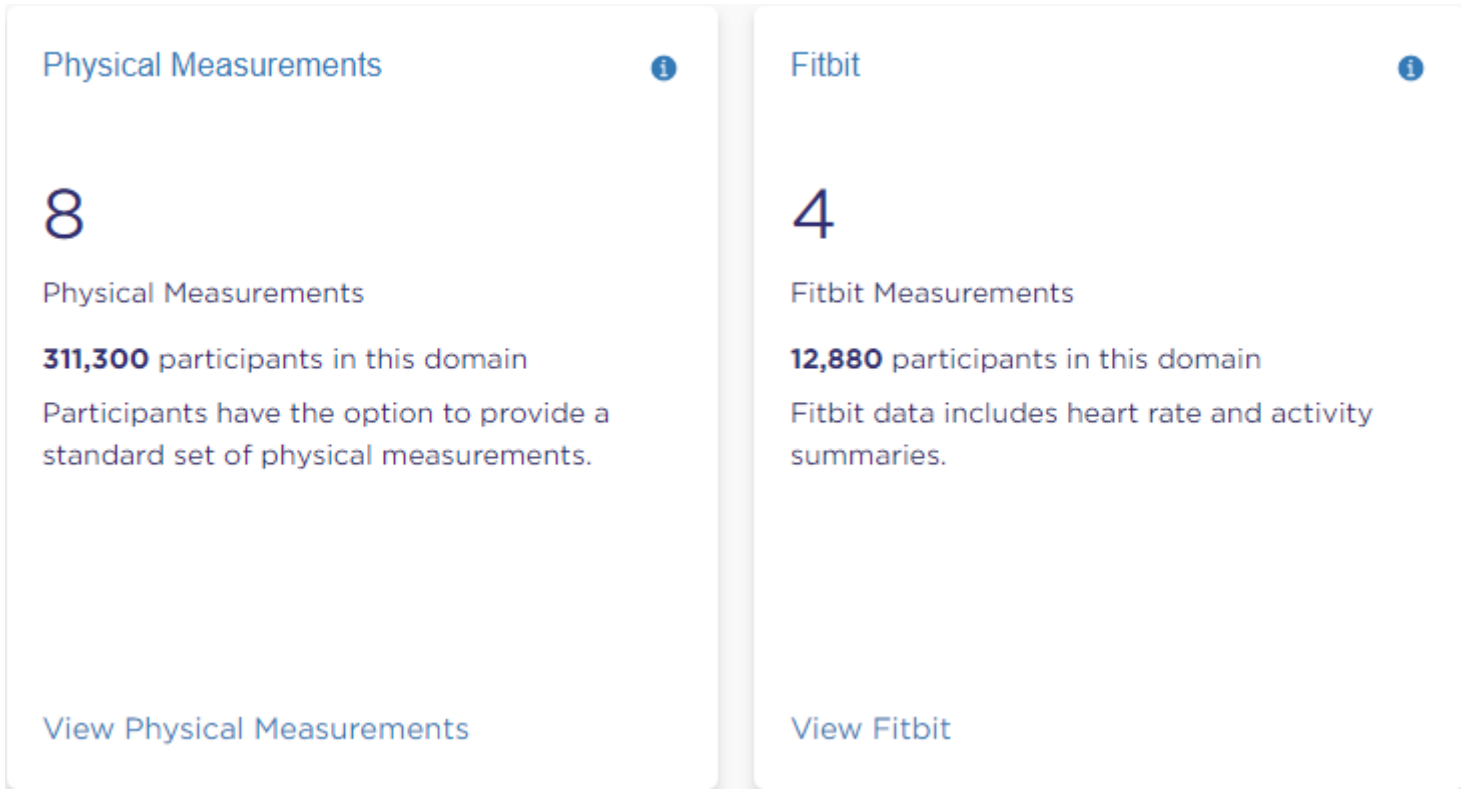
Examples:
Gene: BRCA2, **Variant:** 13-32355250-T-C,
Genomic Region: chr13:32355000-32375000

18,726 variants found

Variant ID	Gene	Consequence	Protein Change	Clinical Significance	Allele Count	Allele Number	Allele Fre
13-32310497-C-T	BRCA2, ZAR1L	intron_variant	-	-	1	197178	0.00000
13-32310503-G-A	BRCA2, ZAR1L	intron_variant	-	-	1	197180	0.00000
13-32310510-T-C	BRCA2, ZAR1L	intron_variant	-	-	262	197178	0.001329
13-32310512-G-A	BRCA2, ZAR1L	intron_variant	-	-	1	197174	0.00000
13-32310514-A-T	BRCA2, ZAR1L	intron_variant	-	-	1	197142	0.00000
13-32310515-C-T	BRCA2, ZAR1L	intron_variant	-	-	7	197170	0.00003
13-32310516-G-A	BRCA2, ZAR1L	intron_variant	-	-	3	197170	0.00001
13-32310516-G-T	BRCA2, ZAR1L	intron_variant	-	-	120243	197170	0.60984
13-32310525-C-A	BRCA2, ZAR1L	intron_variant	-	-	4	197178	0.00002
13-32310527-C-T	BRCA2, ZAR1L	intron_variant	-	-	259	197180	0.001314

Showing at a time 10

Physical Measurements and Wearables



Survey Questions

The Basics

28

questions available

372,380 participants in this domain

This survey includes participant demographic information.

View Complete Survey

Overall Health

21

questions available

372,380 participants in this domain

Survey includes information about how participants report levels of individual health.

View Complete Survey

Lifestyle

26

questions available

372,380 participants in this domain

Survey includes information on participant smoking, alcohol and recreational drug use.

View Complete Survey

Personal Medical History

465

questions available

142,100 participants in this domain

This survey includes information about past medical history, including medical conditions and approximate age of diagnosis.

View Complete Survey

Health Care Access & Utilization

57

questions available

160,880 participants in this domain

Survey includes information about a participant's access to and use of health care.

View Complete Survey

Family Health History

104

questions available

145,620 participants in this domain

Survey includes information about the medical history of a participant's immediate biological family members.

View Complete Survey

COVID-19 Participant Experience (COPE)

191

questions available

105,940 participants in this domain

Survey includes information about the impact of COVID-19 on participant mental and physical health.

View Complete Survey

Minute Survey on COVID-19 Vaccines

141

questions available

101,440 participants in this domain

Survey includes information regarding a participant's COVID-19 vaccination experience.

View Complete Survey

Potential benefits of the Precision Medicine Initiative:

- New approaches for protecting research participants, particularly patients' privacy and the confidentiality of their data.
- Design of new tools for building, analyzing, and sharing large sets of medical data.
- Improvement of FDA oversight of tests, drugs, and other technologies to support innovation while ensuring that these products are safe and effective.
- New partnerships of scientists in a wide range of specialties, as well as people from the patient advocacy community, universities, pharmaceutical companies, and others.
- Opportunity for a million people to contribute to the advancement of scientific research.

Potential long-term benefits of research in precision medicine:

- Wider ability of doctors to use patients' genetic and other molecular information as part of routine medical care.
- Improved ability to predict which treatments will work best for specific patients.
- Better understanding of the underlying mechanisms by which various diseases occur.
- Improved approaches to preventing, diagnosing, and treating a wide range of diseases.
- Better integration of electronic health records (EHRs) in patient care, which will allow doctors and researchers to access medical data more easily.

What are some of the challenges facing precision medicine and the Precision Medicine Initiative?

- Precision medicine is a growing field.
- Many of the technologies that are needed to meet the goals of the Precision Medicine Initiative have only recently been developed.
  - For example, researchers needed to standardize the collection of clinic and hospital data from more than 1 million volunteers around the country. They also needed databases to store large amounts of patient data efficiently.
- The Precision Medicine Initiative also raises ethical, social, and legal issues. It is critical to protect participants' privacy and the confidentiality of their personal and health information. Participants need to understand the risks and benefits of participating in research, which means researchers must have a rigorous process of informed consent.
- Cost is also an issue with precision medicine. The Precision Medicine Initiative itself will cost many millions of dollars in federal funding, and the ongoing initiative will require Congress to approve funding over multiple years. Technologies such as sequencing large amounts of DNA are expensive to carry out (although the cost of sequencing is decreasing). Additionally, drugs that are developed to treat conditions based on molecular or genetic variations are likely to be expensive. Reimbursement from third-party payers (such as private insurance companies) for these targeted drugs is also likely to become an issue.
- If precision medicine approaches are to become part of routine healthcare, doctors and other healthcare providers will need to know more about molecular genetics and biochemistry. They will increasingly need to interpret the results of genetic tests, understand how that information is relevant to treatment or prevention approaches, and convey this knowledge to patients.

Precision Medicine

- Step 1: Gather Data
  - Before crafting a research proposal, investigators can learn what data is available, how to request its use and how to calculate a potential study's risk to patient privacy.
- Step 2: Store and Analyze Data
  - Researchers should save, access and analyze their data remotely in the SAFE. Combining a virtual desktop with a storage network, the SAFE allows researchers to retrieve data from any mobile device or computer using a secure wireless connection and JHED ID. The SAFE is considered the best storage location for research studies as well as the best place to review them.
  - The platform should be approved collaboration and statistical analysis tools, including:
    - Microsoft Office Word and Excel
    - SAS/STAT – a common statistical package that includes more than 100 prewritten analysis procedures
    - STATA – an analytics package that allows researchers to manipulate and visualize statistics and easily produce reports
    - R and R Studio – a software programming language and a software environment that investigators can use for statistical computing and graphics
    - Python – a high-level, general-purpose programming language that is easy for novice and experienced coders to use
- Step 3: Bring to Clinic
  - For precision medicine research to realize its potential, investigators need to convert data insights into clinical applications that can help tailor care plans for patients and predict treatment outcomes.

Precision Medicine in Iran

TCGS

Included Studies

Aim: follow-up major incident & prevalence of cardio-metabolic-related health events

Traits: More than 1000 phenotypes in 22 years and 7 phases

TLGS

Tehran Lipid and Glucose Study

Cohort study

Start: 1999, Ongoing

N: 23265 (Female 49.6%)

Pedigrees: 3029 (Max:56)

N: 19124
Chip:13346
(WGS: 1023)

Aim: Evaluating and comparing several surgical bariatric procedures

Traits: More than 300 phenotypes in 9 years

TOTS

Tehran Obesity Treatment Study

Cohort

Start: 2013, Ongoing

N: 5500 (Female 79.7%)

Pedigrees: 92 (Max:8)

N:469
Chip: 457

Aim: Genetic investigation of Thyroid cancer

Traits: More than 10 phenotypes in 20 years

TCP

Thyroid Cancer Project

Clinical registry

Start: 2001, Ongoing

N: 452 (Female 56.9%)

Pedigrees: 99 (Max:16)

N: 452
Chip: 416

Aim: Genetic investigation of inherited genetic disease

Traits: More than 50 phenotypes in 11 years

CFP

Clinical Follow-up Project

Clinical registry

Start: 2011, Ongoing

N: 328 (Female 51.8%)

Pedigrees: 100 (Max:6)

N:328
Chip: 320
(WGS: 258)

Tehran Cardiometabolic genetic study (TCGS)

Aim: within a longitudinal family-based cohort in response to the lack of fundamental knowledge of the genetic variation diversity pattern in the Iranian population, concentrating on evaluating the genetic basis of Cardiometabolic risk factors.

Data: N:20373 (Female 51.8%); Pedigree: 4452 (Max:56), Chip:14539; WGS: 1281; Genetic Markers (SNP, Indel): ~62 M

Variables: Physical examination, Biochemical markers, Medical history

Start: 2012; Ongoing

مطالعه قند و لیپید تهران

Agreement

2013
Collaboration

ChipTyping

2014

2015

2016

2017

CFP

2th Agreement

2018
Whole Genome Analysis

TOTS

2019

TCP

2020
9

2021
13

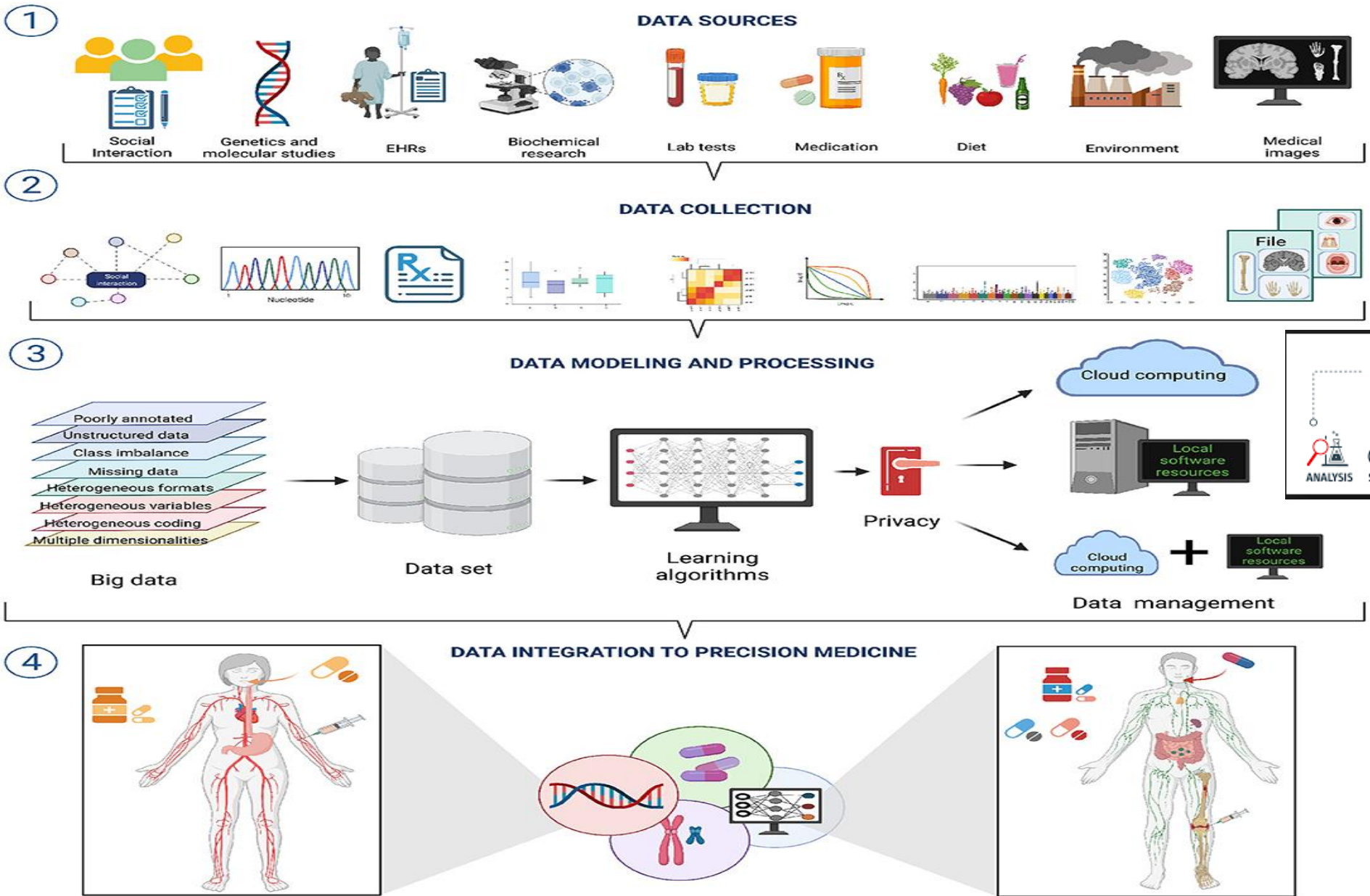
2022
18

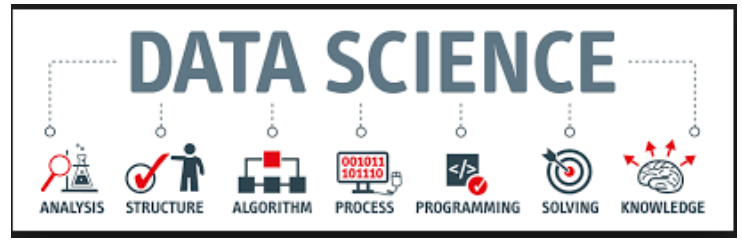
Baseline (1999-2001) → Follow-up (2002-2005) → Follow-up (2006-2008) → Follow-up (2009-2011) → Follow-up (2012-2014) → Follow-up (2015-2017) → Follow-up (2018-2022)

Population diversity

*Region	Province	Population	Total Ethnicity	Chip data Ethnicity
1	Alborz	2712400	**Total N: 2255** **Arab(2), Arab-Persian(5), Gilak(114), Kurd(1), Lur-Persian(1), Persian(2072), Tat(10), Tat-Turk(28), Turk(12), Turkaman(6), Others(4)**	**Total N:1351** **Arab(2), Arab-Persian(3), Gilak(80), Kurd(1), Lur-Persian(1),Persian(1218), Tat(10), Tat-Turk(18), Turk(11), Turkaman(3), Others(4)**
	Golestan	1868819		
	Mazandaran	3283582		
	Qazvin	1273761		
	Qom	1292283		
	Semnan	702360		
	Tehran	13267637		
2	Bushehr	1163400	**Total N: 1221** **Arab(26), Arab-Persian(2), Lur(26), Persian(1157), Qashqai(3), Turk(2), Others(5)**	**Total N: 743** **Arab(22), Arab-Persian(2), Lur(23), Persian(688), Qashqai(2), Turk(1), Others(5)**
	Chaharmahal and Bakhtiari	947763		
	Fars	4851274		
	Hormozgan	1776415		
	Isfahan	5120850		
3	Ardabil	1270420	**Total N: 810** **Gilak(138), Kurd(18), Lur(1), Mix(1), Persian(9), Persian-Turk(2), Tat(1), Turk(633), Others(7)**	**Total N: 504** **Gilak(65), Kurd(17), Lur(1), Mix(1), Persian(6), Persian-Turk(2), Tat(1), Turk(404), Others(7)**
	East Azerbaijan	3909652		
	Gilan	2530696		
	Kordestan	1603011		
	West Azerbaijan	3265219		
	Zanjan	1057461		
4	Hamadan	1738234	**Total N: 1040** **Arab(50), Gilak(3), Kurd(40), Lur(106), Lur-Lak(17), Lur-Persian(3), Persian(803), Tat-Turk(1), Turk(9), Others(8)**	**Total N: 632** **Arab(23), Gilak(2), Kurd(30), Lur(71), Lur-Lak(10), Lur-Persian(3), Persian(476), Tat-Turk(1), Turk(9), Others(7)**
	Ilam	580158		
	Kermanshah	1952434		
	Khuzestan	4710509		
	Lorestan	1760649		
	Markazi	1429475		
5	Kerman	3164718	**Total N: 458** **Balouch(28), Kurd(10), Persian(414), Tat(2), Turk(1), Turkaman(3)**	**Total N: 291** **Balouch(24), Kurd(7), Persian(254), Tat(2), Turk(1), Turkaman(3)**
	North Khorasan	863092		
	Razavi Khorasan	6434501		
	Sistan and Baluchestan	2775014		
	South Khorasan	768898		
	Yazd	1138533		
Foreigner	Afghanistan		**Total N: 36** **Afghan(3), Arab(6), Arab-Persian(5), Mix(4), Russia(7), Turk(6), Turkaman(5)**	**Total N: 24** **Afghan(3), Arab(4), Arab-Persian(4), Mix(4), Russia(6), Turk(3)**
	Azerbaijan			
	Iraq			
	Russia			
	Turkmenistan			
	Yemen			
	Unknown		**Total N: 357** **Persian(274), Turk(83)**	**Total N: 306** **Persian(234), Turk(72)**

Social Intraction	Genetic Markers	EHRs	Biochemical research	Medication	Diet	Enviornment	Demographic information	Ethnicity
75	62M	962	53	47	221	80	13	7





Be in contact

- Gemiran.org

COHORT PROFLE UPDATE: TEHRAN CARDIOMETABOLIC GENETIC STUDY

Abstract

The Tehran cardiometabolic genetic study (TCGS) is a large population-based cohort study that conducts periodic follow-ups. TCGS has created a comprehensive database comprising 20,367 participants born between 1911 and 2015 selected from four main ongoing studies in a family-based longitudinal framework. The study's primary goal is to identify the potential targets for prevention and intervention for non-communicable diseases that may develop in mid-life and late life. TCGS cohort focuses on cardiovascular, endocrine, metabolic abnormalities, cancers, and some inherited diseases. Since 2017, the TCGS cohort has augmented by encoding all health-related complications, including hospitalization outcomes and self-reports according to ICD11 coding, and verifying consanguineous marriage using genetic markers. This research provides an update on the rationale and design of the study, summarizes its fndings, and outlines the objectives for precision medicine.

Excel File PDF1 Download PDF2 Download

Order	Nature of Data	Data Set Name	Unit	Data Type		Cohort				P/E		Laboratory Blood Test								Tra										
																				Self-Reported Questionary										
				Original	Calculated	Tehran lipid and glucose study (TLGS) (n=1319)	Tehran Obesity Treatment Study (TOTS) (n=164)	Thyroid cancer study (TCP) (n=7)	Clinical Follow-up Project (CFP) (n=18)	Antropometric (n=5)	Ankle-arm Blood Pressure (n=8)	Basic metabolic panel (n=11)	Lipid panel (n=4)	Advanced Lipid Panel (n=4)	Inflammatory panel (n=4)	Thyroid Function Test (n=9)	Liver Function Test (n=5)	Hormones (n=4)	Others (n=4)	Demographic information (n=8)	Ethnicity (n=7)	Past Medical History (n=47)	Adoloscents Smoking (n=15)	Adults Smoking (n=40)	Adolescents Physical Activity (n=8)	Adults Physical Activity (n=17)	Obstetric and Gynecology (n=73)	Dietary intake (n=221)	Follow up (n=58)	Health-related quality of life (HRQoL) (n=75)
1	Age based on valid birth date	Age	Years	*		*	*													*										
2	Documented birth date	Date of birth	Y/M/D	*		*	*													*										
3	Gender	Sex	male/female	*		*	*													*										
4	Educational certificate, level	Education	Index	*		*	*													*										
5	Family relationship with spouse and situation in family	Familial relationship	Index	*		*	*													*										
6	Marital status	Marital status	YES / NO	*		*	*													*										
7	Job group and Occupational status past month	Occupation	Index	*		*	*													*										
8	Total and continuous residence in Tehran	Residency period	Years	*		*														*										
9	Report angiography	angiores	YES / NO	*		*																*								
10	Ace inhibitors last month	iface	YES / NO	*		*																*								
11	Angiography history	ifangiog	YES / NO	*		*	*															*								
12	Angioplasty history	ifangiop	YES / NO	*		*	*															*								
13	Report angiography/We do not have this variable	ifangiores/We do not	YES / NO	*		*																								

Order	1-Certain infectious or parasitic diseases (n= 47)		2-Neoplasms (n= 68)		3-Diseases of the blood or blood forming organs (n= 14)		4-Diseases of the immune system (n= 11)		5-Endocrine nutritional or metabolic diseases (n= 49)		6-Mental behavioural or neurodevelopmental disorders (n= 25)		7-Sleep wake disorders (n= 2)		8-Diseases of the nervous system (n= 34)	
	S.R	H.	S.R	H.	S.R	H.	S.R	H.	S.R	H.	S.R	H.	S.R	H.	S.R	H.
1																
2																
3																
4																
5																
6																
7																
8																
9																
10																
11																
12																
13																