

The Role of AI

Algorithm to bedside symposia

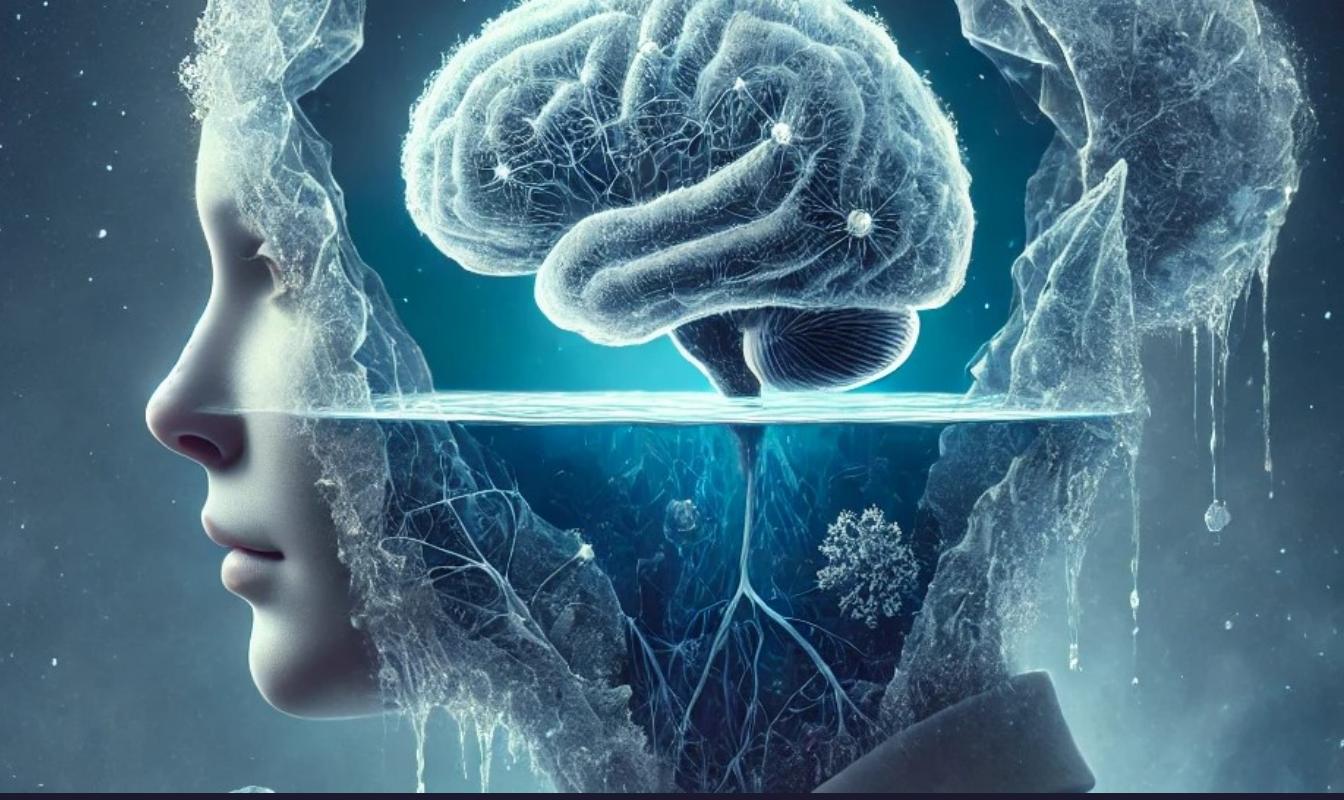
>>>



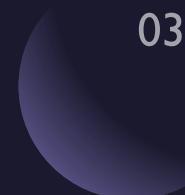


DATA STORIES

From algorithm to bedside



Nermin Ghith, dual B.Sc. in bio/chem.; MPH, PhD
Senior Forsker
Gentofte or Herlev Hospital
03-14 Feb. 2025



Lunch symposia

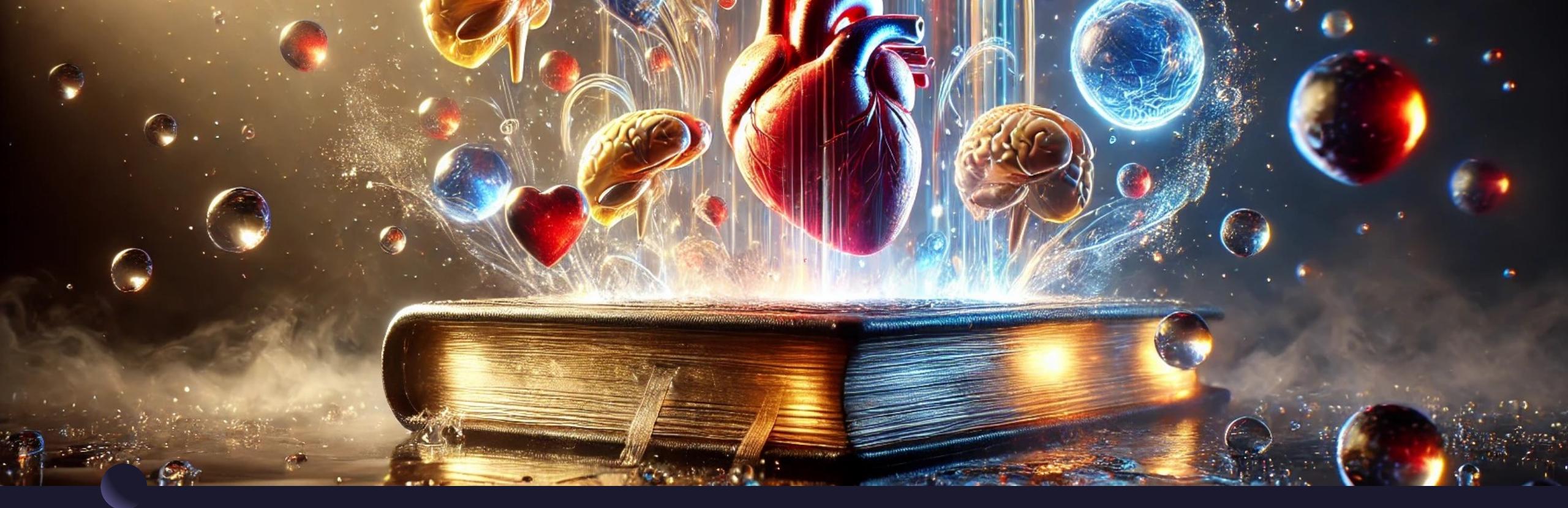
Omics-Based Clinical and Population Studies (session 1:
Mon. 3 Feb. 2025 - Specific Case Studies)

Contextual Epidemiology of Cardiometabolic
Conditions, Polypharmacy and Multi-morbidity
(session 2: Thurs. 6 Feb. 2025)

Scientific Methods in Biomedical Research (sessions 3
and 4: Mon and Thurs. 10 and 13 Feb. 2025)

- Omics in biomedical research
- Evidence and analytics
- Global Burden of Disease, Health Analytics, and Access to Therapeutics
- Economic Evaluations





Session 3

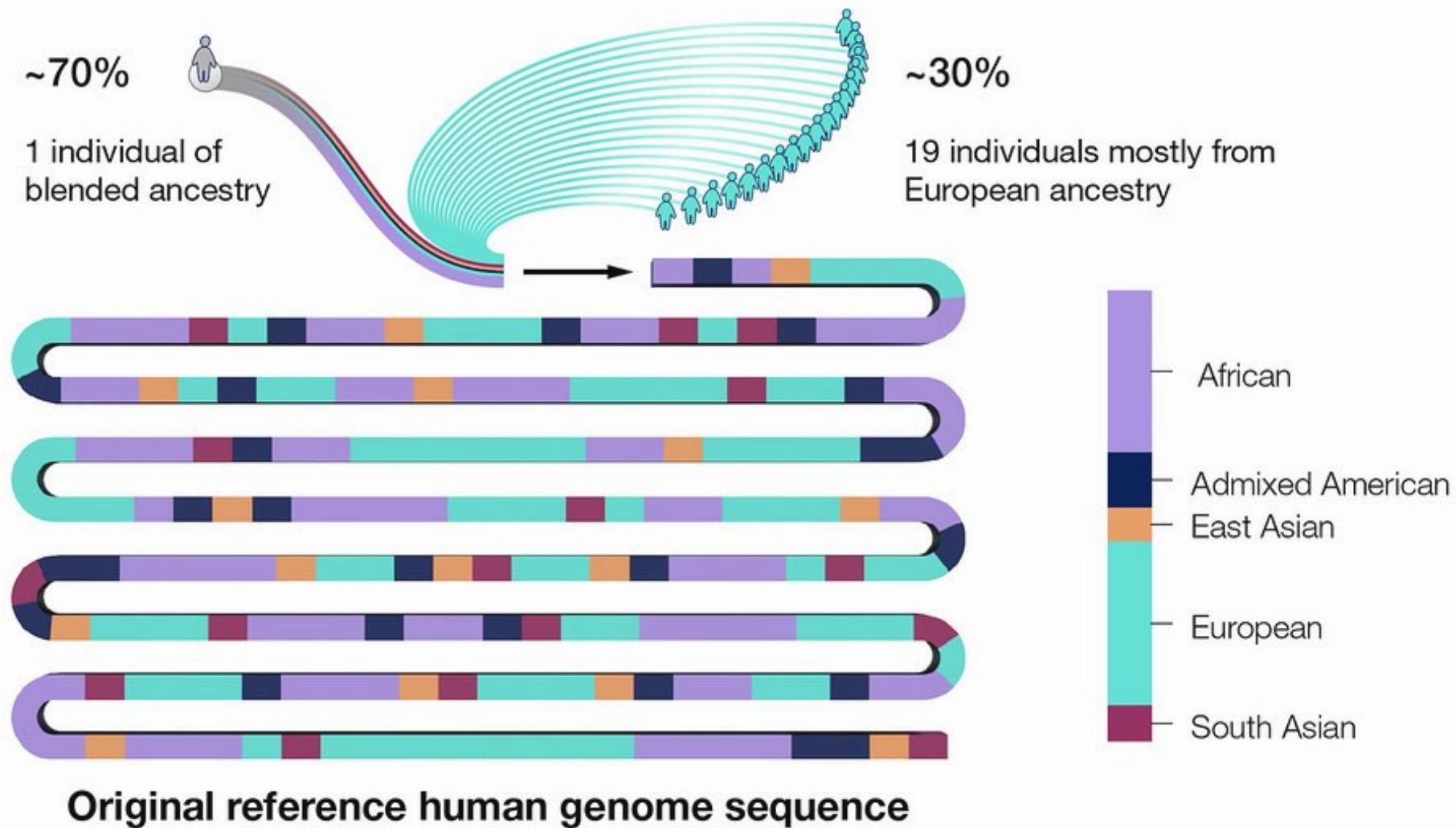
- Omics in biomedical research – **AI-based dry labs.**
- Evidence pyramid and analytics – **clinical vs. AI/omics-based studies**

From algorithm to bedside

Omics
pipelines

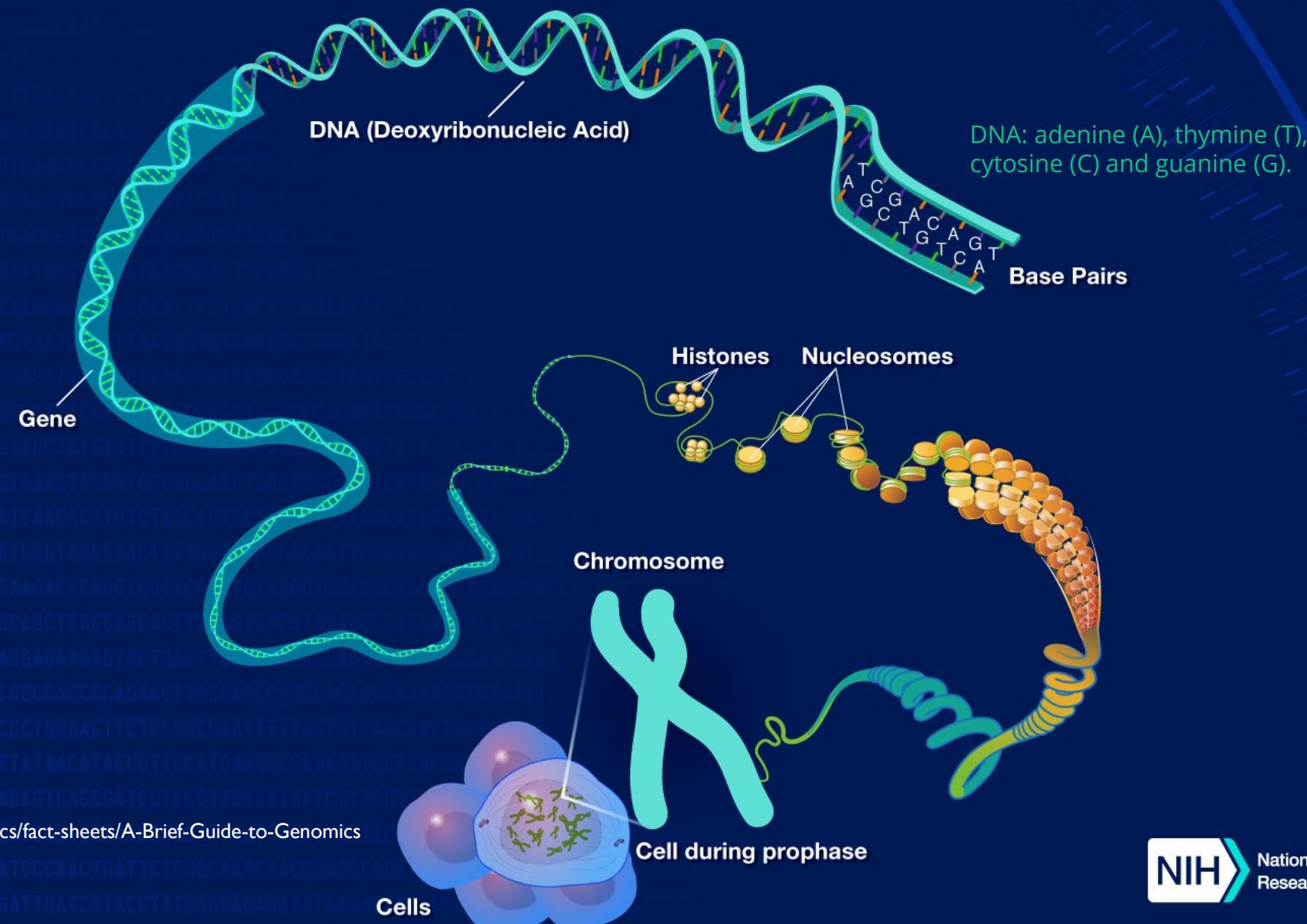


Human genome and diseases



A Brief Guide to Genomics

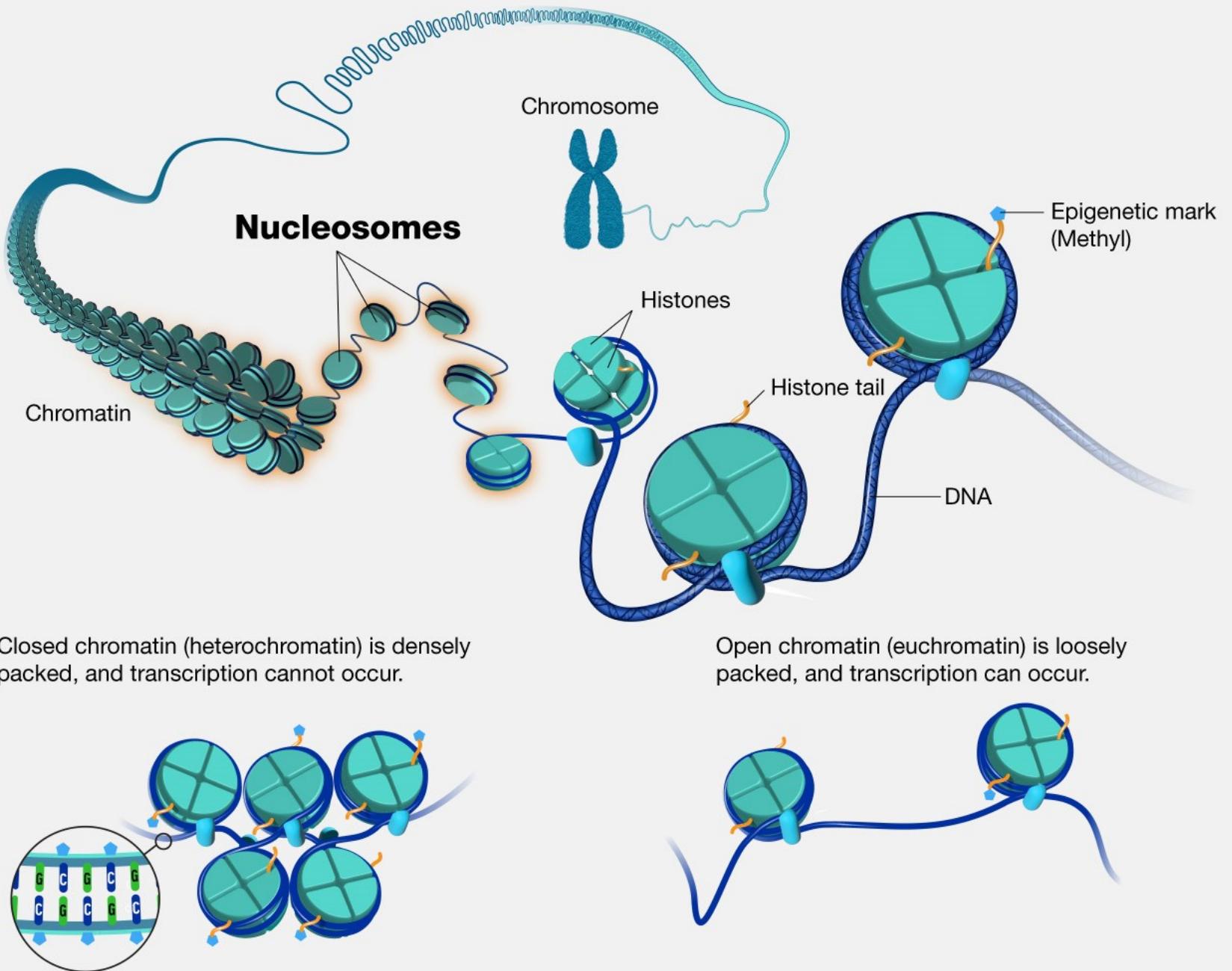
NHGRI FACT SHEETS
genome.gov



Source: <https://www.genome.gov/about-genomics/fact-sheets/A-Brief-Guide-to-Genomics>

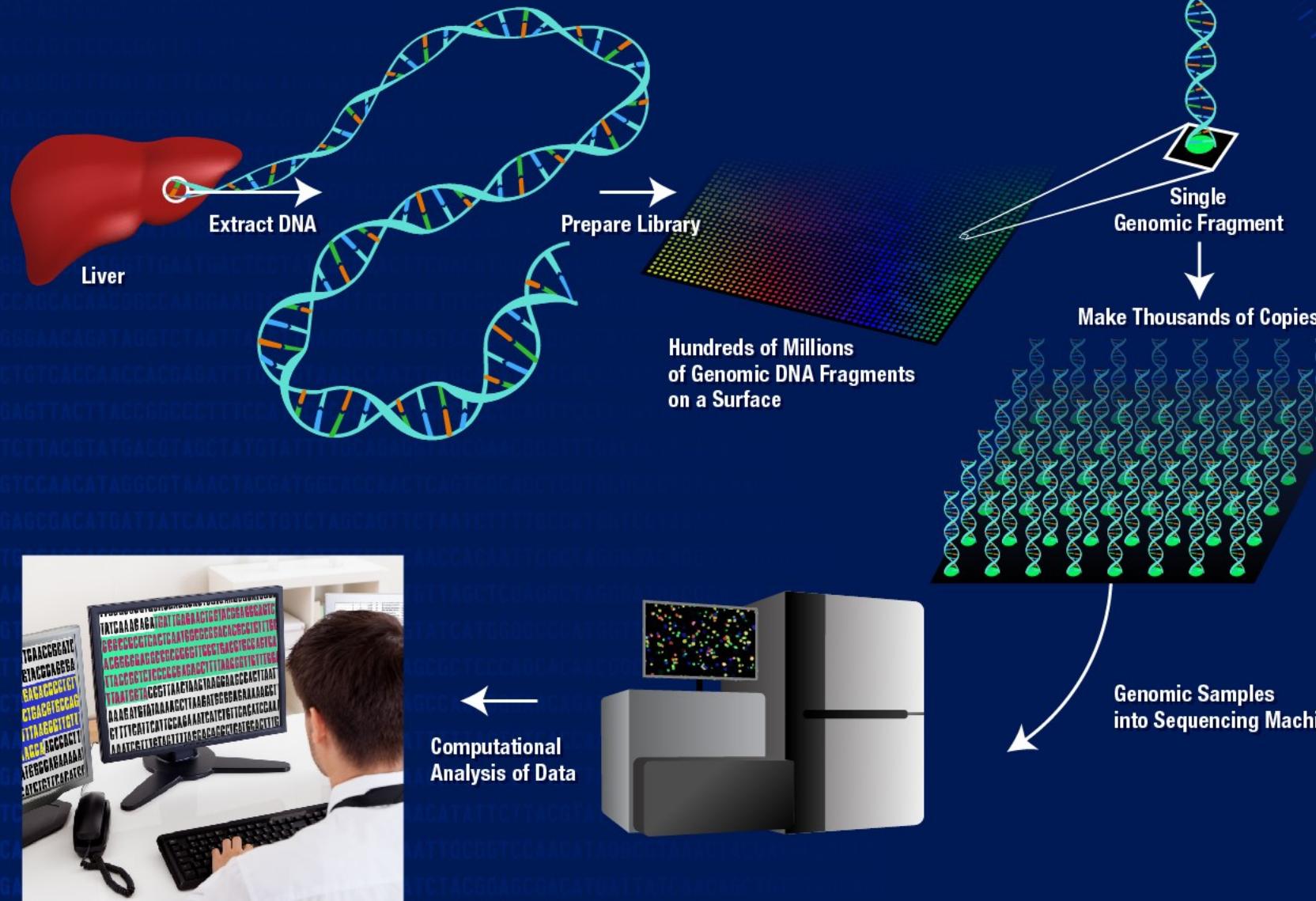


National Human Genome Research Institute



DNA Sequencing

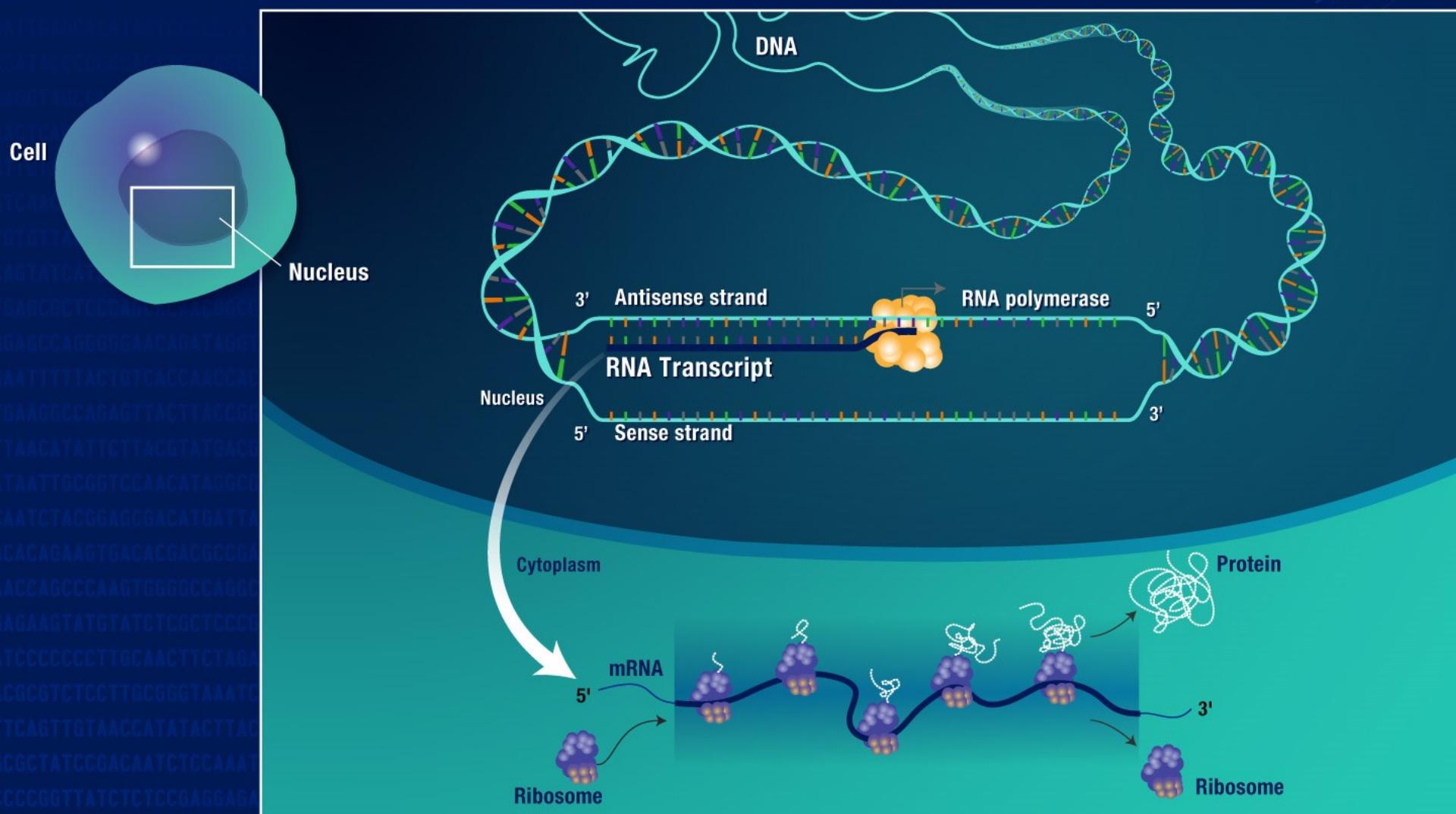
NHGRI FACT SHEETS
genome.gov



National Human Genome
Research Institute

Transcriptome

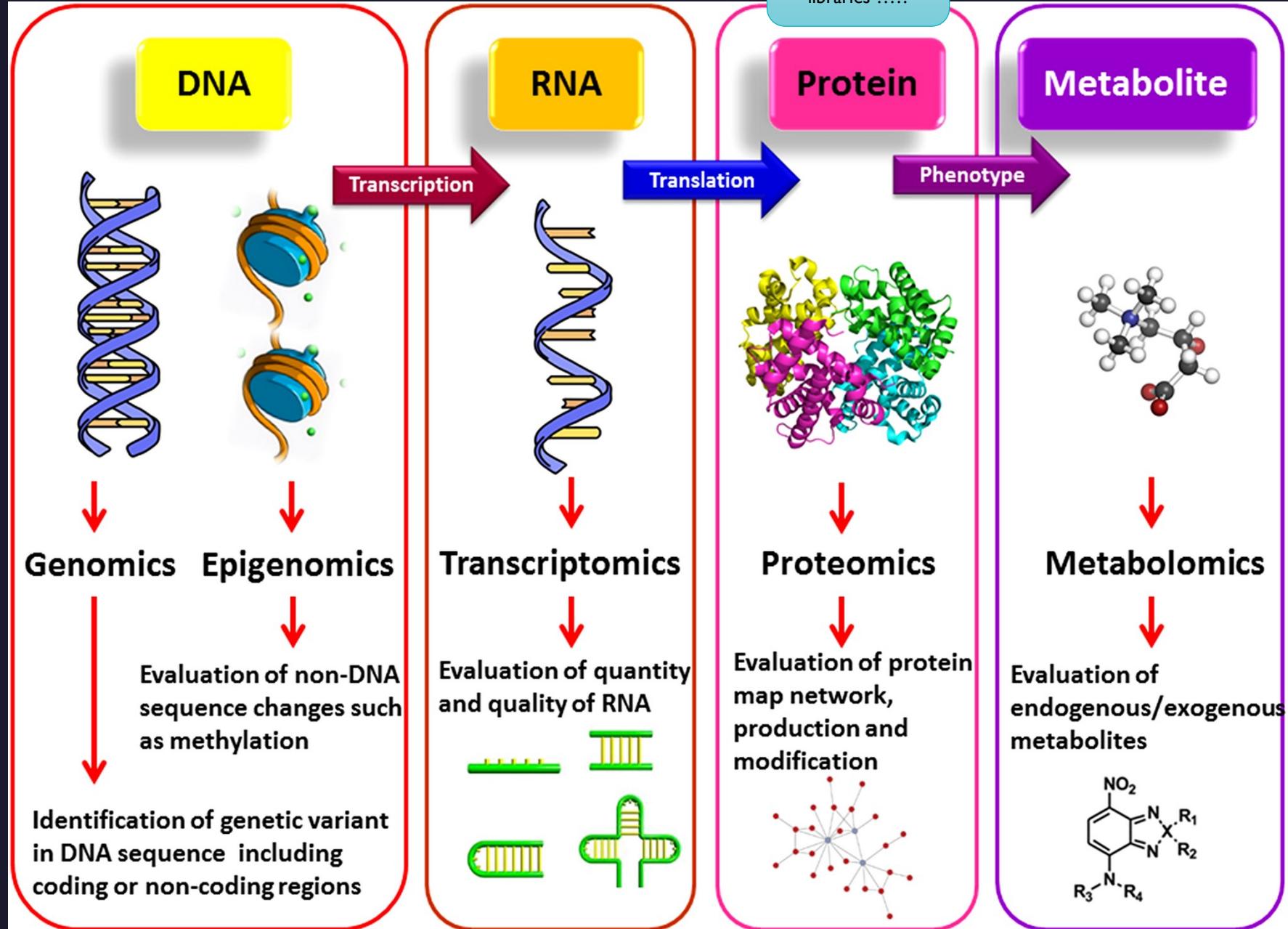
NHGRI FACT SHEETS
genome.gov



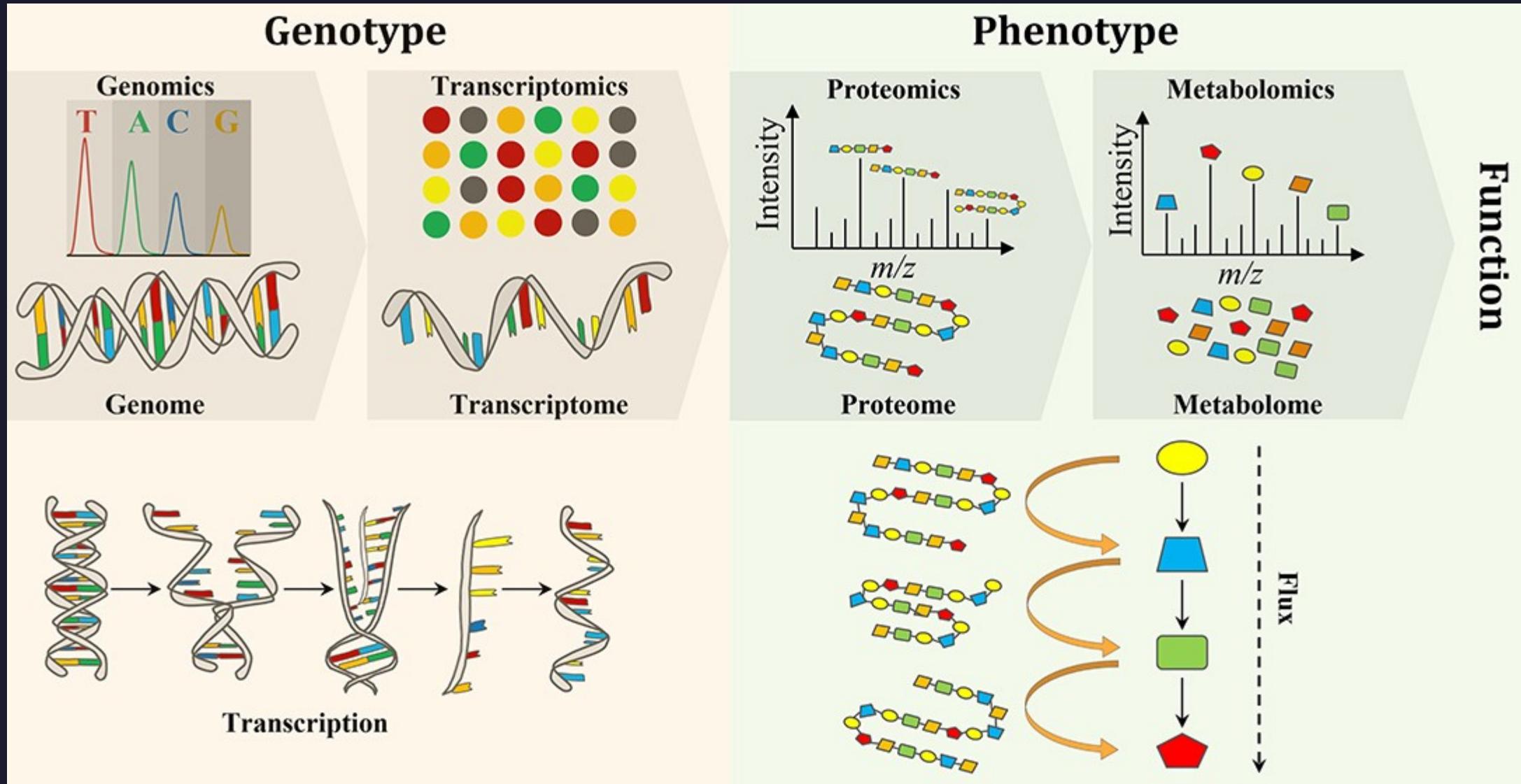
<https://www.genome.gov/about-genomics/fact-sheets/Transcriptome-Fact-Sheet>



National Human Genome
Research Institute



Function



<https://doi.org/10.3389/fbioe.2021.613307>



Omics

NIH definitions



Genomics - The genome is the complete sequence of DNA in a **cell or organism**.

Transcriptomics - The transcriptome is the complete set of RNA transcripts from DNA in a **cell or tissue**.

Proteomics - The proteome is the complete set of proteins expressed by a **cell, tissue, or organism**.

Epigenomics - The epigenome consists of **reversible** chemical modifications to the DNA, or to the histones that bind DNA, and produce changes in the **expression of genes** without altering their base sequence.

Metabolomics - The metabolome is the complete set of small molecule metabolites found within a biological sample (including metabolic intermediates in **carbohydrate, lipid, amino acid, nucleic acid**, and other **biochemical pathways**, along with **hormones** and other **signaling molecules**, as well as exogenous substances such as **drugs** and their **metabolites**).



Genetic mapping

NHGRI FACT SHEETS
genome.gov

Genetic map



Cytogenetic map



Physical map

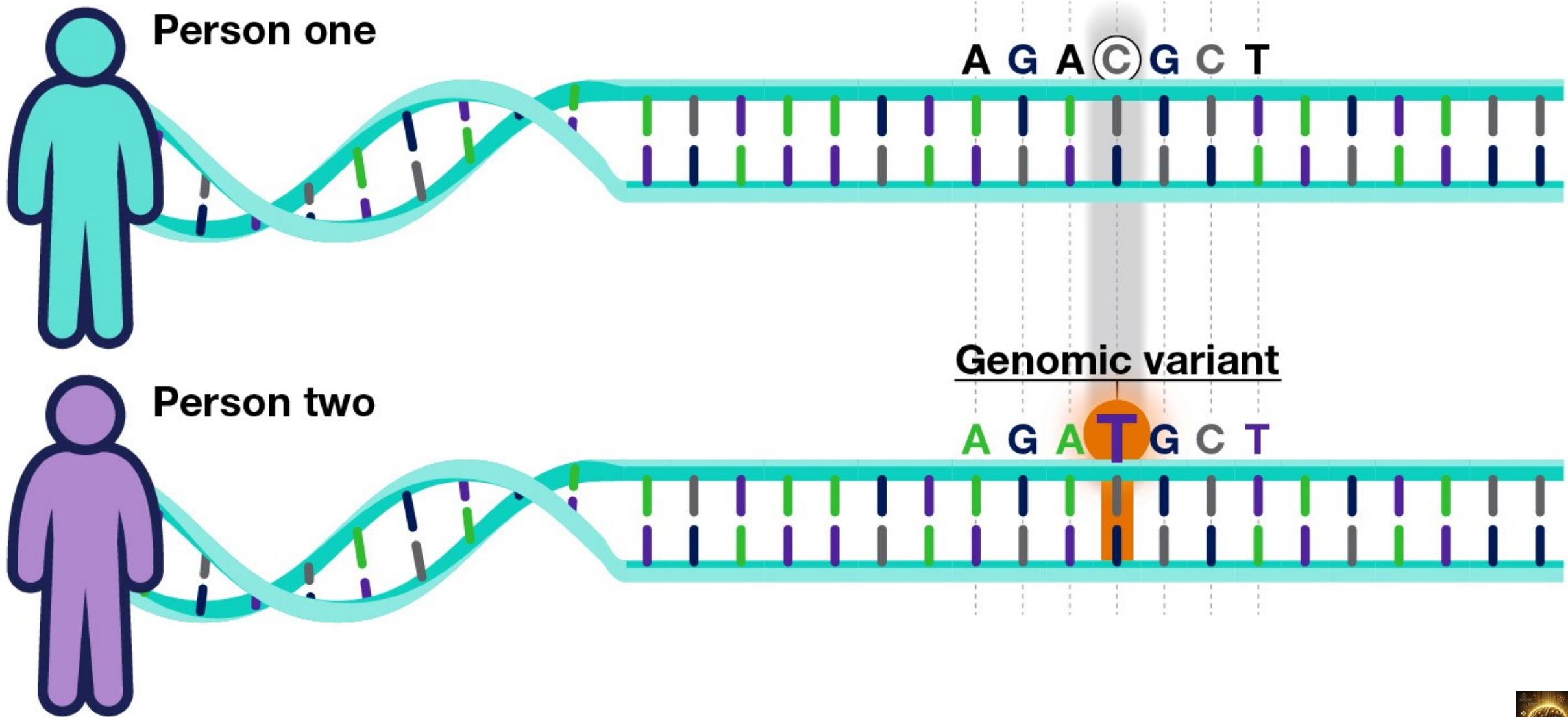


DNA sequence

....GATCTGCATGCATGCTAGCTAGCTAGCTAGAGCTTCG....

Bases

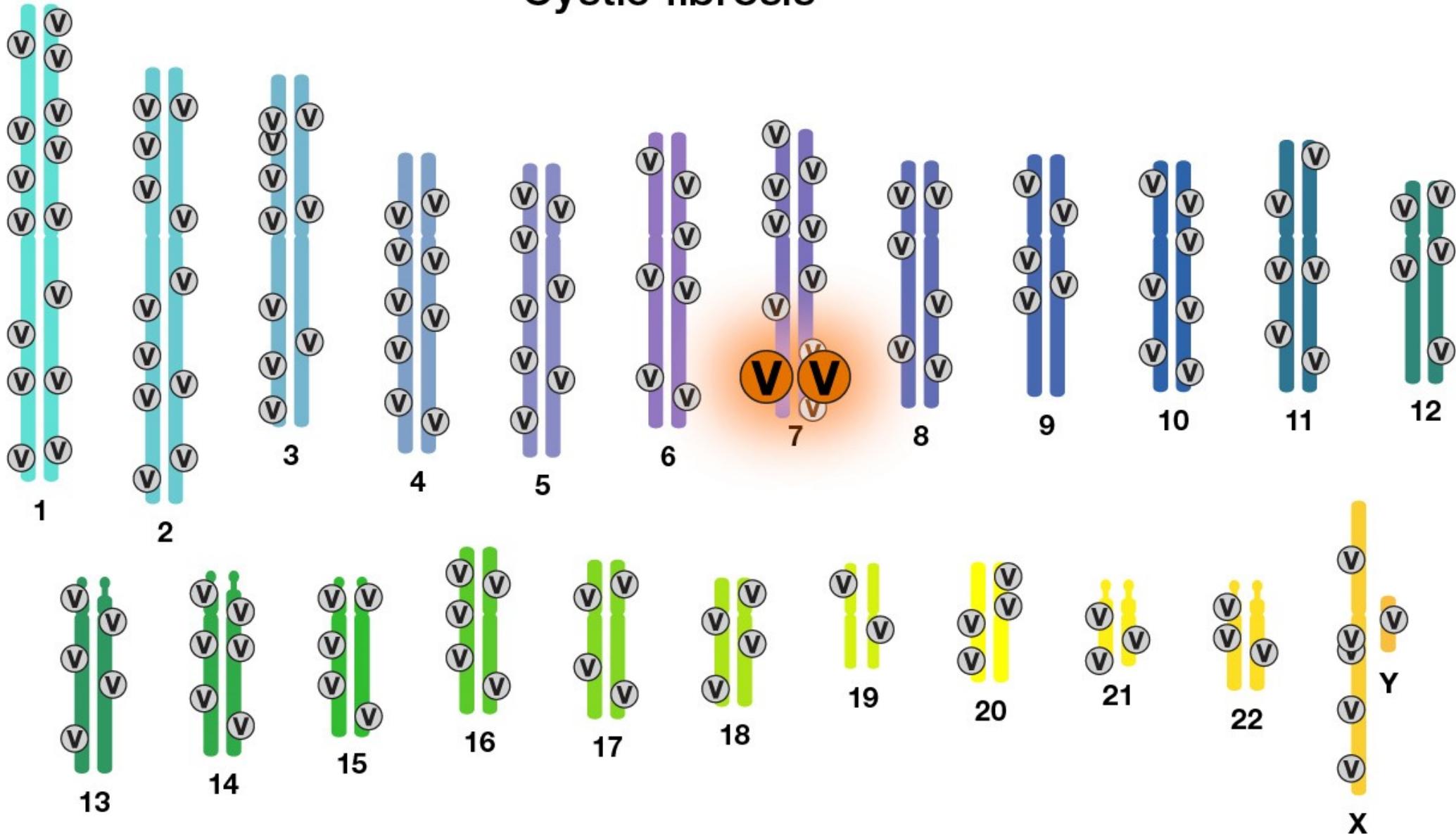
genomic variants and the risk for specific diseases



Single-gene diseases

Cystic fibrosis

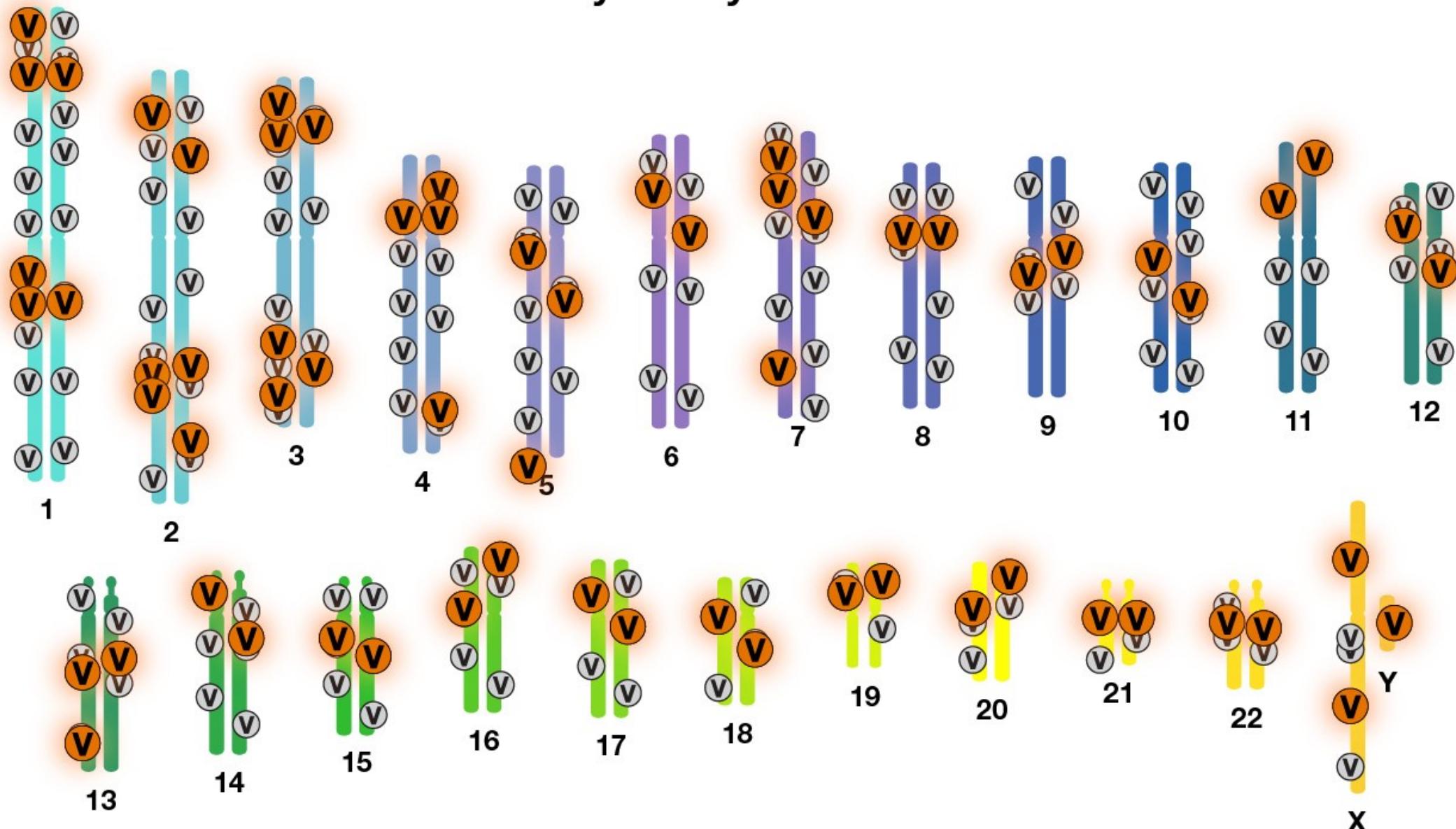
Person
one



Complex diseases - polygenic

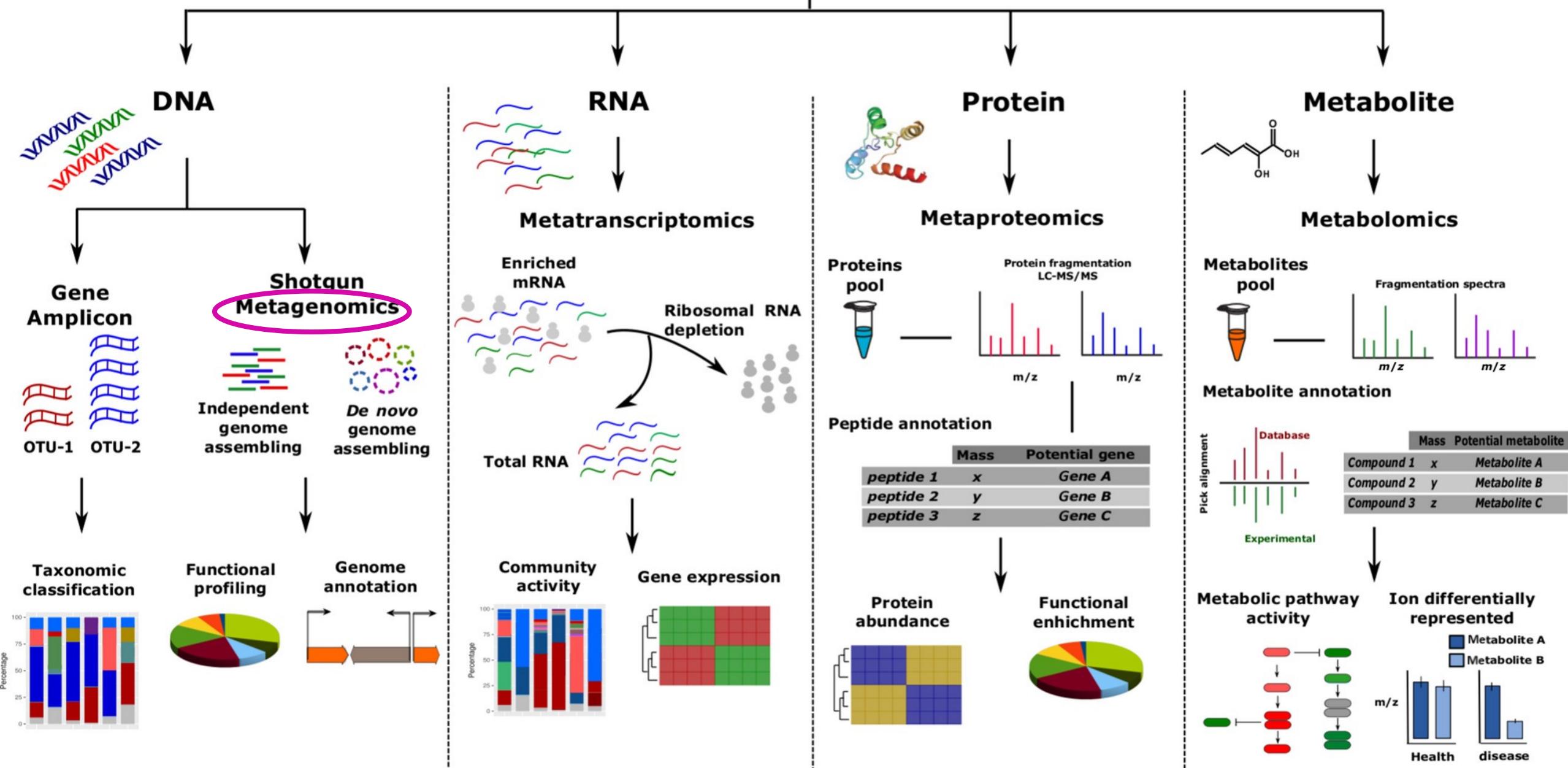
Coronary artery disease

Person
two

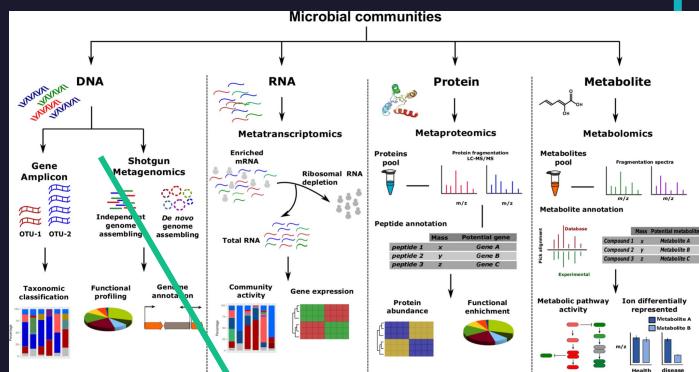


Microbial communities

Reference libraries ?????



Own Project Pipeline



Fragments of DNA

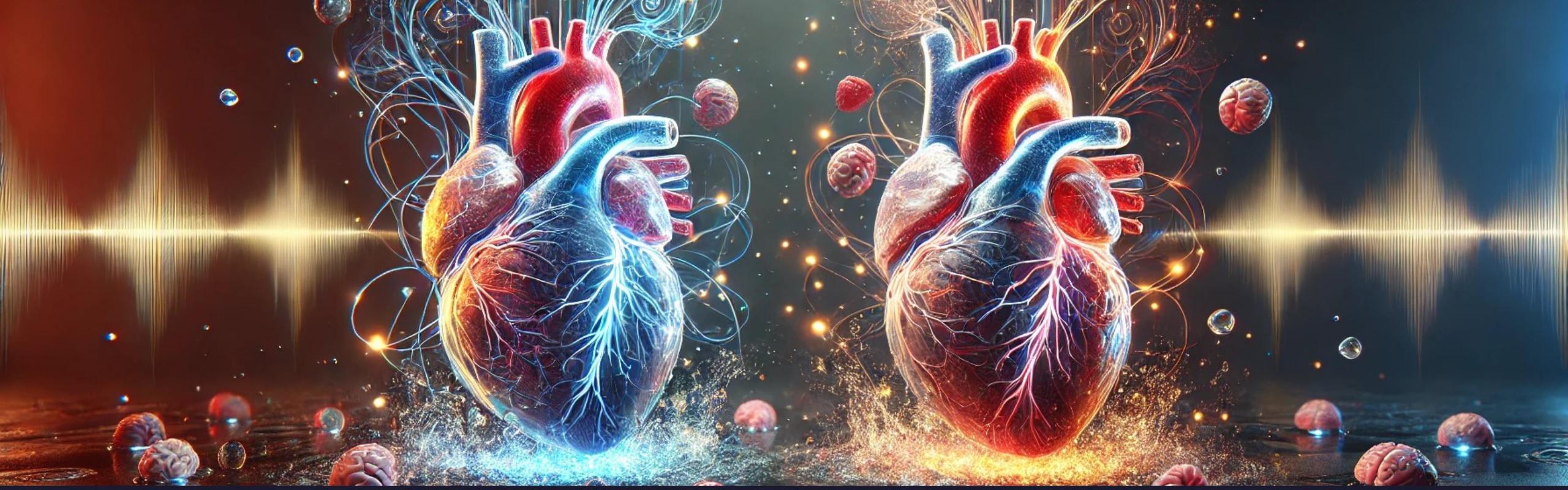
Resistance Genes

19 Classes of Antibiotics

65 Phenotypes of Antibiotics

phenotypic resistance to
different classes of antibiotics

Quantify per sample
(i.e., FPKM)



Omics-based studies

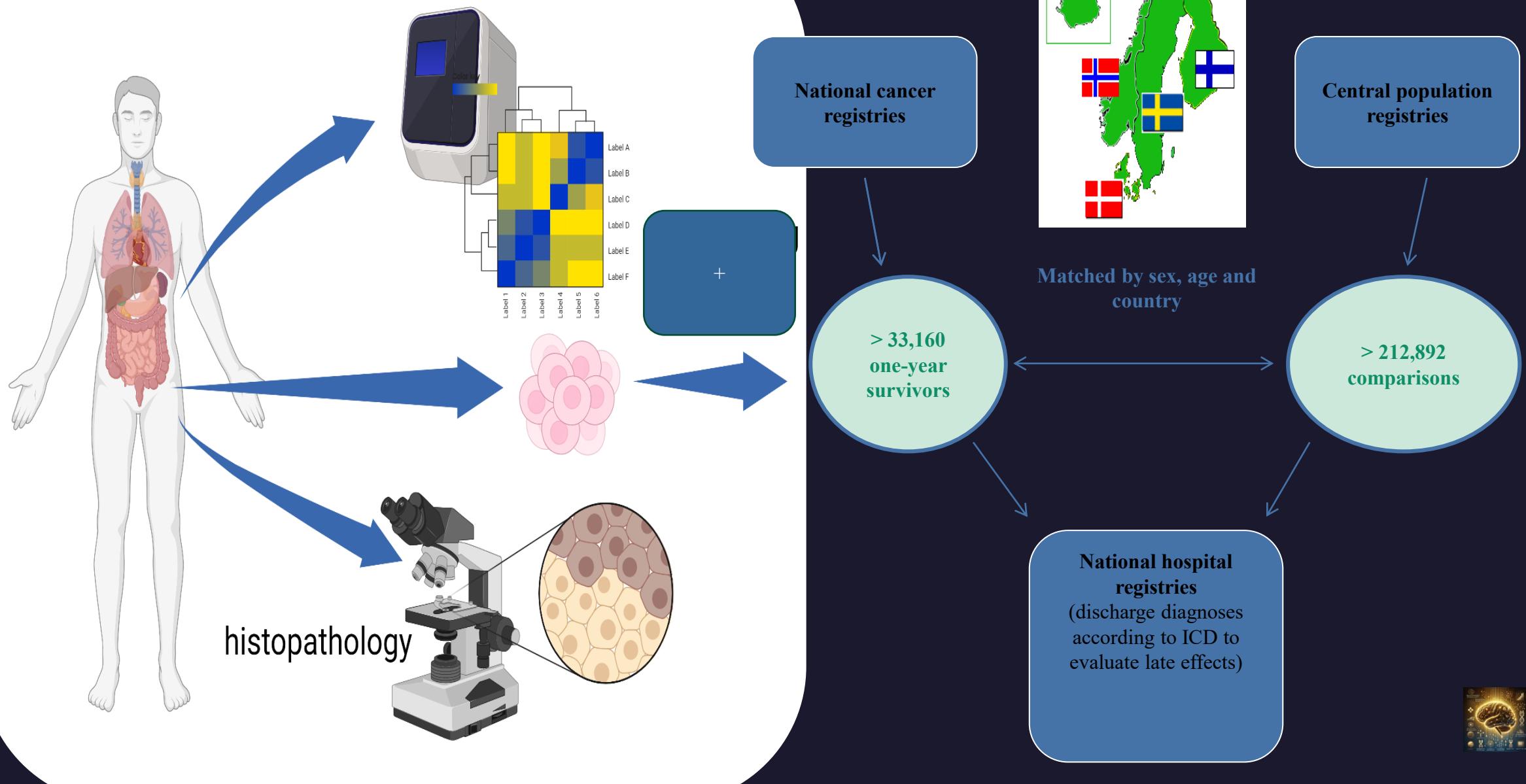
- RCTs?
- Randomization with nested cohorts or combined with an observational omics-based study?
- How can we integrate omics into a traditional RCT?



Integrated (multimodal) Omics

- Data – modeling – classification of health events
 - Health data – models – decisions or/and predictions
- Multidata type – identification of biomarkers – clinical outcomes
- Cancer cohort –intra-tumor heterogeneity – classification of data and/or events - survival estimation – etc
- Many other frames of reference

NGS analysis

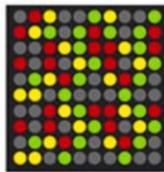


Multi-data type

Genomics



Transcriptomics



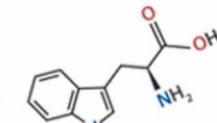
Proteomics



Epigenomics



Metabolomics



e.g., Registry data, EMR



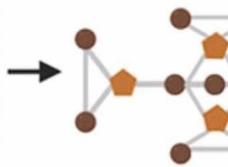
Clinical Metadata

Biomarker Discovery

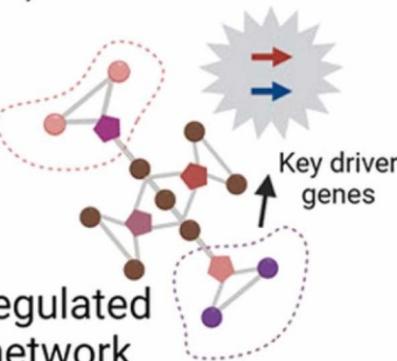
Differential Analysis



Biological Networks
(GCNN, RNN, MLs)

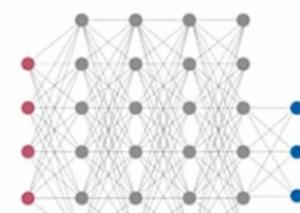


Down-regulated subnetwork

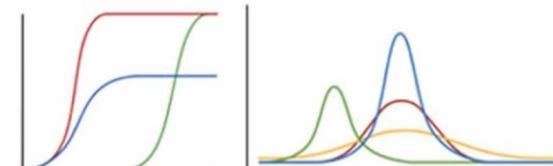


Up - regulated subnetwork

Outcome



Deep Learning

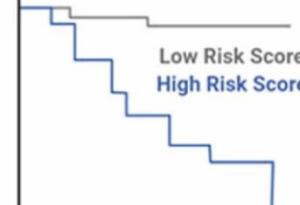
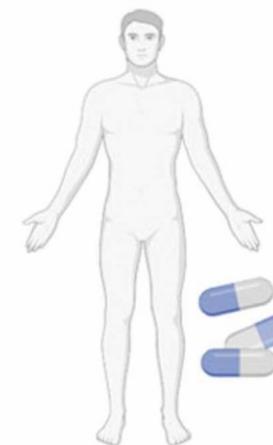


Drug A

Drug X

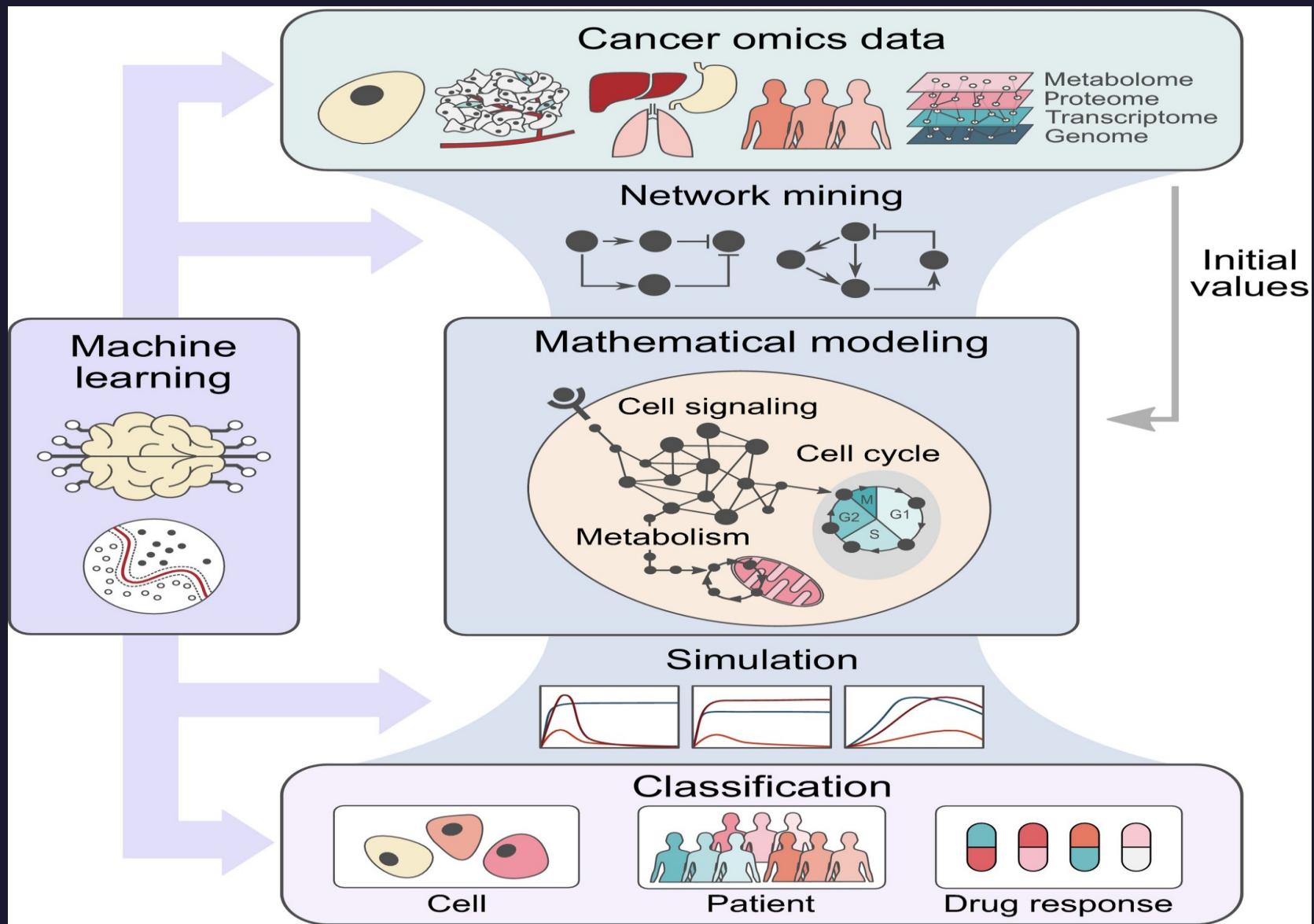
Clinical Outcome

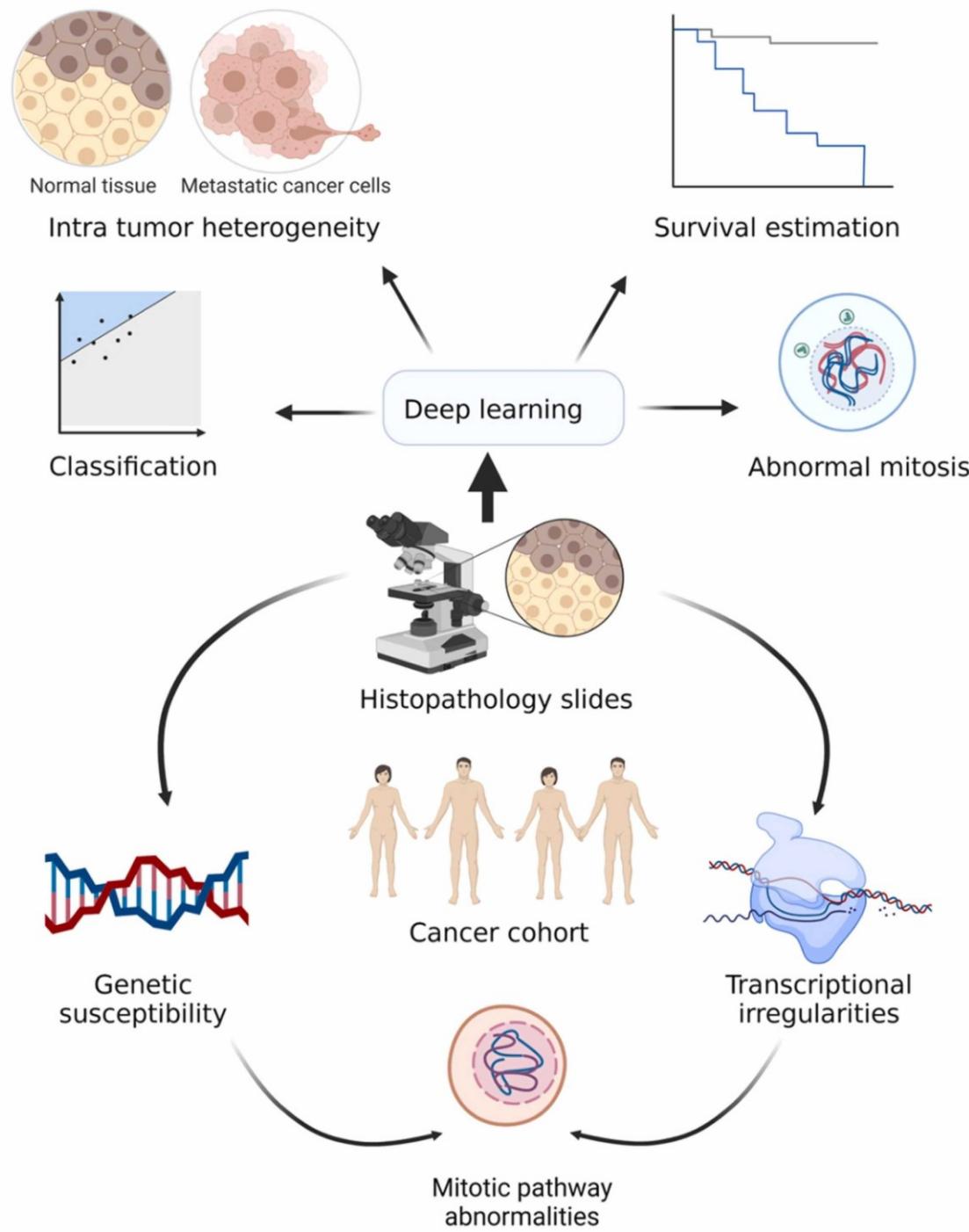
Personalized Medicine



Survival Evaluation







ODE MODEL

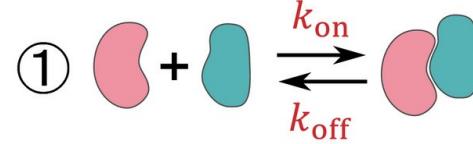
$$\frac{dx}{dt} = S \cdot v(x, \theta), x(0) = x_0$$

Biological network

Build your own model or download from BioModels

 **BioModels**

Kinetic parameters

① 

② 

Initial conditions

mRNA level

Gene	Gene A	Gene B	Gene C	Gene D
mRNA level	Low	Medium	High	Very High

Protein level

Protein	Protein A	Protein B	Protein C	Protein D
Protein level	Medium	Low	High	Medium-High

Estimate from data fitting

Fit the parameters in ODE model to data

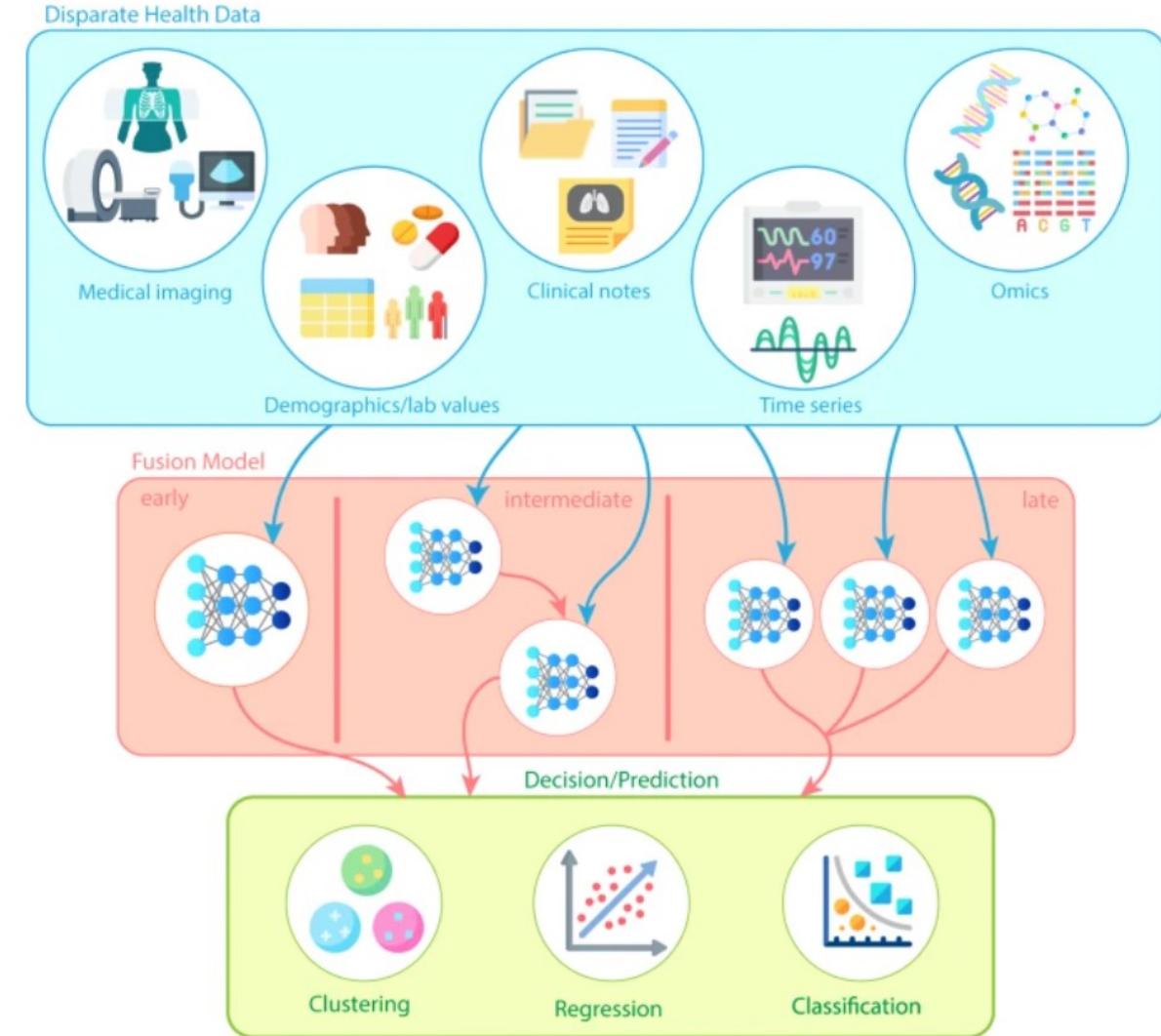
Obtain from transcriptome or proteome data

 **CCLE** Cancer Cell Line Encyclopedia

 **TCGA** THE CANCER GENOME ATLAS



Fig. 2: Early, intermediate, and late fusion; flow of information from information commons to model structure to outcomes.

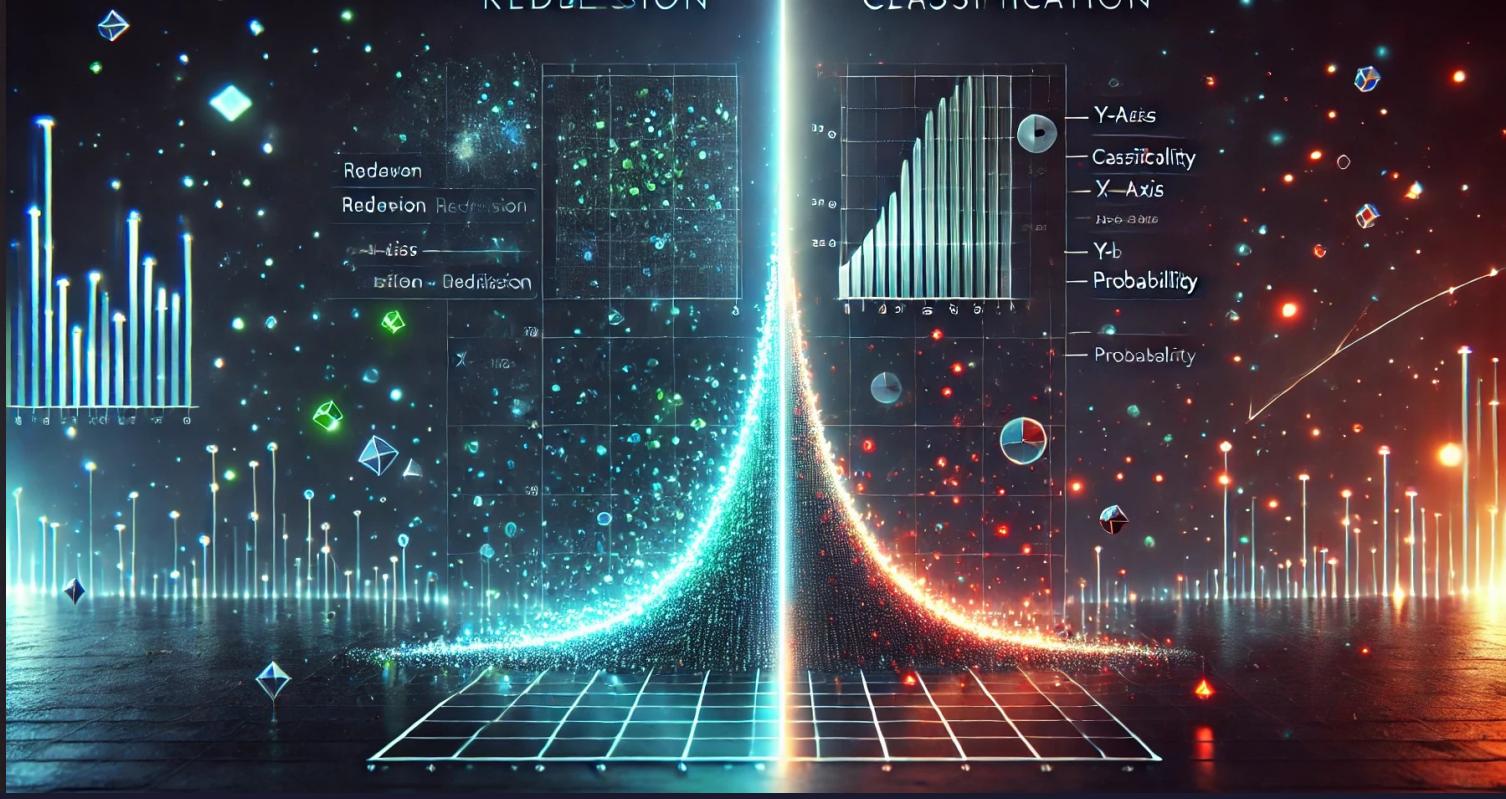
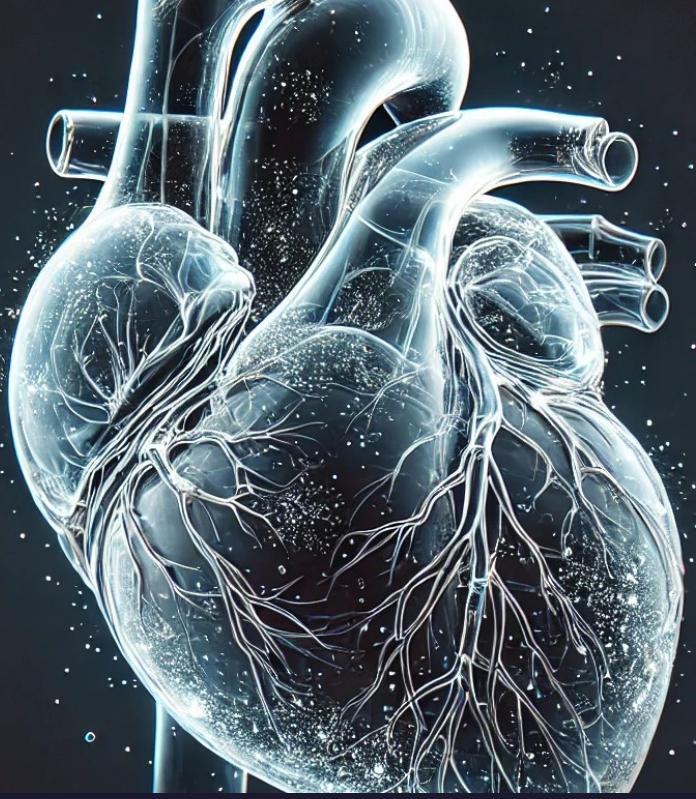


Information fusion can occur in a myriad of ways. In machine learning, early, intermediate, and late fusion is typified by if all the information flows into a single model (early), a step-wise fashion where outputs from one model become inputs for another (intermediate), and lastly, where all unique data types undergo separate modelling after which ensembling and/or voting occurs (late).



Evidence pyramid and analytics



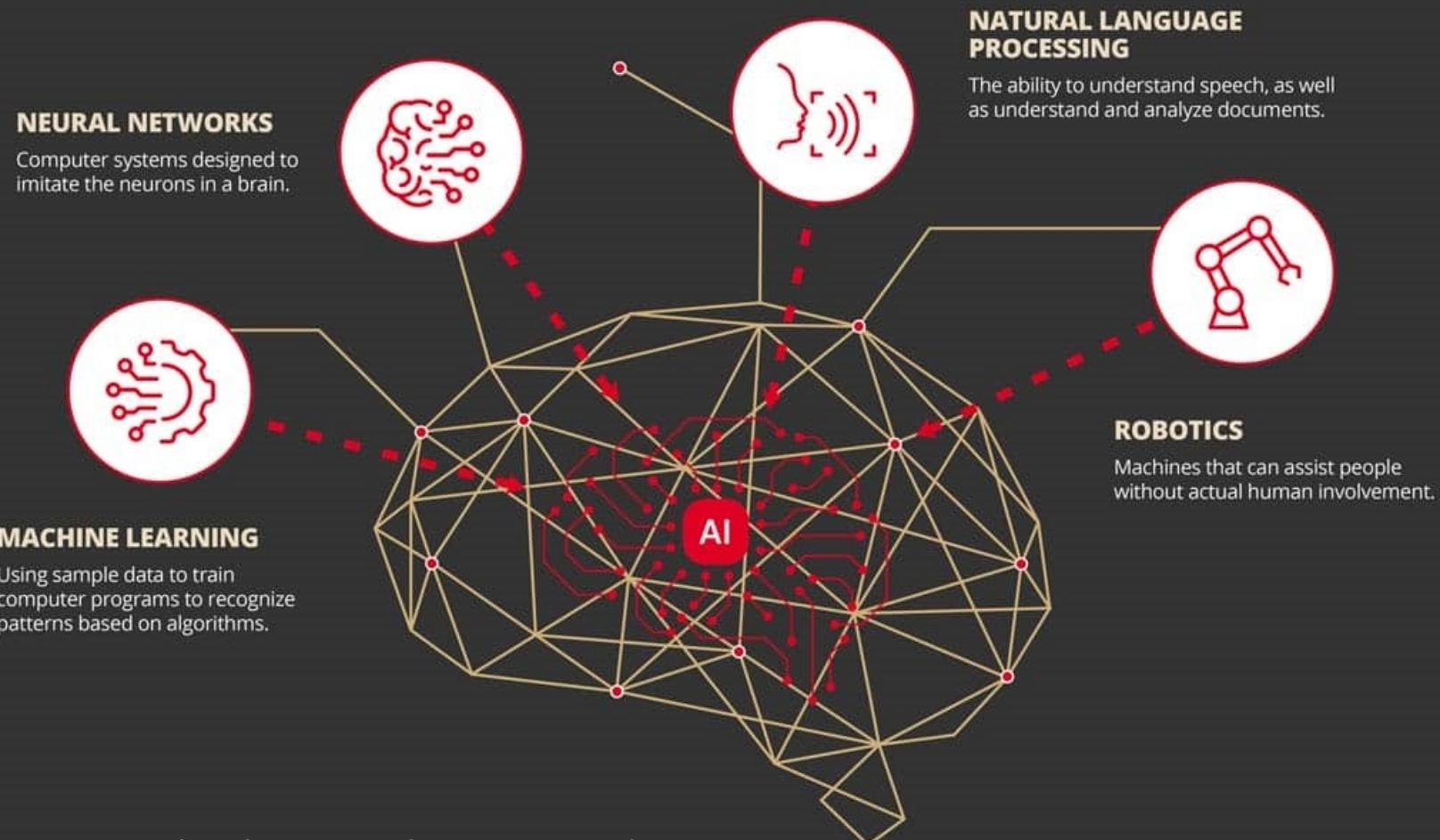


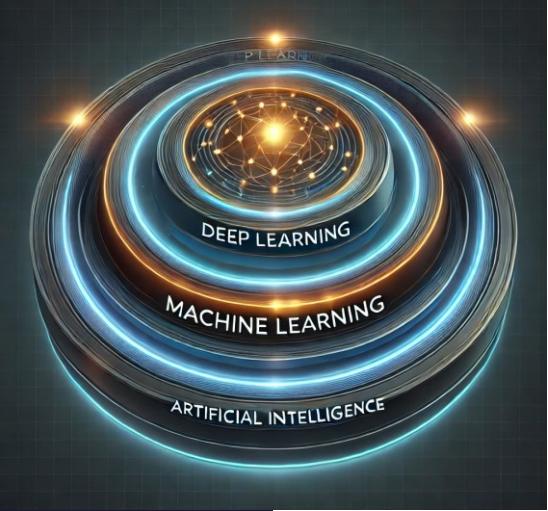
Artificial intelligence (AI)

AI is a science and a set of **computational** technologies that are inspired by — but typically operate quite differently from — the ways people use their **nervous** systems and bodies to **sense**, **learn**, **reason**, and **take action**

genome.gov

WHAT IS ARTIFICIAL INTELLIGENCE?





Artificial Intelligence (AI)

AI refers to computer systems that mimic human intelligence, enabling them to solve problems and understand language.

Machine Learning (ML)

ML is teaching systems to learn from data, enhancing performance without explicit programming.

Deep Learning

Deep Learning employs layered neural networks to find intricate patterns, excelling in tasks like image and speech recognition.



1950's

aic Artificial intelligence (AI)

Human intelligence exhibited by machines

1980's

ai Machine learning

AI systems that learn from historical data

2010's

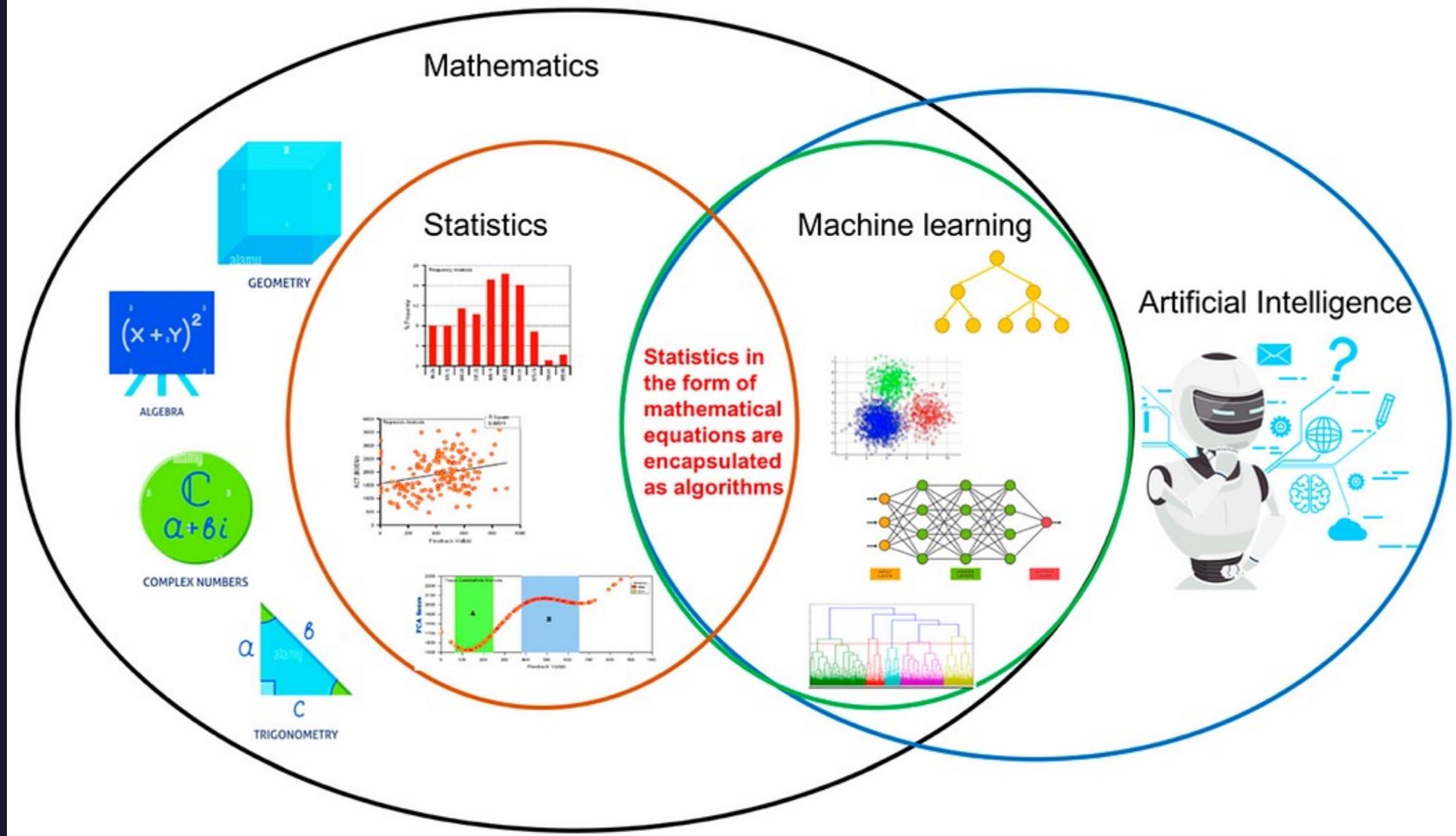
ai Deep learning

Machine learning models that mimic human brain function

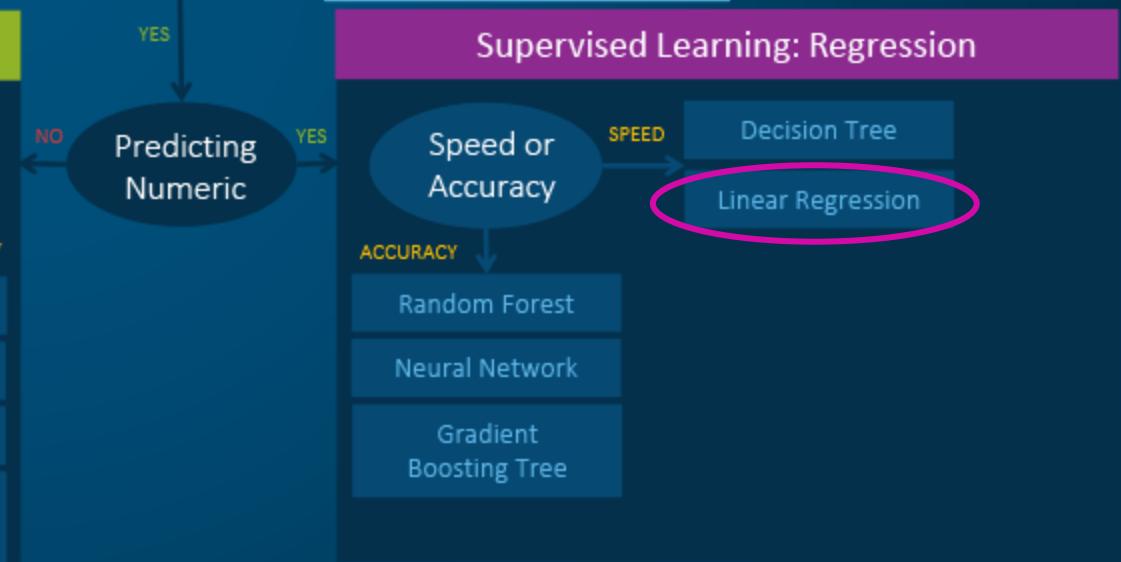
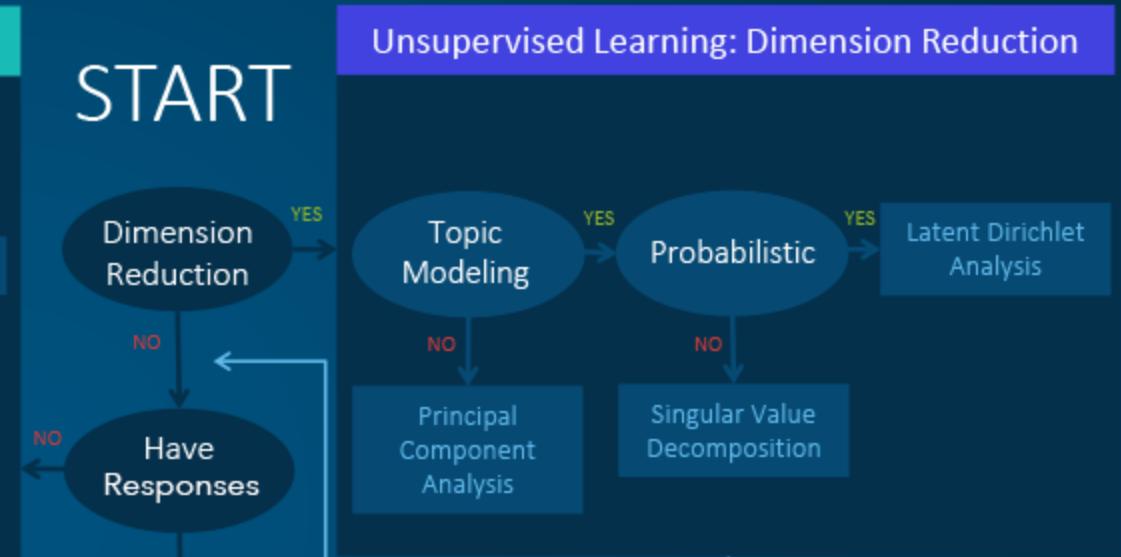
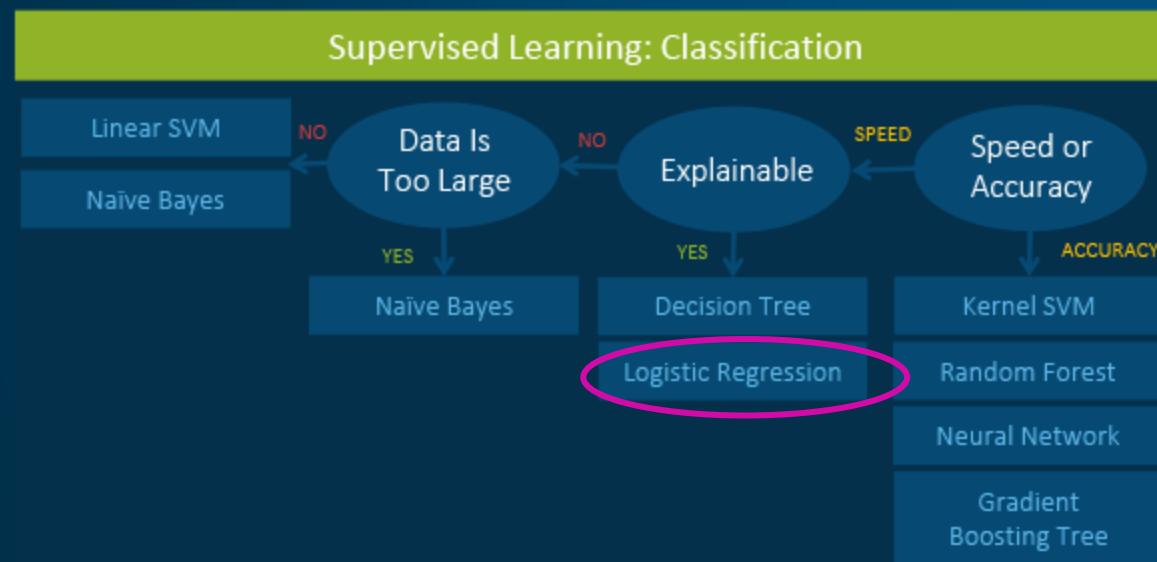
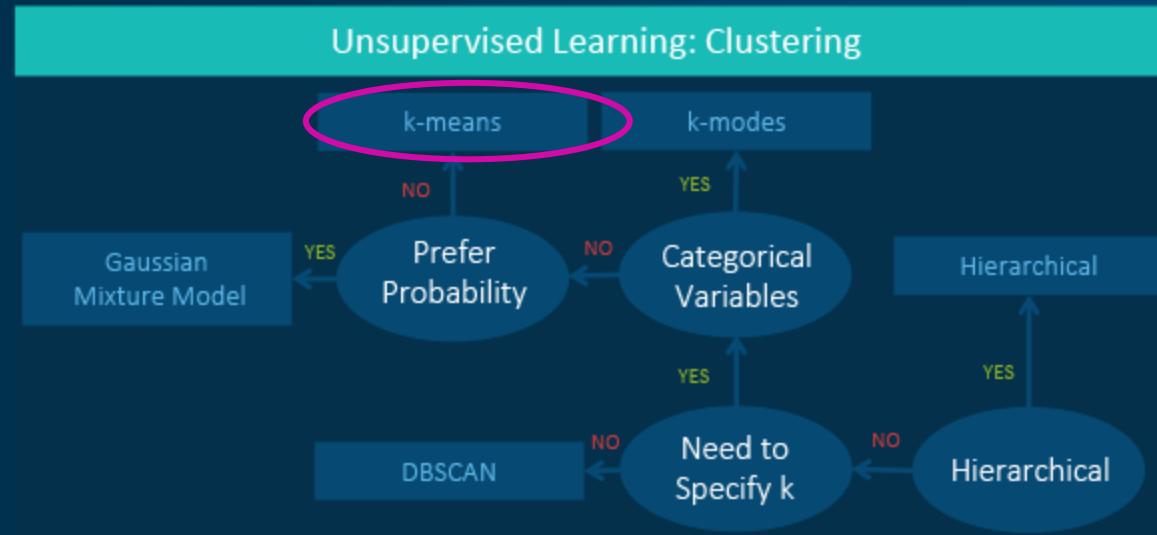
2020's

ai+ Generative AI (Gen AI)

Deep learning models (foundation models) that create original content



Machine Learning Algorithms Cheat Sheet





2018-2019: Researcher, Research Unit of Chronic Conditions, CKFF

- Identify clusters of multimorbidity in the general population: a systematic review (PI-ongoing project).
 - Clusters of multimorbidity in the Danish population.
- Identify utilization of national healthcare services by individuals with multimorbidity in Denmark.

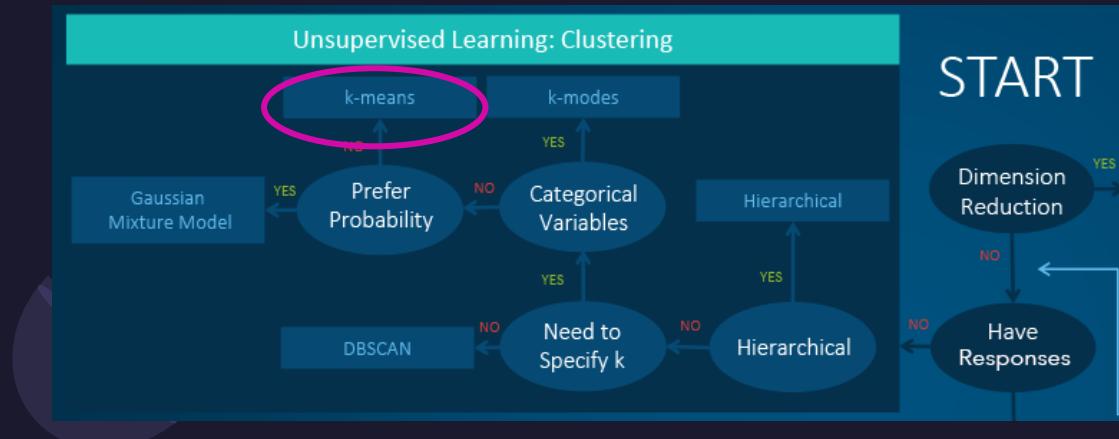


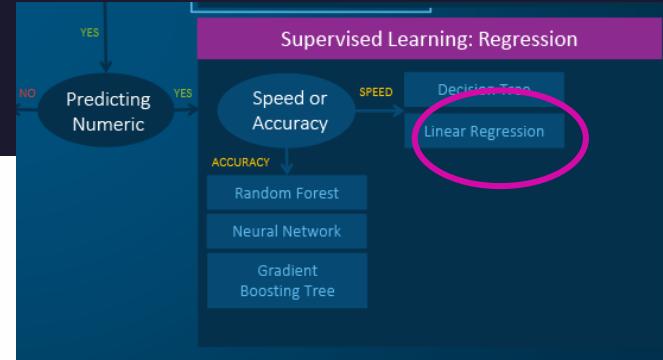
Table 1. Chronic conditions included in definition of multimorbidity (N = 16).

Allergies
Hypertension
High cholesterol
Diabetes (type 1 and type 2)
Heart disease
Stroke
Back pain
Joint disease
Osteoarthritis
Osteoporosis
Chronic obstructive pulmonary disease
Cancer
Dementia
Anxiety
Long-term use of antidepressants
Schizophrenia

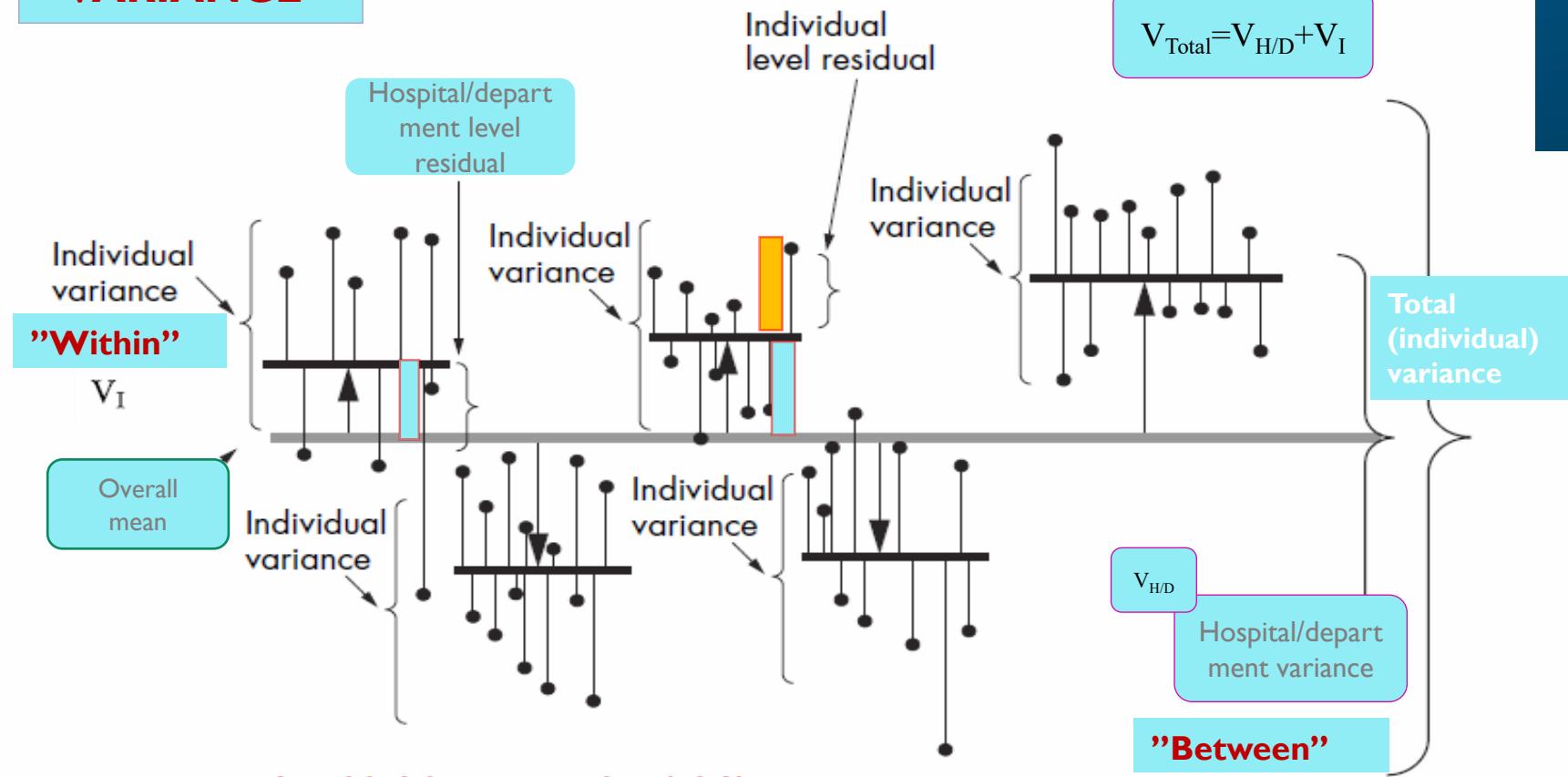
<https://doi.org/10.1371/journal.pone.0214183.t001>

VISUAL STATISTICS...

Multilevel analysis



“VARIANCE”



INTR-CLASS CORRELATION (ICC)

$$VPC/ICC = \frac{\text{Between } V_{\text{Hospital/department}}}{\text{between } V_{\text{Hospital/department}} + \text{within } V_I}$$

Total individual variance = within + between

Ghith and Merlo et al (2016): <https://doi.org/10.1371/journal.pone.0148187>



Machine learning is like
statistics on steroids.



Feature	Statistics	Machine Learning
Goal	Explain relationships, hypothesis testing, and inference	Predict outcomes and optimize models for accuracy
Focus	Understanding causality , estimating parameters, deriving meaning from data	Pattern recognition, generalization , and automation
Causality vs. Correlation	causal inference (e.g., experimental design)	focuses on correlations – difficult to identify ‘why’
Data Needs	Works well with small , structured datasets	Handles massive, unstructured datasets (images, text, etc.)
Data Abundance	Structured manageable size data	Big data : video, audio, text, images,
Methods	Regression models, hypothesis testing, confidence intervals	Neural networks, decision trees, support vector machines
Complexity	Relatively simple models, interpretable results	Deep learning models with millions of parameters
Algorithmic Advances and complex data structures	struggle with high-dimensional data	ML techniques (e.g., random forests, deep learning)
Feature Engineering	Manual	Automated with e.g., Deep learning methods, convolutional networks in image recognition
Black Box Models vs. Interpretability	Interpretable	Sometimes like "black boxes"
Non-parametric vs. Parametric Models	Usually assume a distribution	often uses non-parametric
Computational power	Classical computation	GPUs , TPUs, and distributed computing

From algorithm to bedside

Evidence pyramid



Cochrane SRs Godfathers

Archibald Leman Cochrane was a Scottish physician (12 January 1909 – 18 June 1988)

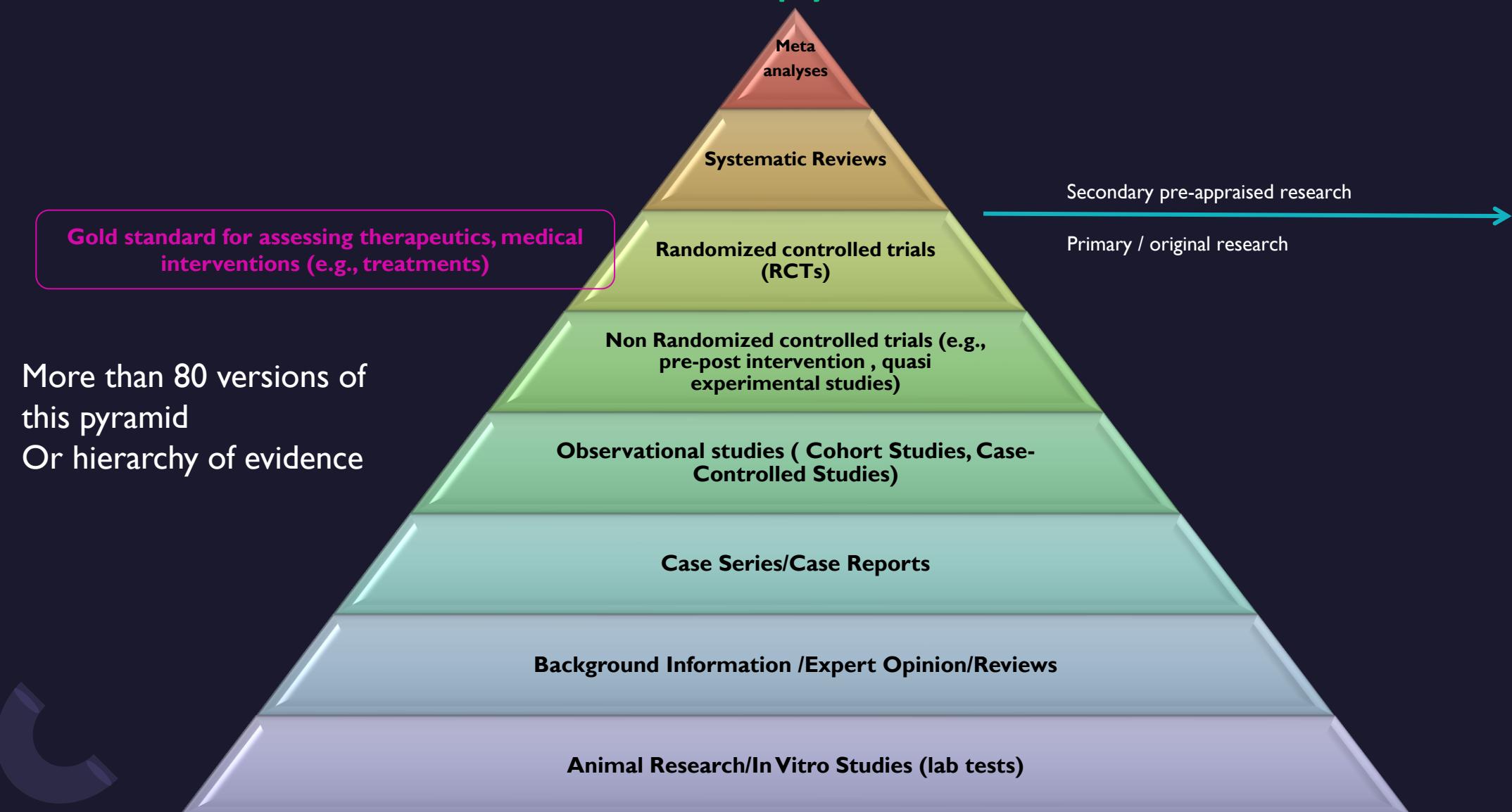


- Effectiveness and Efficiency: Random Reflections on Health Services
- Adopt **randomized** controlled trials to support the development of less biased **clinical trials** and medical interventions.

Sir Iain Geoffrey Chalmers – Cochrane Foundation



Evidence pyramid

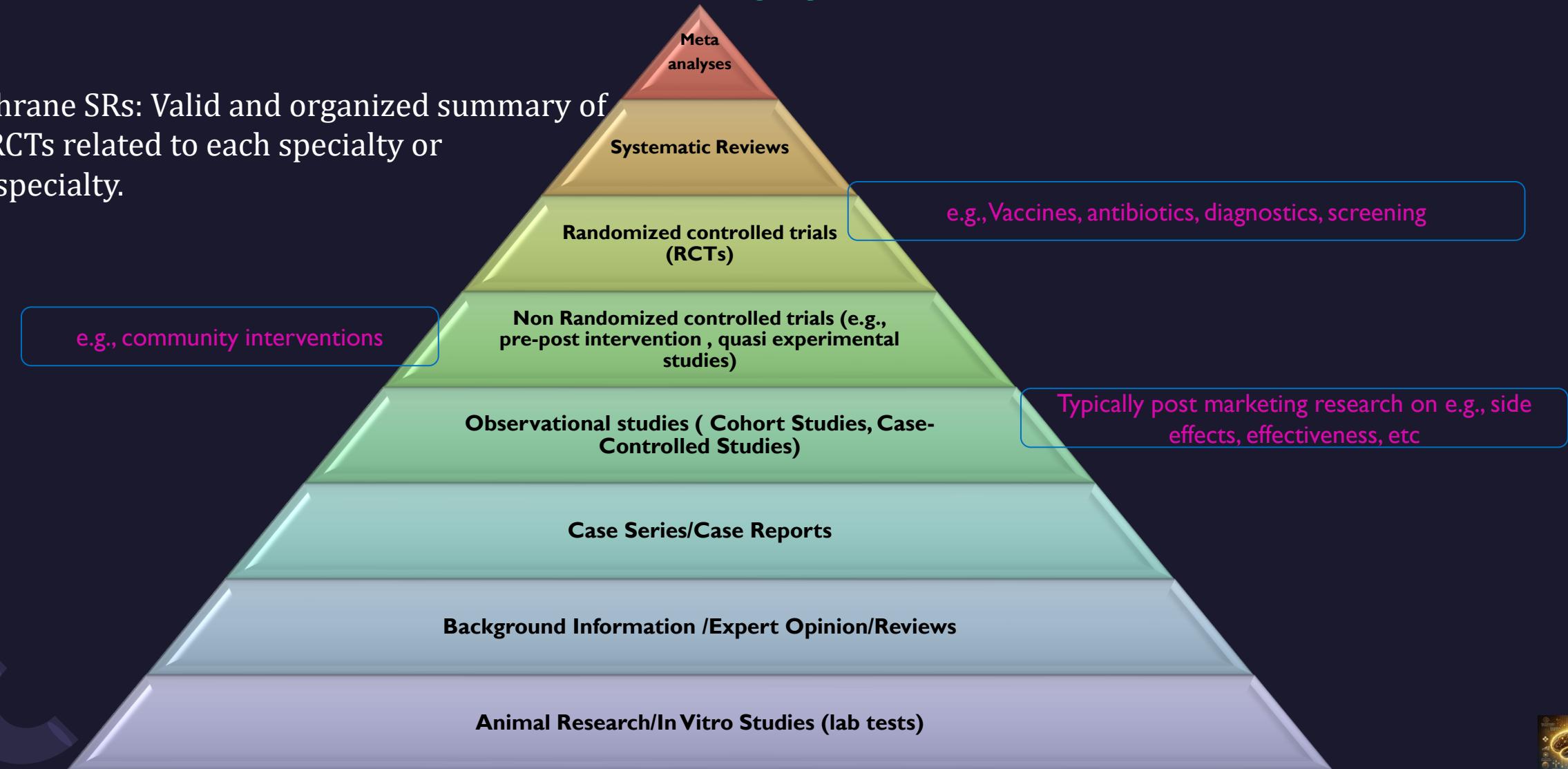


Good reference on critical appraisal tools: <https://www.cebm.ox.ac.uk/resources/ebm-tools/critical-appraisal-tools>



Evidence pyramid

Cochrane SRs: Valid and organized summary of all RCTs related to each specialty or subspecialty.



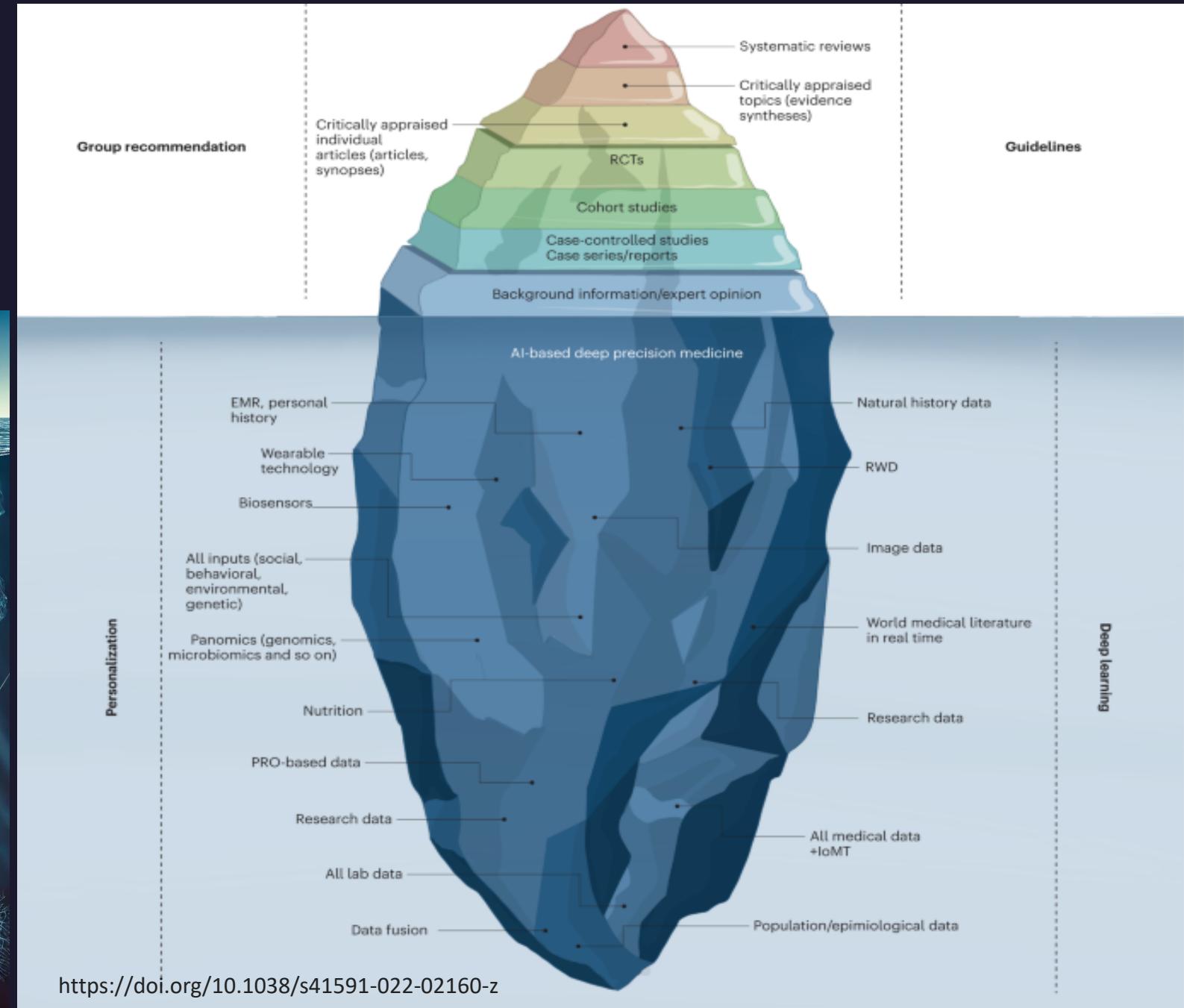


Evidence iceberg



28-01-2025

@NerminGith



43



Thank You

Linkedin connect!

