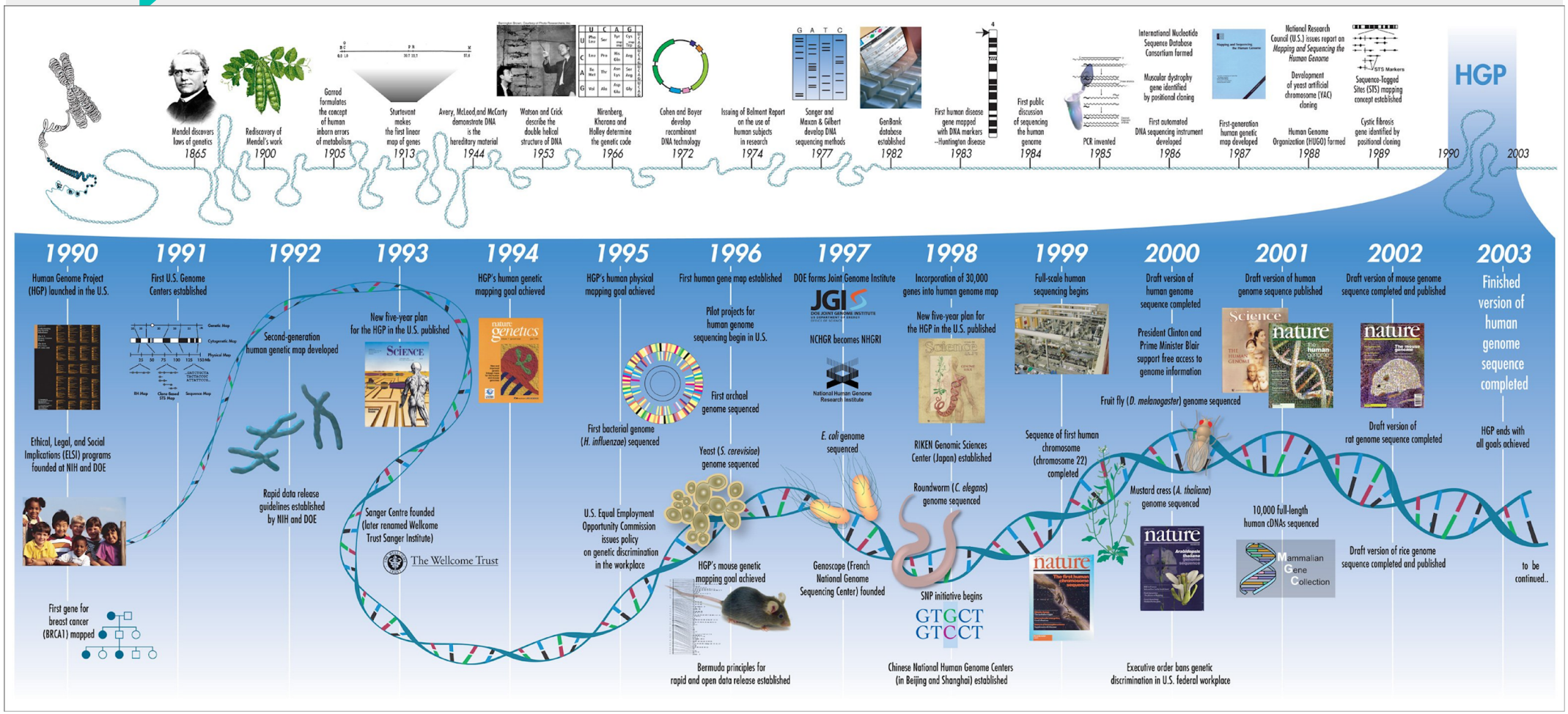
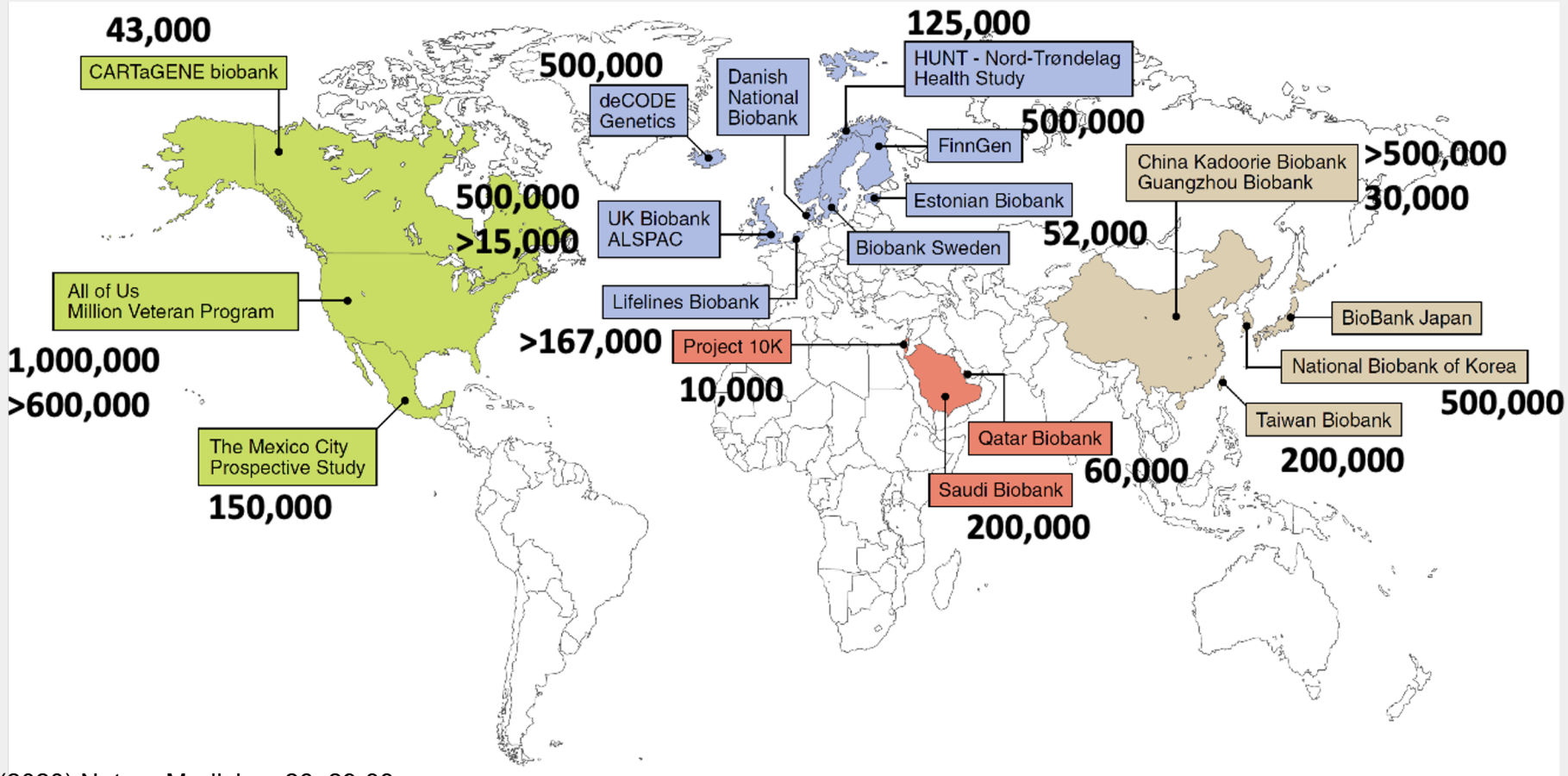
Taiwan Biobank

Su Ming-Wei wei@ibms.sinica.edu.tw

Human Genome Project



Global biobanks



Shilo, et al. (2020) Nature Medicine, 26: 29-38

Why Taiwan ?

Holo Taiwanese
Hakka Taiwanese
Mainianders
Taiwanese indigenous proples
Taiwanese new immigrants

全民健康保險 NATIONAL HEALTH INSURANCE

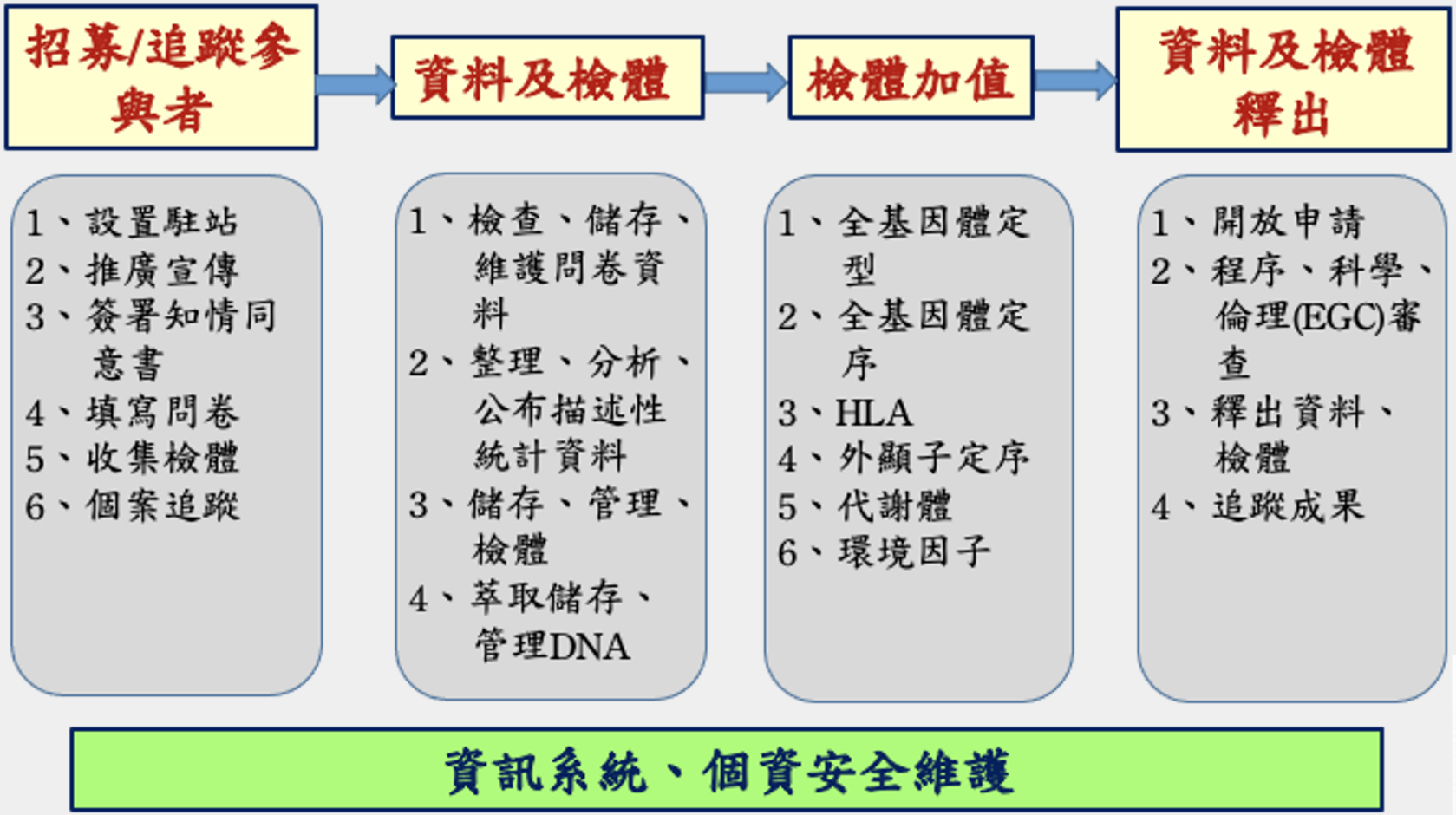
中華民國 總統府 Office of the President Republic of China (Taiwan)
www.president.gov.tw

https://ogme.edu.tw/lc/culturalGroups
https://tendashsix.com/taiwan-medical-service-ranked-first/

TWB

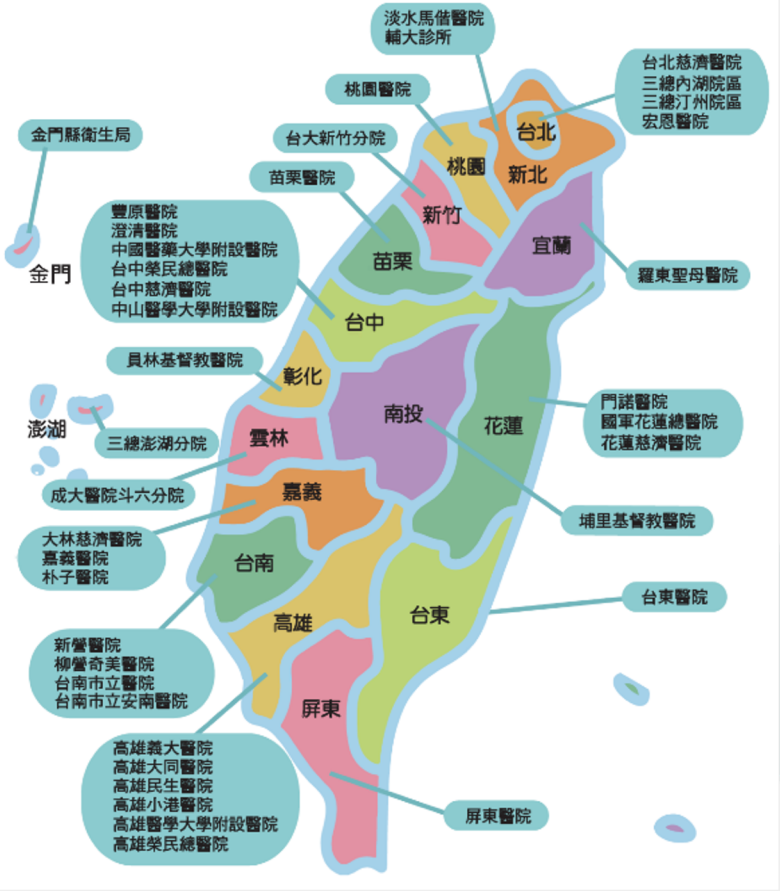
- 主管機關：衛生福利部
- 設置者：中央研究院
- 依2010.2.3 總統府頒布執行之《人體生物資料庫管理條例》設置
- 設有倫理委員會，針對資料庫之管理等有關事項進行審查及監督
- 2012.10.24 取得設置許可
- 第一個主管機關核准設置之人體生物資料庫
- 2012.11.8 開始正式收案 預計目標20萬社區民眾

TWB infrastructure



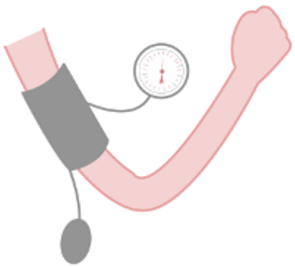
Taiwan Biobank

參與流程

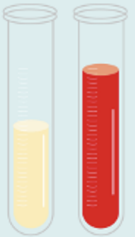


預約您前來駐站

取得您的同意



進行身體檢測



檢體採集

進行問卷訪談

2021.04 avaliable data and samples

200,000
(> 20 years old)
General population

參與個案

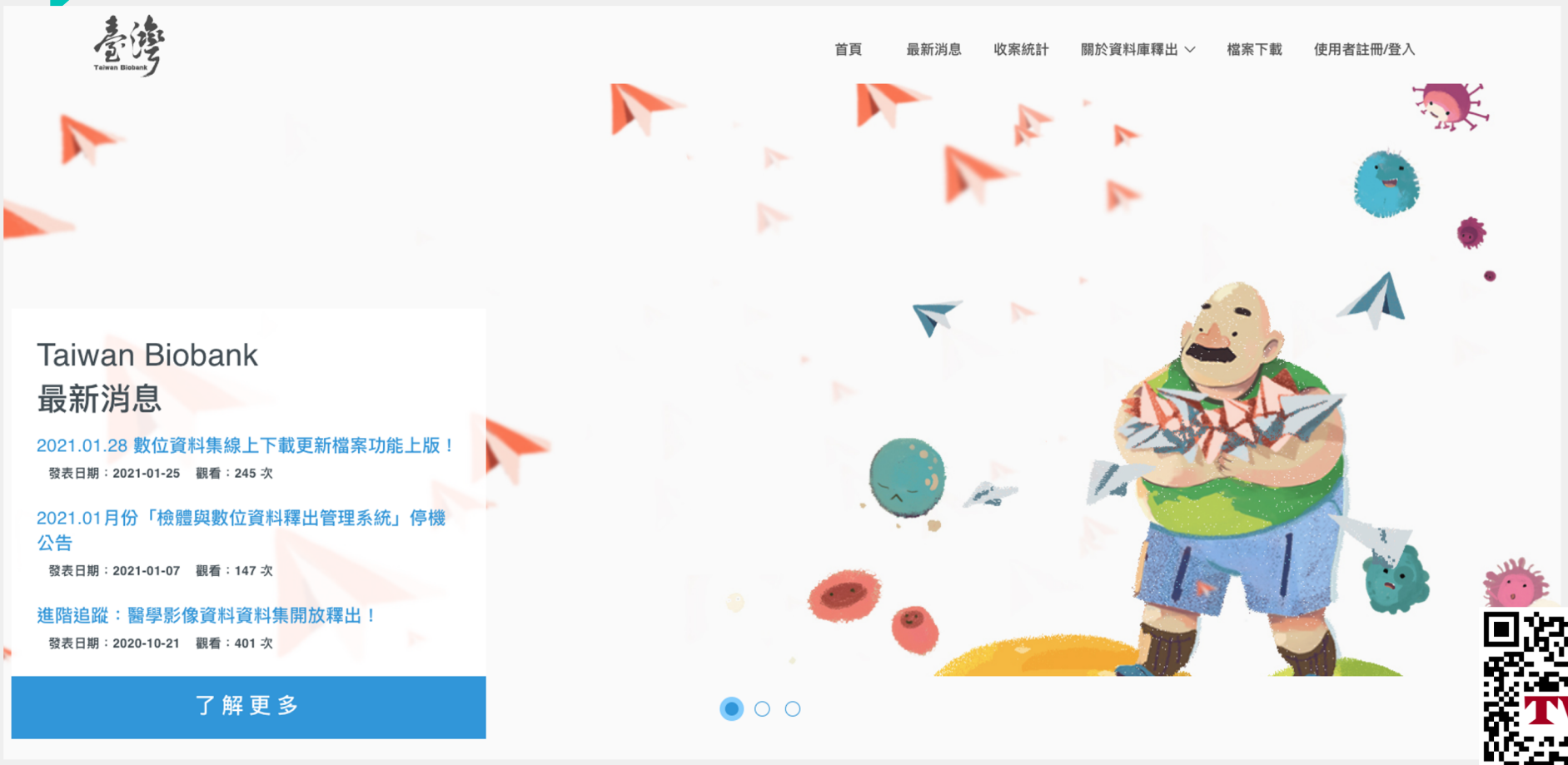
144,990

完成第一輪追蹤個案

35,950

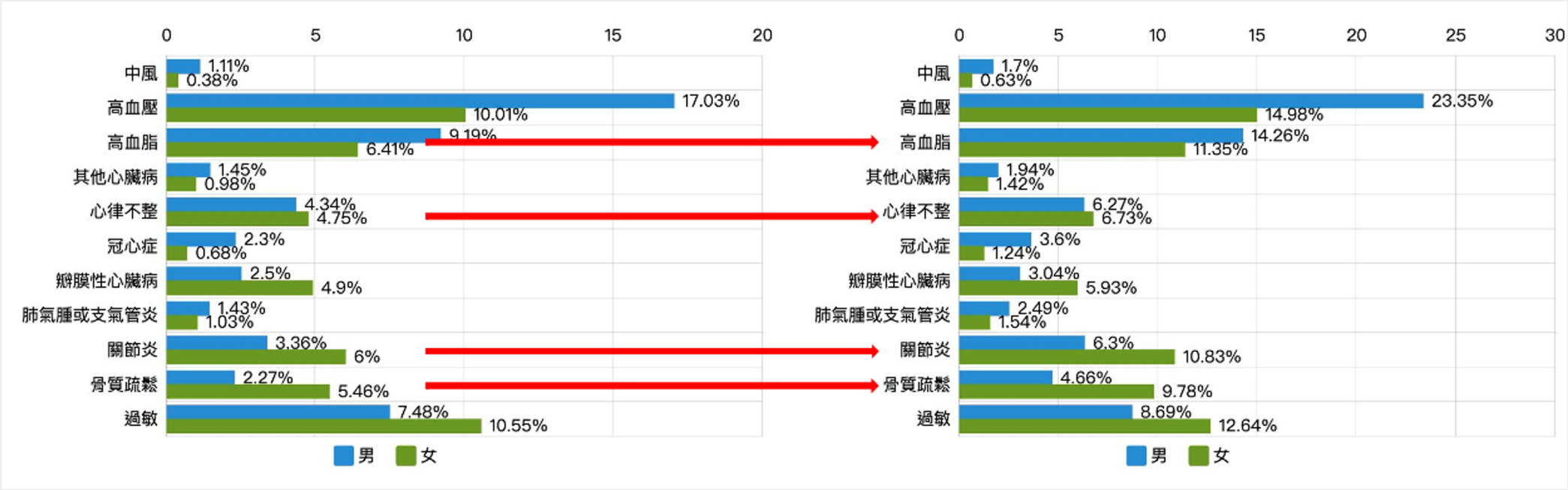
健康資訊	生活習慣問卷	~ 166,438筆
	身體檢查數據	~ 166,409筆
	血清生化數據	~ 166,339筆
	進階追蹤影像	~ 20,408筆
遺傳資訊（筆數）	• 全基因定型	116,066
	• 全基因定序	2,010
	• 人類白血球組織抗原分型	1,102
	• DNA 甲基化定序	2,474
環境暴露資訊	• 血液代謝體核磁共振光譜	869
	• 尿液塑化劑與三聚氰胺	1,353
生物檢體（管數）	• DNA (管)	697,439
	• 血漿 (0.4 c.c. 每管)	1,725,817
	• 尿液 (1.5c.c. 每管)	1,028,789

https://www.biobank.org.tw/index.php



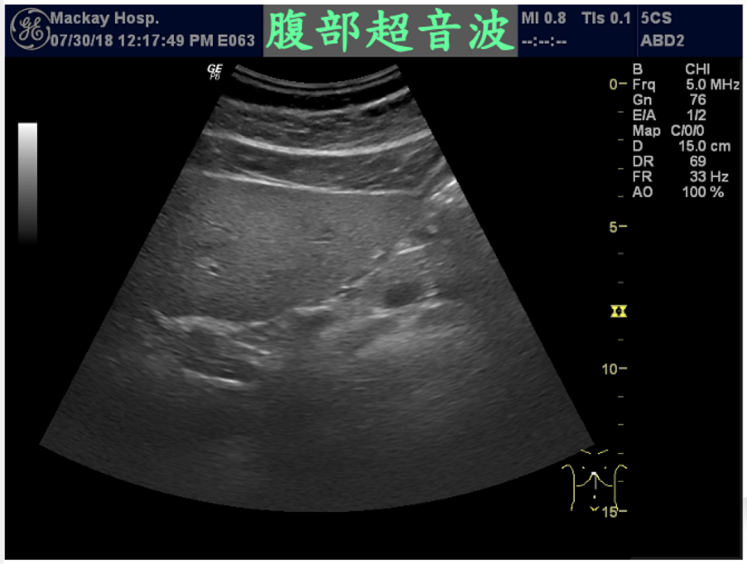
https://reurl.cc/jqjqNZ

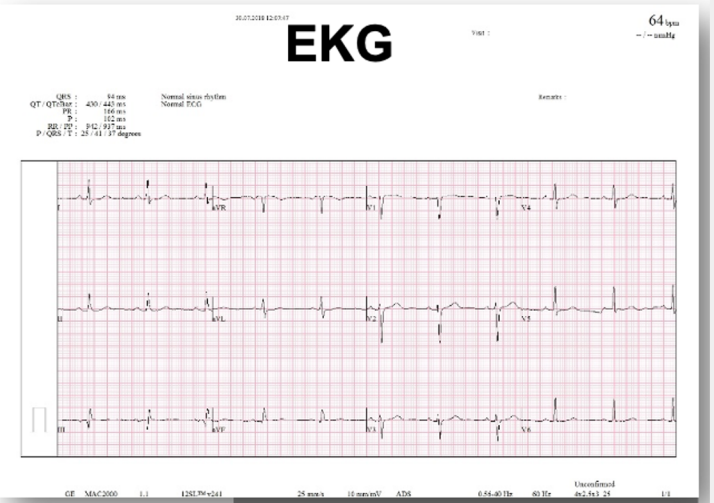
follow-up program

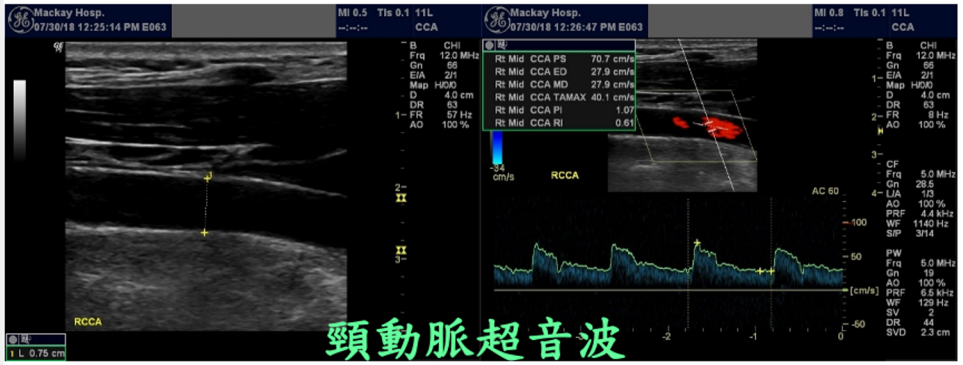


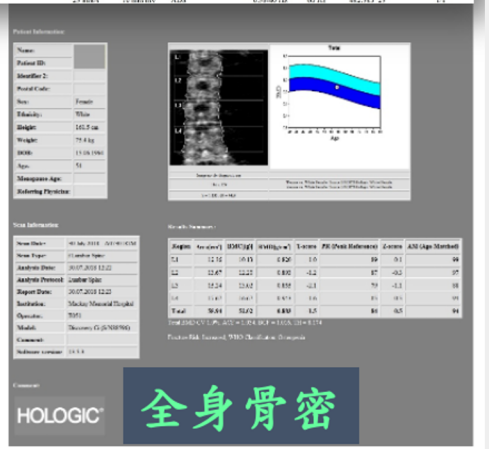
Characteristics	Male (%)	female (%)
BMI		
<18.5	1.29	4.42
18.5-24	35.83	56.96
24-27	35.25	22.58
>27 (Obesity)	27.62	16.05
Waist-Hip Ratio		
M ≥ 0.92; F ≥ 0.88	36.05	30.00
Body Fat Rate (male, female)		
≦17% , ≦20%	12.42	2.02
17-23% , 20-27%	39.11	20.02
23-25%, 27-30%	16.02	17.81
>25% , >30%	32.45	60.11

advance follow-up program









distribution of urinary melamine and phthalate metabolites

Urinary metabolites (μg/g crea.)	TWB (n=1,155)	NHANES (n=2,974)[a]
Melamine (μg/mmol crea.)	0.46 (0.43-0.49)	---
MEHP	11.37 (10.69-12.09)	N.D.[b]
MEOHP	8.17 (7.81-8.54)	3.70 (3.49-3.92)
MEHHP	12.68 (12.06-13.33)	5.86 (5.62-6.12)
MECPP	19.17 (18.36-20.02)	9.14 (8.58-9.74)
MCMHP	3.88 (3.61-4.16)	--
MBzP	1.04 (0.98-1.1)	4.63 (4.06-5.28)
MnBP	20.51 (19.46-21.61)	10.2 (9.53-10.9)
MiBP	8.38 (7.92-8.87)	8.71 (8.06-9.42)
MEP	12.86 (11.82-13.99)	34.7 (31.0-39.0)
MMP	2.22 (2.09-2.36)	---
MiNP	N.D.[b]	N.D.[b]

[a]Data from the urine samples of 2015-2016 for the U.S. population from the National Health and Nutrition Examination Survey. (NHANES, 2019)
[b]N.D., the chemical was analyzed but the proportion of results below limit of detection was too high to provide a valid result.

data and tubes

~1.5 PB

~3 million tubes

ISO certificate

ISO/IEC 27001 for management system

TÜV NORD

CERTIFICATE

Management system as per
ISO/IEC 27001:2013

The Certification Body TÜV NORD CERT GmbH hereby confirms as a result of the audit, assessment and certification decision according to ISO/IEC 17021-1:2015, that the organization

TAIWAN BIOBANK
No. 99, Ln. 130, Sec. 1, Academia Rd., Nangang Dist.,
Taipei City 115202, Taiwan (R.O.C.)
2F,Building B, National Biotechnology Research Park

operates a management system in accordance with the requirements of ISO/IEC 27001 : 2013 and will be assessed for conformity within the 3 year term of validity of the certificate.

The information Security Management System of Information Management Department to provide the operation and maintenance, for Cohort and Dataland Management System Development and Maintenance, Intranet Network Support Service and Server Room

As per the Statement of Applicability (SOA), version 2.0, Effective date 2020-09-28

Certificate Registration No. 44 121 18820001
Audit Report No. IT241-R1

Valid from 2020-11-08
Valid until 2023-11-07
Initial Certification 2017-11-08

Certification Body
at TÜV NORD CERT GmbH

TUV ASIA PACIFIC LIMITED
Taiwan Branch
Room A1, 9F., No.333, Sec.2,
Tun Hua S. Rd.,
Taipei 10669 Taiwan, R.O.C.
2020-11-06

TÜV NORD CERT GmbH Langemarckstrasse 20 45141 Essen www.tuev-nord-cert.com

ISO/IEC 29100 for personal information protection

TÜV NORD

CERTIFICATE

Conformity of Personal Information Protection as Defined per
ISO 29100:2011

In accordance with TÜV NORD CERT procedures, it is hereby certified that

TAIWAN BIOBANK
No. 99, Ln. 130, Sec. 1, Academia Rd., Nangang Dist.,
Taipei City 115202, Taiwan (R.O.C.)
2F,Building B, National Biotechnology Research Park

with the locations according to the annex

applies a management system in line with the above standard for the following scope

Personal information protection of recruitment participant of Taiwan Biobank(working flow of personal information collection, processing and use) including:
(1) The Center of Taiwan Biobank
(2) Tri-Service General Hospital- Neihu Branch Station
(3) Tri-Service General Hospital- Tingjhou Branch Station
(4) Cheng Ching Hospital Chung Kang Branch Station
(5) Kaohsiung Municipal Ta-Tung Hospital Station
(6) Taichung Veterans General Hospital Station
(7) Kaohsiung Veterans General Hospital Station

Certificate Registration No. TW2918001
Audit Report No. IT-241

Valid from 2018-12-13
Valid until 2021-12-12
Issue Date 2020-11-18
Initial Certification 2018-12-13

Certification Body
at TUV ASIA PACIFIC LIMITED Taiwan Branch

TUV ASIA PACIFIC LIMITED
Taiwan Branch
Room A1, 9F., No.333, Sec.2,
Tun Hua S. Rd.,
Taipei 10669 Taiwan, R.O.C.

This certification was conducted in accordance with the TUV ASIA PACIFIC LIMITED Taiwan Branch auditing and certification procedures based on TÜV NORD CERT auditing and certification procedures and is subject to regular surveillance audits.

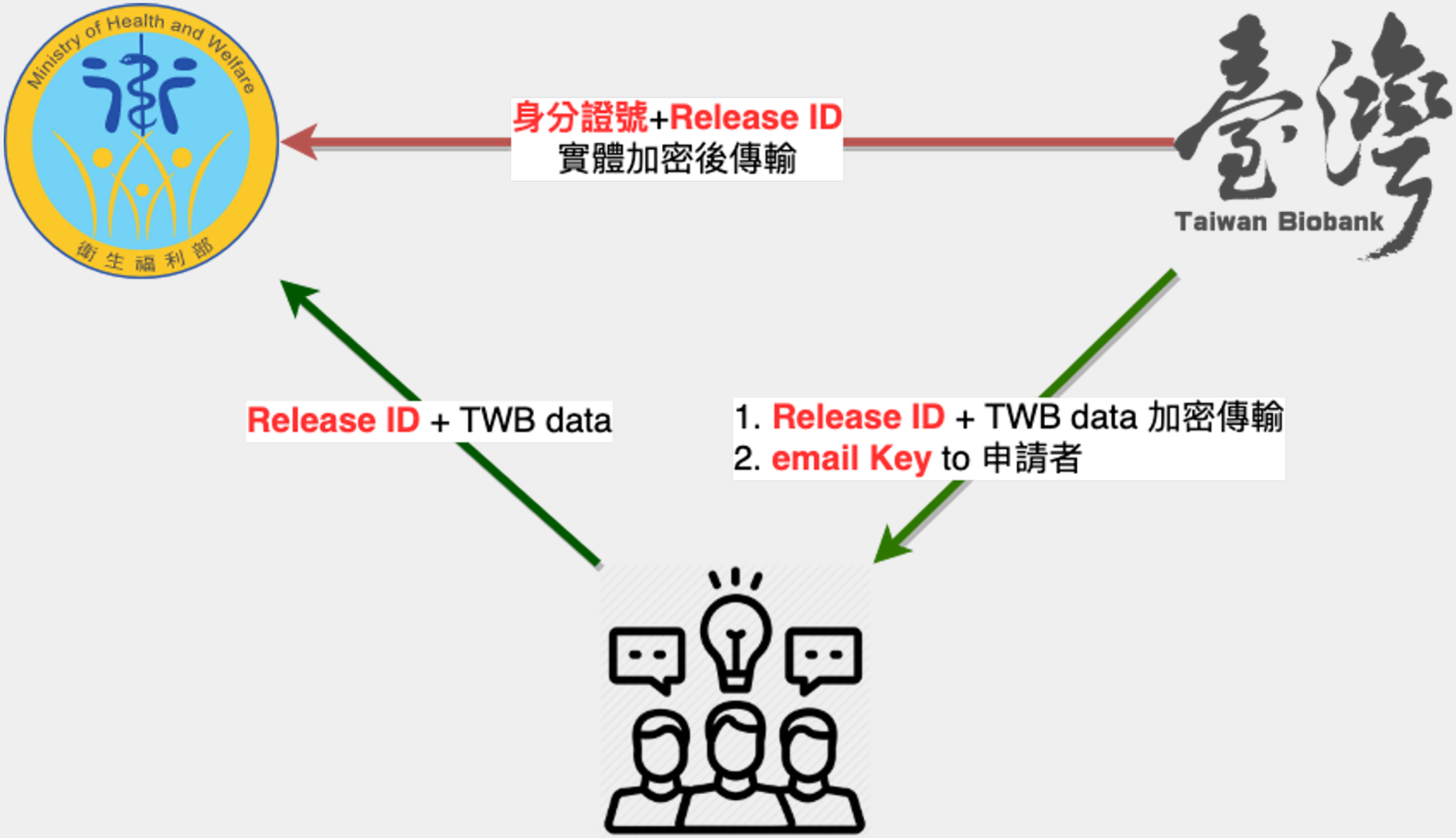
www.tuvnord.com.tw

link NHIRD - TWB projects

通過倫審**共 91 案**	已串連	EGC 倫審中
	18 案	5 案
總計	23 案	

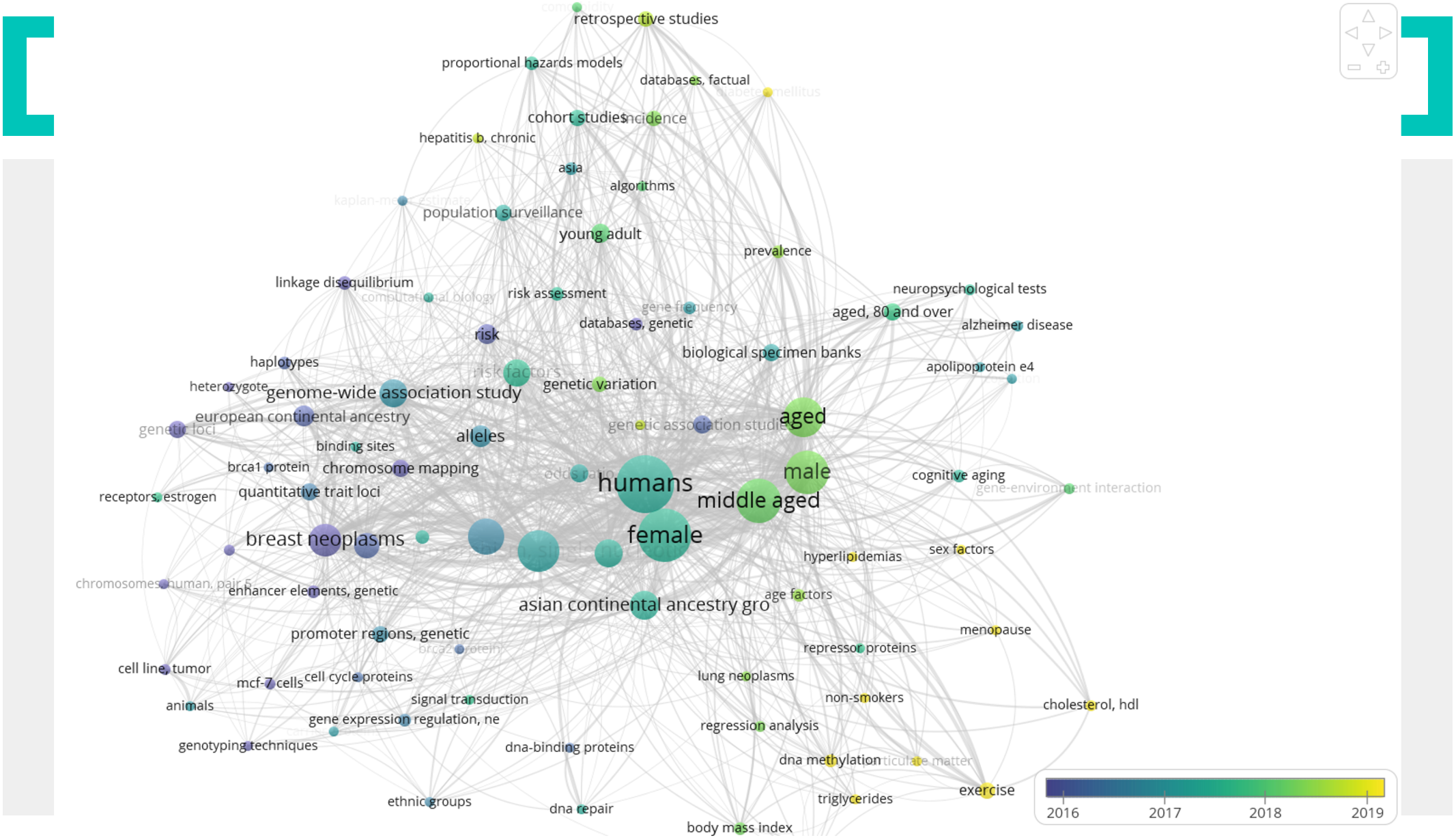
1. **臺灣人體生物資料庫為生物醫學研究的目的而建立，且經參與者「事前同意」及完善的「事後退出」流程，保障參與者權利**
2. **IRB EGC 雙重保障**

TWB-NHIRD


Ministry of Health and Welfare
衛生福利部
身分證號+Release ID
實體加密後傳輸
臺灣
Taiwan Biobank
Release ID + TWB data
1. Release ID + TWB data 加密傳輸
2. email Key to 申請者

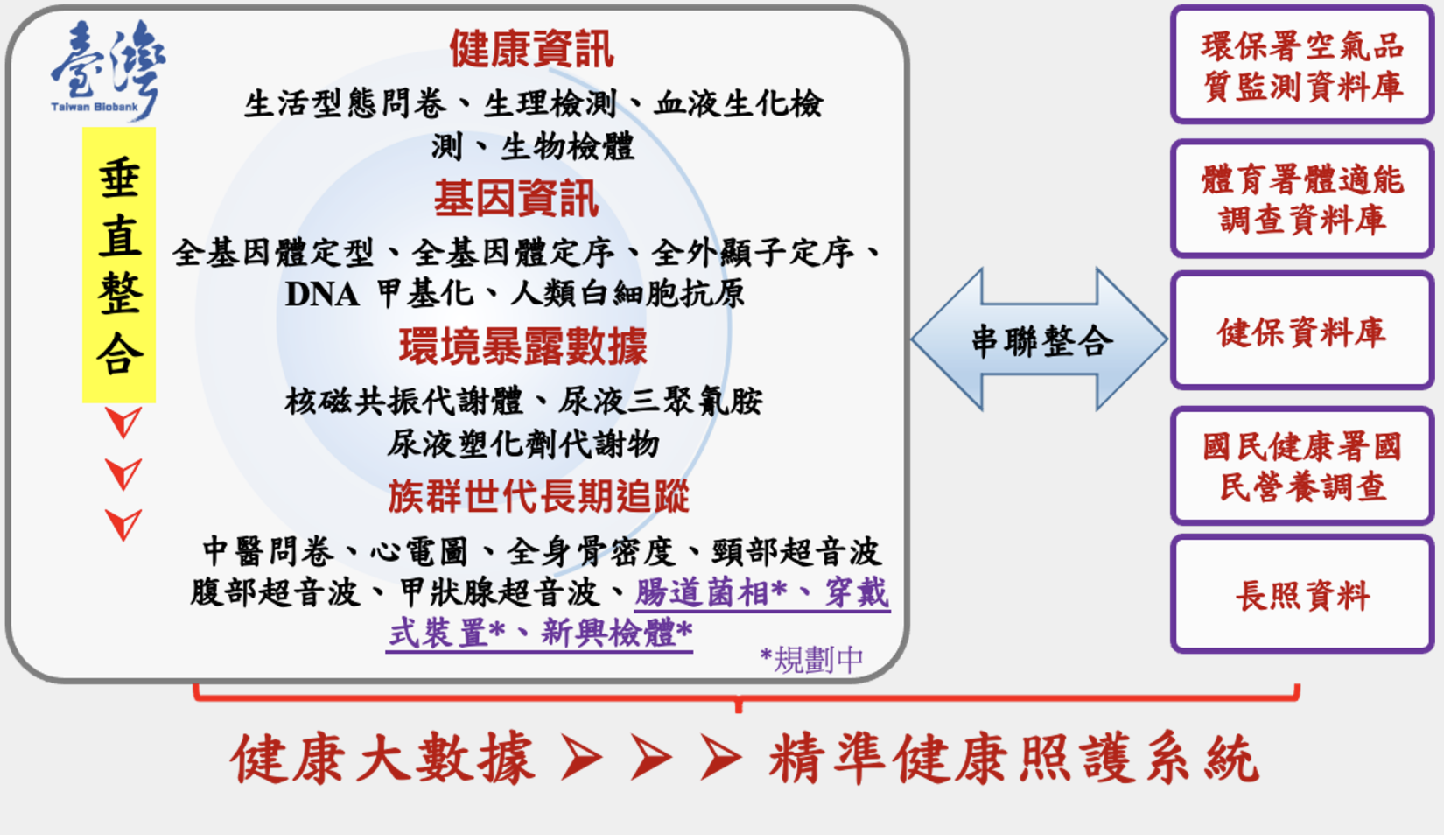
TWB - NHIRD

Diseases	Sex	TWB (%)	NHIRD (%)
Diabetes	Male	10.98	12.53
	Female	7.25	8.84
Hypertension	Male	29.12	26.7
	Female	16.86	17.08
Hyperlipidemia	Male	26.69	26.09
	Female	20.51	21.05

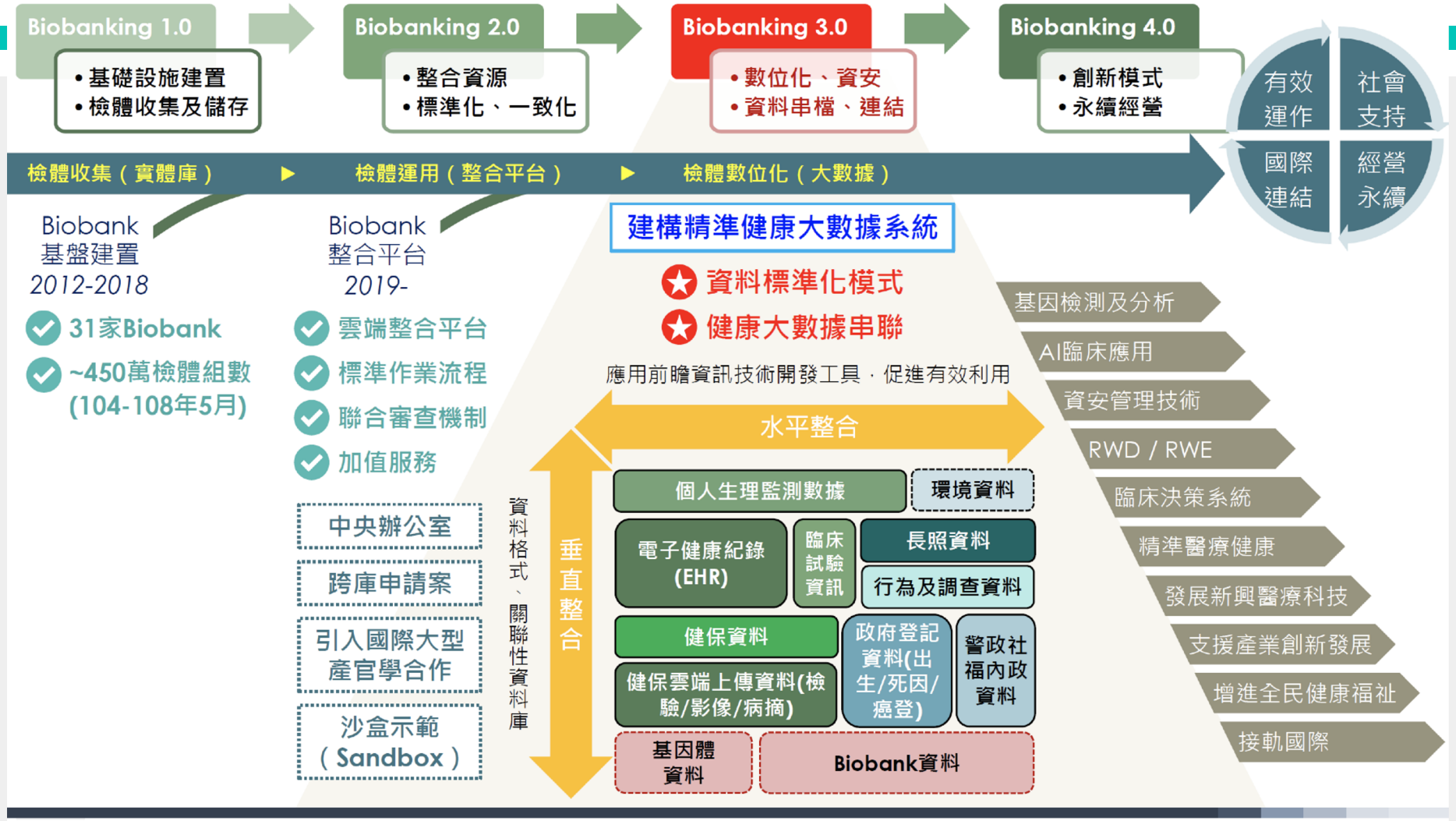

retrospective studies
proportional hazards models
databases, factual
cohort studies
incidence
hepatitis b, chronic
asia
algorithms
population surveillance
young adult
prevalence
linkage disequilibrium
risk assessment
neuropsychological tests
aged, 80 and over
alzheimer disease
gene frequency
databases, genetic
risk
biological specimen banks
haplotypes
risk factors
apolipoprotein e4
heterozygote
genome-wide association study
genetic variation
european continental ancestry
genetic association studies
aged
genetic loci
alleles
binding sites
brca1 protein
chromosome mapping
humans
male
cognitive aging
receptors, estrogen
quantitative trait loci
middle aged
gene-environment interaction
breast neoplasms
female
hyperlipidemias
sex factors
chromosomes, human, pair 5
enhancer elements, genetic
age factors
asian continental ancestry gro
menopause
promoter regions, genetic
repressor proteins
cell line, tumor
lung neoplasms
mcf-7 cells
cell cycle proteins
non-smokers
cholesterol, hdl
animals
signal transduction
gene expression regulation, ne
regression analysis
genotyping techniques
dna-binding proteins
dna methylation
exercise
ethnic groups
dna repair
triglycerides
body mass index
2016
2017
2018
2019


收案
加值
資訊
釋出
全國設置40個駐站
14萬一般民眾參與，追蹤3萬5千例
完整生活問卷、檢測資料、生物檢體
數位資訊 1.2 Petabyte (PB)
生物檢體 >300萬管，
資訊安全及隱私保護獲國際雙認證
2,000 例全基因體序列
開發國人專屬全基因體定型晶片TWB2
10萬筆全基因體定型
建立以GRCh38版人類基因體參考序列
基因體定型晶片之基因體插補
完成逾2,000筆尿液塑化劑代謝物分析
設置Taiwan View公開網站
建立基因體學、表觀基因體學、代謝體學、及環境暴露等資訊
一般民眾之描述性統計分析
申請者成功串聯健保資料庫
成功國際傳輸
支援41個機構之141件計畫
釋出數位資料超過九千萬人次、生物檢體近18萬管
申請者發表相關國際期刊逾200篇
開發釋出申請系統，優化申請流程，提升便捷與安全之異地服務系統

Taiwan Biobank - 垂直整合

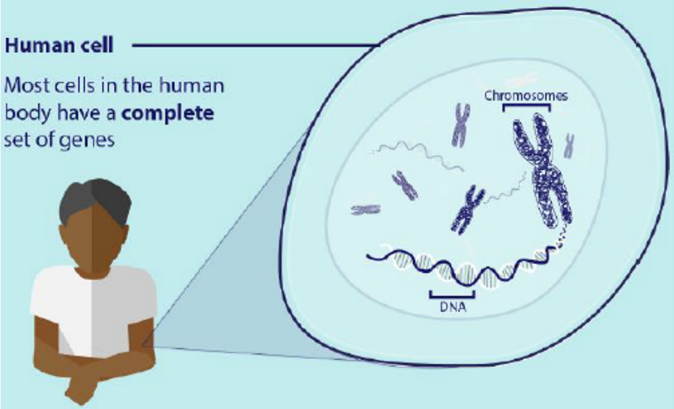


建構以大數據為基礎之國家精準健康架構



Human cell

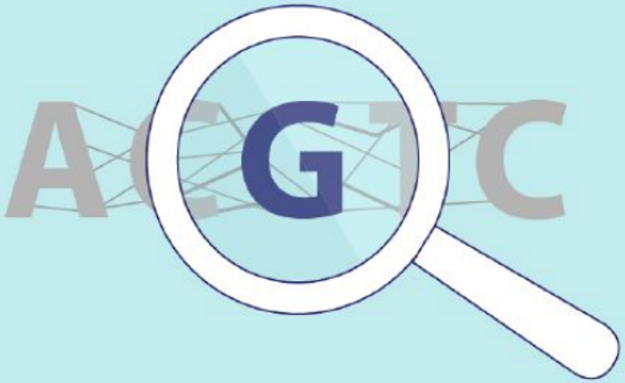
Most cells in the human body have a **complete** set of genes



Your **genome** is one whole set of all your genes plus all the DNA between your genes.

There are around **20,000** genes in your **genome**

Now



We know that the non-gene (non-coding) parts of your genome may have a role to play so we look at the whole thing, every single letter, and how the different parts work together.

5%

were analysed

Only the active genes were looked at

Your genome

with

3 Billion

pairs of letters in the human genome

AAGTAATATGC
TTCTAGGCGTC
TCAAGATGCAT
CTAGCACAGC
GCCCTTTATTA
TCTCTATACTCA
ACTACTAGGGC
TATTTCATATCT
AAATACGCTCG
AGGCTACTGAC
TTATGCTATCG
ATCTCGAGCGC
TDCCGTAATTT
TCGCGAATCAG
AAGTAATATGC
TTCTAGGCGTC
TCAAGATGCAT
CTAGCACAGC
GCCCTTTATTA

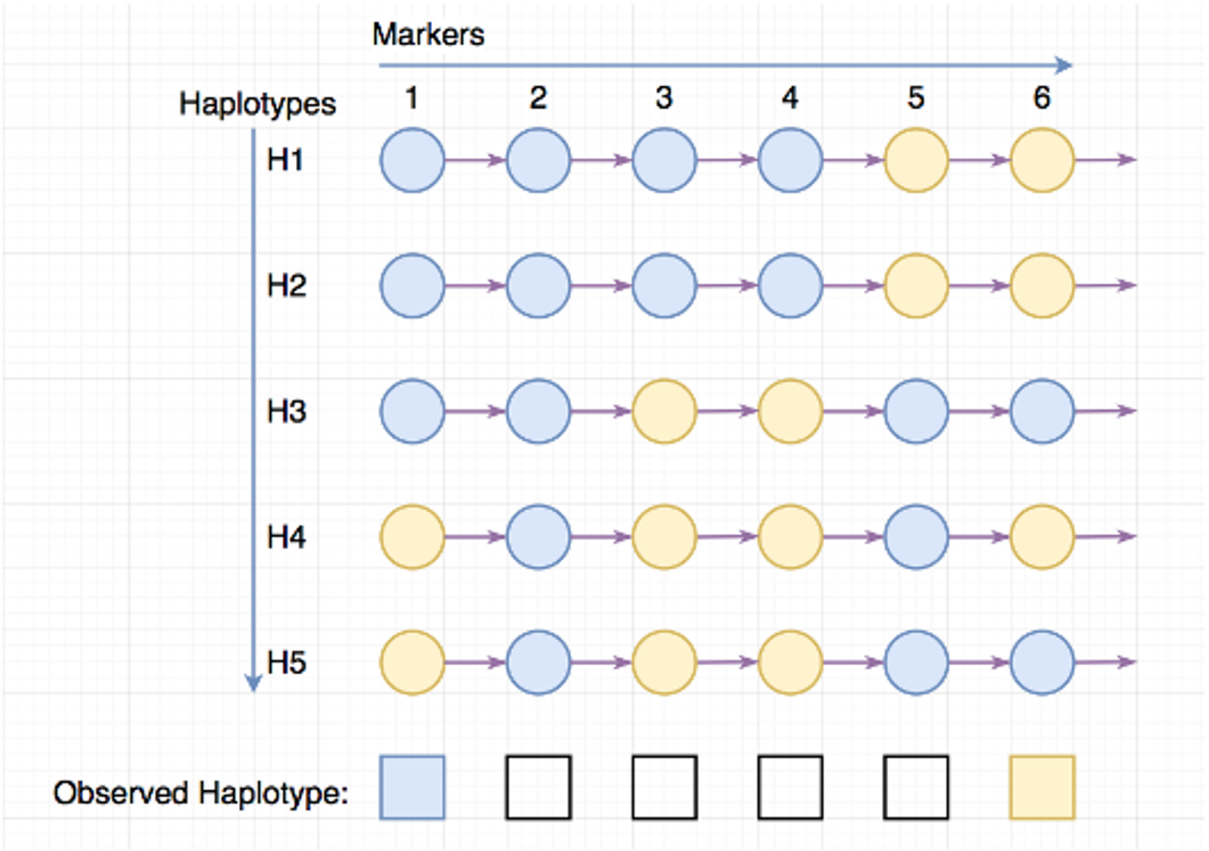
95%

were unused

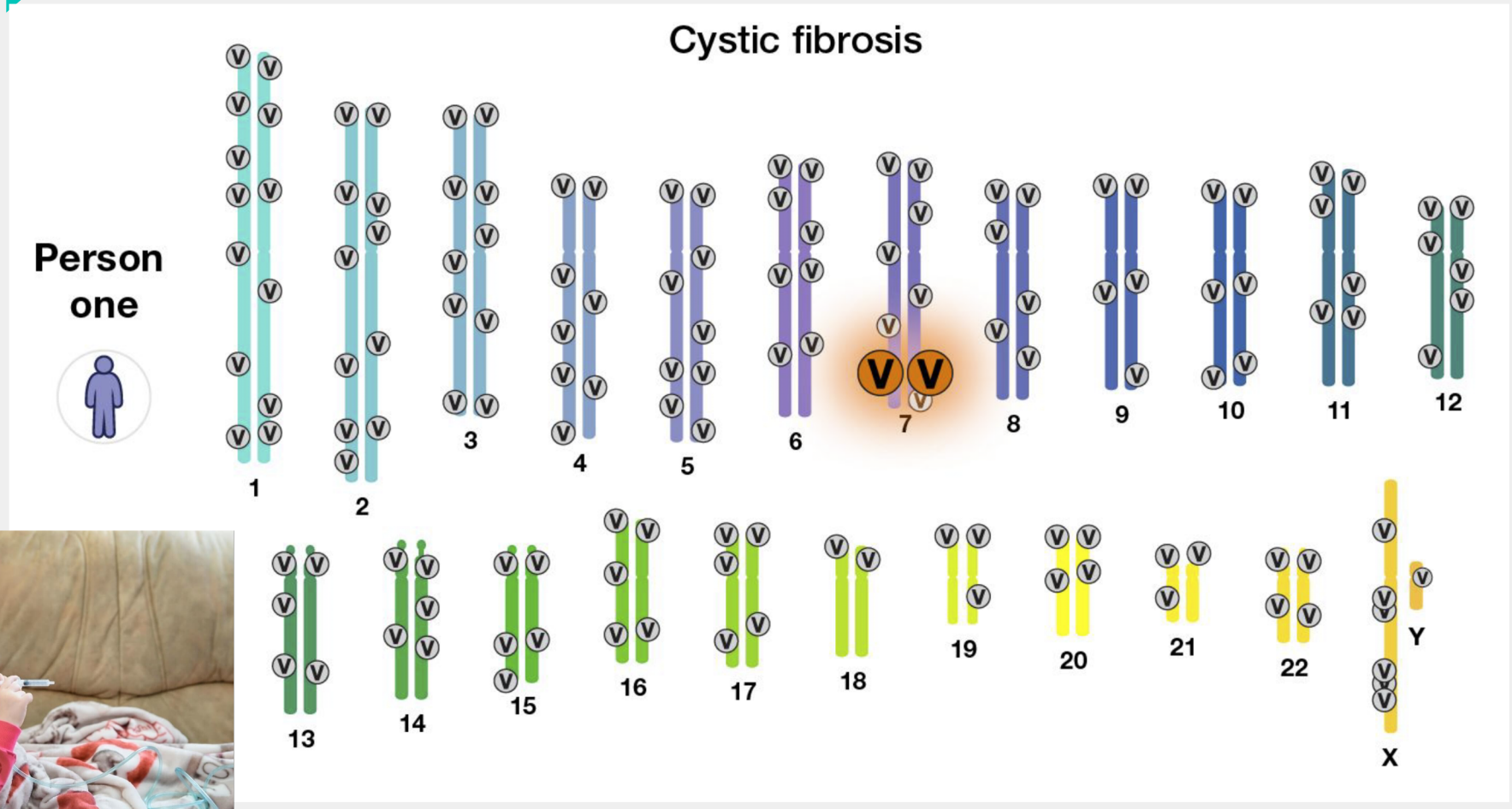
All the non-gene sections that we didn't understand were disregarded as useless

臺灣 BIOBANK • 健康世代

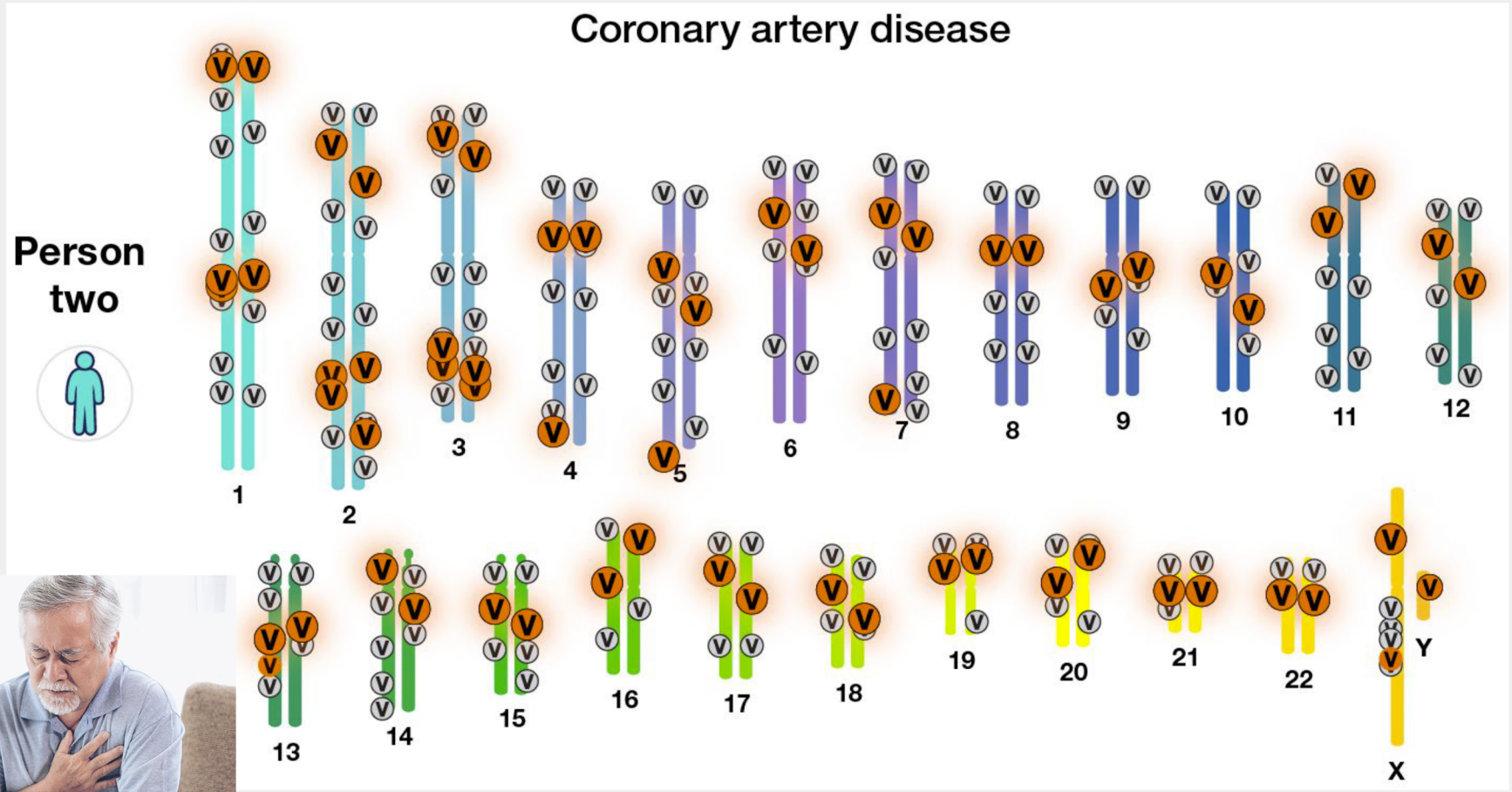
chip design

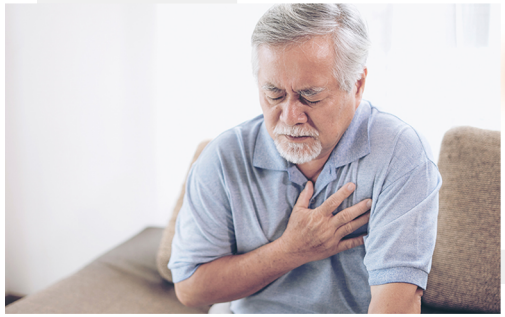

Markers
Haplotypes
1
2
3
4
5
6
H1
H2
H3
H4
H5
Observed Haplotype:

cystic fibrosis (囊狀纖維化) → transmembrane conductance regulator (CFTR) gene on chromosome 7.



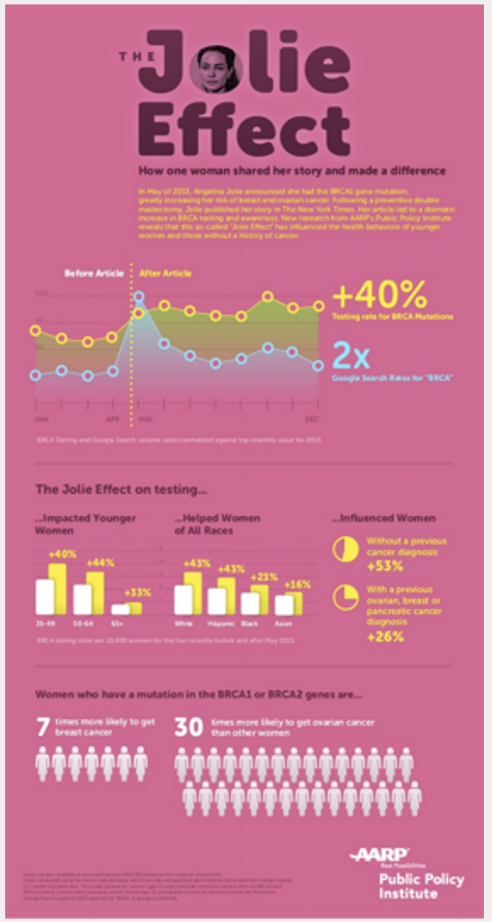
complex diseases (polygenic disease)



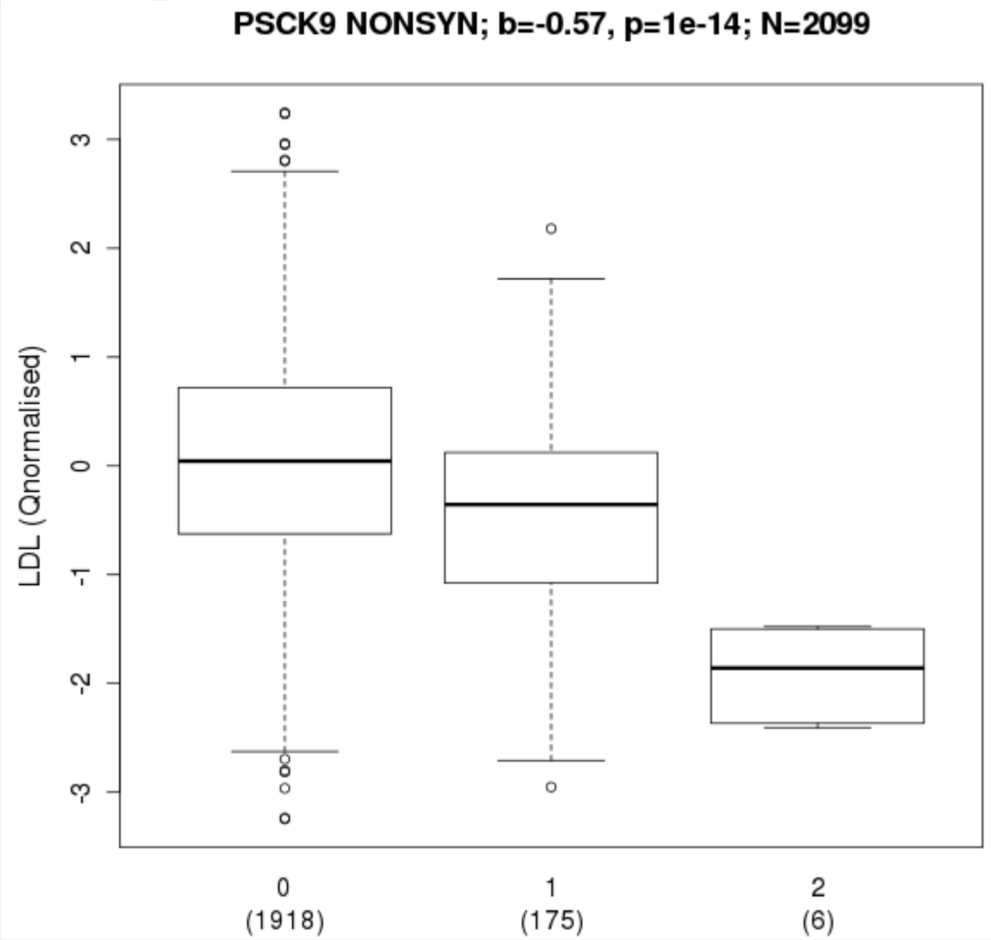


sigle gene

Jolie Effect

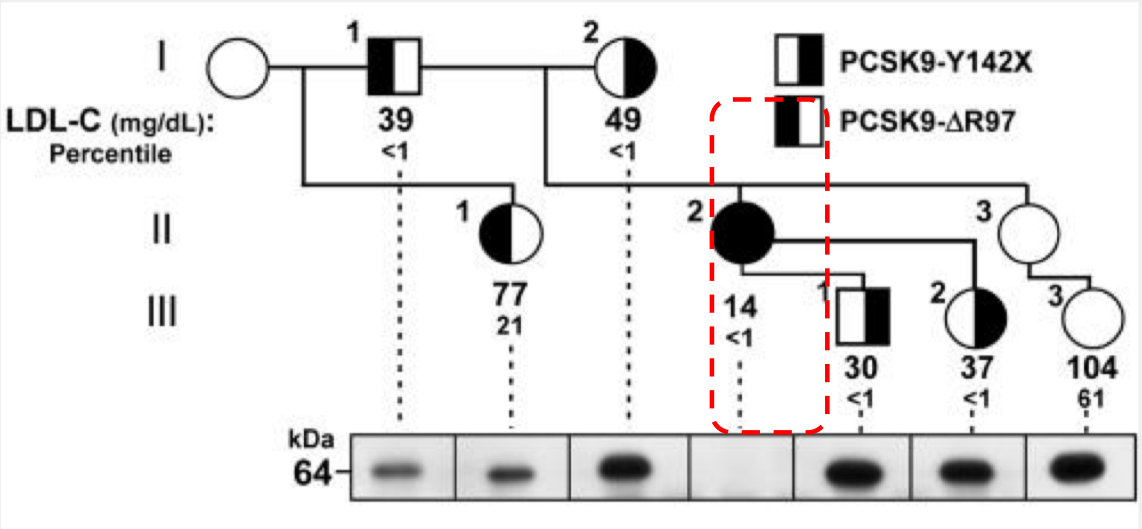


GENETIC VARIANT "RS11591147" IN PCSK9



- Carriers of T variant have lower levels of LDL cholesterol than carriers of G variant
- LDL is a strong risk factor for heart disease

A HUMAN KNOCK-OUT OF PCSK9 (2006)

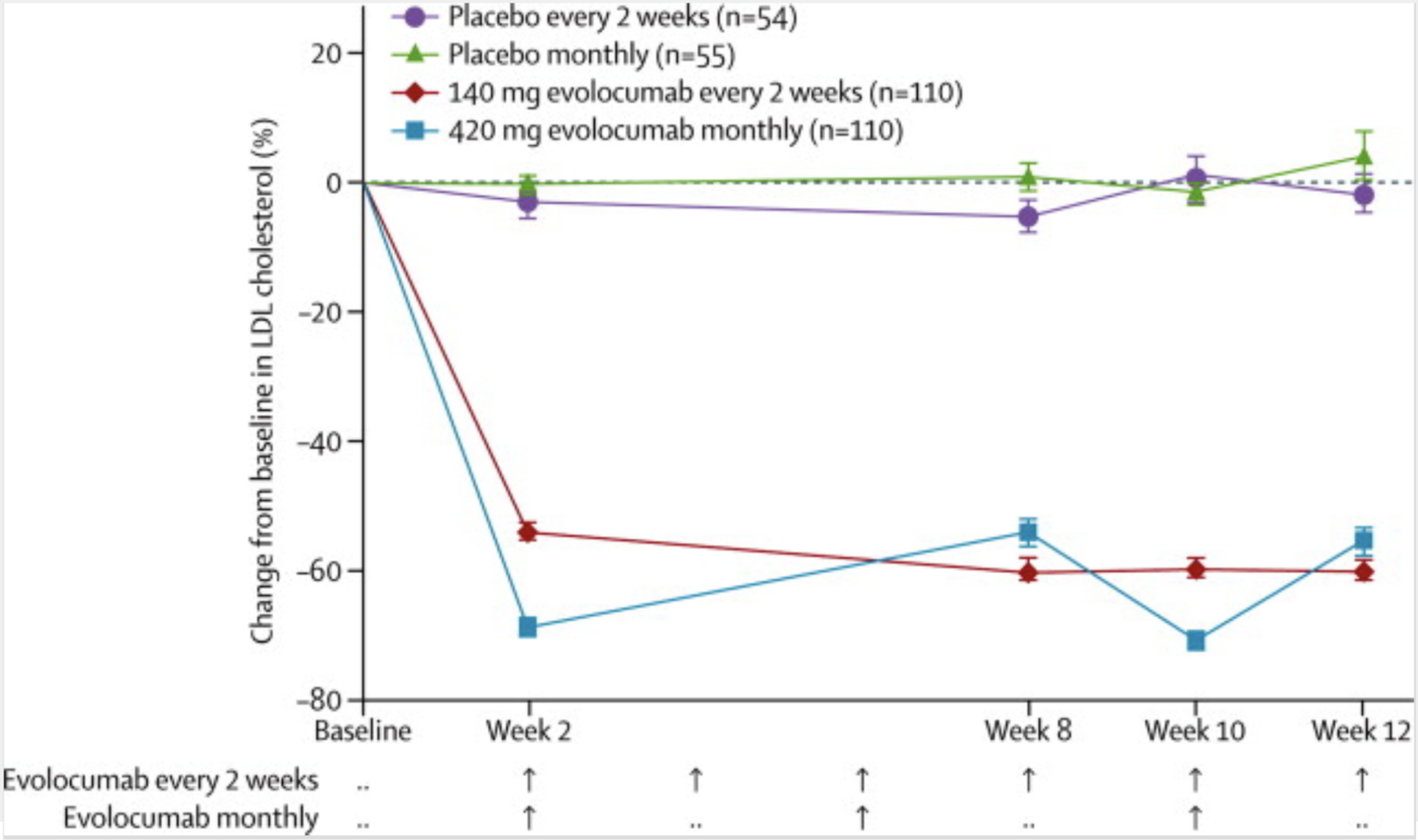


Individual II.2 has zero working copies of PCSK9 gene

- no circulating PCSK9 and an LDL-C of only 14 mg/dL
- apparently healthy, fertile, normotensive, college-educated woman with normal liver and renal function tests who works as an aerobics instructor
- Why is this very interesting observation? Inhibiting PCSK9 might be a safe way to reduce LDL

Zhao et al. AJHG 2006

PCSK9 inhibition with evolocumab (AMG 145) in heterozygous familial hypercholesterolaemia (RUTHERFORD-2): a randomised, double-blind, placebo-controlled trial



Lancet Oct 2014

ORIGINAL ARTICLE

N Engl J Med 2017; 376:1713-1722

Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease

Marc S. Sabatine, M.D., M.P.H., Robert P. Giugliano, M.D., Anthony C. Keech, M.D., Narimon Honarpour, M.D., Ph.D., Stephen D. Wiviott, M.D., Sabina A. Murphy, M.P.H., Julia F. Kuder, M.A., Huei Wang, Ph.D., Thomas Liu, Ph.D., Scott M. Wasserman, M.D., Peter S. Sever, Ph.D., F.R.C.P., and Terje R. Pedersen, M.D. for the FOURIER Steering Committee and Investigators*

FDA Approves Amgen's Repatha (evolocumab) to Prevent Heart Attack and Stroke

Dec 1 2017

In the Repatha cardiovascular outcomes study (FOURIER), Repatha reduced the risk of heart attack by 27%, the risk of stroke by 21% and the risk of coronary revascularization by 22%..

precision medicine in Taiwan

Use of HLA-B*58:01 genotyping to prevent allopurinol-induced severe cutaneous adverse reactions in Taiwan: national prospective study

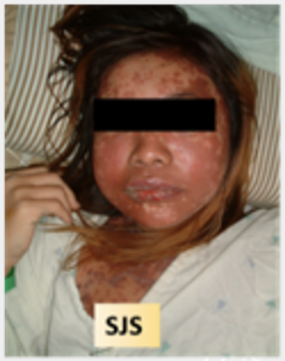
BMJ 2015;351:h4848 Sep 23, 2015

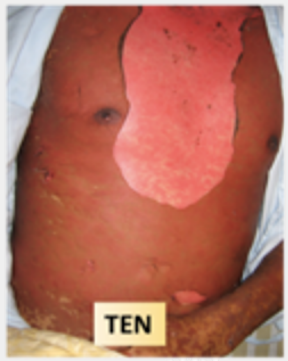
BMJ

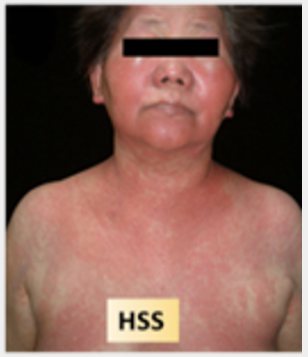
Gout

Allopurinol

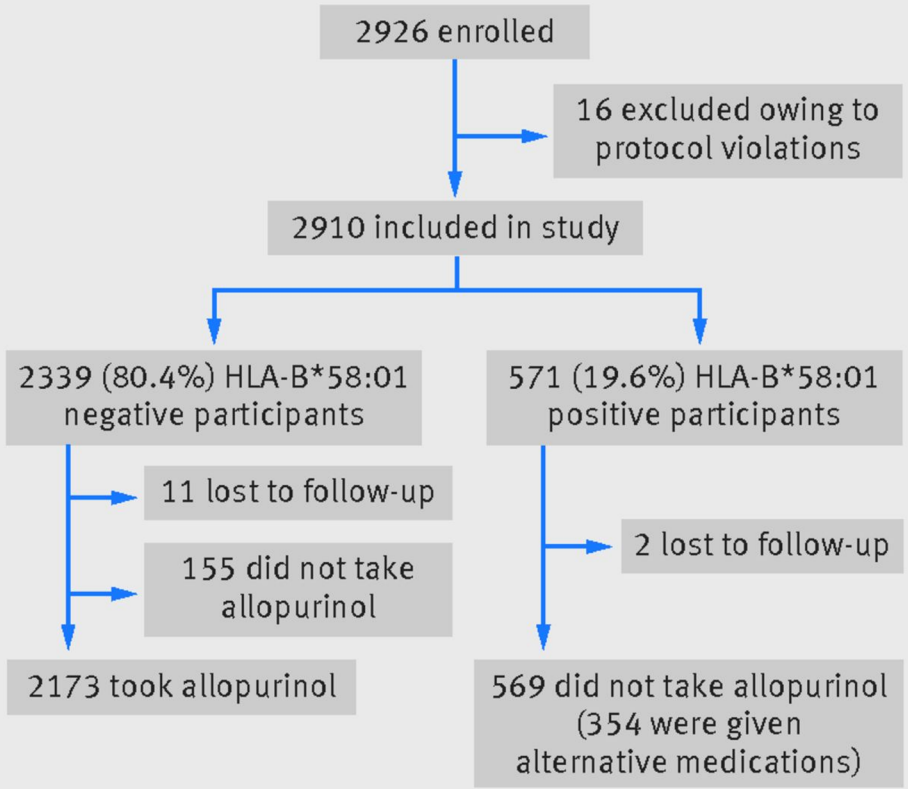
HLA-B*58:01







NHIRD-TWB → reduce health expenditures



Adverse event	HLA-B*58:01 positive participants receiving alternative drug treatment (n=354)	HLA-B*58:01 negative participants receiving allopurinol (n=2173)	Total (n=2910)
Mild cutaneous events			
Rash and itching	3*	94	97
Blisters	0	0	0
Oral ulcers	0	2	2
Rash, itching, oral ulcers, and fever	0	1	1
Rash, itching, and other adverse events	0	22	22

Records from the **National Health Insurance research database**

Incidence of SCARs ~0.3%

BMJ. 2015 Sep 23;351:h4848.

frequency distribution of pharmacogenetic phenotypes predicted by genotypes of TWB cohort

Gene	Drug	Rx[a]/year	EM	IM	PM	ADR[b] carrier rate
CYP2B6	Efavirenz	1,662,525	66.0%	30.5%	3.6%	
CYP2C19	Clopidogrel	63,664,076	39.8%	56.4%	3.8%	
CYP2C9	Celecoxib	65,058,810	93.6%	6.3%	0.1%	
CYP3A5	Tacrolimus	10,272,406	8.1%	40.6%	51.2%	
IL28	Peginterferon	40,941	88.6%	11.1%	0.3%	
NAT2	Isoniazid	7,885,251	28.8%	59.2%	12.0%	
SLCO1B1	Simvastatin	50,695,934	78.9%	19.9%	1.3%	
TPMT	Azathioprine	7,435,217	97.0%	2.9%	0.02%	
UGT1A1	Atazanavir	719,793	53.2%	39.8%	7.0%	
VKORC1	Warfarin	16,121,944	1.1%	19.2%	79.7%	
HLA-A*3101	Carbamazepine	17,078,849				2.0%
HLA-B*1502	Carbamazepine	17,078,849				4.1%
HLA-B*5701	Abacavir	3,049,217				0.2%
HLA-B*5801	Allopurinol	23,888,472				10.5%
MT-RNR1	Amikacin	321,561				4.7%

[a]Rx = prescriptions.

[b]*ADR* = adverse drug reactions.

NPJ Genom Med. 2021 Feb 11;6(1):10.

NOTCH3 cysteine-altering variant is an important risk factor for stroke in the Taiwanese population

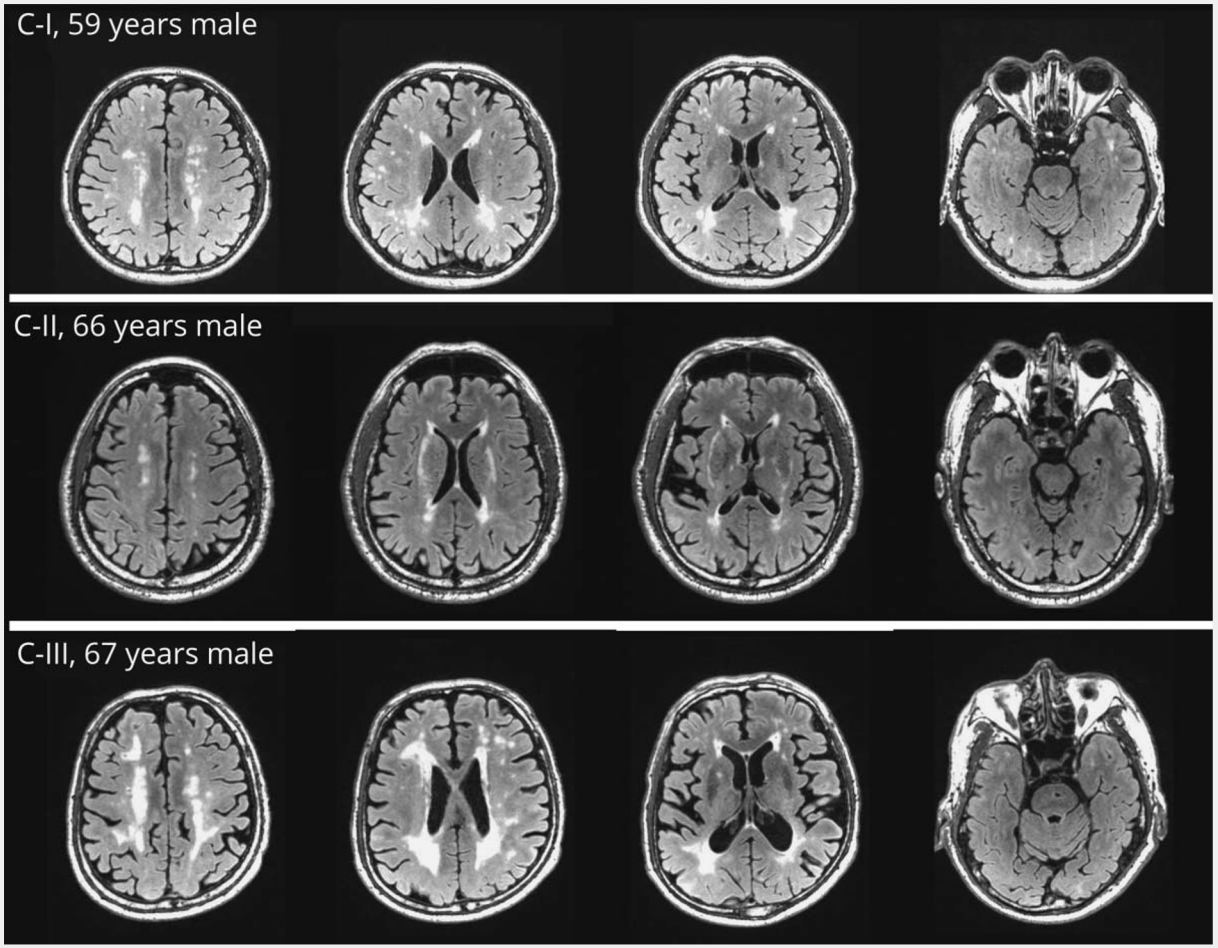
We queried the Taiwan Biobank database for cysteine-altering mutations in exons 2–24 of NOTCH3 within these genomes. The reference coding sequence of NOTCH3, NM_000435.3, was used for annotating the variants. (p.R544C (c.1630G>A), p.C853Y (c.2558G>A), and p.C884Y (c.2651G>A))

The cysteine-altering NOTCH3 variants identified from the Taiwan Biobank database were genotyped in the control participants and patients with stroke using the TaqMan genotyping assay

Only the NOTCH3 p.R544C variant was found in 4 individuals (TP-VGH (n =550))

	Controls			Patients with stroke		
	Taiwan biobank (n = 6,488)	TP-VGH (n = 550)	Biobank + TP-VGH (n = 7,038)	TP-VGH (n = 350)	TC-VGH (n = 450)	TP-VGH + TC-VGH (n = 800)
Male	2,293 (35.3)	230 (41.8)	2,523 (35.8)	248 (70.9)	327 (72.7)	575 (71.9)
Age, y	48.4 ± 10.9	56.7 ± 15.0	49.1 ± 11.5	64.4 ± 13.6	67.5 ± 13.2	66.2 ± 13.5
Hypertension	682 (10.6)	186 (33.8)	868 (12.4)	248 (70.9)	329 (73.1)	577 (72.1)
Diabetes	300 (4.6)	72 (13.1)	372 (5.3)	130 (37.1)	168 (37.3)	298 (37.3)
Hyperlipidemia	388 (6.0)	138 (25.1)	526 (7.5)	129 (36.9)	181 (40.2)	310 (38.8)
Smoking habit	1,228 (18.9)	98 (17.8)	1,326 (18.8)	122 (34.9)	136 (30.4)	258 (32.4)
Alcohol consumption	387 (6.0)	108 (19.6)	495 (7.0)	56 (16.0)	87 (23.2)	143 (19.7)
Family history of stroke	1,146 (17.8)	—	—	40 (14.1)	14 (4.0)	54 (8.5)
***NOTCH3* p.R544C mutation (+)**	56 (0.9)	4 (0.7)	60 (0.9)	8 (2.3)	9 (2.0)	17 (2.1)

Physical examination revealed that they were free of neurologic deficits. Three of them received brain MRI scans, and all had a variable degree of leukoencephalopathy



multi-gene

GWAS study

- **偏頭痛及阿茲海默氏症之全基因體關聯性研究**-台北榮總
- **利用全基因關聯性研究於影響人類身高差異性之基因鑑定與功能性探討**-中國醫藥大學
- **利用台灣生物資料庫探討成人肺功能變化之全基因體關聯研究**-中央研究院

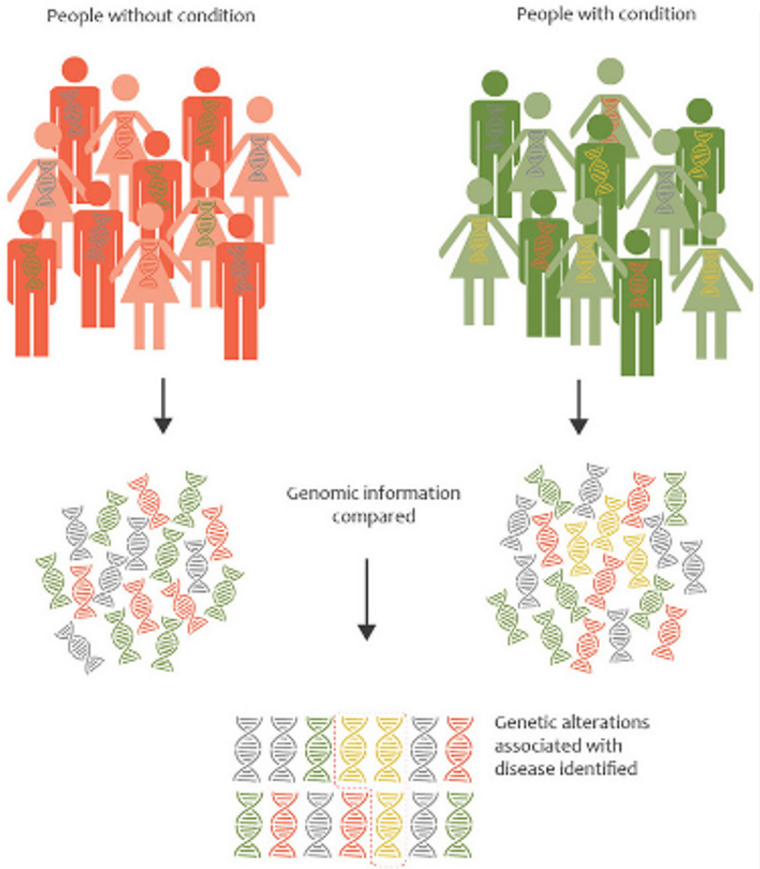
  - Calculate polygenic risk score (PRS) for individual j using m SNPs

$$\mathrm{PRS}_j = \frac{\sum_{i=1}^{m} \ln(OR_i) \times SNP_{ij}}{m}$$

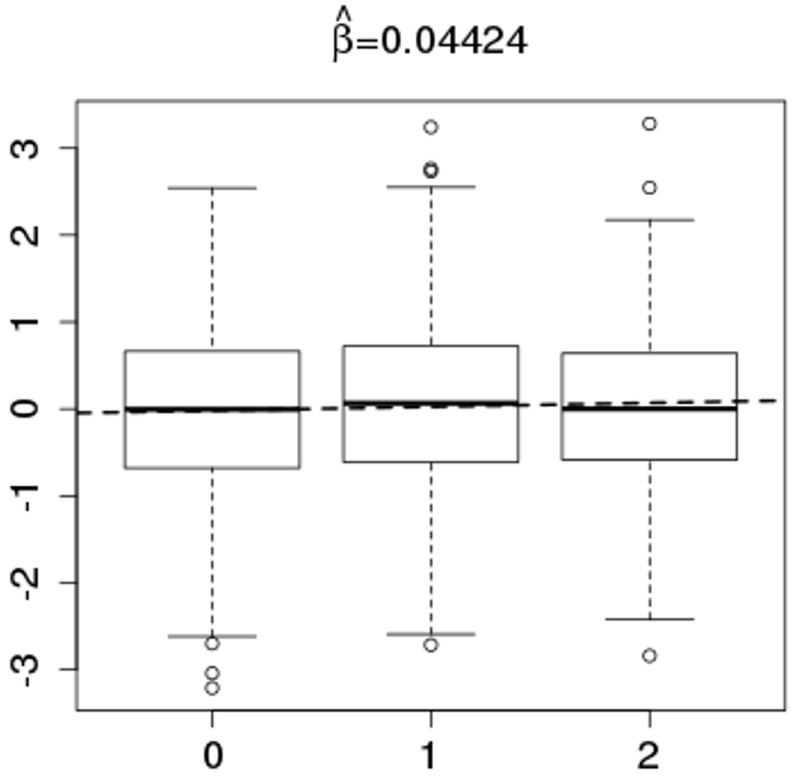
where

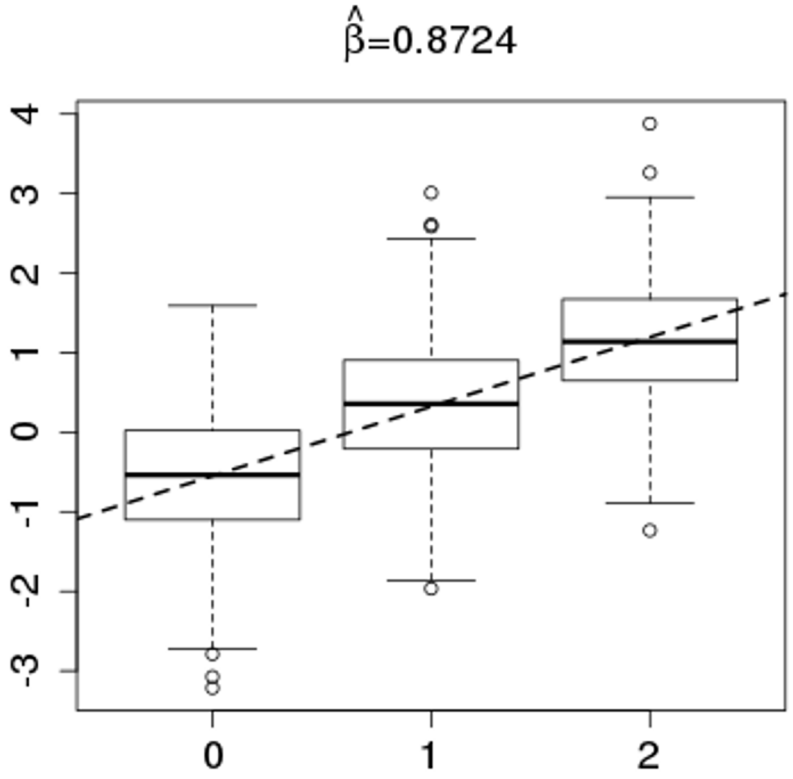
- $\ln(OR_i)$ = effect size for SNP i from discovery sample
- SNP_{ij} = number of risk alleles (0,1,2) for SNP i, individual j in target sample
- m = number of SNPs considered in test set

Purcell / ISC et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder Nature 2009

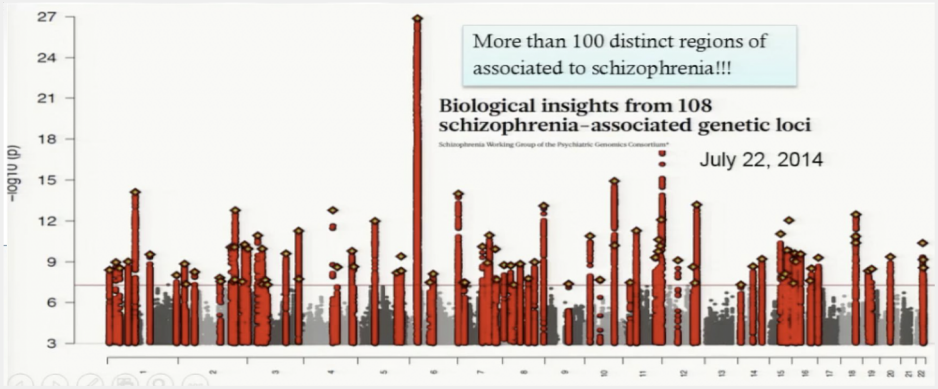


association study





the power of sample size - schizophrenia | psychiatric genomics consortium



polygenic risk score

Discovery GWAS

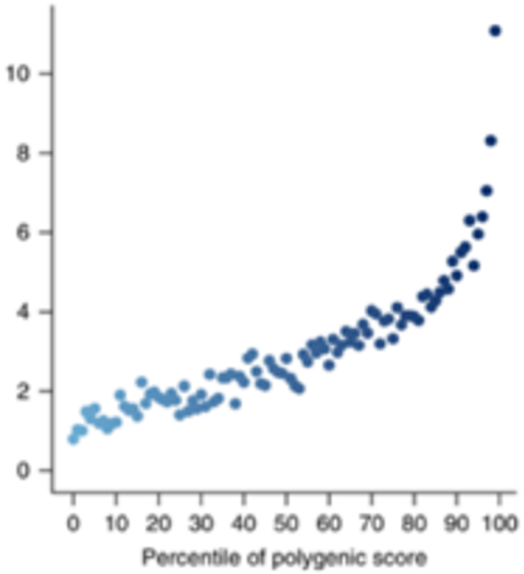
	Weight*	Risk Allele
SNP1	0.2	A
SNP2	-0.3	C
SNP3	0.1	G

Individual	Alleles SNP1	Alleles SNP2	Alleles SNP3
1	AT	AA	CG
2	AA	CA	GG
3	TT	AC	CG
4	TT	AA	GG
5	TA	CA	GC
6	AT	CA	CG
7	AA	AA	GG
8	AA	CC	CG
9	TA	CC	GC
10	AT	AA	CG

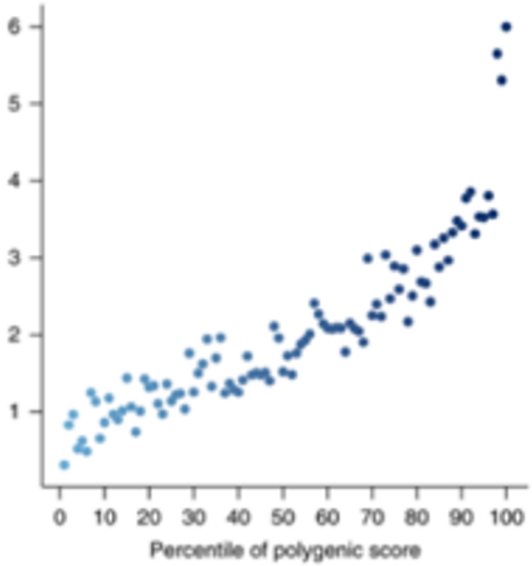
PRS:

Individual	SNP 1	SNP 2	SNP 3	**PRS**
1	0.2+0.0	0.0+0.0	0.0+0.1	**0.3**
2	0.2+0.2	-0.3+0.0	0.1+0.1	**0.3**
3	0.0+0.0	0.0-0.3	0.0+0.1	**-0.2**
4	0.0+0.0	0.0+0.0	0.1+0.1	**0.2**
5	0.0+0.2	-0.3+0.0	0.1+0.0	**0.0**
6	0.2+0.0	-0.3+0.0	0.0+0.1	**0.0**
7	0.2+0.2	0.0+0.0	+0.1+0.1	**0.6**
8	0.2+0.2	-0.3-0.3	0.0+0.1	**-0.1**
9	0.0+0.2	-0.3-0.3	0.1+0.0	**-0.3**
10	0.2+0.0	0.0+0.0	0.0+0.1	**0.3**

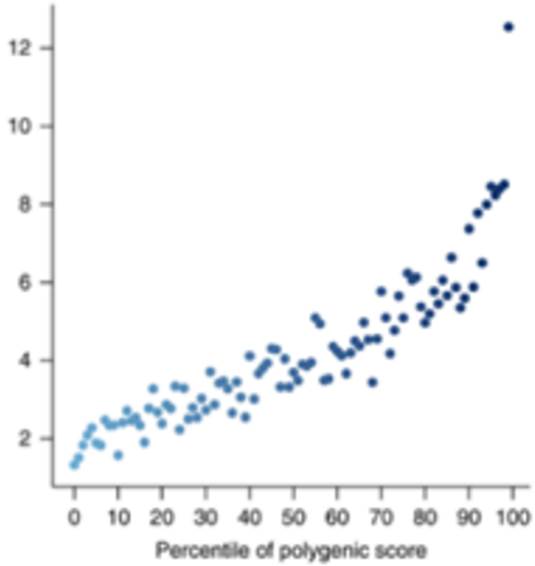
disease risk prediction



Prevalence of Coronary Artery Disease



Prevalence of Type 2 Diabetes

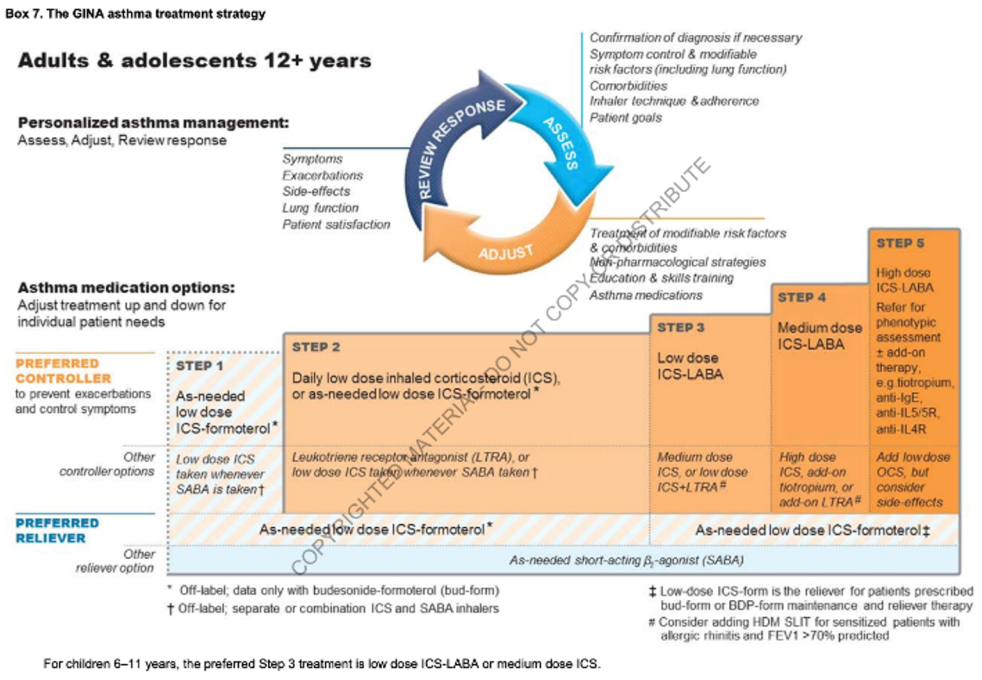


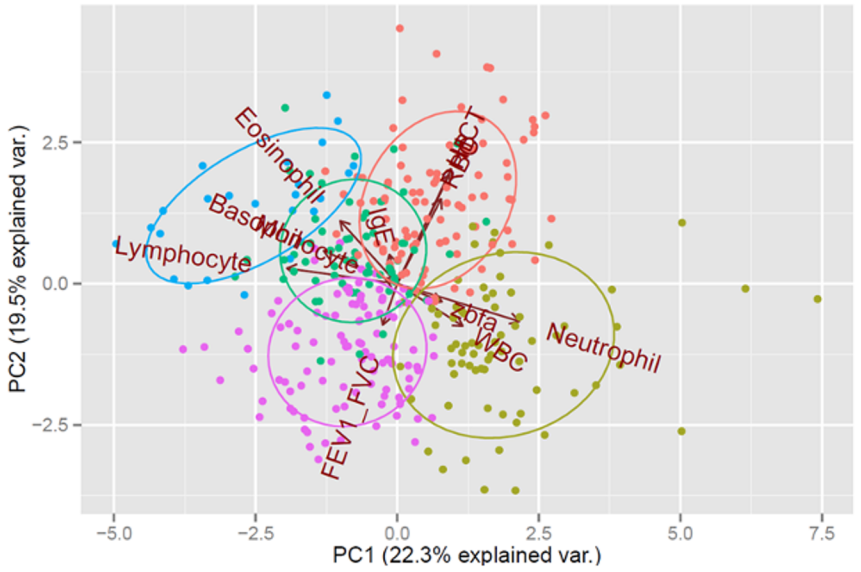
Prevalence of Breast Cancer

"…it is time to contemplate the inclusion of polygenic risk prediction in clinical care, and discuss relevant issues."

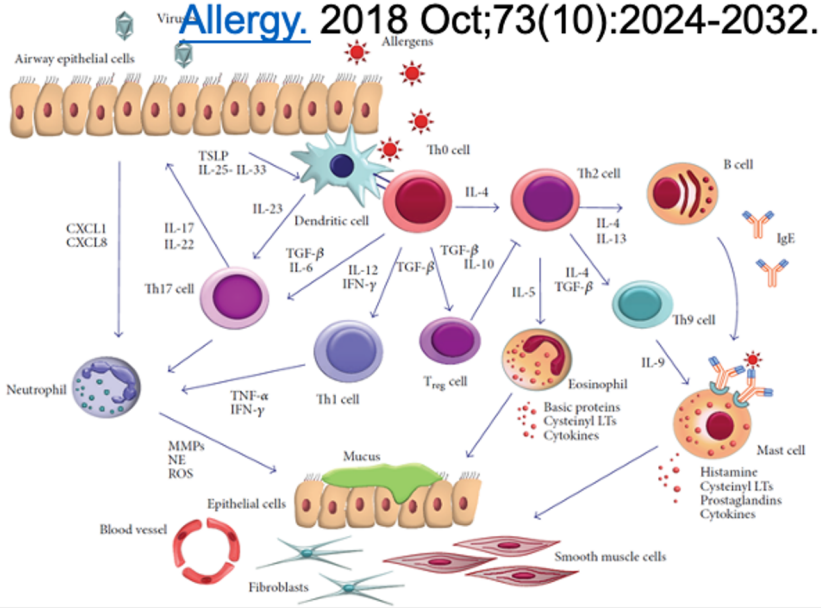
Khera et al. Nature Genetics 50, 1219–1224 (2018)

Precision medicine



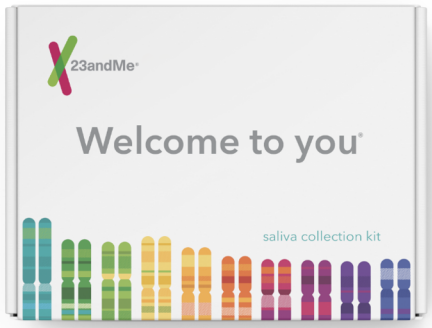


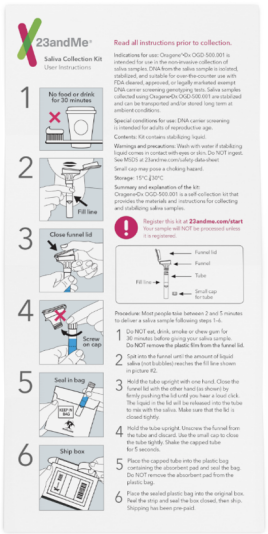
Allergy. 2018 Oct;73(10):2024-2032.



Nature 520, 609–611

Direct to consumer genetic testing (DTG)

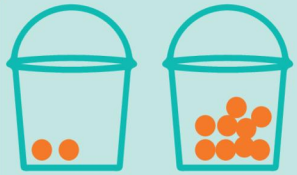




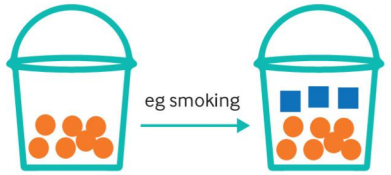
For most health conditions (like diabetes or cancers), each person has some risk of developing the condition during their life. This diagram represents this as a bucket – a person experiences the condition when their bucket gets full up.

Most health conditions arise due to a combination of genetic and other risk factors.

Genetic factors Other factors

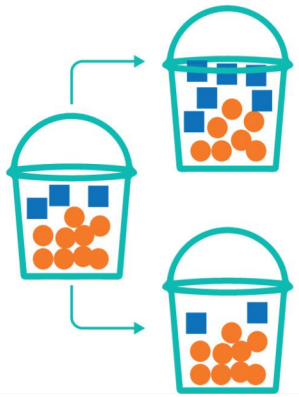


Everyone starts life with some genetic risk factors for the condition. Some people have very few, some people have a lot. You are born with this risk and you can't change it. Direct-to-consumer genetic tests aim to measure this risk (though their measurement is often far from perfect).



Over the course of life, people are exposed to other factors that increase the chance of them experiencing the condition.

If you had more genetic risk factors in your bucket to begin with, it takes fewer other risk factors to fill the bucket and result in the health condition.



Some of the other factors, you can't change, like your age. Some factors, you might be able to work on, like smoking, or alcohol intake.

Even if you have a 'high genetic risk' of a particular condition, for most health conditions, having a high genetic risk does not mean you'll get it.

By working on the factors you can change, you might be able to reduce the chance of developing a health condition, even if you have a high genetic risk.

BMJ 2019;367:l5688

Need for a Precision Health Eco-system

Economically Unsustainable
Current healthcare and human services industries

Treatment Decisions
Gap in accuracy

Evidence-Based Medicine
<50% is actually evidence-based

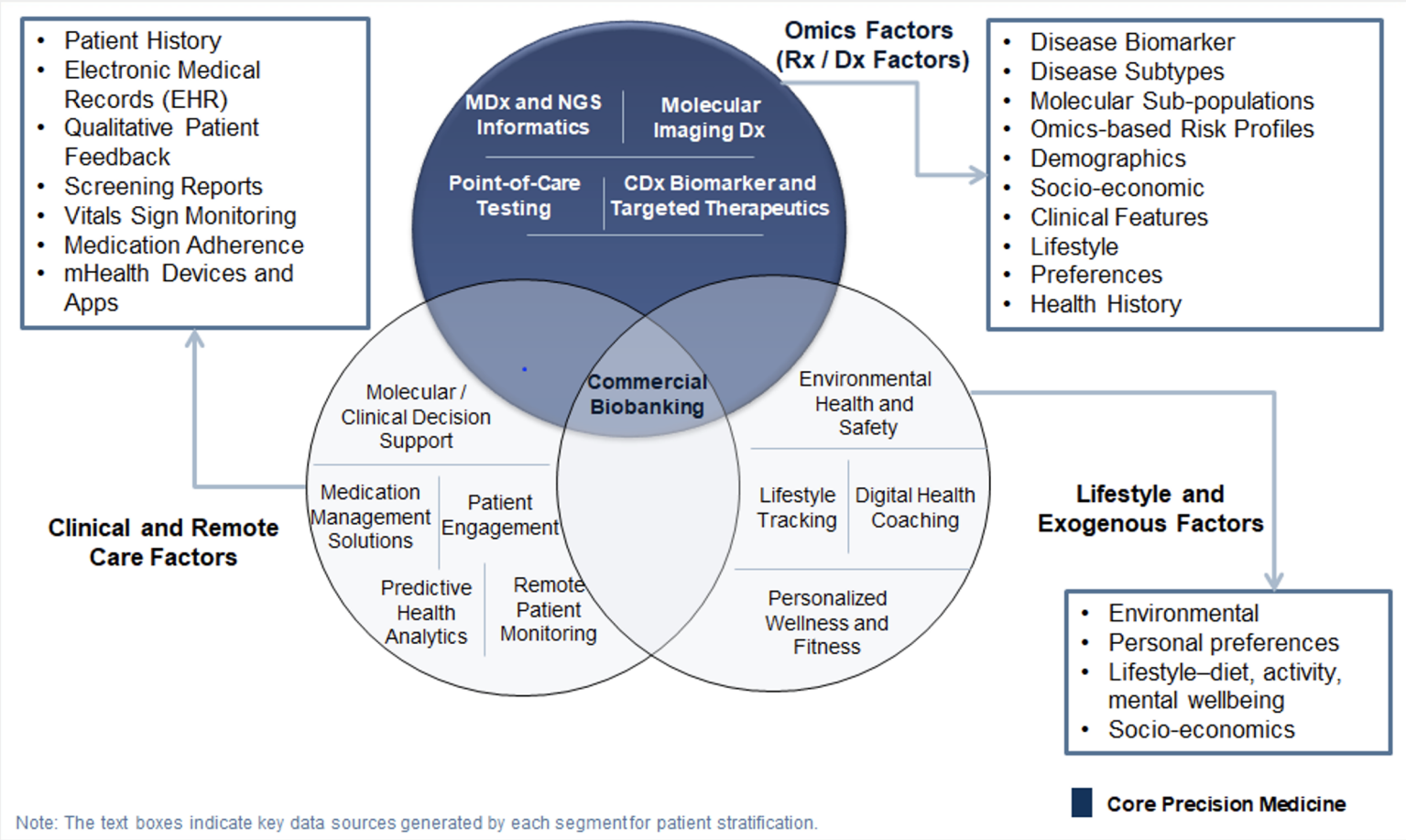
Chronic and Aging Populations
Require more care

Health Spending
Disconnect with outcomes

The problem is one of precision and productivity, not of diseases, gadgets or healthcare

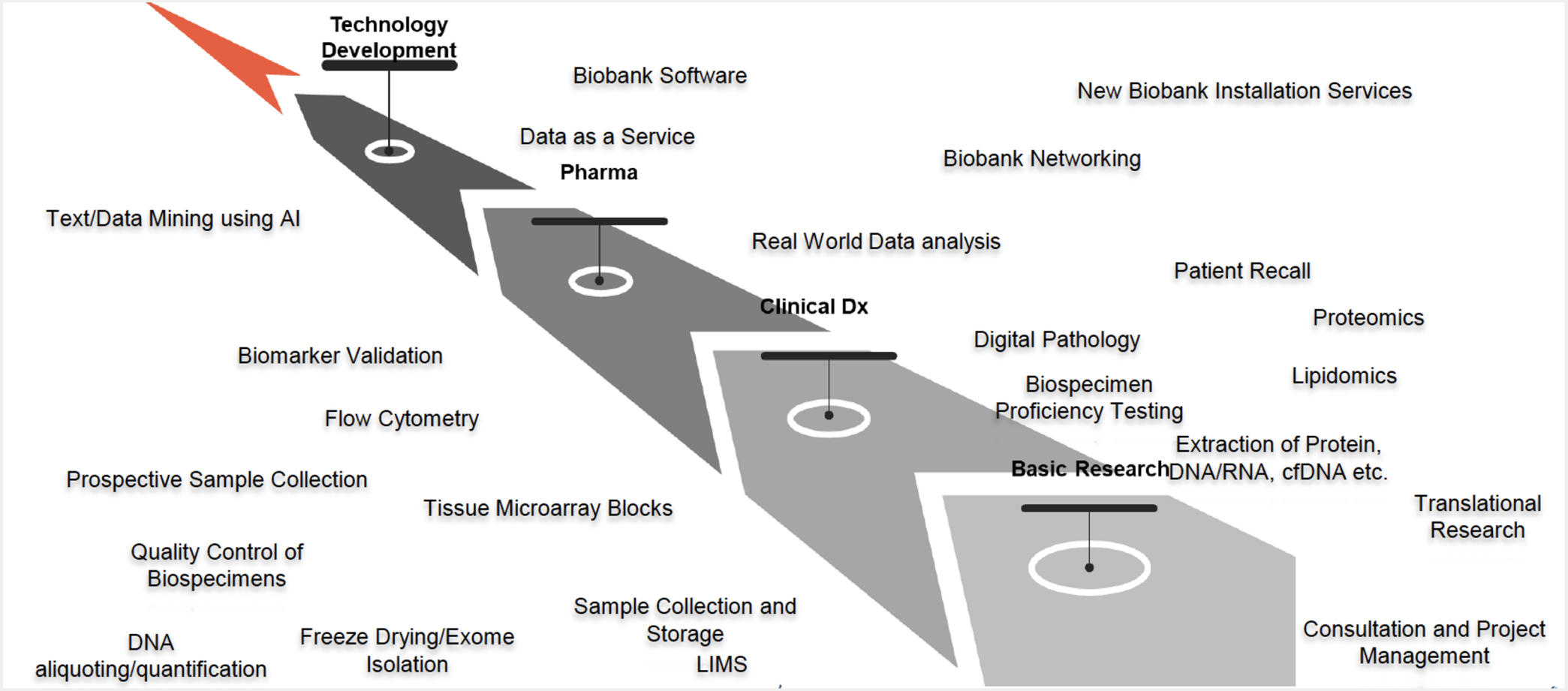
Now the new vision of healthcare is about Precision-based Predictive, Preventive and Participatory care models promoting social and financial inclusion.

Biobank - provide clinical research support that translates into bedside diagnostics and treatments, and advances research technologies into clinical applications



Note: The text boxes indicate key data sources generated by each segment for patient stratification.

Biobank service



Lion man 35,000~40,000 years old