Deciphering Genetically Influenced Metabotypes for Gastric Cancer Risk Stratification and Targeted Primary Prevention

Presenter: Zong-Chao Liu, Ph.D. candidate

Ph.D. mentor: Prof. Wen-Qing Li

Department of Cancer Epidemiology

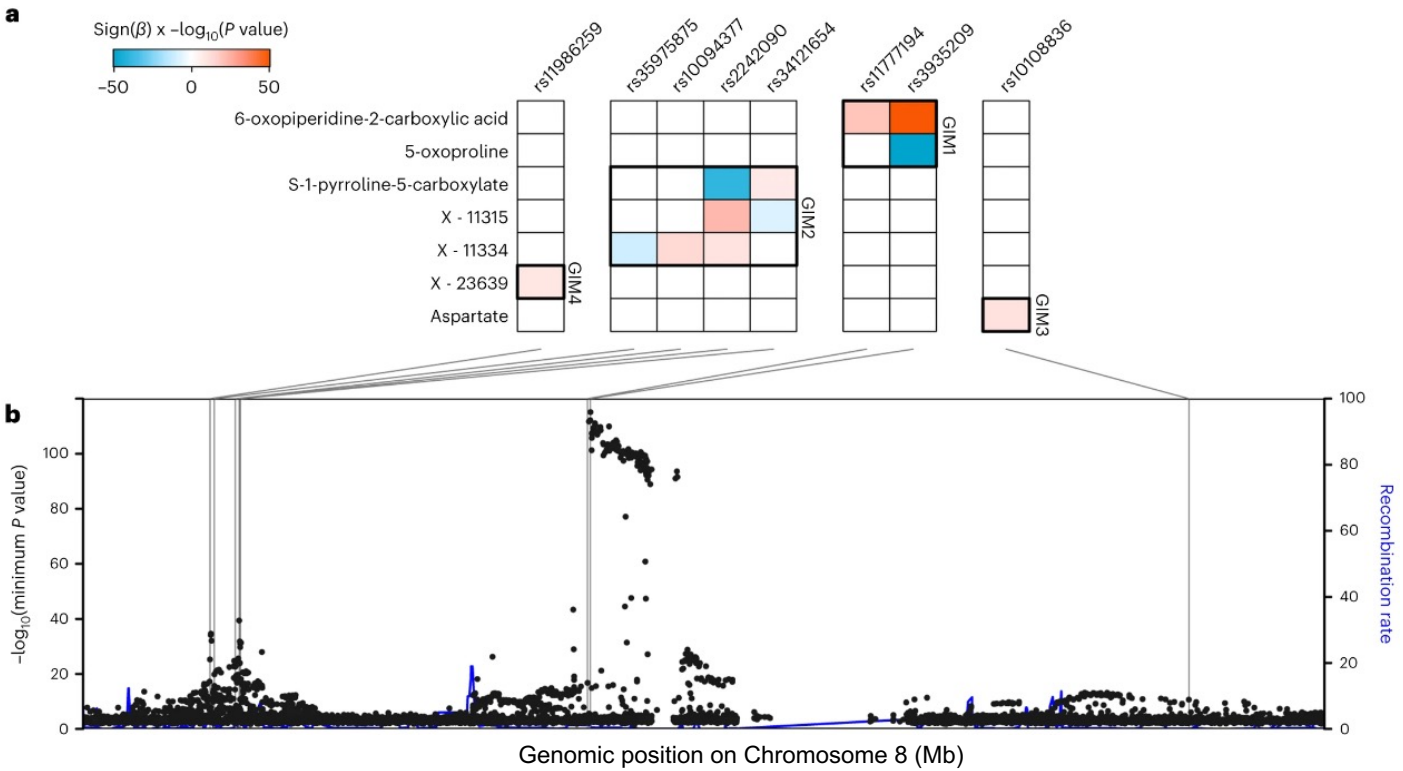
Peking University Cancer Hospital & Institute, China

北京大学
PEKING UNIVERSITY

Genetically influenced metabotypes (GIMs)

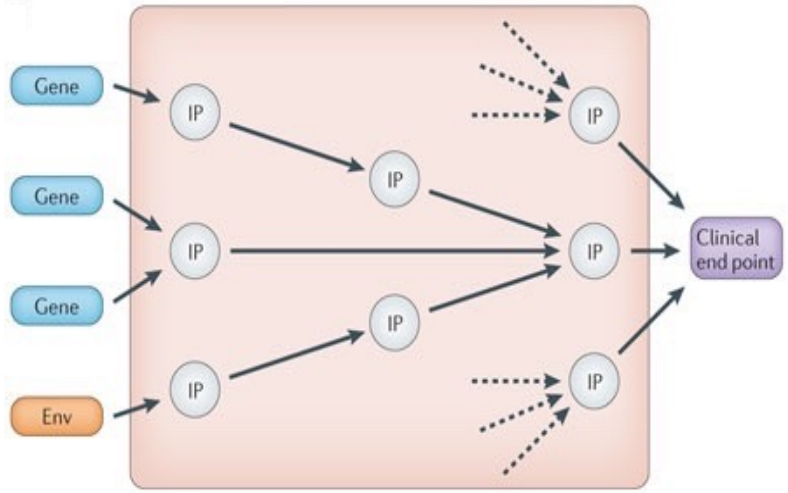
Definition of GIMS

Nat Med 28, 2321–2332 (2022).



Intermediate phenotypes

Nat Rev Genet 13, 759–769 (2012).

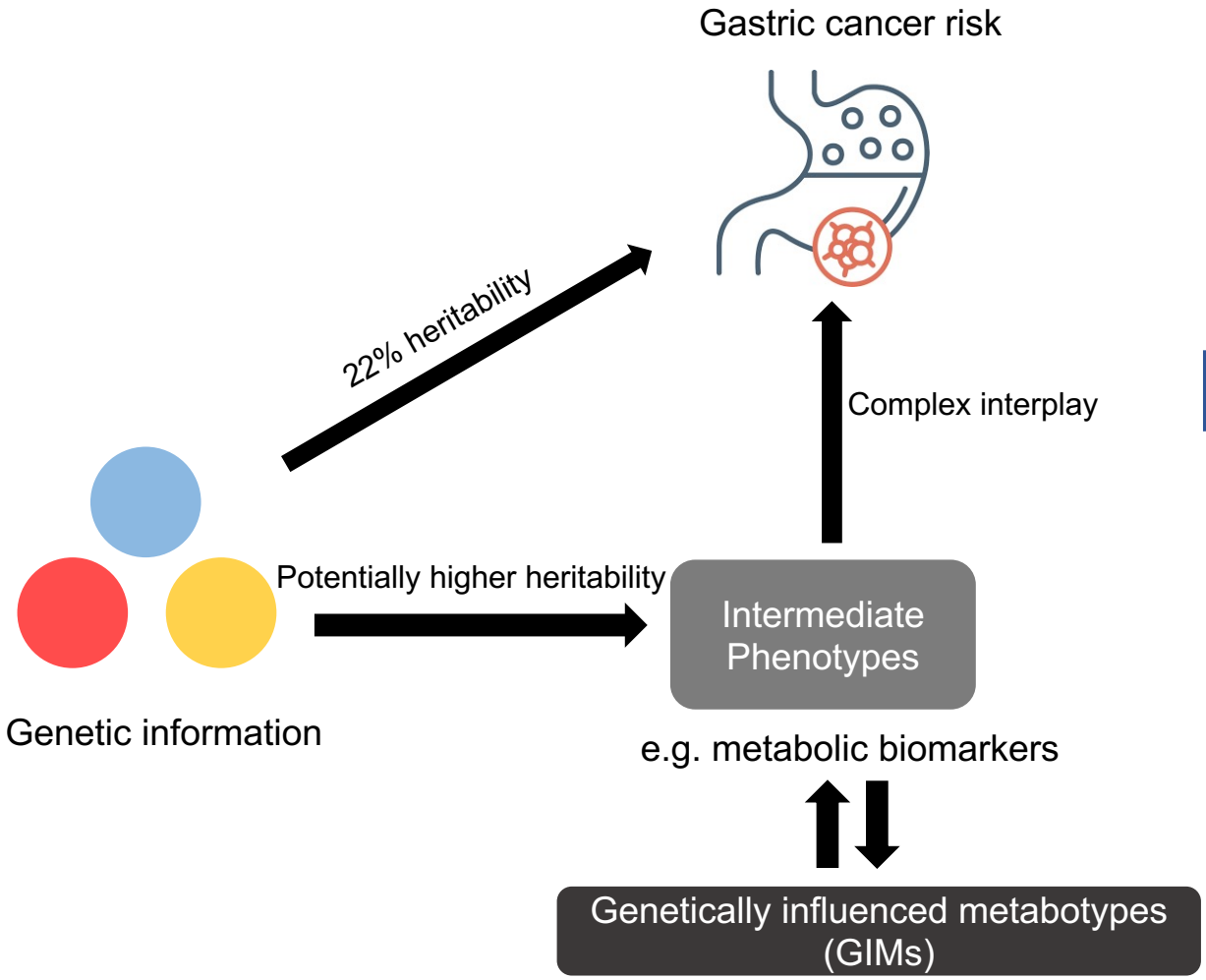


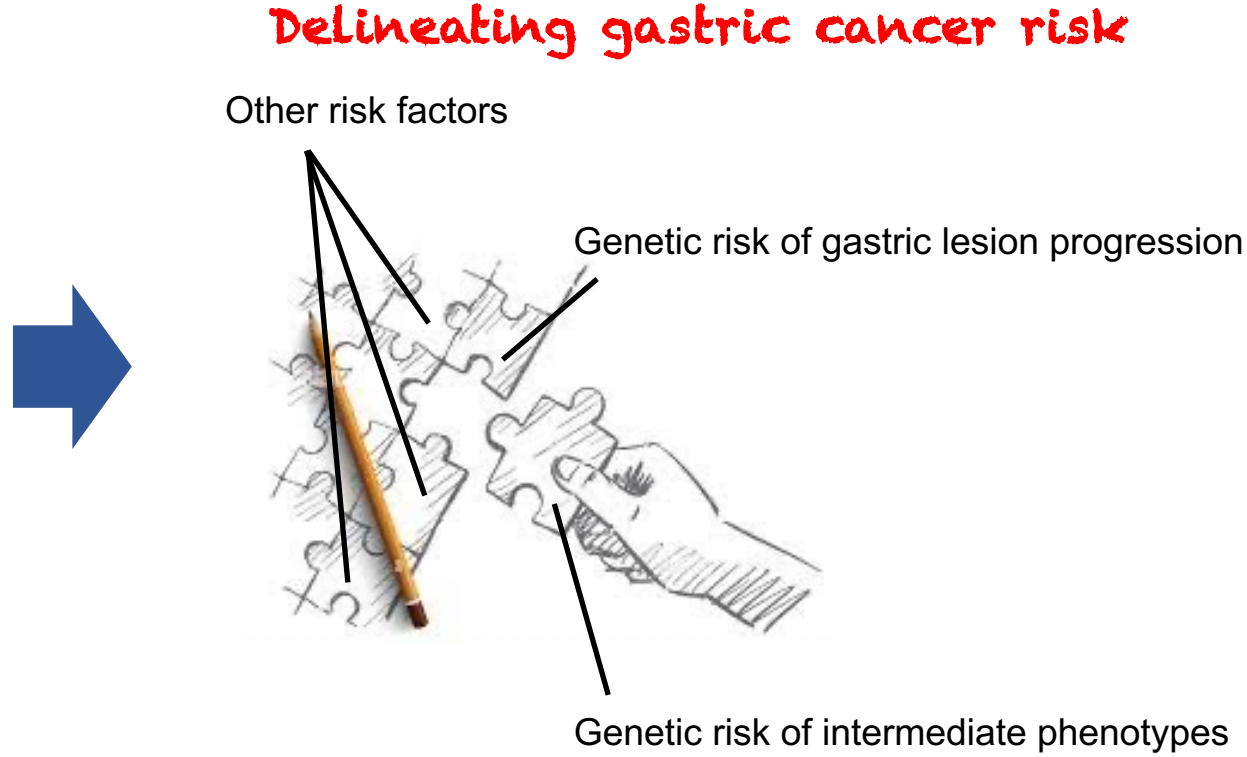
- Metabolic biomarkers are important intermediate phenotypes before diseases occur
- GIMs can be derived from metabolome GWAS

Genetically Influenced Metabotypes (GIMs) are defined by various groups of metabolite quantitative trait loci (mQTLs)

Hum Mol Genet 24(R1):R93-R101 (2015).

A different avenue for gastric cancer risk



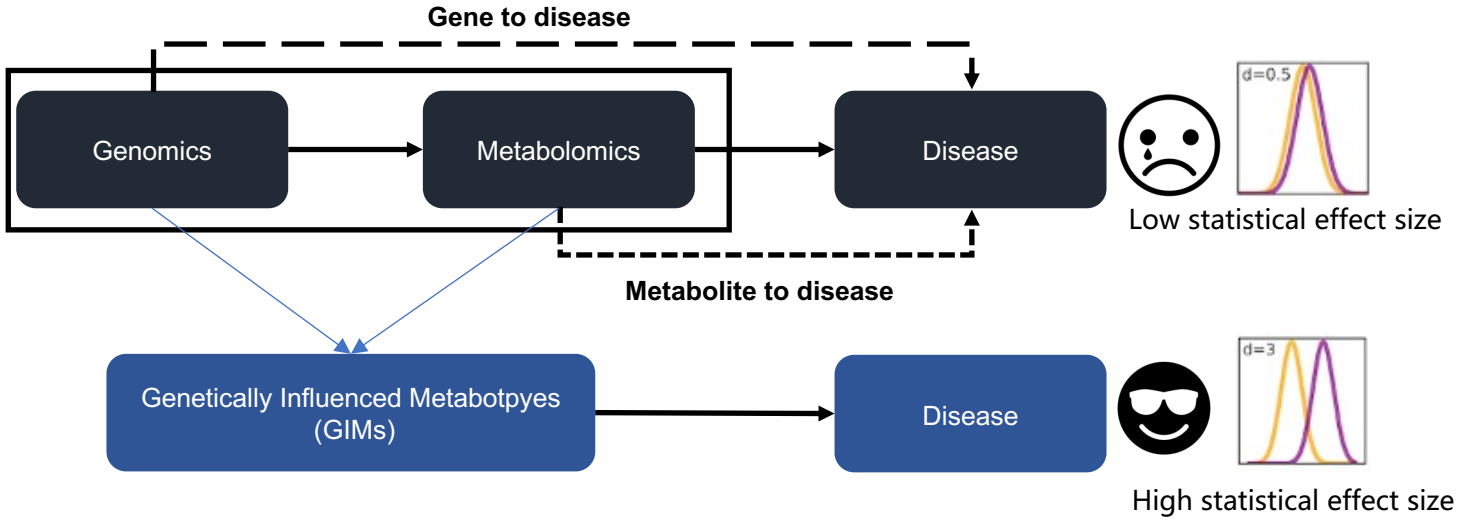


J Intern Med. 2023 Oct; 294(4):378-396.
JAMA. 2016;315(1):68-76

Why GIMs

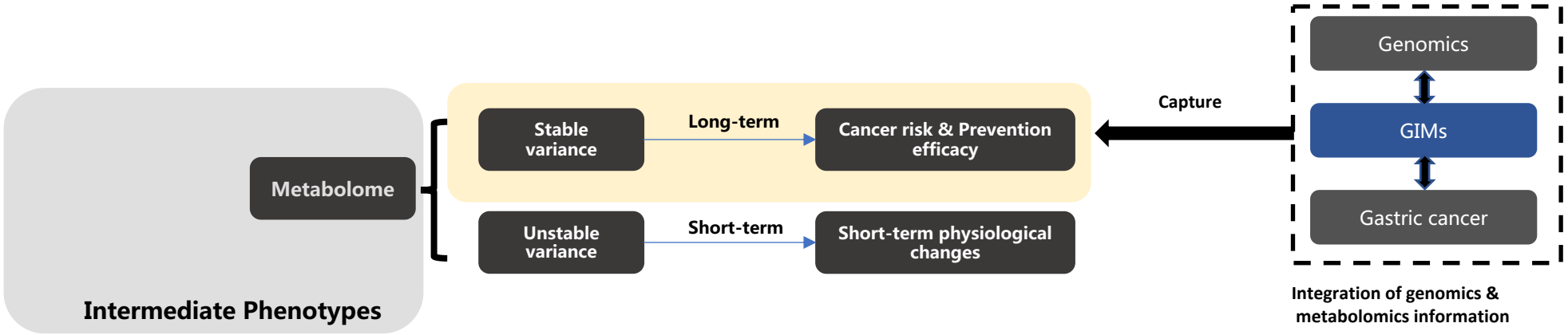
GIMs may exhibit higher statistical power than traditional disease trait loci

Nat Genet 46, 543–550 (2014)
Nat Med 28, 2321–2332 (2022)



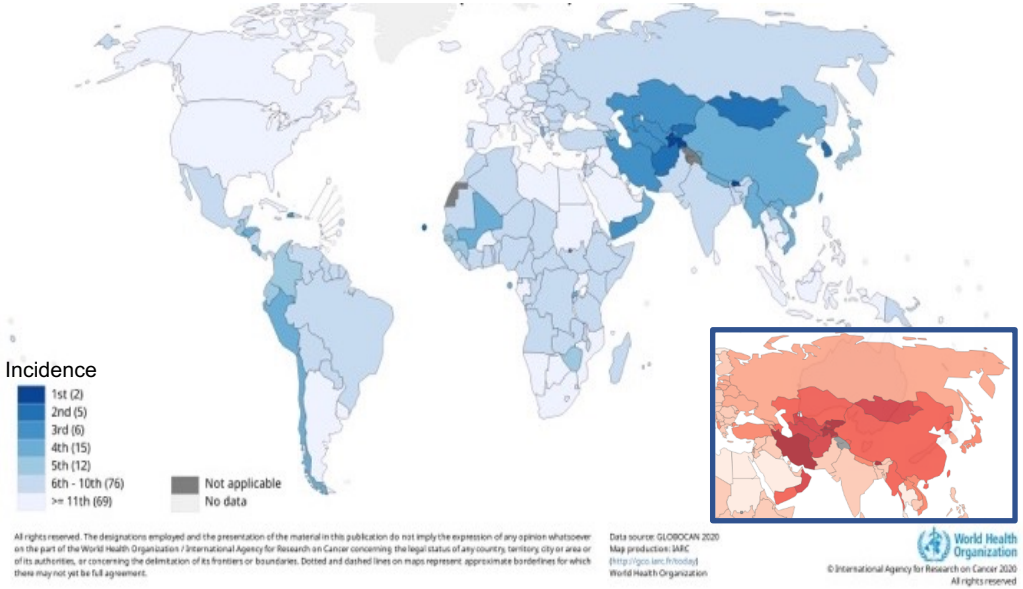
GIMs may capture stable variance of metabolic porfiles

Mol Syst Biol., 2011, 7(1): 525.

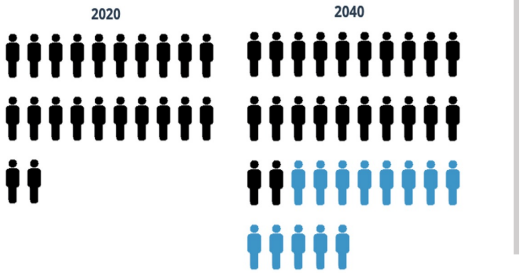


Gastric cancer

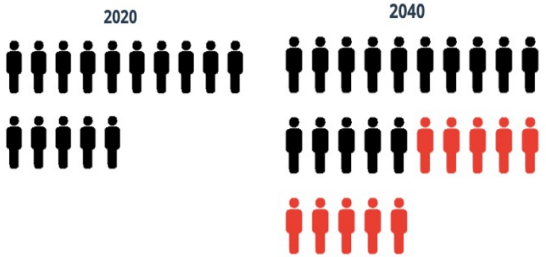
Epidemiologic characteristics



Estimated number of new cases



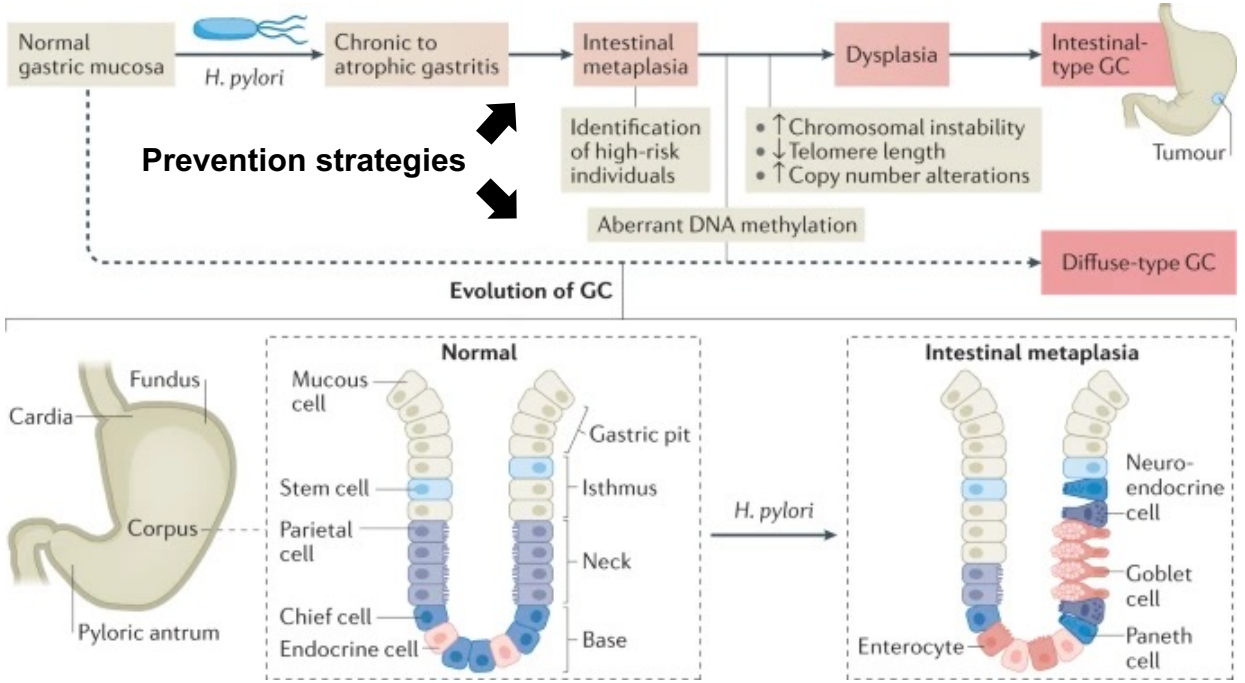
Estimated number of deaths



Risk factors

- *Helicobacter pylori (H. pylori)* infection ★
- Smoking
- Alcoholic consumption
- Low intake of fruit & vegetables
- High salt diet
- High intake of red & processed meats

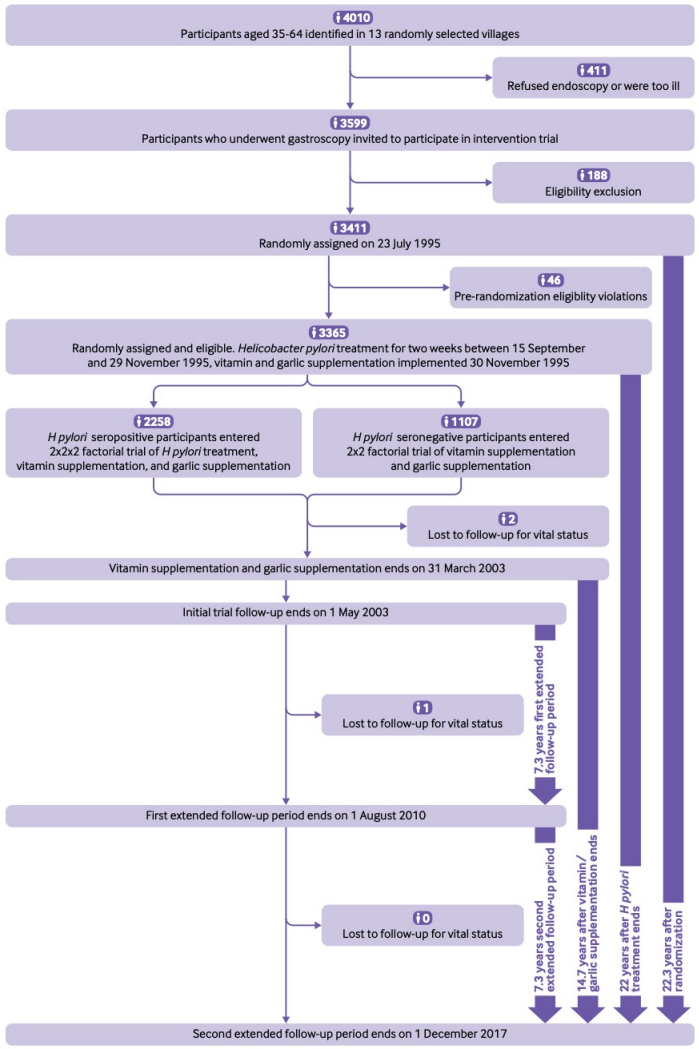
Evolution of gastric cancer (GC)



Nature reviews cancer, 22(2), 71-84.、

Efforts towards gastric cancer prevention

The Shandong Intervention Trial (SIT) Study



***H.pylori* positive**

(n=2,258)

2 x 2 x 2 factorial design

2 weeks *H.pylori* treatment

7.3 years Vitamin supplementation

7.3 years Garlic supplementation

***H.pylori* negative**

(n=1,107)

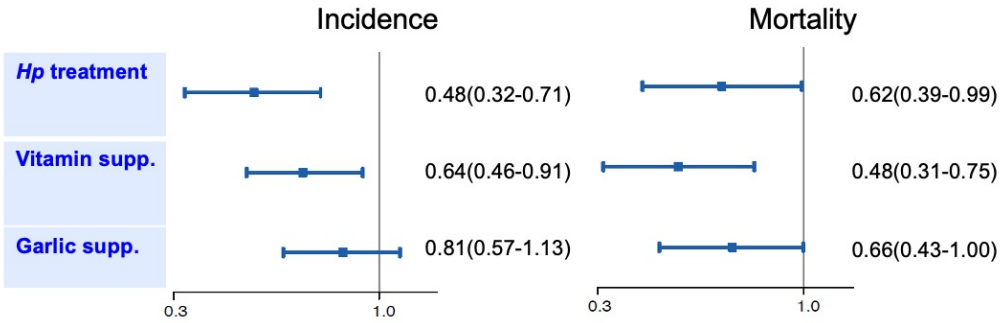
2 x 2 factorial design

7.3 years Vitamin supplementation

7.3 years Garlic supplementation

- High-risk area based
- N=3,365, 2x2x2 factorial design
- 2 weeks of *H.pylori* treatment
- 22.3 years of follow-up

BMJ 2019, 366.



Implications

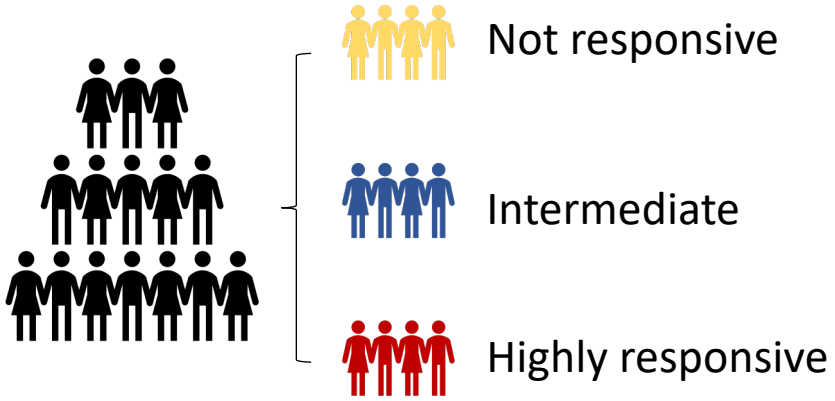
- ✓ ***H.pylori* eradication (2 weeks)**
- ✓ Vitamin supplementation (7 yrs)
- ✓ Garlic supplementation (7 yrs)

Incidence ↓

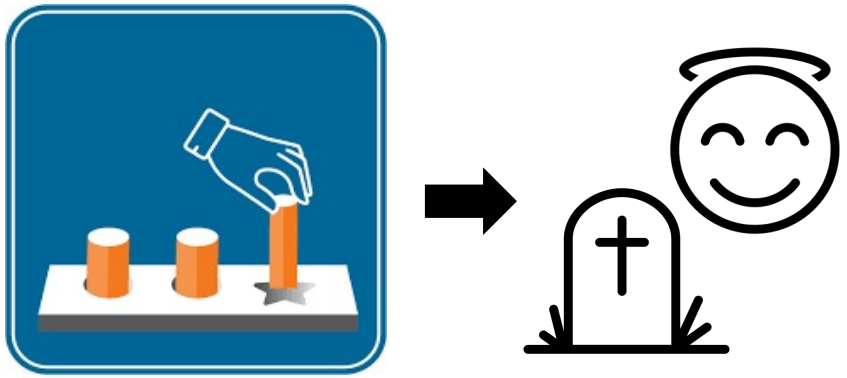
Mortality ↓

H.pylori treatment for two weeks and vitamin or garlic supplementation for seven years were significantly associated with a reduced **incidence** of GC, as well as a reduced risk of **death** due to GC for more than **22 years**.

Problems still exist



Implementation

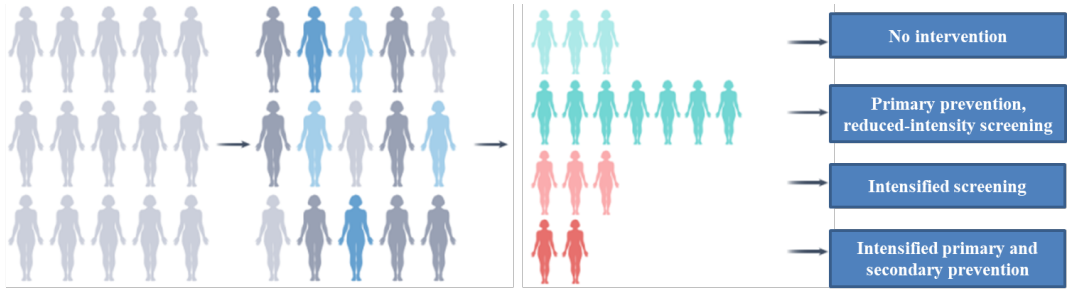


“One-size-fits-all” prevention strategies

Not all of the subjects are completely responsive !

e.g.

- **There's still a possibility of failure in *H.pylori* eradication.**
- **Long-term effects of gastric cancer prevention is only evident in certain subgroups.**



Fine-tuned personalized prevention strategies

- Under high risk of gastric cancer
- Benefits most from prevention

Identification of the target population

Research aim

Profile and identify the high-risk target population that might benefit most from early interventions

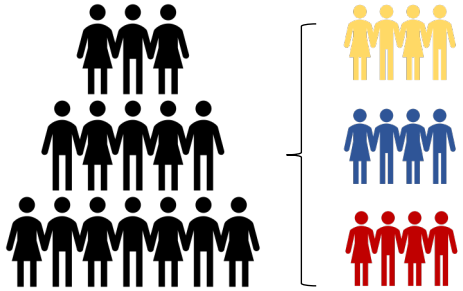
What kind of profiles are we going to depict?

- Genetically Influenced metabotypes (GIM)

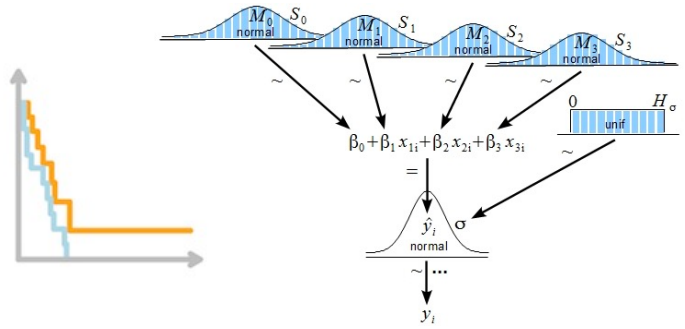
How do we derive the profiles?

- Integration of genomics and metabolomics information from large-scale datasets
- External validation in disease-specific cohorts

Integration of genomics & metabolomics

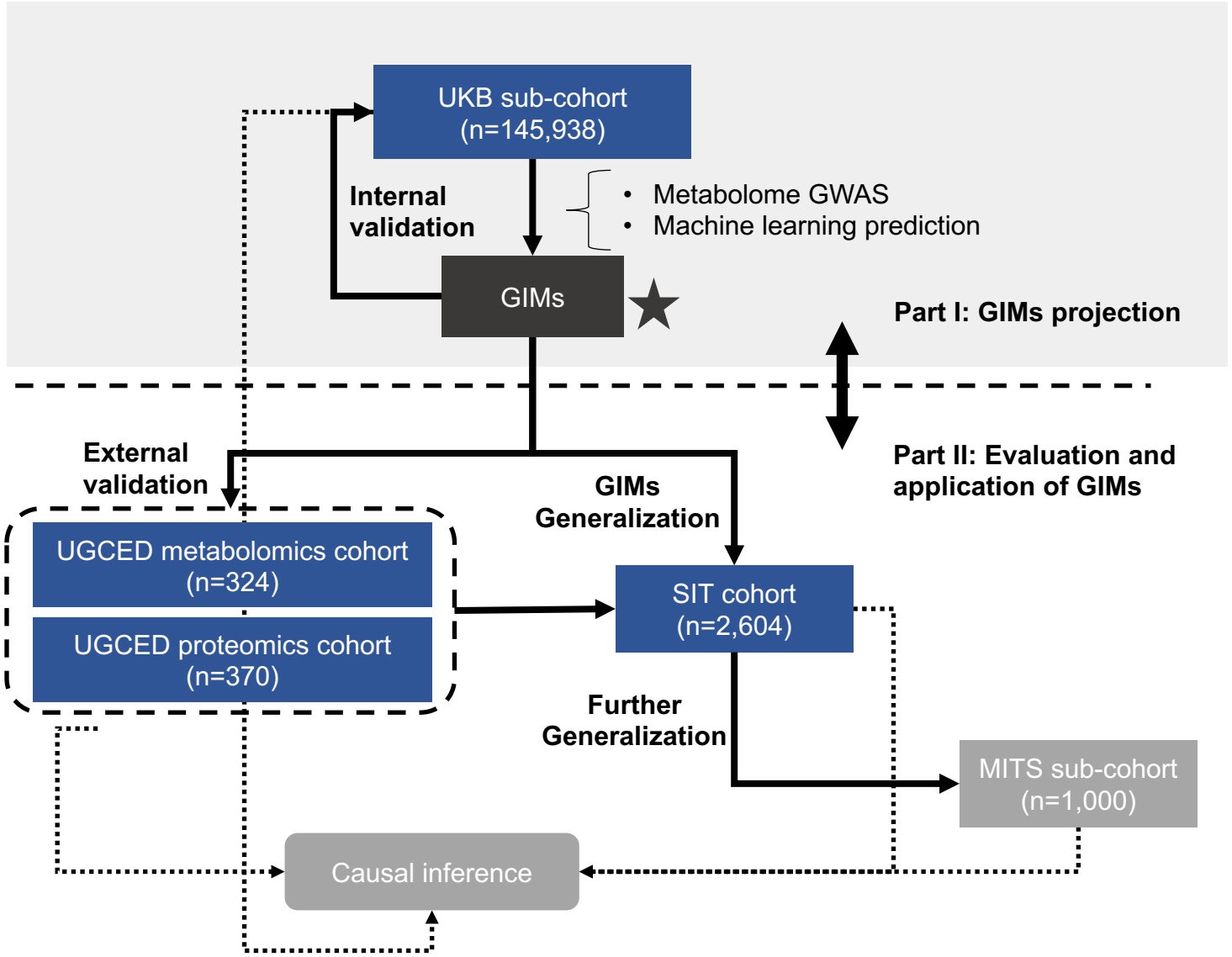


Projection of genetically Influenced metabotypes (GIMs)



Application & evaluation of GIMs for the primary prevention of GC

Study design & Methods



˚Genotypic information is avaliable in the UGCED and SIT and MITS sub-cohort

biobank uk
Enabling scientific discoveries that improve human health

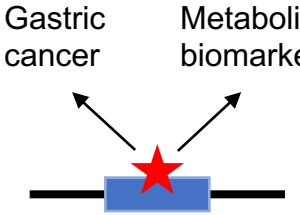
Establishing a cohort for gastric cancer research

- Must have both metabolomics and genotype data with quality control
- Match with ICD-10; Gastric cancer must be primary
- Exclude cases with specified diseases and pregnancy

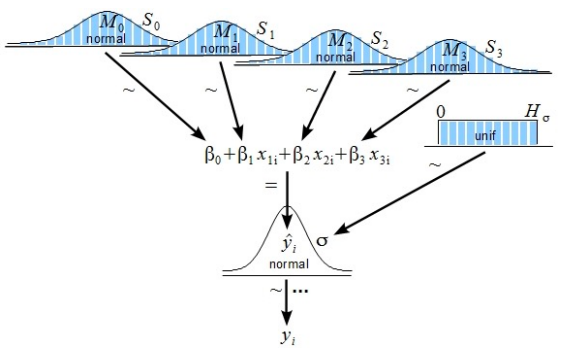
mGWAS & genowide pleiotropic analysis

- **Metabolome GWAS adjusted by multiple potential confounders**
- **Public available GWAS data**

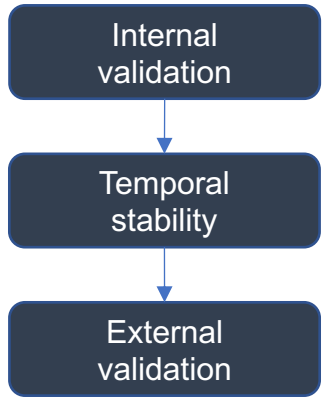
- Nat Genet 53, 1616–1621 (2021).
- Nat Genet 53, 1415–1424 (2021).
- Nat Commun 11, 4423 (2020).



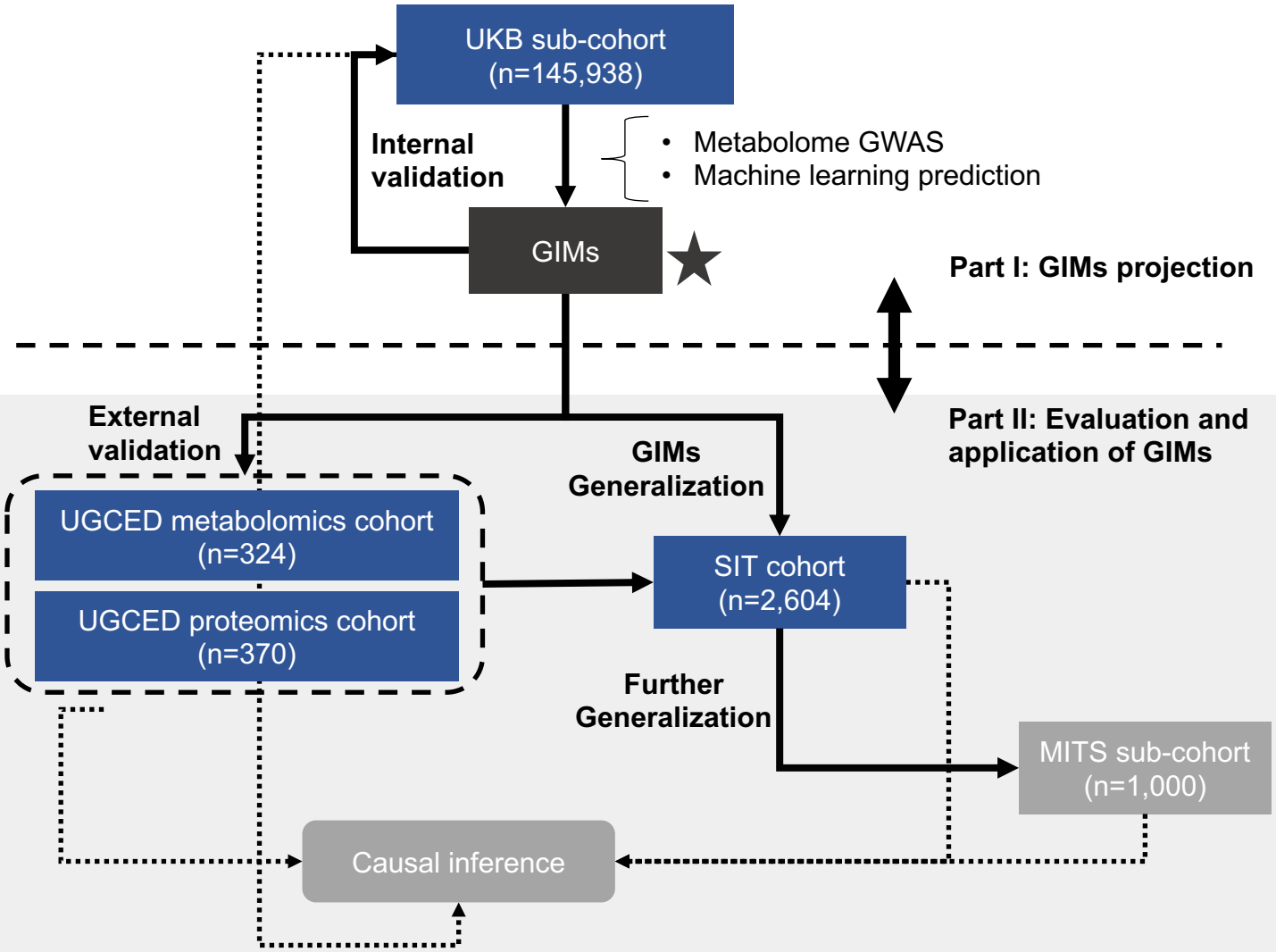
Projecting GIMs: Model training and evaluation



Machine learning predictions



Study design & Methods



BMJ 2019, 366.

JAMA Network Open, 2021, 4(6)

EBioMedicine, 2021, 74.

Theranostics, 2022, 12(10)

- **UGCED:** Upper Gastrointestinal Cancer Early Diagnosis
- **SIT:** Shandong Intervention Trial
- **MITS:** Mass Intervention Trial in Shandong

UGCED Cohort :

- LC-MS metabolomics (n=324) & Proteomics data (n=370)
- Multiple cases of precancerous lesions and early cancer
- Prospective endoscopic follow-up at multiple time points

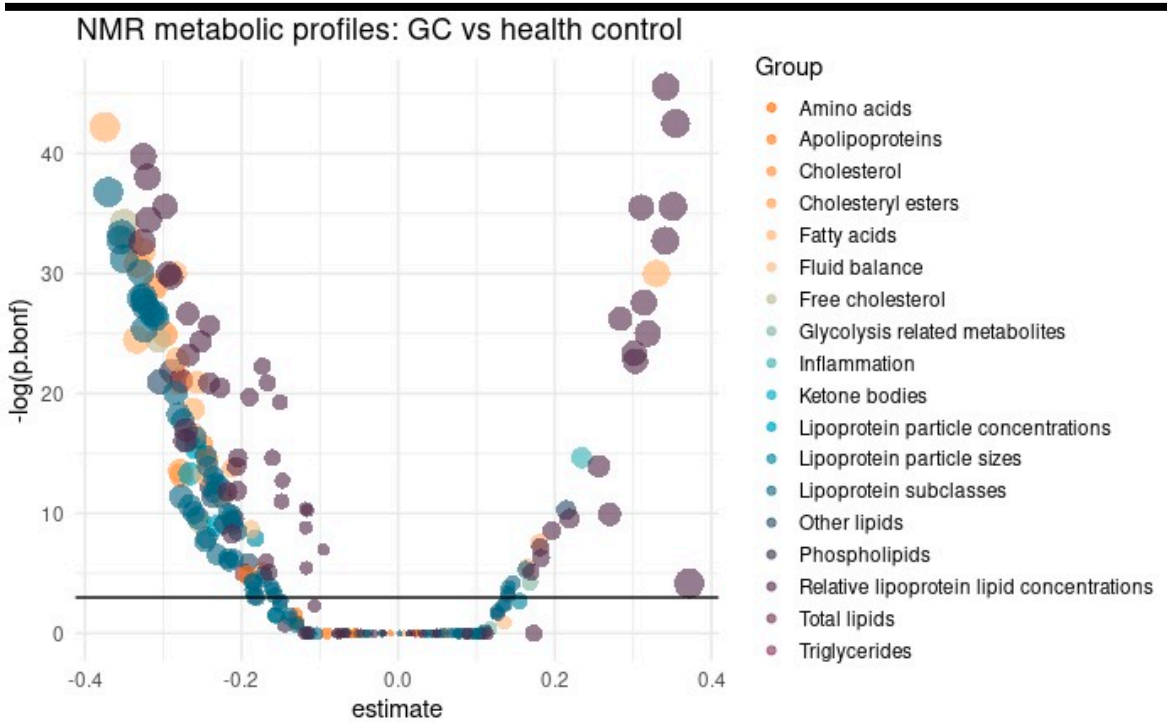
SIT Cohort:

- Randomized, 2x2x2 factorial intervention trial (NCT00339768)
- Based on a high-risk area of gastric cancer
- N=3,365, 2x2x2 factorial design
- 2 weeks of treatment
- 22.3 years of follow-up

˚Genotypic information is avaliable in the UGCED and SIT and MITS sub-cohort

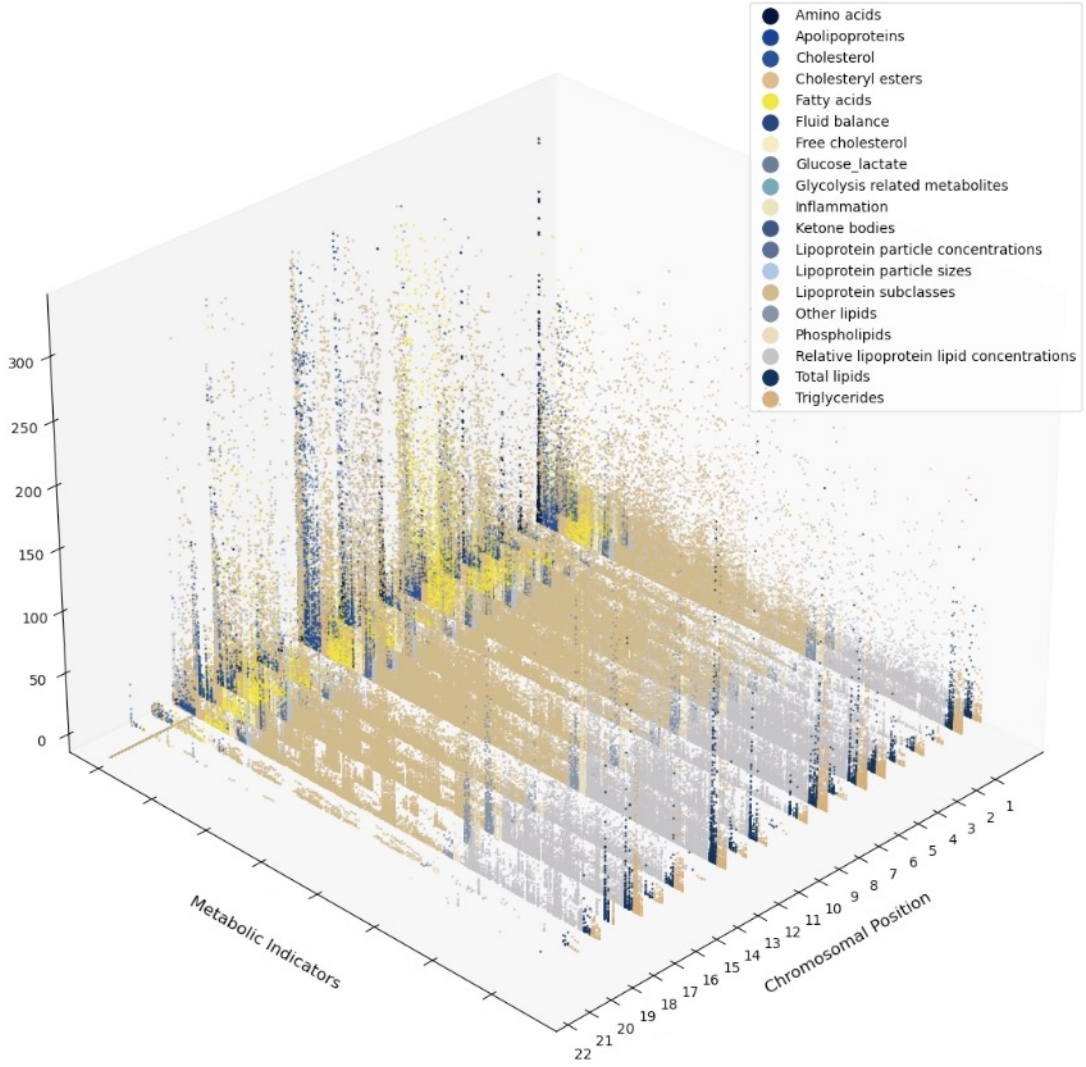
Metabolic biomarkers & polygenic insights

True metabolic profiles are associated with gastric cancer risk



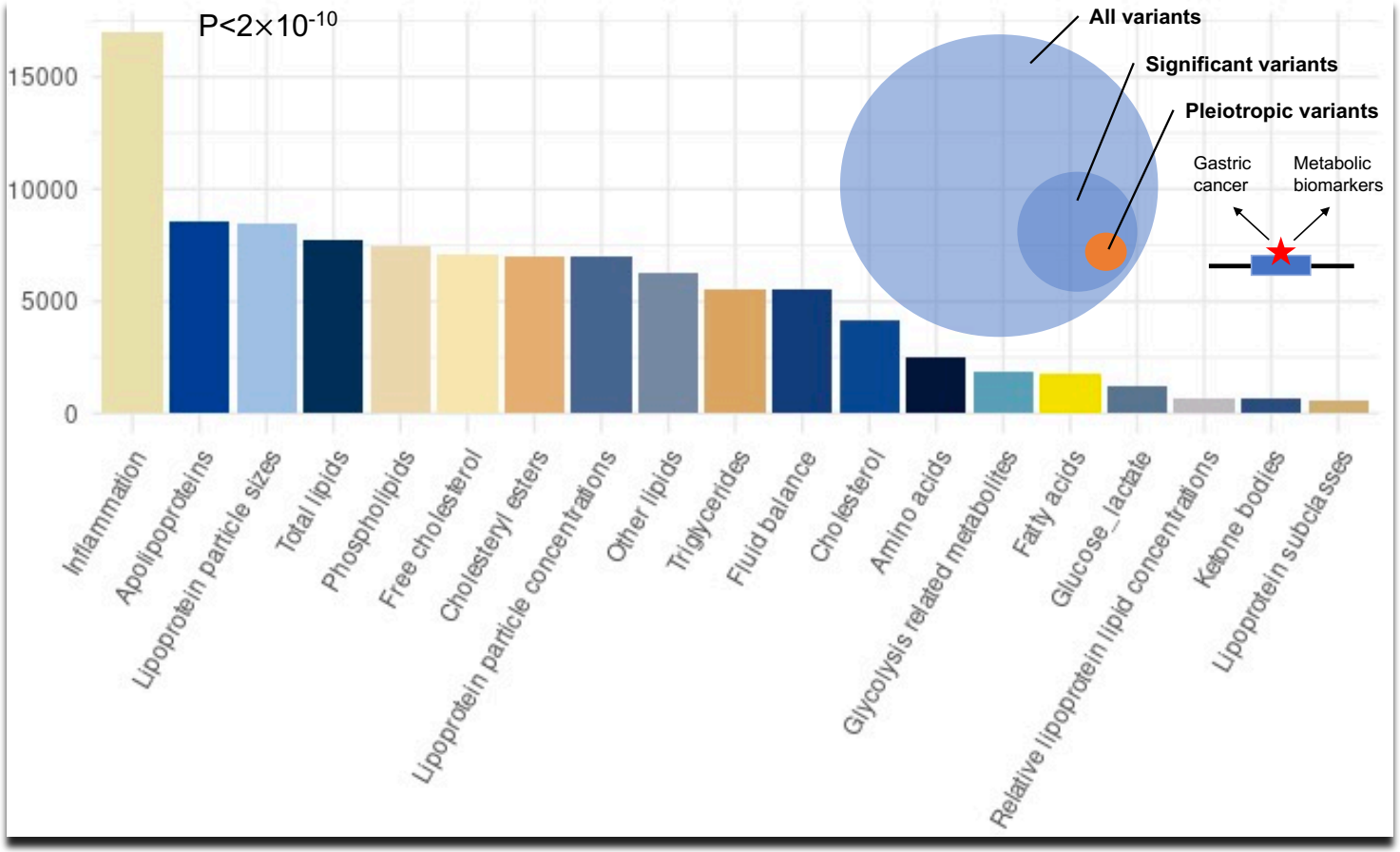
- Blood metabolic profiles are associated with gastric cancer risk
- 70k+ significant variants are identified for variation of 249 biomarkers
- Polygenic effects exist for multiple metabolic biomarkers (traits)

mGWAS for 249 NMR metabolic indicators



Genetic associations with metabolic profiles

Average number of mGWAS signficant variants by metabolic categories



Genome-wide pleiotropic analysis

Metabolome GWAS

- *249 metabolome-GWAS*

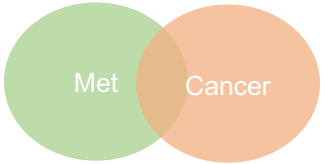
β_1, Z_1

Gastric Cancer GWAS

- *Nat Genet* 53, 1616–1621 (2021).
- *Nat Genet* 53, 1415–1424 (2021).
- *Nat Commun* 11, 4423 (2020).

β_2, Z_2

$$H_0: \beta_1 * \beta_2 = 0$$
$$H_1: \beta_1 * \beta_2 \neq 0$$

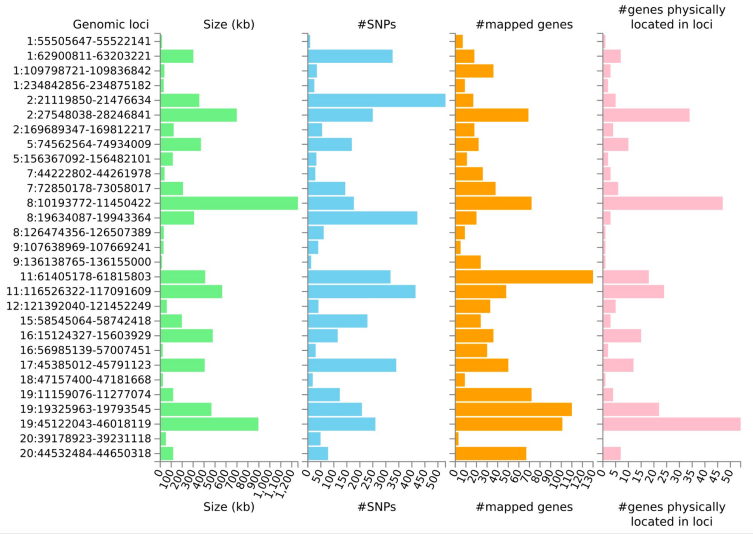


- De-correlate Z values if sample overlap issues exist
- Genome-wide significance threshold: P<5×10-8

Pleiotropic variants and genes

Summary of pleiotropic variants

# Genomic risk loci	29
# Lead SNPs	72
# Ind. Sig. SNPs	166
# Candidate SNPs	4540
# Candidate GWAS tagged SNPs	535



48 potentially pleiotropic genes

Test for phenotypic specificity of the pleiotropic genes

	Alimentary and digestive phenotypes	Non-alimentary and digestive phenotypes
Potentially pleiotropic genes	A1	A2
Non-pleiotropic	A3	A4

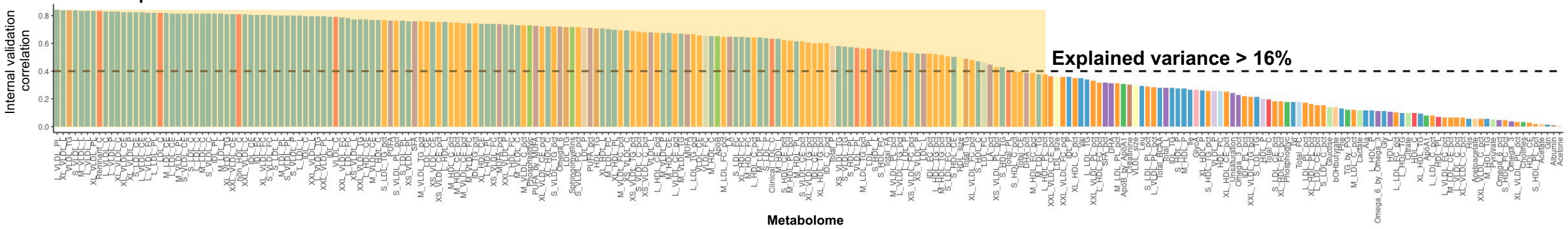
- H_0: Pleiotropic genes do not have phenotype specificity
- H_1: Pleiotropic genes have phenotype specificity

When using digestive/alimentary phenotype (MP:0005381) as the target phenotype:

The odds ratio (OR) for the potentially pleiotropic gene set associated with digestive tract disease phenotype is **2.35 (95% CI: 1.10-4.64)**.

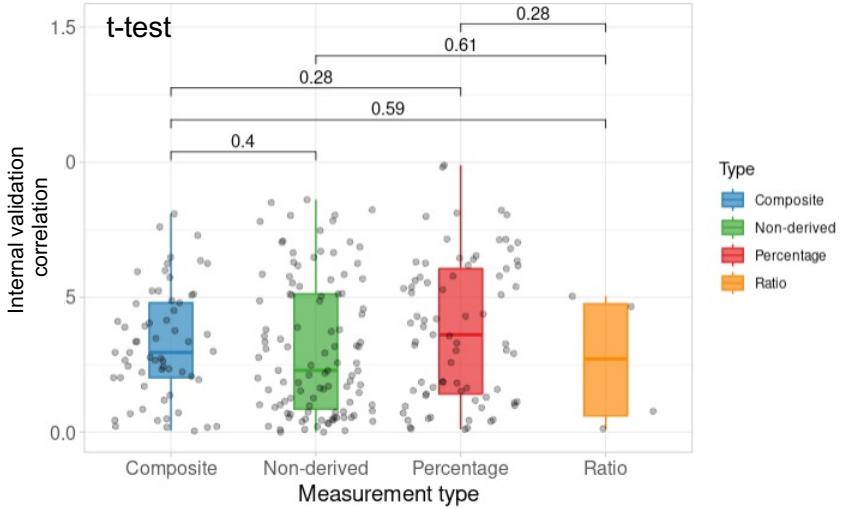
Machine learning models predicting GIMs

Overall performance



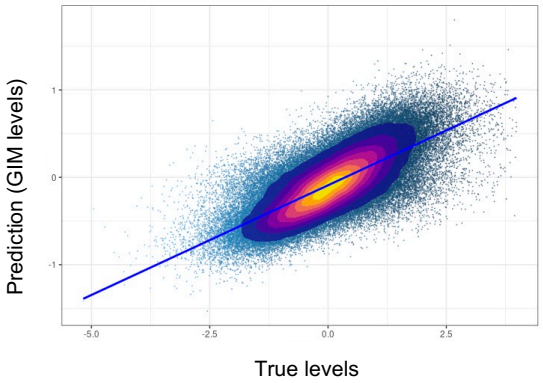
Amino acids
Apolipoproteins
Cholesterol
Cholesteryl esters
Fatty acids
Fluid balance
Free cholesterol
Glycolysis related metabolites
Inflammation
Ketone bodies
Lipoprotein particle concentrations
Lipoprotein particle sizes
Lipoprotein subclasses
Other lipids
Phospholipids
Relative lipoprotein lipid concentrations
Total lipids
Triglycerides

Performance by measurment types

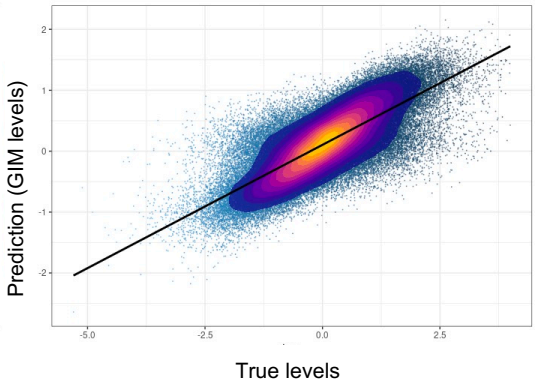


Performance by individual biomarkers (e.g. for fatty acids)

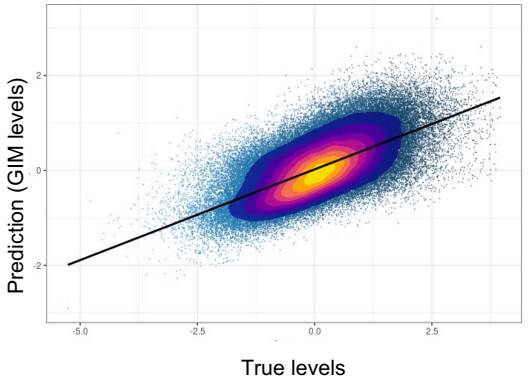
Omega-6 fatty acids



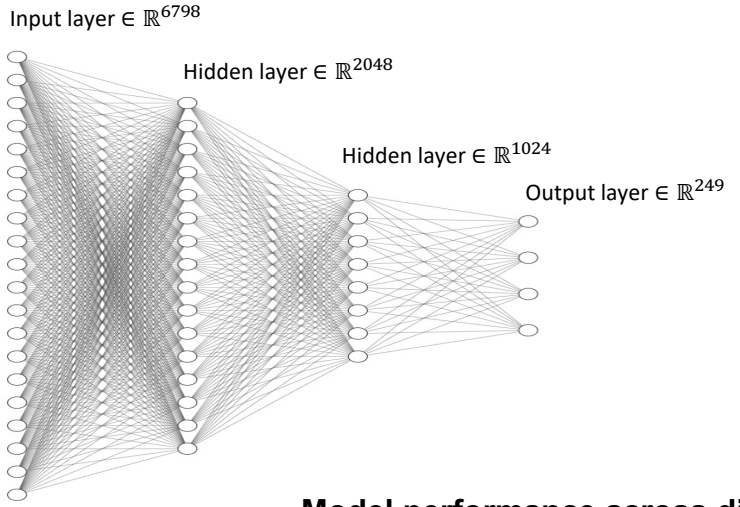
Linoleic acid



Polyunsaturated fatty acids



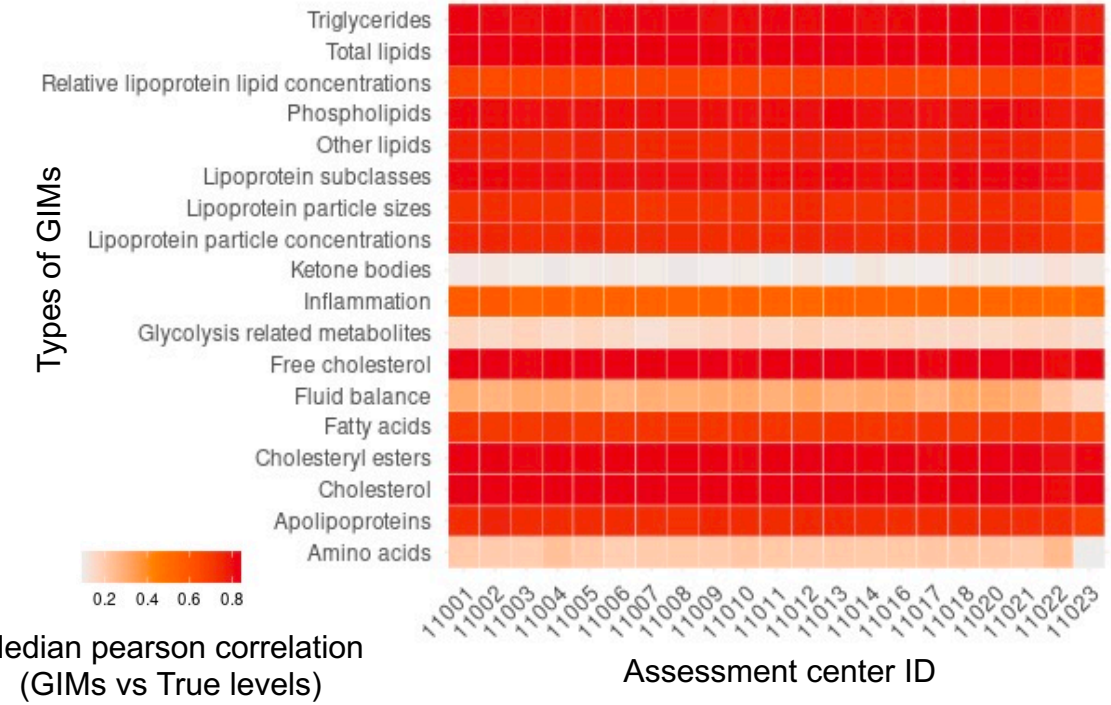
Machine learning models predicting GIMs



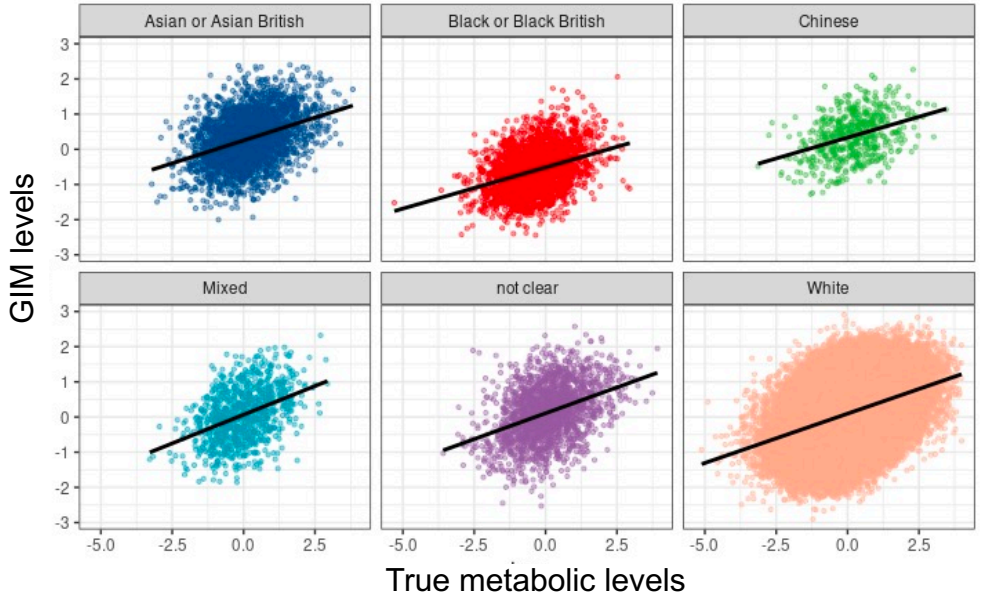
Why neural networks

- Capable of capturing potential interaction effects
- Theoretically can fit any function in nature
- 2048 and 1024 neurons in the 1st and 2nd hidden layer, respectively
- Batch normalization and dropout techniques were applied to reduce overfit

Model performance across different assessment centers

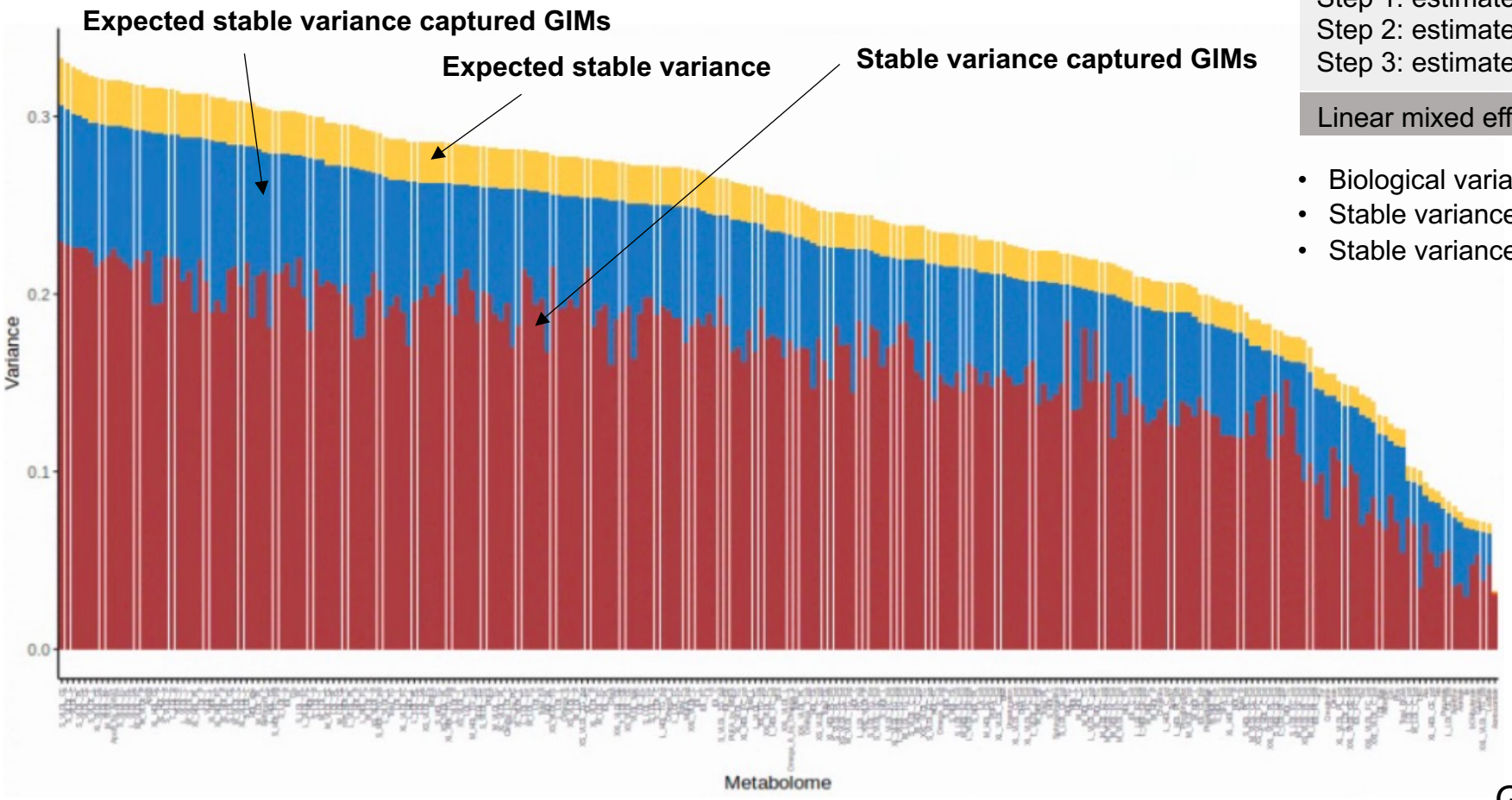


Model performance across different ethnic groups (e.g. for Omega-6 fatty acids)



GIMs are temporarily stable for long-term cancer risk indication

Evaluation on the temporal stability of GIMs

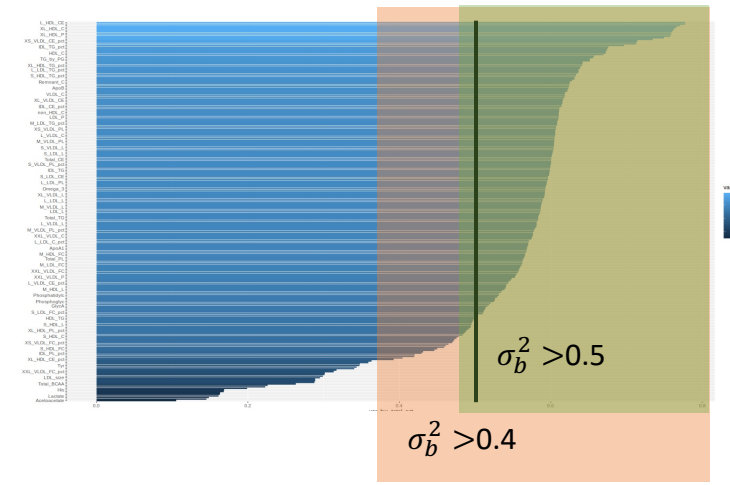


Estimation of stable variance from GIMs

Step 1: estimate the biological variation of each metabolic trait
Step 2: estimate the stable variance of the biological variation
Step 3: estimate the stable variance accounted by genomics

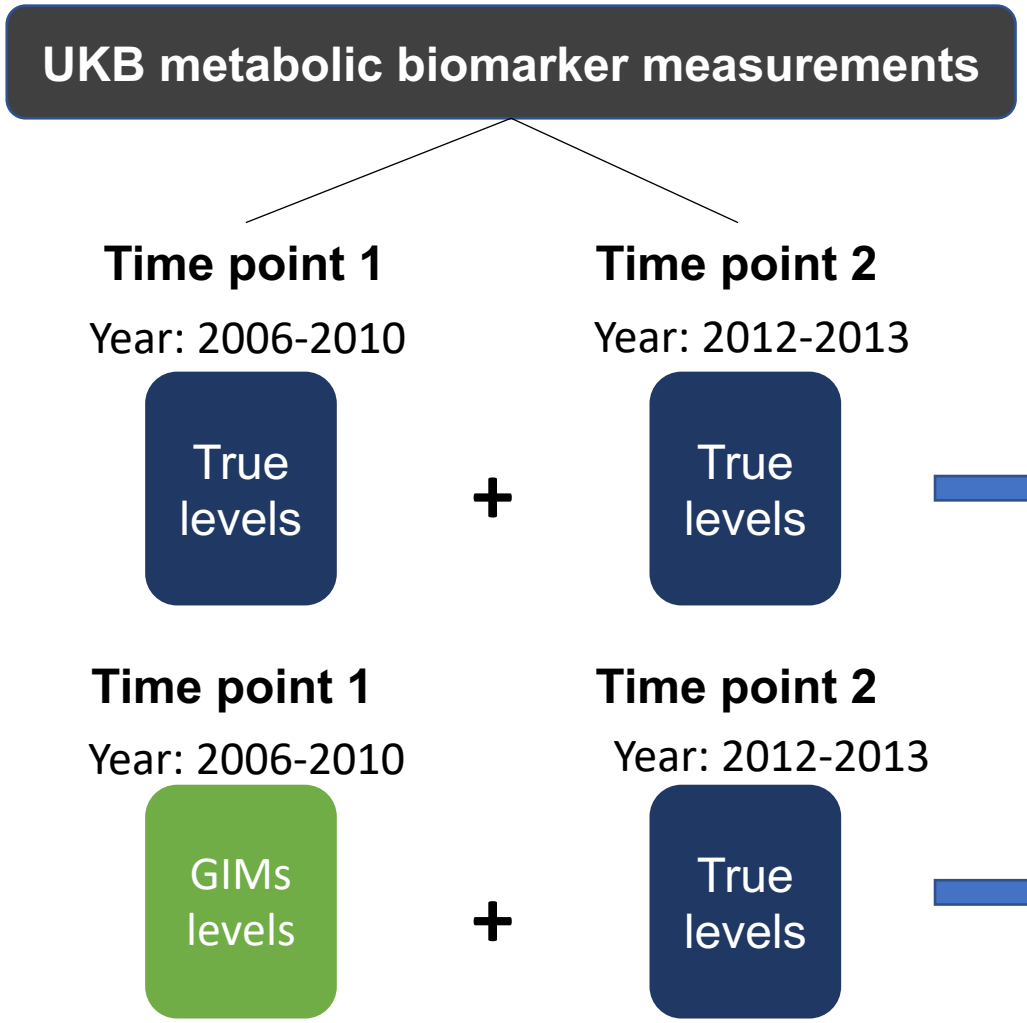
Linear mixed effect model was used for variance decomposition

- Biological variance = $\sigma^2_{between} + \sigma^2_{within}$
- Stable variance = 60% *($\sigma^2_{between} + \sigma^2_{within}$)
- Stable variance accounted by genetics = 42~50% *($\sigma^2_{between} + \sigma^2_{within}$)



Good sign of the ability for risk stratification

GIMs are temporarily stable for long-term cancer risk indication



Estimation of stable variance from GIMs

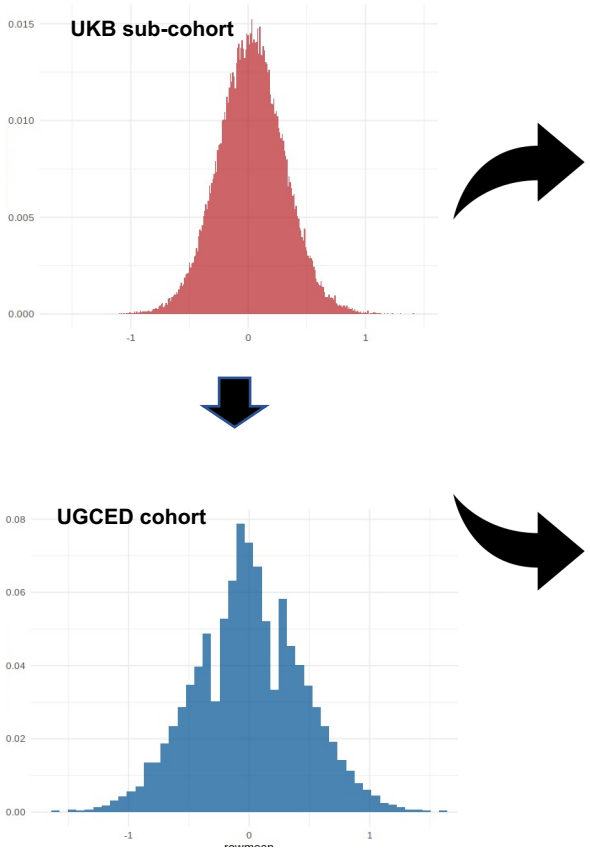
Step 1: estimate the biological variation of each metabolic trait
Step 2: estimate the stable variance of the biological variation
Step 3: estimate the stable variance accounted by genomics

Linear mixed effect model was used for variance decomposition

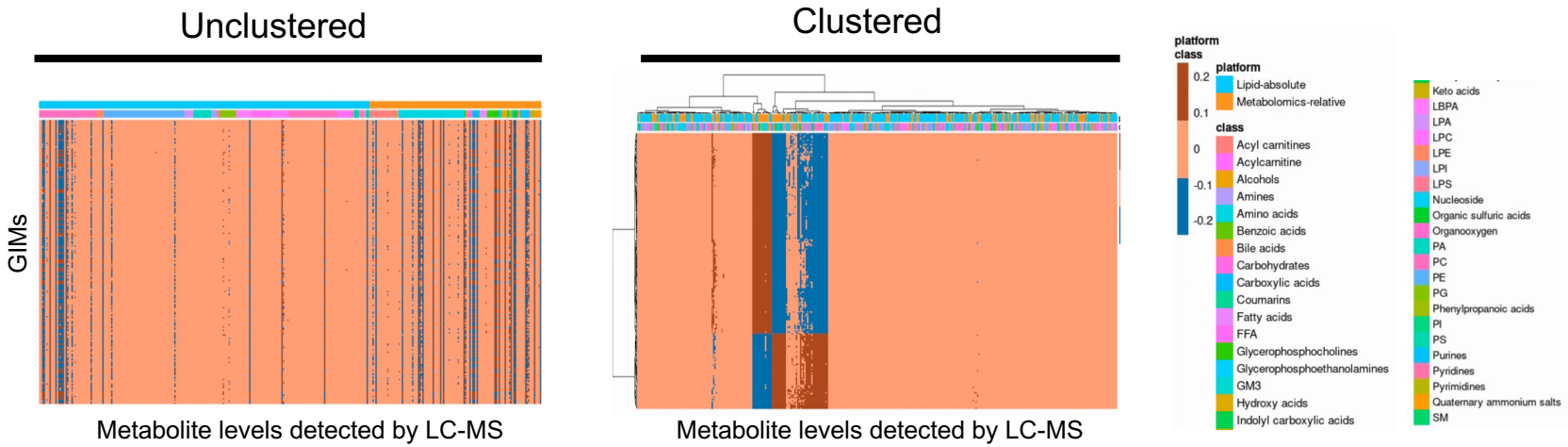
- Biological variance = $\sigma^2_{between} + \sigma^2_{within}$
- Stable variance = 60% *$(\sigma^2_{between} + \sigma^2_{within})$
- Stable variance accounted by genetics = 42~50% *$(\sigma^2_{between} + \sigma^2_{within})$

GIMs coincide with external metabolomics and proteomics profiles

GIMs projection



GIMs display associations with the metabolomic profiles in UGCED cohort



Protein-coding pleiotropic genes associated with risk of early gastric cancer

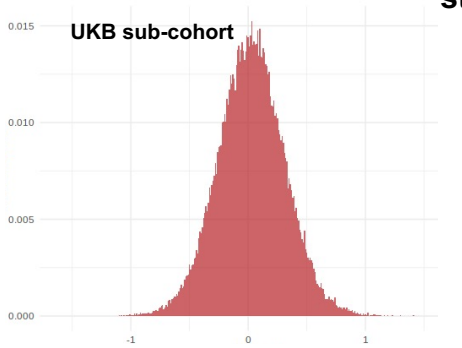
Gene Symbol	OR (95% CI)[a]	nominal p-value	FDR q-value[b]
KPNB1	**5.00 (2.23-14.04)**	**<0.001**	**0.011**
NPEPPS	**2.63 (1.34-5.77)**	**0.008**	**0.033**
APOB	**3.11 (1.78-6.84)**	**0.001**	**0.011**
PDXDC1	**2.49 (1.32-5.14)**	**0.007**	**0.032**
TOMM40	**4.02 (1.98-10.70)**	**0.001**	**0.012**
UBE2L3	0.32 (0.10-0.79)	0.029	0.078
KANK2	2.16 (1.15-4.50)	0.026	0.073

[a]Odds ratios were calcuated by multivariate logistic regression comparing early gastric cancer and chronic atrophic gastritis

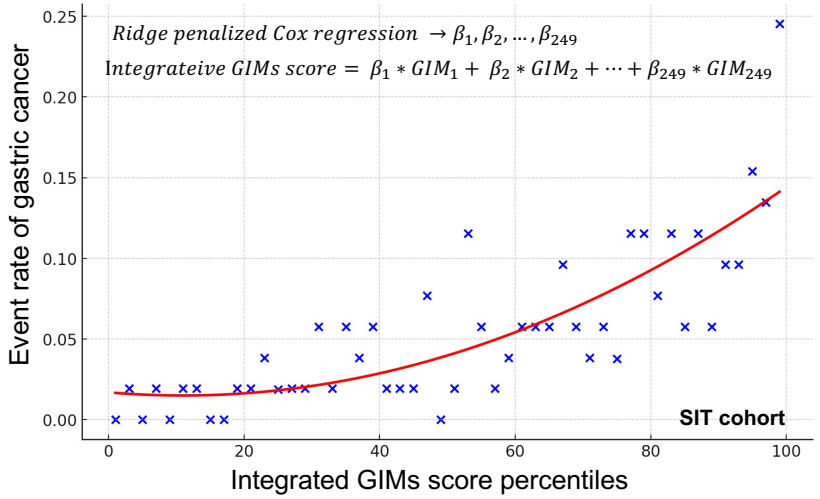
[b]FDR was controlled for the statistical testing procedure for 2682 proteins

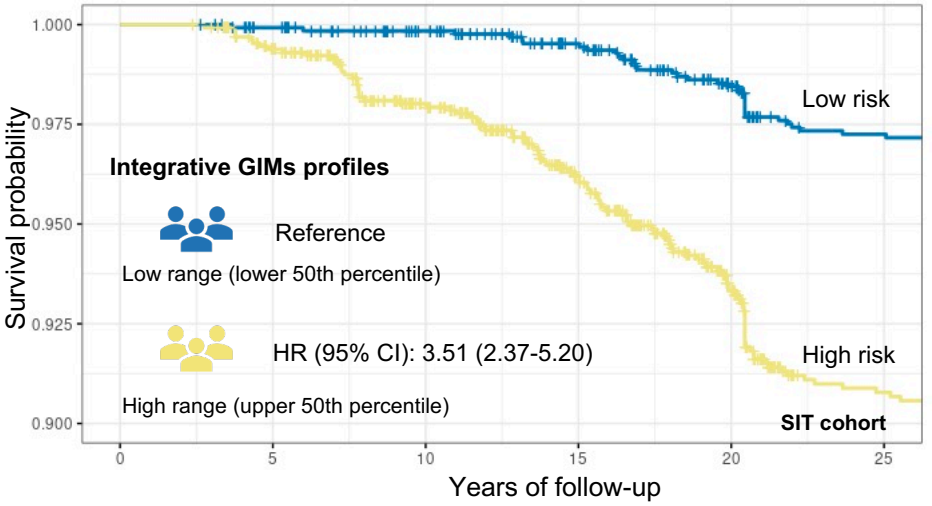
GIMs stratify gastric cancer risk with biological consistency

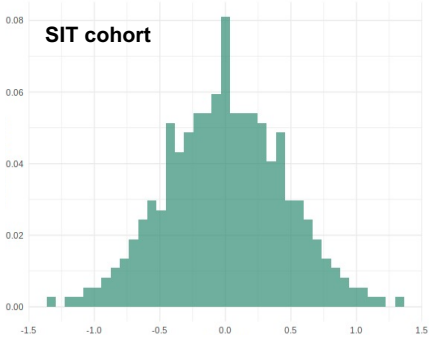
GIMs projection



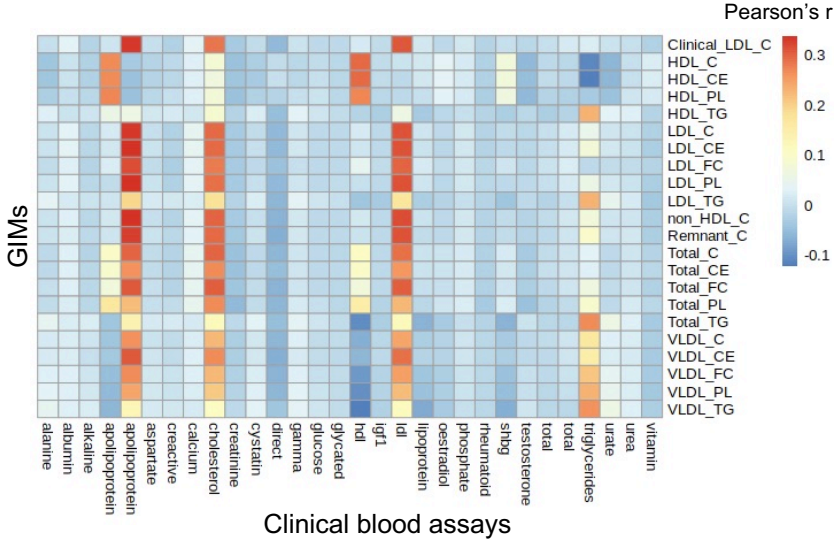
Risk stratification

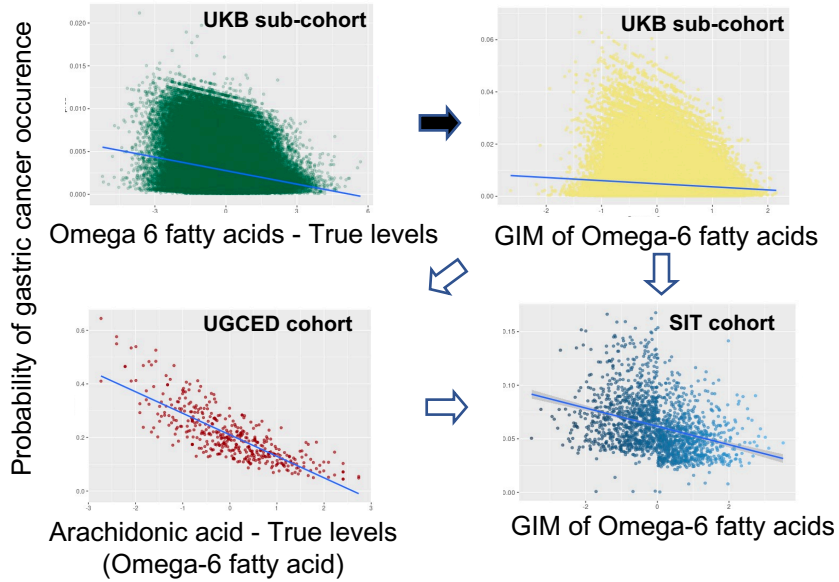






Biological consistency



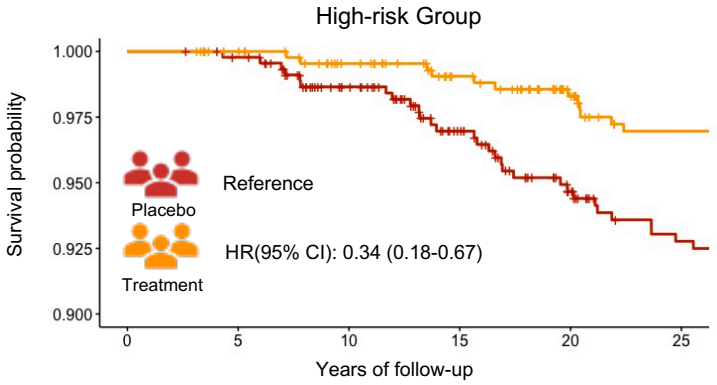


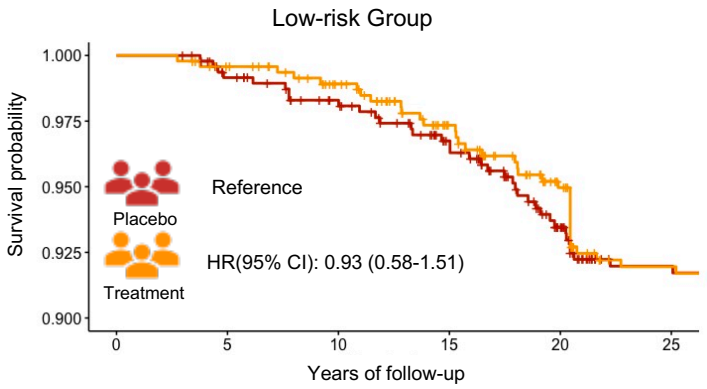
GIMs identify target population for gastric cancer prevention

Efficacy of interventions in preventing gastric cancer across GIMs-defined risk subgroups

Risk group	No. of cancer (Person-years)		HR	95%CI	P for interaction
	Placebo	Treatment			
H.pylori eradication					
High-risk	**35(11282)**	**12(10876)**	**0.34**	**0.18-0.67**	**0.02**
Low-risk	36(113645)	31(10705)	0.93	0.58-1.51	
Vitamin supplementation					
High-risk	28 (15493)	16 (15192)	0.57	0.30-1.03	0.19
Low-risk	48 (15862)	47 (15574)	0.99	0.69-1.55	
Garlic supplementation					
High-risk	28 (15520)	16 (15096)	0.58	0.32-1.08	0.19
Low-risk	48 (15863)	47 (15642)	0.96	0.64-1.44	

For *H.pylori* treatment





- Similar eradication rates noted between the high and low-risk subgroups.
- Higher responsiveness to *H. pylori* treatment for gastric cancer prevention observed in high-risk subjects.

Summary

Conclusion

- GIMs may be indicators of the risk of developing GC, offering new insights into understanding GC etiology.
- GIMs may be an effect modifier for *H.pylori* treatment, thus serving as biomarkers for targeted populations of GC primary prevention.

Ongoing efforts

- Extra external validation by independent cohorts (sub-cohort from MITS)
- Development of causal learning framework for casual inference between the key genetic variants and GIMs

Acknowledgement

We gratefully thank the people contributing to this project

PhD mentor: Wen-Qing Li, Ph.D.

Lab falculty & staff members:

- Kai-Feng Pan, Ph.D.
- Wei-Cheng You, M.D.
- Lian Zhang Ph.D.
- Jun-Ling Ma, B. S.

Current lab students with contribution:

- Heng-Min Xu, Ph.D. candidate
- Meng-Yuan Wang, Master's student

Collaborators:

- Peng Cui, Ph.D. (Tsinghua University)
- Yue He, Ph.D. (Tsinghua University)
- Wei-Dong Liu, B.S. (Linqu County Public Health Bureau, Shandong, China)

Best Oral Presentation Award
in 2023 AACR-KCA Joint Conference on Precision Medicine in Cancer

北京大学 PEKING UNIVERSITY

北京大学肿瘤医院 BEIJING CANCER HOSPITAL

Q & A