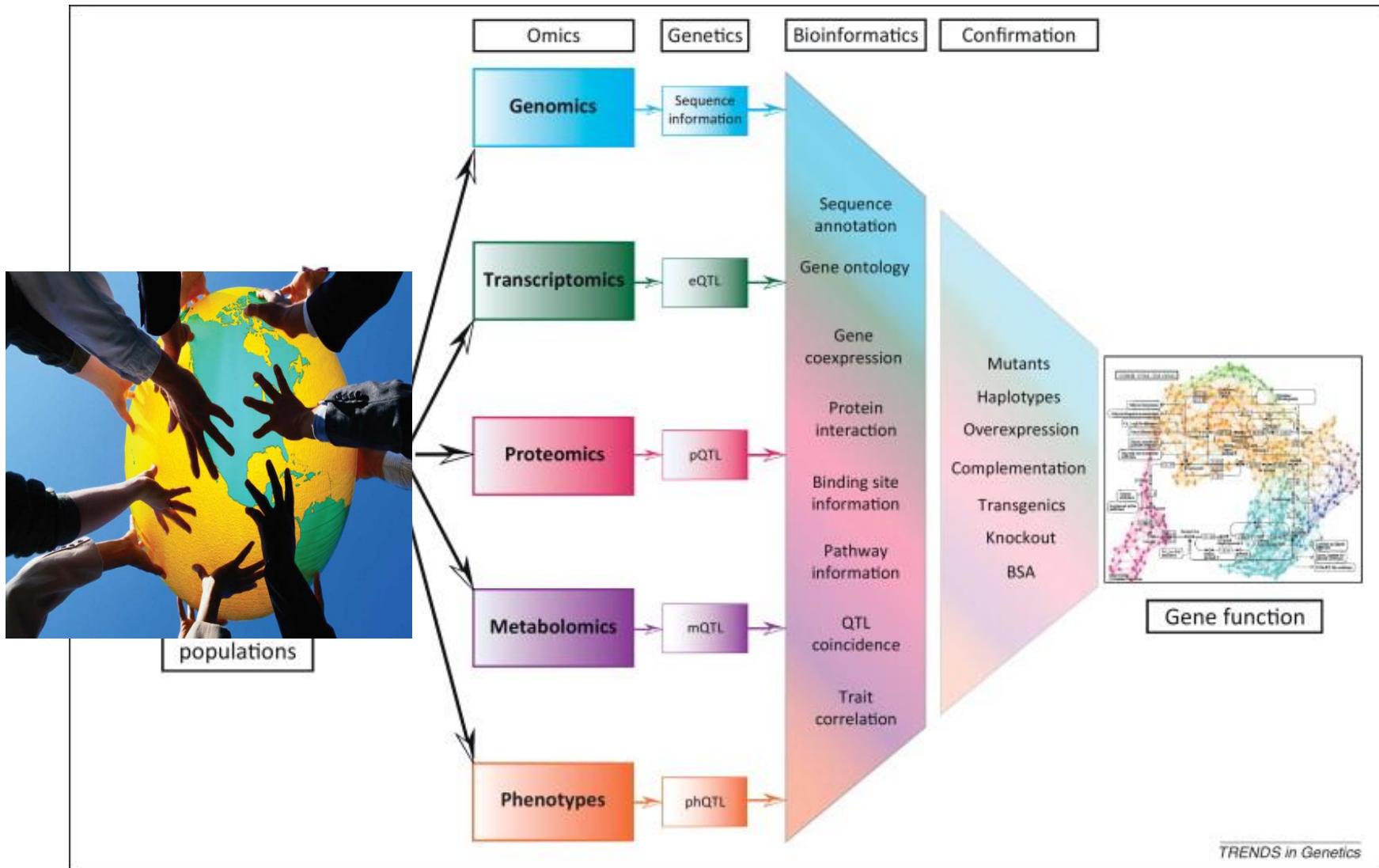


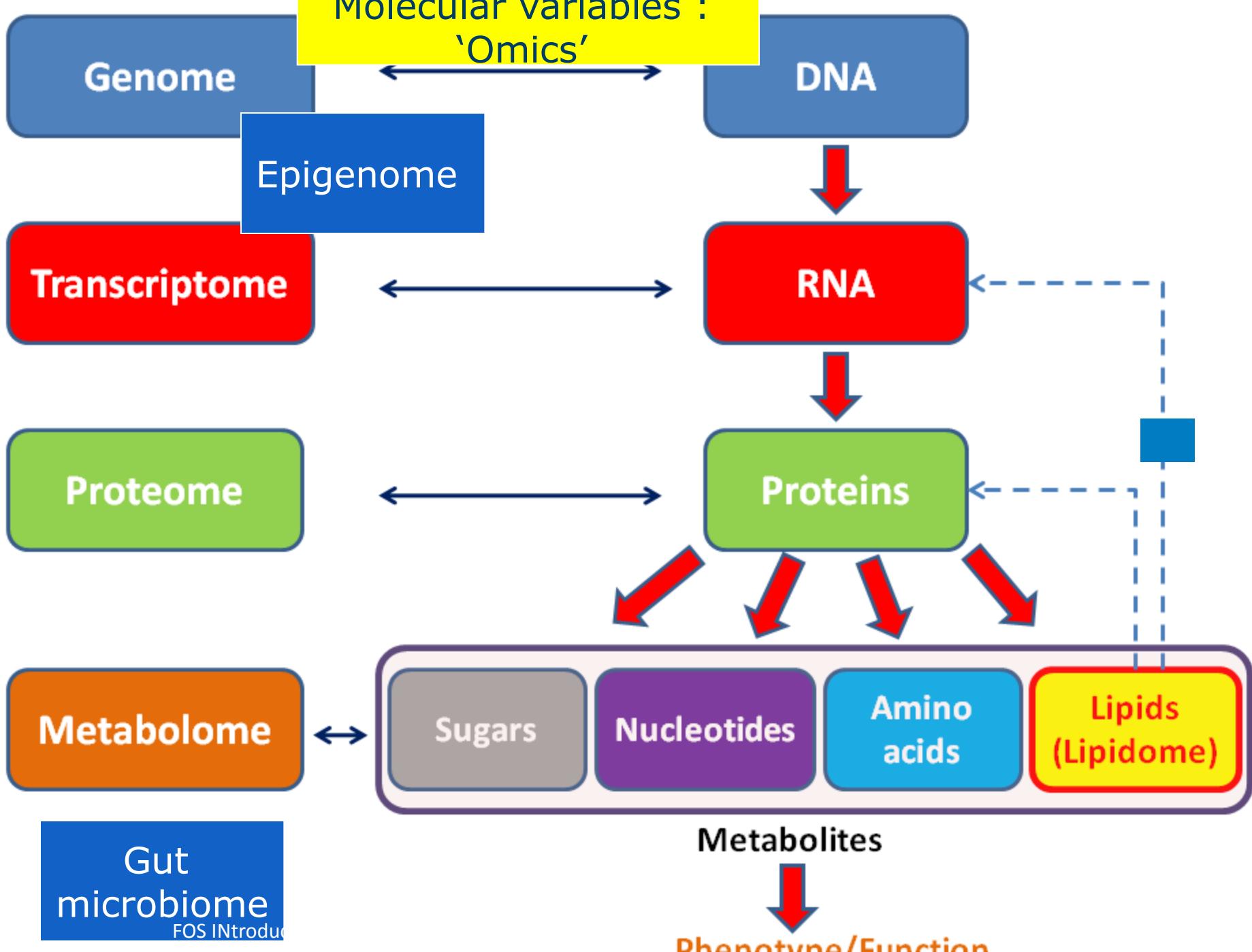
Molecular Epidemiology

- Introduced by Kilbourne (1973), infectious diseases; Schulte and Perera (1993 Principles and Practices)
- Integrates Epidemiology, Medical Sciences and Molecular Biology
- Studies the influence on health of environmental and genetic risk factors measured by (holistic) molecular signatures
- Contributes to
 - prediction/prognosis
 - monitoring exposure, response to interventions
 - etiological understanding (disease mechanisms)**

Etiology and genome biology



Molecular variables :
'Omics'



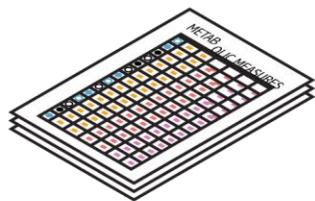
Remark

Now we will focus on the metabolome.

Mostly in blood: 4000 metabolites.

Best platforms measure about 600 in a standardized fashion high throughput

1H-NMR metabolomics platform well-standardized ; affordable (Ala –Korpela Finland)



Fluid balance

- Creatinine (mmol/l)
- Albumin (signal area)

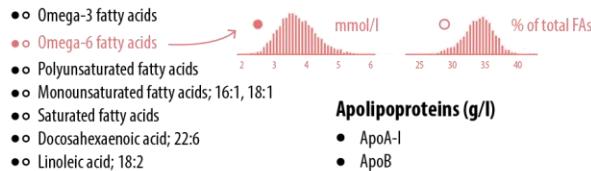
Inflammation (mmol/l)

- Glycoprotein acetyls,
mainly $\alpha 1$ -acid glycoprotein

Fatty acids and saturation

- Total fatty acids
- Estimated degree of unsaturation

Fatty acids (mmol/l and % of total FAs)



14 LIPOPROTEIN SUBCLASSES

12 lipid measures for each subclass

- Esterified cholesterol (mmol/l and % of total lipids)
- Free cholesterol (mmol/l and % of total lipids)
- Triglycerides (mmol/l and % of total lipids)
- Phospholipids (mmol/l and % of total lipids)

- Total cholesterol (mmol/l and % of total lipids)
- Total lipids (mmol/l)
- Particle concentration ($\mu\text{mol/l}$)

Average particle size (diameter in nm)



Average lipid composition (%)



Ketone bodies (mmol/l)

- Acetate
- Acetoacetate
- 3-hydroxybutyrate

Glycolysis related metabolites (mmol/l)

- Glucose
 - Lactate
 - Pyruvate
 - Citrate
 - Glycerol*
- (* not available for EDTA plasma samples)

Amino acids (mmol/l)

- Alanine
- Glutamine
- Glycine*
- Histidine

Branched-chain amino acids

- Isoleucine
- Leucine
- Valine

Aromatic amino acids

- Phenylalanine
- Tyrosine

Glycerides & phospholipids (mmol/l)

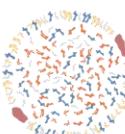
- VLDL triglycerides
- LDL triglycerides
- HDL triglycerides
- Triglycerides
- Phosphoglycerides
- Ratio of triglycerides to phosphoglycerides
- Phosphatidylcholine and other cholines
- Sphingomyelins
- Total cholines

Apolipoproteins (g/l)

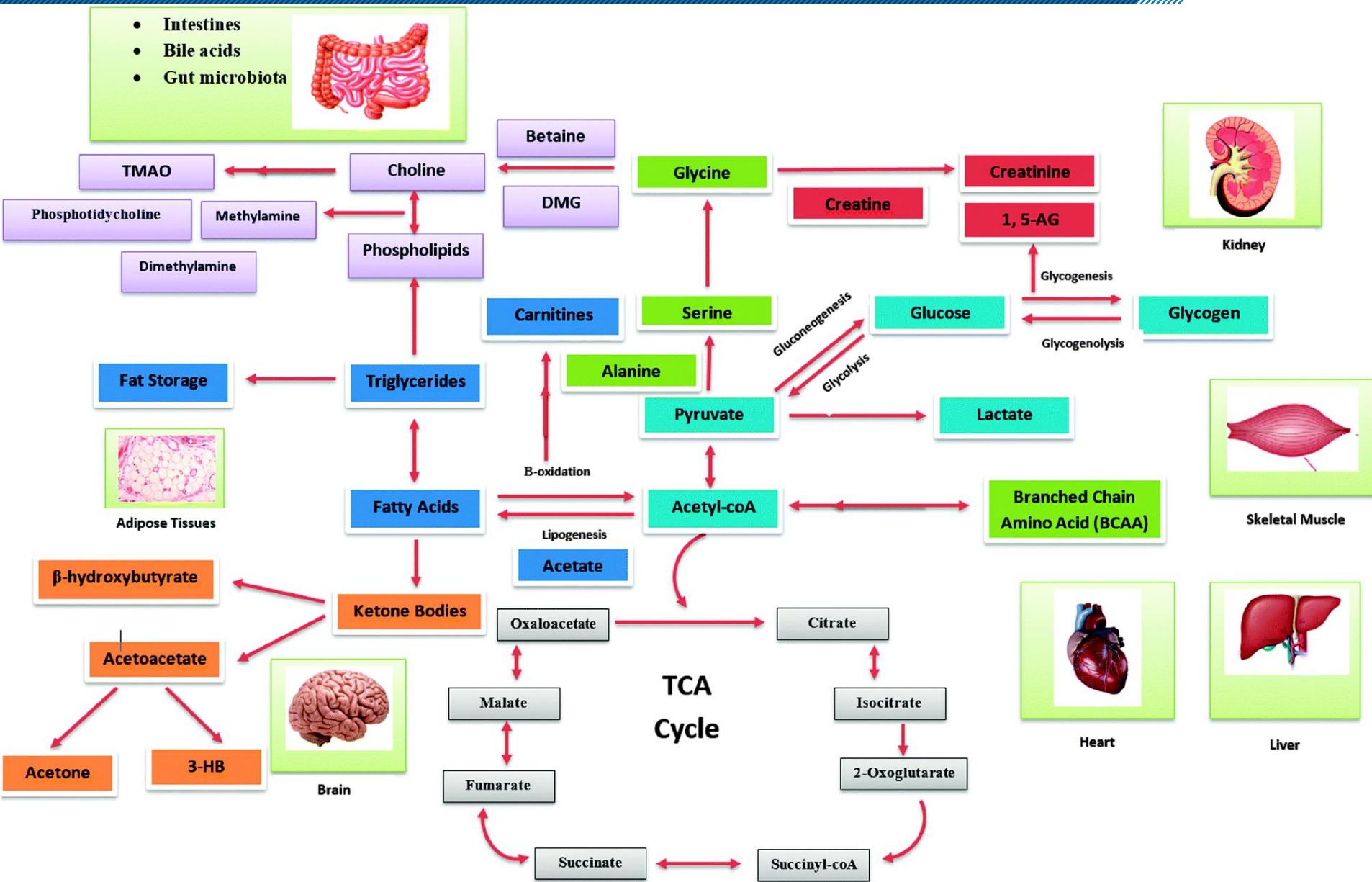
- ApoA-I
- ApoB
- ApoB/ApoA-I

Lipoprotein particle size (nm)

- Mean diameter of VLDL particles
- Mean diameter of LDL particles
- Mean diameter of HDL particles



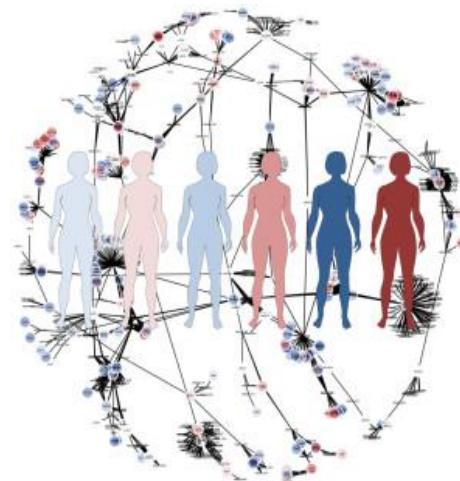
Which tissue functions do the ^1H NMR metabolites represent?



BBMRI Biobanking consortium

Multi-level omics data

N=100,000 GWAS
N=750 Go.NIL

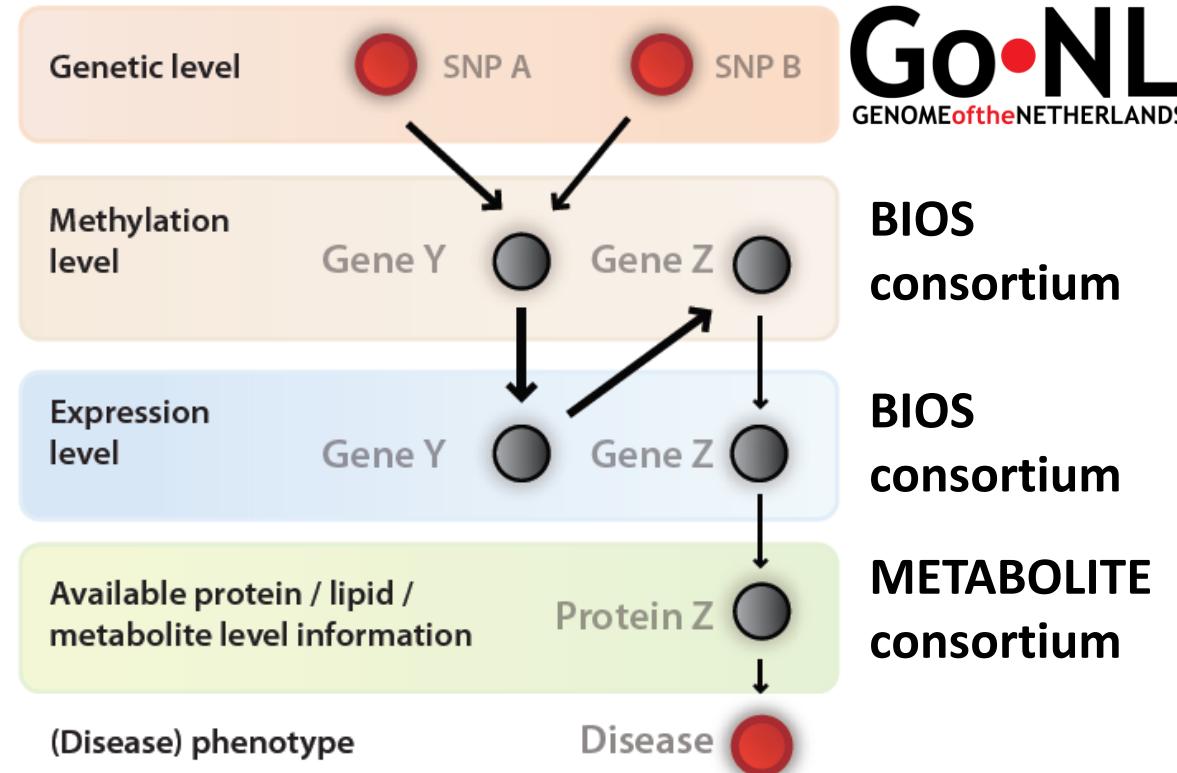


N=4,000

N=4,000

N=50,000

N>250,000



Go•NL
GENOMEofthe NETHERLANDS

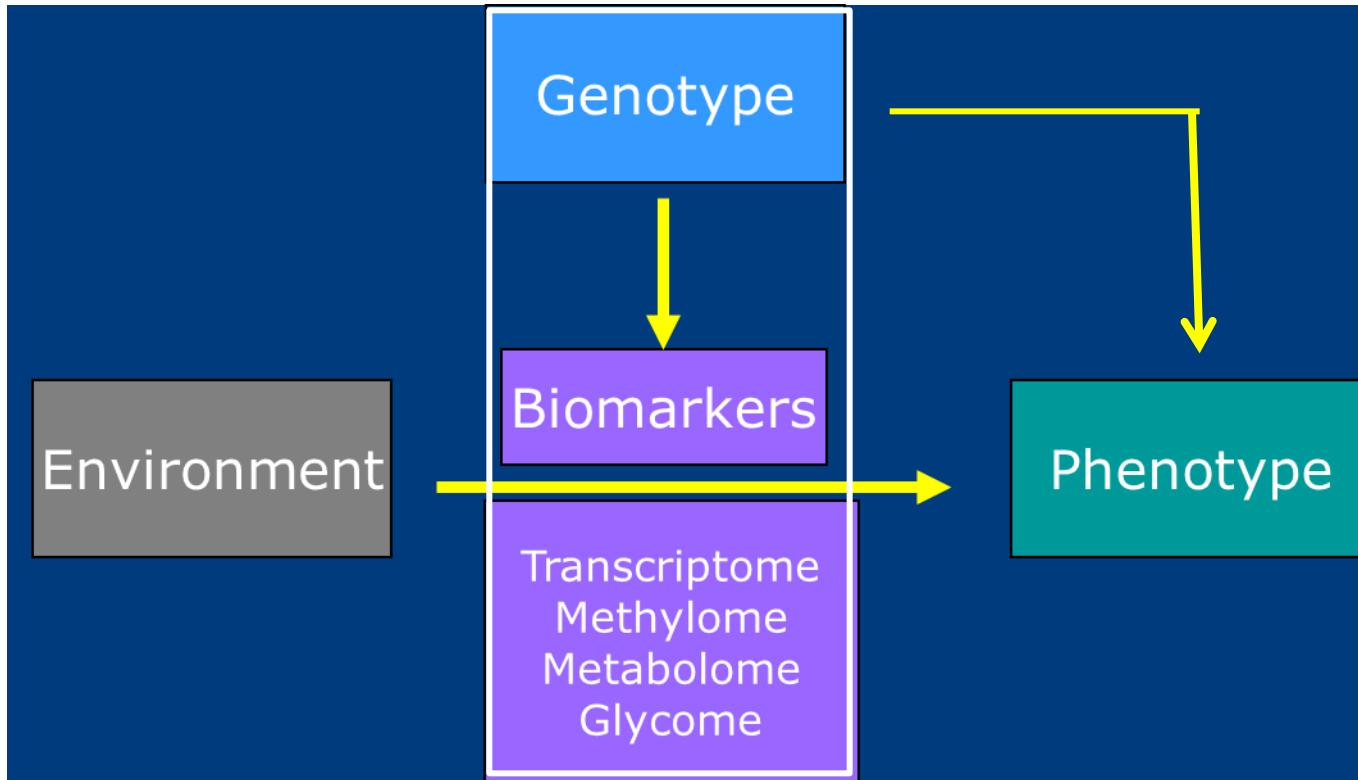
BIOS
consortium

BIOS
consortium

METABOLITE
consortium

Central databases for the research community

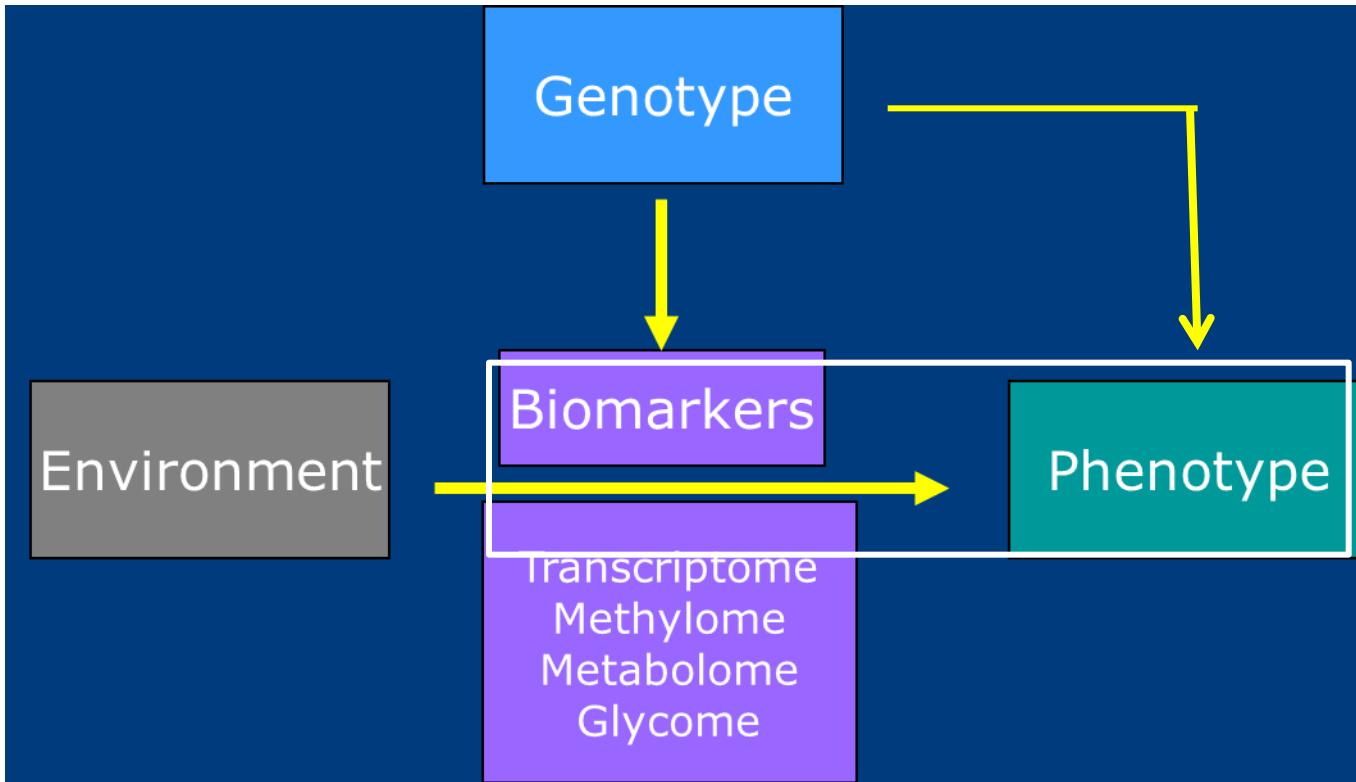
What relations do scientists investigate with molecular profiling data



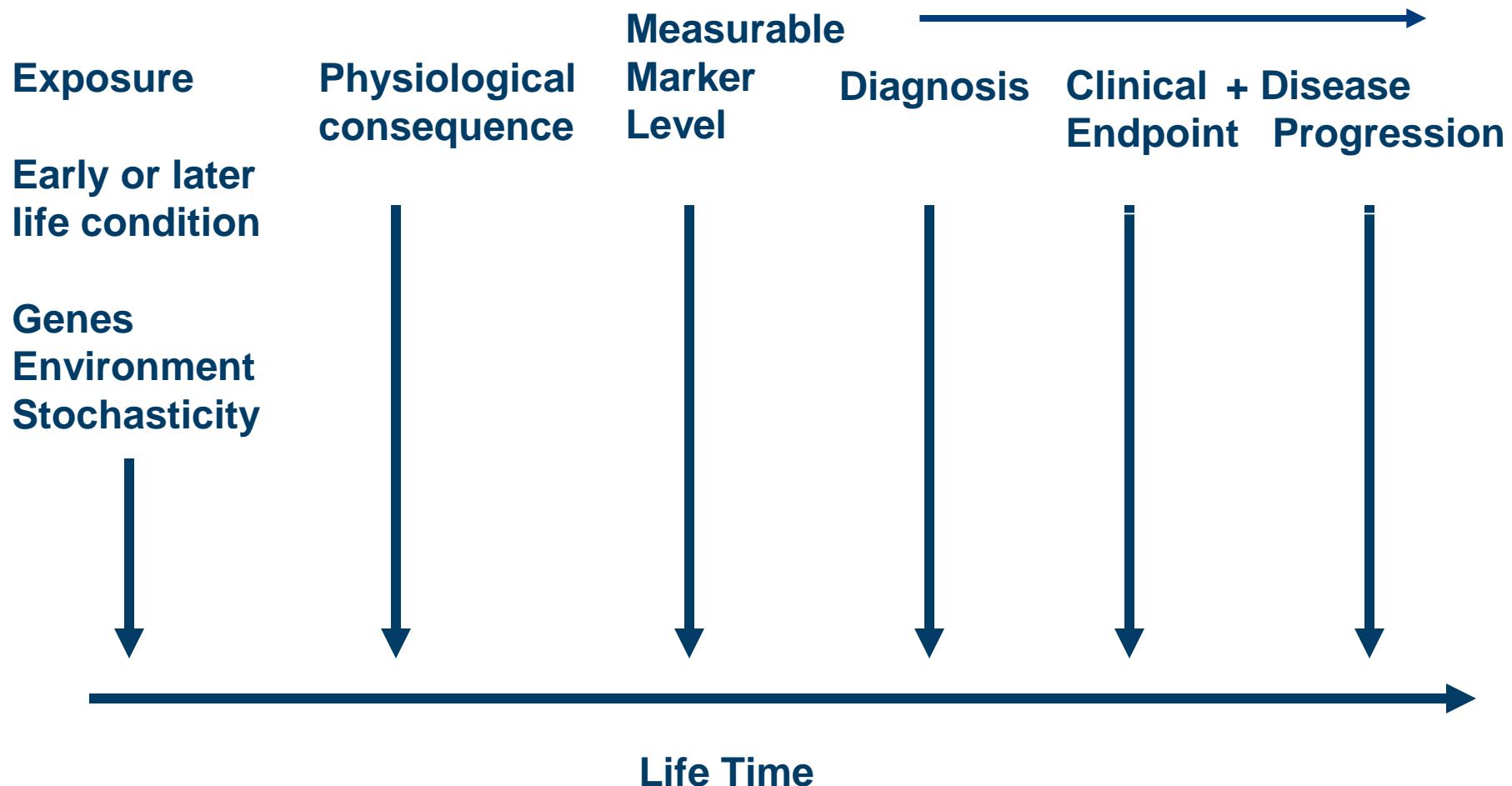
Molecular Epidemiology

- Introduced by Kilbourne (1973), infectious diseases; Schulte and Perera (1993 Principles and Practices)
- Integrates Epidemiology, Medical Sciences and Molecular Biology
- Studies the influence on health of environmental and genetic risk factors measured by (holistic) molecular signatures
- Contributes to
 - prediction/prognosis***
 - monitoring exposure, response to interventions
 - etiological understanding (disease mechanisms)

What relations do scientists investigate with molecular profiling data



Exposure Events in lifetime perspective

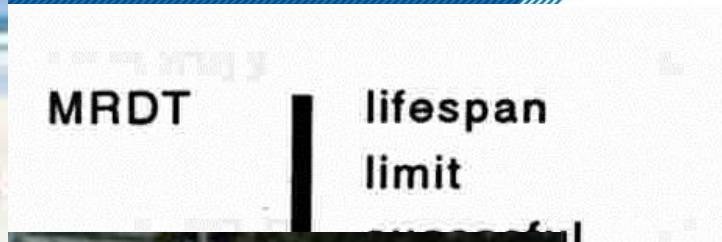


Biomarkers I

- Relation of Exposure /determinant and outcome
- Exposure: environment (early, late, diet, lifestyle, chemicals, geography), host (genetic background, age), health change over time (disease, biological ageing process); outcome = phenotype
- Biomarker (WHO): a substance or biological structure that can be measured in the human body and may influence, explain or predict the incidence or prognosis of outcome of disease

Biomarkers II

or NIH biomarker working group:
a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention



Classical Algorithms

Definition Metabolic syndrome

- Abdominal obesitas >94 cm men, >80 cm women and \geq two of the next:
- Fasting glucose >100 mg/dL (5,6 mmol/L) or diabetes
- Triglycerides >150 mg/dL (1,7 mmol/L) or treatment
- HDL-cholesterol <40 mg/dL for men, < 50 mg/dl for women or treatment
- Blood Pressure > 130/85 mmHg SBP or treatment



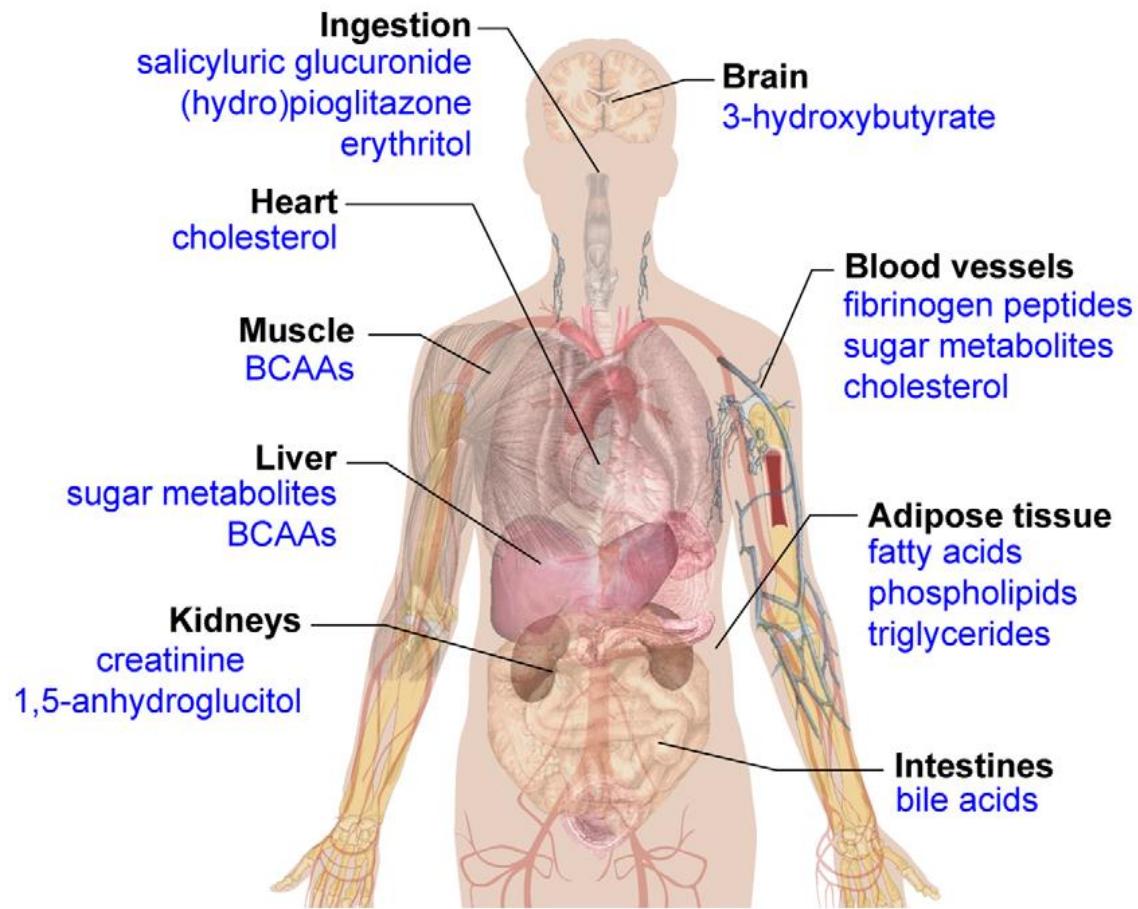
Definition Framingham risk score (10 years CVD risk)

- Age
- Gender
- Smoking
- Diabetes
- Total cholesterol
- HDL-cholesterol
- Systolic Blood Pressure



Molecular Biomarkers

Type II Diabetes prediction 1H NMR metabolites



By what steps can you make a novel predictor

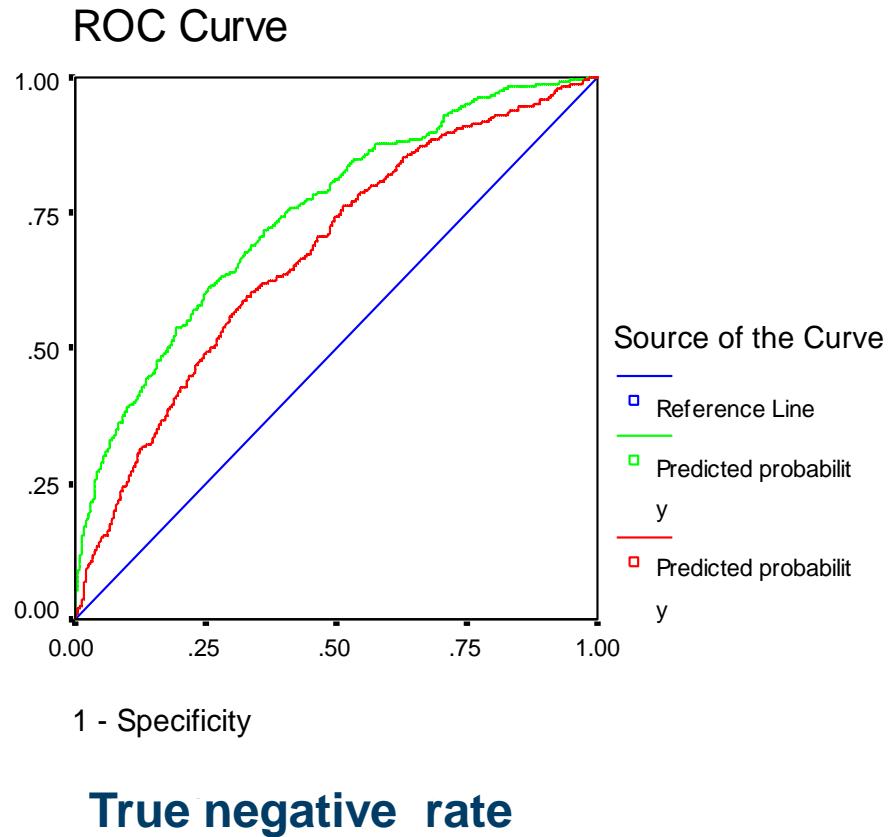
Generate a predictor : factor identification and model development :

- 1. Exploration of associations between metabolites and diverse endpoints.**
- 2. Cross sectional → Prospective/longitudinal follow-up studies**
- 3. Univariate (single metabolites) , multivariate**
- 4. Replication in independent studies**
- 5. Meta-analysis in multiple studies, create predictors (for example of mortality risk) and compare to existing predictors**



Receiver Operator Characteristic (ROC) curves to compare novel and traditional predictors (example 10 y CVD risk)

True positive rate



Blue : 50%-50%

→ AUC=0.5

Red : model with Age ,
Diabetes, Smoking

→ AUC=0.67

Green: model with Age,
Diabetes, Smoking
HDLcholesterol and systolic
blood pressure

→ AUC=0.75

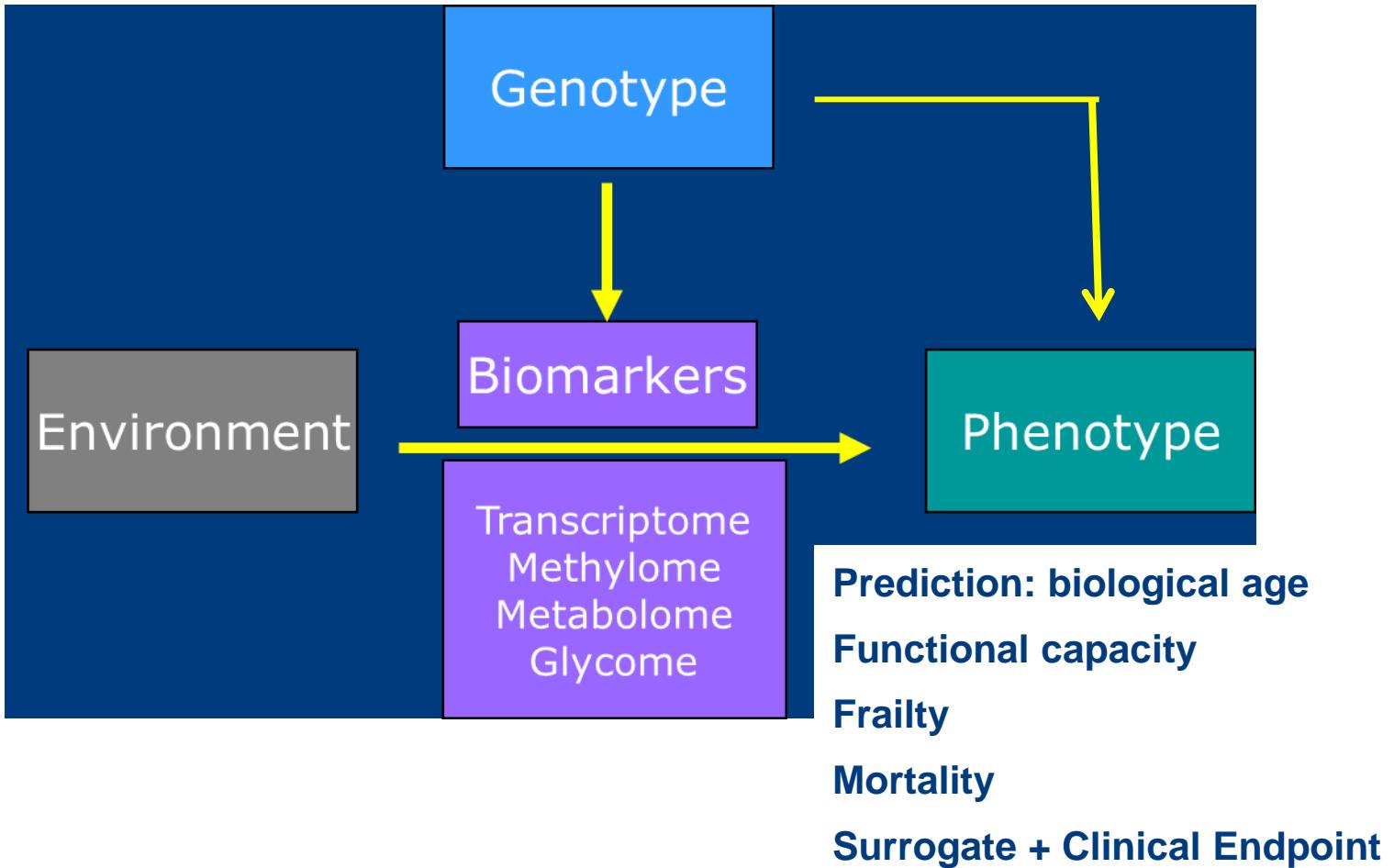
Prediction of disease risk

Using omics variables; how to understand the literature

Classification Table (Counts)

		Predicted Condition		Total
		Positive	Negative	
True Condition	Positive	True Positive (A)	False Negative (C)	A + C
	Negative	False Positive (B)	True Negative (D)	B + D
Total	A + B	C + D	A + B + C + D	

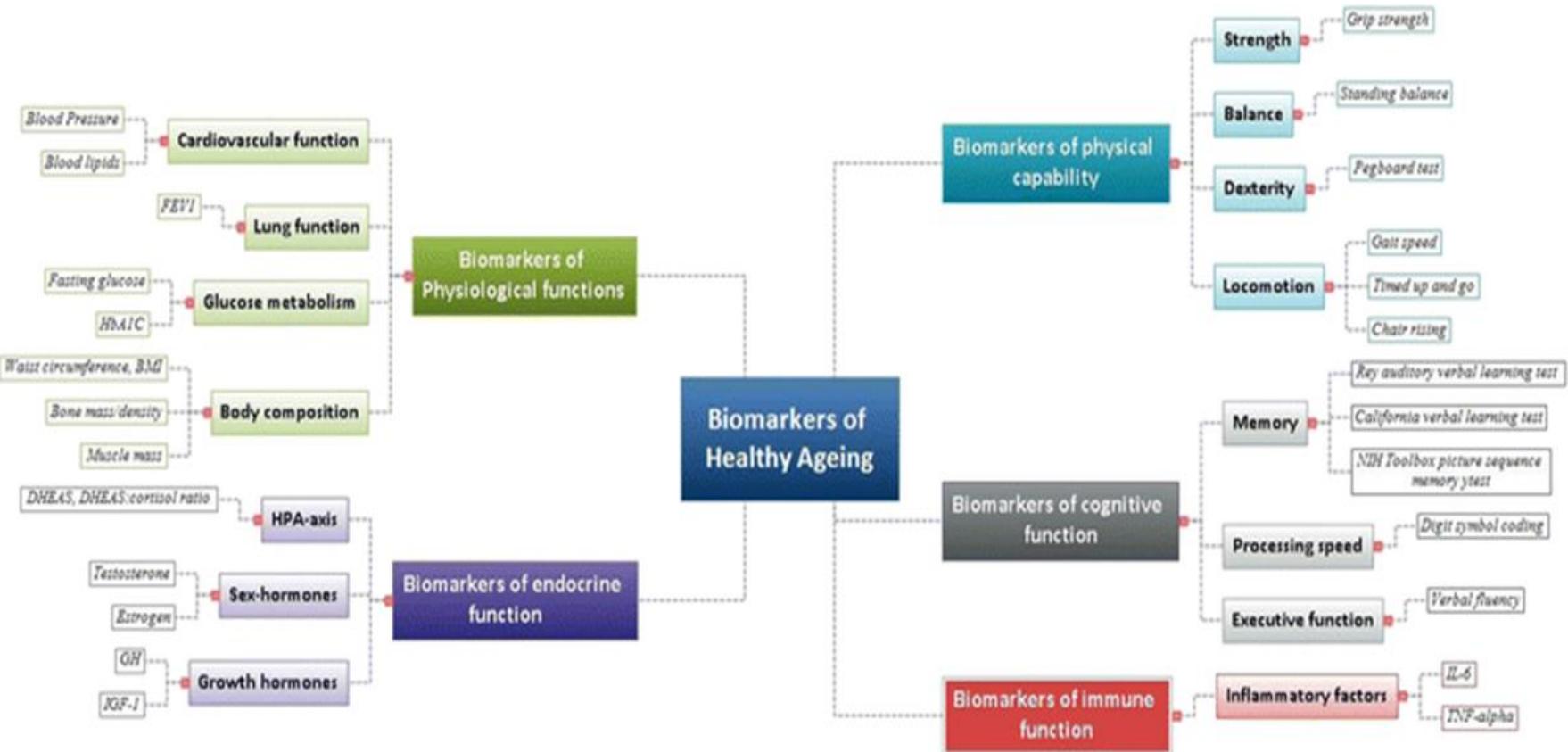
What types of phenotype or outcome ? In Ageing research



Biological Age Prediction (a score per individual) WHY

- to measure an individual's overall health status
- use as phenotype in genetic studies
- predict the risk of death , surrogate endpoint
- predict the risk of age-related disease incidence
- evaluate the effect of a health care management program
- evaluate the effect of lifestyle/management interventions

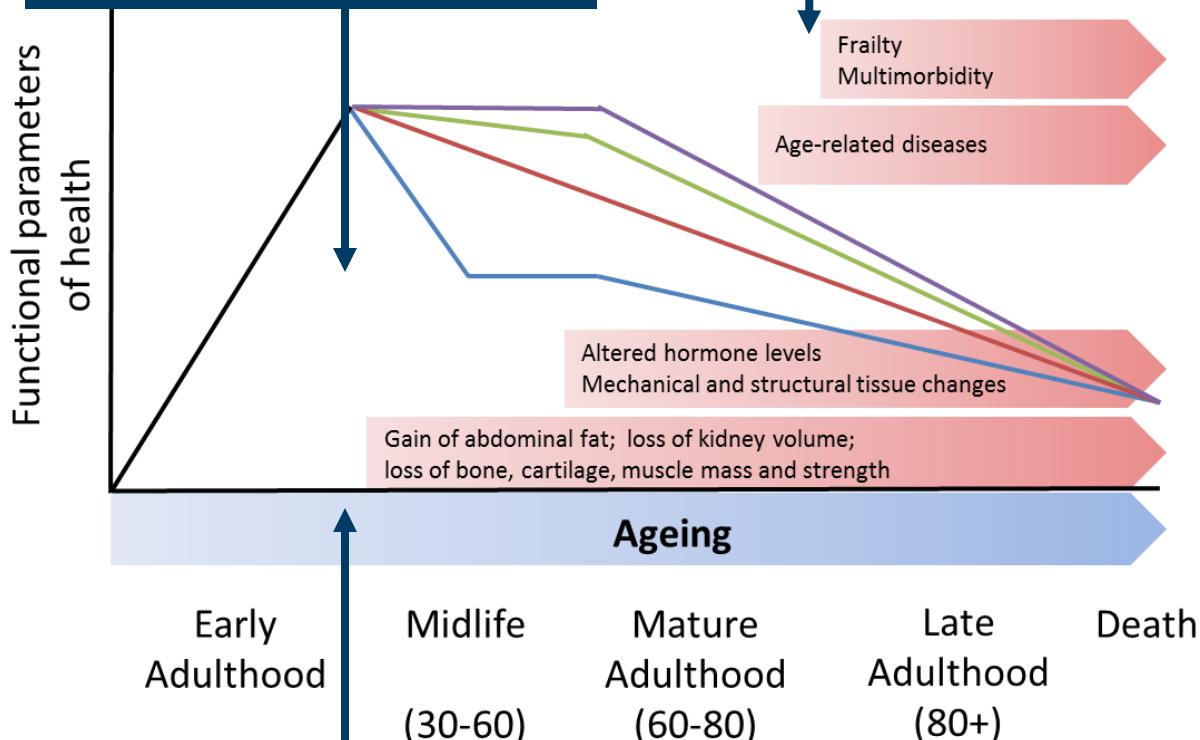
Domains of ageing organs/systems as proposed by Lara et al., BMC Medicine 2015



Lara proposes that biological age includes at least these domains.
Can we make a single score that covers these various domains ?

Multivariate combinations of biomarkers

NHANES Biological Age
Levine and Crimmins
Am J Hum Biol. 2014
Age >30



Frailty: an accumulation of deficits
Mitnitski AB, Mogilner AJ, Rockwood K.
Sci World J. 2001
Now: electronic health records
for hospital based frailty scores

Pace of Aging Belsky et al., PNAS 2015
Age 26-38, three timepoints
Read outs multiple organ systems
Gum health; also telomeres

Biological age – markers

Clinical variables (clocks) tracking age

Output multiple organ systems. Belsky et al., PNAS 2015 .

1000 subjects 38 year old. NHANES Biological Age index: 28-61 years

Longitudinal measures at 26, 32 and 38 years: personal rates of physiological decline

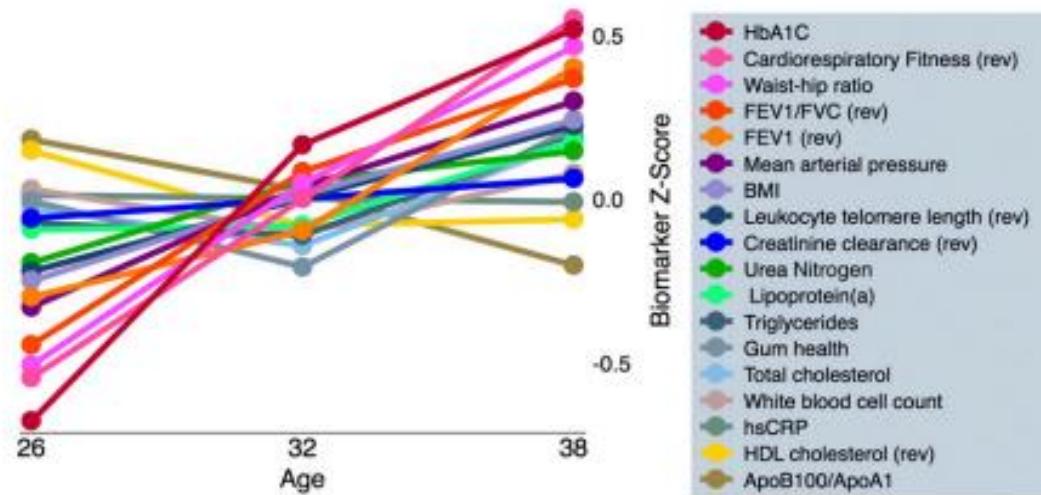
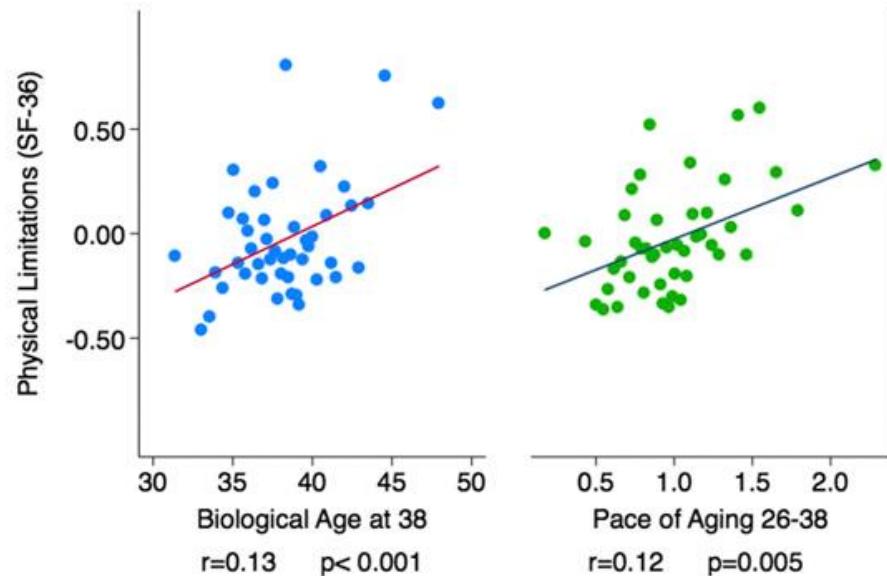
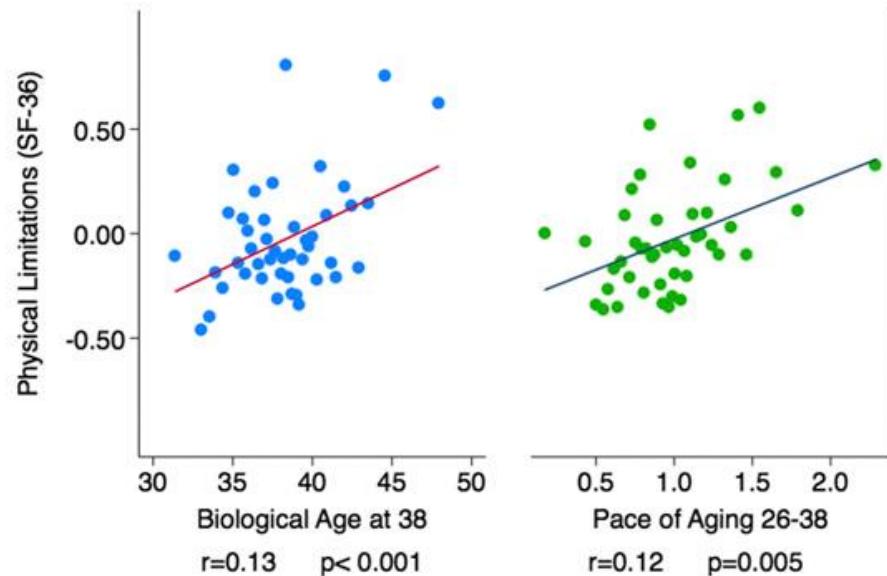
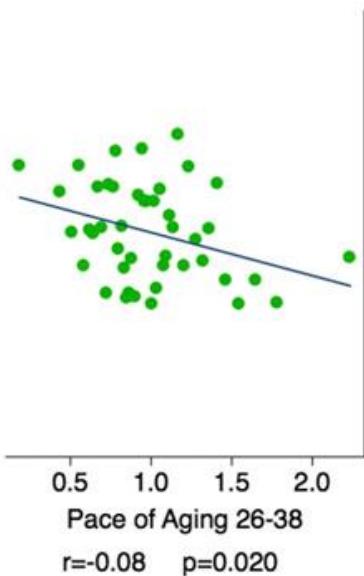
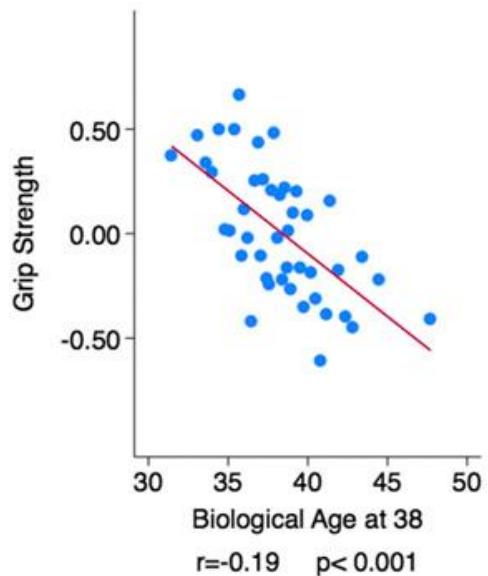
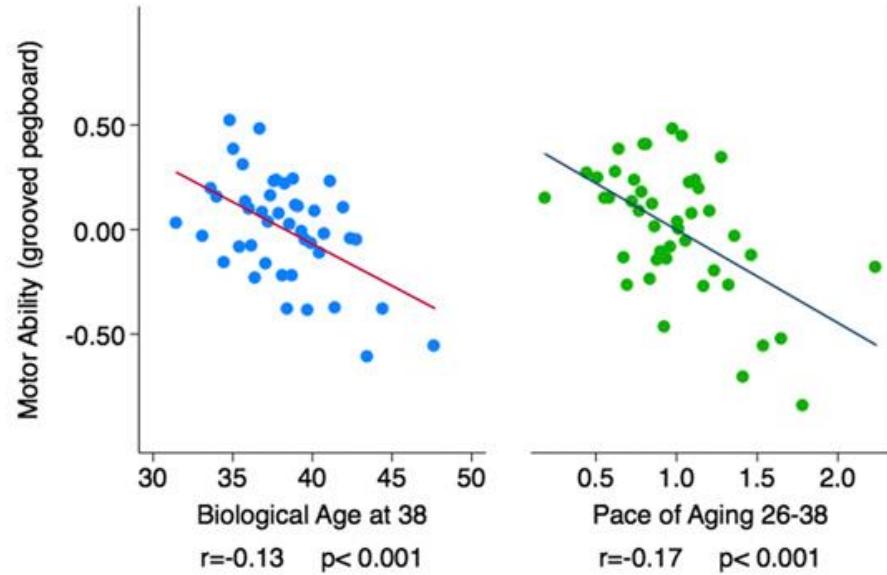
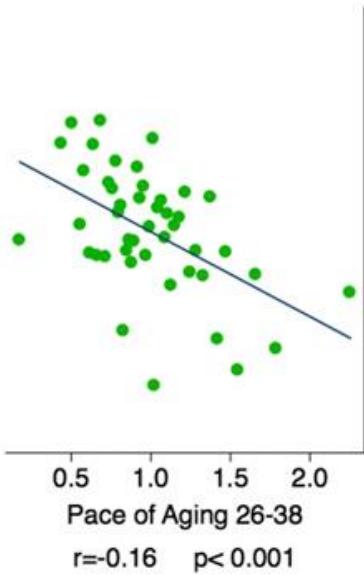
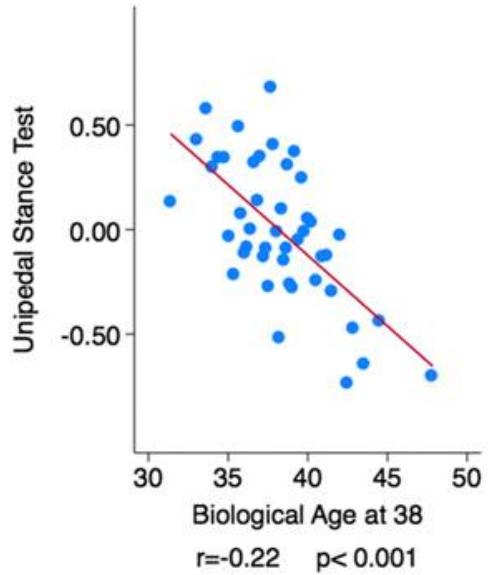
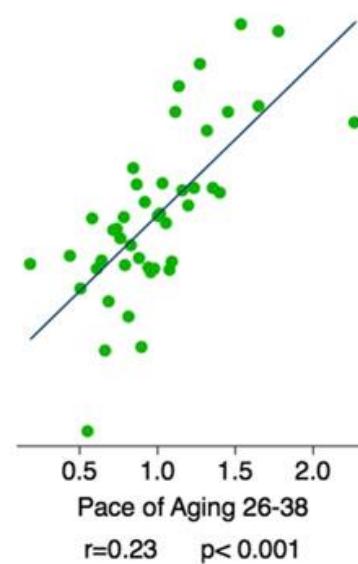
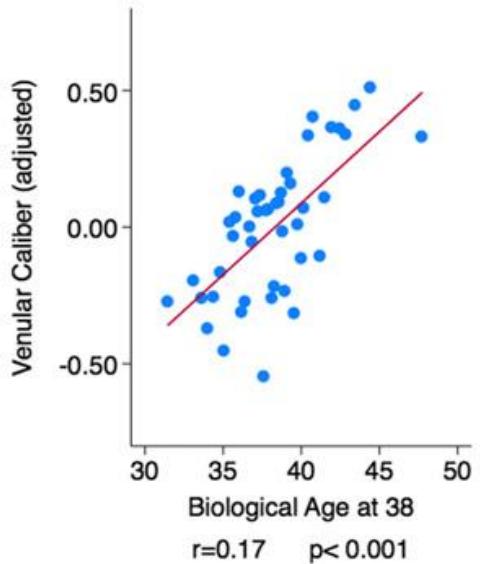
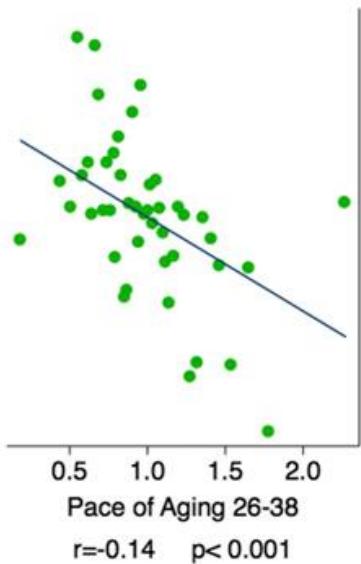
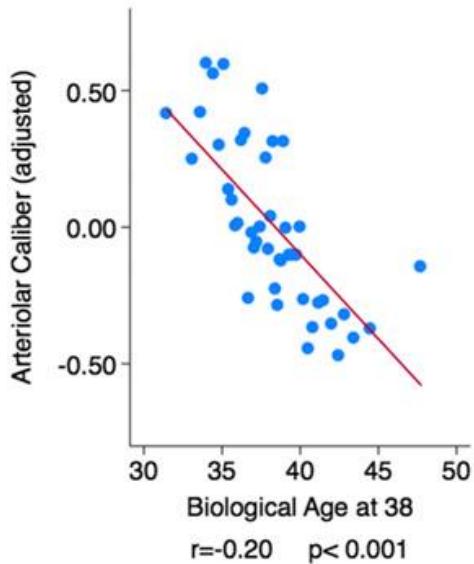
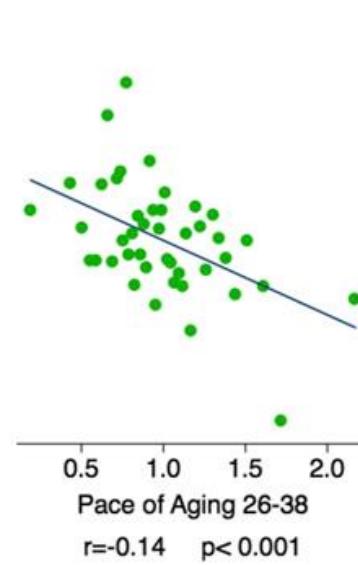
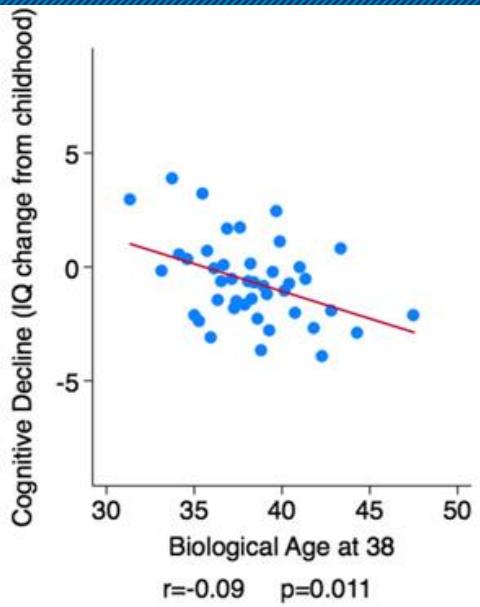
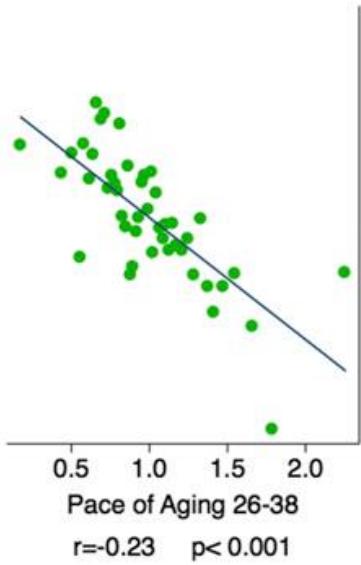
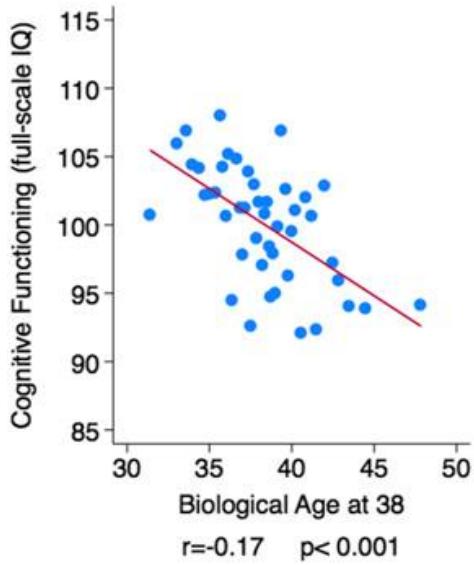


Fig. 3. Healthy adults exhibit biological aging of multiple organ systems over 12 y of follow-up. Biomarker values were standardized to have mean = 0 and SD = 1 across the 12 y of follow-up (Z scores). Z scores were coded so that higher values corresponded to older levels of the biomarkers; i.e., Z scores for cardiorespiratory fitness, lung function (FEV₁ and FEV₁/FVC), leukocyte telomere length, creatinine clearance, and HDL cholesterol, which decline with age, were reverse coded so that higher Z scores correspond to lower levels.

Physical capability



Cognition

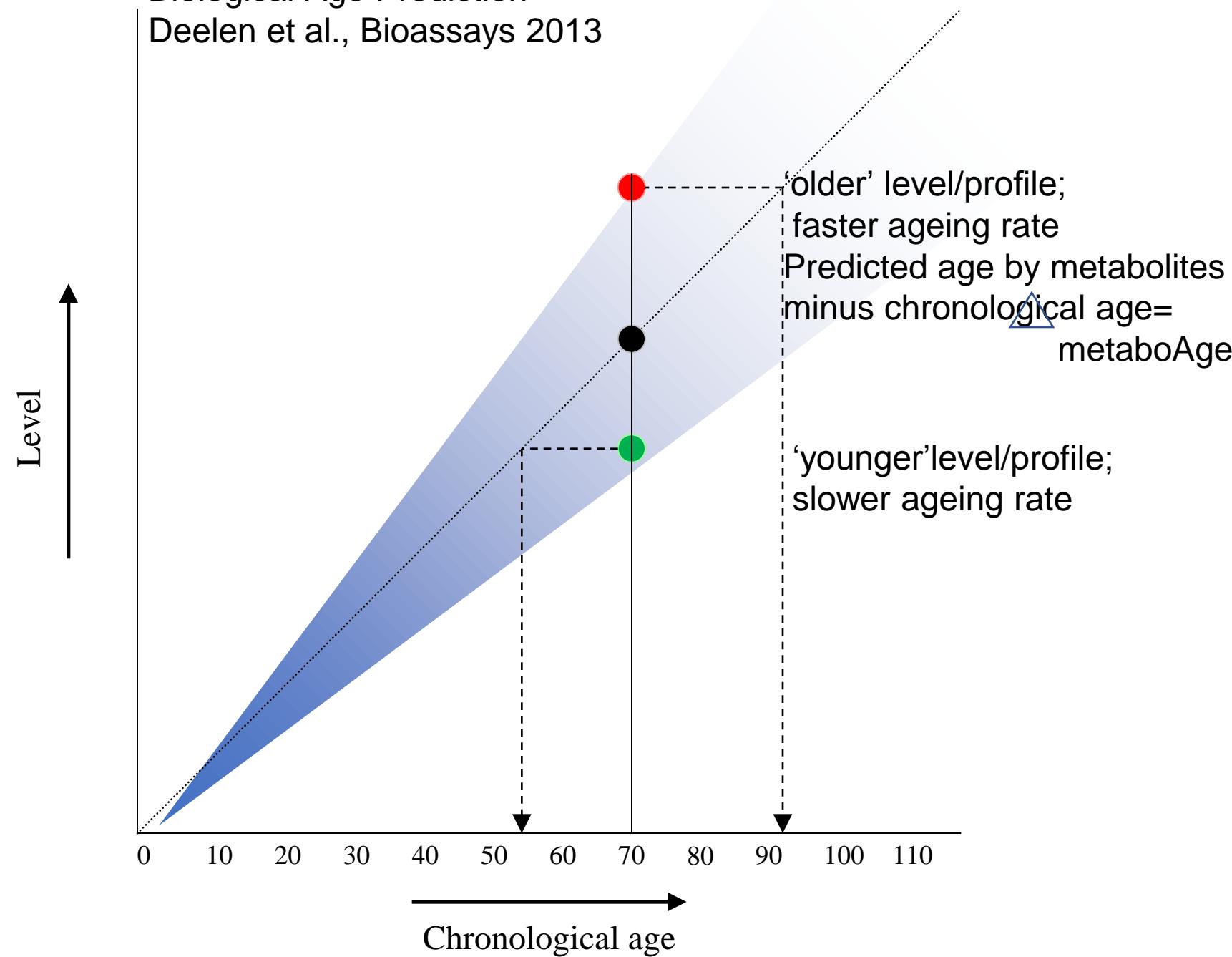


Remark

Others have tried to make predictors of biological age on the basis of the relation of molecular patterns and chronological age (mostly in cross sectional studies). This was done for DNA methylation, transcriptome and for metabolome.

Biological Age Prediction

Deelen et al., Bioassays 2013



Remark

You must prove that a predictor predicts a relevant outcome: if the predictor indicates biological age than it should associate with and predict onset of disease and mortality, for example.

Regularities with chronological age. DNA methylation (Hannum, Horvath)

DNAage predicts:

Mortality

Cancer risk

Obesity

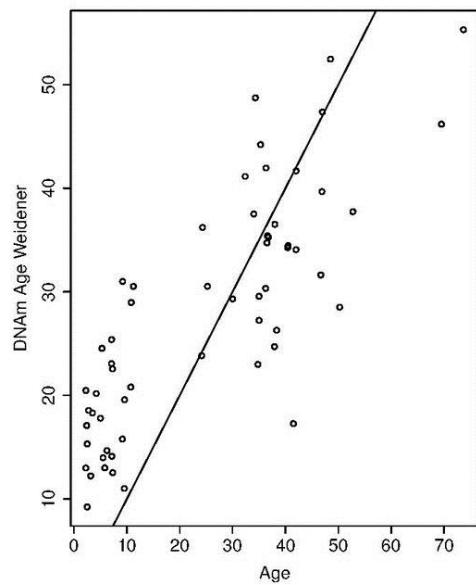
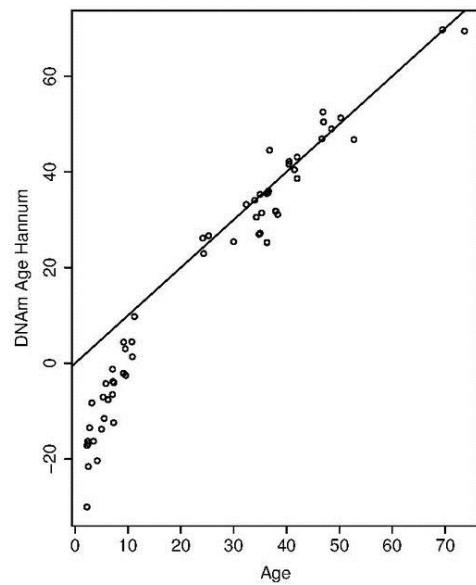
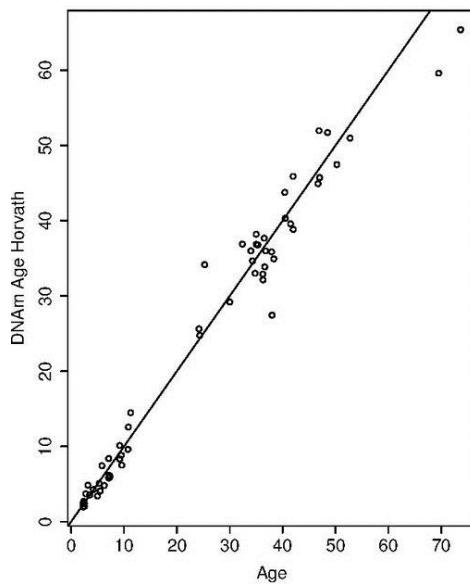
Gestational age

Brain tissue ageing

HR 1.02-1.04 Effect size for mortality:



C Weidener cor=0.81, p=7.9e-15



Remark

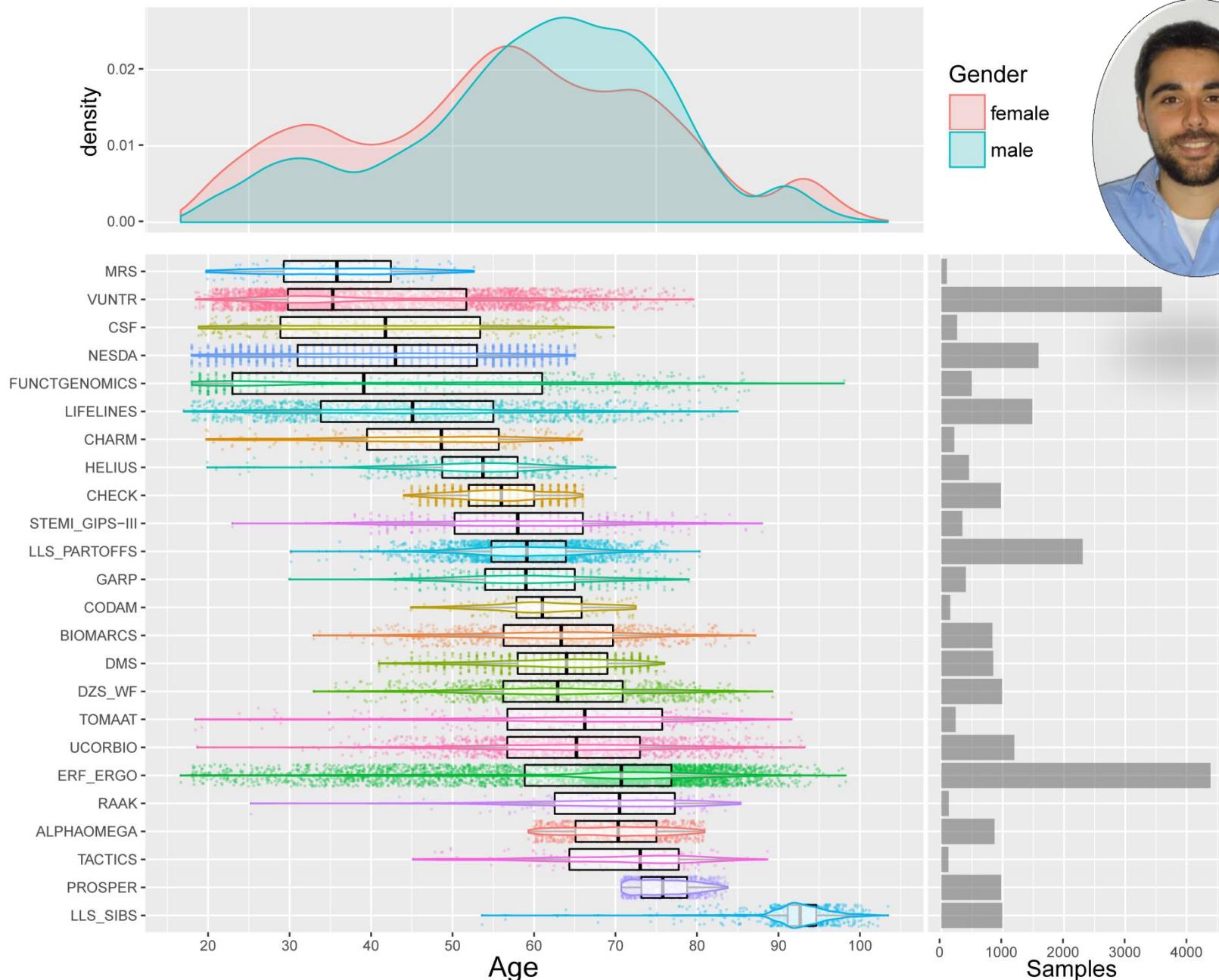
Two attempts are being made to make predictors from the metabolome.

1. Based on chronological age
2. Based on mortality

Study Population after Quality Control

BBMRI: 25 453 samples from 26 biobanks (under review)

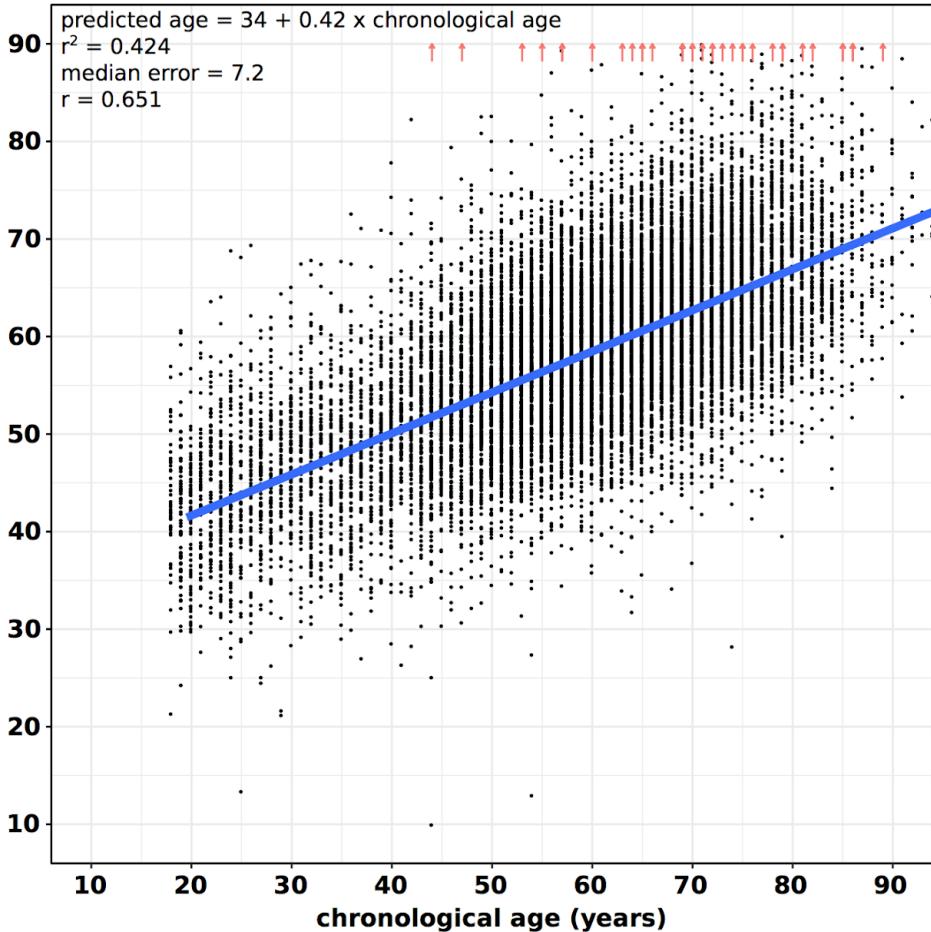
Erik van den Akker
Jurriaan Barkey-Wolf



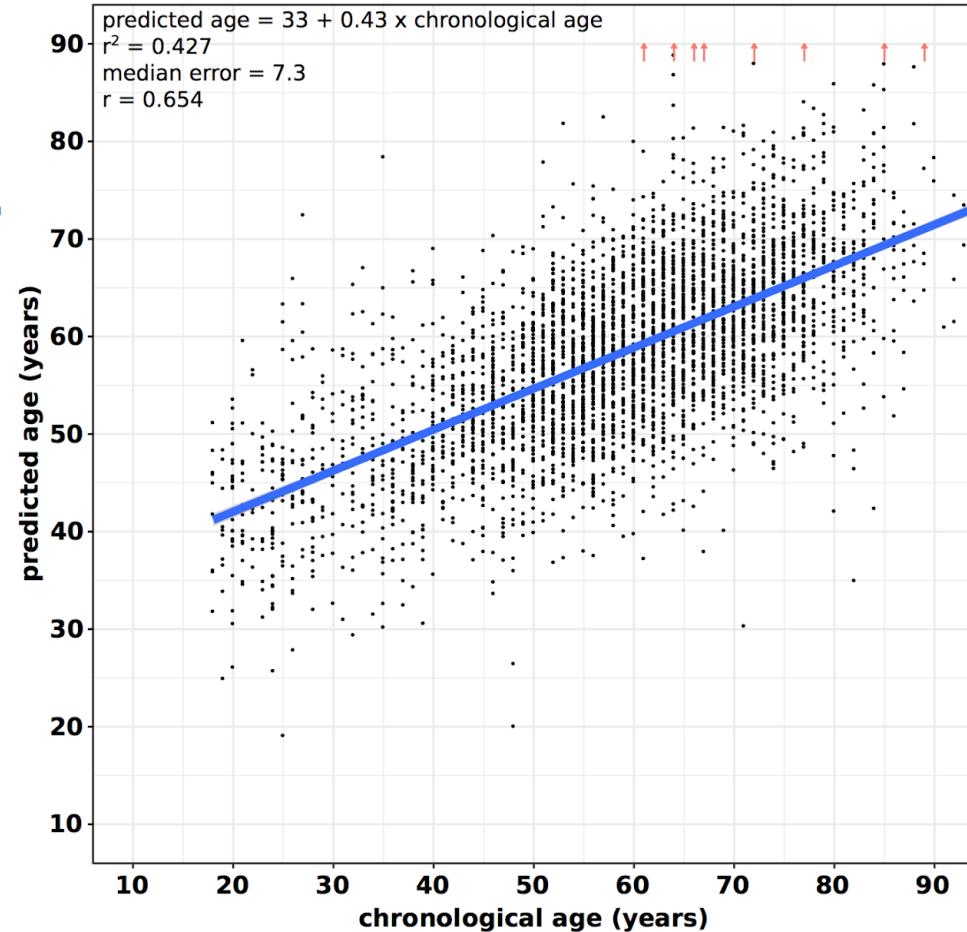
Training and evaluating the age-predictor

5 cross-fold validation

Training set (n = 15 208)



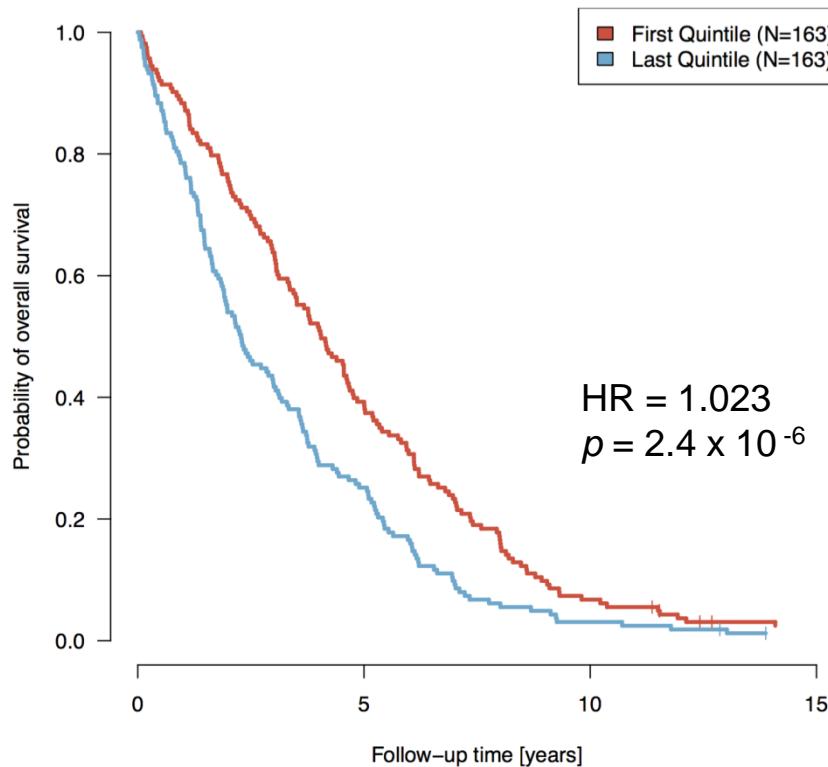
Test set (n = 3 802)



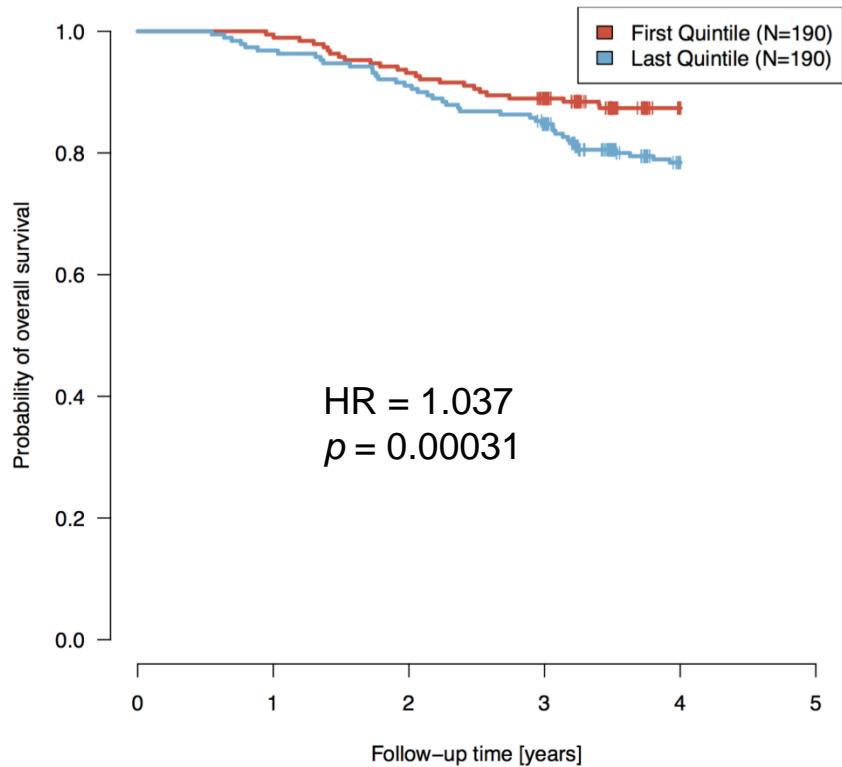
$$\text{predicted age} = \beta_0 + \beta_1 m_1 + \beta_2 m_2 + \dots + \beta_{56} m_{56} + \varepsilon$$

Associations with Future Mortality in longitudinal studies

Mortality analysis in the LLS



Mortality analysis in PROSPER



With every year older metaboAge: 2-4% increased mortality risk
Comparable to DNAmAge

Effect independent of BMI, smoking, diabetes and hypertension and medication

Remark

You can use the >200 metabolites in the ^1H NMR platform to make a predictor based on chronological age, but also based on prediction of mortality

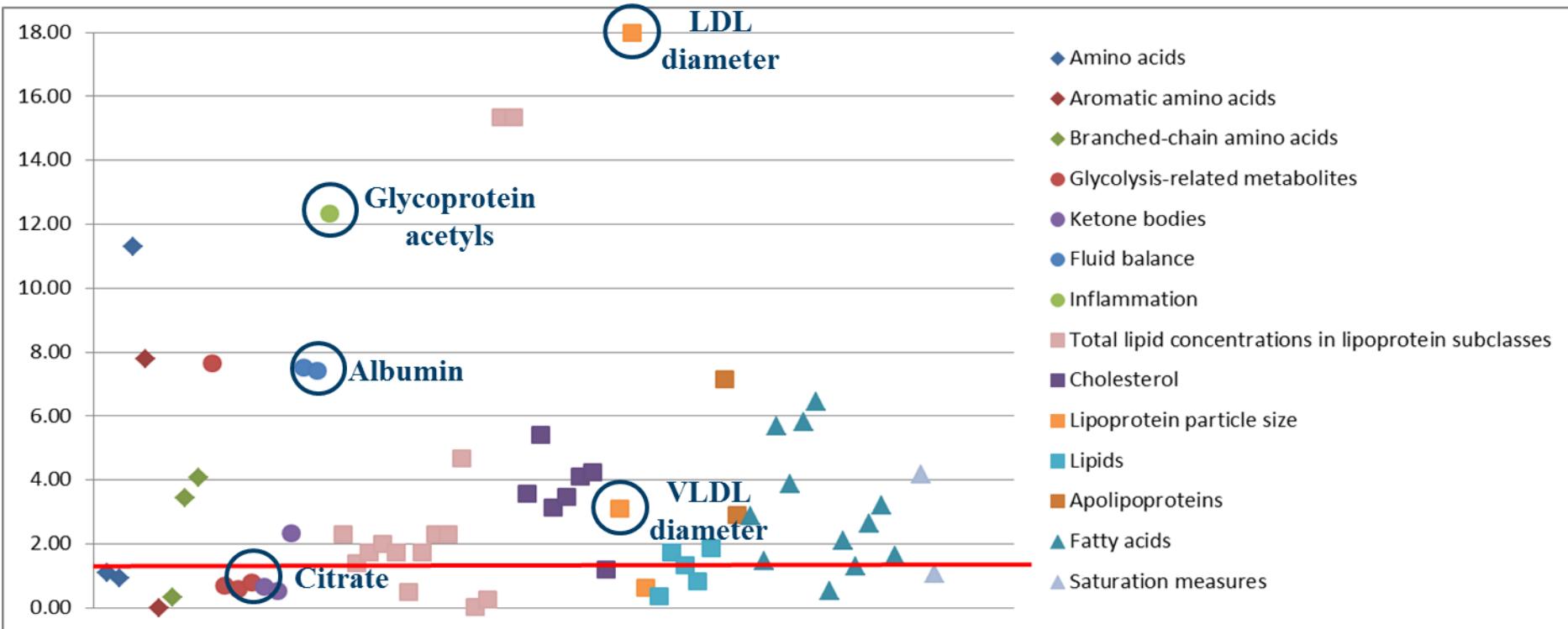
MORTALITY: ^1H NMR based analysis

44.000 subjects and 5.500 deaths

Study	N	Males (%)	Deaths
AlphaOmega	568	428 (75.4%)	157
ALSPAC	4,351	0 (0%)	17
EGCUT	10,988	4,106 (37.4%)	912
ERF	680	307 (45.1%)	107
FINRISK97	7,603	3,778 (49.7%)	1,213
FINRISK07	4,816	2,256 (46.8%)	190
KORA	1,790	871 (48.7%)	123
LLS nonagenarians	843	326 (38.7%)	823
LLS offspring + partners	2,241	999 (44.6%)	191
PROSPER	5,329	2,583 (48.5%)	467
Rotterdam Study	2,963	1,241 (41.9%)	1,254
TwinsUK	1,996	0 (0%)	58
	44,168		5,512

Metabolites associating to Mortality

Stepwise regression analysis following Fischer et al. 2014



¹H NMR metabolites independently predicting mortality

Metabolite	Full name	Hazard ratio	95% CI	P-value	I ²	P-value (heterogeneity)
XXLVLDLL	Extremely large VLDL	0.80	0.75 - 0.85	1.53 x 10 ⁻¹³	0.08	0.363
SHDLL	Small HDL	0.87	0.84 - 0.90	5.98 x 10 ⁻¹⁹	0.52	0.018
VLDLD	VLDL diameter	0.85	0.80 - 0.90	8.51 x 10 ⁻⁸	0.21	0.241
PUFAFA	Polyunsaturated fatty acids (%)	0.78	0.75 - 0.80	1.06 x 10 ⁻⁴⁷	0.71	8.65 x 10 ⁻⁵
Glc	Glucose	1.16	1.13 - 1.19	2.22 x 10 ⁻²⁹	0.56	0.008
Lac	Lactate	1.06	1.03 - 1.10	6.24 x 10 ⁻⁵	0.28	0.173
His	Histidine	0.93	0.90 - 0.96	1.15 x 10 ⁻⁵	0.24	0.213
Ile	Isoleucine	1.23	1.14 - 1.32	2.14 x 10 ⁻⁸	0.39	0.078
Leu	Leucine	0.82	0.76 - 0.89	7.34 x 10 ⁻⁷	0.35	0.109
Val	Valine	0.87	0.82 - 0.92	1.04 x 10 ⁻⁶	0.07	0.376
Phe	Phenylalanine	1.13	1.09 - 1.17	2.39 x 10 ⁻¹²	0.44	0.052
AcAce	Acetoacetate	1.08	1.05 - 1.11	1.73 x 10 ⁻⁸	0.35	0.108
Alb	Albumin	0.89	0.87 - 0.92	9.96 x 10 ⁻¹³	0.52	0.017
Gp	Glycoprotein acetyls	1.32	1.27 - 1.38	7.45 x 10 ⁻⁴¹	0.45	0.046

1H NMR metabolites independently predicting mortality

Traditional vs 14 metabolites score :

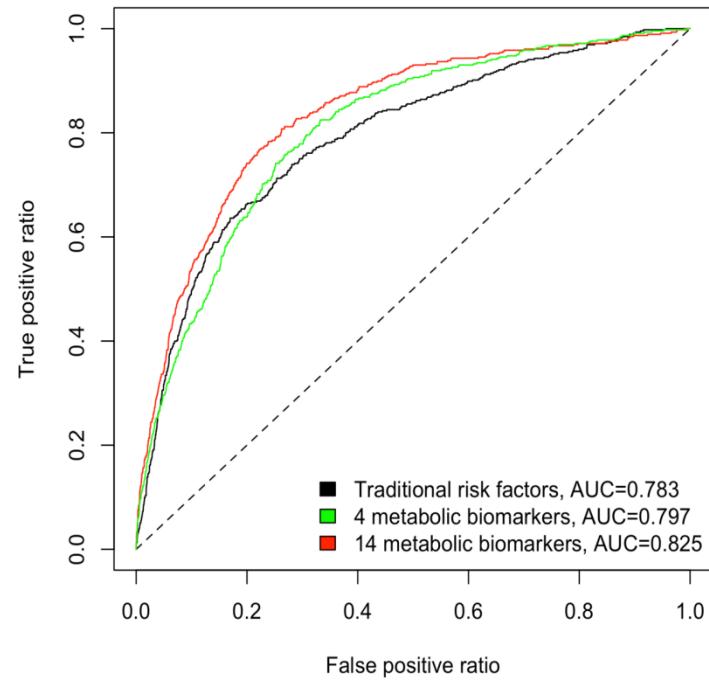
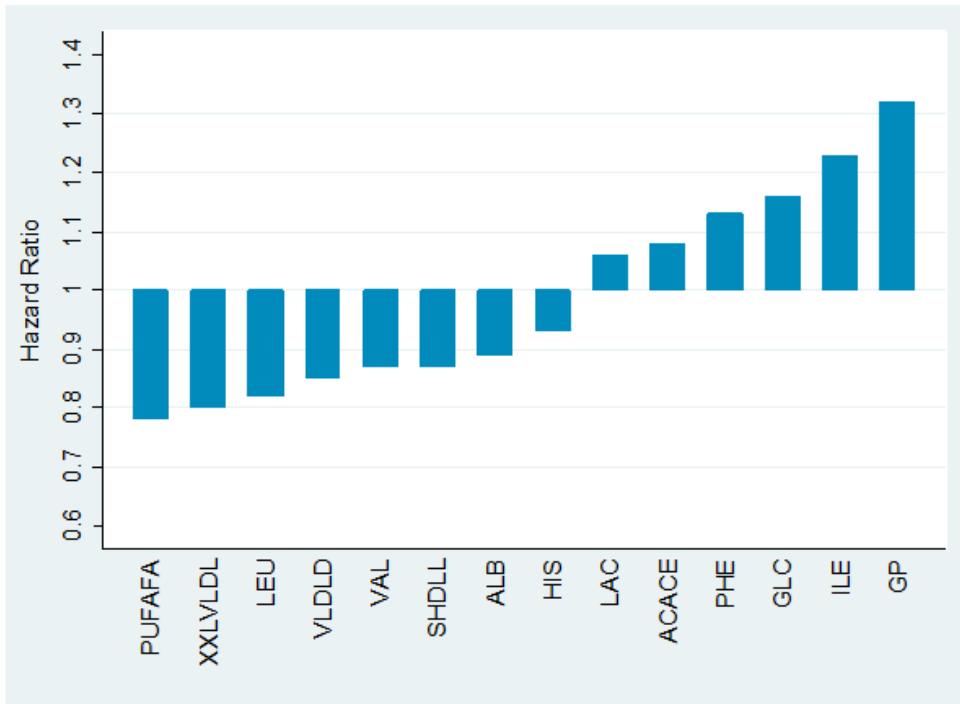
5 years >60

0,617 → 0,812

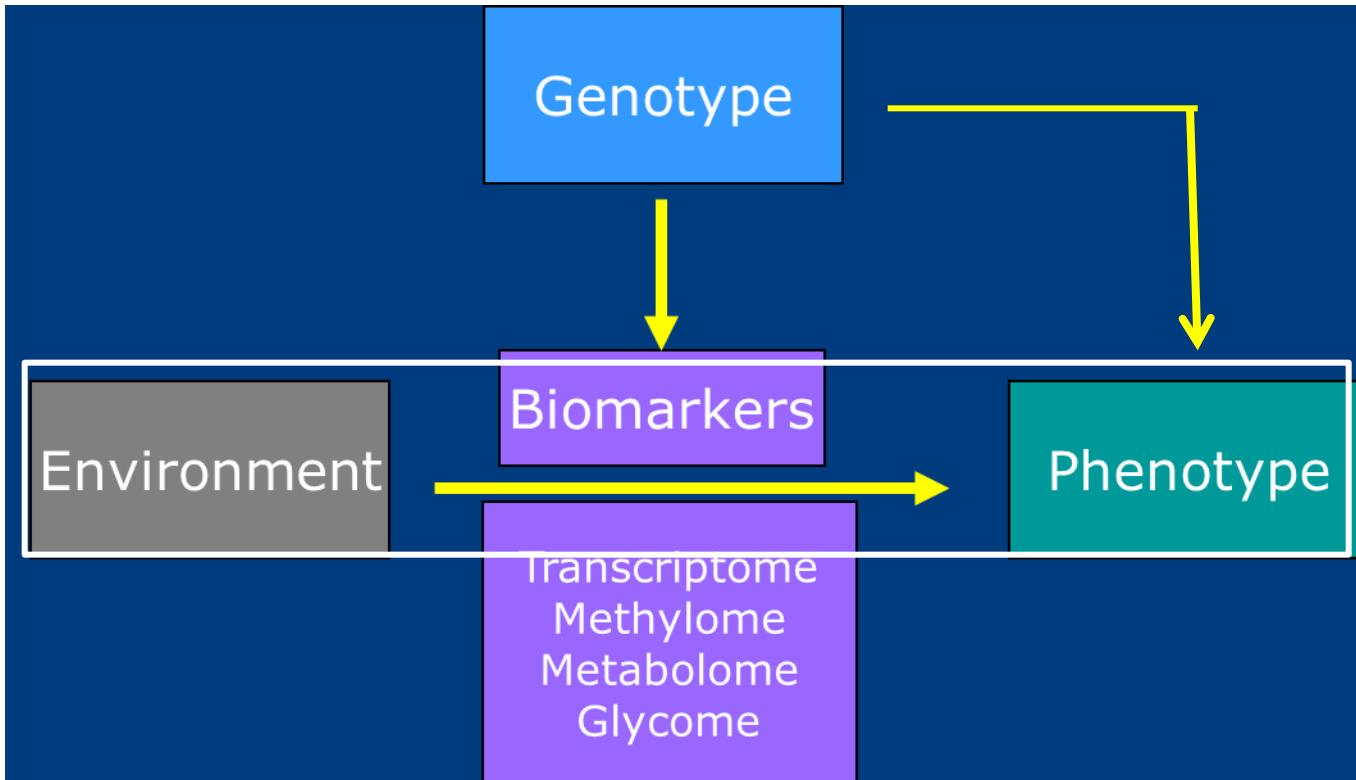
0.195±0.051 P=1.44 x 10-4

Cancer/CVD/infections

Independent effects of 14 metabolites on mortality



What relations do scientists investigate with molecular profiling data



Molecular Epidemiology

- Introduced by Kilbourne (1973), infectious diseases; Schulte and Perera (1993 Principles and Practices)
- Integrates Epidemiology, Medical Sciences and Molecular Biology
- Studies the influence on health of environmental and genetic risk factors measured by (holistic) molecular signatures
- Contributes to
 - prediction/prognosis
 - monitoring exposure, response to interventions**
 - etiological understanding (disease mechanisms)

Biomarkers II

or NIH biomarker working group:
a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention

Does the Metabolite profile Monitor Lifestyle intervention v.d. Rest et al., 2016

13-weeks intervention (164 subjects)

- 12.5% reduced dietary intake
- 12.5% increased physical activity

v.d. Rest et al., 2016

Baseline

End



Before baseline:

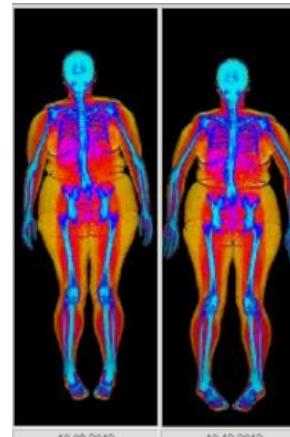
- FFQ
- IPAQ
- Accelerometer
- 24 hr blood glucose

During intervention

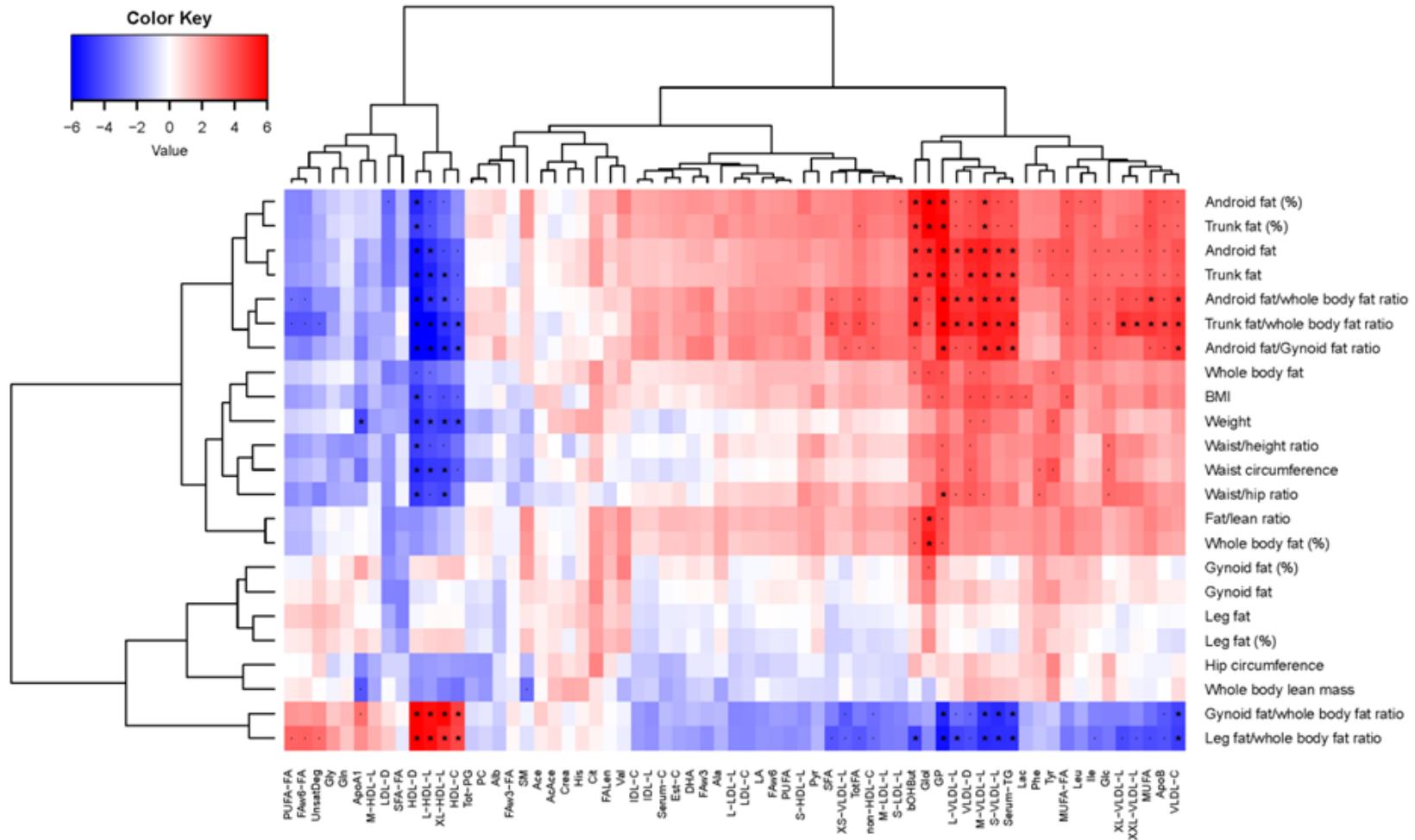
- 4 x 24h recalls
- Diary
- Hunger and satiety questionnaires

One week before end:

- Accelerometer
- 24 hr blood glucose



Which metabolites associate at baseline with DEXA scan variables of fat distribution. Heat map



Exercise

Biomarkers of biological age:

DNA methylation

PACE of Ageing

Telomeres

Questions for the Discussion on Biological Age Predictors Belsky and Marioni paper

- 1. What is the main aim of the paper.**
- 2. Why, what is the problem?**
- 3. What new research strategy would solve the problem**
- 4. What questions are being asked specifically. What is measured in which study population and design for these questions.**

Question 1:

Question 2:, etc.

- 5. What answers were found to the questions**

Answer to Q 1:

Answer to Q2: , etc.

- 6. What do you think of the paper ?**

- 7. What are the main conclusions and the main discussion points the authors bring up.**

Biological age – markers

Clinical variables (clocks) tracking age

Output multiple organ systems. Belsky et al., PNAS 2015 .

1000 subjects 38 year old. NHANES Biological Age index: 28-61 years

Longitudinal measures at 26, 32 and 38 years: personal rates of physiological decline

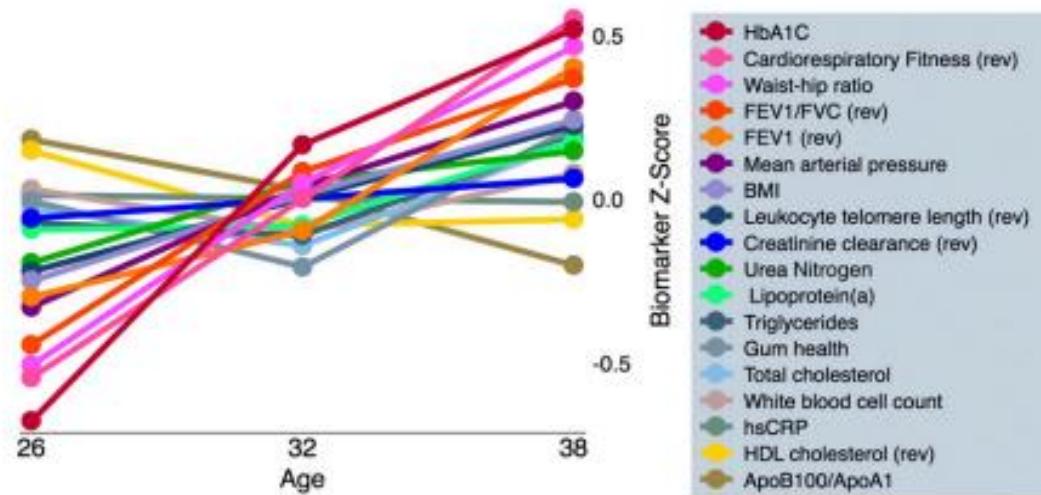


Fig. 3. Healthy adults exhibit biological aging of multiple organ systems over 12 y of follow-up. Biomarker values were standardized to have mean = 0 and SD = 1 across the 12 y of follow-up (Z scores). Z scores were coded so that higher values corresponded to older levels of the biomarkers; i.e., Z scores for cardiorespiratory fitness, lung function (FEV₁ and FEV₁/FVC), leukocyte telomere length, creatinine clearance, and HDL cholesterol, which decline with age, were reverse coded so that higher Z scores correspond to lower levels.

Regularities with chronological age. DNA methylation (Hannum, Horvath)

DNAage predicts:

Mortality

Cancer risk

Obesity

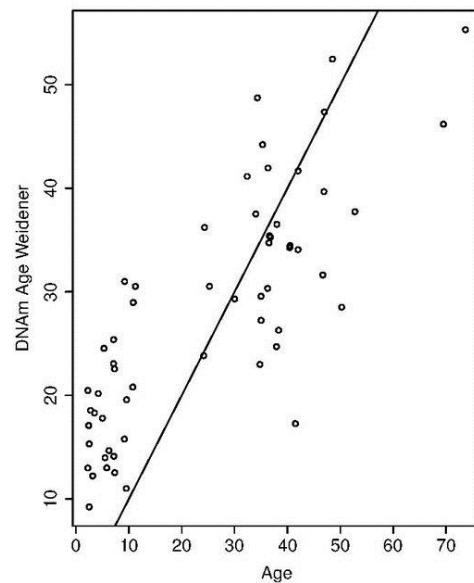
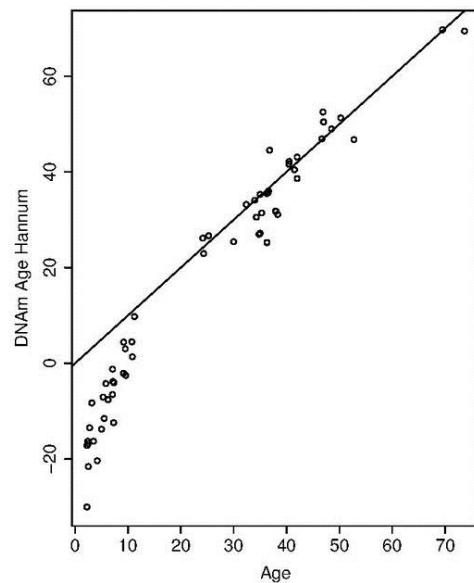
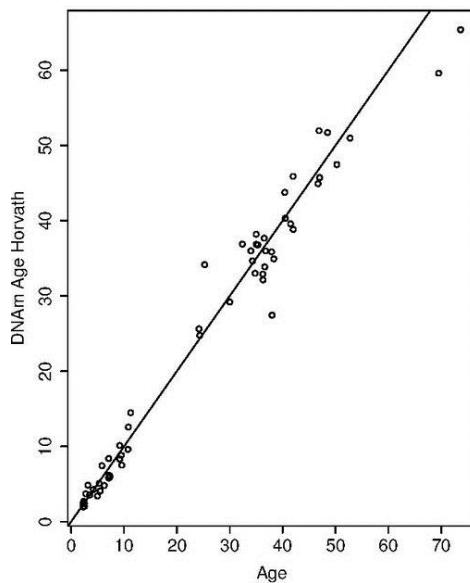
Gestational age

Brain tissue ageing

HR 1.02-1.04 Effect size for mortality:



C Weidener cor=0.81, p=7.9e-15



Telomere shortening, age and mortality prediction

