LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Universität Augsburg Medizinische Fakultät

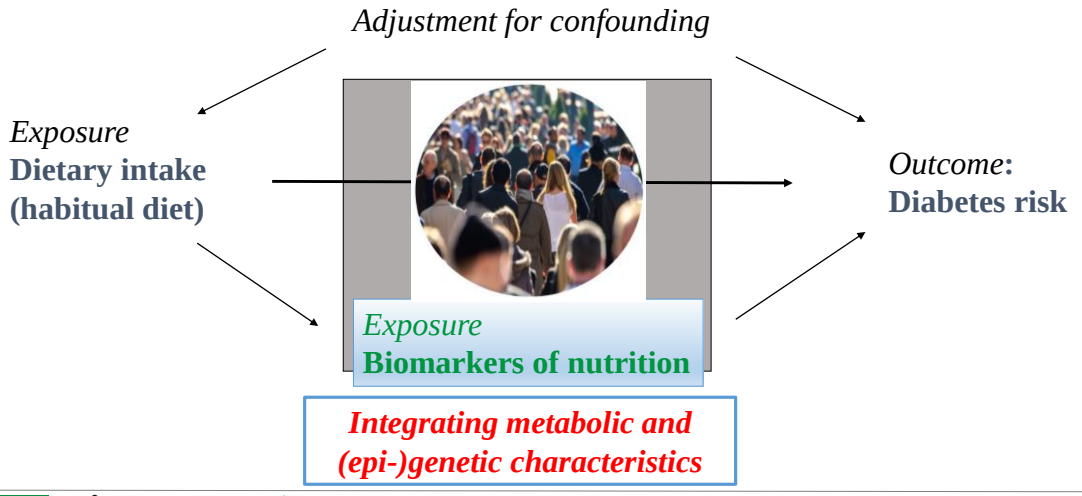
UNIVERSITÄTSKLINIKUM AUGSBURG
Lehrstuhl für Epidemiologie

Identification of diet-disease associations in metabolically homogenuous subgroups

Jakob Linseisen

National Academies of Sciences, Engineering, Medicine;
Standing Committee on Evidence Synthesis, and Communications in Diet and Chronic Disease Relationships;
August 31, 2021

Considering biologic variation in nutritional epidemiology



Gene-diet interaction analysis gives limited information

Gene-lifestyle interaction on risk of type 2 diabetes: A systematic review

Stefan Dietrich[1] | Simone Jacobs[1] | Ju-Sheng Zheng
Lukas Schwingshackl[5] | Matthias B. Schulze[1,4,6]

Summary

The pathophysiological influence of gene-lifestyle interactions on the risk to develop type 2 diabetes (T2D) is currently under intensive research. This systematic review summarizes the evidence for gene-lifestyle interactions regarding T2D incidence. MEDLINE, EMBASE, and Web of Science were systematically searched until 31 January 2019 to identify publication with (a) prospective study design; (b) T2D incidence;

on; 28 ary and ne- RS, ys- and R2,

ADRA2B, *TNFα*, and *LIPC* variants. However, most findings represent single study findings obtained in European ethnicities. Although some interactions have been reported, their conclusiveness is still low, as most findings were not yet replicated across multiple study populations.

Conclusions:

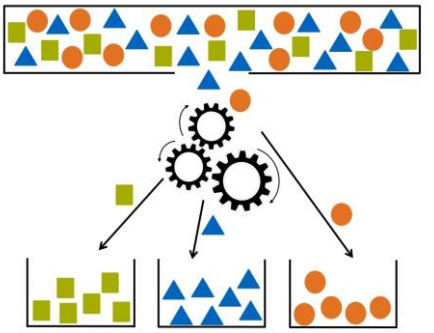
- Most findings represent single study findings obtained in European ethnicities.
- Although some interactions have been reported, their conclusiveness is still low, as most findings were not yet replicated across multiple study populations.

(Obesity Reviews 2019)

Metabolic phenotyping

Alternatively, define phenotypes which are likely to reflect specific genetic and metabolic characteristics

→ e.g., „Metabotypes“



Metabolically homogenuous subgroups

The starting point: Identification of comprehensive metabotypes (KORA F4)

Database for the definition of 'metabotypes'

- Standard clinical markers (e.g. plasma lipids, liver enzymes)
- Data on metabolomics
- Genetic data (SNPs from GWAs)
- Epigenetic data (CpG methylation from GWAs)
- Proteomics data

} -omics data

Grouping methods

- K-means cluster analysis
- Hierarchical cluster analysis
- Self-organizing maps
- Gaussian mixture models
- DBSCAN/OPTICS, other bioinformatic methods
- Etc.

(Riedl, (…), Linseisen, Mol Nutr Food Res 2018)

Identification of metabotypes,

based on 32 biochemical parameters

by K-means

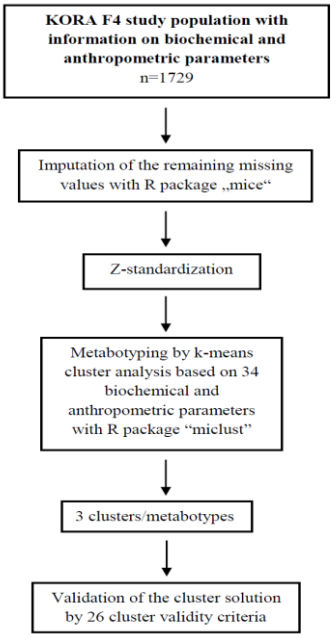
3 clusters

Whole EDTA blood	Serum	Plasma	Urine
- **Glycated hemoglobin** - **Leukocytes** - Average telomere length in leukocytes	- **Glucose** - **Total cholesterol** - **High density lipoprotein** - **Total cholesterol/high density lipoprotein ratio** - **Low density lipoprotein** - **Triglycerides** - **Uric acid** - **Insulin** - **High-sensitive C-reactive protein** - **Gamma-glutamyltransferase** - **Glutamate-pyruvate transaminase** - **Glutamate-oxaloacetate transaminase** - **Alkaline phosphatase** - Cystatin C - Interleukin-18 - Thyroid peroxidase antibodies - Thyroid-stimulating hormone - Free thyroxine	- Lipoprotein(a) - Apolipoprotein A-IV - Afamin - Leptin - Sex-hormone-binding globulin - Renin - Aldosterone - Non-esterified fatty acids - Insulin-like growth factor-I - Insulin-like-growth-factor-binding-protein-3	- Creatinine - Albumin

All parameters were measured in KORA F4 participants; only the **bold parameters** were measured in KORA FF4 participants.

(Riedl, (…), Linseisen, Mol Nutr Food Res 2018)

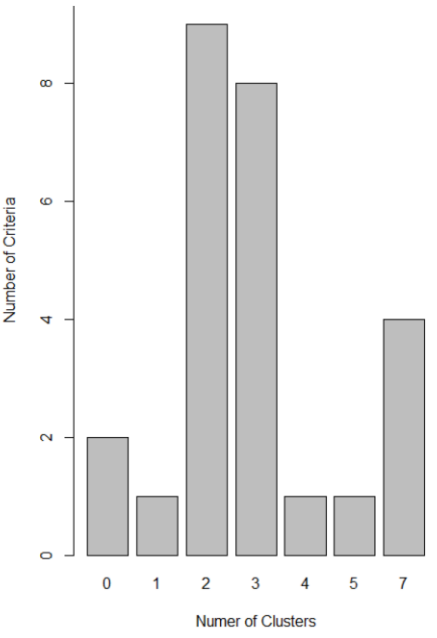
Overview of the analysis approach, KORA F4 study



(Riedl, (…), Linseisen, Mol Nutr Food Res 2018)

Number of clusters

chosen by 26 cluster validity criteria using the R package ‘NbClust’, v. 3.0



(Riedl, (…), Linseisen, Mol Nutr Food Res 2018)

Description of metabotype subgroups:
Medians (25th, 75th percentile) of selected biochemical parameters in KORA FF4

	Total N=3001	Cluster 1 N=1189	Cluster 2 N=1440	Cluster 3 N=372	P value
BMI [kg/m²]	26.9 (5.93)	24.19 (4.31) [a]	28.24 (4.85) [b]	**33.17** (6.47) [c]	<0.001
Uric Acid [µmol/l]	299.41 (114)	243.5 (79) [a]	334.12 (94) [b]	**375** (114) [c]	<0.001
Glucose [mg/dl]	94 (14)	89 (10.80) [a]	96 (12.40) [b]	**122.5** (28.40) [c]	<0.001
HDLc [mmol/l]	1.39 (0.52)	1.70 (0.45) [a]	1.26 (0.37) [b]	**1.18** (0.39) [c]	<0.001
Non-HDLc [mmol/l]	4.05 (1.32)	3.59 (1.21) [a]	**4.49** (1.28) [b]	4.01(1.25) [c]	<0.001
TG [mmol/ l]	1.19 (0.90)	0.82 (0.56) [a]	1.46 (0.87) [b]	**1.76** (1.20) [c]	<0.001
AP [µmol/l]	1.10 (0.42)	1 (0.43) [a]	1.14 (0.39) [b]	**1.21** (0.43) [c]	<0.001
GPT [µkat/l]	0.35 (0.23)	0.28 (0.16) [a]	0.40 (0.24) [b]	**0.48** (0.32) [c]	<0.001
GOT [µkat/l]	0.41 (0.14)	0.38 (0.12) [a]	0.43 (0.14) [b]	**0.45** (0.19) [c]	<0.001
GGT [µkat/l]	0.43 (0.4)	0.31 (0.26) [a]	0.50 (0.42) [b]	**0.63** (0.53) [c]	<0.001
HbA1c [%]	5.50 (0.5)	5.30 (0.42) [a]	5.50 (0.5) [b]	**6.10** (0.98) [c]	<0.001
hs-CRP [mg/L]	1.18 (2.03)	0.76 (1.38) [a]	1.38 (2.08) [b]	**2.49** (3.27) [c]	<0.001
Leukocytes (n/L)	5.70 (2)	5.30 (1.84) [a]	5.90 (1.93) [b]	**6.30** (2.00) [c]	<0.001

LMU · UNIVERSITÄTSKLINIKUM AUGSBURG · HelmholtzZentrum münchen, Deutsches Forschungszentrum für Gesundheit und Umwelt

Description of metabotype subgroups:
Incidence of metabolic diseases (during 7 years of follow-up)

	Hypertension	Diabetes	Hyperuricemia/ Gout	Dyslipidemia	At least one metabolic disease
Cluster 1	18.2%	2.0%	0.7%	5.6%	19.7%
Cluster 2	25.3%	5.1%	2.3%	18.3%	36.8%
Cluster 3	41.2%	28.3%	10.8%	19.2%	54.5%

→ Metabotype 3 is associated with a very high risk of incident diabetes

LMU · UNIVERSITÄTSKLINIKUM AUGSBURG · HelmholtzZentrum münchen, Deutsches Forschungszentrum für Gesundheit und Umwelt

(Riedl, (…), Linseisen, Mol Nutr Food Res 2018)

Exploring diet-diabetes associations (KORA FF4)

Exposure: habitual food or nutrient intake (blended approach)

- fruits
- vegetables
- potatoes
- total meat
- red meat
- poultry
- processed meat
- eggs
- total dairy
- milk
- yogurt
- cheese
- coffee
- fruit and vegetable juice
- sugar sweetened beverages (SSB)
- alcohol
- total fiber

Diabetes/glucose tolerance status in cross-sectional analyses

Normal glucose tolerance + Prediabetes (IFG, IGT, or both) vs. Undetected type 2 diabetes + Prevalent type 2 diabetes

Diet-diabetes associations, by metabotypes

	Cluster 1/Cluster 2 N=1217		Cluster 3 N=300		
Food or Nutrient	OR	95% CI	OR	95% CI	P-value interaction[b]
Fruits (50 g/d)	0.91	0.75 - 1.11	**0.83** ↓	**0.68 - 0.99**	0.24
Total meat (50 g/d)	**1.67** ↑	**1.04 - 2.67**	1.51	0.95 - 2.41	0.16
Processed meat (50 g/d)	**2.23** ↑	**1.24 - 4.04**	1.79	1.00 - 3.26	0.11
Sugar sweetened beverages (50 g/d)	0.92	0.73 - 1.08	**1.21** ↑	**1.09 - 1.35**	**0.01**

[b] P-value of likelihood ratio test for comparison of models with and without the interaction term of metabotype and the respective food item or nutrient; Significant results ($p<0.05$) printed in **bold**. Fully adjusted models

Specific advice to cluster-3 participants:
Decrease consumption of sugar-sweetened beverages

(Riedl, (…), Linseisen, Eur J Nutr 2020)

Dietary pattern analysis

Usual intake estimates of food groups were used to derive dietary pattern:

- Alternate Healthy Eating Index (AHEI)
- Mediterranean Diet Score (MDS)

(Wawro, (…), Linseisen, Nutrients 2020)

Dietary pattern and diabetes risk, by metabotype (KORA FF4, n=1287)

	Diabetes Mellitus								
	Total (n=145)[d]		Cluster 1 (n=25)[e]		Cluster 2 (n=32)		Cluster 3 (n=88)		
AHEI	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**p-interaction**
Model 1[a]	**0.75**	**0.57 - 0.97**	1.47	0.82 - 2.62	0.71	0.47 - 1.18	**0.61**	**0.41 - 0.90**	**0.031**
Model 2[b]	0.79	0.60 - 1.04	1.52	0.83 - 2.78	0.83	0.49 - 1.41	**0.60**	**0.39 - 0.90**	**0.022**
Model 3[c]	0.87	0.65 - 1.15	1.73	0.93 - 3.23	0.95	0.55 - 1.63	**0.60**	**0.39 - 0.92**	**0.035**
MDS									
Model 1[a]	**0.72**	**0.55 - 0.94**	1.29	0.75 - 2.21	0.61	0.37 - 1.01	**0.59**	**0.39 - 0.90**	**0.038**
Model 2[b]	0.784	0.58 - 1.03	1.40	0.76 - 2.55	0.75	0.43 - 1.29	**0.62**	**0.40 - 0.96**	**0.029**
Model 3[c]	0.82	0.61 - 1.09	1.49	0.80 - 2.79	0.78	0.45 - 1.35	0.65	0.41 - 1.03	0.056

[a] Adjusted for age, sex, energy intake

[b] Adjusted for age, sex, energy intake, education, physical activity, smoking

[c] Adjusted for age, sex, energy intake, education, physical activity, smoking, waist circumference, hypertension

[d] Further adjusted for metabotypes

MDS: models for 2 points increase in the MDS score; AHEI: models for 10 points increase in the AHEI score

Diabetes: known and unknown DM (vs. NGT and prediabetes)

(Wawro, (…), Linseisen, Nutrients 2020)

Replicating the findings in the Irish NANS study

Applying our metabotyps to participants of the NANS study,
and applying their metabotype to our KORA cohort sample

→ and it worked!
With 4 parameters.

ADVANCED SCIENCE NEWS
www.advancedsciencenews.com

Molecular Nutrition & Food Research
www.mnf-journal.com

Table 2. Variation of the grouping variables across the three metabotype clusters of the KORA F4 study.

	Metabotypes			*p*-value
	Cluster 1 ($N = 590$)	Cluster 2 ($N = 813$)	Cluster 3 ($N = 341$)	
Grouping variables (serum fasting levels in mmol L^{-1})	Median (25%, 75%)	Median (25%, 75%)	Median (25%, 75%)	
TAG	**0.98** (0.73, 1.28)	1.21 (0.89, 1.60)	2.59 (2.07, 3.18)	**<0.0001**
TC	6.18 (5.66, 6.77)	**5.09** (4.60, 5.58)	6.33 (5.79, 6.98)	**<0.0001**
HDL cholesterol	1.81 (1.65, 2.02)	1.29 (1.14, 1.45)	**1.16** (1.03, 1.34)	**<0.0001**
Glucose	**5.11** (4.83, 5.50)	5.39 (5.00, 5.89)	5.89 (5.39, 6.72)	**<0.0001**

Median (25th, 75th percentile); Kruskal–Wallis test; Significant results ($p < 0.05$) are highlighted in bold; Underlined values represent the highest value across the clusters; the bold values represent the lowest; KORA, Cooperative Health Research in the Region of Augsburg; TAG, triacylglycerol; TC, total cholesterol.

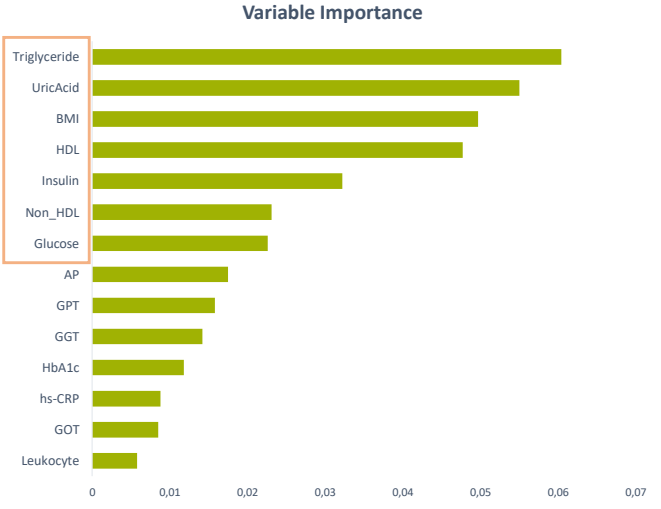
→ Develop kind of a 'best model' in terms of broad useability and specificity.

(Riedl et al., Mol Nutr Food Res 2020)

Further reduction of parameters aiming for a most practical classification

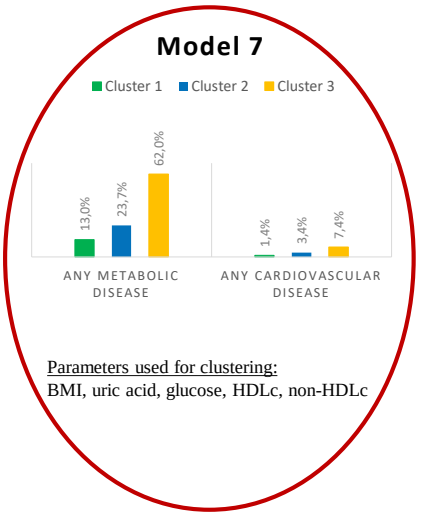
Variable importance of the clustering parameters using
(1) the permutation variable importance method,

confirmed by
(2) cross-validated permutation variable importance and
(3) gradient boosted feature selection method.

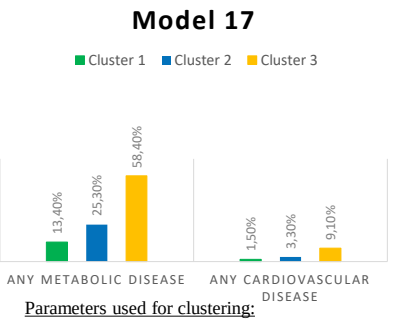


(Dahal, (…), Linseisen, submitted)

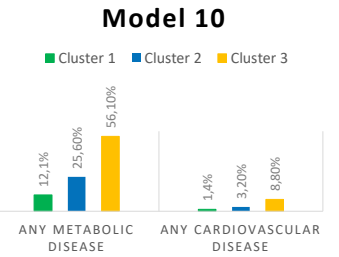
Best metabotyping models with respect to incident disease occurence (KORA FF4)



Parameters used for clustering:
BMI, uric acid, glucose, HDLc, non-HDLc



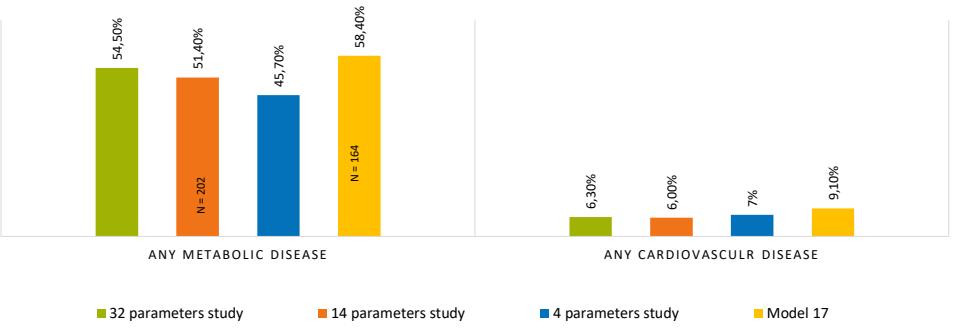
Parameters used for clustering:
glucose, triglycerides, HDLc, non-HDLc



Parameters used for clustering:
BMI, glucose, triglycerides, HDLc, non-HDLc

(Dahal, (…), Linseisen, submitted)

Comparison of incidence of diseases in cluster 3, model 17, to previously identified metabotypes in the KORA study



32 parameters study: **comprehensive parameters**: BMI, Glucose, TC/HDL, HDL, Uric acid, LDL, Leukocytes, Average telomere length in leukocytes, Afamin, Leptin, TPO antibodies, Cystain, Albumin, Creatine, IL-18, Renin, Aldosterone, FT4, Insulin, Alkaline phosphatase; Apolipoprotein A-IV, GGT, GOT, GPT,HbA1c, hs-CRP, IGF-I, IGFBP3, Lp(a), NEFAs, Sex-hormone-binding globulin, Total cholesterol, Triglycerides, Thyroperoxidase, Thyroid-stimulating hormone (Riedl, A. *et al. Mol. Nutr. Food Res.* (2018))

16 parameters study:BMI, CRP, AP, GGT, GOT, GPT, HbA1c, Leukocytes, Triglyceride, HDL, Uric acid, Total cholesterol, LDL, Total cholesterl/HDL, Glucose, Insulin (Riedl, A. et al.. Eur. J. Nutr. (2019))

4 parameters study: Metabotypes based on NANS study:
Total cholesterol, Triglycerides, HDL, Glucose (Riedl, A. et al.. Mol. Nutr. Food Res. 2020)

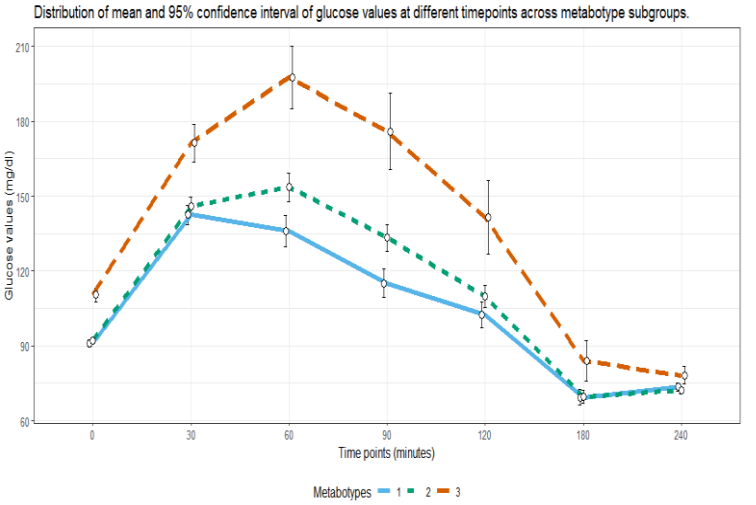
Model 17: 4 parameters:
Triglycerides, HDL, non-HDL, Glucose

Applying metabotypes

- in diet intervention studies
- in diet-EWAS analyses
- in diet-gut microbiota analyses

Assignment of participants in other studies to the metabotype clusters by **using the smallest total Euclidean distance** of the, e.g., four z-standardized values for TAG, TC, HDL cholesterol, and glucose to the respective z-standardized cluster centers (means) of these variables.

Applying metabotypes in intervention studies. here: oral glucose tolerance test (OGTT), with 75 g glucose bolus

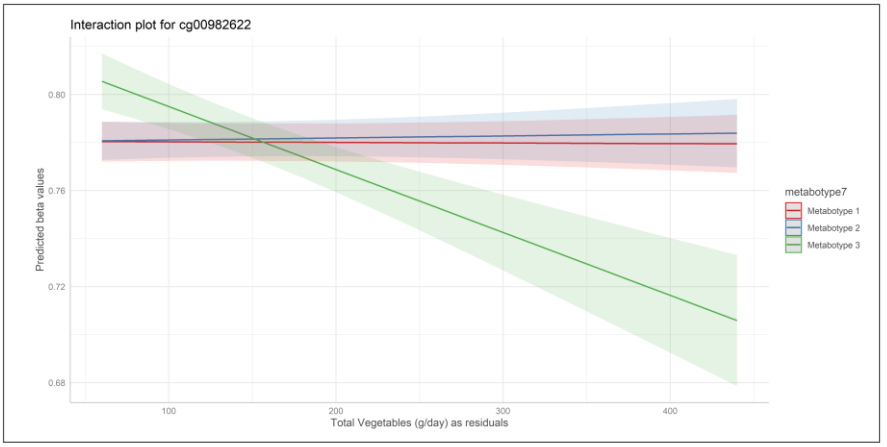


Diet and EWAS:
few associations, e.g. with cabbage intake

CpG	TE.fixed	lower.fixed	upper.fixed	pval.fixed	CHR	UCSC_RefGene_Name	UCSC_RefGene_Group
a	0,00019	0,000133	0,000247	7,51422E-11	20		
b	-0,000239	-0,000313	-0,000164	3,01882E-10	7		
c	-0,000527	-0,000693	-0,000361	4,68482E-10	12		
d	-0,001025	-0,001348	-0,000702	4,77208E-10	1		
e	-0,000235	-0,000314	-0,000156	5,28363E-09	12		
f	-0,000388	-0,000528	-0,000248	5,20904E-08	19		
g	-0,000499	-0,000681	-0,000318	6,59077E-08	6		
h	0,000543	0,000345	0,000741	7,54485E-08	6		
i	0,000408	0,000259	0,000557	8,54606E-08	17		

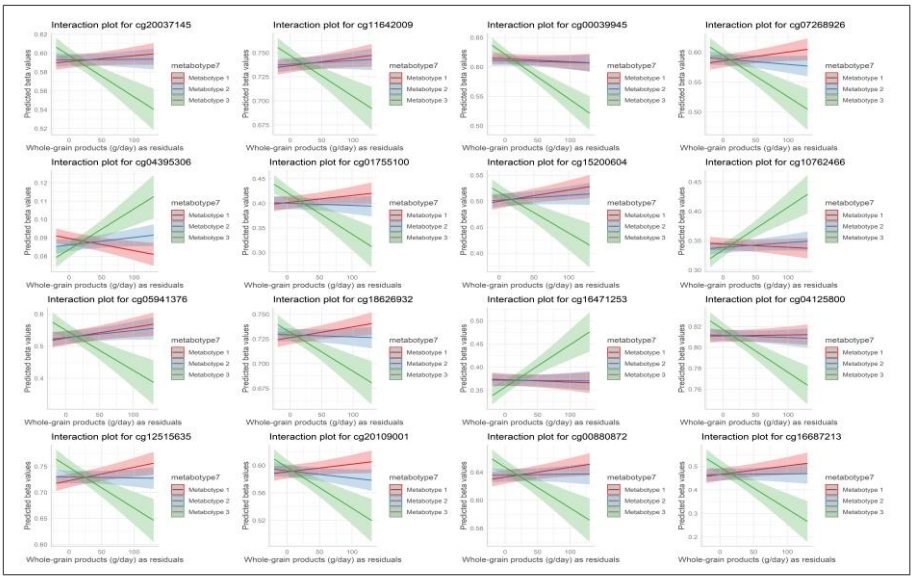
Linear regression model (fixed-effect model, 400k CpG's)
Methylation(beta) ~ exposure + sex + age + age² + bmi + bmi² + smoking + total energy intake + alcohol(g/day)
+ monocytes + basophile granulocytes + eosinophile granulocytes + lymphocytes + plate; Bonferroni adjusted threshold

Diet and EWAS, by metabotype:
significant interaction effects, e.g. with vegetable intake



Linear regression model (fixed effect model; EPIC chip, Illumina: 850k)
Methylation(beta) ~ exposure + sex + age + age² + bmi + bmi² + smoking + total energy intake + alcohol(g/day) + monocytes +
basophile granulocytes + eosinophile granulocytes + lymphocytes + plate **+ metabotype + metabotype*exposure**; FDR < 0.1

Diet and EWAS, by metabotype: e.g. whole-grain products



Diet and EWAS, by metabotype

Conclusion:

If confirmed in other studies,
dietary modification of DNA methylation depends on metabolic derangement.

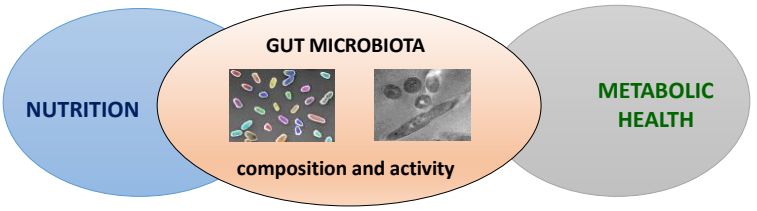
Metabotyping can help identify such associations.

Diet and gut microbiota association

→ Characterize subjects according to their microbiota composition (16s rRNA data)
Latenent Dirichlet Allocation (20 subgroups)

→ Identify subgroups that are associated to diet

→ Identify subgroups that are associated with metabolic health



Latent Dirichlet Allocation (LDA)

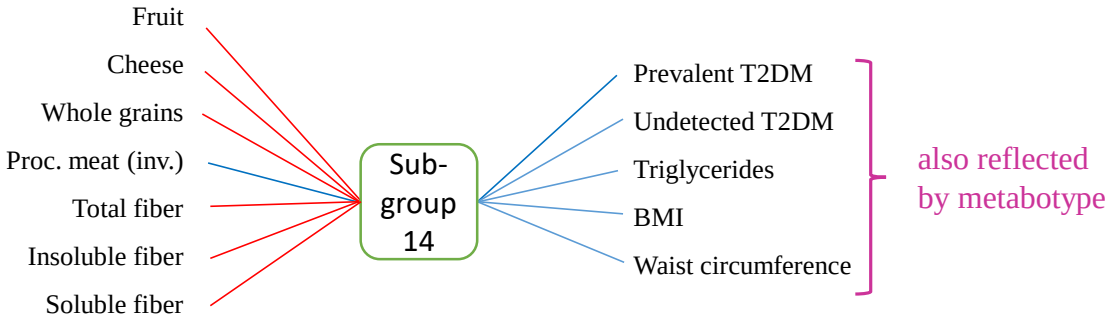
- Machine learning method developed in population genetics
- Determines latent (hidden) groups that explain patterns in the data
- Similar to soft clustering in that it allows partial membership

- Based on perplexity and log likelihood score, 20 subgroups were selected
- Each subgroup represents a sub-community of bacteria that tend to appear together
- Each person has a probability for each topic. Each person‘s microbiome is made up of different percentages of the 20 topics

→ Dirichlet regression analysis (multi-variable adjusted)

Subgroup 14: associations with both diet and diseases

No interaction effects of the metabotype!!



→ **Subjects with a low proportion of subgroup 14 may benefit from dietary modifications in terms of metabolic risk reduction**

(Breuninger, (…), Linseisen, Microbiome 2021)

Diet and gut microbiota: no role of metabotypes

Conclusion:

We found no interaction beteen diet and gut microbiota subgroups by metabotypes,

indicating that metabolic derangements do not modify diet-microbiota associations.

Gut microbiota composition has to be considered in addition to metabotypes.

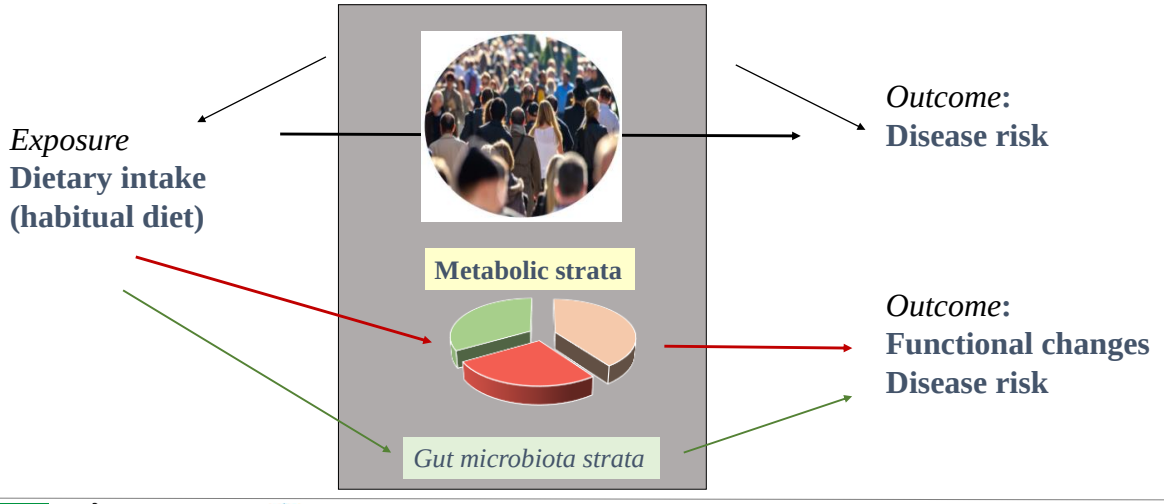
Summary: Metabolic Phenotyping in Nutrition Research

Here: Definition of „Metabotypes“ by few metabolic parameters as a simple and practical way of population stratification

Consider metabolic phenotypes
which integrate specific genetic and metabolic characteristics

- **for future scientific research**
- **for developing more effective intervention strategies (with higher compliance?)**
- **for developoments towards personalized nutrition.**

Added value of metabolic phenotyping in nutrition research



Thank you for your attention!!