The Challenges of Meta-Analysing Metabolomics Data; Experiences from the Consortium Of METabolomics Studies (COMETS)

Rachel Kelly PhD
Channing Division of Network Medicine
BWH, HMS

Thursday 23rd April 2020
CEDAR MRC Epidemiology Seminar

THE·UNIVERSITY·OF
WARWICK

The University of
Nottingham

ICR The Institute of Cancer Research

Imperial College London

HARVARD
T.H. CHAN
SCHOOL OF PUBLIC HEALTH

VE RI TAS
HARVARD MEDICAL SCHOOL

CHANNING DIVISION OF NETWORK MEDICINE

BWH BRIGHAM AND WOMEN'S HOSPITAL

SAMUEL ADAMS
BOSTON LAGER

Edwin Silverman
Division Chief

BRIGHAM HEALTH
BRIGHAM AND WOMEN'S HOSPITAL

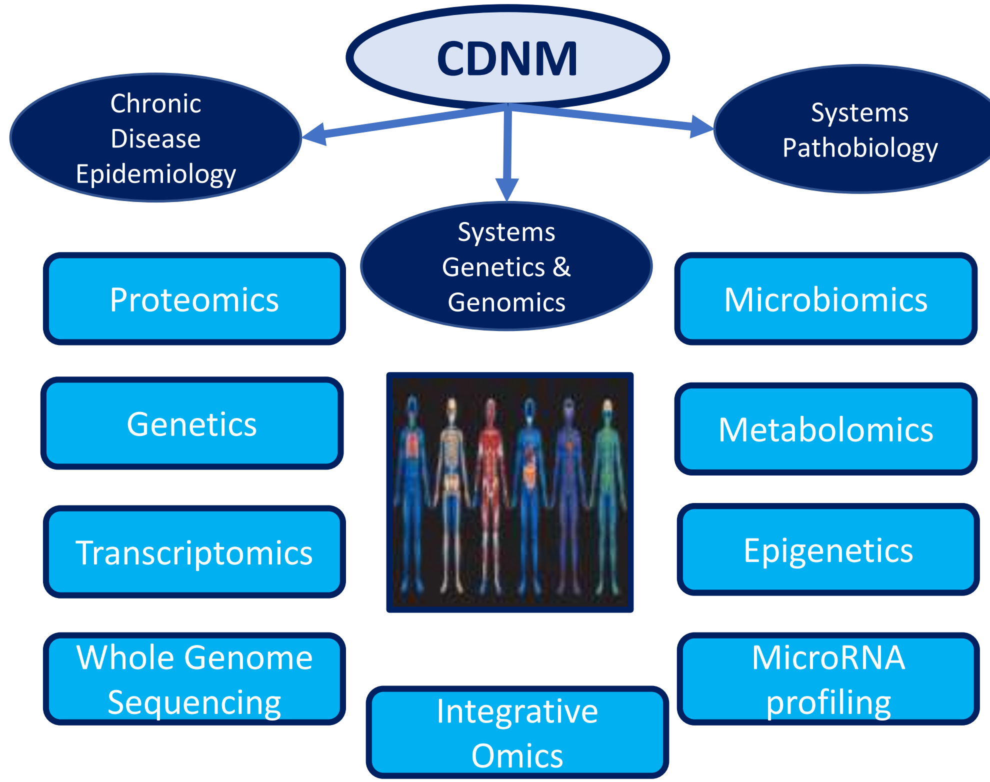
HARVARD MEDICAL SCHOOL

Our mission is to

(i) use an **integrated, network-based, systems biology-driven approach** to define the etiology of complex diseases;

(ii) reclassify complex diseases based on **systems pathobiological mechanisms**;

(iii) to develop new treatments and preventive strategies based on these new disease classifications using **systems pharmacology approaches**”



https://www.brighamandwomens.org/research/departments/channing-division-of-network-medicine

CHANNING DIVISION OF NETWORK MEDICINE

Edwin Silverman
Division Chief

BRIGHAM HEALTH
BRIGHAM AND WOMEN'S HOSPITAL

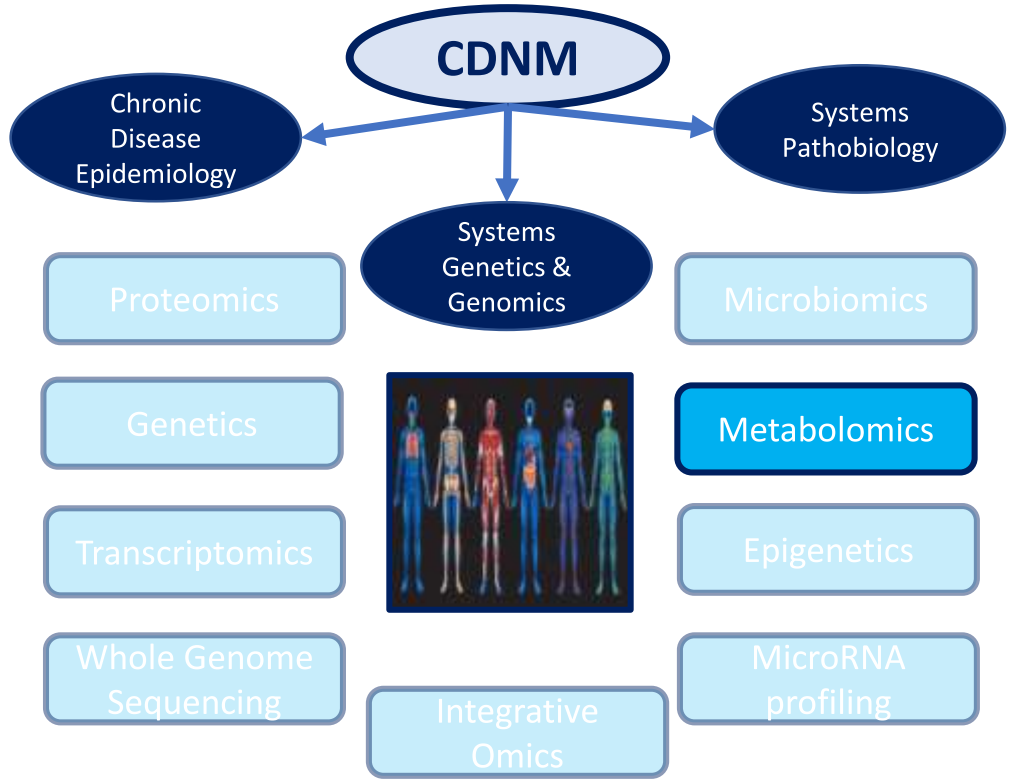
HARVARD MEDICAL SCHOOL

Our mission is to

(i) use an **integrated, network-based, systems biology-driven approach** to define the etiology of complex diseases;

(ii) reclassify complex diseases based on **systems pathobiological mechanisms**;

(iii) to develop new treatments and preventive strategies based on these new disease classifications using **systems pharmacology approaches**"

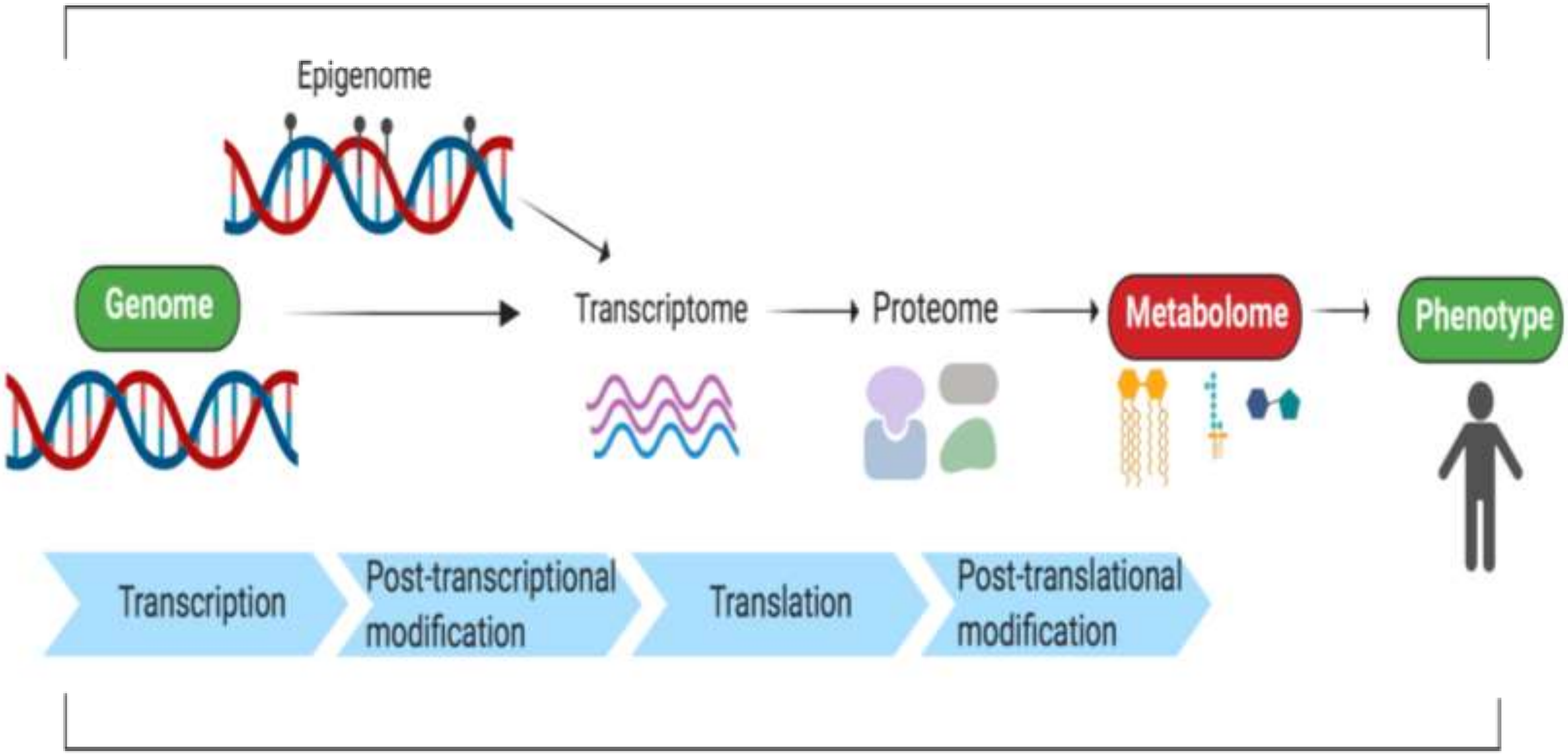


https://www.brighamandwomens.org/research/departments/channing-division-of-network-medicine

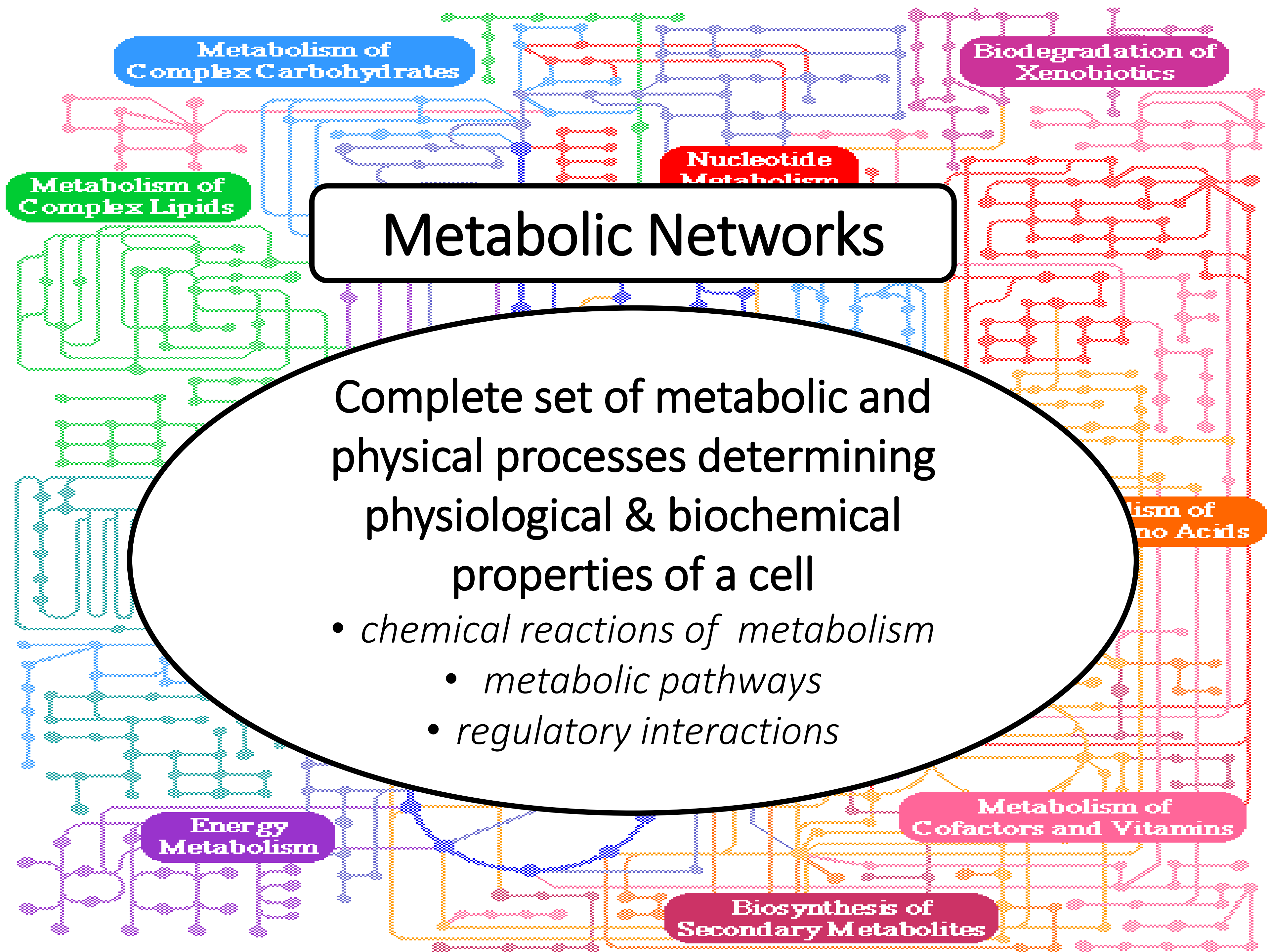
Metabolomics

The Central Biological Dogma

Environment; Diet; Air pollution; Geography; Demographics



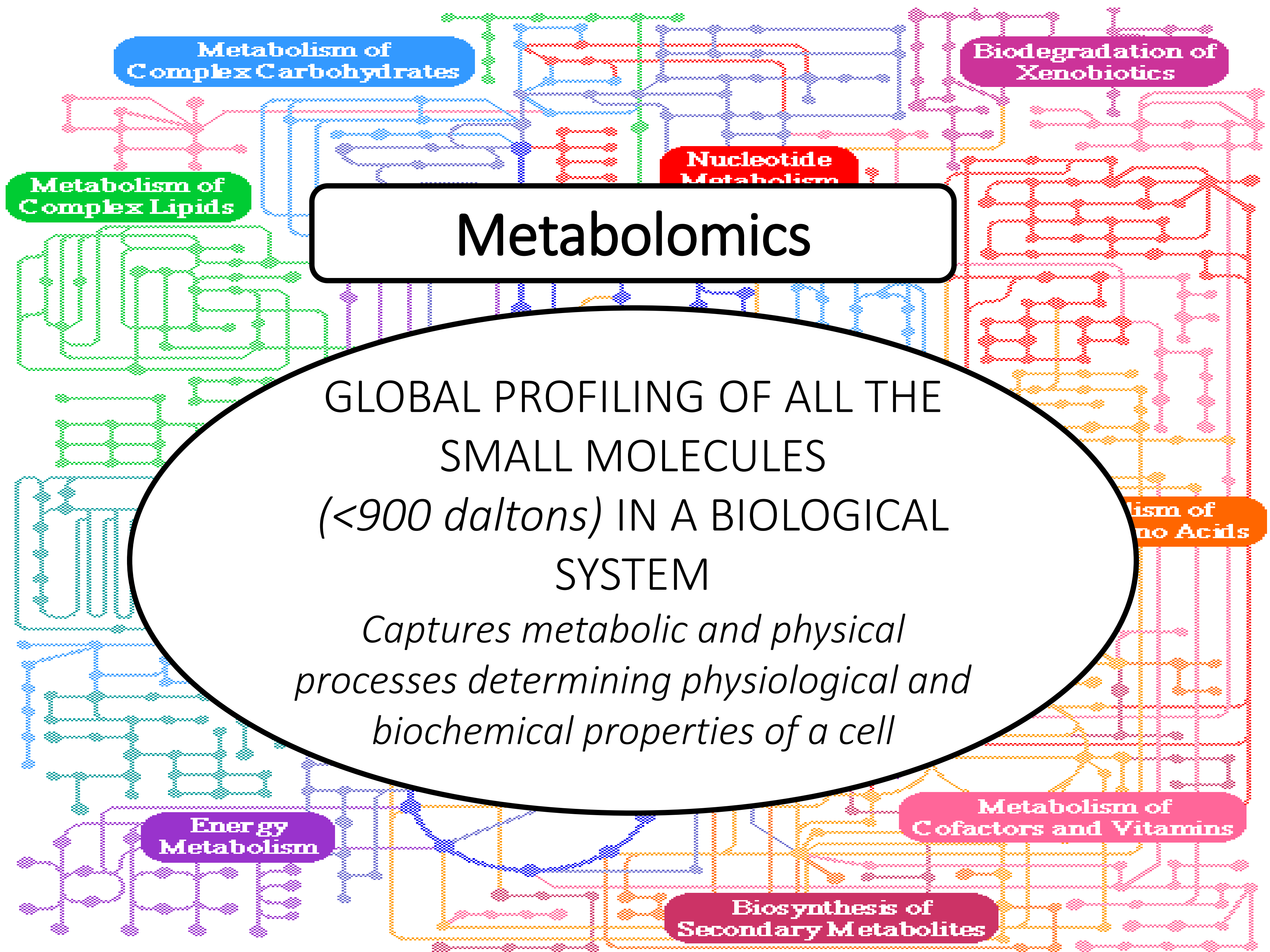
Microbiome



Metabolic Networks

Complete set of metabolic and physical processes determining physiological & biochemical properties of a cell

- *chemical reactions of metabolism*
- *metabolic pathways*
- *regulatory interactions*

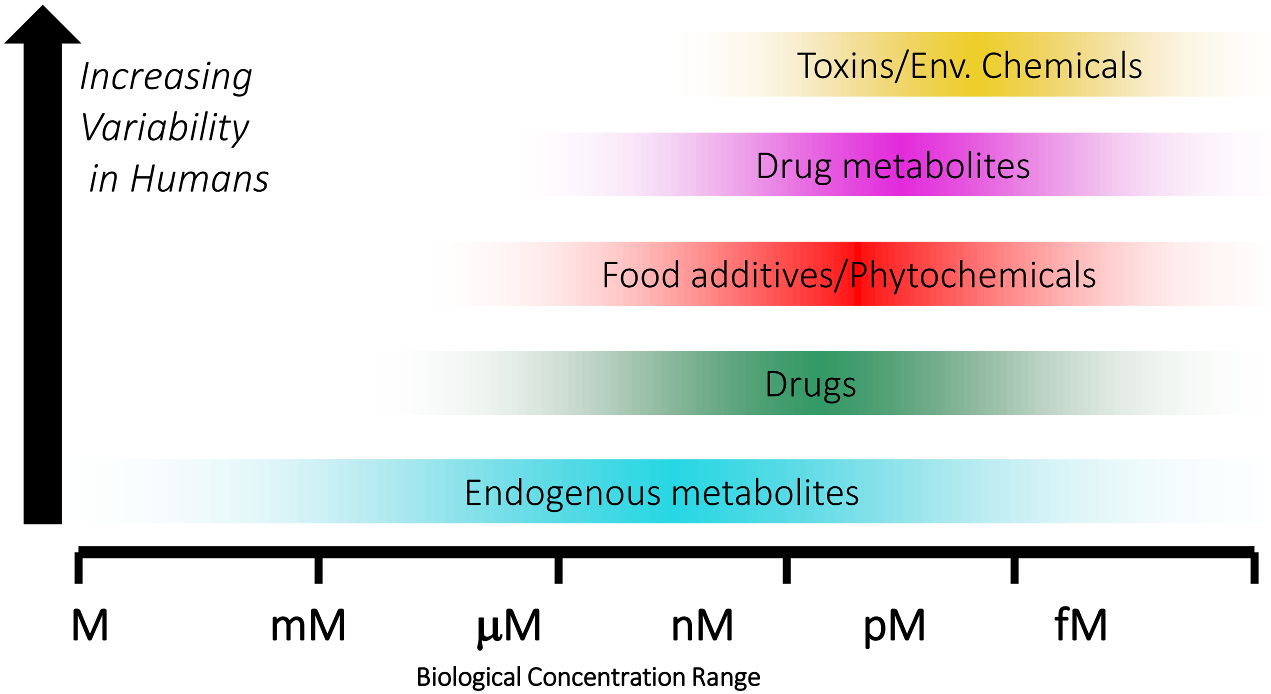


Metabolomics

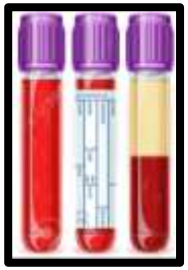
GLOBAL PROFILING OF ALL THE SMALL MOLECULES *(<900 daltons)* IN A BIOLOGICAL SYSTEM

Captures metabolic and physical processes determining physiological and biochemical properties of a cell

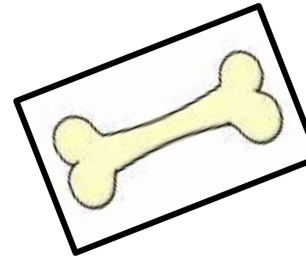
The Human Metabolomes



Courtesy of bioinformatics.ca

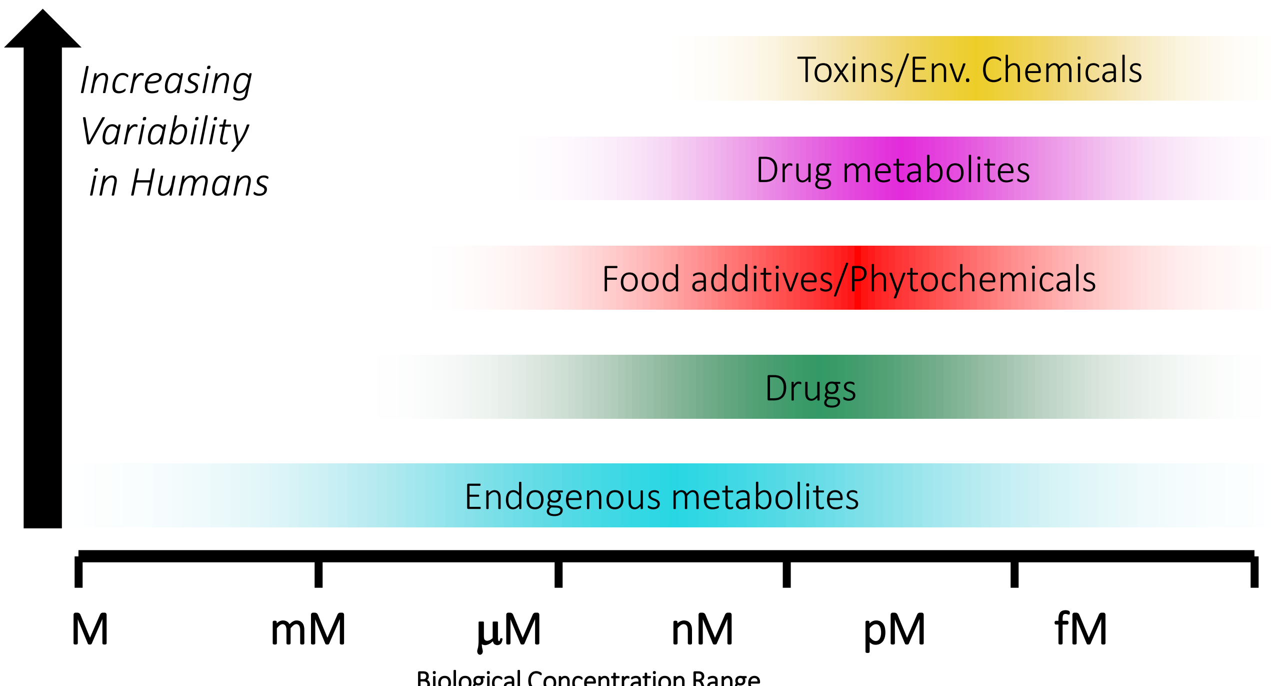






The Human Metabolomes



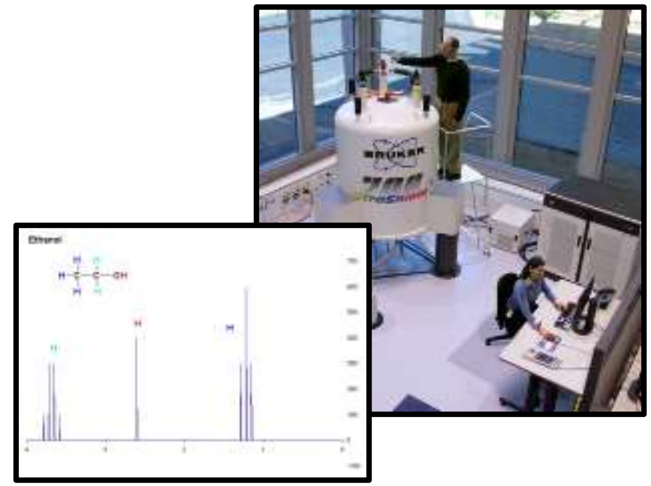


Courtesy of bioinformatics.ca

Metabolomic Profiling Platforms

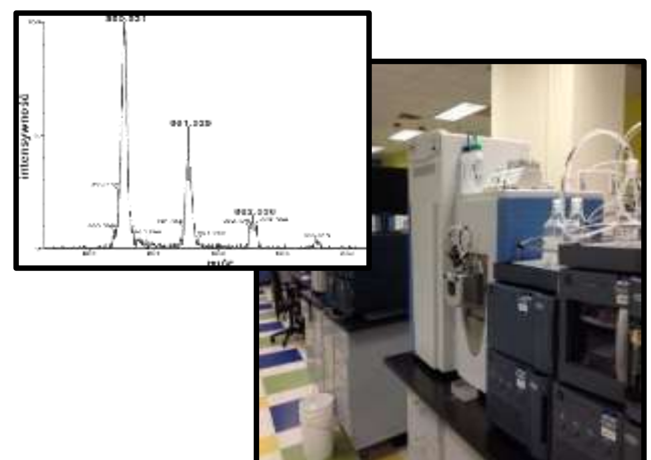
Nuclear Magnetic Resonance

- Most commonly H^1 or C^{13}
- Spectra based on the chemical shift induced by strong magnetic field
- Quantitative
- Highly reproducible
- Provides structural information



Mass spectrometry

- Ionization followed by assessment of mass-to-charge ratio
- More Sensitive
- Can measure more metabolites



Metabolomic Profiling Data: *Targeted* versus *Untargeted*

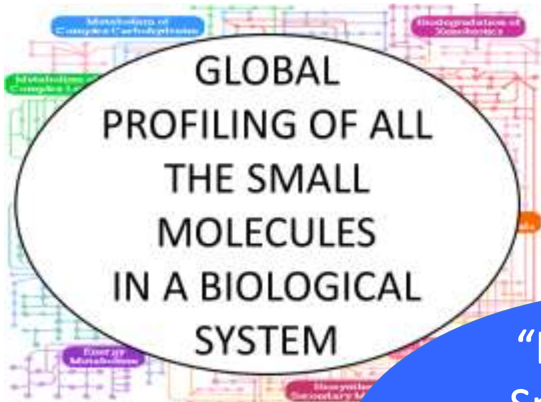
Untargeted

- Comprehensive 'Global" analysis of all measurable analytes in a biological sample
- Includes metabolites of unknown identity
- *Relative* abundance

Targeted

- Measurement of defined groups of chemically characterized and biochemically annotated metabolites
- Fewer Metabolites
- *Absolute* quantification

What can the Metabolome tell us?



GLOBAL PROFILING OF ALL THE SMALL MOLECULES IN A BIOLOGICAL SYSTEM

"Instantaneous Snapshot of the Physiological Status of a Biological System"

FUTURE

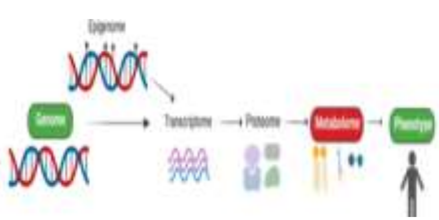
Pre/Early Stage Pathogenesis

PRESENT

Environment | Phenotype

PAST

Environment | Preceeding "omes"



What can the Metabolome do for us?

Metabolomic Biomarkers

HEALTHY

IMPROVED UNDERSTANDING OF BIOLOGICAL MECHANISMS

DISEASED

SUSCEPTIBILITY

PREDICTION

DIAGNOSIS

ENOTYPING

PROGNOSIS

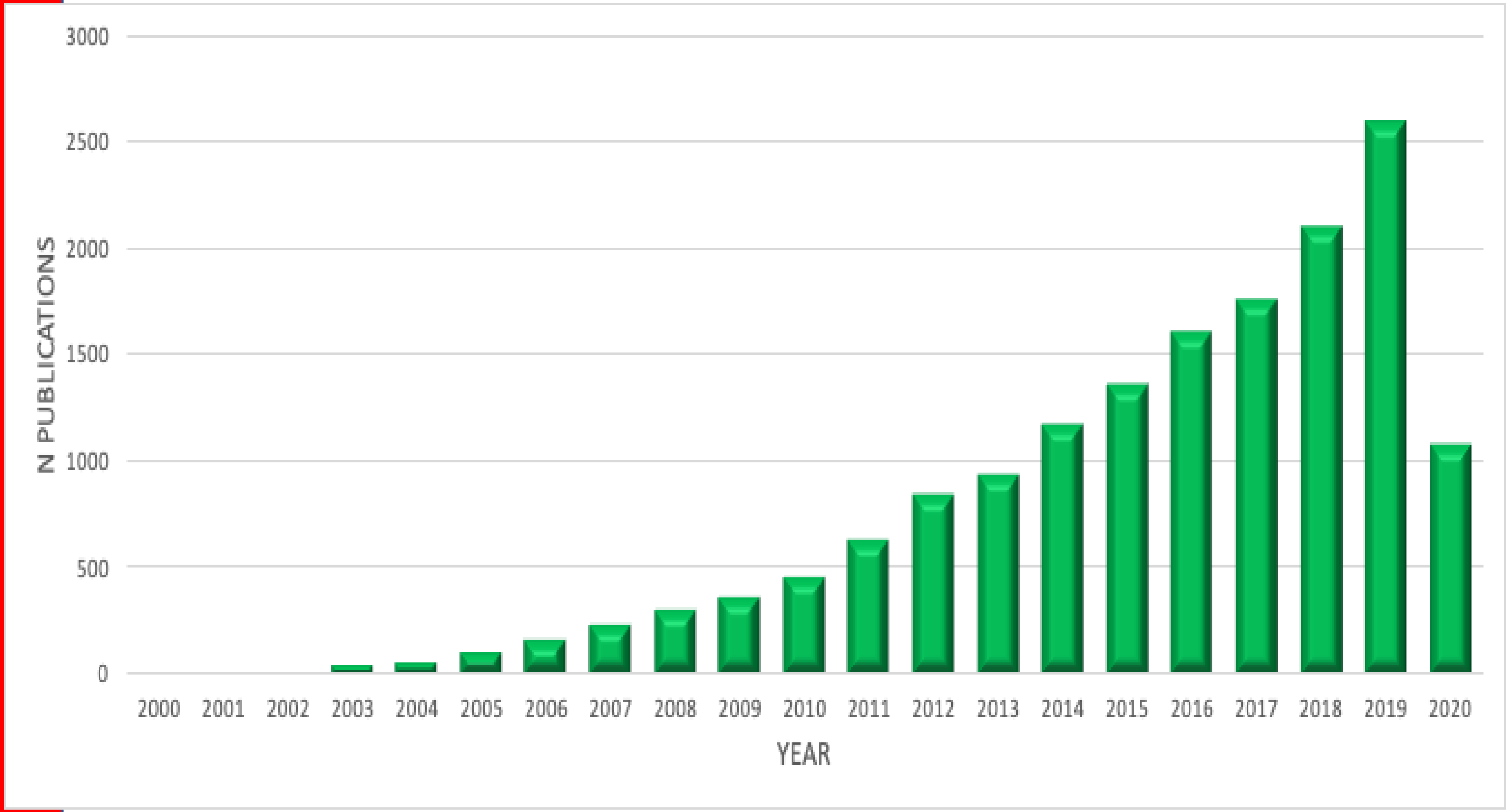
THERAPEUTIC RESPONSE

RECURRENCE

EXPOSURE

Metabolomics is a Rapidly Growing Field

Number of Metabolomics Related Publications Per Year Since 2000



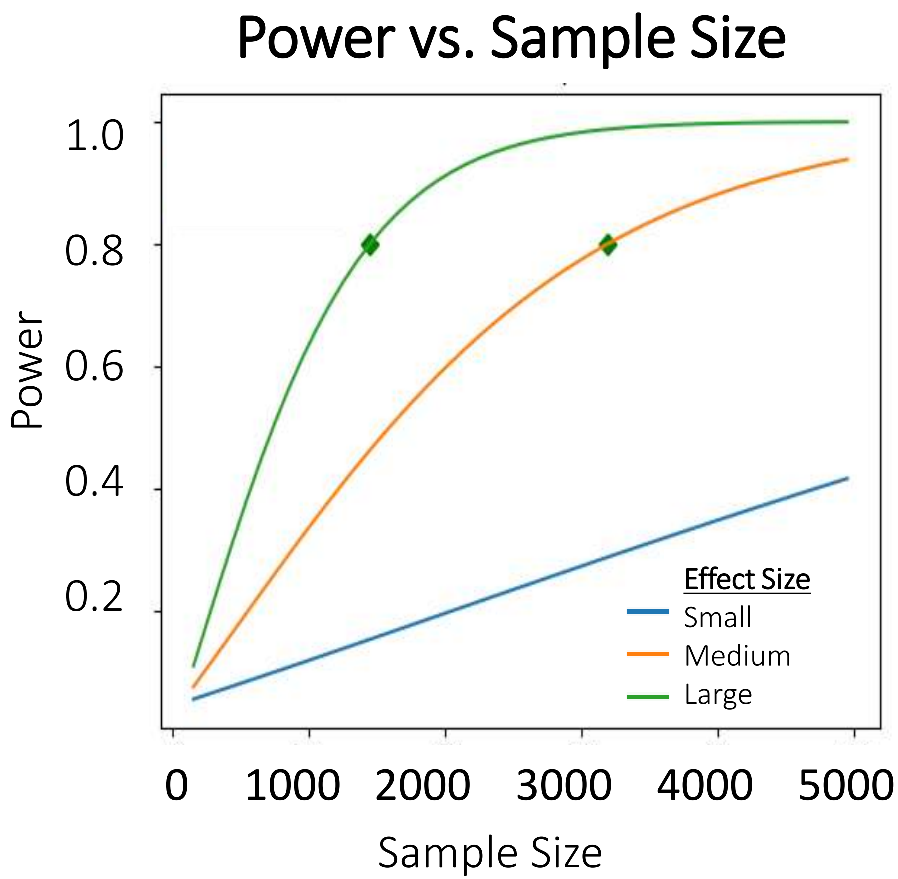
PUBMED SEARCH TERMS: (metabolom*[title] OR (Metabolite profil*[title]) OR (metabolite signature[title]))
SEARCH DATE: 22/4/2020

Metabolomics Can be Noisy

Biological Heterogeneity

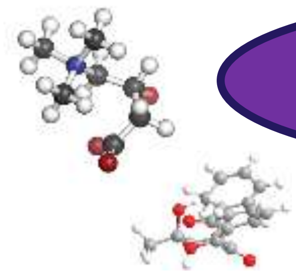
Technical Heterogeneity

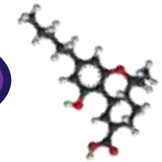
Obtaining the Strongest Metabolomic Findings



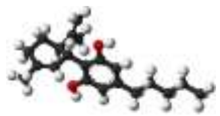
- Many existing metabolomic cohorts have limited sample sizes
- Meta-analyses provide:
  - More power to detect an effect
  - More precise and accurate effect estimates
  - More generalizable findings

Meta-Analyzing Metabolomic data Can be Complex





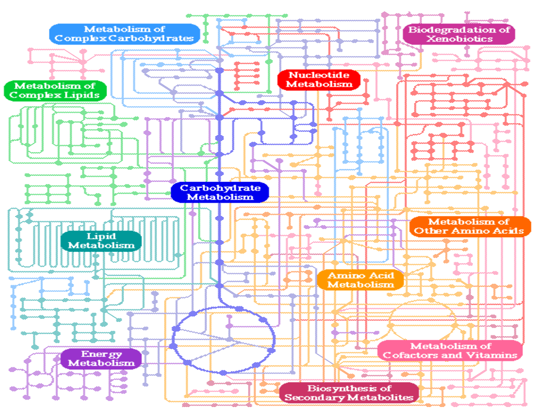
Metabolomic Data



Incomplete Coverage of the Metabolome

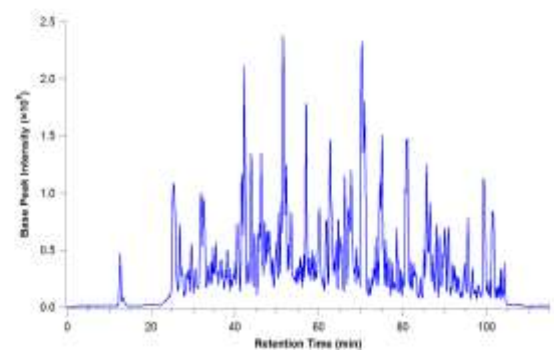
Inconsistent Nomenclature

Relative Nature of Measurement



(2S)-2-amino-3-(1H-indol-3-yl)Propanoic acid
(2S)-2-amino-3-(1H-indol-3-yl)Propanoate
(S)-a-amino-1H-Indole-3-propanoate
(S)-alpha-amino-1H-Indole-3-propanoic acid
(S)-alpha-amino-beta-(3-Indolyl)-propionic acid
(S)-Tryptophan
1H-Indole-3-alanine
L-(-)-Tryptophan
L-beta-3-Indolylalanine
L-β-3-indolylalanine

Plus 71 other synonyms…



The <u>CO</u>nsortium of <u>MET</u>abolomic <u>S</u>tudies (COMETS)

COnsortium of METabolomics Studies

- *Extramural-intramural partnership promoting collaboration among prospective <u>metabolomic epidemiology</u> studies*

- <u>Mission and Objectives</u>
- Provide framework to foster international collaborations among studies sharing common objectives;
- Provide forum for the discussion, development & pursuit of new research
- Advance knowledge of the metabolome

- <u>Membership Eligibility</u>
- Prospective cohort, 100+ participants with blood metabolomics
- Phenotype follow-up
- MS or NMR

Krista Zanetti:
COMETS Program Officer

Jessica Lasky-Su
COMETS Chair

https://epi.grants.cancer.gov/comets/

Yu *et al*. ***AJE***, (2019) 188; 6

COMETS Analytics

http://www.comets-analytics.org

"a freely-accessible cloud-based, self-serviced analytic platform developed for consortium-based metabolomics analyses"

Built-in Metabolite Harmonization: continually updating master list of metabolites with multiple levels of information

Yu *et al*. ***AJE***, (2019) 188; 6

Methods paper in Development (Temprosa et al.)

Steven C. Moore

Marinella Temprosa

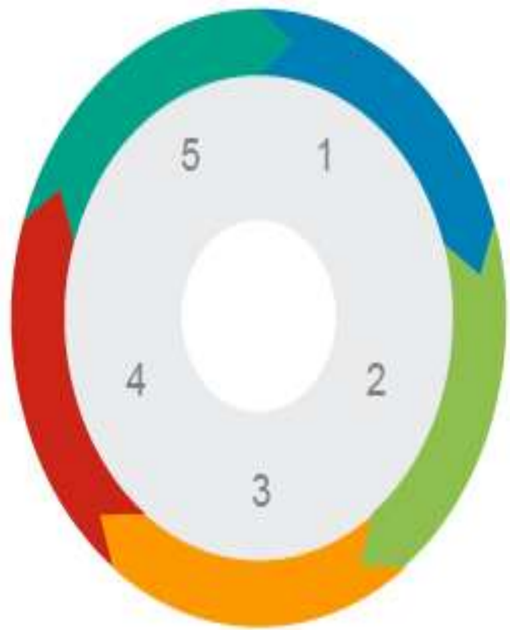
Ewy Mathé

Analysis Pipeline

COMETS Analytics
http://www.comets-analytics.org

Meta-Analyses
Centralized analyses of aggregated cohort data

Metabolite Harmonization
Create mapping of metabolites across cohorts and platforms



Cohort-Specific Analyses
Conduct patient level analyses for data exploration, and approved, manuscript proposals

Data Preparation
Facilitate creation of common input file

Data Integrity
Identify possible data problems prior to analyses

Courtesy of Ella Temprosa

Metabolite Harmonization: All available metabolite information

COMETS Analytics maintains a dynamic continually updating master list of metabolites with multiple levels of information on every metabolite submitted for analysis, Including:

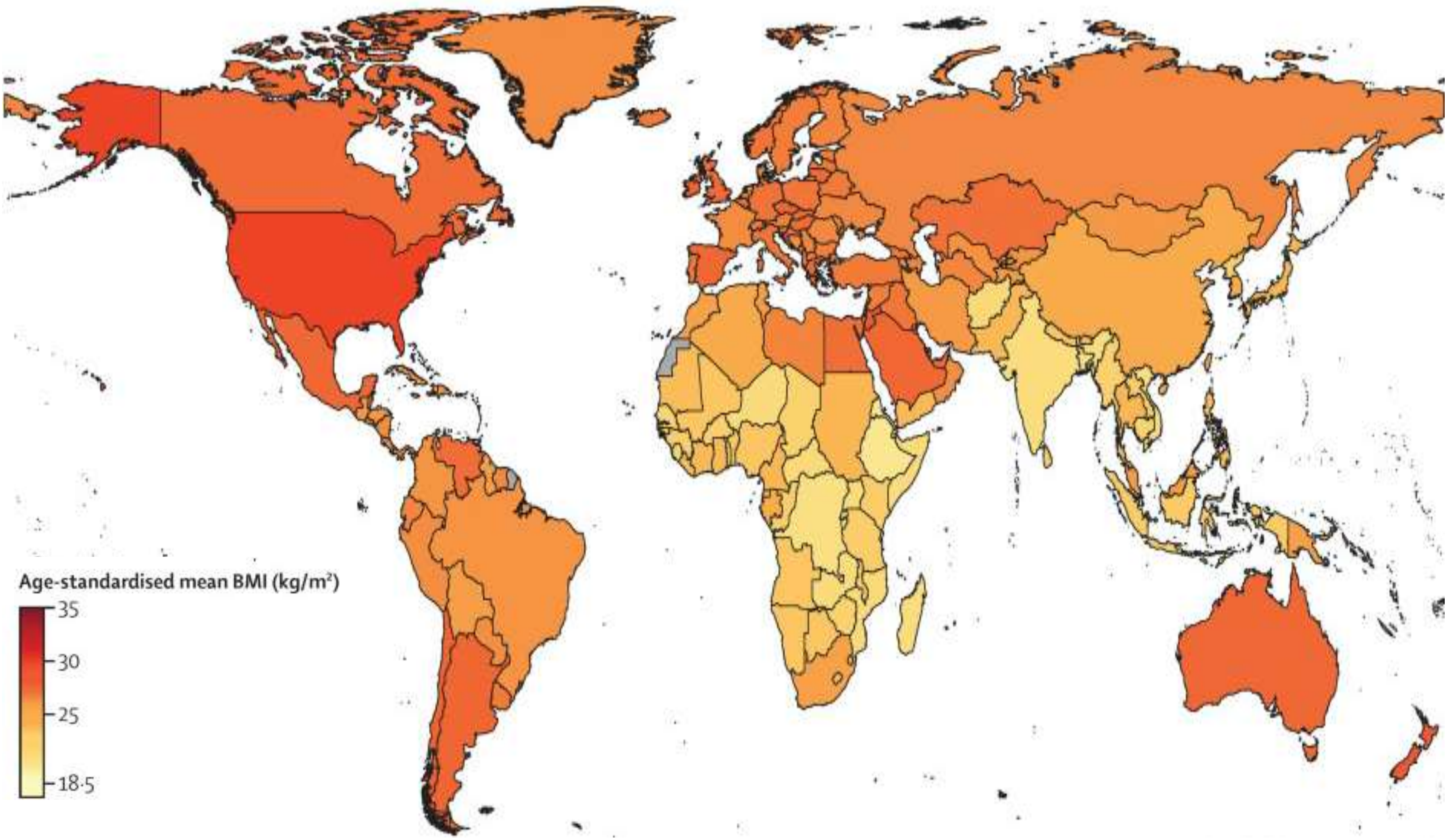
- *Chemical Id (HMDB, KEGG, CHEBI, etc.)*
- *Platform assigned super pathways*
- *m/z and retention time*
- *Metabolite classification via metabolomics work bench and metabolon*

Automatic and manual curation

metabid	metabolite_name	SUPER_PATHWAY	SUB_PATHWAY	COMP_ID	PLATFORM	PUBCHEM	HMDB_ID	KEGG ID	chEBI ID	m/z	rt
ACETOACETATE	acetoacetate	Lipid	Ketone Bodies	33963	GC/MS	96	HMDB00060	NA	NA	NA	NA
ACETYLCARNITINE	acetylcarnitine	Lipid	Fatty Acid Metabolism(Acyl Carnitine)	32198	LC/MS Pos	1	HMDB00201	NA	NA	NA	NA
ACISOGA	acisoga	Amino Acid	Polyamine Metabolism	43258	LC/MS Pos	129397	NA	NA	NA	NA	NA
ADENINE	adenine	Nucleotide	Purine Metabolism, Adenine containing	554	LC/MS Pos	190	HMDB00034	NA	NA	NA	NA

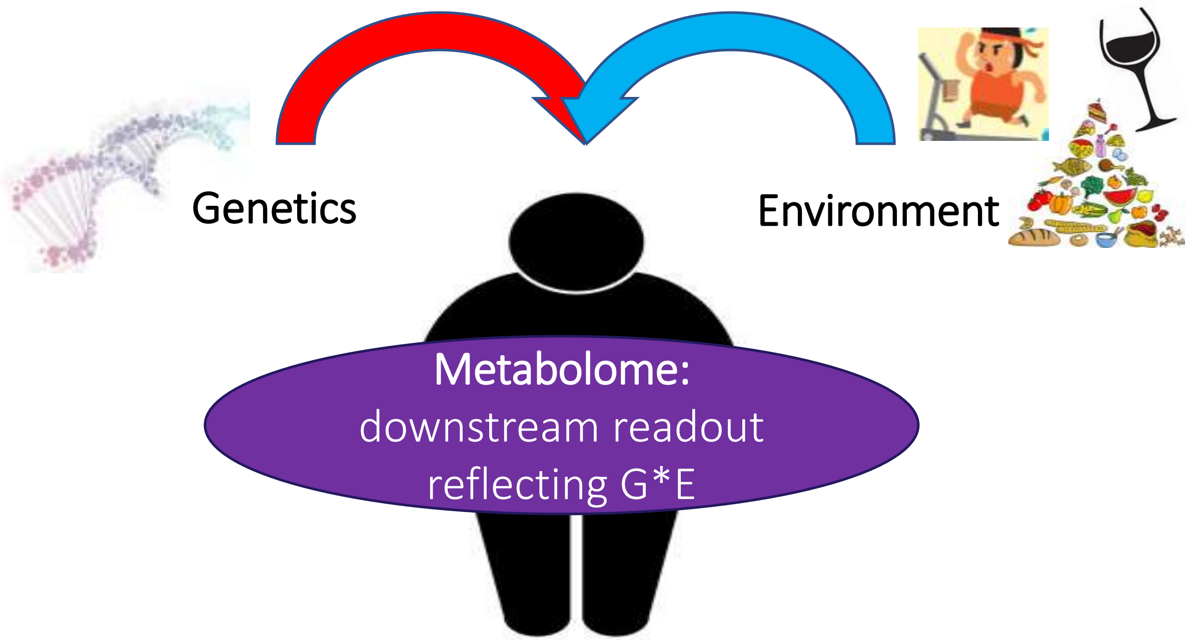
The Metabolome of BMI:
A COMETS Meta-analysis

Age Standardized Mean BMI in Men by Country in 2014



NCD Risk Factor Collaboration, The Lancet (2016)

BMI is perfectly suited to Metabolomic Exploration



COMETS Proposal: The BMI of Obesity

- *Aim 1:* To evaluate relationships between blood metabolite concentrations and BMI across multiple cohorts utilizing a meta-analysis approach within the ***COMETS*** consortium

- *Aim 2:* To evaluate heterogeneity of associations by participant characteristics and by study characteristics

Methods

COMETS Analytics
http://www.comets-analytics.org

Cohort Specific Analyses
Metabolite~BMI

↓

Meta-Analysis

↓

False Discovery Correction

↓

Evaluation of Heterogeneity

↓

Pathway analysis/Metabolite visualization

Investigated Models

COMETS Analytics

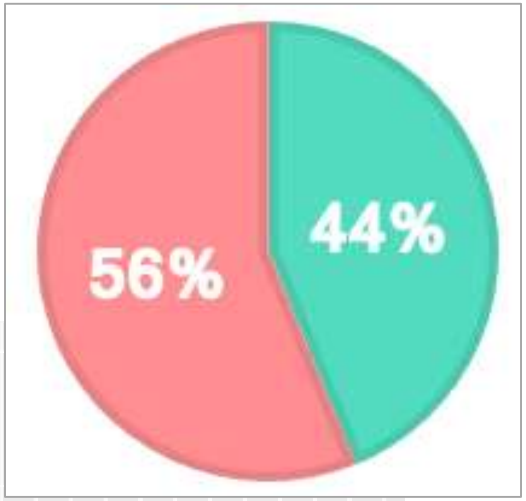
http://www.comets-analytics.org

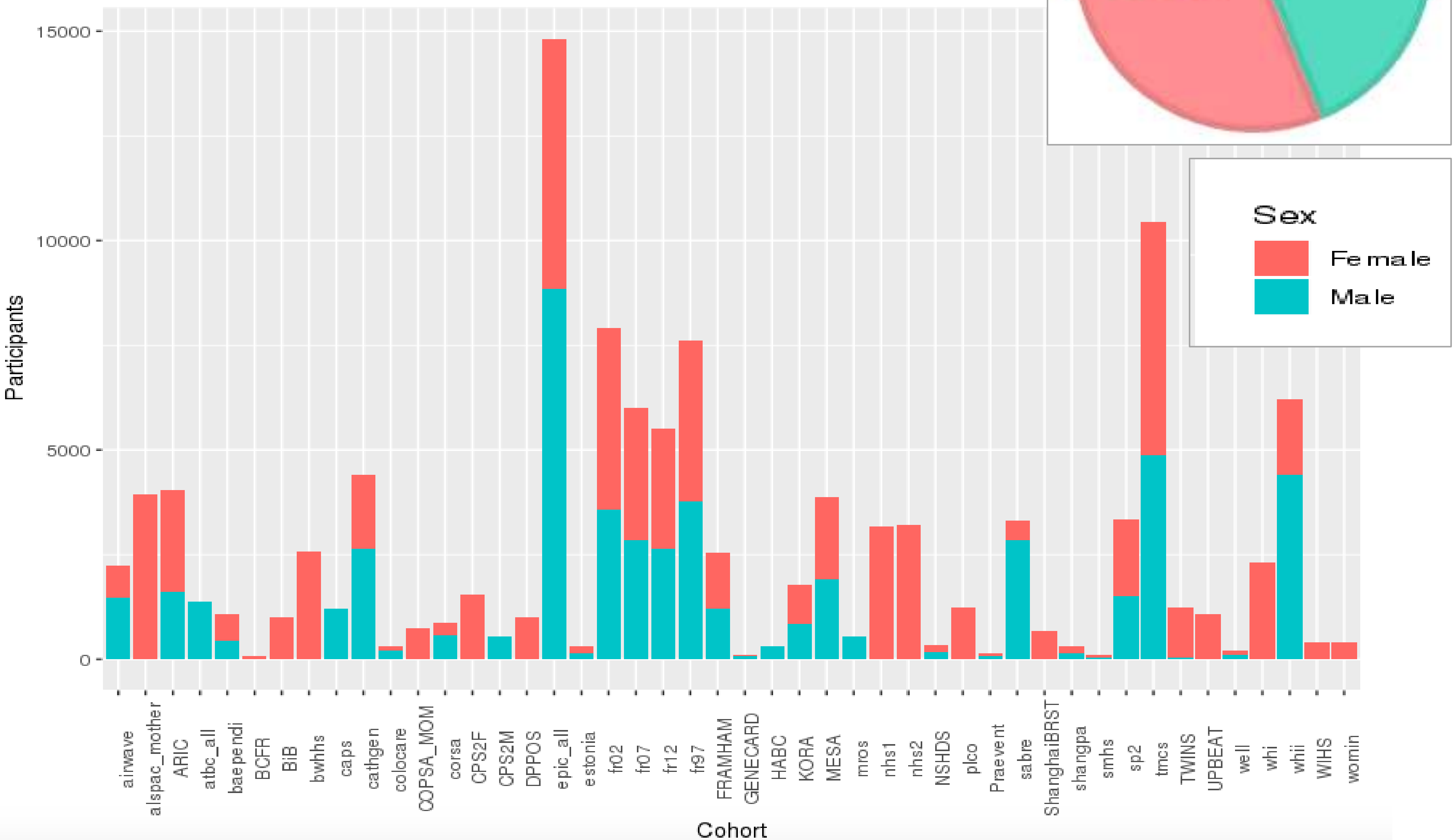
	MODEL	OUTCOMES	EXPOSURE	ADJUSTMENT	STRATIFICATION	WHERE
2	BMI.1.0 Basic adjustment	All metabolites	bmi	age female race_grp nested_case		bmi>=1
3	BMI.1.1 Multivariable adjusted	All metabolites	bmi	age female race_grp educ_grp smk_grp alc_grp fasted nested_case		bmi>=1
4	BMI.1.2 Multivariable and diabetes	All metabolites	bmi	age female race_grp educ_grp smk_grp alc_grp fasted prev_diabetes nested_case		bmi>=1
5	BMI.2.0 Gender stratified	All metabolites	bmi	age race_grp educ_grp smk_grp alc_grp fasted nested_case	female	bmi>=1
6	BMI.2.1 Gender stratified diabetes	All metabolites	bmi	age race_grp educ_grp smk_grp alc_grp fasted prev_diabetes nested_case	female	bmi>=1
7	BMI.3.0 Race stratified	All metabolites	bmi	age female educ_grp smk_grp alc_grp fasted nested_case	race_grp	bmi>=1
8	BMI.3.1 Race stratified diabetes	All metabolites	bmi	age female educ_grp smk_grp alc_grp fasted prev_diabetes nested_case	race_grp	bmi>=1
9	BMI.4.0 Fasted stratified	All metabolites	bmi	age female race_grp educ_grp smk_grp alc_grp nested_case	fasted	bmi>=1
10	BMI.4.1 Fasted stratified diabetes	All metabolites	bmi	age female race_grp educ_grp smk_grp alc_grp prev_diabetes nested_case	fasted	bmi>=1
11	BMI.5.0 Diabetes stratified	All metabolites	bmi	age female race_grp educ_grp smk_grp alc_grp fasted nested_case	prev_diabetes	bmi>=1
12	BMI.6.0 nested_case stratified	All metabolites	bmi	age female race_grp educ_grp smk_grp alc_grp fasted	nested_case	bmi>=1
13	BMI.7.0 Age stratified	All metabolites	bmi	female race_grp educ_grp smk_grp alc_grp fasted nested_case	age_grp	bmi>=1
14	BMI.7.1 Age stratified diabetes	All metabolites	bmi	female race_grp educ_grp smk_grp alc_grp fasted prev_diabetes nested_case	age_grp	bmi>=1
15	BMI.7.2 Age stratified age-adj	All metabolites	bmi	age female race_grp educ_grp smk_grp alc_grp fasted nested_case	age_grp	bmi>=1
16	BMI.7.3 Age stratfied age-adj diabetes	All metabolites	bmi	age female race_grp educ_grp smk_grp alc_grp fasted prev_diabetes nested_case	age_grp	bmi>=1

Metabolites | SubjectMetabolites | SubjectData | VarMap | Models

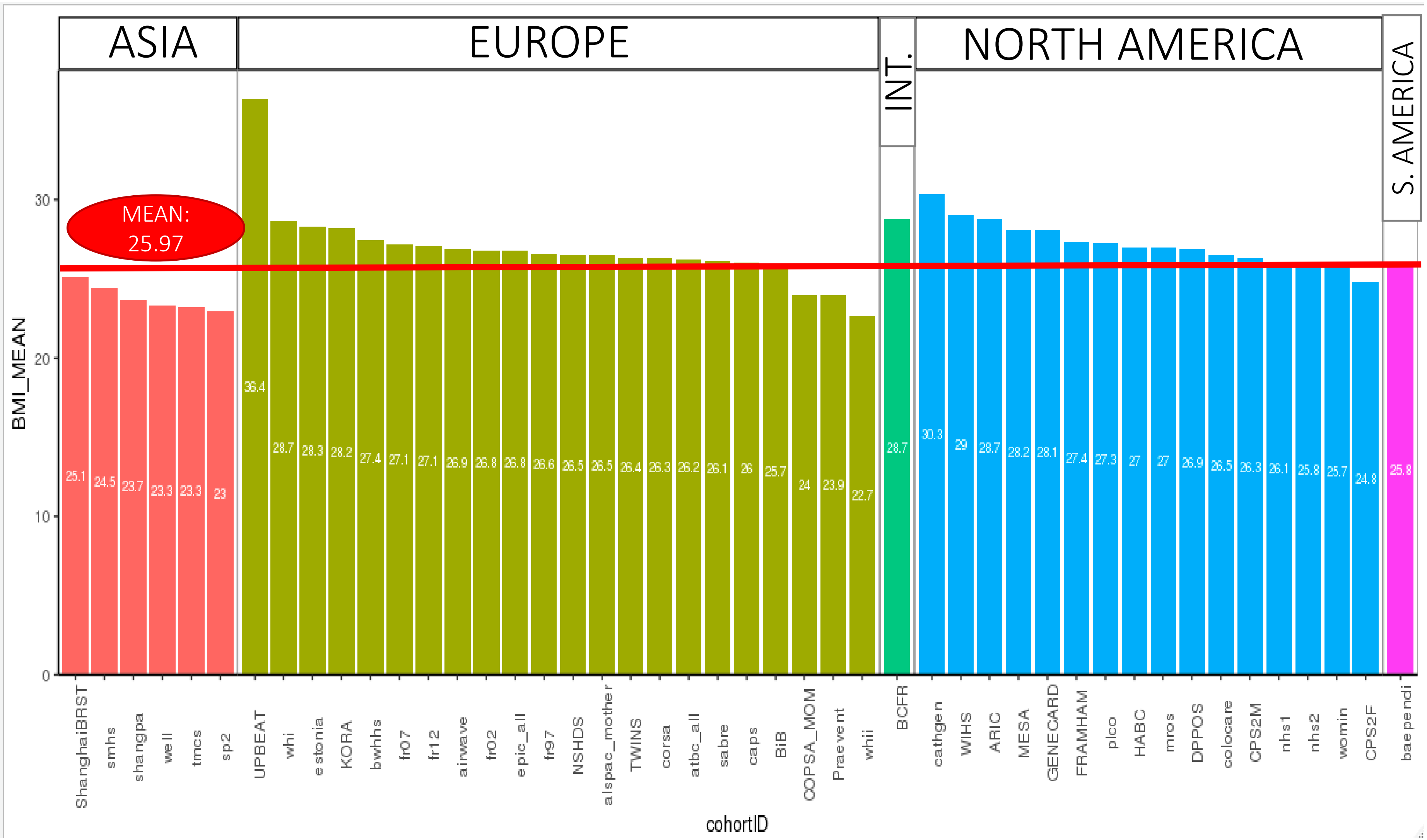
Pre-specified Models in the User Input File

126,423 Participants from 46 Cohorts

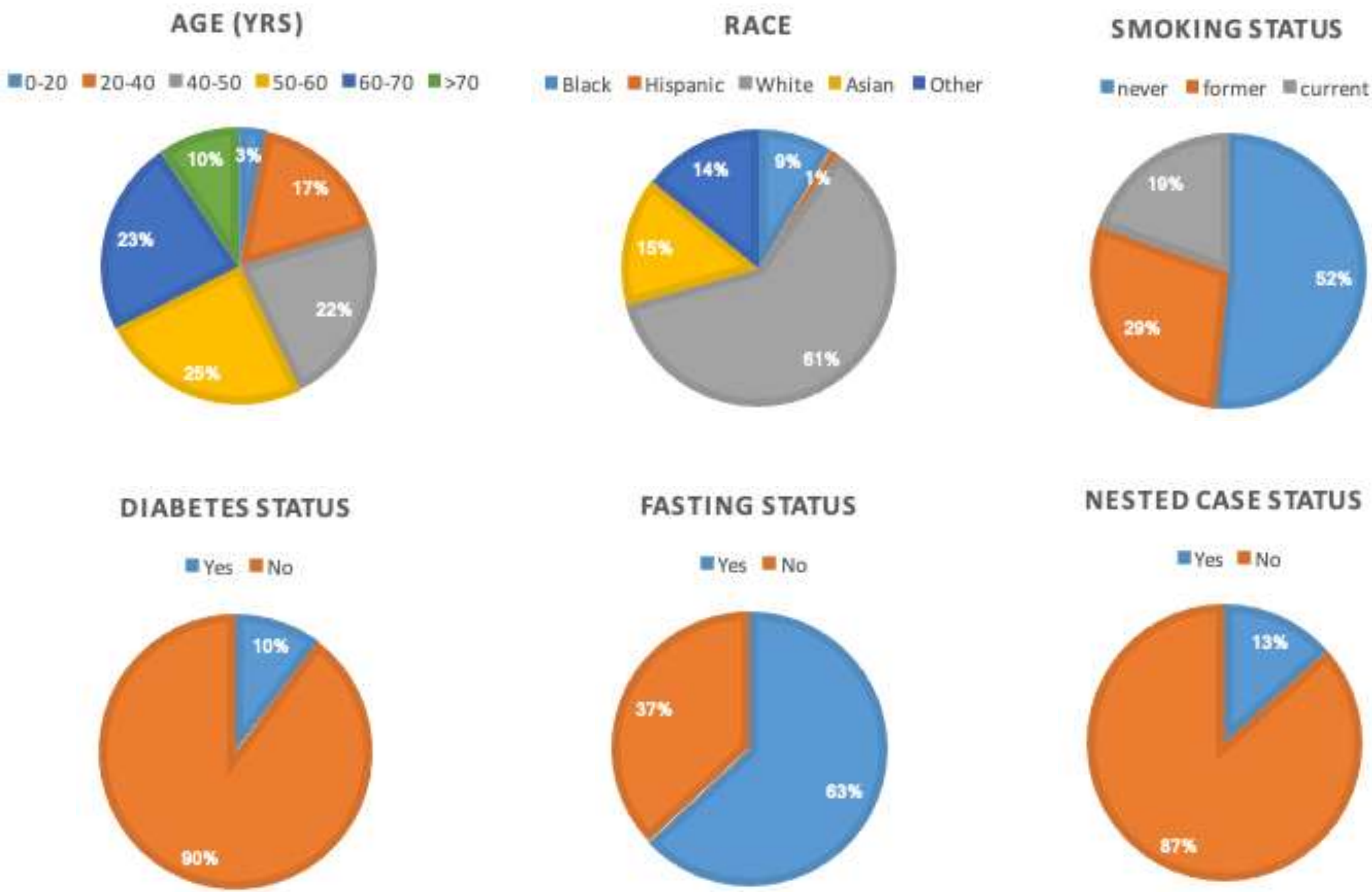




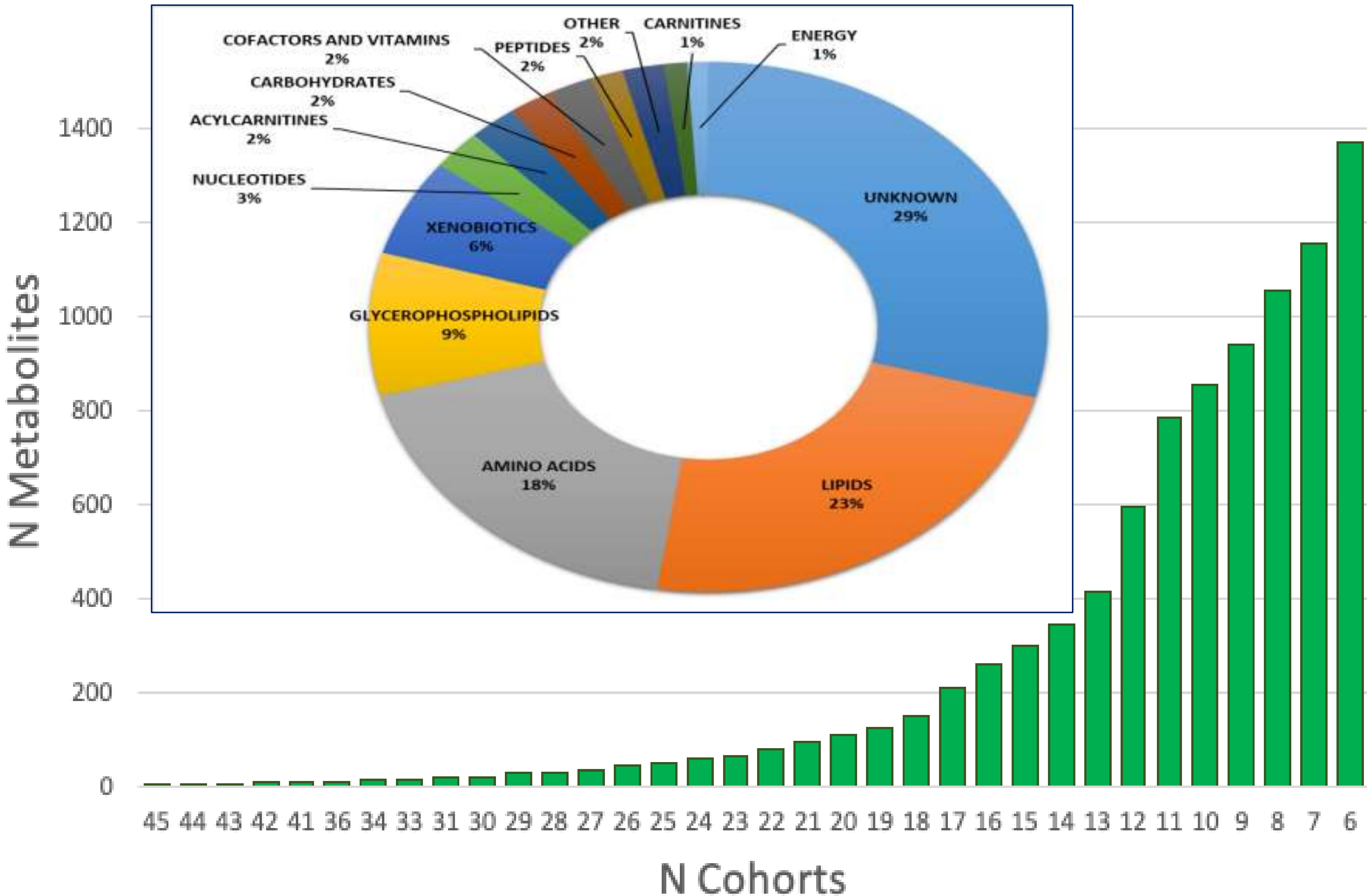
Participant Mean BMI



Participant Characteristics

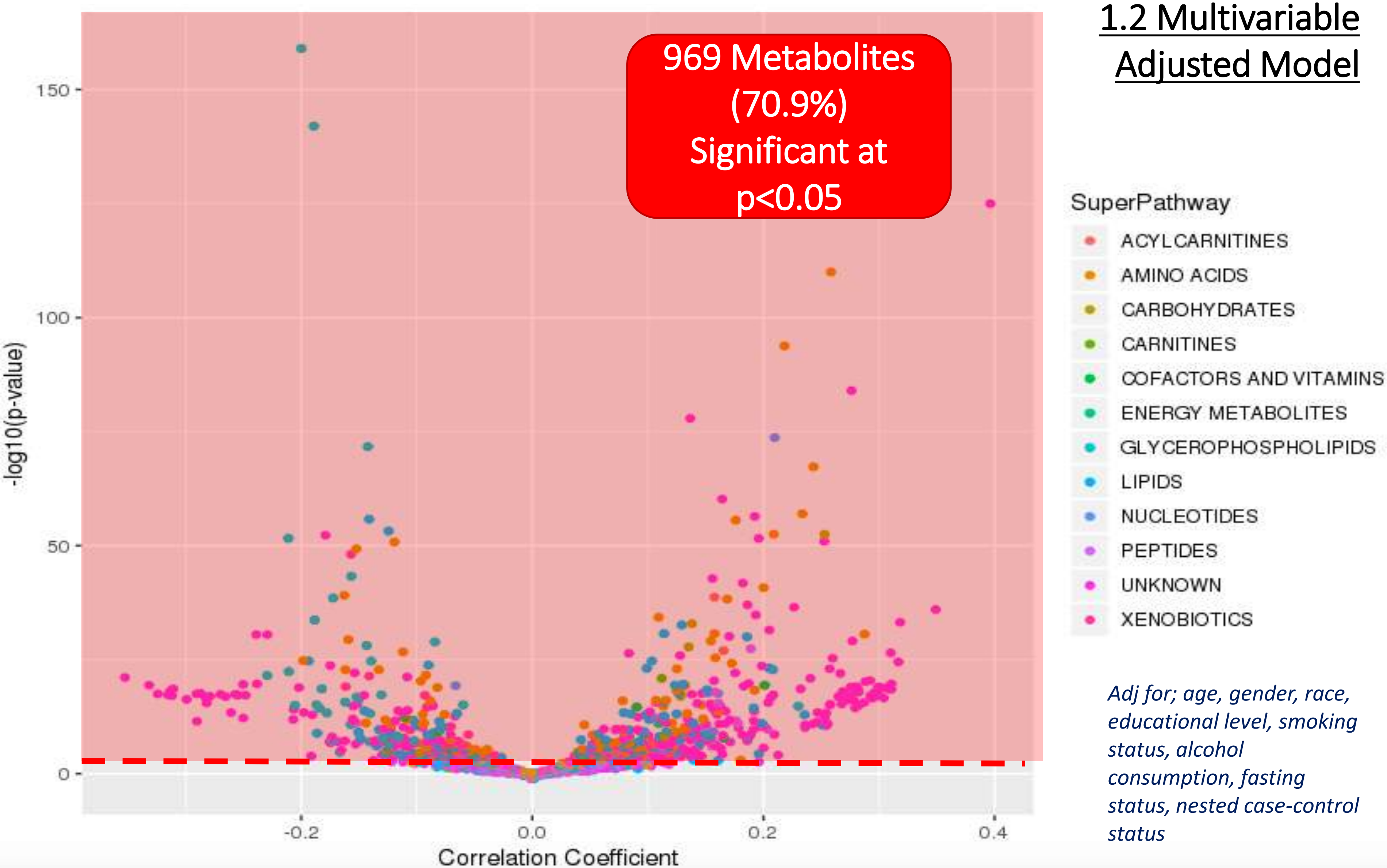


1367 Harmonized Metabolites



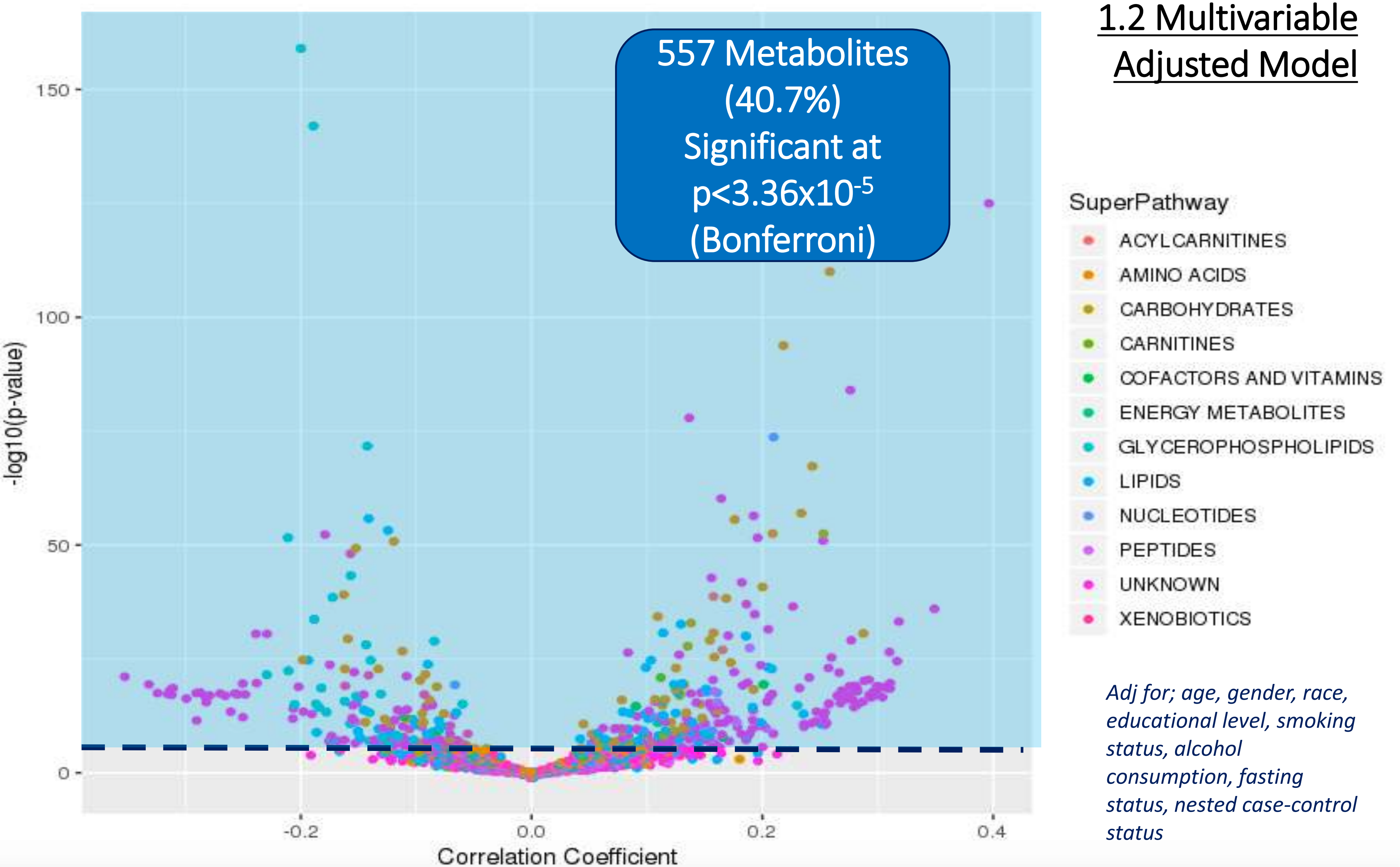
BMI~Metabolite Correlations

1.2 Multivariable Adjusted Model



Adj for; age, gender, race, educational level, smoking status, alcohol consumption, fasting status, nested case-control status

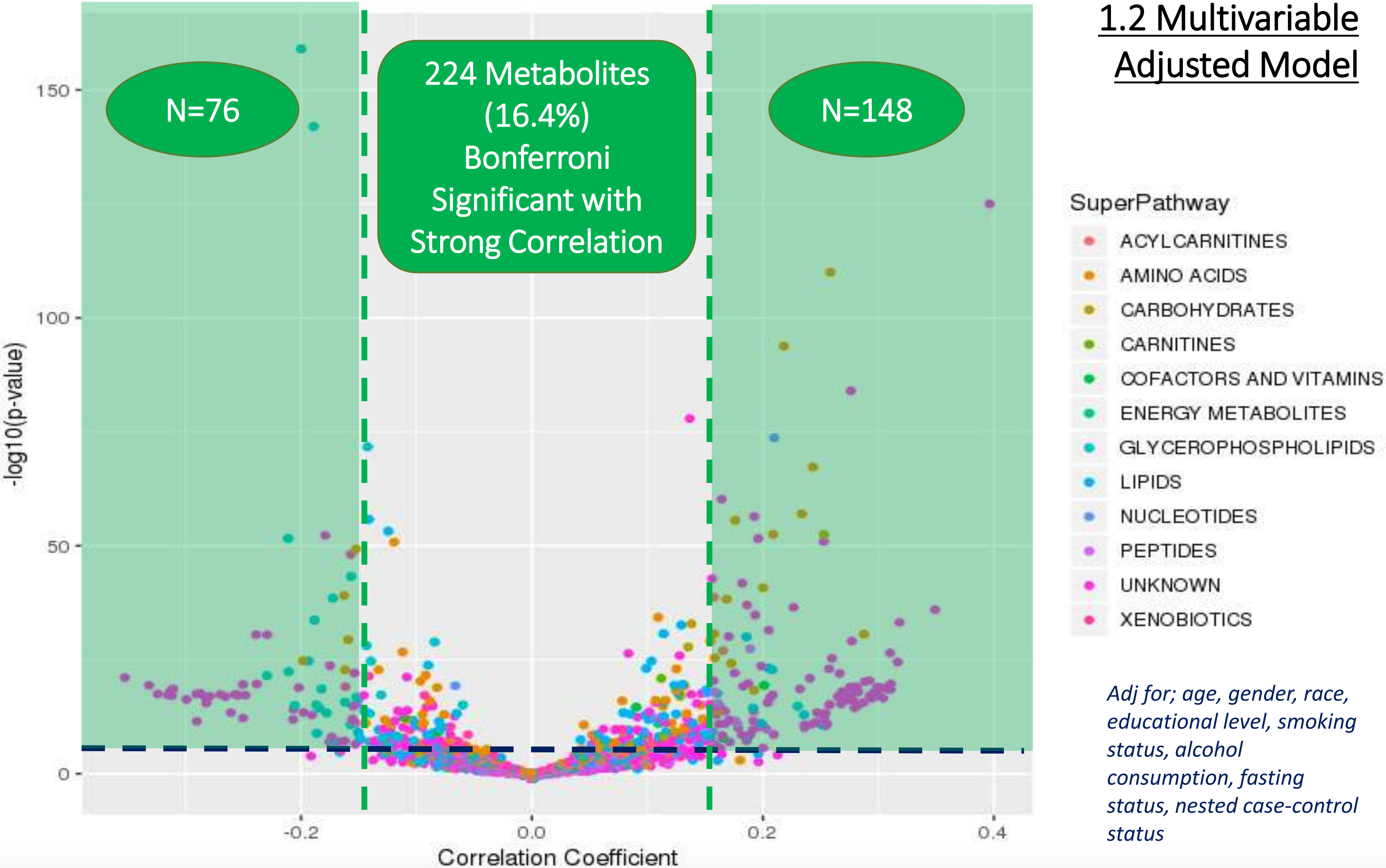
BMI~Metabolite Correlations



1.2 Multivariable Adjusted Model

Adj for; age, gender, race, educational level, smoking status, alcohol consumption, fasting status, nested case-control status

BMI~Metabolite Correlations



1.2 Multivariable Adjusted Model

Adj for; age, gender, race, educational level, smoking status, alcohol consumption, fasting status, nested case-control status

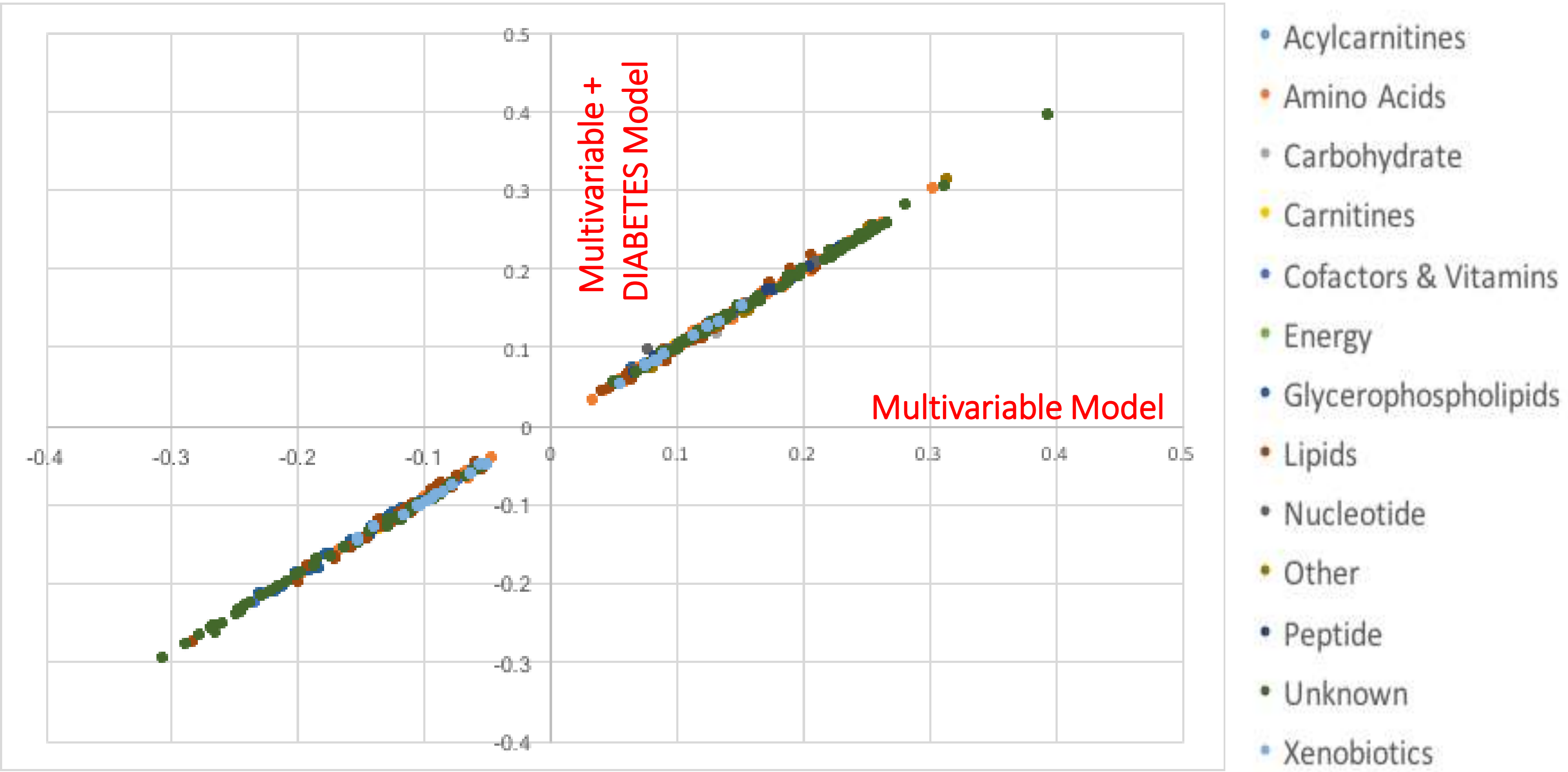
Top Metabolite Hits

Metabolite	Superpathway	Coef. (95%CI)	P-Value
PC(o-22:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	GLYCERO-PHOSPHOLIPIDS	-0.2 (-0.21,-0.19)	6.41E-160
PC(18:0/24:0)	GLYCERO-PHOSPHOLIPIDS	-0.19 (-0.2,-0.17)	5.90E-143
X - 11315	UNKNOWN	-0.18 (-0.2,-0.16)	4.57E-53
LysoPC(18:2(9Z,12Z))	GLYCERO-PHOSPHOLIPIDS	-0.21 (-0.24,-0.18)	2.56E-52
Guanidinosuccinic acid	AMINO ACIDS	-0.15 (-0.17,-0.13)	4.91E-50
Cinnamoylglycine	XENOBIOTICS	-0.16 (-0.18,-0.14)	8.32E-49
2-3-PROPYL2-(TRIMETHYLAMMONIO)ETHYLPHOSPHATE	GLYCERO-PHOSPHOLIPIDS	-0.16 (-0.18,-0.13)	5.50E-44
L-Asparagine	AMINO ACIDS	-0.16 (-0.19,-0.14)	7.42E-40
PC(18:1(9Z)/24:0)	GLYCERO-PHOSPHOLIPIDS	-0.17 (-0.2,-0.15)	3.20E-39
LysoPC(18:1(9Z))	GLYCERO-PHOSPHOLIPIDS	-0.19 (-0.22,-0.16)	1.85E-34

Metabolite	Superpathway	Coef. (95%CI)	P-Value
cortolone glucuronide	UNKNOWN	0.4 (0.37,0.43)	2.06E-125
L-Valine	AMINO ACIDS	0.26 (0.24,0.28)	1.39E-110
L-Tyrosine	AMINO ACIDS	0.22 (0.2,0.24)	1.76E-94
X - 17340	UNKNOWN	0.28 (0.25,0.3)	1.05E-84
N2,N2-Dimethyl guanosine	NUCLEOTIDES	0.21 (0.19,0.23)	1.83E-74
L-Isoleucine	AMINO ACIDS	0.24 (0.22,0.27)	5.57E-68
2-Hydroxy butyric acid	ALPHA HYDROXY ACIDS	0.16 (0.15,0.18)	6.64E-61
L-Leucine	AMINO ACIDS	0.23 (0.21,0.26)	1.12E-57
X - 17357	UNKNOWN	0.19 (0.17,0.22)	3.61E-57
L-Kynurenine	AMINO ACIDS	0.18 (0.15,0.2)	2.71E-56

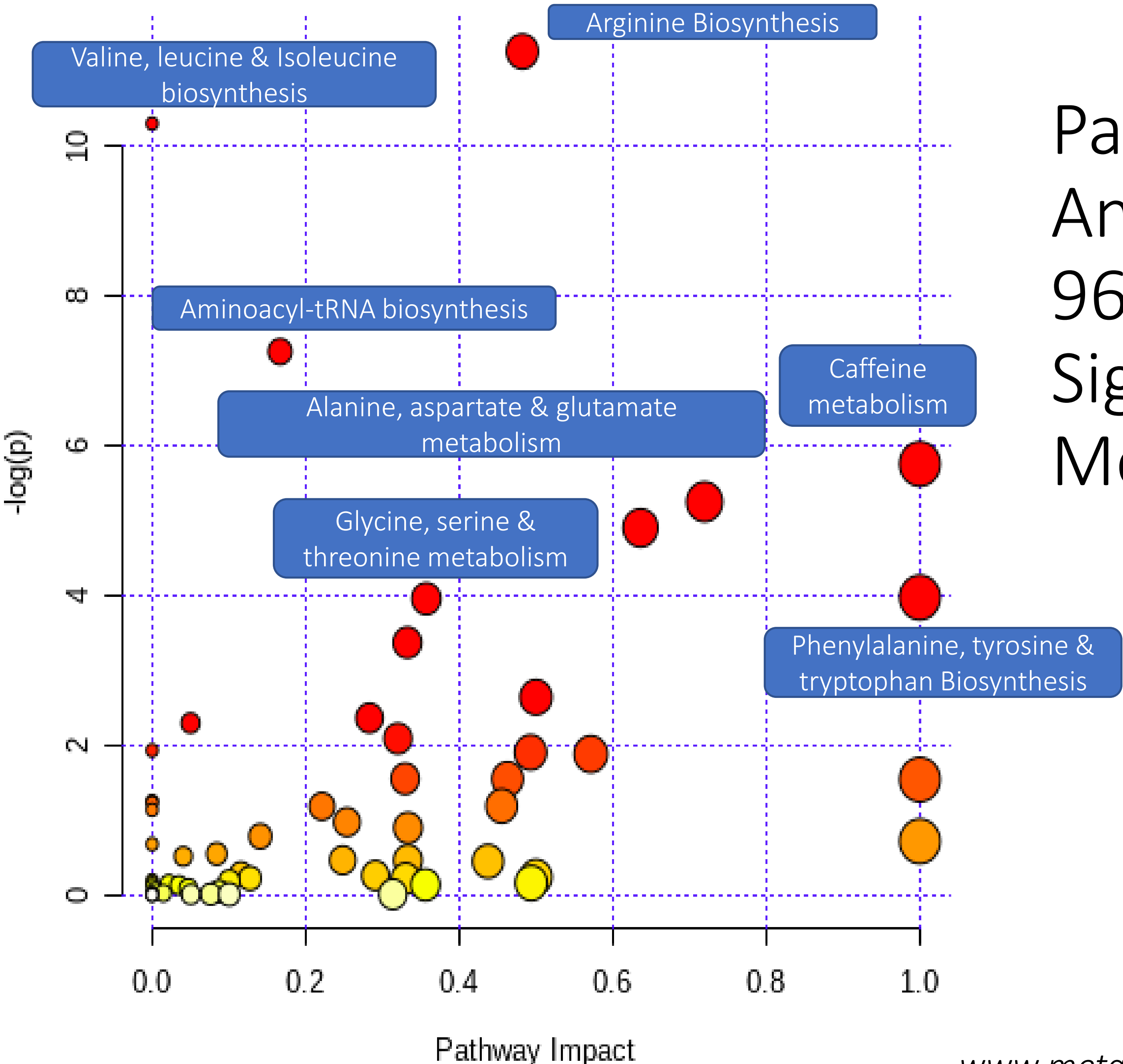
Additional Adjustment for Diabetes

Comparison of Correlation Coefficients for significant metabolites in the two models



Adj for; age, gender, race, educational level, smoking status, alcohol consumption, fasting status, nested case-control status (and Diabetes)

Pathway Analysis of 969 Significant Metabolites

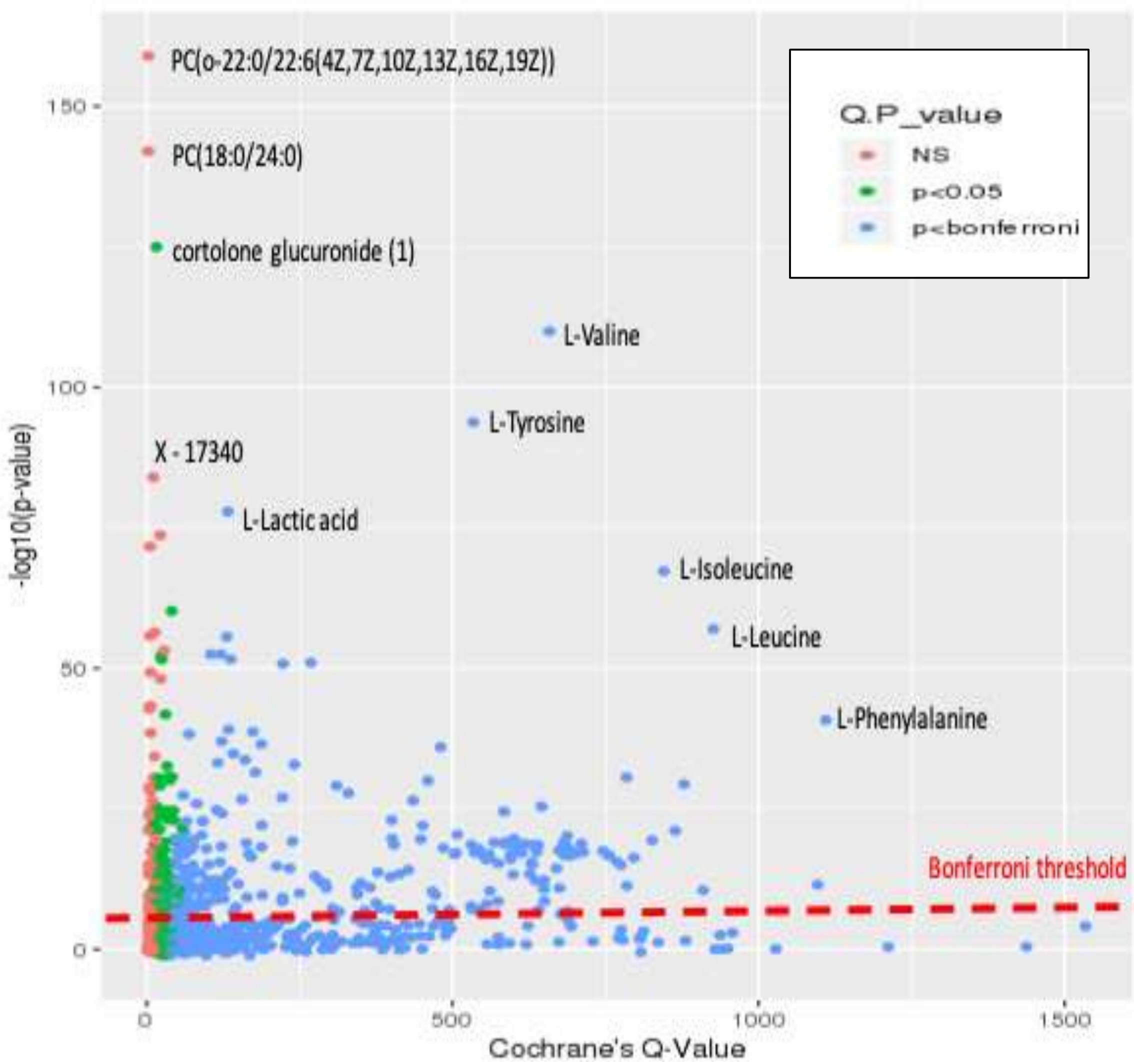


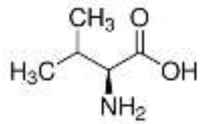
www.metaboanalyst.ca

Cochrane's Q-Value

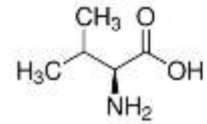
- ❖ Used to assess *<u>between Study Heterogeneity</u>*
- ❖describes % variability in effect estimates due to heterogeneity rather than chance
- ❖Larger Q value (& smaller Q p-value) means more likely there is heterogeneity between studies
- ❖460/557 Bonferroni significant metabolites, (82.6%) were 'significantly' heterogeneous
- ❖Sensitive to large numbers

Cochrane's Q-Value versus –*log10(p-value)* for Each Metabolite

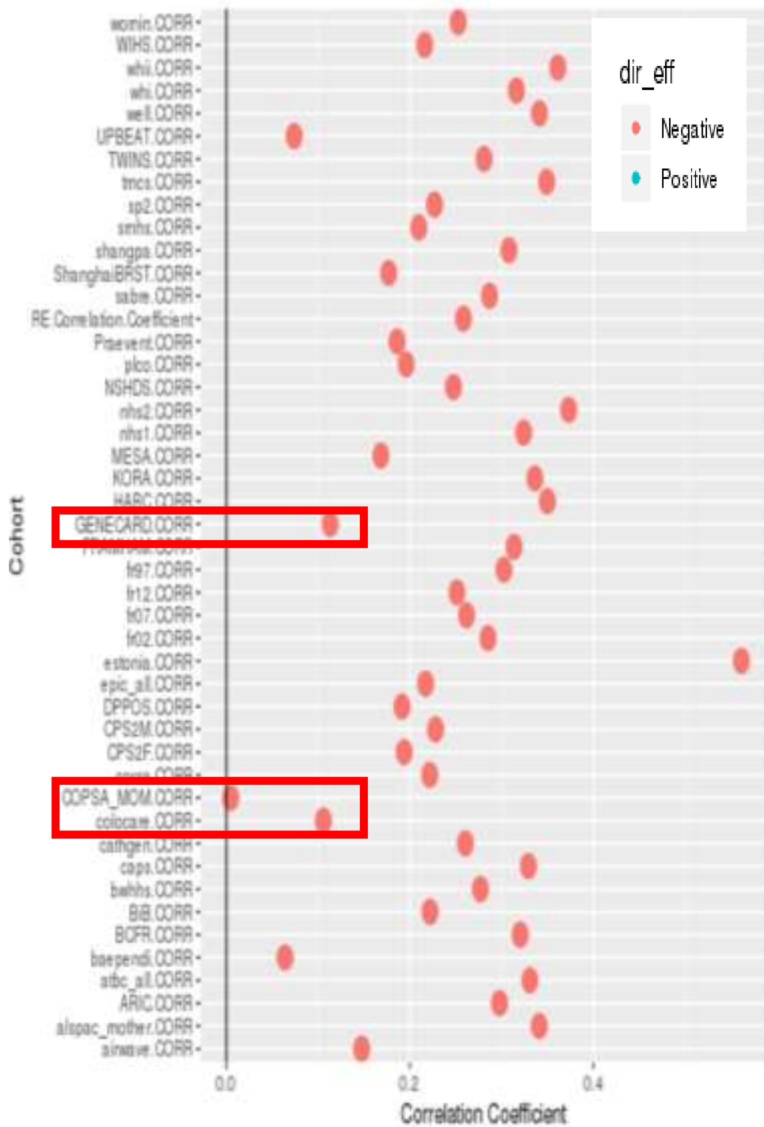




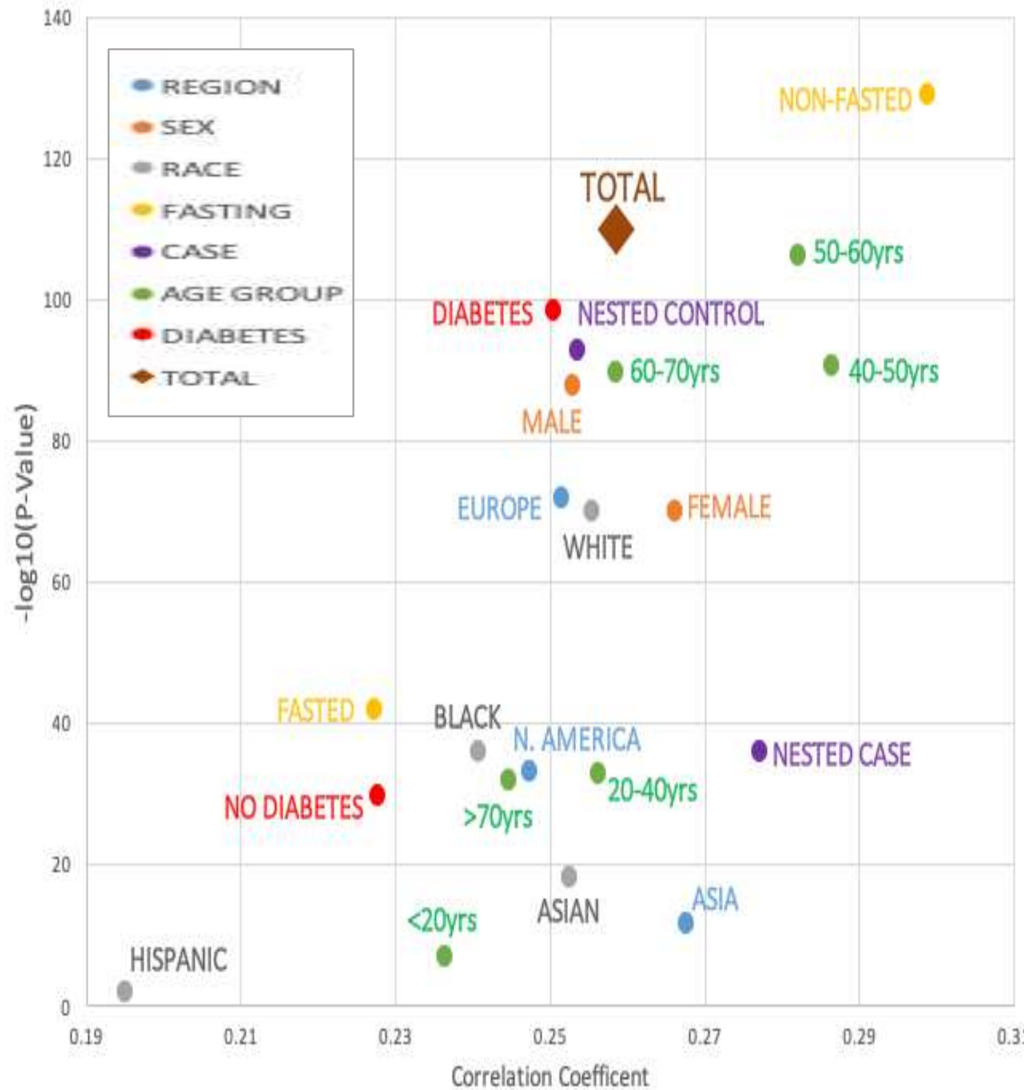
Valine



Correlation Coefficient by Cohort



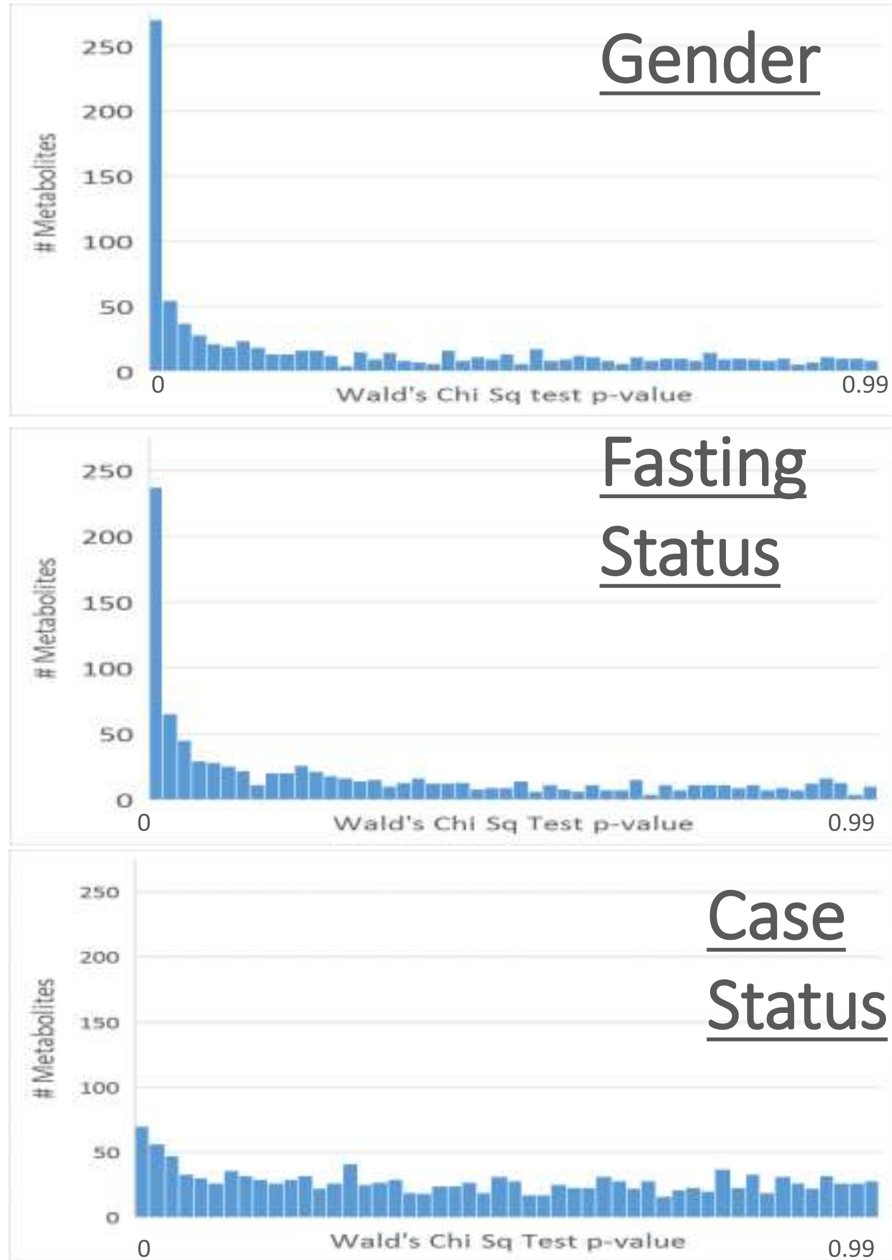
Correlation Coefficient by Strata

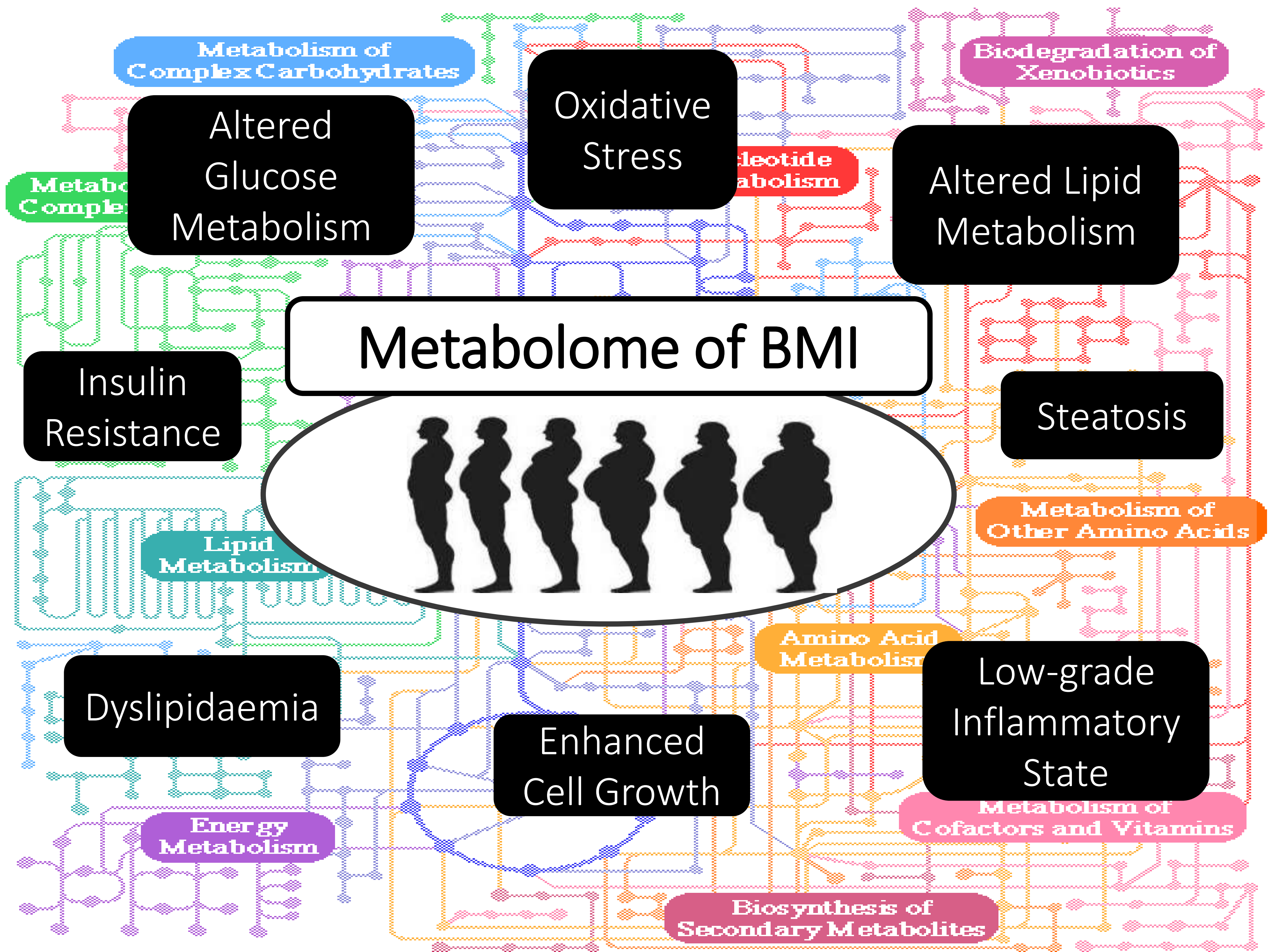


Sources of Heterogeneity

- Wald's Chi Square Test used to assess <u>between Strata heterogeneity</u>
- Fasting Status and Gender are the biggest sources of heterogeneity
- Case status does not appear to be a significant source of heterogeneity

N Metabolites Exhibiting Significant Heterogeneity by Strata





Metabolome of BMI
Altered Glucose Metabolism
Oxidative Stress
Altered Lipid Metabolism
Insulin Resistance
Steatosis
Dyslipidaemia
Enhanced Cell Growth
Low-grade Inflammatory State
Metabolism of Complex Carbohydrates
Biodegradation of Xenobiotics
Lipid Metabolism
Metabolism of Other Amino Acids
Amino Acid Metabolism
Metabolism of Cofactors and Vitamins
Energy Metabolism
Biosynthesis of Secondary Metabolites

Conclusions & Future Directions

COMETS Analytics
http://www.comets-analytics.org

Conclusions

- It is **feasible** to perform large scale meta-analyses across multiple diverse metabolomics cohorts
  - *Different populations*
  - *Different profiling platforms*
  - *Targeted and untargeted*
  - *NMR and Mass Spec*
  - *Serum and plasma*

- We demonstrate that such meta-analyses can provide **robust and biologically informative results**

- An increased BMI is associated with increased levels of amino acids, in particular branched chain amino acids, and with decreased levels of cholesterol esters and High Density Lipoproteins

Next Steps

- ❖Incorporate remaining studies
- ❖Further exploration of heterogeneity
- ❖Assessing correlation between “top hits”
- ❖Identification of metabolite profiles/signatures in addition to individual metabolites
- ❖Considerations of extremes of BMI
- ❖Consideration of adiposity measures in a subset
- ❖Pathway/Network Interpretation

- Recruitment of cohorts is ongoing

- Any participating investigator can submit a project proposal for a meta-analysis across the cohorts

- A wealth of data is waiting to be explored!

https://epi.grants.cancer.gov/comets/

Acknowledgements

Jessica Lasky Su (BWH, HMS)

Steven Moore (NIH)**

David Ruggieri

Kaitlyn Mazzilli

Ella Temprosa (George Washington University)

Ewy Mathe (Ohio State)

NIH: Krista Zanetti, Gwendolyn Alexandre

***IMS**: Nathan Appel; Adam Risch*

***COMETs harmonization committee:** Eric Boerwinkle, Eoin Fahy, Michael Gunter, Lucas Lotta, Alexandre Pereira, Mary Playdon, Adam Risch, Josh Sampson, Bing Yu*

CBIIT web development Team

The BMI writing group & Diabetes working group

All participating COMETS Cohorts

COnsortium of METabolomics Studies

https://epi.grants.cancer.gov/comets/

COMETS Analytics

http://www.comets-analytics.org

NIH National Institutes of Health

Questions: Rachel Kelly (hprke@channing.harvard.edu)