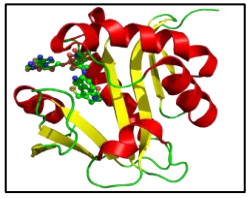
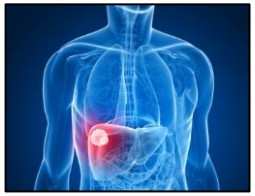
Frontiers in Air Quality Science

Celebrating 21 years of the Environmental Research Group

The exposome approach to the study of exposure-disease associations

Paul Elliott

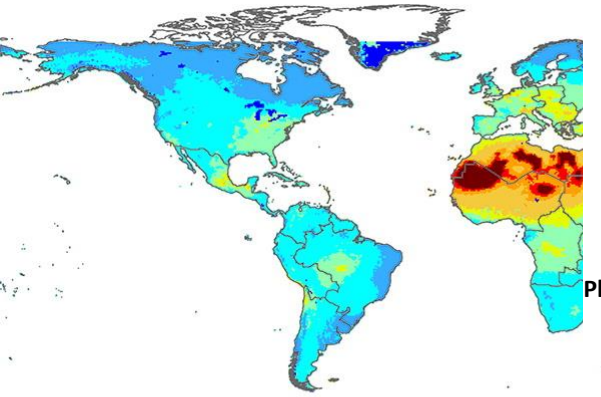
Director of MRC-PHE Centre for Environment and Health

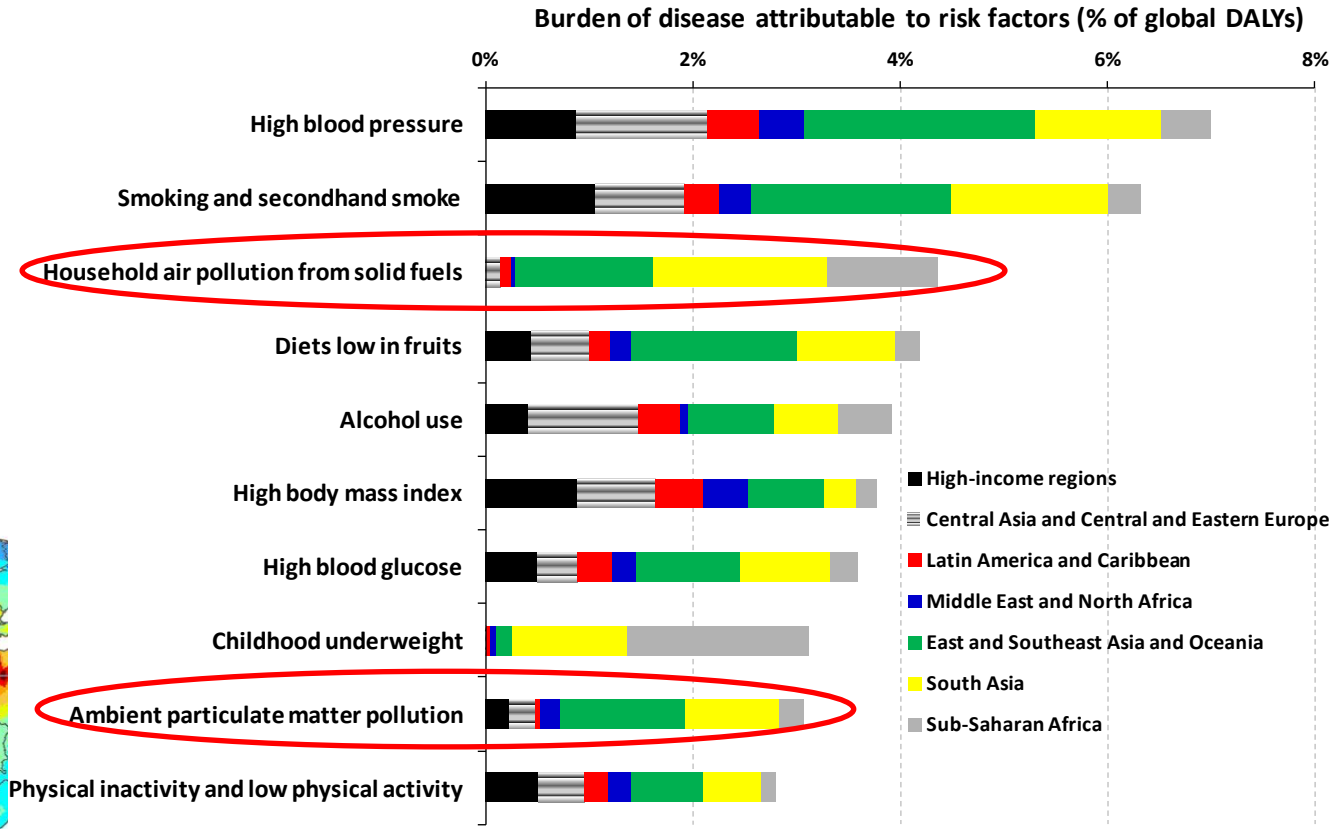
Imperial College London

MRC-PHE
Centre for Environment & Health

Public Health England | MRC Medical Research Council | Imperial College London | King's College London

The Global Burden of Air Pollution and Other Preventable Risk Factors





Lim et al *Lancet* 2012; Ezzati and Riboli *NEJM* 2013

Brauer et al *Environmental Science and Technology* 2012

MRC-PHE
Centre for Environment & Health

Public Health England | MRC Medical Research Council | Imperial College London | King's College London

“Genetics loads the gun, but Environment pulls the trigger”

After Elliott Proctor Joslin MD,

A pioneering American diabetologist and founder of the Joslin Diabetes Center, 1869-1962

Br Med J 1991; 302: 1231

Genome-Wide Association Studies (GWAS) Paradigm

- **Untargeted** analysis
- Has led to discovery of hundreds of new replicated genetic associations in past 5 years
- Typically examine 2M+ common variants (measured and imputed) across the genome
  - Problem of multiple testing → Bonferroni correction (genome-wide 5×10^{-8})
- Large sample size (consortia, meta-analyses)
- Replication
- ***Typically explains only a small part of the population variance in disease/trait***

While genetic data are a (fixed) digital read-out...

Environmental exposure data vary over time, are continuously distributed, with wide dynamic range...

And difficult to measure....

Such that ... the quality of the exposure data has been called the Achilles heel of environmental epidemiology

New approaches required!

The Exposome refers to the totality of environmental "exposures" from conception onwards (Lifecourse)

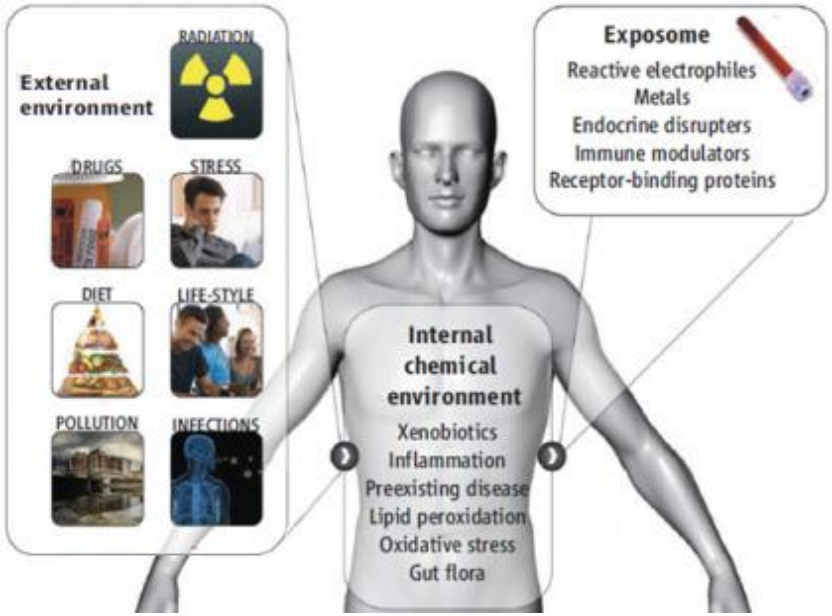
The ***internal exposome*** is based on measurements in biological material of ***complete sets of biomarkers***, especially ***during critical life stages***, through use of high throughput ***omics techniques*** (Rappaport & Smith, *Science* 2010)

Biomarkers cover a wide range of molecules, ranging from e.g., xenobiotics, metabolites in blood or urine (***metabonomics***), proteins (***proteomics***), mRNA (***transcriptomics***), covalent complexes with DNA and proteins (***adductomics***)

Exposome Concept

The Exposome is a new paradigm for studying the environmental causes of disease

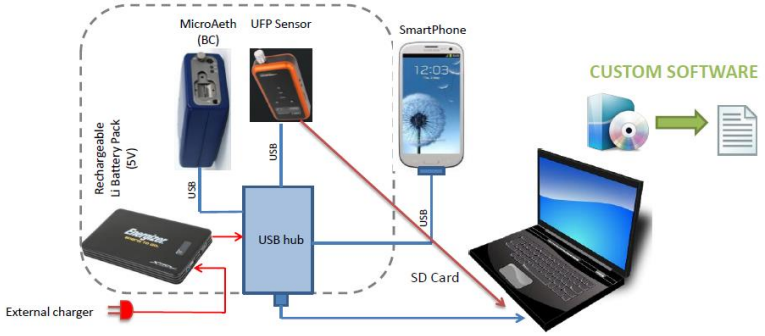
Coverage	All chemicals (exogenous and endogenous) and mixtures
Focus	Stratified population
Study design	Hypothesis generating (discovery)
Measurement	Better and quantitative measurements, understanding of mechanisms
Monitoring	External and internal environment
Time frame	Conception through late adulthood, life-course

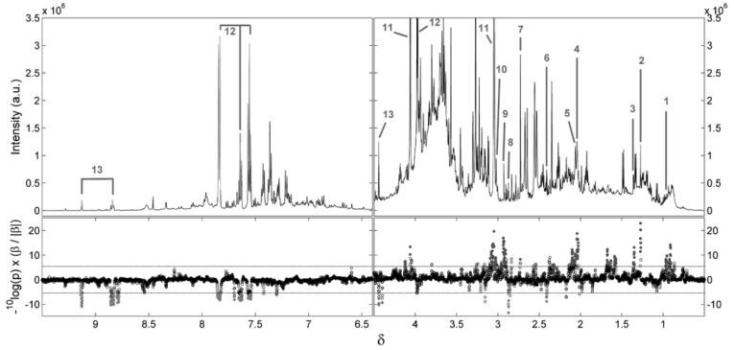


Rappaport and Smith, *Science*, 2010

Exposome - New Opportunities

- New technologies for measurement and modelling of environmental exposures
- New approaches (*omics*) to capitalise on population cohorts and biobanks
- New insights on pathways and mechanisms linking environmental exposures to disease (*exposome*)

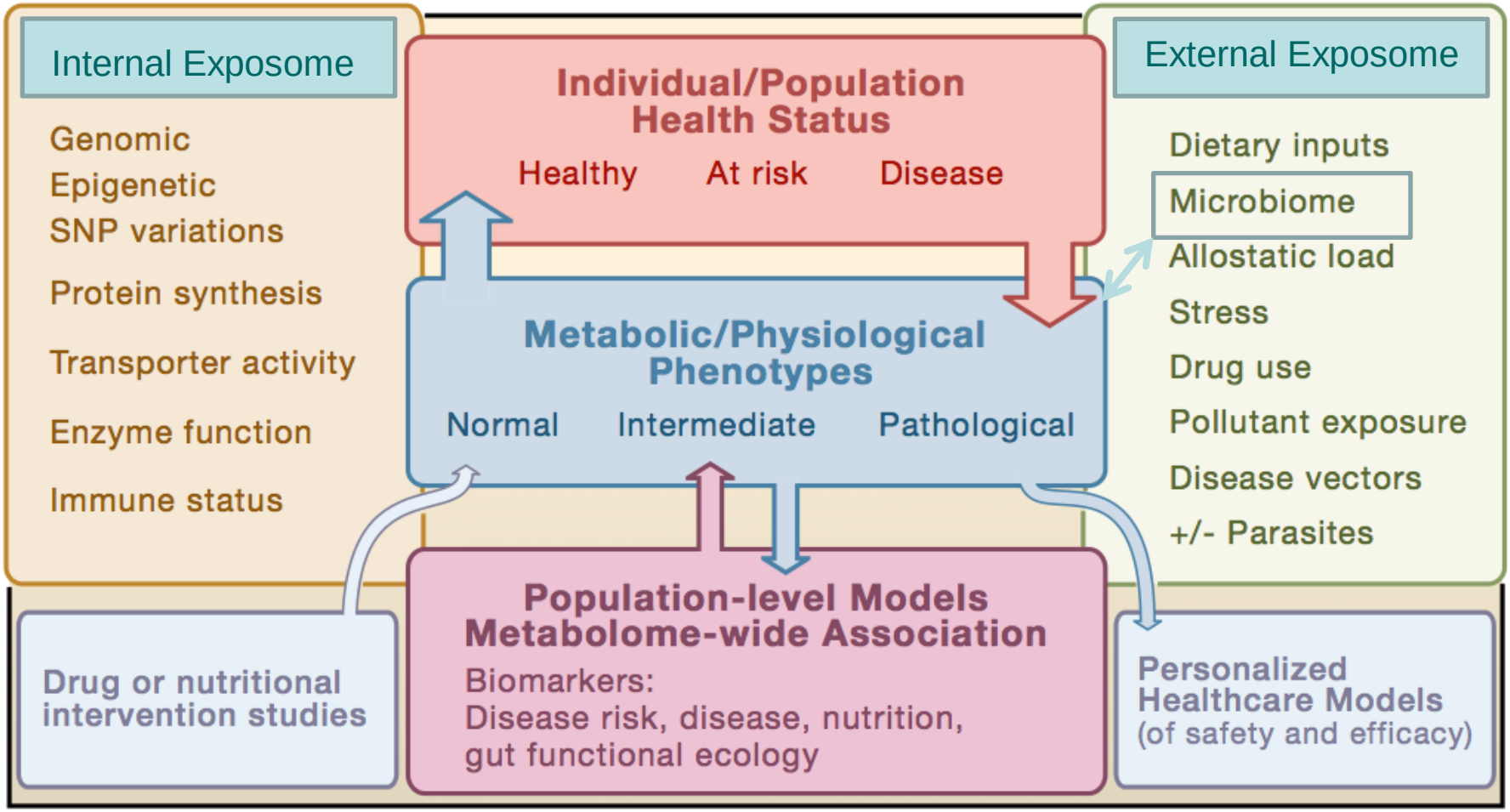




MRC-PHE
Centre for Environment & Health

Public Health England | MRC Medical Research Council | Imperial College London | King's College London

“Individual Exposomes”



Metabolic Phenotyping in Health and Disease

Cell (2008)134: 714-717.

Cell

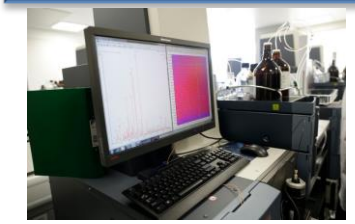
E. Holmes,[1] I.D. Wilson,[2] and J.K. Nicholson[1,*]

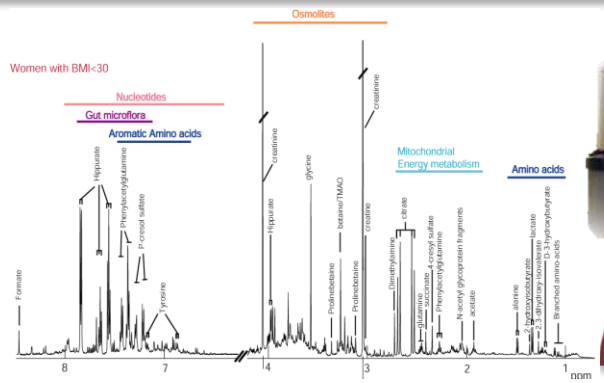
MRC-PHE
Centre for Environment & Health

Public Health England | MRC Medical Research Council | Imperial College London | King's College London

Exposome Concept 2

Epidemiological study

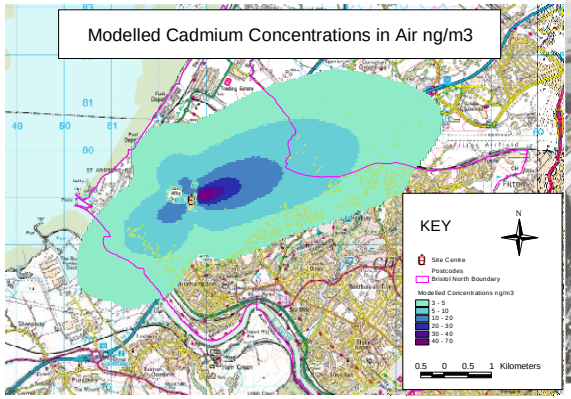


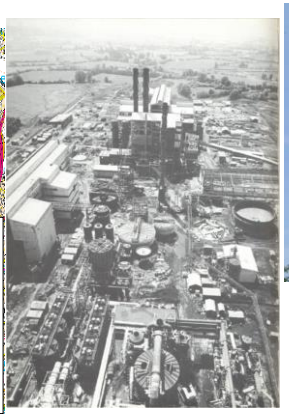


"Lifecourse"

Exposure → **Biomarkers of exposure** → **Biomarkers of early effect** → **Disease**

Environmental sensors & modelling

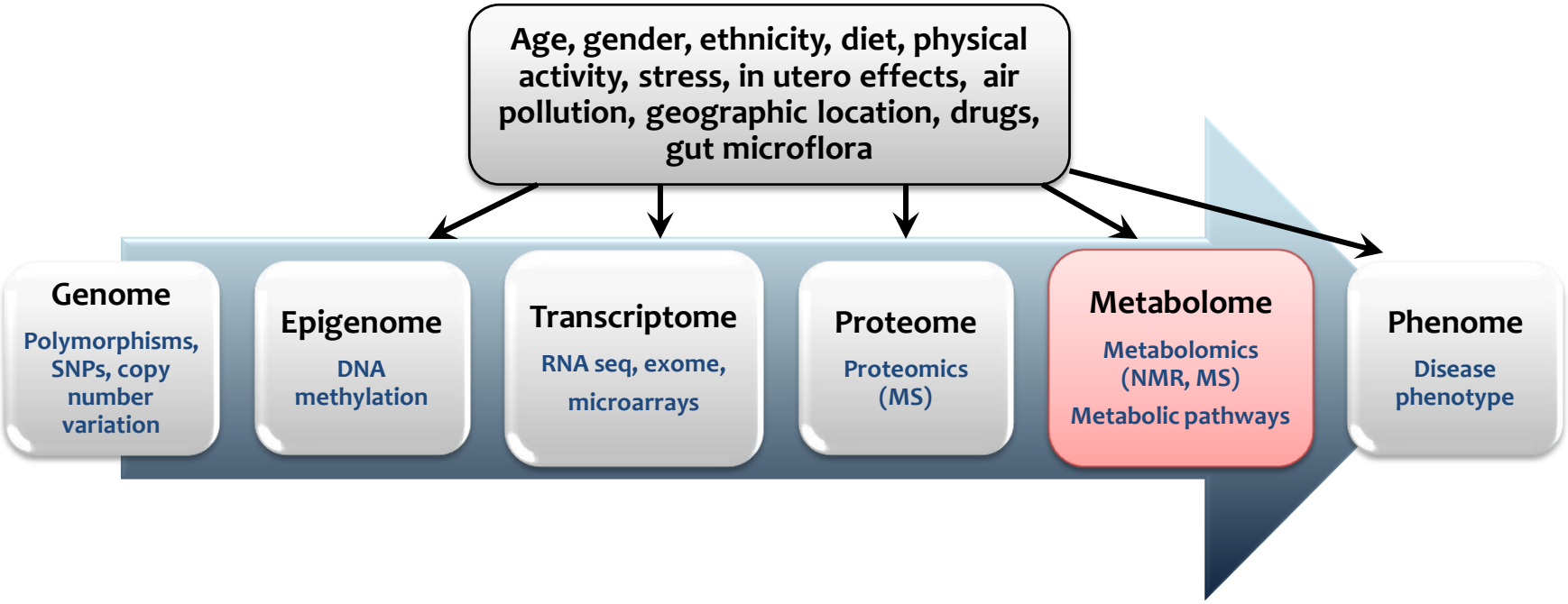




MRC-PHE
Centre for Environment & Health

Public Health England | MRC Medical Research Council | Imperial College London | King's College London

Relationships and interactions between different *omic* measurements



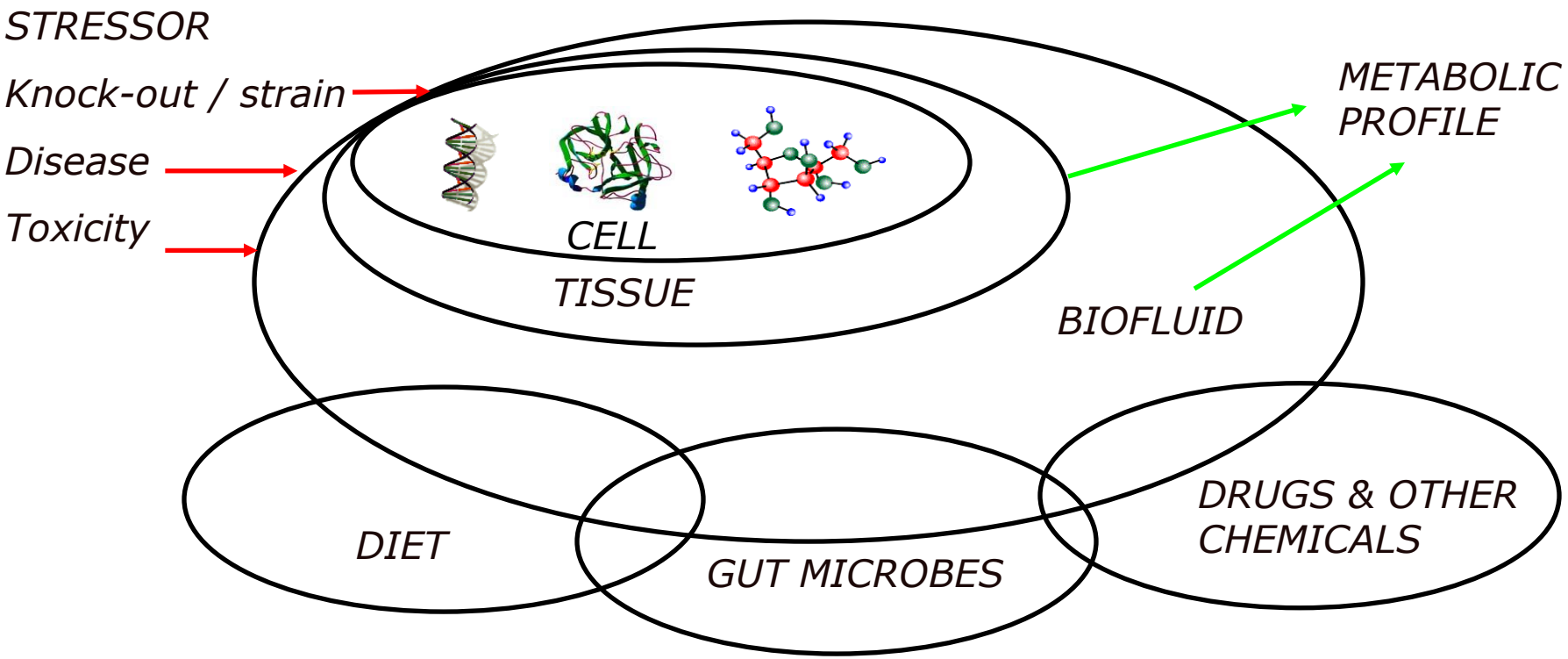
Tzoulaki et al *AJE* 2014 (in press)

Metabolic Profiling

Study of the complement of small molecules within biological systems

Also known as metabolomics or metabonomics

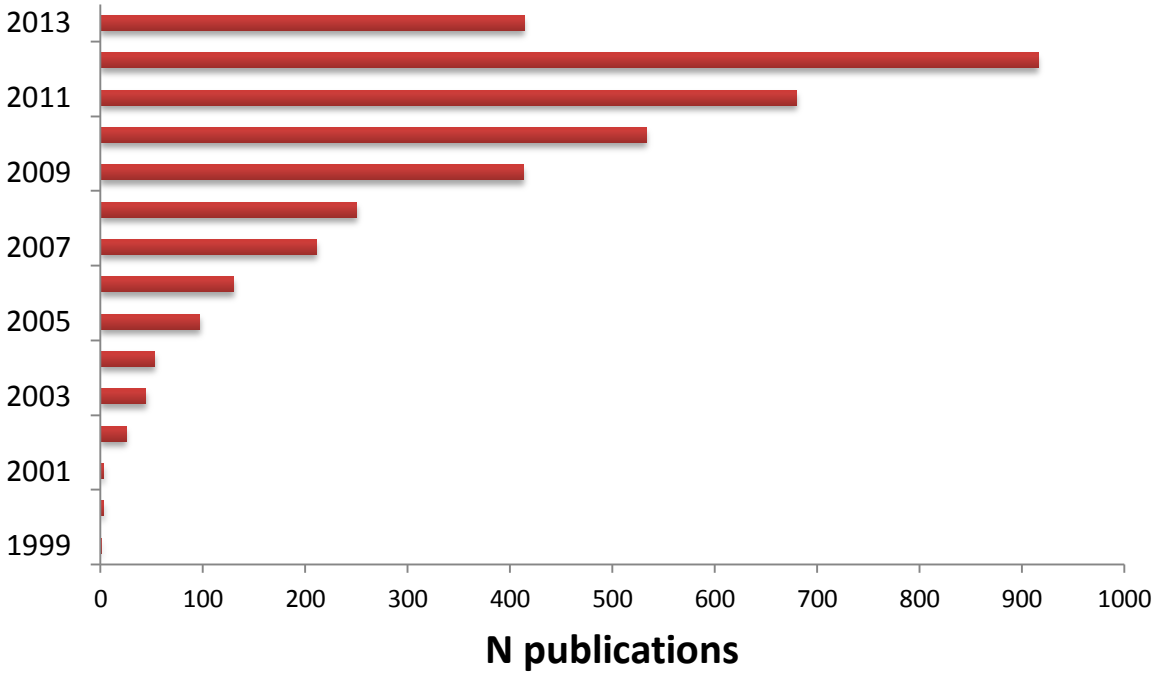
Untargeted: No prior hypothesis of specific metabolites involved



System ⟷ Environment

MRC-PHE
Centre for Environment & Health
Public Health England
MRC Medical Research Council
Imperial College London
KINGS College LONDON

Metabolomics publications

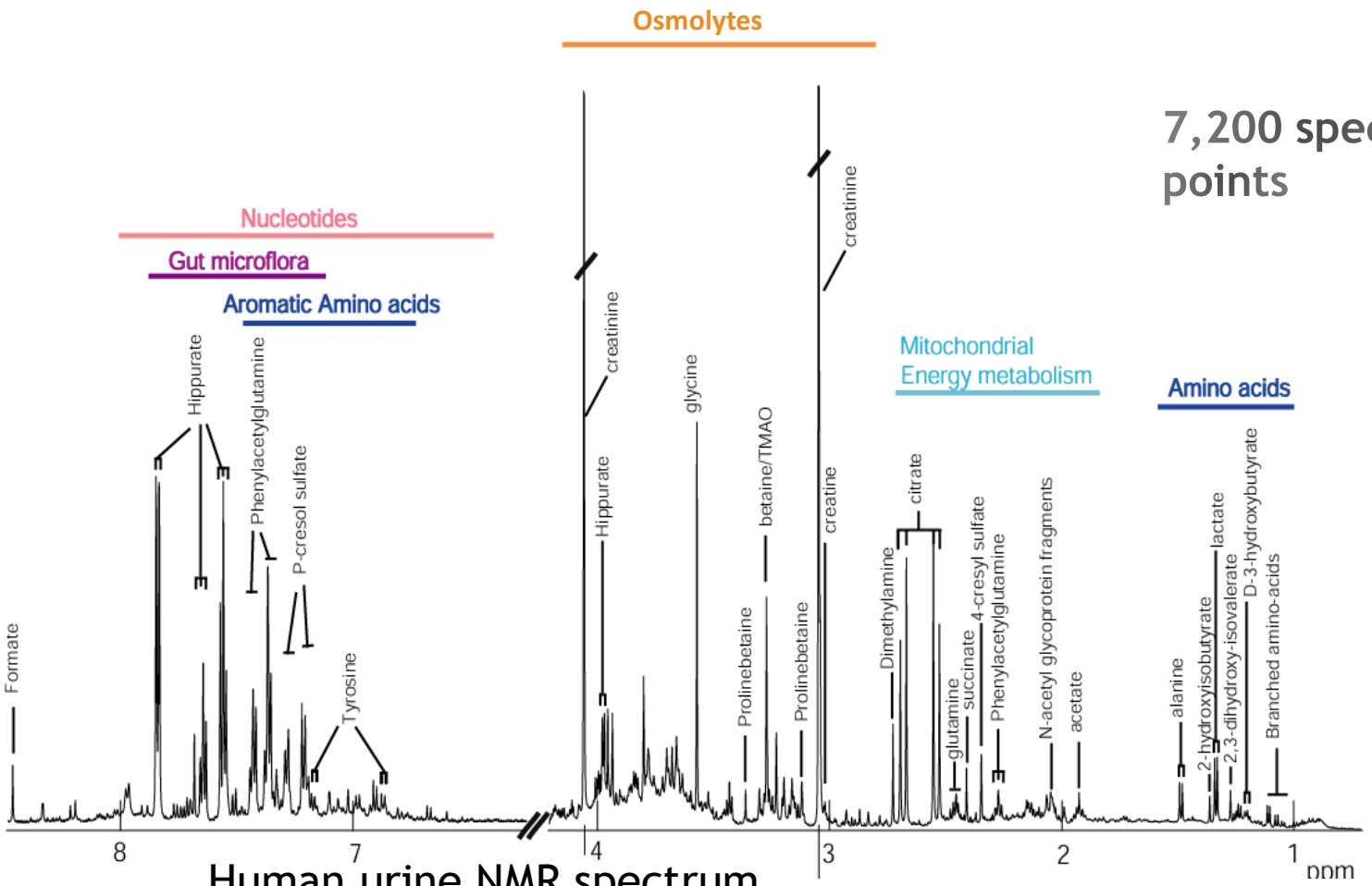


- Publications in PUBMED with terms (metabolomic* OR metabonomic* OR metabolome) by year of publication

Tzoulaki et al *AJE* 2014 (in press)

NMR metabolomics

High-throughput Analytical Analysis of metabolite content of biological samples



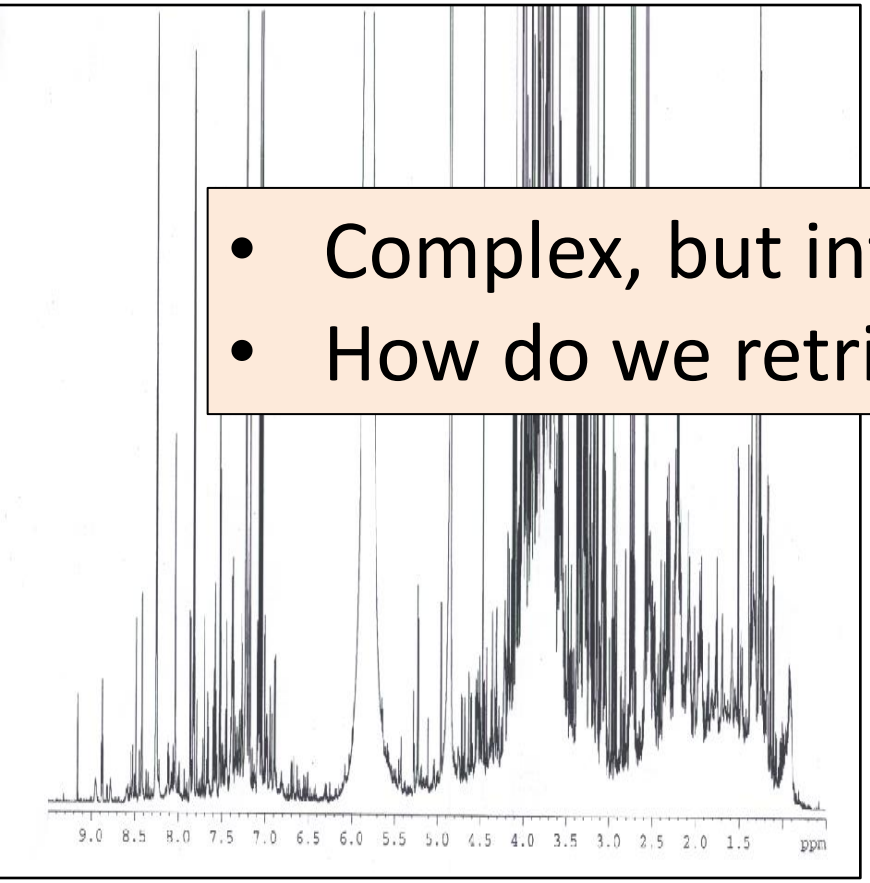
7,200 spectral points

Human urine NMR spectrum

^{1}H Nuclear Magnetic Resonance (NMR) Spectroscopy (600 MHz, Bruker)

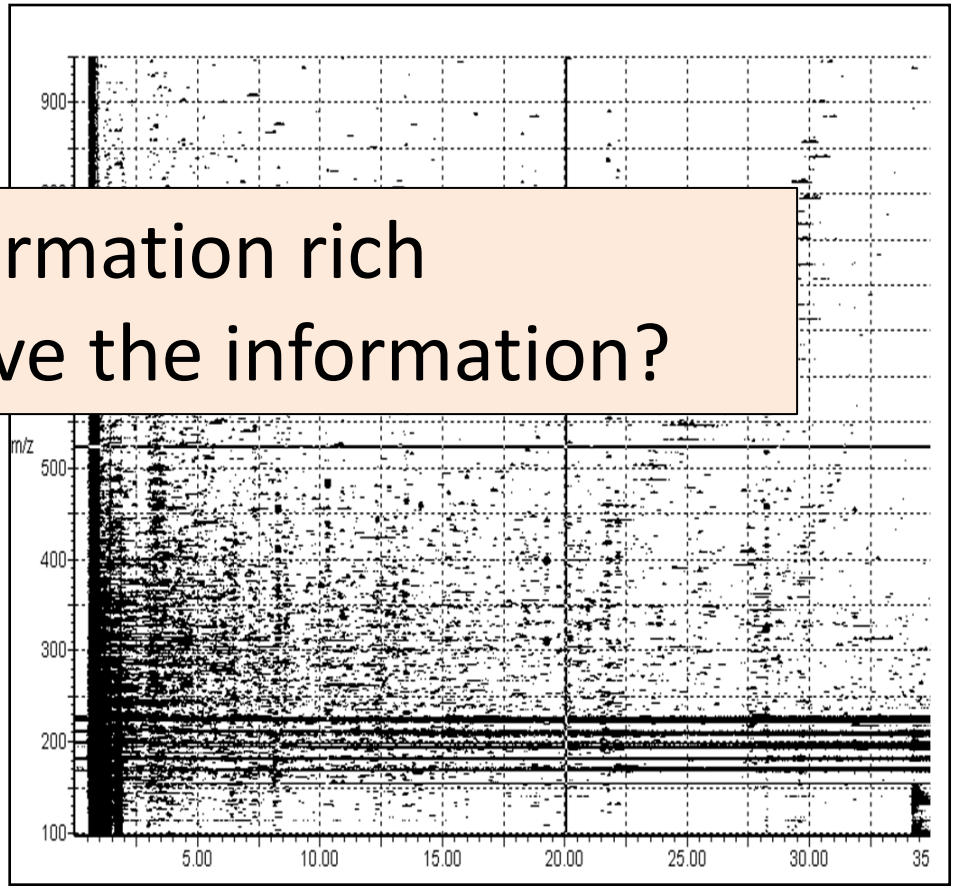
The Challenge of Metabolomic Data

NMR



~1000s signals,100s metabolites

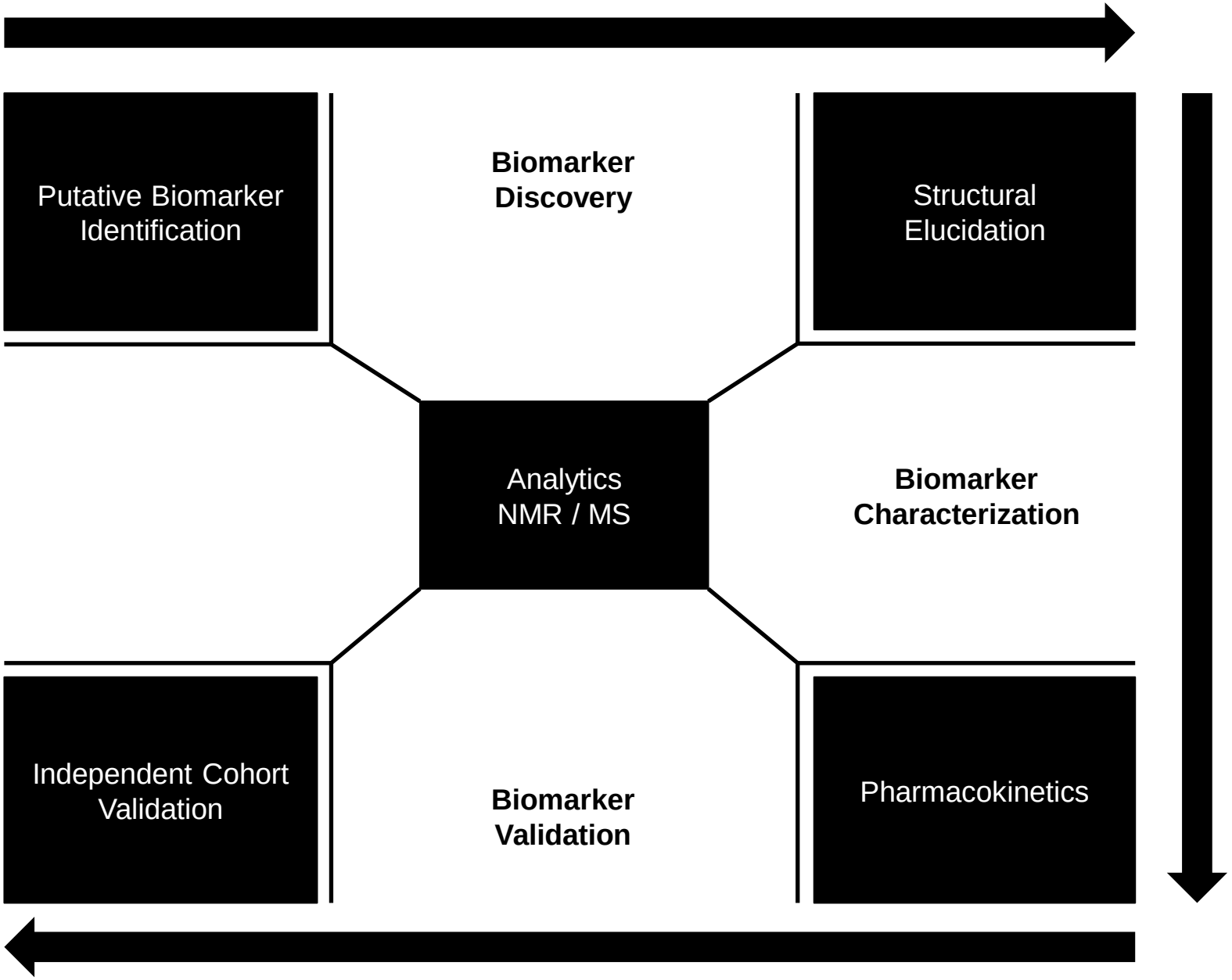
LC-MS



~10,000s signals,1000s (?) metabolites

- Complex, but information rich
- How do we retrieve the information?

Measuring The Metabolome



Human metabolic phenotype diversity and its association with diet and blood pressure

Elaine Holmes[1]*, Ruey Leng Loo[1,2]*, Jeremiah Stamler[3], Magda Bictash[1,2], Ivan K. S. Yap[1,2], Queenie Chan[2], Tim Ebbels[1], Maria De Iorio[2], Ian J. Brown[2], Kirill A. Veselkov[1], Martha L. Daviglus[3], Hugo Kesteloot[4], Hirotsugu Ueshima[5], Liancheng Zhao[6], Jeremy K. Nicholson[1] & Paul Elliott[2]

Nature 2008

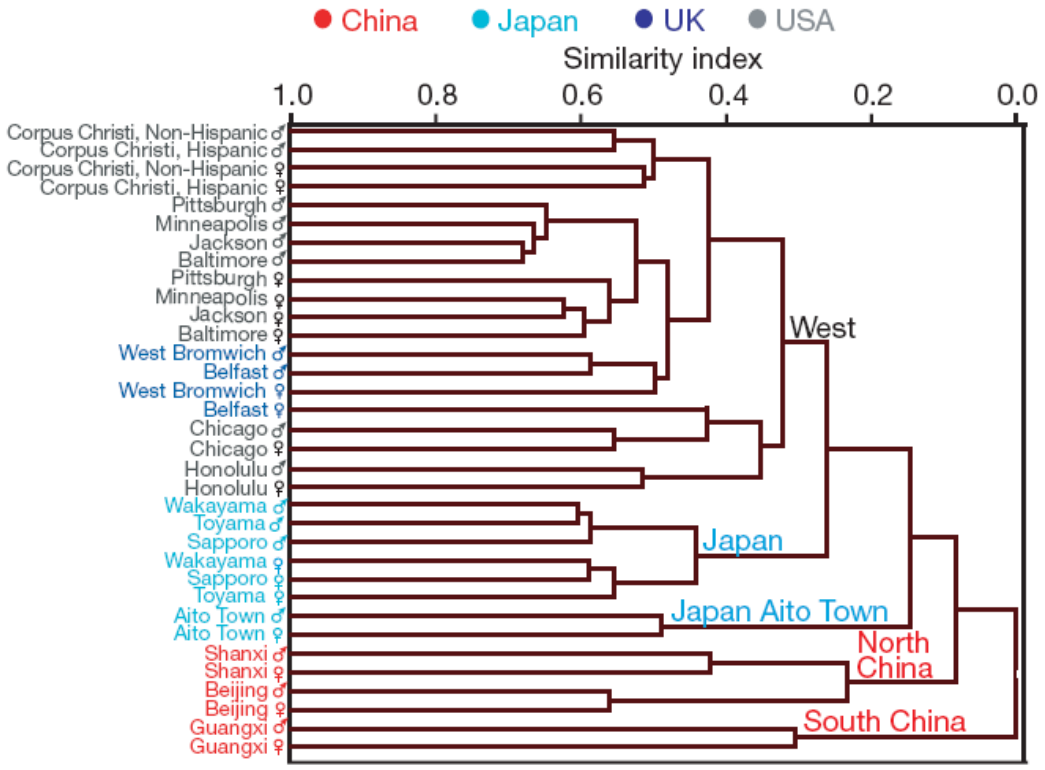


Figure 1 | Hierarchical cluster analysis using group average linkage based on median ^{1}H NMR urine spectra, by population sample and gender ($n = 4,630$). Data for first 24-h urinary specimens.

Metabolites responsible for discriminating populations and diet, also showed significant associations to blood pressure in individuals

Table 1 | Estimated mean differences in systolic and diastolic BP

Urinary metabolite	A*	
	Not adjusted for BMI‡	
Alanine	2.69	(6.06)
Formate	−1.19	(−2.62)
Hippurate	−2.10	(−4.85)
N-methylnicotinate	−0.09	(−0.21)
Alanine	1.57	(5.17)
Formate	−0.90	(−2.96)
Hippurate	−0.98	(−3.33)
N-methylnicotinate	−0.07	(−0.25)

MRC-PHE
Centre for Environment & Health

Public Health England | MRC Medical Research Council | Imperial College London | King's College London

Discriminatory metabolites - across populations

	NMR Spectroscopic Data	
Urinary Metabolites	**Moieties**	**Signal directly observed, δ ¥**
2-aminoisobutyric acid	CH; CH_2	1.20 (d); **2.61(m); 3.11(m)**
2-hydroxyibuprofen	CH_3	**1.12(s);** 1.52(d); 2.78(s); 3.56(t); 3.83(q); 3.98(q); 7.22-7.37(m)
2-oxoglutarate	CH_2	**2.45(t),** 3.00(t)
3-hydroxyisovalerate	CH_3	**1.27(s);** 2.37(s)
Acetylcarnitine	CH_3	2.17 (s); 2.52(dd); 2.65 (dd); **3.21(s);** 3.61 (d); 3.85(dd)
Alanine	CH_3	**1.48 (d)**; 3.79 (qt)
Citrate	CH_2	**2.54(d); 2.66(d);**
Creatine	CH_2	3.04(s); **3.94(s)**
Ethanol	CH_2; CH_3	**1.2(t); 3.65(q)**
Ethylglucoside	CH_3	**1.24(t)**;3.41(dd);3.55(dd);3.69(m);3.71(dd);3.77(dd);3.80(q);3.86(dd)
Formate	CH	**8.46(s)**
Guanidinoacetate	CH_2	**3.80(s)**
Hippurate	CH_2; CH; CH; CH	**3.97 (d); 7.55 (t); 7.64 (t); 7.84(d)**
Lysine	CH_2;CH_2	1.47(m); 1.72 (m); **1.89 (m); 3.01 (t)**; 3.77(t)
N-acetyl-glycoproteins	CH_3	**1.98 - 2.06**
N-methyl-2-pyridone-5 carboxamide	CH; CH	**8.34(m)**
N-methylnicotinate	CH_3;CH; CH;CH	**4.44(s); 8.08(t); 8.84(t); 9.13(s)**
NNN-trimethylysine	N-(CH_3)	**3.12(s)**; 3.35(m); 3.76(m)
Phenylacetyglutamine	CH; CH	**1.98 (m)**; 2.13 (m); 4.20(m); **7.37(t); 7.43(t)**
Protein envelope (unresolved)‡	-	**0.90 - 0.98**
Sarcosine	CH_3	**2.75 (s)**
Suberate	CH_2	**1.31(m); 1.55(m); 2.18(t)**
Succinate	CH_3	**2.41(s)**
Taurine	CH_2	3.26 (t); **3.43(t)**
Trimethylamine-*N*-oxide	N-(CH_3)$_3$	**3.27 (s)**
Unknown 1	-	**2.80 (s)**
Unknown 2	-	**3.36(s)**
Unknown 3	-	**1.27**
Unknown 4	-	**7.96 (m)**
Unknown 5	CH_3	**2.78 (s)**

¥ chemical shifts in bold indicated signals found to be significantly different in the OPLS-DA pairwise comparisons (Suppl Fig. 1A) and those not in bold were directly observed in 1D ^{1}H-NMR and their connectivities confirmed by statistical total correlation spectroscopy (STOCSY)

‡ protein envelope is overlapped with superimposed signals from isoleucine, leucine, valine

Criteria for discriminatory metabolites included $P<7x10^{-6}$ (Bonferroni) and ranked in top 5% of all coefficients *for both 1st and 2nd 24-h urine specimens*

Holmes et al, ***Nature****, 2008*

Dietary and Urinary Metabonomic Factors Possibly Accounting for Higher Blood Pressure of Black Compared With White Americans

Results of International Collaborative Study on Macro-/Micronutrients and Blood Pressure

Hypertension 2013 ;62:1074-80

Jeremiah Stamler, Ian J. Brown, Ivan K.S. Yap, Queenie Chan, Anisha Wijeyesekera, Isabel Garcia-Perez, Marc Chadeau-Hyam, Timothy M.D. Ebbels, Maria De Iorio, Joram Posma, Martha L. Daviglus, Mercedes Carnethon, Elaine Holmes, Jeremy K. Nicholson, Paul Elliott, for the INTERMAP Research Group

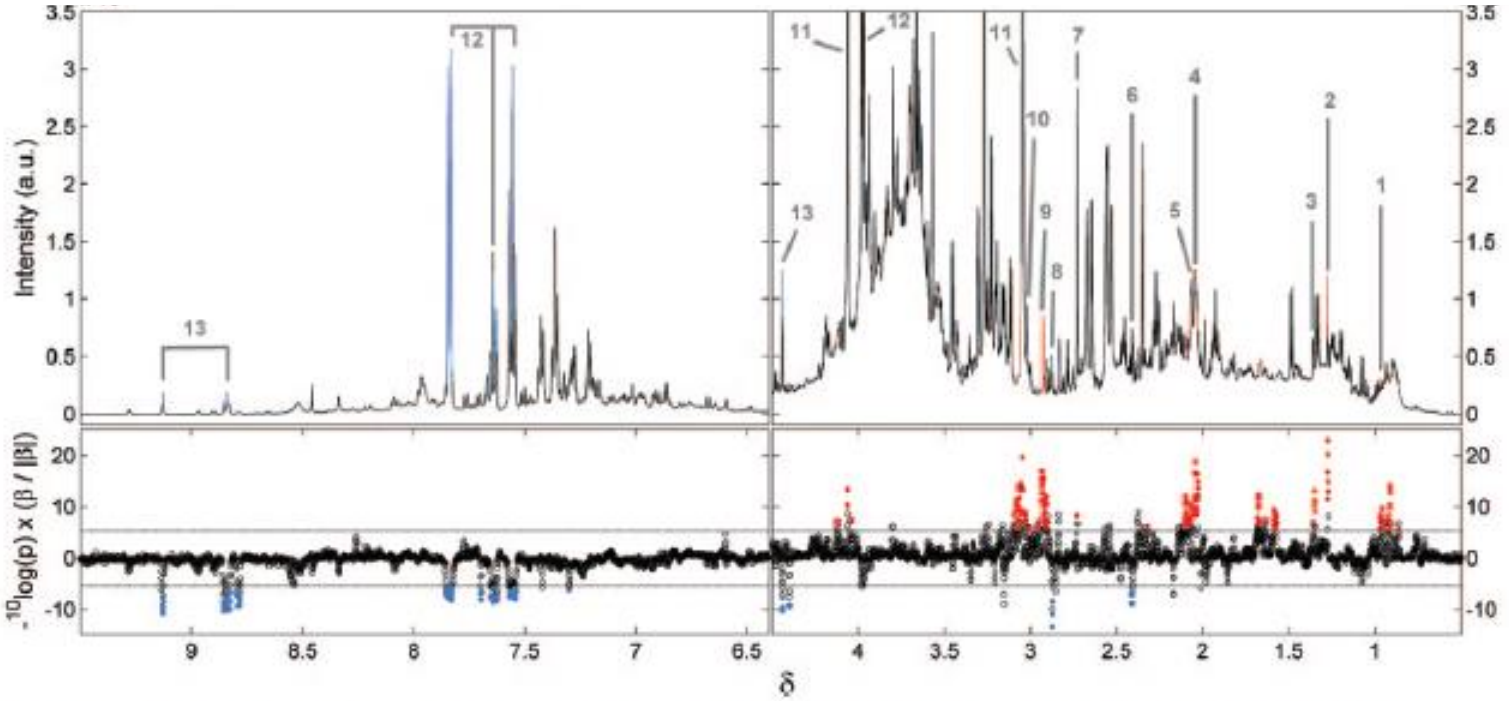


Figure. Median urinary proton nuclear magnetic resonance spectrum of INTERMAP US black and non-Hispanic white American (NHWA) participants (**top**), based on the first urine collection (n=1455). Manhattan plot indicating the significant spectral variables (**bottom**). Metabolites higher in black individuals compared with NHWA are shown in red; metabolites higher in NHWA individuals compared with blacks are shown in blue. 1 indicates leucine; 2, 3-hydroxyisovalerate; 3, 2-hydroxyisobutyrate; 4, *N*-acetyls of glycoprotein fragments; 5, *N*-acetyl neuraminic acid; 6, succinate; 7, dimethylamine; 8, trimethylamine; 9, dimethylglycine; 10, lysine; 11, creatinine; 12, hippurate; and 13, *N*-methyl nicotinic acid.

Exposome Studies – Avonmouth Zn Smelter

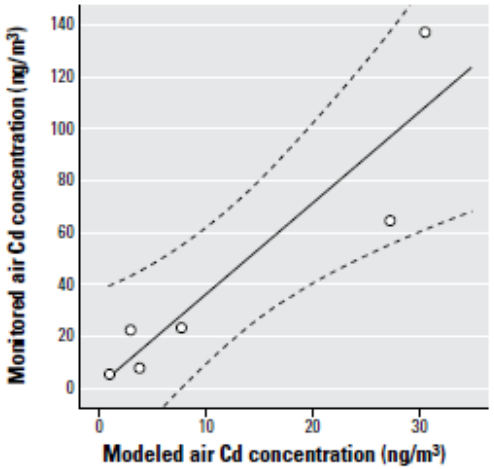
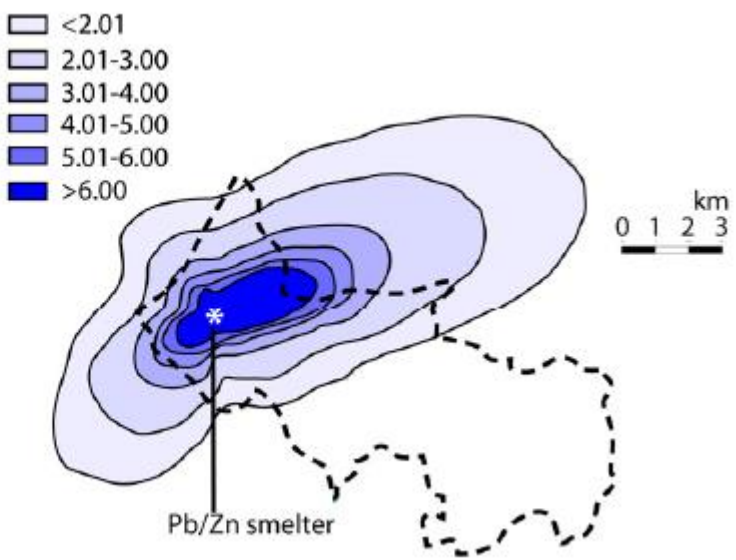
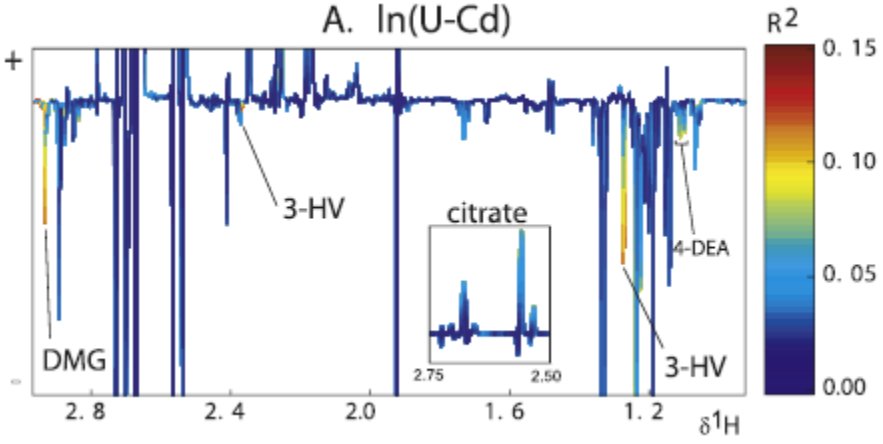


Figure 2. Monitored Cd levels as a function of modeled data (note different scales). Lines indicate linear regression and 95% CI (R^2 = 0.84).





3-HV, 3-hydroxyisovalerate; DMG, dimethylglycine; 4-DEA, 4-deoxyerythronic acid.

Mitochondrial & 1 C metabolism

Thomas et al *EHP* 2009; Ellis et al BMC Med 2012

EXPOsOMICS: Mission

- Exposomics aims to predict **individual disease risk** related to the environment, by characterizing the **external and internal exposome** during critical periods of life, including in utero, based on established European cohorts

  - Focus on air and water pollution
  - Focus on the external exposome in adults

Experimental Short Term Studies (STS)

New measurements

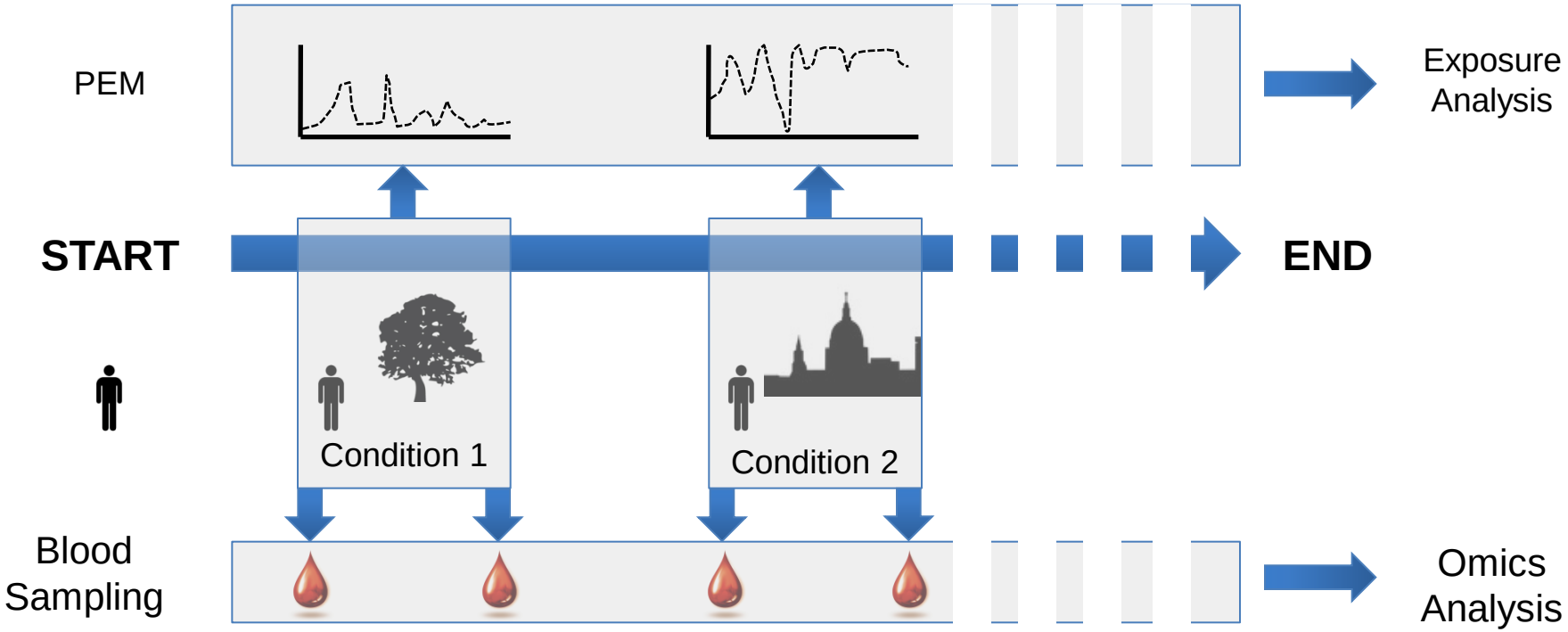
- Oxford Street 2
- TAPAS

Existing resources

- Oxford Street 1
- RAPTES

Outline

- Contrasting levels of AP (high/low) during exposure periods
- PEM measurements made during exposure periods
- Blood samples for omics before and after each exposure period



Mother-Child Cohort Studies and Adult Long-Term Studies

New measurements

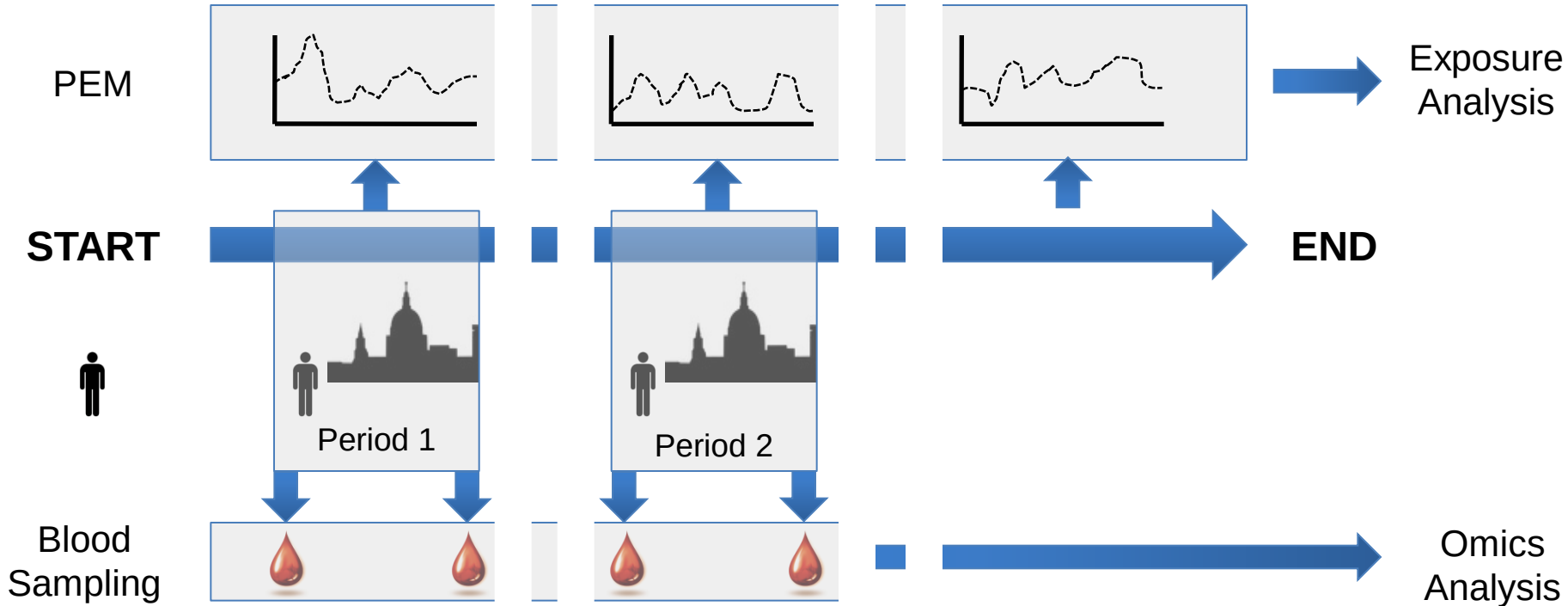
- INMA
- EPIC-ESCAPE and East Anglia
- SAPALDIA

Existing resources

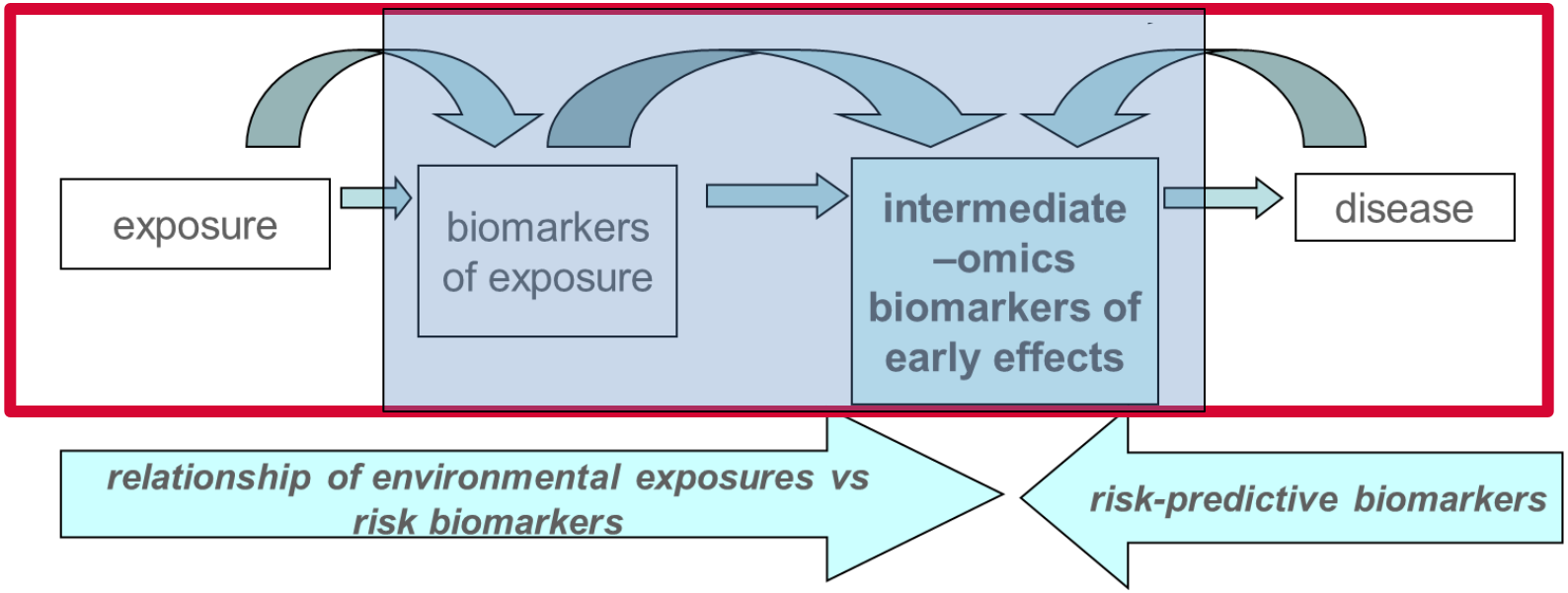
- INMA
- EPIC-ESCAPE and East Anglia
- Rhea
- ALSPAC
- Piccoli+

Outline

- Participants in existing cohort studies with stored blood samples
- PEM measurements will be made during three periods at five sites



Meet-in-the-middle Research Paradigm



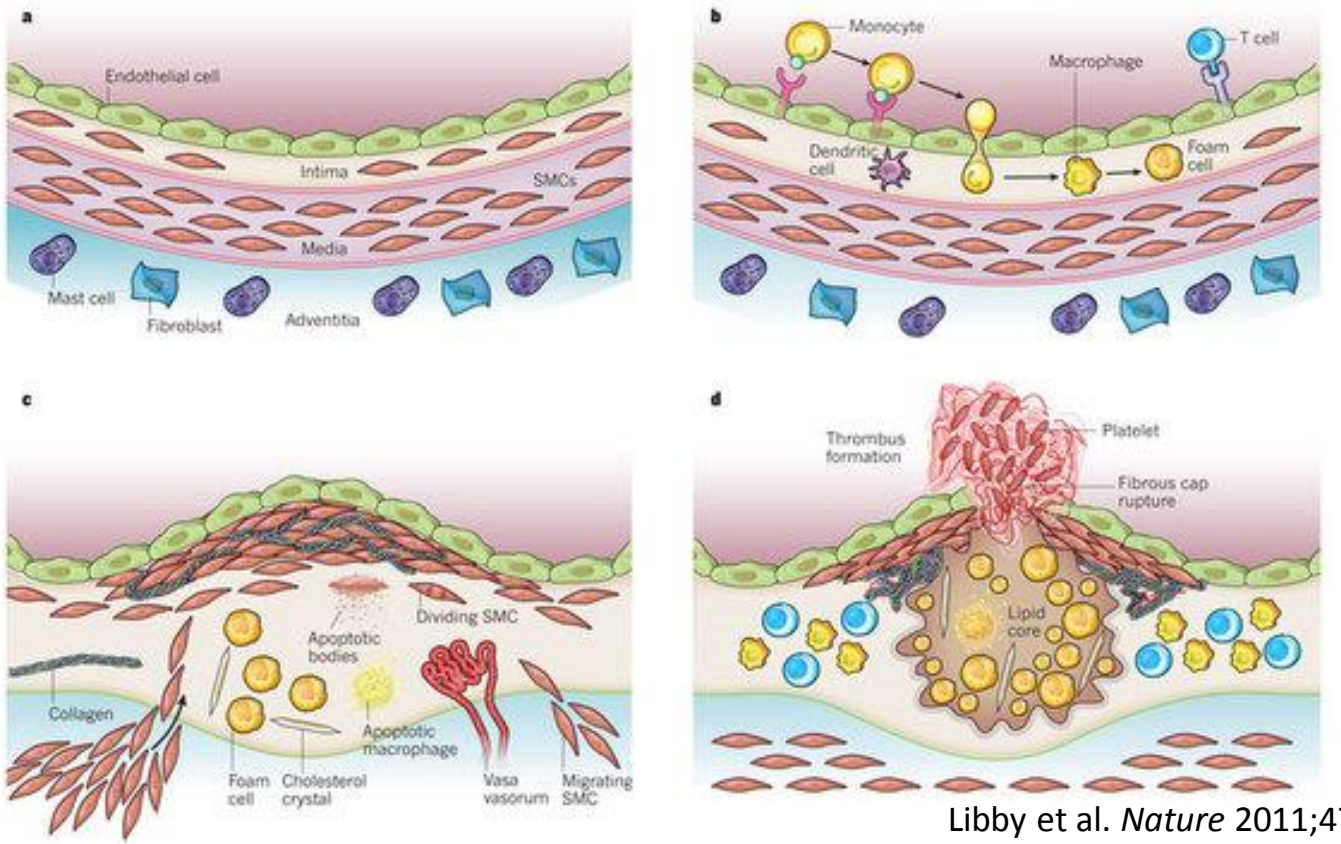
Chadeau-Hyam et al *Biomarkers* 2011; 16(1): 83–88

COMBI-BIO: Specific Aims

- To use systems biology approaches, based on metabolic phenotyping and computational medicine, to discover, test and validate novel biomarkers for subclinical atherosclerosis;
- To use cross-platform (NMR, MS) and multi-omics (genome-wide, metabolome-wide) analyses to investigate underlying biochemical pathways, and hence to advance understanding of aetiopathogenesis of atherosclerosis development and progression;
- To develop prognostic combinatorial biomarkers and risk scores to improve early prediction and patient stratification/management for subclinical atherosclerosis.

Atherosclerosis

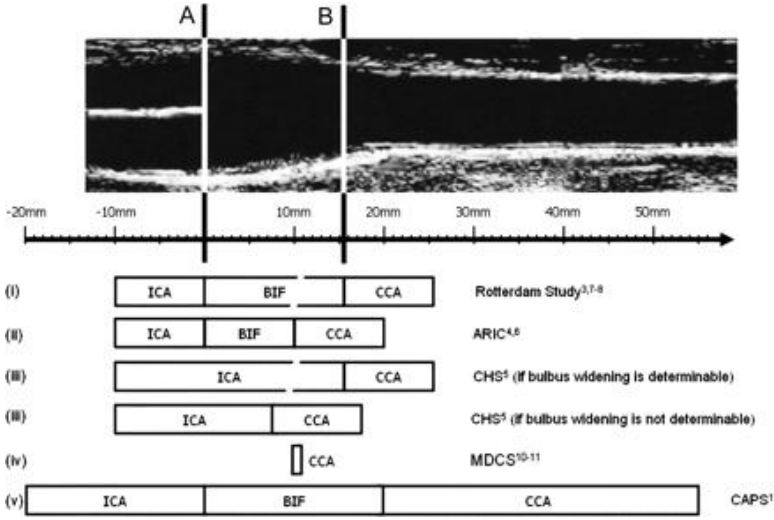
- Systemic chronic disease of the arterial wall
- Remains subclinical for many decades
- Main underlying cause of MI
- Pathophysiology of atherosclerotic lesion formation and of complications still poorly understood



Libby et al. *Nature* 2011;473, 317–325

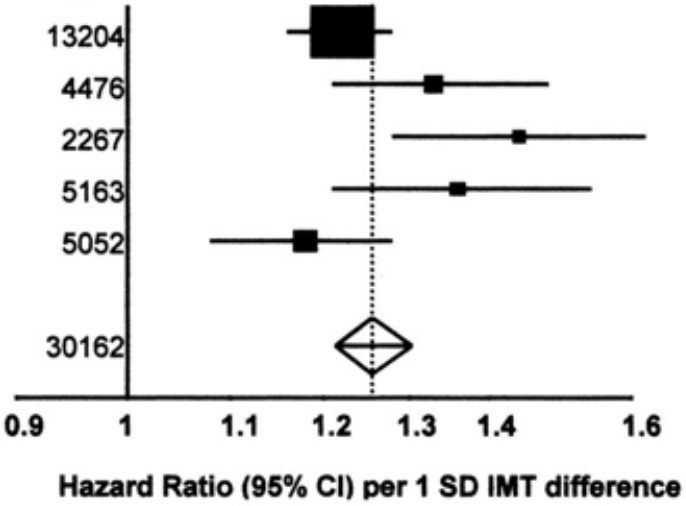
Measurement of atherosclerosis

- Intima-media thickness of carotid arteries
- Established measurement of early disease



A Hazard ratio (HR) for MI per 1 SD difference in CCA-IMT, adjusted for age and sex

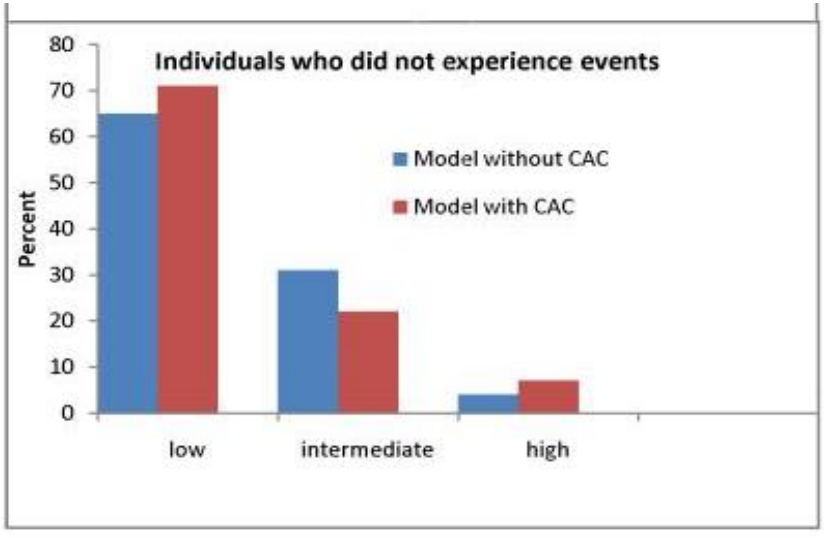
Study	HR	[95% CI]	n
Atherosclerosis Risk in Communities Study (ARIC)	1.22	[1.16-1.28]	13204
Cardiovascular Health Study (CHS)	1.33	[1.21-1.48]	4476
Rotterdam Study	1.44	[1.28-1.62]	2267
Malmö Diet and Cancer Study subcohort (MDCS)	1.36	[1.21-1.54]	5163
Carotid Atherosclerosis Progression Study (CAPS)	1.18	[1.08-1.28]	5052
TOTAL	1.26	[1.21-1.30]	30162
I^2 for heterogeneity	65.2%		

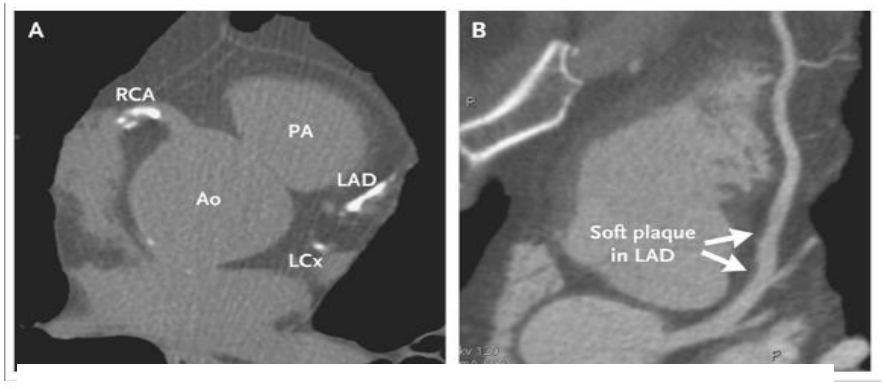


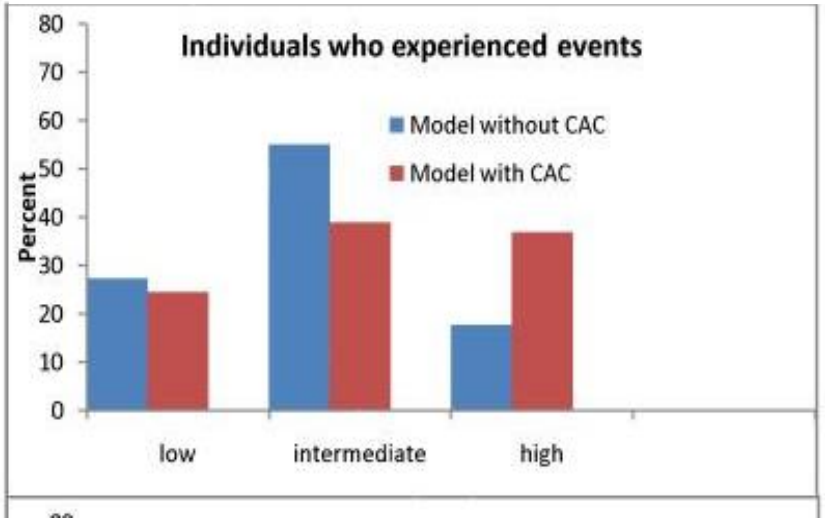
Lorenz et al. *Circulation* 2007, 115 (4) 459-467

Measurement of atherosclerosis

- Coronary artery calcium
- In MESA, addition of CAC to traditional risk factors improved classification of individuals according to risk







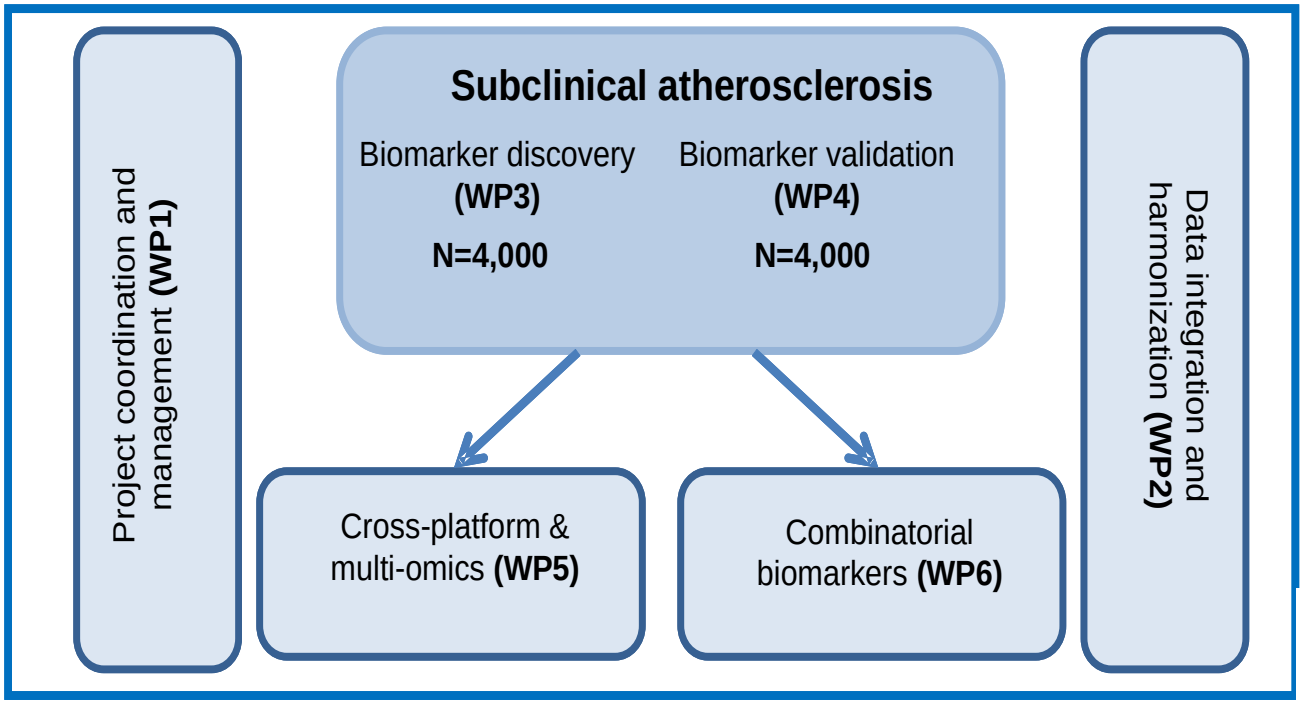
Polonsky *et al. JAMA* 2010; 303(16): 1610–1616.

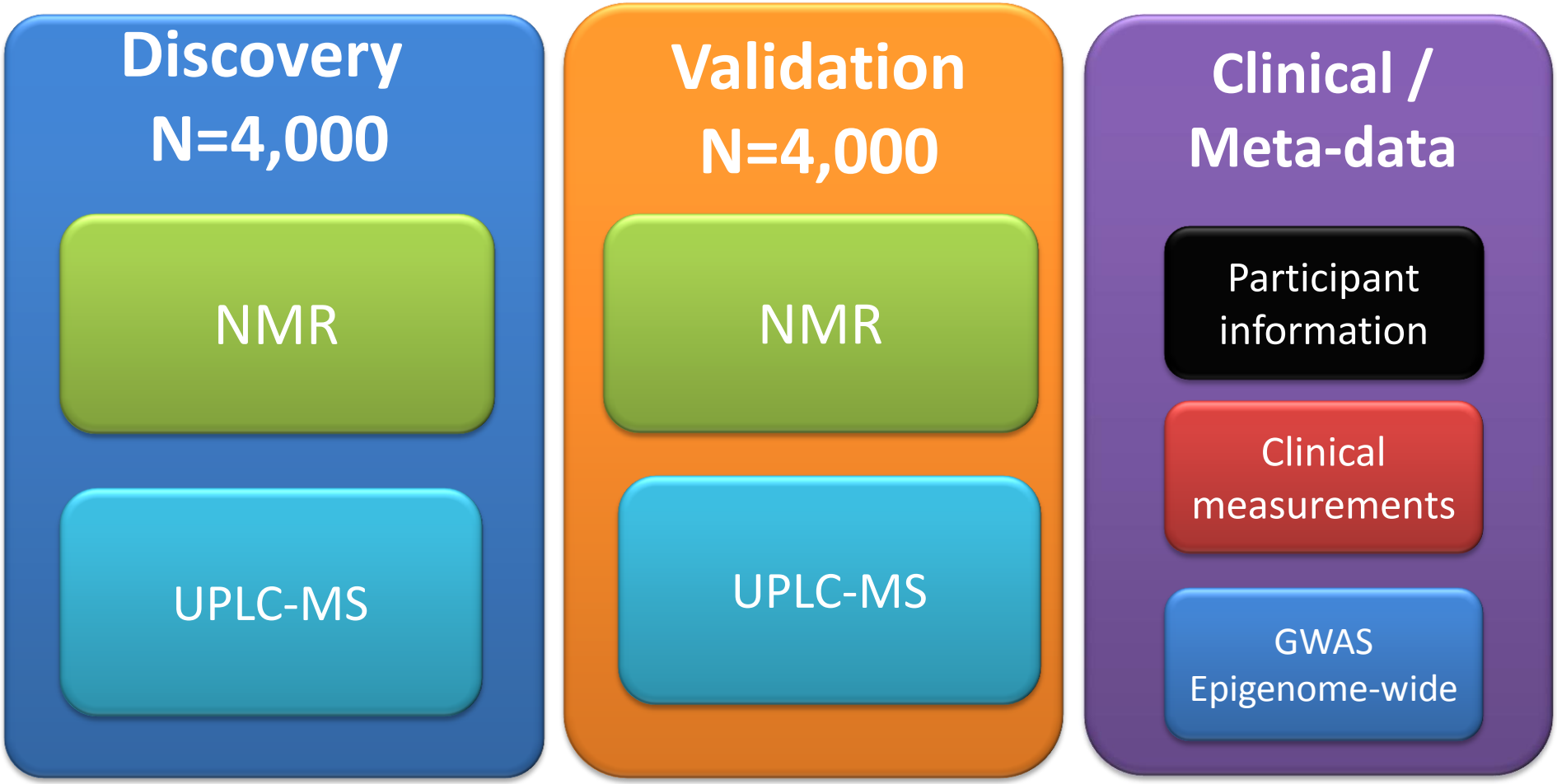
COMBI-BIO Cohorts and Study design

Table 1.1b Cohorts and numbers of participants for discovery and validation of novel metabolic biomarkers for subclinical atherosclerosis* in COMBI-BIO

	Discovery	Validation
West London LOLIPOP cohort	1,000	1,000
Rotterdam Study	1,000	1,000
Multi-Ethnic Study of Atherosclerosis (MESA)	2,000	2,000
TOTAL	***4,000***	***4,000***

*Subclinical atherosclerosis measurements of both coronary artery calcium (CAC) and carotid intima medial thickness (IMT) available for each participant





Metabolic characterization of phenotypes (CAC, IMT) measured within cohorts

MRC-PHE
Centre for Environment & Health

Public Health England | MRC Medical Research Council | Imperial College London | King's College London

Wellcome Trust Sustaining Health Pilot (China)
Study concept

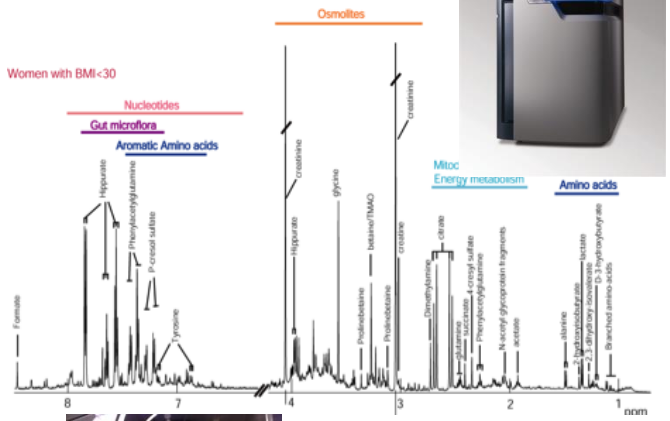
Intervention

Risk factor exposure

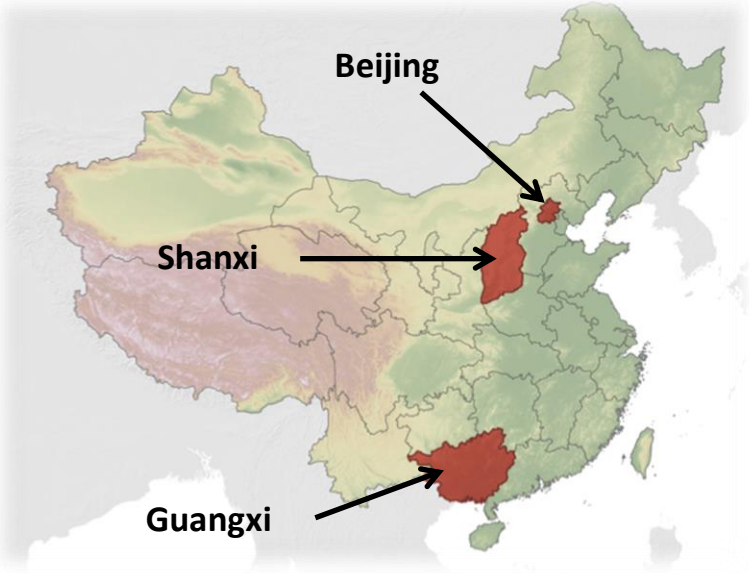
Biomarkers of exposure

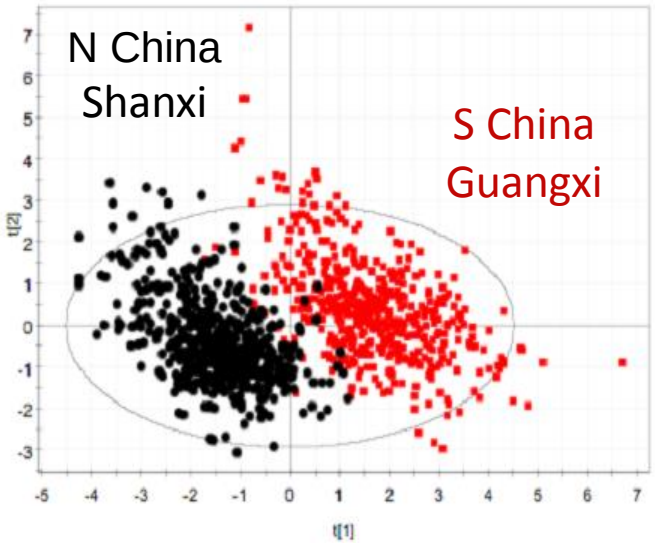
Biomarkers of early effect

Clinical measures



Study platform: INTERMAP China sites and participants





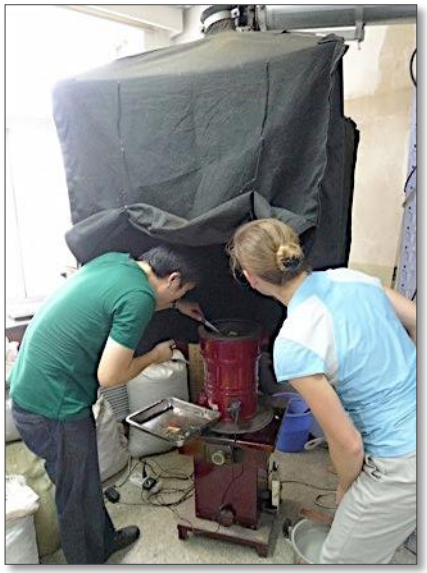
- 839 men and women, 40-59 years
- Multiple visits and measurements of diet, alcohol, weight, height, blood pressure, and socio-demographic variables
- Two 24-h urine collections

	Mean (SD)	
Trait	**N China (N=523)**	**S China (N=244)**
SBP mm Hg	123.8 (18.6)	115.4 (13.0)
Ur Na mmol/24h	271.4 (88.3)	139.2 (55.5)
Ur Na/K ratio	7.8 (2.4)	3.7 (1.5)
Ca mg/1000 kcal	136.5 (48.4)	175.0 (62.5)
Mg mg/1000 kcal	133.2 (38.7)	198.2 (27.2)

Yap et al. *J Proteome Res* 2010;9(12):6647-54

Stove modification, testing, and adaptation to household use

Technology development



Laboratory emissions testing

Testing in the rural energy lab

Pilot at village demonstration sites

Summary

- Major chronic diseases (e.g., heart disease, metabolic disease) at epidemic proportions worldwide and largely environmental in origin
- MWAS via NMR, MS and other high-throughput omics technologies applied to large-scale epidemiological studies offer new opportunity for discovery on the causes and mechanisms of disease
- Dedicated high throughput facilities are required to carry out this work (like the MRC-NIHR National Phenome Centre at Imperial/King's College)

Thanks to.....

Toby Athersuch

Jill Baumgartner

Majid Ezzati

Joanna Tzoulaki

Paolo Vineis