Modelling the MITM principle with metabolomics data

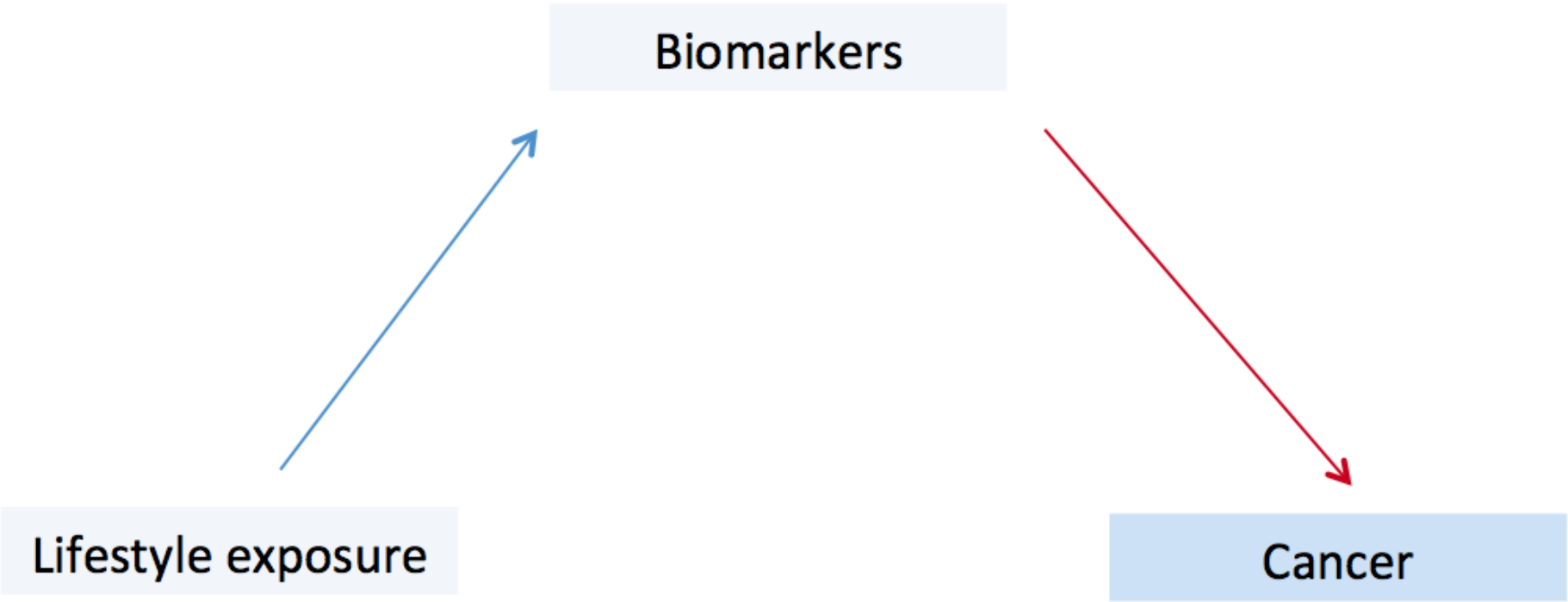
Pietro Ferrari
ferrarip@iarc.fr

International Agency for Research on Cancer

Outline

- The meeting-in-the-middle principle
- Metabolomics as an emerging resource
- Developing analytical frameworks to handle overwhelming information
- An application from the EPIC study

The MITM principle



(Vineis & Perera, *CEBP*, 2007)

Metabolomics

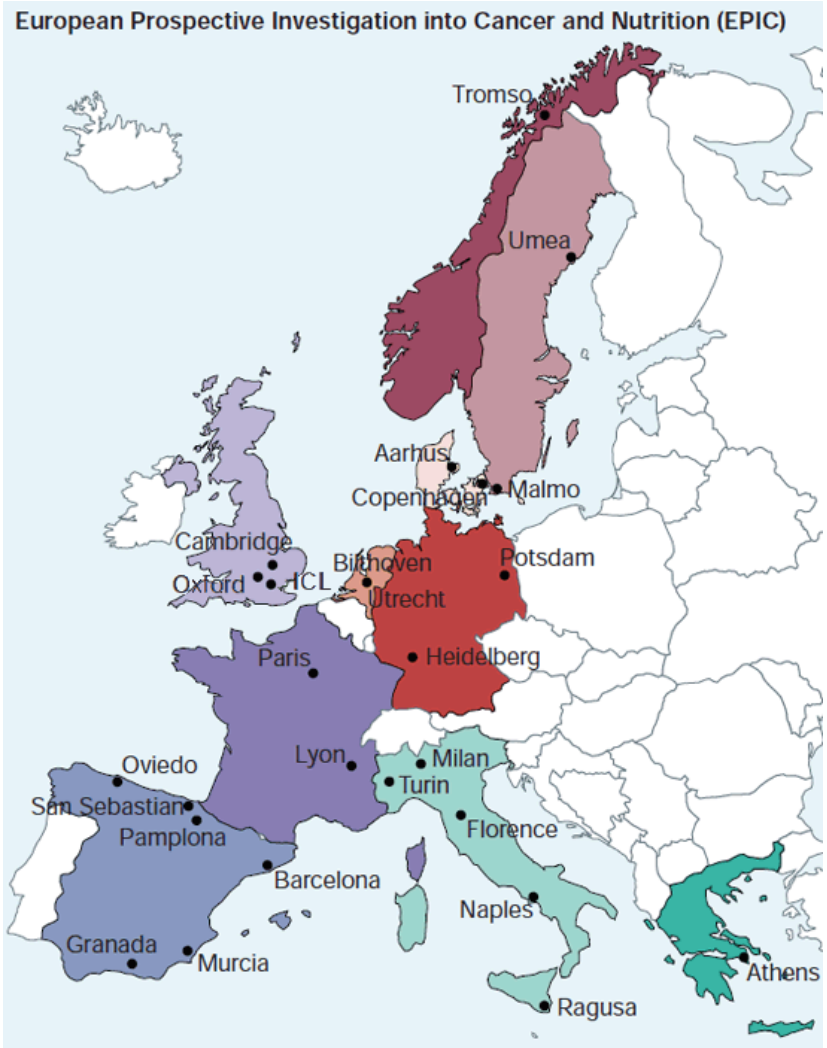
- Comprehensive and quantitative analysis of wide arrays of metabolites in biological samples

- Powerful tool for potential identification of causal pathways in the development of (morbid) conditions

- Among other expertise, need of statistical methodologies to fully exploit this potential

The EPIC Study

- Prospective cohort with 500,000 participants from 23 centres
  - Dietary and lifestyle exposures assessed at baseline
  - Biological samples collected at baseline from 80% disease-free participants



International Agency for Research on Cancer

World Health Organization

EPIC study on HCC

- Nested case-control study on hepatocellular carcinoma (HCC)

- 147 cases and 147 controls matched on

  age, fasting status, date and time of the day at blood collection, sex, study centre; in women also on menopausal status and HRT use

Lifestyle exposure

- Components of the Healthy Lifestyle Index (HLI): Diet[1], BMI, physical activity (PA), alcohol consumption at recruitment, smoking, and Hepatitis status, Type-II diabetes

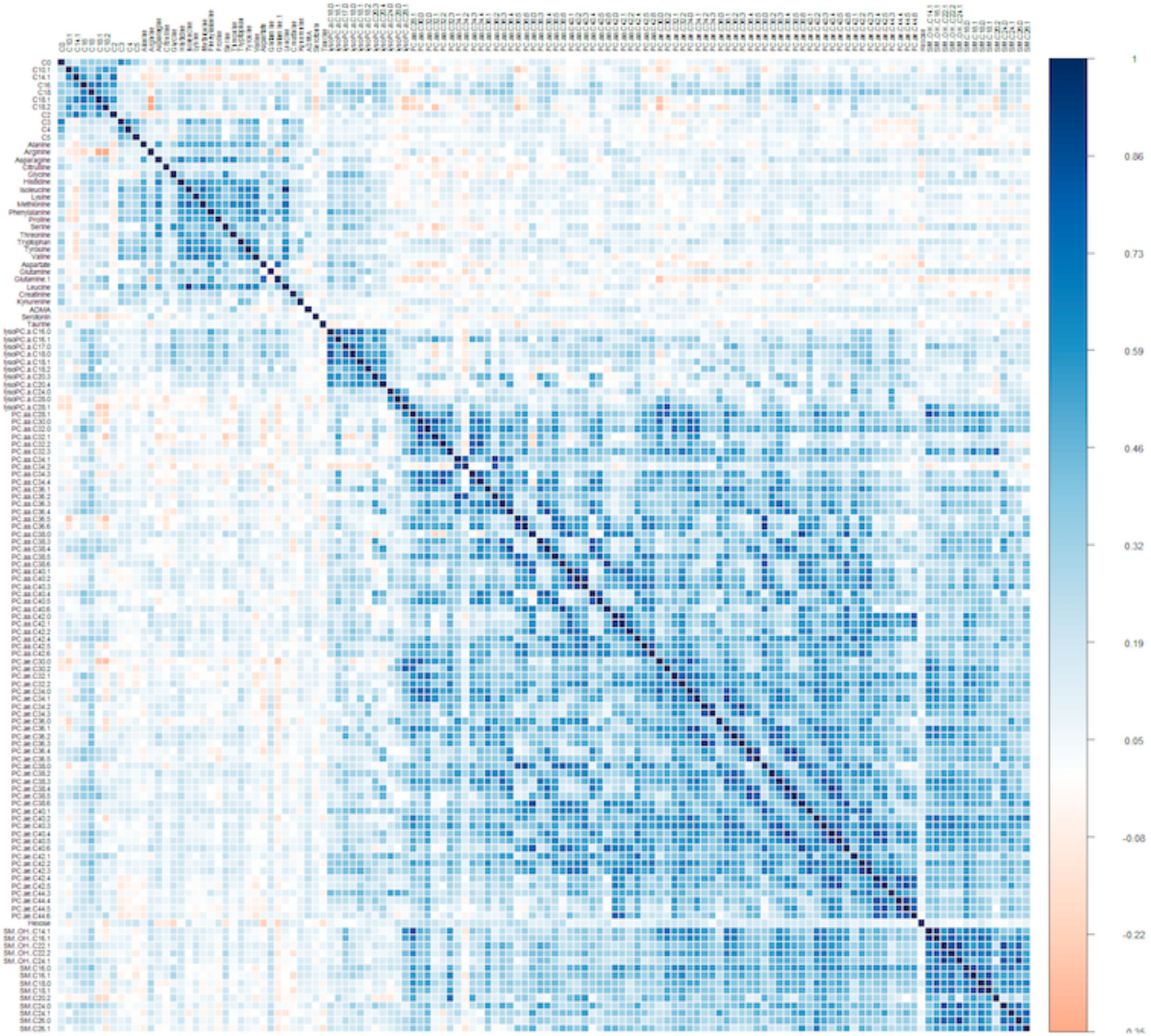
[1] *A priori* score using cereal fibre, folate, PUF/SF, fatty fish, margarine, glycemic load and fruits & vegetables

Metabolomics exposure

- 132 blood metabolites acquired in cancer-free study participants

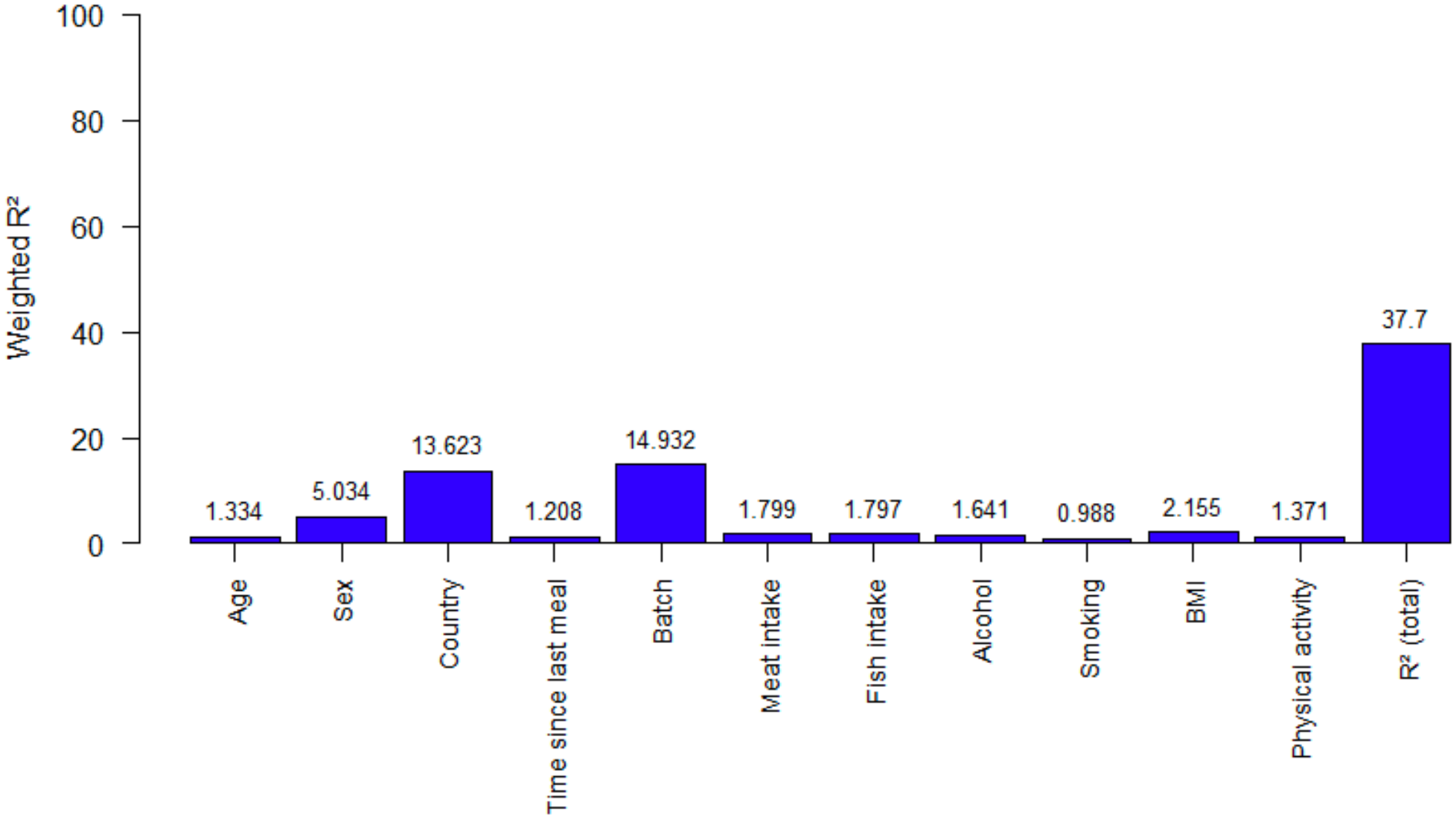
- A targeted metabolite profiling approach (Biocrates AbsoluteIDQ-p180 Kit) using UPLC coupled to a Q-Trap mass spectrometer

Biocrates metabolites



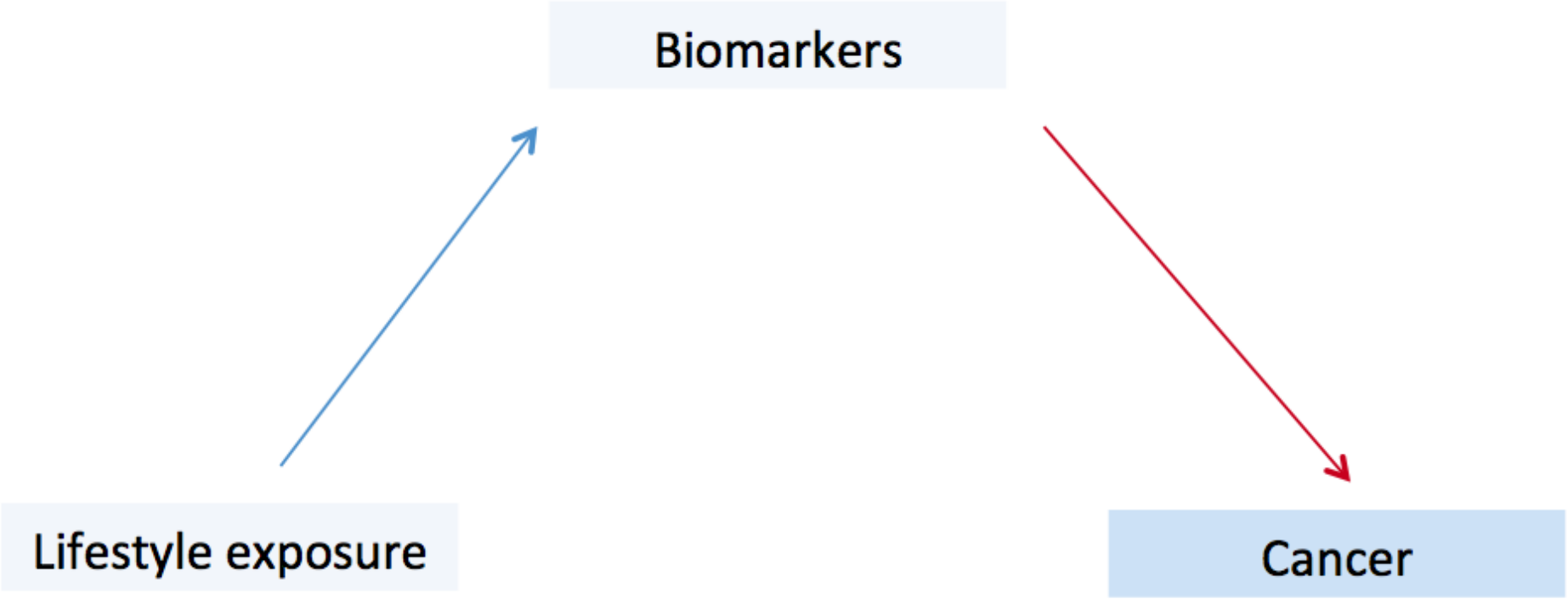
International Agency
World Health Organization

The PC-PR2 analysis



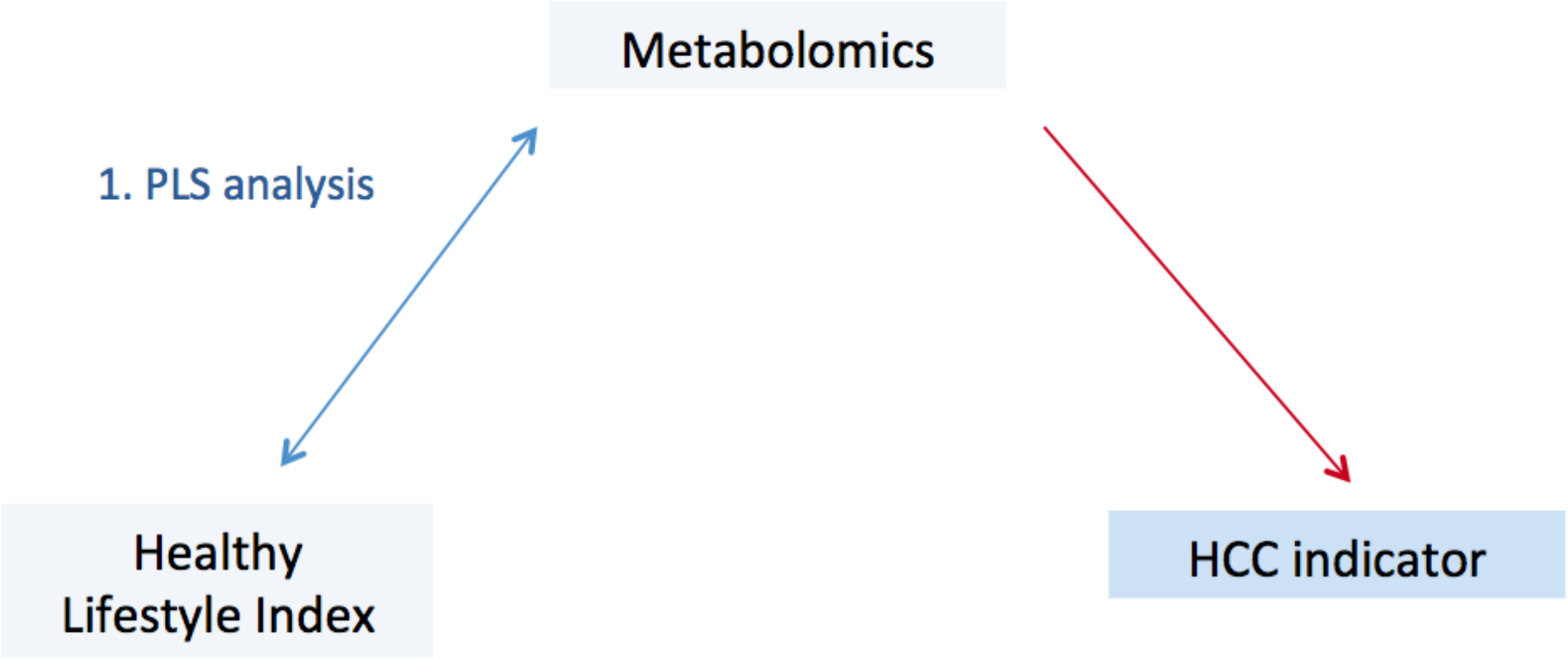
(Fages & Ferrari *et al.*, Metabolomics, 2014)

The MITM principle

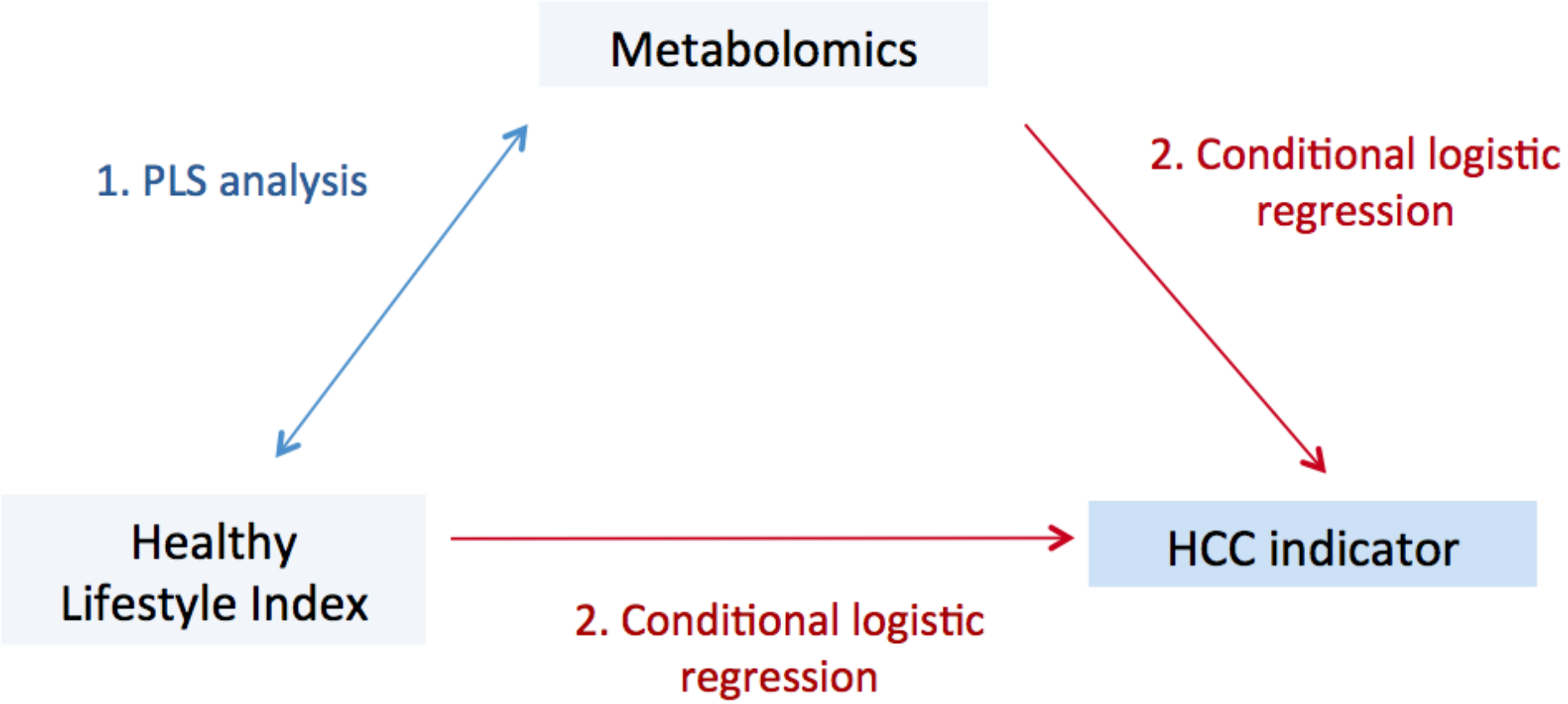


(Vineis & Perera, *CEBP*, 2007)

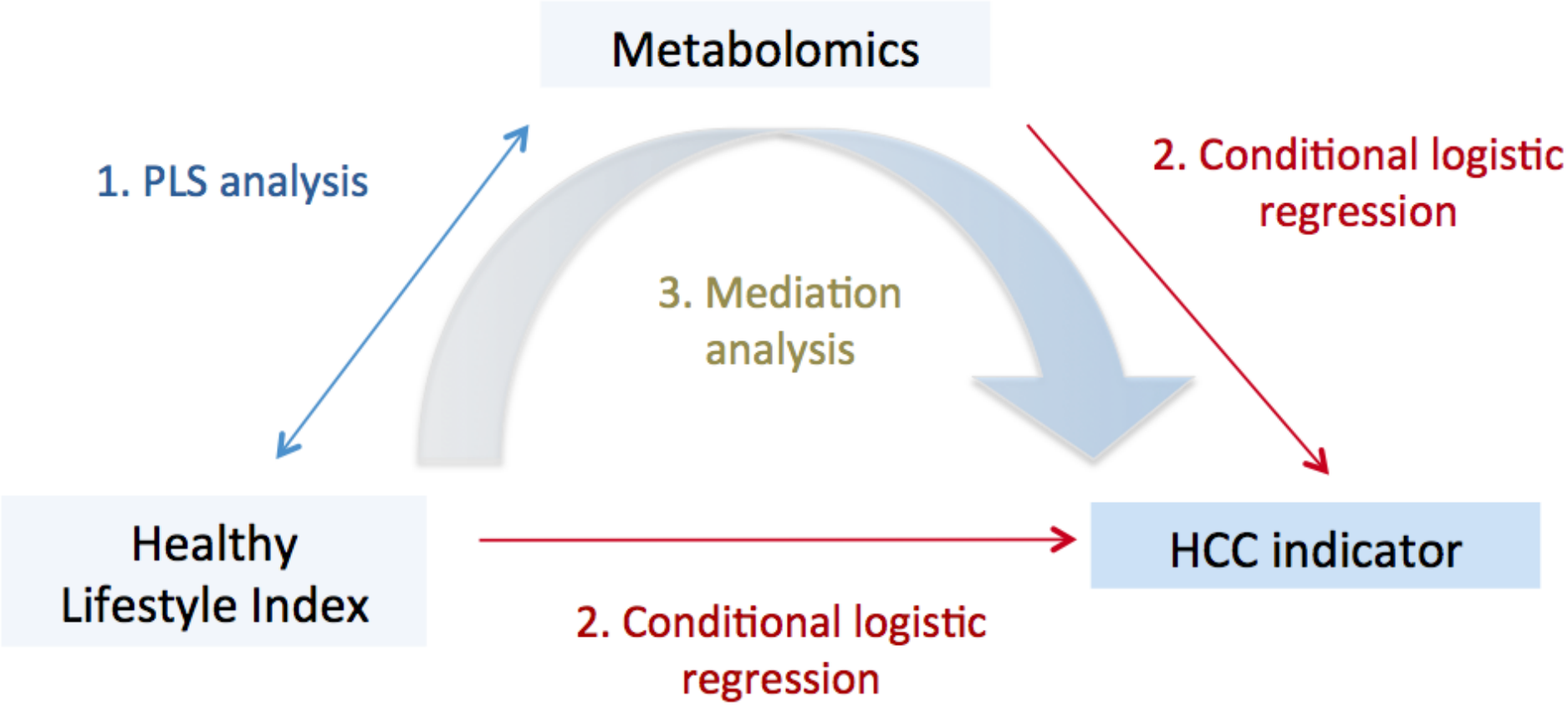
The MITM in EPIC



The MITM in EPIC



The MITM in EPIC (1.)



(Assi *et al.*, *Mutagenesis*, 2015)

International Agency for Research on Cancer
World Health Organization

1. The PLS factors

HLI components	Loadings	Metabolites	Loadings
BMI (kg/m^2)	-0.22	SM(OH) C14:1	0.22
Diet score	0.03	SM(OH) C16:1	0.21
PA (MET-h/week)	0.22	SM(OH) C22:2	0.21
Smoking status	-0.48	LysoPC a C16:1	-0.19
Alcohol (g/day)	-0.76	PCaa C32:1	-0.25
Hepatitis status	-0.27	PCae C30:2	0.18
Type-II diabetes	-0.10	PCae C36:2	0.18

1. The PLS factors (ii)

HLI components	Loadings	Metabolites	Loadings
BMI (kg/m^2)	-0.22	SM(OH) C14:1	0.22
Diet score	0.03	SM(OH) C16:1	0.21
PA (MET-h/week)	0.22	SM(OH) C22:2	0.21
Smoking status	-0.48	LysoPC a C16:1	-0.19
Alcohol (g/day)	-0.76	PCaa C32:1	-0.25
Hepatitis status	-0.27	PCae C30:2	0.18
Type-II diabetes	-0.10	PCae C36:2	0.18

1. HLI and HCC in EPIC

Exposure	OR[1] (95%CI)	p_{value}
HLI	0.89 (0.79, 0.99)	0.047
HLI, PLS score	0.82 (0.57, 1.17)	0.272
Metabolites, PLS score	0.79 (0.69, 0.90)	0.00004

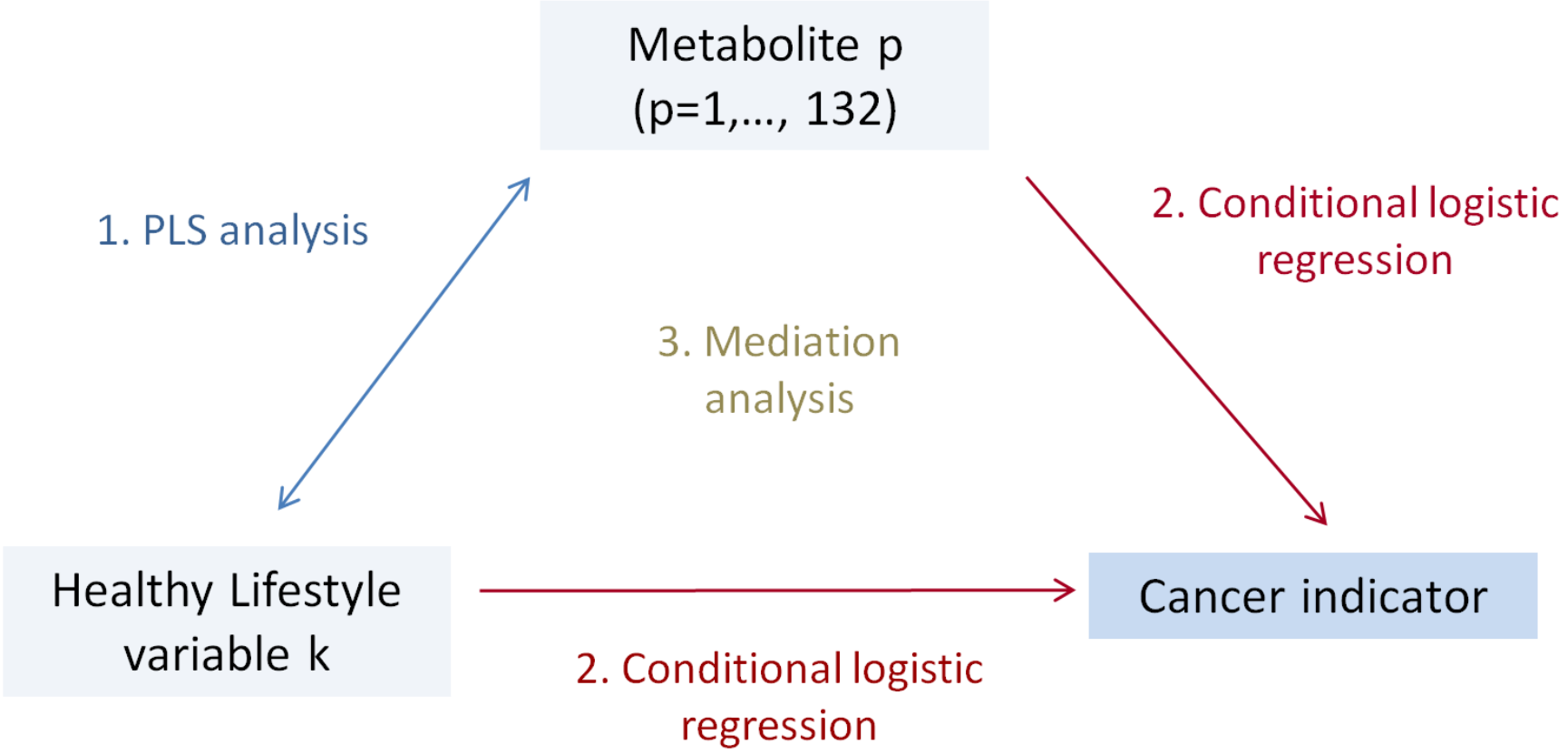
[1] Expressing HCC risk associated to 1-SD increase exposure; models adjusted for C-reactive protein and a composite score for liver damage.

1. HLI and HCC in EPIC

Exposure	OR[1] (95%CI)	p_{value}
HLI	0.89 (0.79, 0.99)	0.047
HLI, PLS score	0.82 (0.57, 1.17)	0.272
Metabolites, PLS score	0.79 (0.69, 0.90)	0.00004

[1] Expressing HCC risk associated to 1-SD increase exposure; models adjusted for C-reactive protein and a composite score for liver damage.

The MITM in EPIC (2.)



Separately for each k=1,...,7

2. The PLS factor for BMI

HLI component	Loading	Metabolites	Loadings
		Glutamate	-0.23
		Glutamine	0.19
BMI	1	Tyrosine	-0.25
		LysoPC a C17:0	0.22
		LysoPC a C18:1	0.16
		LysoPC a C18:2	0.24
		PCae C36:2	0.21

2. BMI and HCC

Exposure	OR[1] (95%CI)	p_{value}
BMI	0.97 (0.84, 1.11)	0.634
Metabolites, PLS score	1.15 (1.06, 1.25)	0.0007

[1] Expressing HCC risk associated to 1-SD increase exposure; models adjusted for HLI (*minus* BMI), C-reactive protein and a composite score for liver damage.

2. Alcohol at recruitment and HCC

Exposure	OR[1] (95%CI)	p_{value}
Alcohol	0.97 (0.83, 1.13)	0.673
Metabolites, PLS score	1.10 (1.03, 1.16)	0.002

[1] Expressing HCC risk associated to 1-SD increase exposure; models adjusted for HLI (*minus* alcohol), C-reactive protein and a composite score for liver damage.

Concluding remarks

- A flexible analytical framework for the MITM
  - Key to create multi-disciplinary settings, with a common language

Concluding remarks

- A flexible analytical framework for the MITM
  - Key to create multi-disciplinary settings, with a common language
- But again, clearly biology is way more complex than statistical modeling
  - A good way to render fashionable models useful is to make them sensible

Acknowledgments

- Nada Assi, Micheal Leitzman, Isabelle Romieu, Mazda Jenab (NEP Group, IARC)
- David Achaintre, Augustin Scalbert (BMA Group, IARC)
- Paolo Vineis (ICL, UK) and all EPIC PIs
- Terry for his perseverance