SOCIAL DETERMINANTS AND HEALTHY AGING: LIFEPATH A H2020 PROJECT

Application: a life course approach to explore the biological embedding of socioeconomic position

M. Chadeau-Hyam, P Vineis

March 17, 2016

Imperial College London

Lifepath HEALTHY AGEING FOR ALL

SOCIAL INEQUALITIES IN HEALTH

Educational inequalities in mortality in Europe (EPIC): 371,295 participants across 9 countries

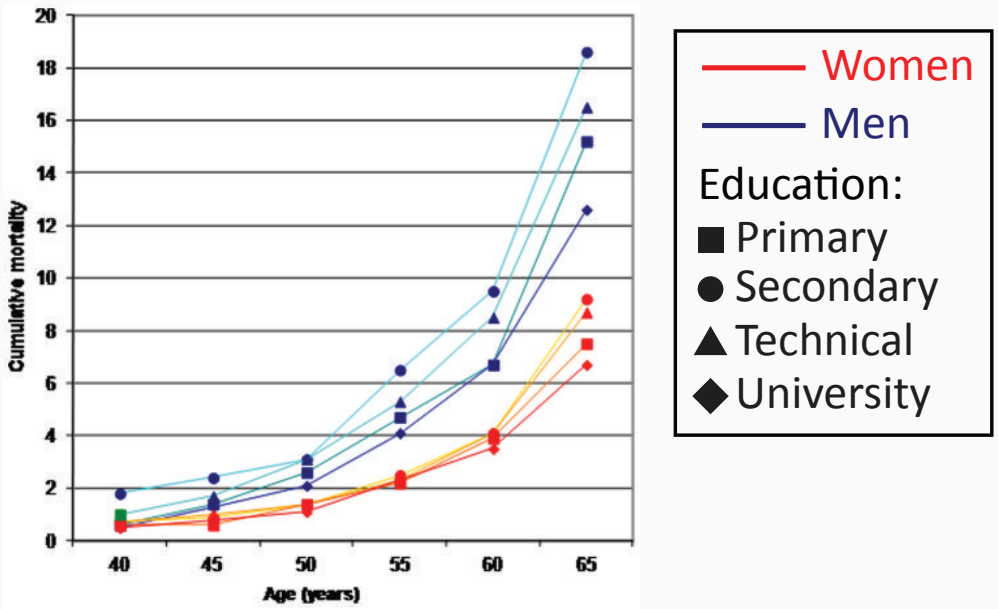


Figure: Cumulative mortality by age classes, education level and sex.

⇒ Higher mortality in men and decreasing mortality with education

Gallo *et al.*, 2012

Psychosocial factors and social inequalities

	Education (in years)				Income (in dollars)			
Variable	<12	12–15	>15	*p*-value	<10K	10K–29K	30K+	*p*-value
Mean, marital stress scale	.29	.17	–.03	<.05	.31	.22	.06	.05
Mean, parental stress scale	.18	.10	–.05	<.05	.11	.15	.03	.12
Mean, financial stress scale	.61	.15	–.04	<.001	1.12	.35	–.24	<.001
Mean number of events past 3 years	1.09	.92	.87	<.05	1.11	1.02	.81	<.001
Mean number of lifetime events	.76	.53	.33	<.001	.61	.57	.44	<.001
Ever widowed (%)	2.5	1.3	.8	.16	5.3	1.0	.5	<.001
Ever divorced (%)	31.6	28.4	13.3	<.001	25.0	28.5	22.1	.01
Ever had child die (%)	11.3	5.0	1.4	<.001	6.1	5.6	4.2	.28
Ever assaulted (%)	31.1	18.8	17.1	<.001	25.0	22.0	17.0	.005

Figure: Stress - Life Event Variables by Socioeconomic Indicators

→ Individuals with lower education levels and lower income experience more and stronger of stressfull events

Lantz *et al.*, 2005

Main health behaviours and social inequalities: low *vs.* high (ref) occupation

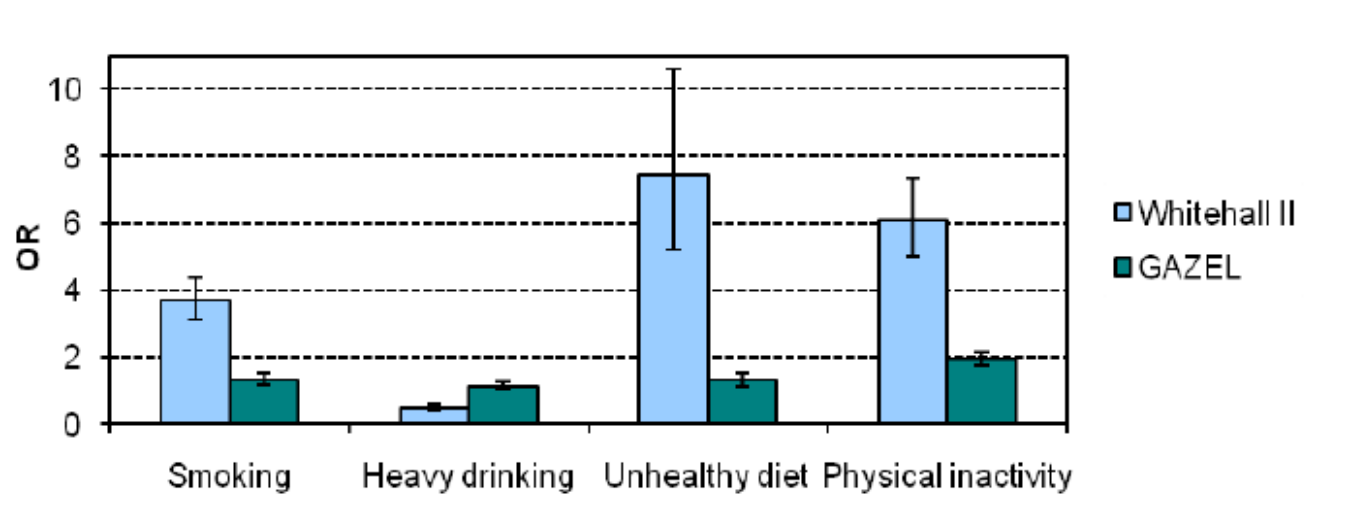


Figure: Occupational position and health behaviours in the British Whitehall II (N=9,771) and the French GAZEL (N=17,760 at first) cohort studies.

→ Country-specific socioeconomic gradient in smoking, unhealthy diet, and physical inactivity

Stringhini *et al.*, 2011

Socioeconomic position and inflammation: CRP in a US population-based sample, NHANES IV

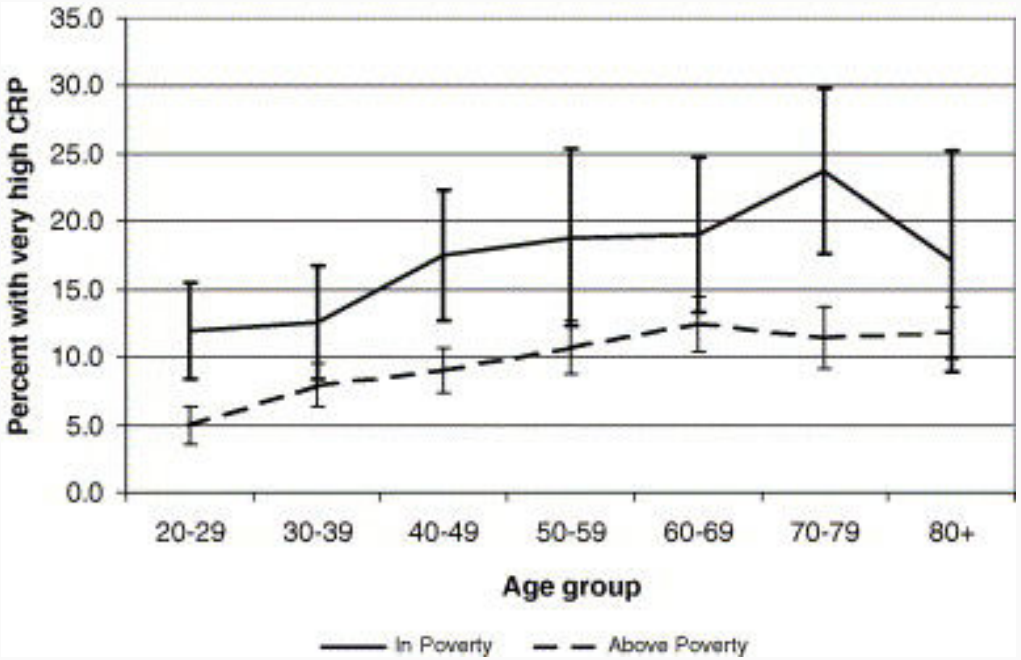
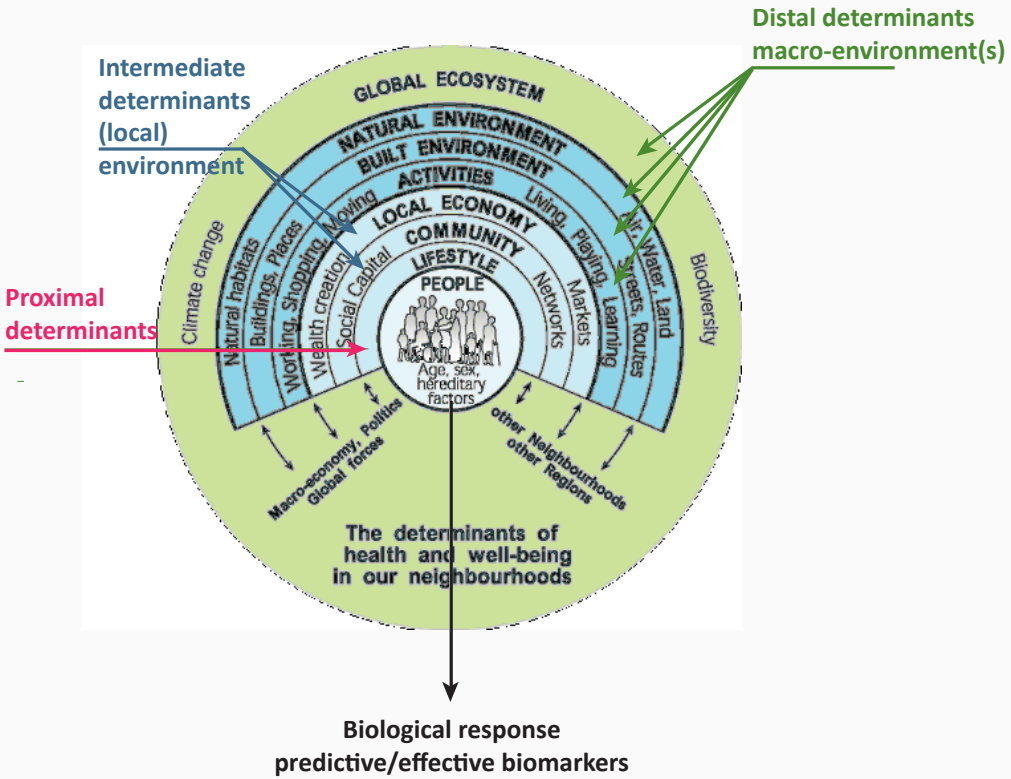


Figure: Prevalence of very highCRP (>10 mg/L) CRP by age group and poverty status (N = 7634).

→ Strikingly higher prevalence of (clinically indicative of infection) levels of CRP in deprived populations.

Alley *et al.*, 2006

Typology of the health determinants: 3 main classes



⇒ within each class socio-economic factors may play a role

⇒ need to investigate molecular markers of SEP experiences and their health consequences

Barton and Grant, 2006

The Strachan-Sheikh Model: build-up and decline stages

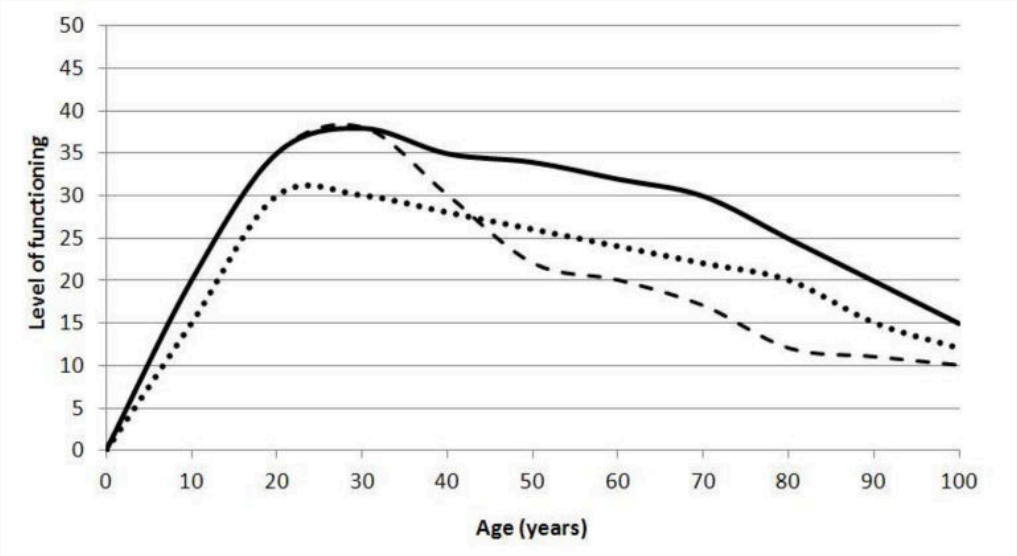


Figure: Life course representation of growth and decline of **levels of functioning**.

→Adverse socio-economic experience in early life can affect the mode of the build-up phase (dotted line)

→ Adverse socio-economic experience later in life can affect the decline rate (dashed line)

Strachan-Sheikh, 2004; Blane *et al.*, 2013

The build-up/decline model:

- **The 2 main stages:**
  1. **Build-up**: from conception and early intra-uterine life to late adolescence or early twenties, characterised by rapid successions of developmentally and socially sensitive periods (potentiation)
  2. **Decline**: starting in early adulthood, is a period of 'decline' from maximum attained health towards loss of function, overt disease and death
- Build-up stage strongly determines subsequent ageing trajectories as it influences the maximum attained level of health
- SE exposures can affect the potentiation and the decline rate
- Under this model healthy ageing can be achieved by:
  1. Maximising the build-up phase: preventing adverse (effects of) early life exposures
  2. Slowing down the decline phase: preventing adverse (effects of) later exposures

→ Need to identify these SE exposures and understand their **drivers and effects**

LIFEPATH PROJECT: HEALTHY AGEING FOR ALL

Specific and orignial objectives of LIFEPATH

1. Demonstrate that healthy ageing is strongly uneven in society, due to multiple environmental, behavioural and social circumstances that affect individuals' life trajectories
2. Improve the understanding of the mechanisms through which healthy ageing pathways diverge by social circumstances, by investigating **life course biological pathways using omic technologies**
3. Provide evidence on the reversibility of the poorer ageing trajectories experienced by individuals exposed to the strongest adversities, by using an experimental approach; and to analyse the health consequences of the current economic recession in Europe
4. Provide updated, relevant and innovative evidence for underpinning future policies

---

http://www.lifepathproject.eu/

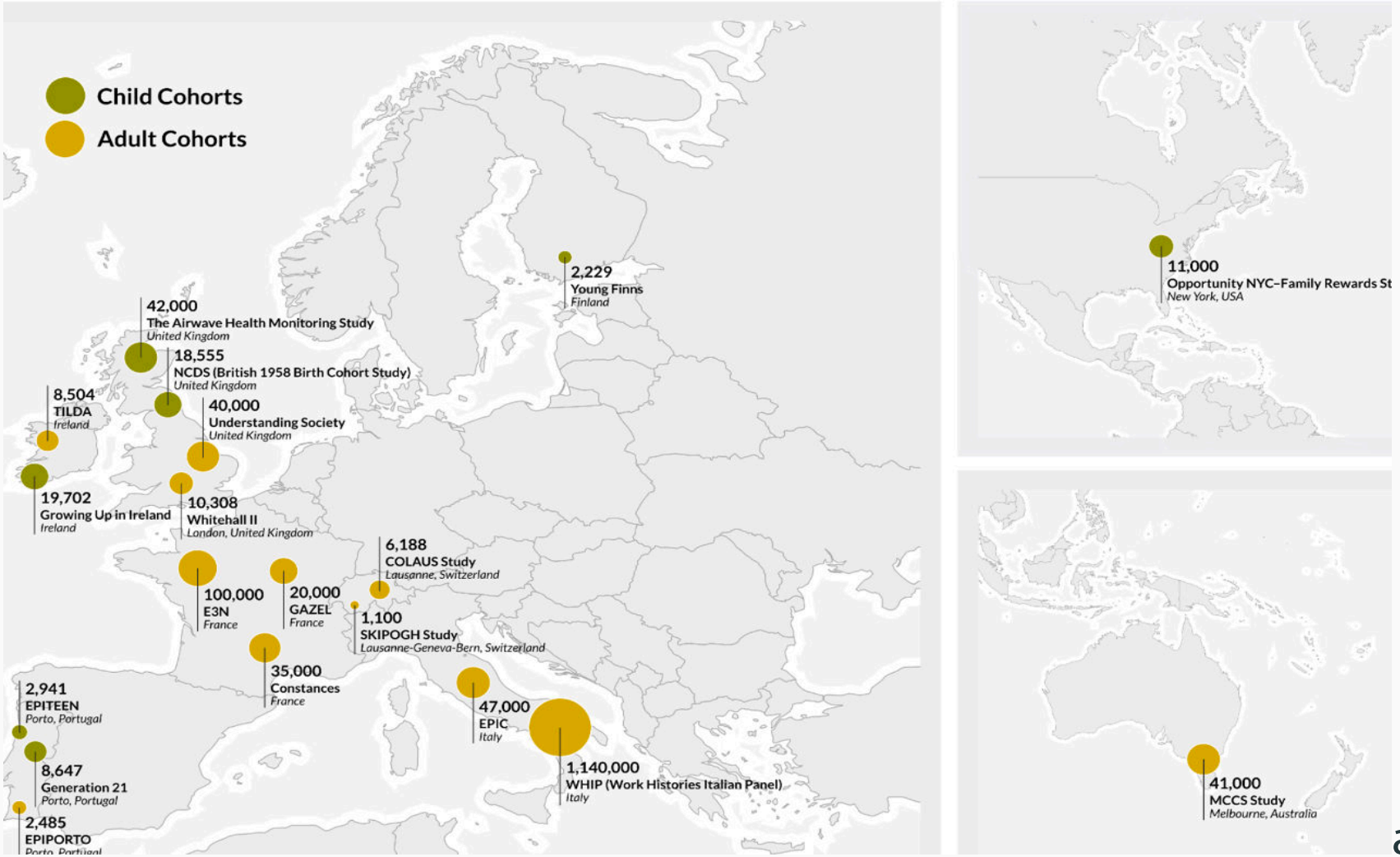
Data sources

- Europe-wide and national surveys (updated up to 2010)
- Longitudinal cohorts (across Europe) with deep phenotyping and repeat biological samples (total population >33,000)
- Other large cohorts with bio-samples (total population >202,000 and a large registry dataset with over a million individuals with very rich information on work trajectories and health)
- A randomized experiment on conditional cash transfer for poverty reduction in New York City

---

http://www.lifepathproject.eu/

Data sources



alert—

LIFEPATH covers numerous regions, age ranges, and exposures

`http://www.lifepathproject.eu/`

OMICs Markers already measured or whose measurement is funded/on-going

http://www.lifepathproject.eu/

Partners

- Imperial College London (Lead)
- INSERM & Université Paul Sabatier, Toulouse
- Lausanne University Hospital
- HuGeF, Torino

Four main Objectives

1. Mediation analysis of SES, risk factors and health outcomes
2. Defining a statistical suite to investigate omic signatures of SES factors
3. Mechanistic models integrating omic data from different platforms
4. Longitudinal models for healthy ageing

Aim: to devise/develop and apply a novel analytical strategies to

- gain knowledge on structures governing SES - risk factors - health outcomes relationships
- identify relevant and stable (within and across populations) structures
- identify potential mediators and effect modifiers

→ gain mechanistic/causal knowledge on which and how SES factors mediate their effect

→ generation of prior knowledge to inform subsequent analyses

Methods

- sequentially-adjusted regression approaches
- structural equation models
- Bayesian hierarchical models (including use of DAGs)
- causal inference models (including g-computation)

Main candidate approaches

- Univariate approaches & multiple testing correction
- (Supervised) dimensionality reduction techniques
- Variable selection approaches (penalized regression and Bayesian alternatives)

LIFEPATH specific constraints:

- subtle and complex effect of SES factors
- heterogeneous sets of exposures (conventional risk factors - and complex SES measures, partially redundant)
- possible joined (and non additive) effects of exposures
- generalisability: how findings replicate outside the social context under investigation (interaction)

→ Generalise methods to accommodate multiple and correlated exposures and incorporate structures/interactions

Aim: explore regulatory cascades triggered by SE exposures and affect health

- mechanistic and causal models are restricted to very few drivers
- profiling techniques don't incorporate structures

  → combine both approaches

Ways forward:

- Use variable selection approaches to identify key nodes (as defined as scores of OMICs markers and/or SES and/or risk factors) and run reduced dimension models as in Task 1

  →interpretability and reliability of the 'nodes'?
- Two step procedure: sequential profiling techniques to order OMIC markers WRT their 'proximity' to exposure. Use networks to draw a typology within and across (ordered) classes of markers

  →how to integrate prior information in these models?
- Generalise the BVS paradigm and identify the best causal graphes

Need to incorporate a longitudinal and dynamic component

- SES measures are dynamic: calendar time and age related
- Crucial role of SES trajectories
- Existence of age-driven susceptibility
- Volatility of OMICS signals

  → models will depend on available data

Some candidate approaches

- Estimate a a 'volatility map' (pooling profiles from different cohorts & using repeated samples)
- Cross sectional data: define composite scores and sequential adjustments on time ordered covariates; interaction models
- Longitudinal data: trajectory classification algorithms (warping models), explicit mechanistic modelling (multi-state models)

  → how to integrate a causal component in a longitudinal setting?

2 Pilot studies: Proteomics and Transcriptomics

R. Castagné[1,2], M. Kelly-Irving[2], C. Delpierre[2], P. Vineis[1] & M. Chadeau-Hyam[1]; [1]Imperial College; [2] INSERM, Toulouse

Data: Italian component of the EPIC study

Biological measures

EpiGenomics

- Illumina Infinium Human Methylation 450 BeadChip
- 485,512 Methylation sites
- 1,716 samples

Transcriptomics

- Agilent 44k
- 29,662 probes
- 268 samples

Proteins

- Luminex Multianalyte Profiling
- 28 inflammatory-related proteins

Life course socioeconomic position (SEP)

Childhood SEP

- Father's occupation
- 2 classes: 'Manual' and 'Non-manual'

Young adulthood SEP

- Participant's education
- 2 classes: 'High' (above the minimum legal education level) and 'Low'

Adulthood SEP

- Highest household occupation
- 2 classes: 'Manual' and 'Non-manual'

Model formulation, for individual i:

- Variable of interest: X^i (SEP, 2 classes)
- Predictors: Y^i, Proteins, Gene expression or Methylation level
- Fixed effects: FE^i, age and gender, phase and centre, case-control status
- Random effect variables: u^{A^i} where A^i are nuisance variables (i.e. sample position on the array, ...)
- Full model

Methodology: likelihood ratio test

1. Run the model **with** and **without** the variable of interest (X^i)
2. Compare both models

$\rightarrow$ for each biomarker we obtain a p-value testing the association between the proteins and the SES classes

- **Hypothesis:** consistent positive direction of the association between biomarkers and SEP

Definition

1. Get the denoised data to remove noise variation of different batches
2. Split each biomarker level into quartiles
3. Assign 0 for quartile 1 to 3
4. Assign 1 for quartile 4
5. **Global score**: Sum across biomarkers

Continuous alternative:

- First PC from a principal component analyses based on 'de-noised' biomarker levels

Life-course multivariate linear regression: sequential adjustment on time-ordered SEP-inidcators

- **Model A:** Age, gender, case-control status, phase and center and father job
- **Model B-1:** Model A + education
- **Model B-2:** Model A + highest household's occupation
- **Model C:** Model B-1 + highest household's occupation
- **Model D:** Model C + BMI + Smoking status + Alcohol

Father's occupational position; ref. 'non-manual'

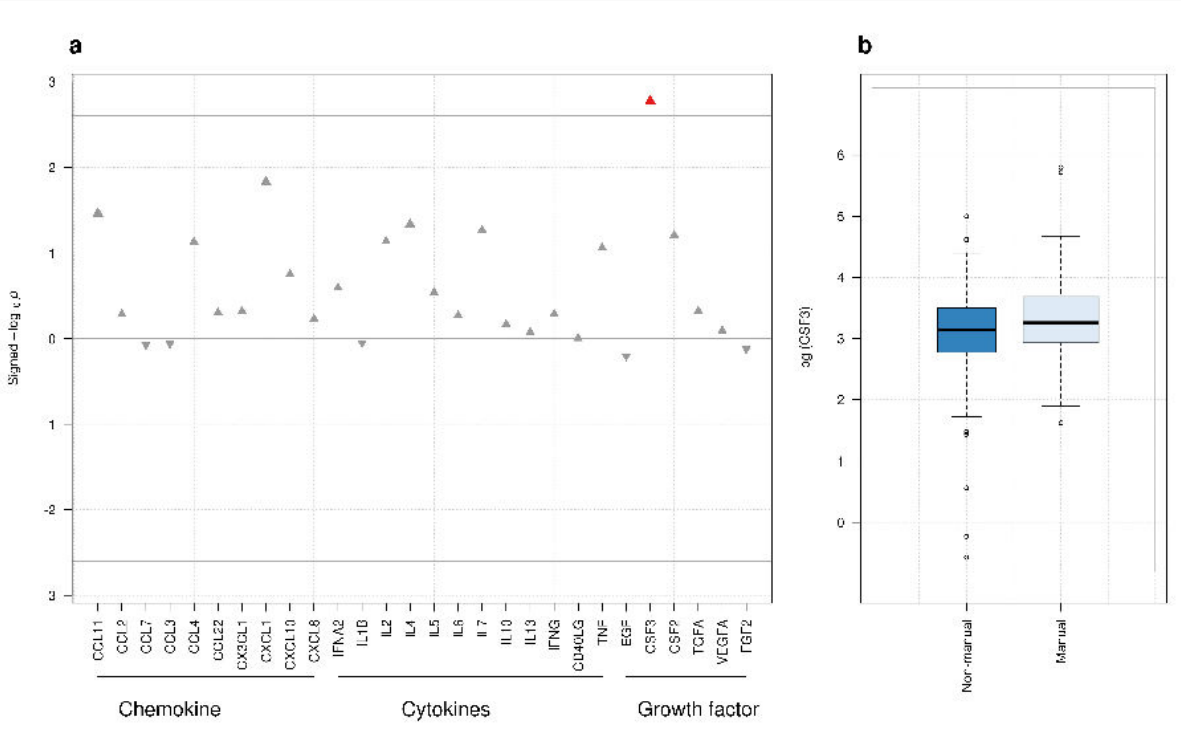


Figure: (a) Signed 'Manhattan plot' for the 28 proteins. **(b)** Boxplot of log transformed CSF3 plasma levels per father occupational position group.

→ General increased inflammation for lower paternal occupation and other SEP indicators

→ Only CSF3 remains significant after multiple testing correction

Life course linear regression with CSF3 blood levels

Table: Life course multiple regression analyses for plasma concentration of CSF3. Estimates are based on 230 participants with full SEP and lifestyle information.

		Model A		Model B-1		Model B-2		Model C		Fully Adjusted Model	
(A) Plasma concentration of CSF3											
Variables	**Levels**	β **(SE)**	**P-value**	β **(SE)**	**P-value**	β **(SE)**	**P-value**	β **(SE)**	**P-value**	β **(SE)**	**P-value**
Father's occupational position	Manual	0.29 (0.09)	0.002	0.27 (0.10)	0.008	0.29 (0.10)	0.004	0.28 (0.10)	0.008	0.28 (0.10)	0.007
Participant's education	Low			0.03 (0.10)	0.733	-	-	0.04 (0.11)	0.742	0.02 (0.12)	0.868
Household's highest occupation	Manual					0.01 (0.10)	0.916	-0.01 (0.11)	0.957	-0.02 (0.11)	0.862
BMI										0.02 (0.01)	0.230
Smoking status	Former									-0.01 (0.12)	0.917
	Current									-0.01 (0.12)	0.608
Alcohol										0.001 (0.003)	0.644

$\rightarrow$ Adjusting on later life SEP indicators do not affect CSF3-father's occupation association

$\rightarrow$ association seem not to be affected by later experiences

$\rightarrow$ Addition of the lifestyle factors do not affect the association

$\Rightarrow$ results suggest a biologically imprinted early-life exposure leading to higher inflammatory burden, and seems independant of SES-related exposures

Life course linear regression with the inflammatory score

Table: Life course multiple regression analyses for father's occupational using the inflammatory score **(B)** and the first PC **(C)**.

		Model A		Model B-1		Model B-2		Model C		Fully Adjusted Model	
(B) Inflammatory score											
Variables	**Levels**	**β (SE)**	**P-value**	**β (SE)**	**P-value**	**β (SE)**	**P-value**	**β (SE)**	**P-value**	**β (SE)**	**P-value**
Father's occupational position	Manual	1.96 (0.89)	0.029	2.88 (0.97)	0.003	2.64 (0.93)	0.005	3.08 (0.98)	0.002	2.93 (1.00)	0.004
Participant's education	Low			-2.22 (0.98)	0.024	-	-	-1.54 (1.08)	0.156	-1.5 (1.10)	0.174
Household's highest occupation	Manual					-2.22 (0.97)	0.023	-1.56 (1.07)	0.149	-1.49 (1.09)	0.174
BMI										-0.07 (0.13)	0.617
Smoking status	Former									-0.62 (1.16)	0.594
	Current									-0.57 (1.16)	0.621
Alcohol										-0.02 (0.03)	0.433
(C) Principal component 1											
Variables	**Levels**	**β (SE)**	**P-value**	**β (SE)**	**P-value**	**β (SE)**	**P-value**	**β (SE)**	**P-value**	**β (SE)**	**P-value**
Father's occupational position	Manual	-0.60 (0.45)	0.182	-1.05 (0.49)	0.031	-0.84 (0.47)	0.074	-1.10 (0.49)	0.026	-1.06 (0.50)	0.034
Participant's education	Low			1.10 (0.49)	0.025	-	-	0.93 (0.54)	0.088	0.95 (0.55)	0.086
Household's highest occupation	Manual					0.79 (0.49)	0.104	0.39 (0.54)	0.466	0.40 (0.55)	0.462
BMI										-0.01 (0.07)	0.856
Smoking status	Former									0.27 (0.58)	0.639
	Current									0.48 (0.58)	0.411
Alcohol										0.01 (0.01)	0.637

→ The association with early life SEP is detected by the score

→ For PC1 the association is significant upon adjustment on father's occupation

→ Results are robust to behavioural factors

→ For both scores, association with participant's education in model B-1 only

⇒ role of SEP trajectories?

Early childhood x young adulthood social mobility

Table: Multiple regression analyses of social mobility through the interaction term between father's occupation and participant highest household position. Results are presented for the inflammatory score **(A)** and the first PC **(B)**.

	Social mobility		
(A) Inflammatory score			
Variables	β	**SE**	**P-value**
Intercept (stable Non-manual)	8.40	3.18	0.009
Manual to Non-manual	2.38	1.04	0.023
Non-Manual to Manual	-3.36	2.16	0.122
stable Manual	0.42	1.11	0.705
(B) Principal component 1			
Variables	β	**SE**	**P-value**
Intercept (stable Non-manual)	-1.89	1.59	0.236
Manual to Non-manual	-0.72	0.52	0.170
Non-Manual to Manual	1.33	1.08	0.222
stable Manual	-0.05	0.56	0.933

→ The score reveals differential inflammatory status in 'stable Non Manual' and 'Non-Manual to Manual' participants

→ No association found using PC1

Social mobility

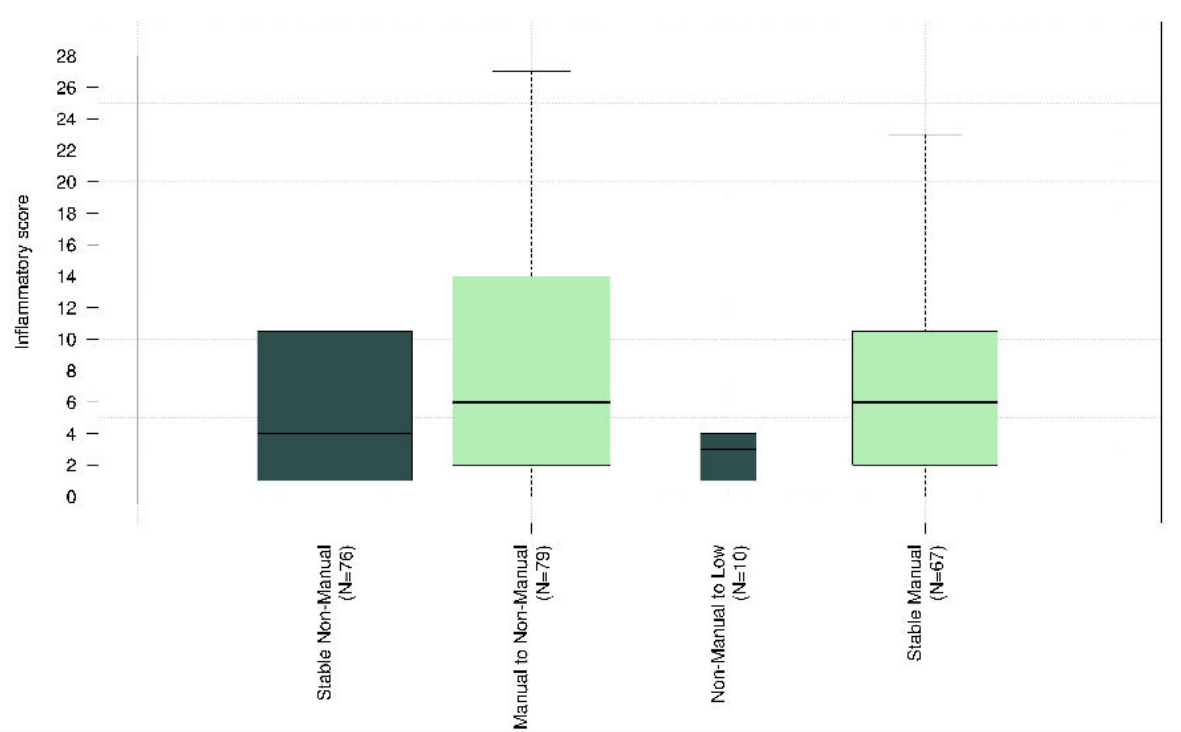
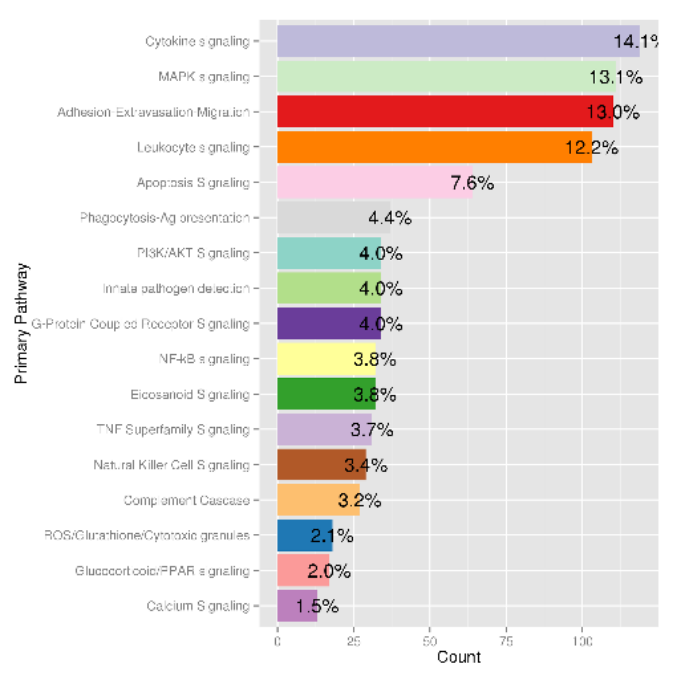


Figure: Box-and-whisker plot summarising the distribution of the inflammatory score across the four categories of the social mobility index.

→ Stronger effect of the upward social mobility

Choosing genes

- Pathways were build using Ingenuity Pathways Analysis
- Genes chosen were assigned to one of the functionnal pathways:
- 1027 genes in the paper, 845 genes present in our dataset



Loza *et al.*, 2007

Lifecourse linear regression with the inflammatory transcriptome score

	Inflammatory transcriptome score							
	Model A		Model B		Model C		Model D	
	β (se)	pval	β (se)	pval	β (se)	pval	β (se)	pval
father.job	21.81 (10.32)	3.59E-02	26.25 (11.26)	2.07E-02	25.41 (11.39)	2.68E-02	25.37 (11.46)	2.79E-02
education			-11.21 (11.35)	3.24E-01	-14.09 (12.61)	2.65E-01	-9.81 (12.67)	4.40E-01
hho					6.59 (12.49)	5.99E-01	9.14 (12.5)	4.65E-01
bmi							-3.23 (1.53)	3.62E-02
former							19.23 (13.44)	1.54E-01
current							10.19 (13.33)	4.45E-01
alcohol							-0.06 (0.34)	8.68E-01
	Inflammatory transcriptome PC 1							
	Model A		Model B		Model C		Model D	
	β (se)	pval	β (se)	pval	β (se)	pval	β (se)	pval
father.job	-4.03 (2.14)	6.11E-02	-4.36 (2.34)	6.34E-02	-4.12 (2.36)	8.28E-02	-3.93 (2.39)	1.02E-01
education			0.85 (2.35)	7.20E-01	1.68 (2.62)	5.22E-01	1.42 (2.64)	5.91E-01
hho					-1.9 (2.59)	4.64E-01	-1.98 (2.61)	4.49E-01
bmi							0.14 (0.32)	6.57E-01
former							-0.61 (2.8)	8.29E-01
current							4.51 (2.78)	1.06E-01
alcohol							0.04 (0.07)	5.35E-01

$\rightarrow$ Inflammatory transcriptome global score and PC1 are associated with father occupational position

$\rightarrow$ Association remains significant after adjusting for bmi, smoking status and alcohol

$\rightarrow$ No association with education after adjusting on early life sep

$\Rightarrow$ maybe consider alternative scores?

Can we replicate the association between the father's occupationnal position and the inflammatory transcriptome score?

Dataset GSE15180

- **Overall design:** Samples from 30 adults with low early-life SES and 30 adults with high early-life SES
- **Summary:** This study conducted transcriptional profiling of PBMC in healthy adults who were low vs. high in early-life SES to explore the long-lasting genomic effects of early experience
- **Platform:** Illumina HumanRef-8 v3.0 expression beadchip

Kobor MS *et al.*, 2009

One SES (high/low early SES) and no confounders

	β	β (se)	P-val
Global score	24.50	10.21	1.97E-02
PC1	-2.80	2.86	3.32E-01

$\rightarrow$ The association between the inflammatory transcriptome global score and early life SEP is replicated in the dataset GSE15180

Promising pilot results

- **Inflammation results:** SEP-inflammation associations were detected and involved SEP-trajectories
- **Power:** these associations were detected with limited size
- **Integration:** using prior knowledge we were able to integrate OMICs data from different platforms and to replicate results

Next steps

- **Methodological developments:** tested on existing data
- **Generalisation of the approach to other OMICS:** methylation (on-going) / adductomics
- **OMICS integration:** insight into cross-omics effect mediation
- **Harmonisation:** considerable effort is on-going to ensure data comparability across LIFEPATH study