Composite scores, social embodiment and risk of CVD: Evidence from the UK Biobank cohort

SoT – Risk Assessment Specialty Section webinar – 10^{th} February, 2021

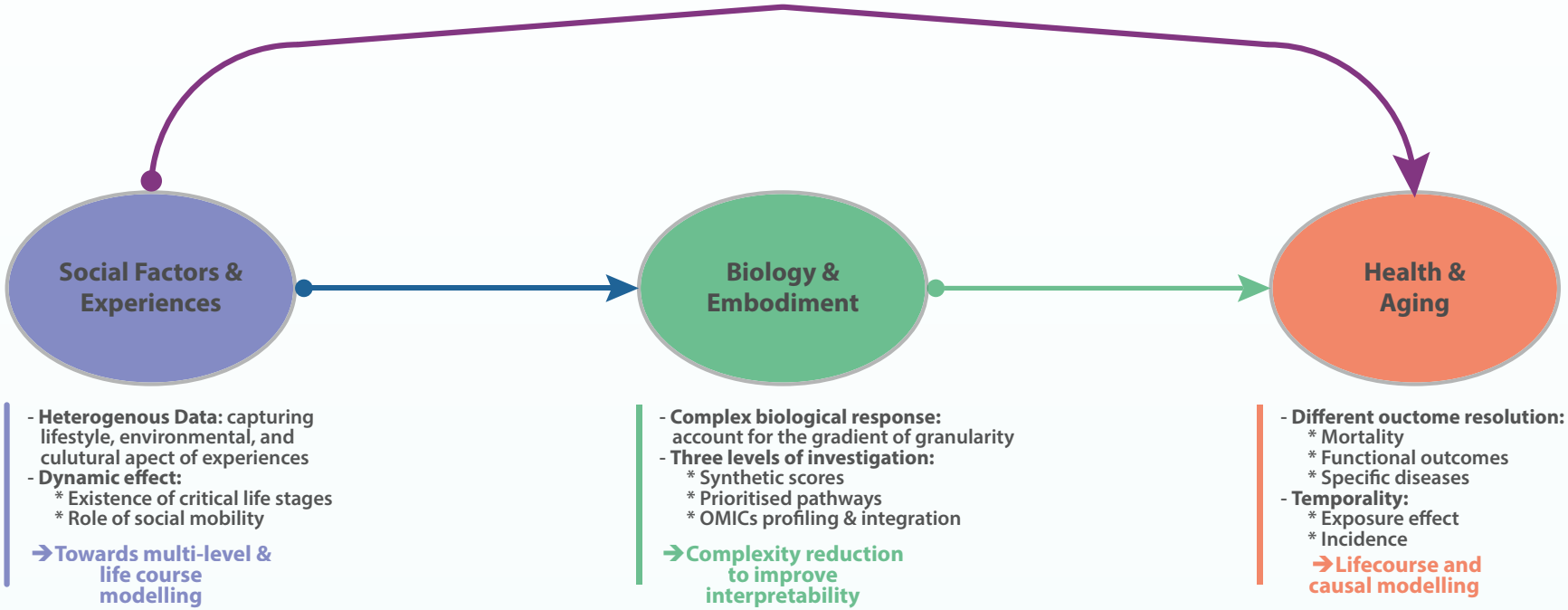
Marc Chadeau-Hyam

m.chadeau@imperial.ac.uk

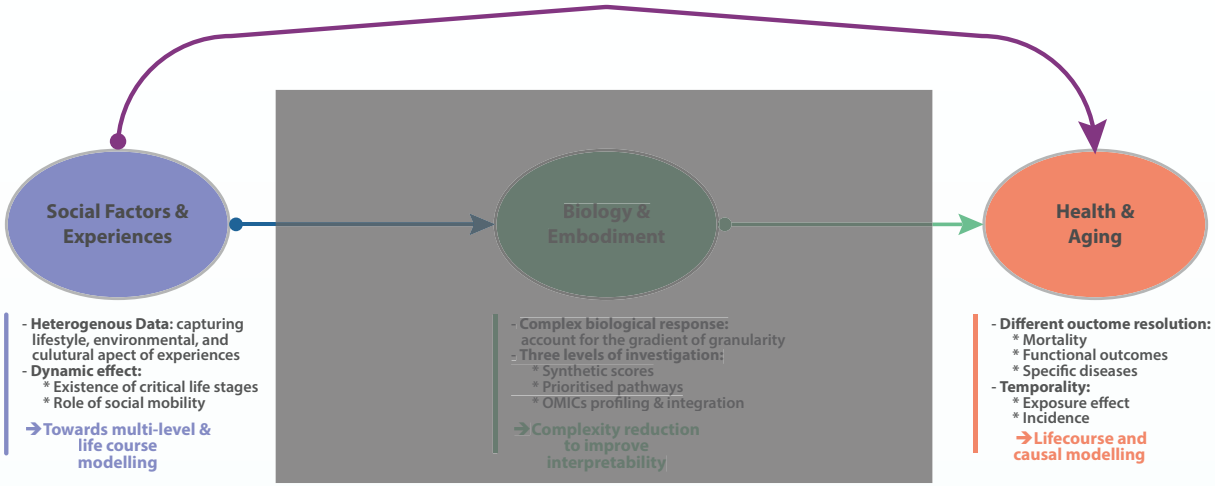
Imperial College London

LIFEPATH: Aims and overall approach

- Aim: Investigate mechanisms involved in the quality of ageing and health risk
- Data types: Social factors, biological markers, and health outcomes
- Overarching framework:



Stream 1: Linking Social Factors and health outcomes



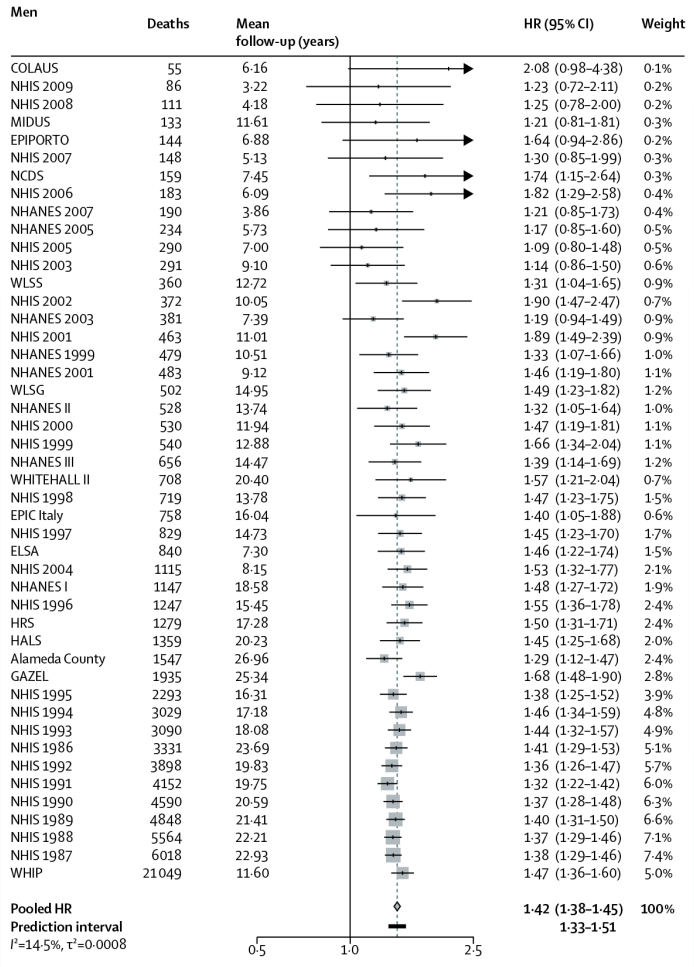
- Robust Evidence that social factors affect health risk


Articles

Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1·7 million men and women

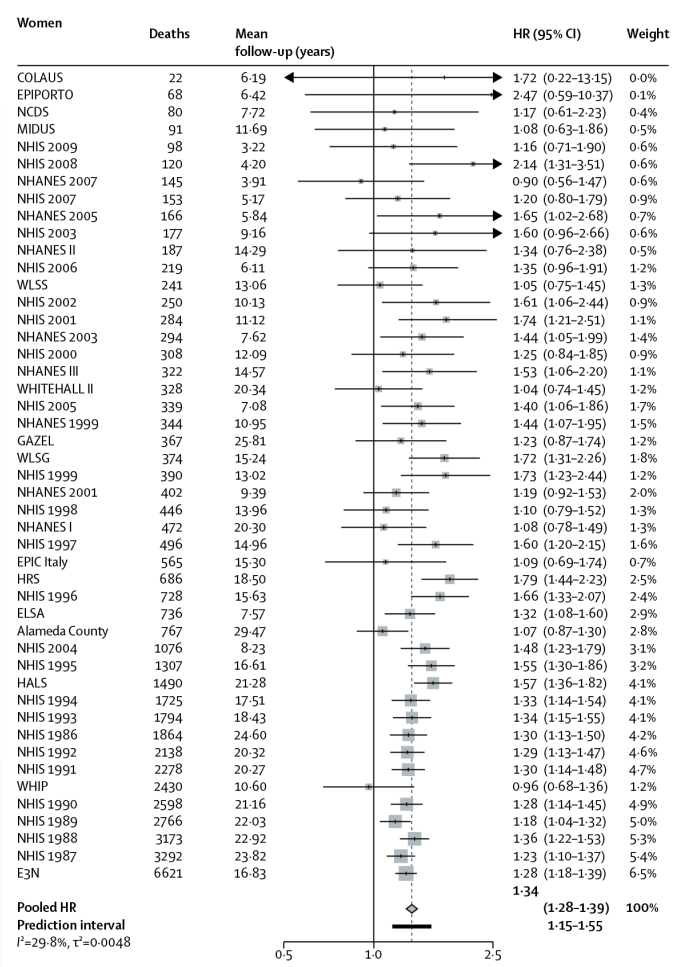
Silvia Stringhini*, Cristian Carmeli*, Markus Jokela*, Mauricio Avendaño*, Peter Muennig, Florence Guida, Fulvio Ricceri, Angelo d'Errico, Henrique Barros, Murielle Bochud, Marc Chadeau-Hyam, Françoise Clavel-Chapelon, Giuseppe Costa, Cyrille Delpierre, Silvia Fraga, Marcel Goldberg, Graham G Giles, Vittorio Krogh, Michelle Kelly-Irving, Richard Layte, Aurélie M Lasserre, Michael G Marmot, Martin Preisig, Martin J Shipley, Peter Vollenweider, Marie Zins, Ichiro Kawachi, Andrew Steptoe, Johan P Mackenbach, Paolo Vineis†, Mika Kivimäki†, for the LIFEPATH consortium‡

Stream 1: Linking Social Factors and health outcomes



Men	Deaths	Mean follow-up (years)	HR (95% CI)	Weight
COLAUS	55	6·16	2·08 (0·98–4·38)	0·1%
NHIS 2009	86	3·22	1·23 (0·72–2·11)	0·2%
NHIS 2008	111	4·18	1·25 (0·78–2·00)	0·2%
MIDUS	133	11·61	1·21 (0·81–1·81)	0·3%
EPIPORTO	144	6·88	1·64 (0·94–2·86)	0·2%
NHIS 2007	148	5·13	1·30 (0·85–1·99)	0·3%
NCDS	159	7·45	1·74 (1·15–2·64)	0·3%
NHIS 2006	183	6·09	1·82 (1·29–2·58)	0·4%
NHANES 2007	190	3·86	1·21 (0·85–1·73)	0·4%
NHANES 2005	234	5·73	1·17 (0·85–1·60)	0·5%
NHIS 2005	290	7·00	1·09 (0·80–1·48)	0·5%
NHIS 2003	291	9·10	1·14 (0·86–1·50)	0·6%
WLSS	360	12·72	1·31 (1·04–1·65)	0·9%
NHIS 2002	372	10·05	1·90 (1·47–2·47)	0·7%
NHANES 2003	381	7·39	1·19 (0·94–1·49)	0·9%
NHIS 2001	463	11·01	1·89 (1·49–2·39)	0·9%
NHANES 1999	479	10·51	1·33 (1·07–1·66)	1·0%
NHANES 2001	483	9·12	1·46 (1·19–1·80)	1·1%
WLSG	502	14·95	1·49 (1·23–1·82)	1·2%
NHANES II	528	13·74	1·32 (1·05–1·64)	1·0%
NHIS 2000	530	11·94	1·47 (1·19–1·81)	1·1%
NHIS 1999	540	12·88	1·66 (1·34–2·04)	1·1%
NHANES III	656	14·47	1·39 (1·14–1·69)	1·2%
WHITEHALL II	708	20·40	1·57 (1·21–2·04)	0·7%
NHIS 1998	719	13·78	1·47 (1·23–1·75)	1·5%
EPIC Italy	758	16·04	1·40 (1·05–1·88)	0·6%
NHIS 1997	829	14·73	1·45 (1·23–1·70)	1·7%
ELSA	840	7·30	1·46 (1·22–1·74)	1·5%
NHIS 2004	1115	8·15	1·53 (1·32–1·77)	2·1%
NHANES I	1147	18·58	1·48 (1·27–1·72)	1·9%
NHIS 1996	1247	15·45	1·55 (1·36–1·78)	2·4%
HRS	1279	17·28	1·50 (1·31–1·71)	2·4%
HALS	1359	20·23	1·45 (1·25–1·68)	2·0%
Alameda County	1547	26·96	1·29 (1·12–1·47)	2·4%
GAZEL	1935	25·34	1·68 (1·48–1·90)	2·8%
NHIS 1995	2293	16·31	1·38 (1·25–1·52)	3·9%
NHIS 1994	3029	17·18	1·46 (1·34–1·59)	4·8%
NHIS 1993	3090	18·08	1·44 (1·32–1·57)	4·9%
NHIS 1986	3331	23·69	1·41 (1·29–1·53)	5·1%
NHIS 1992	3898	19·83	1·36 (1·26–1·47)	5·7%
NHIS 1991	4152	19·75	1·32 (1·22–1·42)	6·0%
NHIS 1990	4590	20·59	1·37 (1·28–1·48)	6·3%
NHIS 1989	4848	21·41	1·40 (1·31–1·50)	6·6%
NHIS 1988	5564	22·21	1·37 (1·29–1·46)	7·1%
NHIS 1987	6018	22·93	1·38 (1·29–1·46)	7·4%
WHIP	21049	11·60	1·47 (1·36–1·60)	5·0%
Pooled HR			**1·42 (1·38–1·45)**	**100%**
Prediction interval			**1·33–1·51**	
I^2=14·5%, τ^2=0·0008				

0·5 1·0 2·5

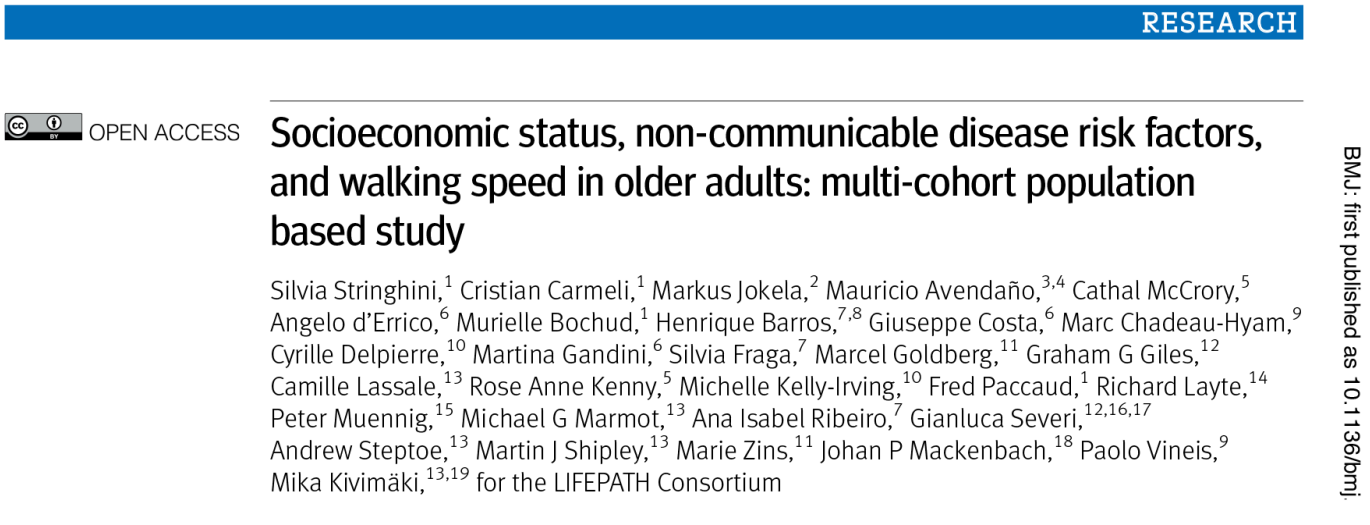


Women	Deaths	Mean follow-up (years)	HR (95% CI)	Weight
COLAUS	22	6·19	1·72 (0·22–13·15)	0·0%
EPIPORTO	68	6·42	2·47 (0·59–10·37)	0·1%
NCDS	80	7·72	1·17 (0·61–2·23)	0·4%
MIDUS	91	11·69	1·08 (0·63–1·86)	0·5%
NHIS 2009	98	3·22	1·16 (0·71–1·90)	0·6%
NHIS 2008	120	4·20	2·14 (1·31–3·51)	0·6%
NHANES 2007	145	3·91	0·90 (0·56–1·47)	0·6%
NHIS 2007	153	5·17	1·20 (0·80–1·79)	0·9%
NHANES 2005	166	5·84	1·65 (1·02–2·68)	0·7%
NHIS 2003	177	9·16	1·60 (0·96–2·66)	0·6%
NHANES II	187	14·29	1·34 (0·76–2·38)	0·5%
NHIS 2006	219	6·11	1·35 (0·96–1·91)	1·2%
WLSS	241	13·06	1·05 (0·75–1·45)	1·3%
NHIS 2002	250	10·13	1·61 (1·06–2·44)	0·9%
NHIS 2001	284	11·12	1·74 (1·21–2·51)	1·1%
NHANES 2003	294	7·62	1·44 (1·05–1·99)	1·4%
NHIS 2000	308	12·09	1·25 (0·84–1·85)	0·9%
NHANES III	322	14·57	1·53 (1·06–2·20)	1·1%
WHITEHALL II	328	20·34	1·04 (0·74–1·45)	1·2%
NHIS 2005	339	7·08	1·40 (1·06–1·86)	1·7%
NHANES 1999	344	10·95	1·44 (1·07–1·95)	1·5%
GAZEL	367	25·81	1·23 (0·87–1·74)	1·2%
WLSG	374	15·24	1·72 (1·31–2·26)	1·8%
NHIS 1999	390	13·02	1·73 (1·23–2·44)	1·2%
NHANES 2001	402	9·39	1·19 (0·92–1·53)	2·0%
NHIS 1998	446	13·96	1·10 (0·79–1·52)	1·3%
NHANES I	472	20·30	1·08 (0·78–1·49)	1·3%
NHIS 1997	496	14·96	1·60 (1·20–2·15)	1·6%
EPIC Italy	565	15·30	1·09 (0·69–1·74)	0·7%
HRS	686	18·50	1·79 (1·44–2·23)	2·5%
NHIS 1996	728	15·63	1·66 (1·33–2·07)	2·4%
ELSA	736	7·57	1·32 (1·08–1·60)	2·9%
Alameda County	767	29·47	1·07 (0·87–1·30)	2·8%
NHIS 2004	1076	8·23	1·48 (1·23–1·79)	3·1%
NHIS 1995	1307	16·61	1·55 (1·30–1·86)	3·2%
HALS	1490	21·28	1·57 (1·36–1·82)	4·1%
NHIS 1994	1725	17·51	1·33 (1·14–1·54)	4·1%
NHIS 1993	1794	18·43	1·34 (1·15–1·55)	4·1%
NHIS 1986	1864	24·60	1·30 (1·13–1·50)	4·2%
NHIS 1992	2138	20·32	1·29 (1·13–1·47)	4·6%
NHIS 1991	2278	20·27	1·30 (1·14–1·48)	4·7%
WHIP	2430	10·60	0·96 (0·68–1·36)	1·2%
NHIS 1990	2598	21·16	1·28 (1·14–1·45)	4·9%
NHIS 1989	2766	22·03	1·18 (1·04–1·32)	5·0%
NHIS 1988	3173	22·92	1·36 (1·22–1·53)	5·3%
NHIS 1987	3292	23·82	1·23 (1·10–1·37)	5·4%
E3N	6621	16·83	1·28 (1·18–1·39)	6·5%
			1·34	
Pooled HR			**(1·28–1·39)**	**100%**
Prediction interval			**1·15–1·55**	
I^2=29·8%, τ^2=0·0048				

0·5 1·0 2·5

- 46 cohorts support higher mortality in lower SES in men, HR=1.4
- 44 of the cohorts provide consistent results in women, HR=1.3

Stream 1: Linking Social Factors and Functional outcomes



RESEARCH

OPEN ACCESS

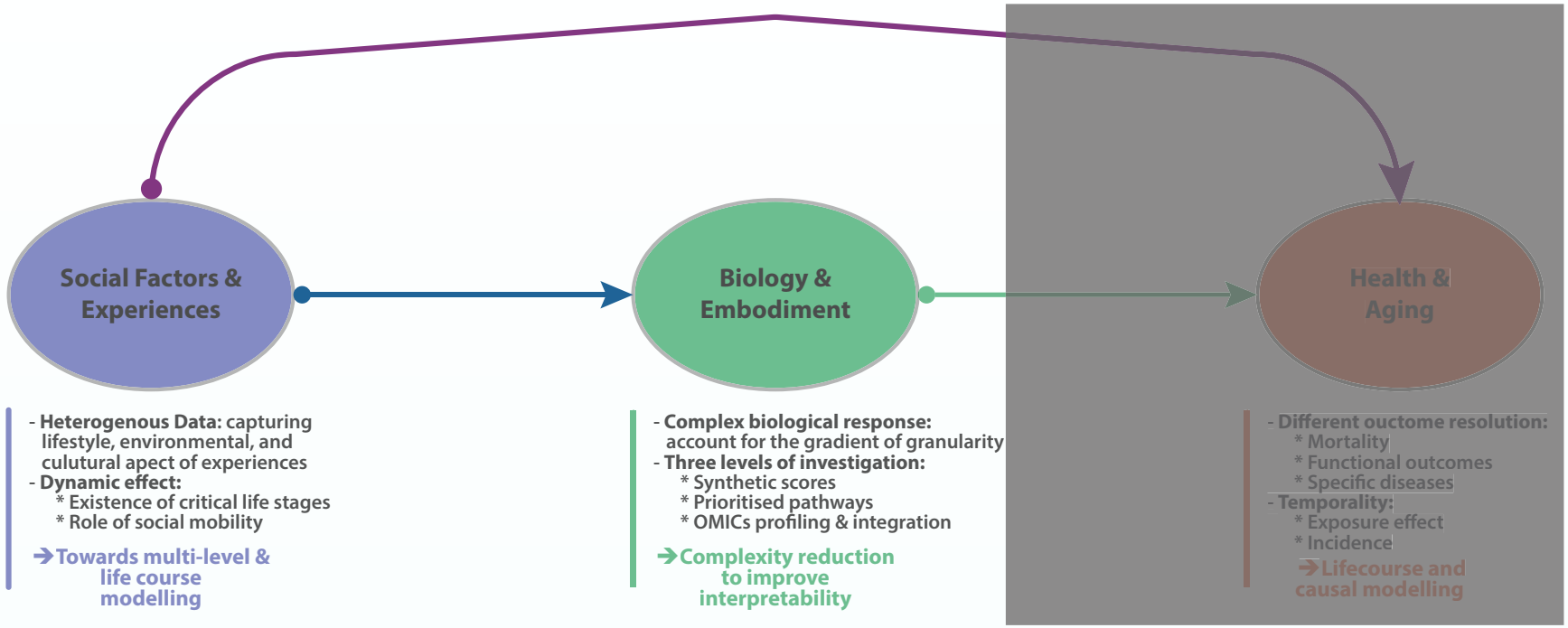
Socioeconomic status, non-communicable disease risk factors, and walking speed in older adults: multi-cohort population based study

Silvia Stringhini,[1] Cristian Carmeli,[1] Markus Jokela,[2] Mauricio Avendaño,[3,4] Cathal McCrory,[5] Angelo d'Errico,[6] Murielle Bochud,[1] Henrique Barros,[7,8] Giuseppe Costa,[6] Marc Chadeau-Hyam,[9] Cyrille Delpierre,[10] Martina Gandini,[6] Silvia Fraga,[7] Marcel Goldberg,[11] Graham G Giles,[12] Camille Lassale,[13] Rose Anne Kenny,[5] Michelle Kelly-Irving,[10] Fred Paccaud,[1] Richard Layte,[14] Peter Muennig,[15] Michael G Marmot,[13] Ana Isabel Ribeiro,[7] Gianluca Severi,[12,16,17] Andrew Steptoe,[13] Martin J Shipley,[13] Marie Zins,[11] Johan P Mackenbach,[18] Paolo Vineis,[9] Mika Kivimäki,[13,19] for the LIFEPATH Consortium

BMJ: first published as 10.1136/bmj

- By age 85, low vs high SES is associated with a loss of 7-11 years of functioning.
- These cannot be attributed to health risk factors.
- Most risk factors (except physical activity) are related to smaller losses of functioning

Stream 2: Investigating biomarkers of SEP



- 3 approaches corresponding to a gradient of granularity
  1. Focus on prioritized pathways: Inflammation
  2. Synthetic scores
  3. OMICs approaches

Focus on Inflammation

ARTICLE

https://doi.org/10.1038/s41467-019-08732-x OPEN

Multi-cohort study identifies social determinants of systemic inflammation over the life course

Eloïse Berger[1], Raphaële Castagné[1], Marc Chadeau-Hyam[2], Murielle Bochud[3], Angelo d'Errico[4], Martina Gandini[4], Maryam Karimi[2], Mika Kivimäki[5,6], Vittorio Krogh[7], Michael Marmot[5], Salvatore Panico[8], Martin Preisig[3], Fulvio Ricceri[4], Carlotta Sacerdote[9], Andrew Steptoe[5], Silvia Stringhini[10], Rosario Tumino[11], Paolo Vineis[2,12], Cyrille Delpierre[1] & Michelle Kelly-Irving[1]

- Data: 6 LIFEPATH cohorts (N=23,008), with SEP factors in the life-course & CRP measurement in blood as a proxy for inflammatory status.
- Aim: Explore the CRP-SEP association across country; in the life-course; evaluate the role of lifestyle factors and behaviors
- Results: overall higher inflammatory burden in lower SEP group
  - Consistent gradient irrespective of the SEP metric
  - Stronger associations in women
  - Lifestyle factors marginally attenuate the associations

Complexity Reduction approach: Allostastic Load


Psychoneuroendocrinology 67 (2016) 76–85

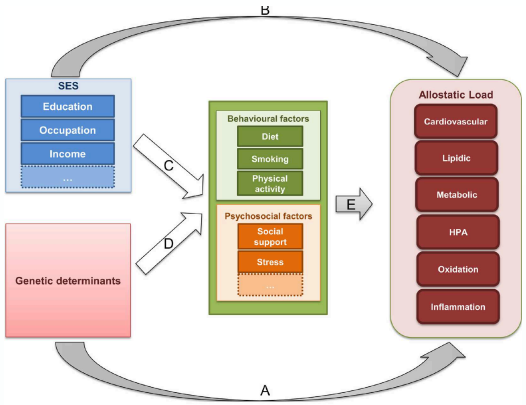
Contents lists available at ScienceDirect

Psychoneuroendocrinology

journal homepage: www.elsevier.com/locate/psyneuen

Sociodemographic, behavioral and genetic determinants of allostatic load in a Swiss population-based study

Dusan Petrovic[a], Edward Pivin[a], Belen Ponte[a,b], Nasser Dhayat[c], Menno Pruijm[d], Georg Ehret[e], Daniel Ackermann[c], Idris Guessous[a,f], Sandrine Estoppey Younes[a], Antoinette Pechère-Bertschi[b], Bruno Vogt[c], Markus Mohaupt[c], Pierre-Yves Martin[b], Fred Paccaud[a], Michel Burnier[d], Murielle Bochud[a], Silvia Stringhini[a,*]



- Definition: multi-system synthetic score capturing physiological wear-and-tear (6 systems included)
- Data: SKIPOGH study (N=1,128), with SEP &14 blood-derived biomarkers
- Results: SEP-AL associations by gender
  - Main trend: higher AL for lower SEP categories
  - Stronger associations for education, and in women
  - Lifestyle factors marginally attenuate the associations

$\Rightarrow$ stronger gradient for early life SEP

The BHS an extension over the allostatic load

Research report

OPEN ACCESS

Early-life inequalities and biological ageing: a multisystem Biological Health Score approach in *Understanding Society*

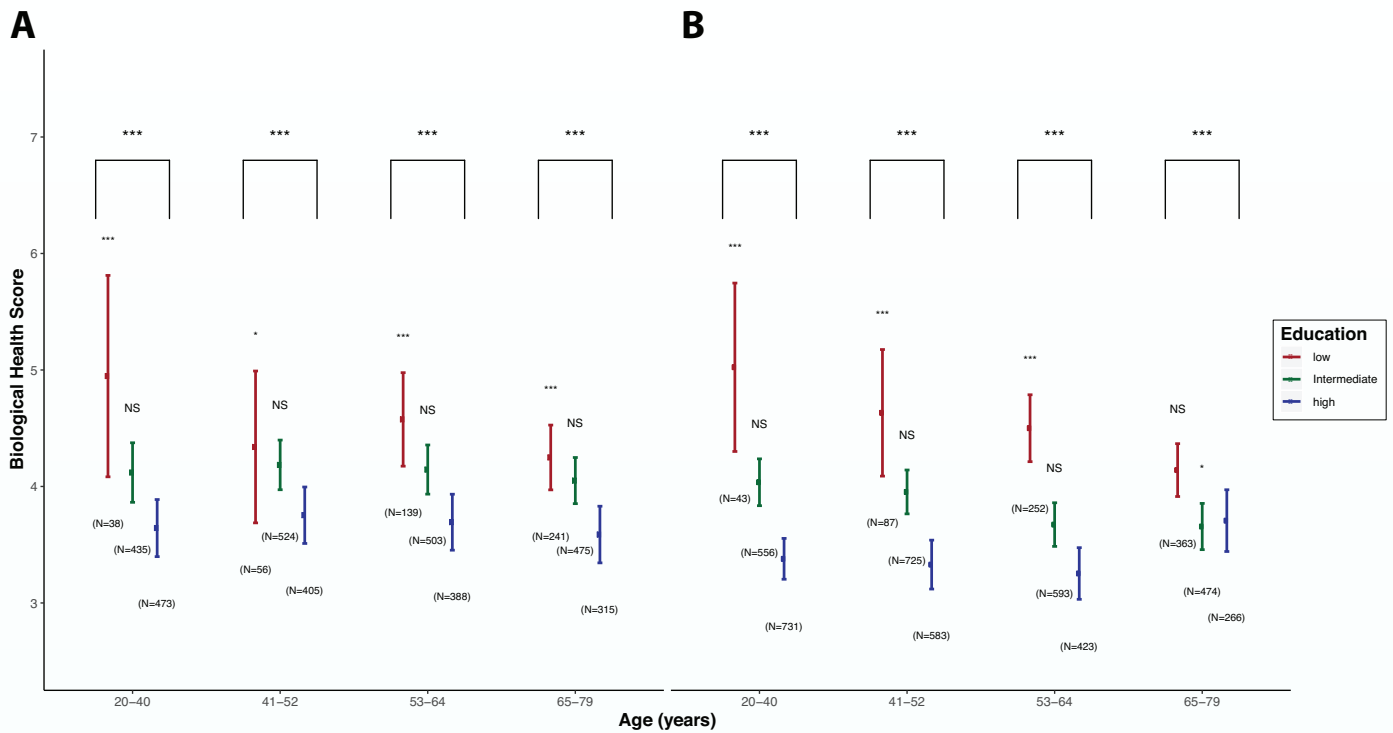
Maryam Karimi,[1,2] Raphaële Castagné,[3] Cyrille Delpierre,[3] Gaëlle Albertus,[3] Eloïse Berger,[3] Paolo Vineis,[1,2,4] Meena Kumari,[5] Michelle Kelly-Irving,[3] Marc Chadeau-Hyam[1,2]

BMJ

Karimi M, *et al*. *J Epidemiol Community Health* 2019;**0**:1–10. doi:10.1136/jech-2018-212010 1

- Data: Understanding Society (N=9,088), with educational attainment &16 blood-derived biomarkers capturing 6 physiological systems (including liver and kidney functions)
- Aims: Define BHS as an extension from the AL
  1. Explore BHS gradient across SEP groups and age classes
  2. Quantify the relative contribution of each system

Investigating social gradients in composite scores

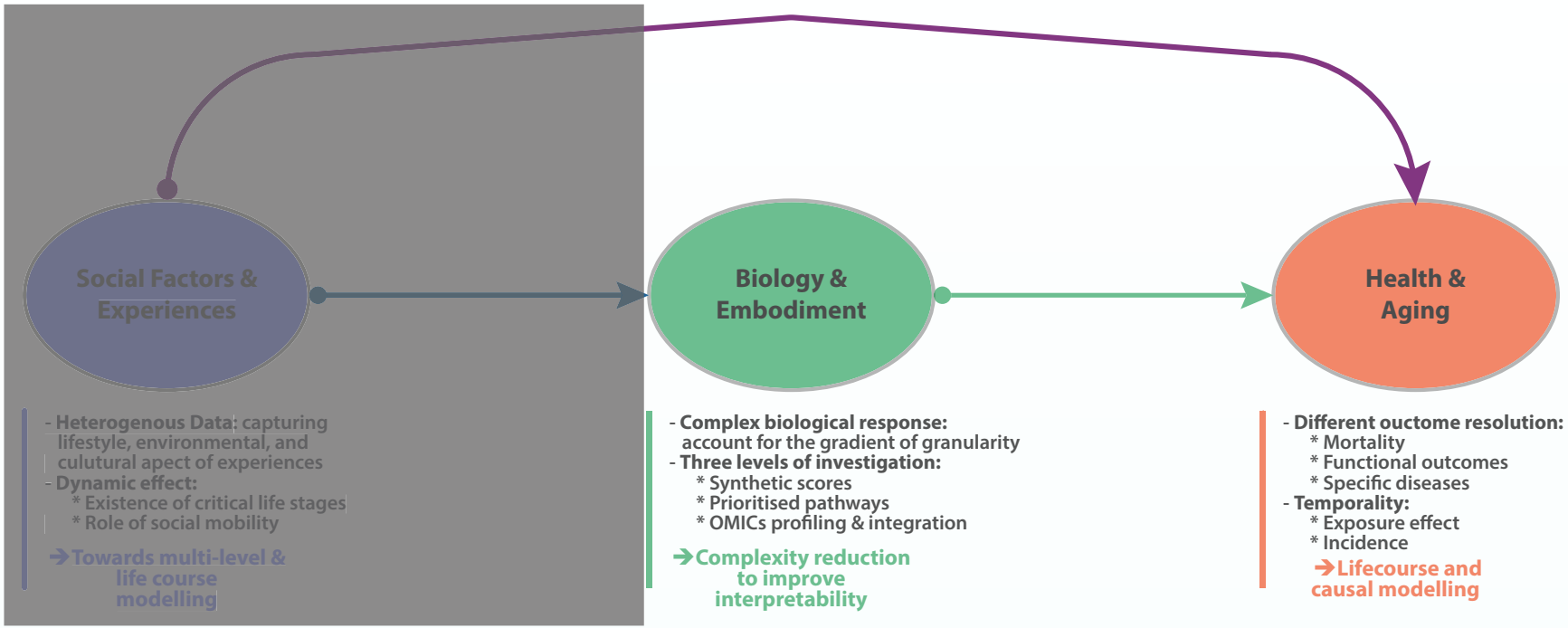


Results: systematic SEP-related gradient (higher scores in disadvantaged pop)

- Consistent results in men (A) and women (B)
- Gradient is observed in all age groups
- Gradient is not affected by adjustment on lifestyle factors

$\Rightarrow$ effects are detected in early adulthood and persist

Step 3: From biology to Health Outcomes



- Gradient of resolution:
  1. Low resolution biological factors
  2. Pathways
  3. Composite Scores
  4. Full-resolution OMICs profiles

Allostatic Load and mortality

European Journal of Epidemiology (2018) 33:441–458
https://doi.org/10.1007/s10654-018-0364-1

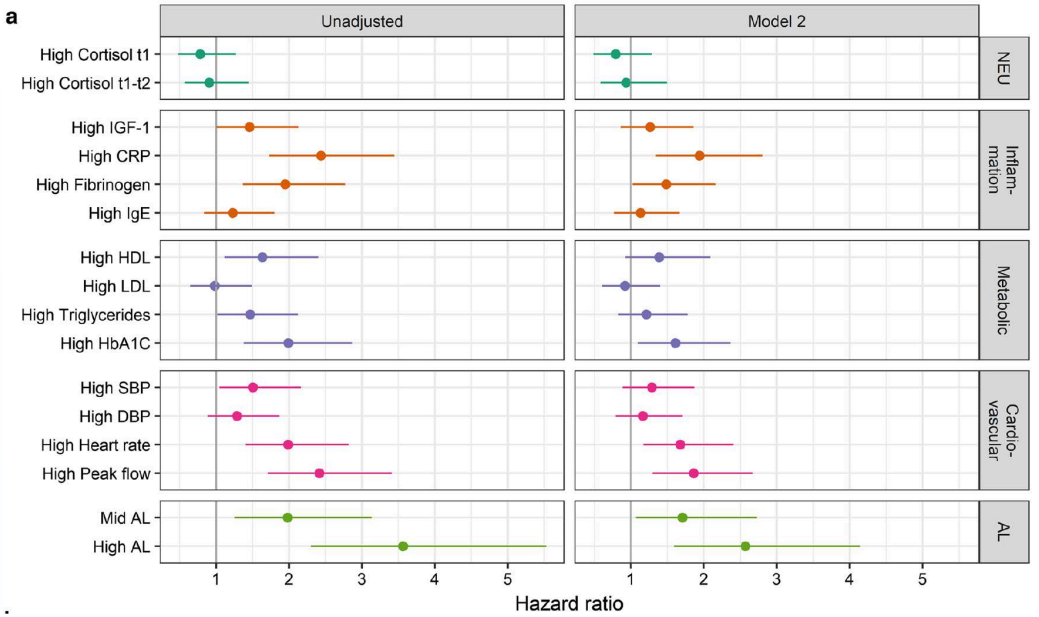
MORTALITY

Allostatic load and subsequent all-cause mortality: which biological markers drive the relationship? Findings from a UK birth cohort

Raphaële Castagné[1] · Valérie Garès[1] · Maryam Karimi[2] · Marc Chadeau-Hyam[2] · Paolo Vineis[2,3] · Cyrille Delpierre[1] · Michelle Kelly-Irving[1] · for the Lifepath Consortium

- Data: 1958 British Birth cohort (N=8,113) 14 blood-derived biomarkers. 132 deaths
- Aims
  1. Evaluate the effect of AL and its constituents on mortality
  2. Investigate the role of SEP (education) and behaviors in these associations.

Allostatic Load and mortality



Results: Hazard Ratio by system and biomarker and for AL

- Positive contribution of all markers/system to mortality (except NE)
- Attenuation upon adjustment on adulthood confounders
- Effect of AL, and most system remain significant after adjustment for behaviors and SEP (not shown).

$\Rightarrow$AL at 44 predicts mortality irrespective of subsequent SE experiences

$\Rightarrow$ the multi-system AL predicts better than each system separately

BHS, mortality and incident pathologies: UKBiobank

- Study Overview: 502,536 volunteers from the UK aged 37-73 years at entry between 2006 and 2010.
- Questionnaire data: computer-based questionnaire on life-course exposures, medical history and treatments.
- Anthropometric/clinical data: from clinical assessment centres computer-based questionnaire on life-course exposures, medical history and treatments.
- Mortality Outcomes: linkage to death registers
- Health status Follow-up: range from 0.2 to 12.04 years
- Incident pathologies identified through linkage to NHS central registers, cancer and hospital registers, and/or nurse-administered questionnaire.
- Biosampling: participants donated one blood sample at baseline
  - Genome-wide scans were measured (N=672,345 genotyped SNPs in 488,377 participants)
  - Panel of 30 prioritised biomarkers

Health Outcomes of interest in UK Biobank

- Cancer Outcomes: all sites
- CVD Outcomes: coronary arterial disease, angina, stroke, and related outcomes
- External cause mortality: suicide and accident
- Mortality Outcomes: all-cause, cancer, CVD and external cause

	All-cause	Cancer	CVD	External causes	Other causes
Males	8,735	4,316	1,547	328	2,544
Females	5,661	3,698	352	160	1,451
Total	**14,396**	**8,014**	**1,899**	**488**	**3,995**

- Incident pathologies: cancer and CVD incidence

	Cancer	CVD
Males	26,123	10,114
Females	26,320	5,539
Total	**52,443**	**15,653**

Biomarkers selection

- UK Biobank Biomarkers: 13 measuring 5 systems:
  1. ***Metabolic system (N=4):*** Glycated haemoglobin (HbA1c), High-density lipoprotein cholesterol (HDL), Low-density lipoprotein cholesterol (LDL), Triglycerides (Tri);
  2. ***Cardiovascular system (N=3):*** Systolic and diastolic blood pressure (SBP, and DBP, respectively), pulse (Pulse);
  3. ***Inflammatory / immune system (N=2):*** C-Reactive Protein (CRP), and Insulin-like growth factor 1 (IGF-1);
  4. ***Liver function (N=3):*** Alanine transaminase (ALT), Aspartate transaminase (AST), Gamma glutamyltransferase (GGT);
  5. ***Kidney function (N=1):*** Creatinine (Cre);

Complexity Reduction: Composite score

- Biomarkers dichotomisation: we define the 'at-risk' quartile as the lowest quartile for HDL and IGF-1, and the highest quartile for all remaining 11 biomarkers. Quartiles were defined for each gender and age group (<50, 50-64, and >64 years old) separately.
- Scores derivation: For a given system s, and individual i

$$\text{sub-BHS}_S^i = \sum_{k=1}^{13} \frac{I_k^i}{n_s}, \text{ where}$$

I_k^I is the binary score for biomarker k, and n_s: # of biomarkers in system s

- BHS definition:

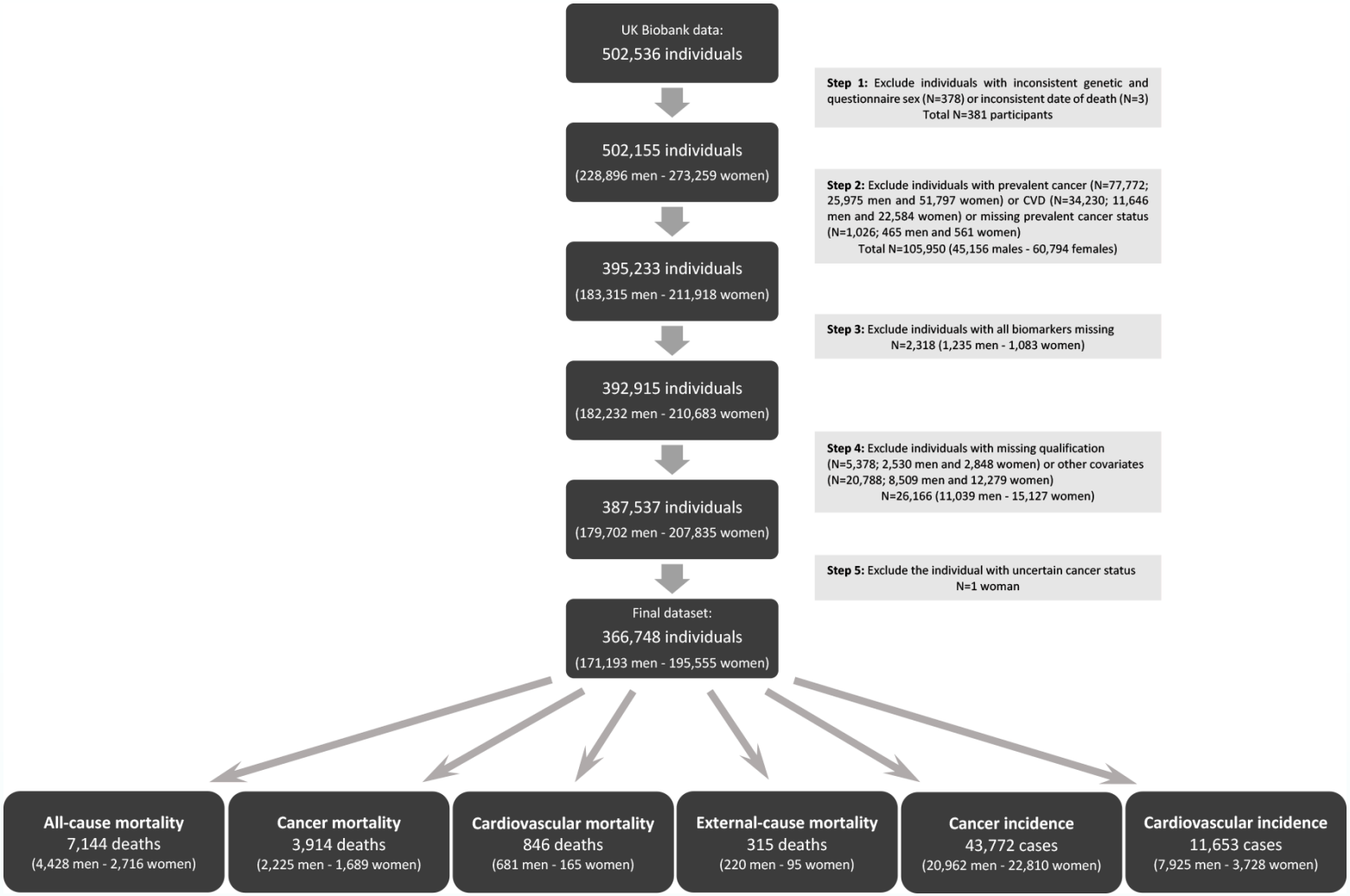
$$BHS^i = \frac{\sum_{s=1}^{5} \text{sub-BHS}_s^i}{n_s} \text{ where } n_s \text{ is the number of systems in the BHS}$$

$\Rightarrow$ the BHS and sub scores are all on the same scale ($\in [0, 1]$)

Main Analytical Plan

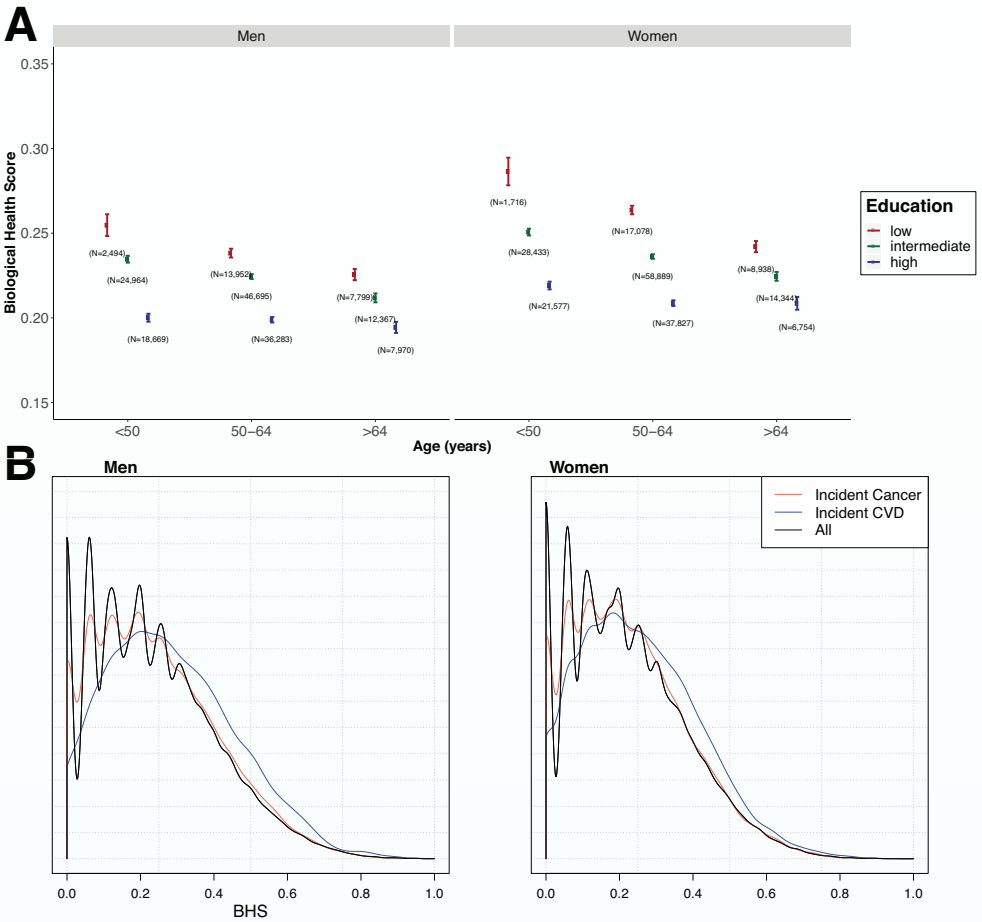
- Descriptive Analyses: Investigate social gradients in BHS
  - Compare BHS levels by education level (low, intermediate, high)
  - Investigate the role of socially-patterned exposures and behaviours in these gradients (medical status, smoking, physical activity, alcohol, and BMI)
- Survival Analyses: proportional hazards Cox models
  - Investigate the role of the BHS (and sub-scores) in mortality and incident pathologies: setting the BHS (or sub-scores as predictor)
  - Attenuation analyses: sequential adjustment for (i) Education, (ii) Behaviours, (iii) BMI, (iv) Medical History
- Investigate the role of Education:
  - Similar survival analyses, setting the Education level as predictor
  - Attenuation analyses: sequential adjustment for (i) Behaviours, (ii) BMI, (iii) Medical History, and (iv) BHS

Participants Selection



$\Rightarrow$ We selected a total of 366,748 participants (171,193 men and 195,555 women) who were free of cancer and CVD at baseline

BHS Distribution in UK Biobank



- Unambiguous social gradient in the BHS from UK Biobank
- Slightly right shifted BHS distributions in incident CVD cases

$\Rightarrow$ has the BHS an effect on mortality and incidence, independent of education?

Survival analyses: Univariate Models

MEN	Mortality				Incidence	
	All-cause N=4,428	Cancer N=2,225	CVD N=681	External cause N=220	Cancer N=20,962	CVD N=7,925
	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value
BHS	1.14 [1.12-1.16] $7.63x10^{-44}$	1.11 [1.09-1.14] $1.00x10^{-16}$	1.25 [1.20-1.31] $2.70x10^{-24}$	0.99 [0.91-1.08] $8.49x10^{-01}$	1.02 [1.01-1.03] $1.01x10^{-04}$	1.15 [1.13-1.16] $1.28x10^{-93}$
System-specific sub-score						
Metabolic	1.05 [1.04-1.06] $2.66x10^{-14}$	1.04 [1.02-1.06] $1.27x10^{-05}$	1.14 [1.11-1.18] $5.80x10^{-20}$	0.97 [0.92-1.03] $3.45x10^{-01}$	1.01 [1.00-1.01] $4.20x10^{-03}$	1.12 [1.11-1.13] $6.43x10^{-138}$
Cardiovascular	1.05 [1.04-1.06] $2.99x10^{-26}$	1.04 [1.02-1.05] $1.32x10^{-07}$	1.10 [1.07-1.12] $1.06x10^{-15}$	1.04 [1.00-1.08] $6.13x10^{-02}$	1.00 [1.00-1.01] $2.22x10^{-01}$	1.05 [1.05-1.06] $1.31x10^{-51}$
Inflammatory	1.07 [1.06-1.08] $3.20x10^{-57}$	1.06 [1.05-1.08] $7.09x10^{-22}$	1.09 [1.07-1.11] $3.51x10^{-15}$	1.01 [0.97-1.05] $7.00x10^{-01}$	1.01 [1.01-1.02] $1.20x10^{-06}$	1.04 [1.03-1.05] $6.72x10^{-32}$
Liver	1.03 [1.02-1.04] $1.89x10^{-11}$	1.03 [1.01-1.04] $2.06x10^{-04}$	1.04 [1.02-1.07] $3.61x10^{-04}$	1.02 [0.98-1.06] $3.73x10^{-01}$	1.00 [1.00-1.01] $1.63x10^{-01}$	1.02 [1.02-1.03] $1.42x10^{-10}$
Kidney	0.99 [0.98-0.99] $1.24x10^{-03}$	1.00 [0.98-1.01] $4.20x10^{-01}$	0.99 [0.97-1.01] $5.75x10^{-01}$	0.95 [0.91-0.99] $2.05x10^{-02}$	1.00 [1.00-1.00] $8.60x10^{-01}$	1.00 [0.99-1.01] $8.59x10^{-01}$

Survival analyses: Univariate Models

MEN	Mortality				Incidence	
	All-cause N=4,428	Cancer N=2,225	CVD N=681	External cause N=220	Cancer N=20,962	CVD N=7,925
	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value
BHS	1.14 [1.12-1.16] $7.63\text{x}10^{-44}$	1.11 [1.09-1.14] $1.00\text{x}10^{-16}$	1.25 [1.20-1.31] $2.70\text{x}10^{-24}$	0.99 [0.91-1.08] $8.49\text{x}10^{-01}$	1.02 [1.01-1.03] $1.01\text{x}10^{-04}$	1.15 [1.13-1.16] $1.28\text{x}10^{-93}$
System-specific sub-score						
Metabolic	1.05 [1.04-1.06] $2.66\text{x}10^{-14}$	1.04 [1.02-1.06] $1.27\text{x}10^{-05}$	1.14 [1.11-1.18] $5.80\text{x}10^{-20}$	0.97 [0.92-1.03] $3.45\text{x}10^{-01}$	1.01 [1.00-1.01] $4.20\text{x}10^{-03}$	1.12 [1.11-1.13] $6.43\text{x}10^{-138}$
Cardiovascular	1.05 [1.04-1.06] $2.99\text{x}10^{-26}$	1.04 [1.02-1.05] $1.32\text{x}10^{-07}$	1.10 [1.07-1.12] $1.06\text{x}10^{-15}$	1.04 [1.00-1.08] $6.13\text{x}10^{-02}$	1.00 [1.00-1.01] $2.22\text{x}10^{-01}$	1.05 [1.05-1.06] $1.31\text{x}10^{-51}$
Inflammatory	1.07 [1.06-1.08] $3.20\text{x}10^{-57}$	1.06 [1.05-1.08] $7.09\text{x}10^{-22}$	1.09 [1.07-1.11] $3.51\text{x}10^{-15}$	1.01 [0.97-1.05] $7.00\text{x}10^{-01}$	1.01 [1.01-1.02] $1.20\text{x}10^{-06}$	1.04 [1.03-1.05] $6.72\text{x}10^{-32}$
Liver	1.03 [1.02-1.04] $1.89\text{x}10^{-11}$	1.03 [1.01-1.04] $2.06\text{x}10^{-04}$	1.04 [1.02-1.07] $3.61\text{x}10^{-04}$	1.02 [0.98-1.06] $3.73\text{x}10^{-01}$	1.00 [1.00-1.01] $1.63\text{x}10^{-01}$	1.02 [1.02-1.03] $1.42\text{x}10^{-10}$
Kidney	0.99 [0.98-0.99] $1.24\text{x}10^{-03}$	1.00 [0.98-1.01] $4.20\text{x}10^{-01}$	0.99 [0.97-1.01] $5.75\text{x}10^{-01}$	0.95 [0.91-0.99] $2.05\text{x}10^{-02}$	1.00 [1.00-1.00] $8.60\text{x}10^{-01}$	1.00 [0.99-1.01] $8.59\text{x}10^{-01}$

- BHS is associated to increased mortality, from all-cause, cancer and CVD: HR range 1.11 to 1.25 $p < 10^{-16}$ in men
- None of the scores are related to external cause mortality
- Unlike other systems, kidney weakly contributes to mortality

Survival analyses: Univariate Models

MEN	Mortality				Incidence	
	All-cause N=4,428	Cancer N=2,225	CVD N=681	External cause N=220	Cancer N=20,962	CVD N=7,925
	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value
BHS	1.14 [1.12-1.16] $7.63\text{x}10^{-44}$	1.11 [1.09-1.14] $1.00\text{x}10^{-16}$	1.25 [1.20-1.31] $2.70\text{x}10^{-24}$	0.99 [0.91-1.08] $8.49\text{x}10^{-01}$	1.02 [1.01-1.03] $1.01\text{x}10^{-04}$	1.15 [1.13-1.16] $1.28\text{x}10^{-93}$
System-specific sub-score						
Metabolic	1.05 [1.04-1.06] $2.66\text{x}10^{-14}$	1.04 [1.02-1.06] $1.27\text{x}10^{-05}$	1.14 [1.11-1.18] $5.80\text{x}10^{-20}$	0.97 [0.92-1.03] $3.45\text{x}10^{-01}$	1.01 [1.00-1.01] $4.20\text{x}10^{-03}$	1.12 [1.11-1.13] $6.43\text{x}10^{-138}$
Cardiovascular	1.05 [1.04-1.06] $2.99\text{x}10^{-26}$	1.04 [1.02-1.05] $1.32\text{x}10^{-07}$	1.10 [1.07-1.12] $1.06\text{x}10^{-15}$	1.04 [1.00-1.08] $6.13\text{x}10^{-02}$	1.00 [1.00-1.01] $2.22\text{x}10^{-01}$	1.05 [1.05-1.06] $1.31\text{x}10^{-51}$
Inflammatory	1.07 [1.06-1.08] $3.20\text{x}10^{-57}$	1.06 [1.05-1.08] $7.09\text{x}10^{-22}$	1.09 [1.07-1.11] $3.51\text{x}10^{-15}$	1.01 [0.97-1.05] $7.00\text{x}10^{-01}$	1.01 [1.01-1.02] $1.20\text{x}10^{-06}$	1.04 [1.03-1.05] $6.72\text{x}10^{-32}$
Liver	1.03 [1.02-1.04] $1.89\text{x}10^{-11}$	1.03 [1.01-1.04] $2.06\text{x}10^{-04}$	1.04 [1.02-1.07] $3.61\text{x}10^{-04}$	1.02 [0.98-1.06] $3.73\text{x}10^{-01}$	1.00 [1.00-1.01] $1.63\text{x}10^{-01}$	1.02 [1.02-1.03] $1.42\text{x}10^{-10}$
Kidney	0.99 [0.98-0.99] $1.24\text{x}10^{-03}$	1.00 [0.98-1.01] $4.20\text{x}10^{-01}$	0.99 [0.97-1.01] $5.75\text{x}10^{-01}$	0.95 [0.91-0.99] $2.05\text{x}10^{-02}$	1.00 [1.00-1.00] $8.60\text{x}10^{-01}$	1.00 [0.99-1.01] $8.59\text{x}10^{-01}$

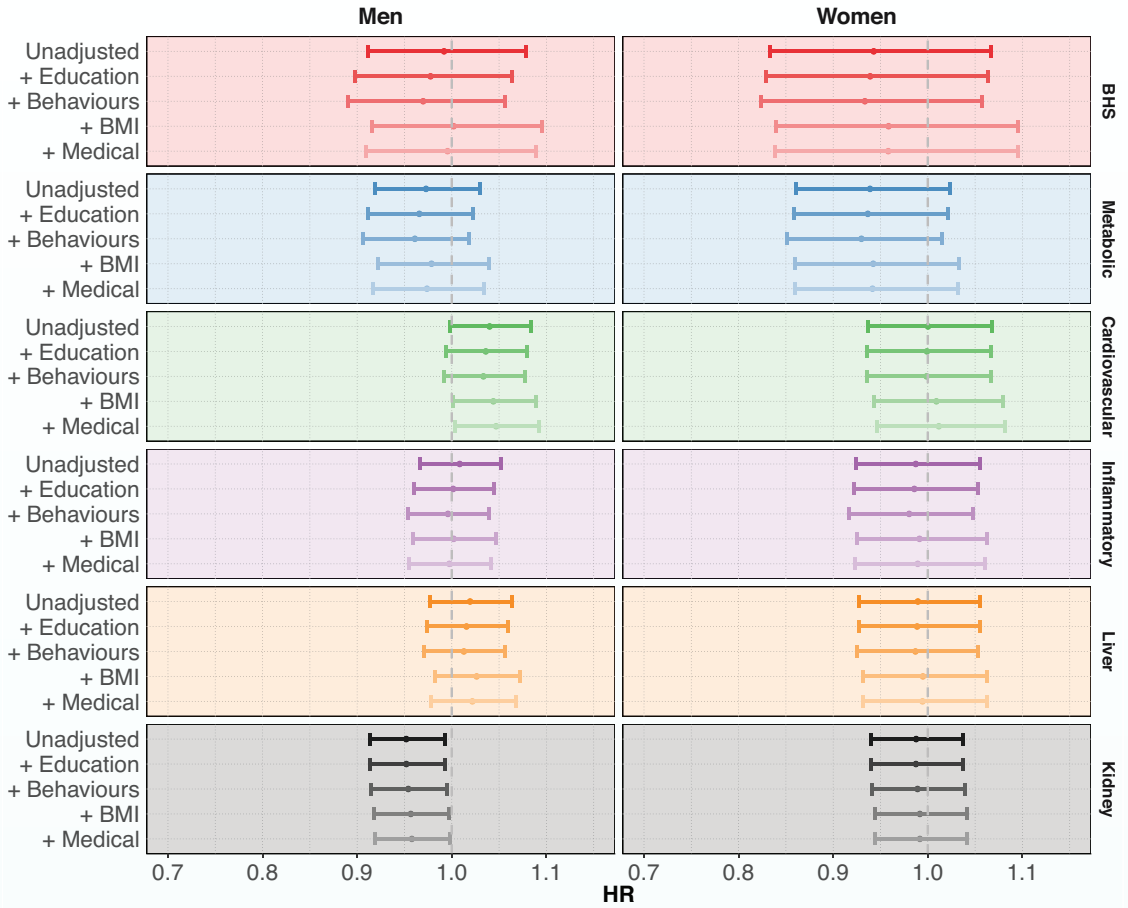
- BHS is associated to increased cancer and CVD incidence
- Weaker effects for cancer incidence: HR 1.02 ($p<10^{-4}$)
- All systems but kidney are associated to CVD incidence
- Only Metab and Inflamm systems are associated to cancer incidence

Survival analyses: Univariate Models

WOMEN	Mortality				Incidence	
	All-cause N=2,716	Cancer N=1,689	CVD N=165	External cause N=95	Cancer N=22,810	CVD N=3,728
	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value	HR [95% CI] p-value
BHS	1.09 [1.07-1.12] $8.38\text{x}10^{-16}$	1.07 [1.04-1.10] $8.54\text{x}10^{-06}$	1.21 [1.11-1.31] $1.21\text{x}10^{-05}$	0.94 [0.83-1.07] $3.51\text{x}10^{-01}$	1.02 [1.01-1.03] $1.07\text{x}10^{-05}$	1.17 [1.15-1.19] $6.84\text{x}10^{-65}$
System-specific sub-score						
Metabolic	1.04 [1.03-1.06] $3.91\text{x}10^{-08}$	1.03 [1.01-1.05] $4.44\text{x}10^{-04}$	1.18 [1.12-1.24] $9.45\text{x}10^{-10}$	0.94 [0.86-1.02] $1.54\text{x}10^{-01}$	1.01 [1.01-1.02] $9.69\text{x}10^{-06}$	1.12 [1.11-1.14] $3.56\text{x}10^{-87}$
Cardiovascular	1.03 [1.02-1.04] $1.29\text{x}10^{-06}$	1.01 [1.00-1.03] $8.94\text{x}10^{-02}$	1.08 [1.03-1.13] $1.47\text{x}10^{-03}$	1.00 [0.94-1.07] $9.96\text{x}10^{-01}$	1.01 [1.01-1.01] $7.99\text{x}10^{-06}$	1.05 [1.04-1.06] $6.72\text{x}10^{-26}$
Inflammatory	1.04 [1.03-1.05] $2.81\text{x}10^{-12}$	1.03 [1.02-1.05] $2.49\text{x}10^{-05}$	1.07 [1.02-1.12] $4.15\text{x}10^{-03}$	0.99 [0.92-1.06] $7.05\text{x}10^{-01}$	1.00 [1.00-1.01] $3.59\text{x}10^{-02}$	1.05 [1.04-1.06] $1.15\text{x}10^{-19}$
Liver	1.03 [1.02-1.04] $1.51\text{x}10^{-07}$	1.02 [1.01-1.04] $3.06\text{x}10^{-03}$	1.04 [0.99-1.09] $8.38\text{x}10^{-02}$	0.99 [0.93-1.06] $7.49\text{x}10^{-01}$	1.00 [1.00-1.01] $8.45\text{x}10^{-02}$	1.04 [1.03-1.05] $5.31\text{x}10^{-15}$
Kidney	1.00 [0.99-1.01] $9.93\text{x}10^{-01}$	1.00 [0.99-1.01] $9.05\text{x}10^{-01}$	0.98 [0.95-1.02] $4.20\text{x}10^{-01}$	0.99 [0.94-1.04] $6.23\text{x}10^{-01}$	1.00 [1.00-1.00] $9.83\text{x}10^{-01}$	1.01 [1.00-1.01] $1.16\text{x}10^{-01}$

- Similar conclusions in Women
- Weaker effects mortality than in men
- Stronger effect size estimated for incidence
- Weaker p-values in women

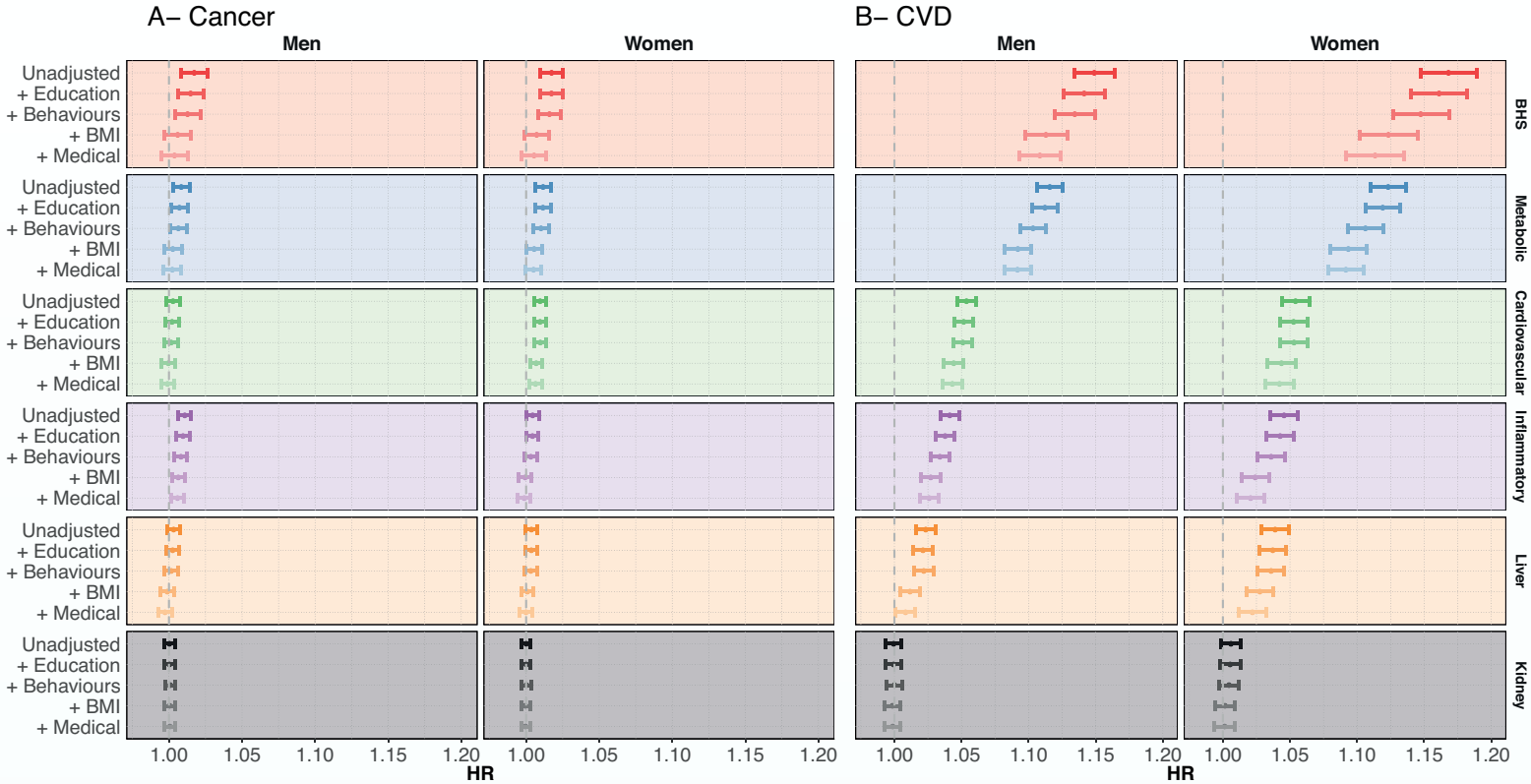
BHS and external cause mortality



- None of the BHS, or system-specific sub-scores were associated to external cause mortality, irrespective of the adjustment & gender

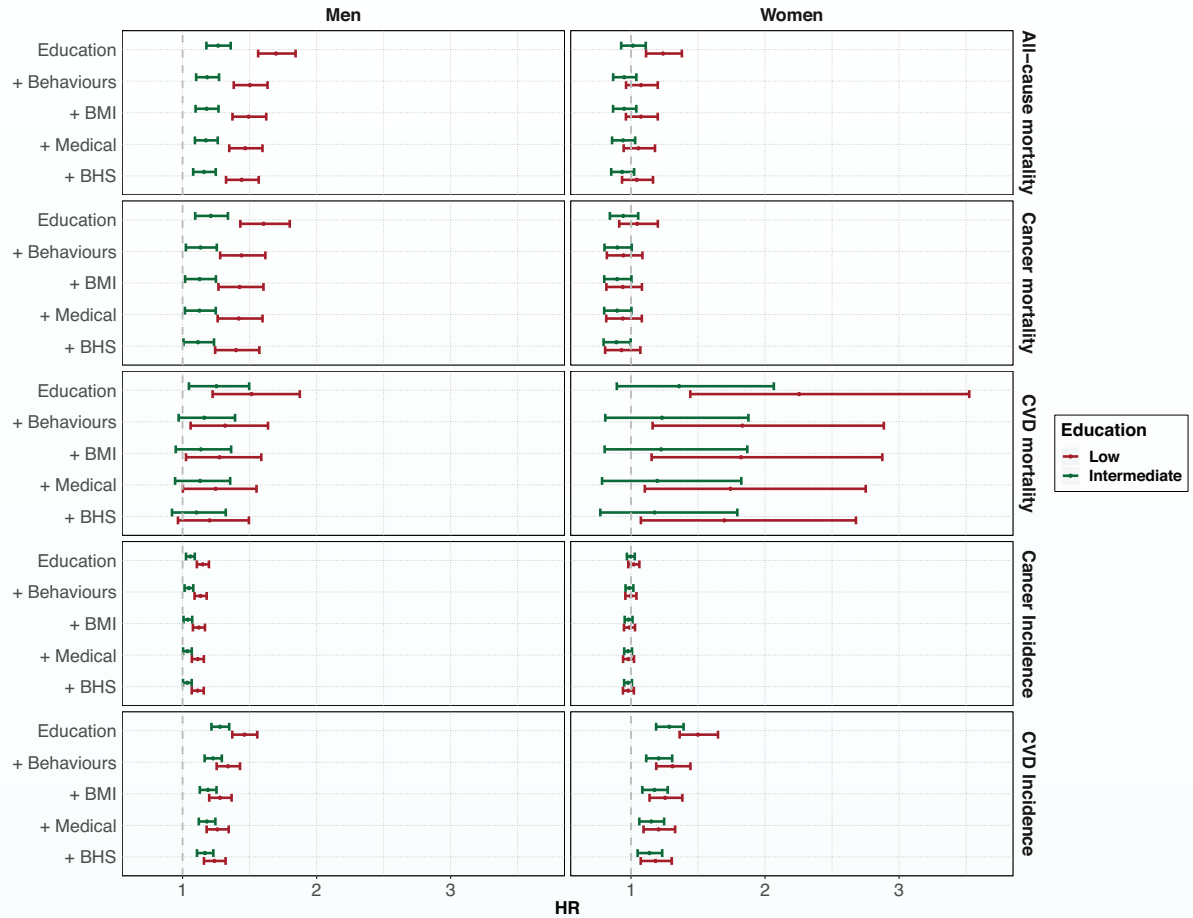
  $\Rightarrow$ External cause mortality serves as a negative control outcome

BHS and cancer & CVD incidence



- Contribution of BHS & all systems except kidney to cancer incidence
- Modest effect attenuation by education; stronger attenuation by BMI
- Fully adjusted HR for CVD incidence: HR 1.11 ($p<10^{-46}$)
- None of the scores remain associated to cancer incidence in the fully adjusted model

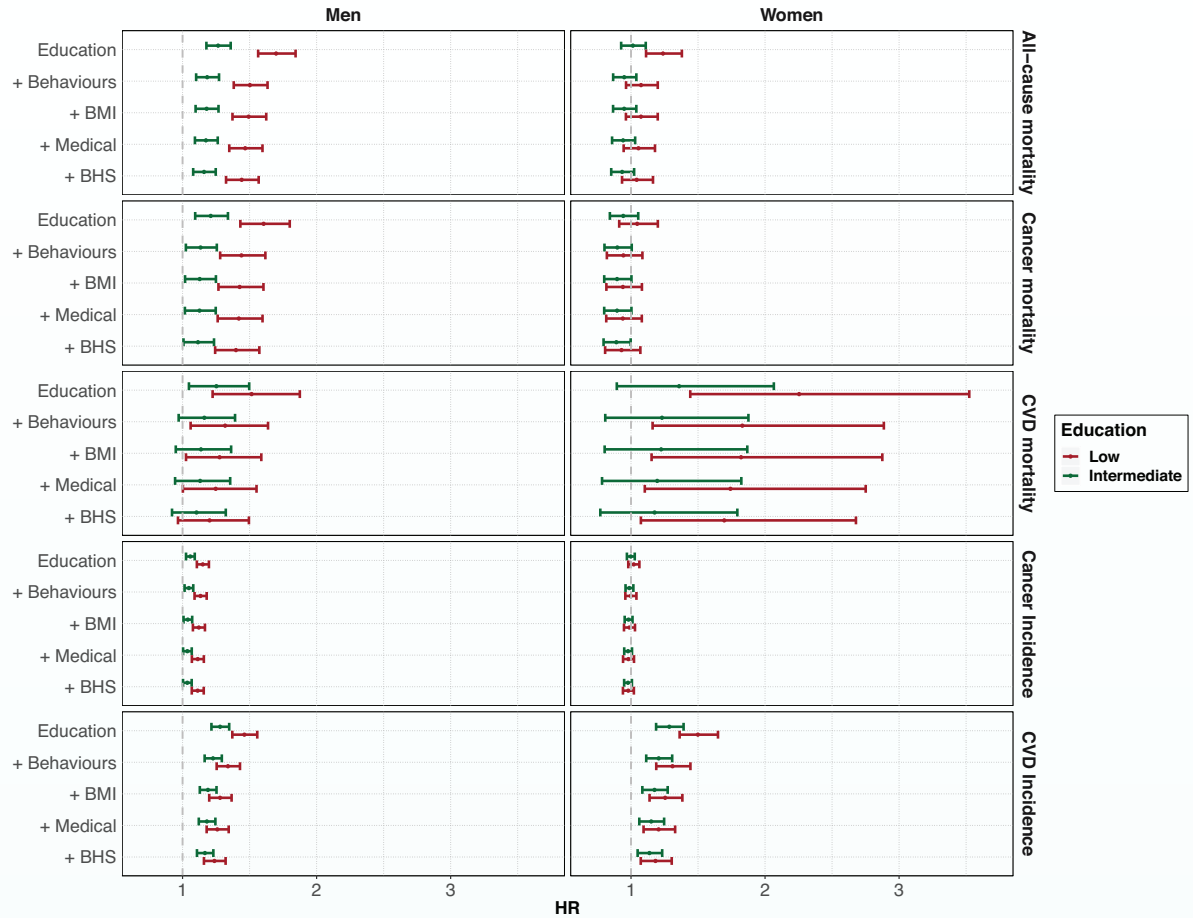
Explore the Role of Education



Education is associated with:

- Increased CVD Incidence
- All-cause and Cancer (men only)
- Cancer incidence & CVD Mortality (men only for low education)

Explore the Role of Education

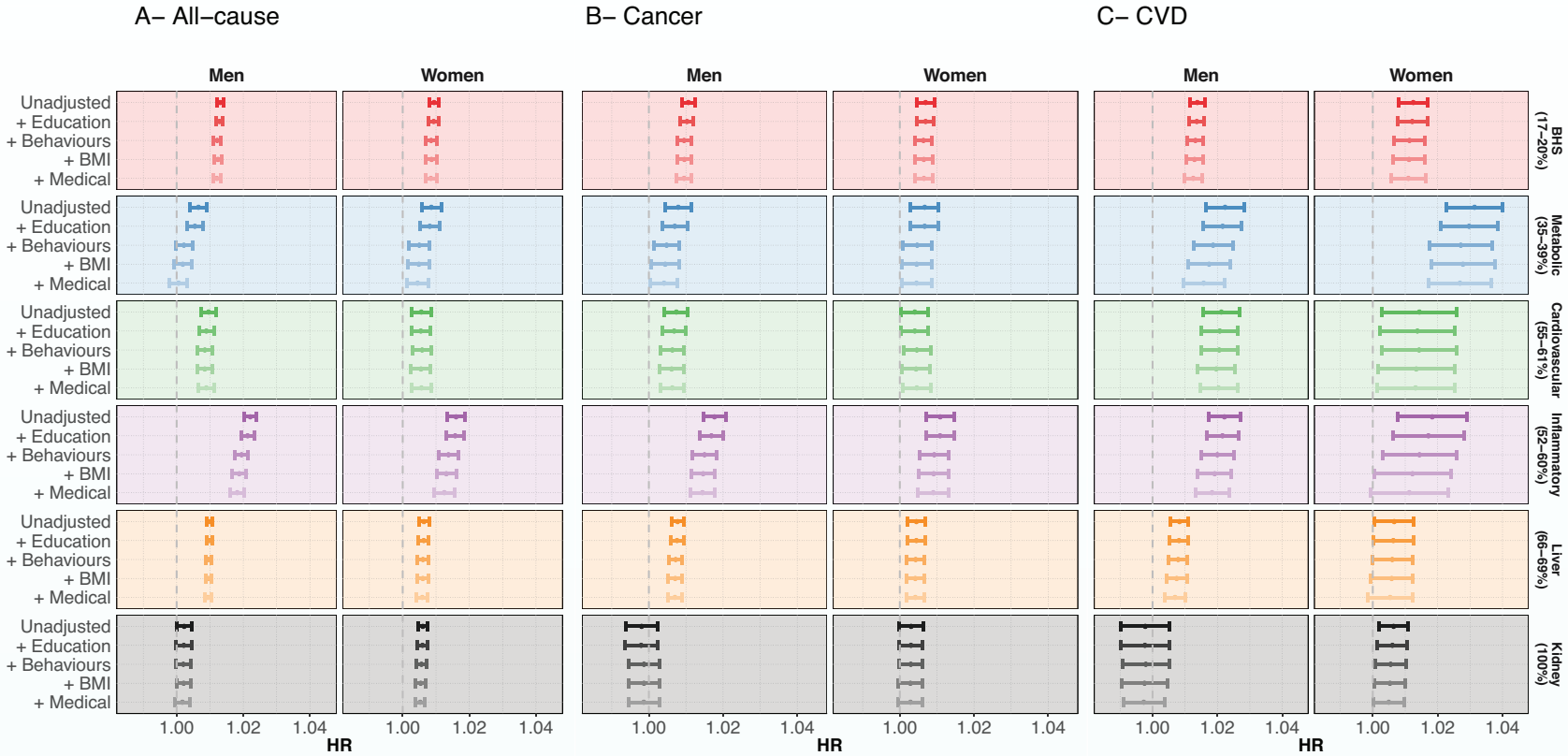


Education is associated with:

- Modest effect attenuation by BHS; stronger attenuation by behaviours
- Education is only associated with CVD incidence in both gender and all education groups the fully adjusted model

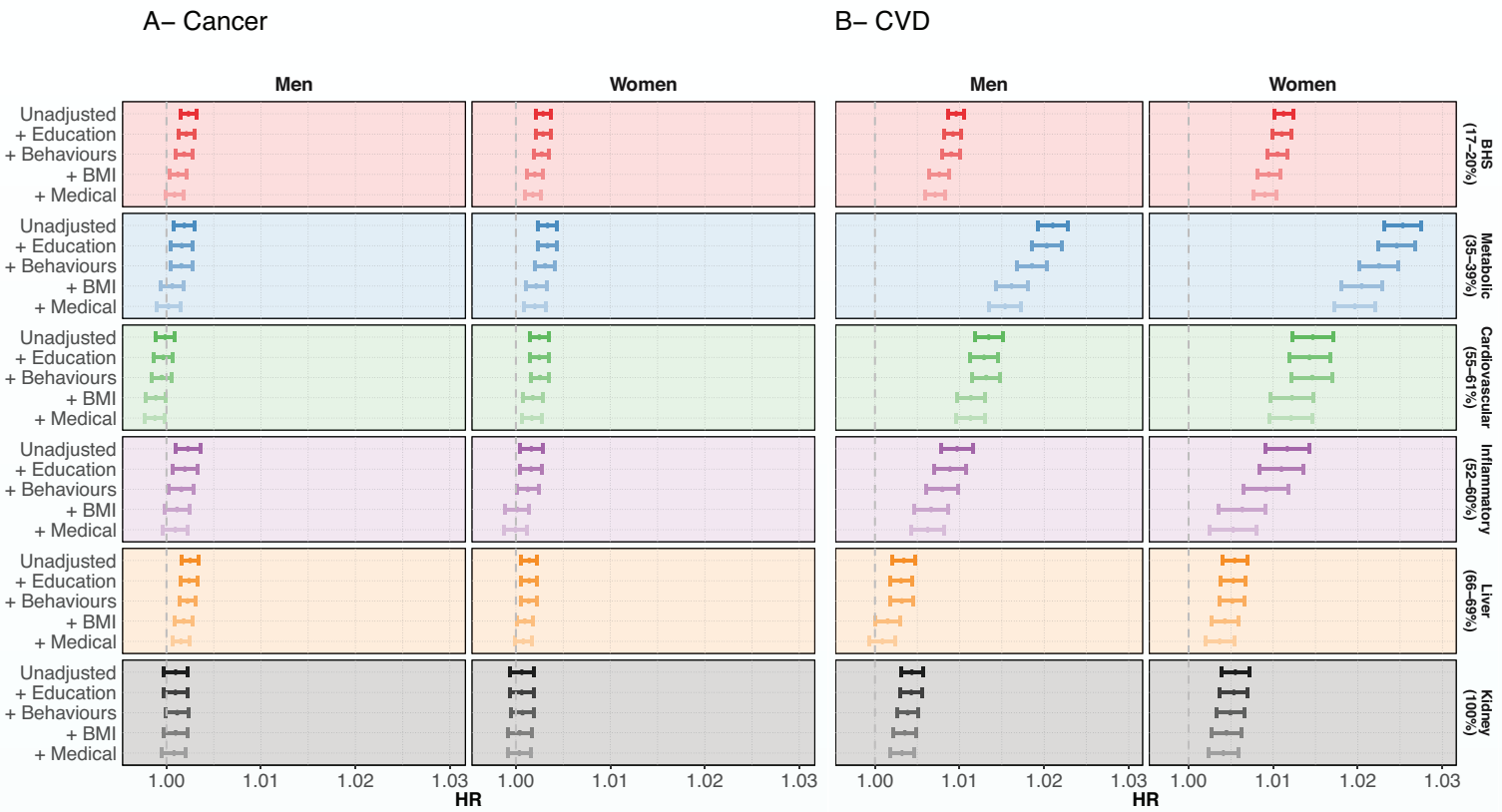
Sensitivity analyses: unsupervised score

Approach: consider the first PC from a PCA for the 13 biomarkers or all biomarkers in each system



- similar conclusions for mortality than when using the BHS

Sensitivity analyses: unsupervised score

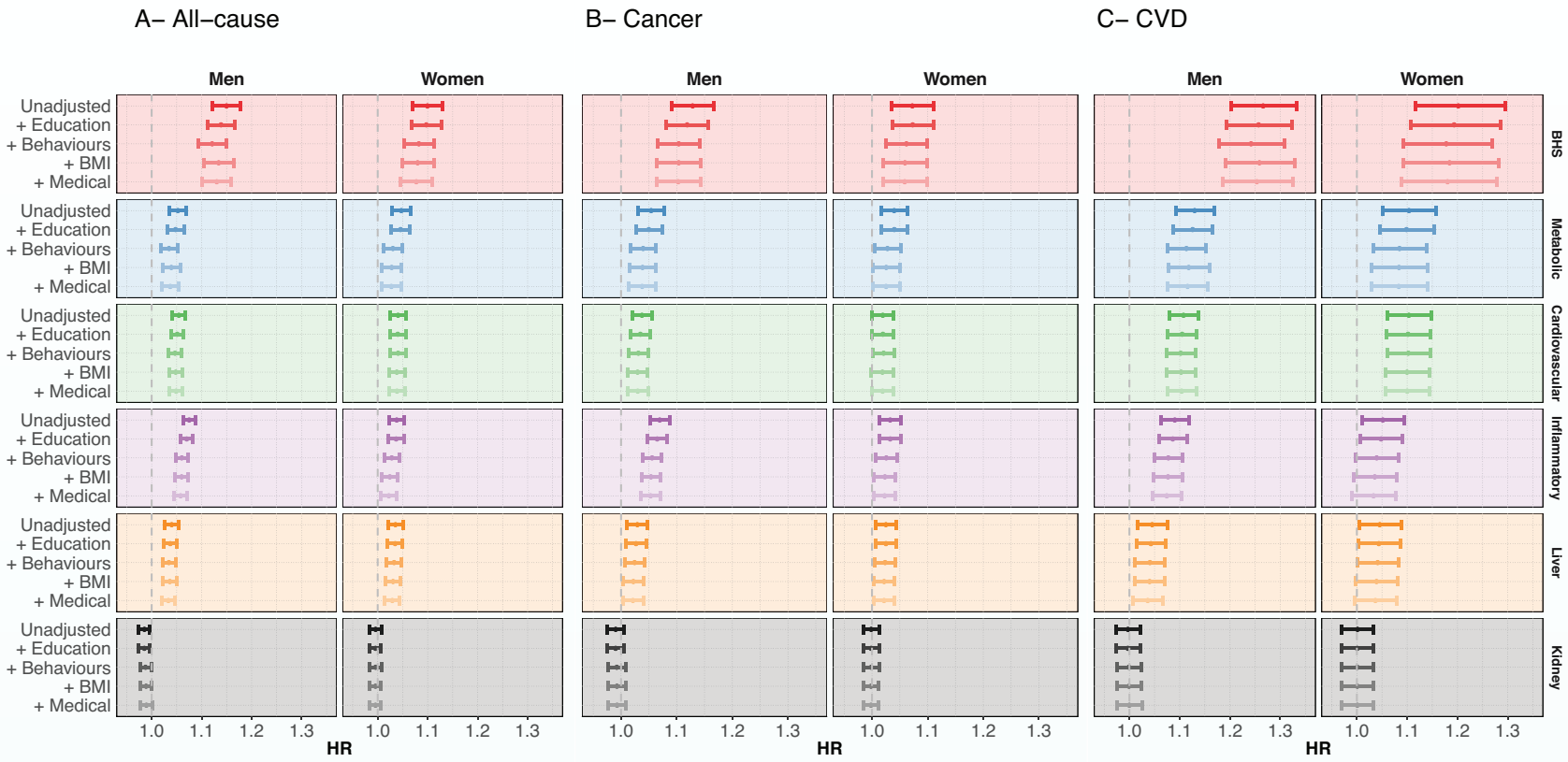


- similar conclusions for mortality than when using the BHS
- similar conclusions for incidence
- However, much smaller effect size estimates

$\Rightarrow$ possible scale effect

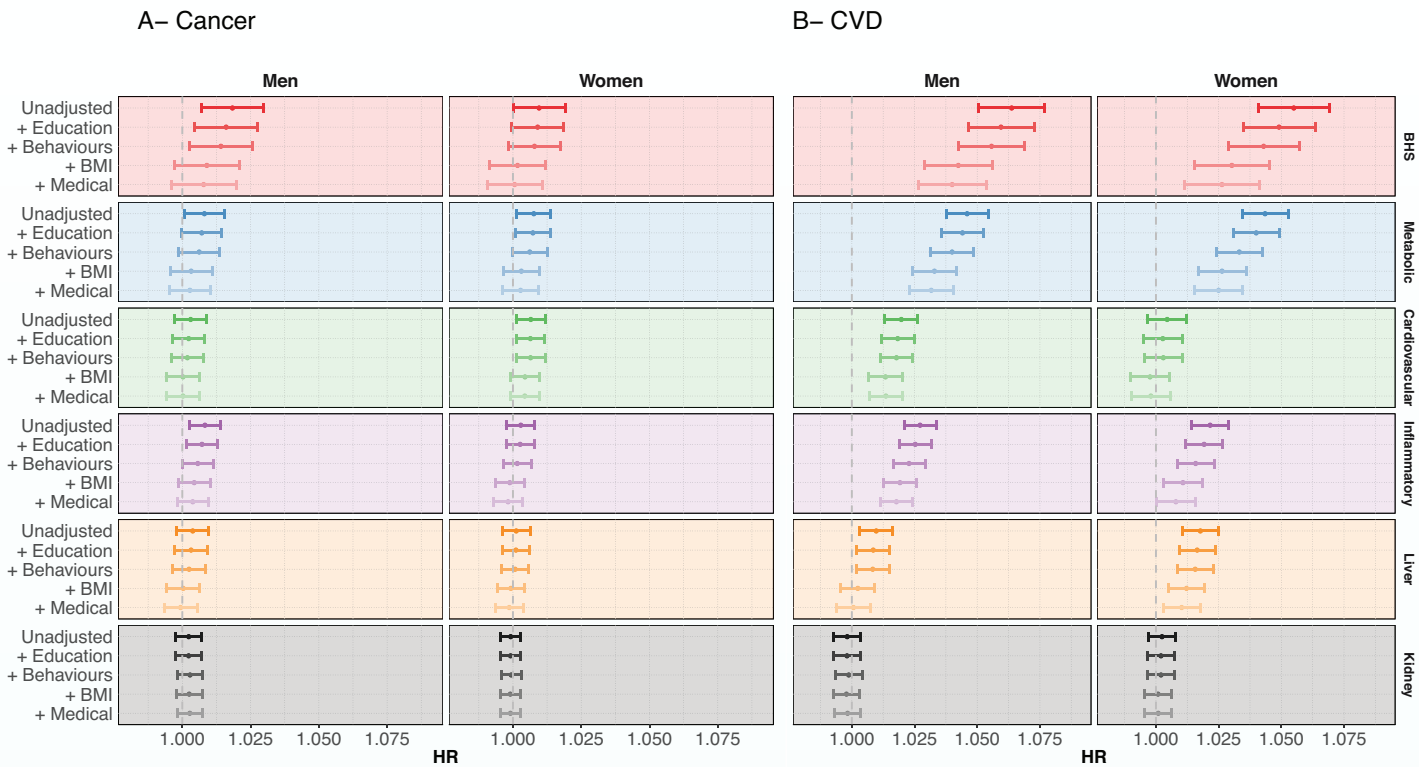
Sensitivity analyses: Revisiting CVD definition

Approach: including in CVD systemic and circulatory diseases



- Weaker associations for mortality
- Liver and Inflammatory scores are no longer associated to CVD mortality

Sensitivity analyses: Revisiting CVD definition



- As for mortality, effect on CVD mortality and incidence are weaker than when using the BHS
- Weaker results especially in women

$\Rightarrow$ BHS and subscores seem to predict better CAD, than systemic and circulatory diseases

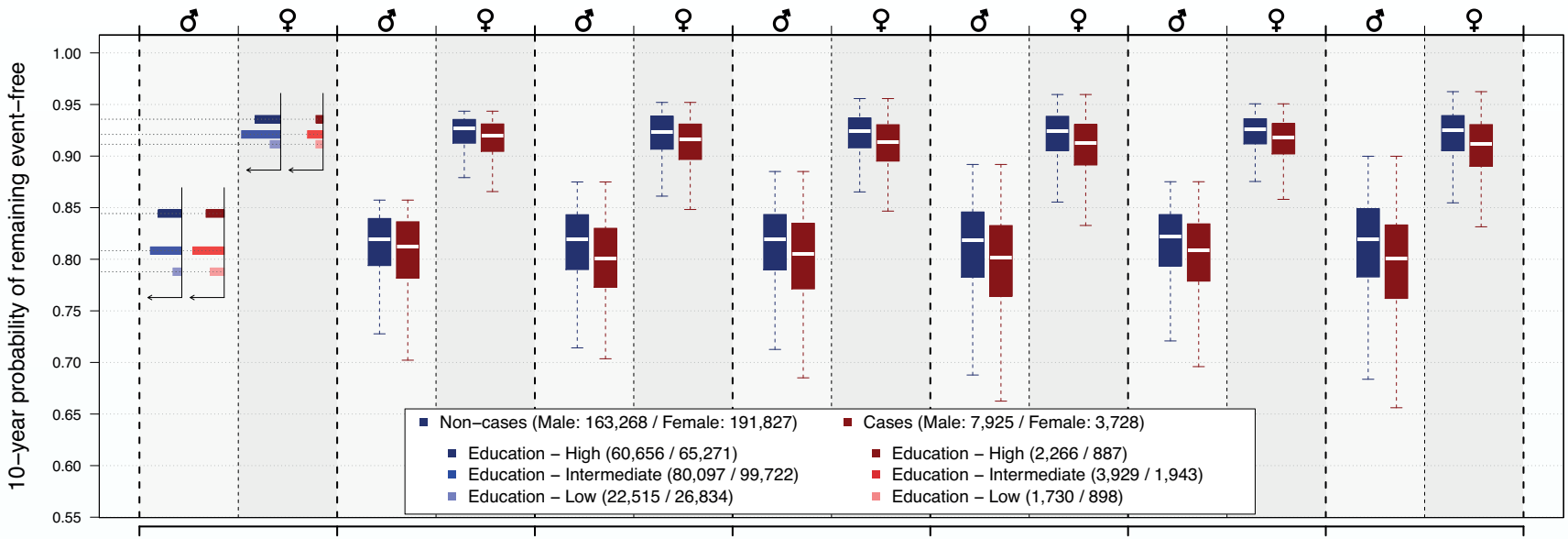
BHS and mortality & morbidity: causal assessment

- Approach: one sample, two-step least square Mendelian randomisation
  1. Identify the genetic instrument for the BHS (N=172 SNPs)
  2. Infer the instrumentally-explained BHS (2% of BHS explained)
  3. Infer the causal effect using Cox models
  4. Adjust the model for education
  5. Adopt a multivariable MR approach (including the instrumentally-explained education)
- Data: N=672,345 SNPs assayed in all 366,748 participants

	Base model		Base model + Education		Base model + $\widehat{\text{Education}}$	
	HR	p-value	HR	p-value	HR	p-value
All-cause mortality	1.03	6.09×10^{-1}	1.00	9.39×10^{-1}	1.02	7.33×10^{-1}
Cancer mortality	0.99	8.91×10^{-1}	0.96	5.61×10^{-1}	0.98	8.14×10^{-1}
CVD mortality	1.12	4.43×10^{-1}	1.11	5.16×10^{-1}	1.11	4.82×10^{-1}
Cancer incidence	1.01	6.29×10^{-1}	1.01	7.33×10^{-1}	1.01	6.68×10^{-1}
CVD incidence	1.31	3.32×10^{-11}	1.30	3.18×10^{-10}	1.30	1.23×10^{-10}

$\Rightarrow$ Results are suggestive of a causal link between BHS and CVD incidence only that is independent of education

BHS and cancer & CVD incidence



Harrell's C−statistic	Education	BHS	Health Behaviours	Education + Health Behaviours	BHS + Health Behaviours	BHS + Education	BHS + Education + Health Behaviours
Age−scale (adjusted for age) ♂	0.54 (0.53−0.54)	0.57 (0.56−0.57)	0.57 (0.57−0.58)	0.58 (0.57−0.58)	0.59 (0.58−0.59)	0.57 (0.57−0.58)	0.59 (0.58−0.60)
♀	0.54 (0.53−0.55)	0.58 (0.57−0.59)	0.59 (0.58−0.60)	0.59 (0.58−0.60)	0.60 (0.60−0.61)	0.59 (0.58−0.60)	0.61 (0.60−0.62)
Time−scale (including age) ♂	0.67 (0.66−0.67)	0.65 (0.65−0.66)	0.66 (0.65−0.67)	0.67 (0.66−0.68)	0.67 (0.67−0.68)	0.68 (0.67−0.68)	0.67 (0.66−0.68)
♀	0.69 (0.68−0.70)	0.68 (0.67−0.69)	0.68 (0.67−0.69)	0.70 (0.69−0.71)	0.70 (0.69−0.71)	0.71 (0.70−0.71)	0.70 (0.69−0.71)

- BHS is as predictive as Behaviours
- Complementarity of BHS, Education and Behaviours

⇒ Can we find biological support for this complementarity

BHS in UK Biobank: Conclusions

- Main results:
  - BHS is strongly associated to increased mortality (all-cause, cancer and CVD)
  - Effect attenuation mainly through behaviours
  - Strong effect of BHS and all sub-scores on CVD incidence surviving adjustment for socially-patterned exposures and behaviours
  - All scores except kidney are contributing to this association
  - MR is suggestive of a causal link between BHS and CVD incidence, independent of education
  - for both mortality and incidence analyses: limited role of education
  - Stronger effect of BHS than that of all subscores

$\Rightarrow$ BHS captures complementary physiological features that are disease relevant

$\Rightarrow$ such features are independent/complementary to education

BHS in UK Biobank: Conclusions

Strength & Limitations

- UK Biobank data: unique resource
- First analysis of Biological ageing in relation to mortality and morbidity
- Thorough investigation of the (lack of) role of education
- Representativity: UK Biobank suffers from Healthy Volunteer Bias
- Limited number of biomarkers (in particular for some systems)
- Outcome definitions are wide

$\Rightarrow$ investigate site-specific cancer and investigate different CVD outcomes

$\Rightarrow$ Explore common effects across outcomes

BHS in UK Biobank: Conclusions

- Interpretation:
  - An 10% increase in the BHS will increase the risk of incident CVD by 10 (over 4 years follow-up)
  - Of the (N=21,311) UK Biobank participants with 2 serial biomarker measurements, 25% (N=5,126) are exposed to that excess risk

$\Rightarrow$ Explore the mechanisms that are independent/complementary to education and disease-relevant

OMICs and composite scores: the NFBC cohort

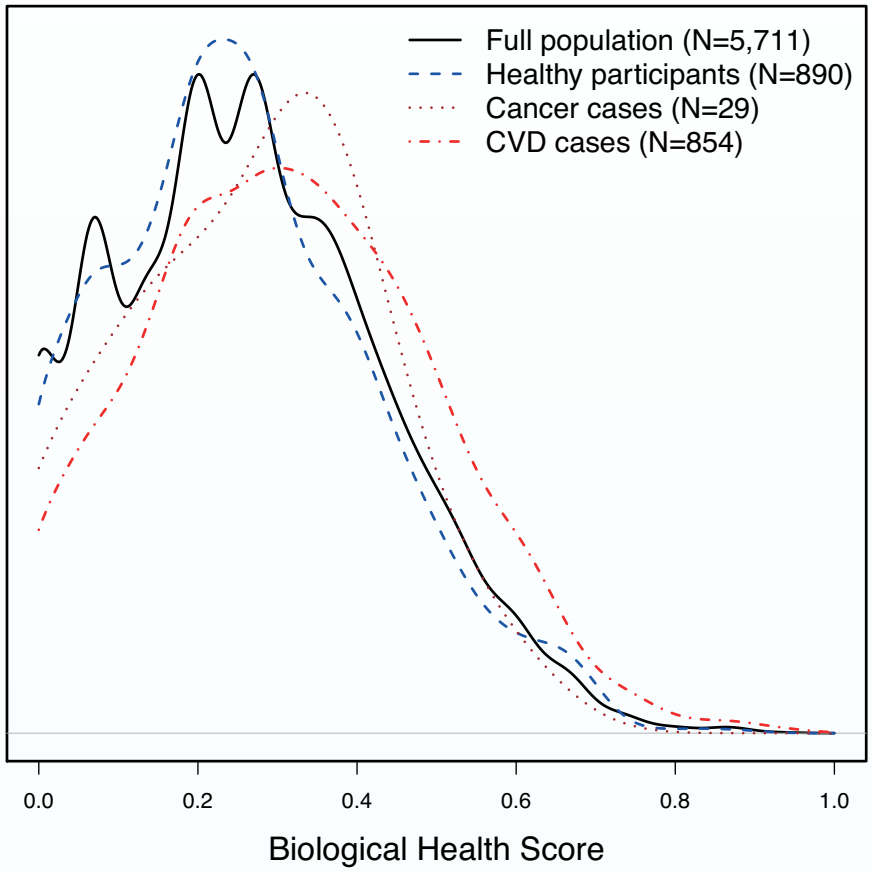
NFBC is a birth cohort (1966) including N=12,000 mother-child pairs:

- Individual characteristics including lifestyle and social factors, and comorbidities
- Biomarker measurements (including HDL, LDL and total cholesterol, triglycerides)
- Other measurements (blood pressure, pulse, spirometry)

$\Rightarrow$ 5 physiological systems assayed

  - Metabolic system: HDL cholesterol, total cholesterol, triglycerides
  - Cardiovascular system: systolic and diastolic blood pressure, pulse
  - Inflammatory/immune system: C-reactive protein, protein acetyls
  - Kidney function: creatinine
  - Liver function: albumin

The BHS in the NFBC cohort



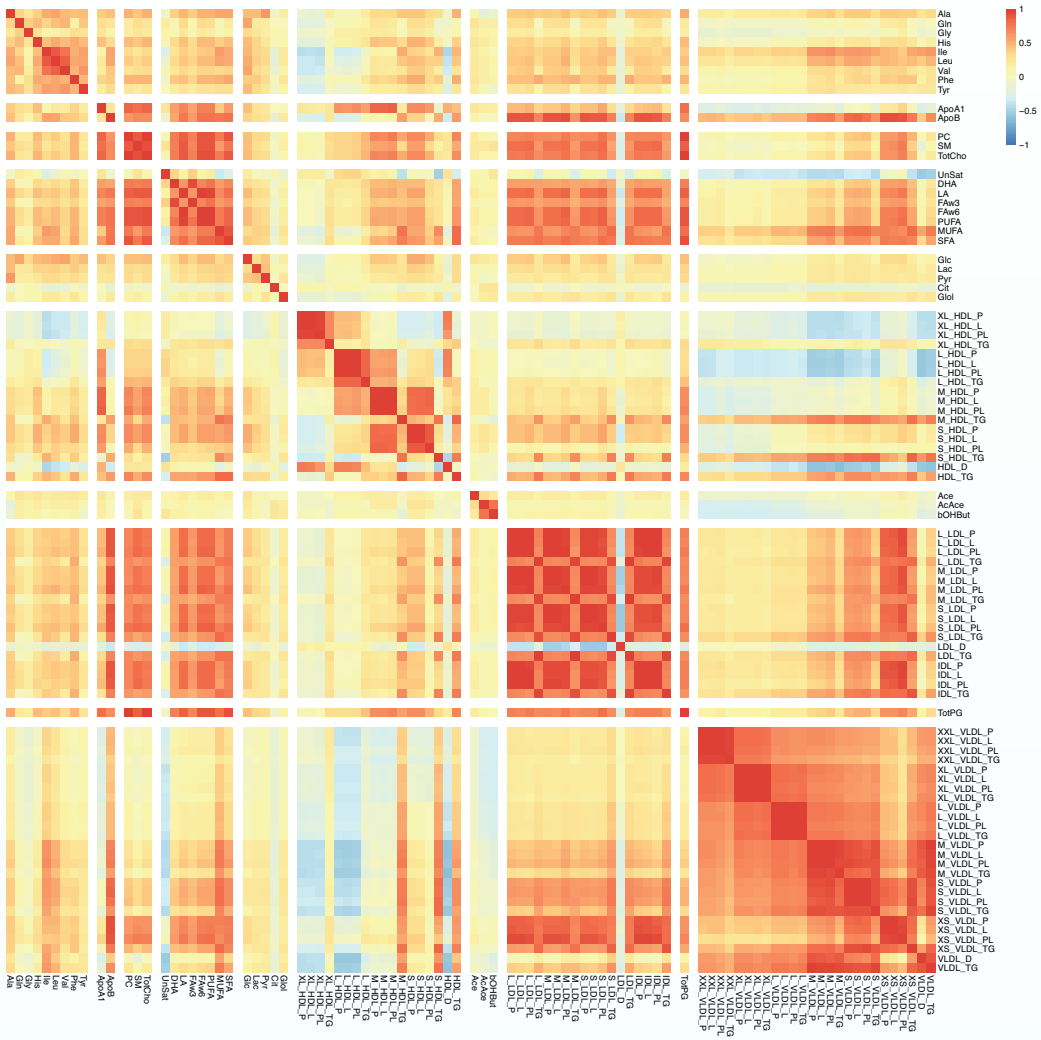
- As expected, lower score in healthy participants
- As observed in UK Biobank, higher BHS in CVD cases

Annotated NMR data in the NFBC cohort

After filtration 93 NMR Variables: 9 families and one pathway

- Amino acids (N=9)
- Apolipoproteins (N=2)
- Cholines (N=3)
- Fatty acids (N=8)
- Glycolysis and gluconeogenesis (N=5)
- Ketone bodies (N=3)
- Phosphoglycerides (N=1)
- Very low density lipoproteins (N=26)
- Low and intermediate density lipoproteins (N=18)
- High density lipoproteins (N=18)

Annotated NMR data in the NFBC cohort

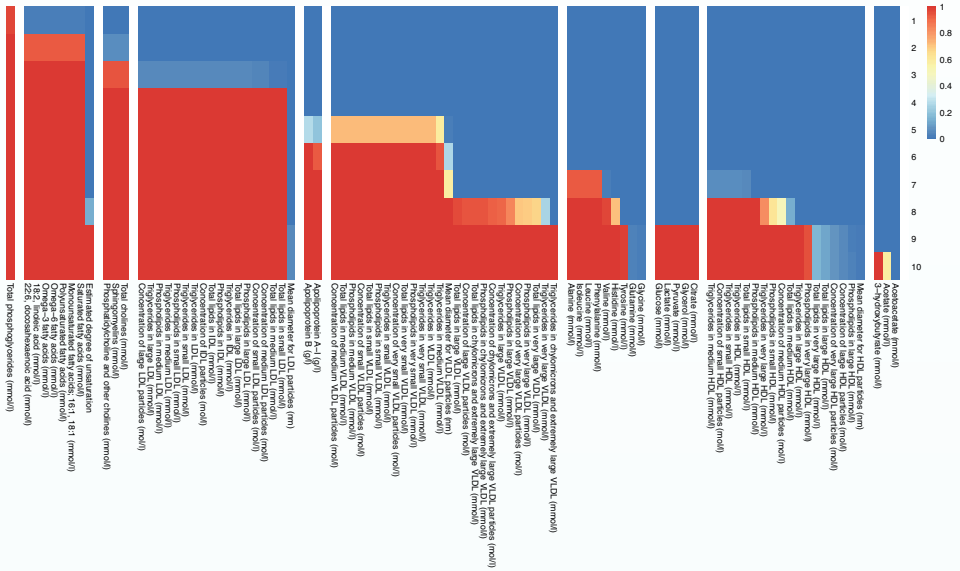


⇒ Some correlations within groups (cholines, fatty acids, LDL, VLDL)

An sgPLS model for the BHS

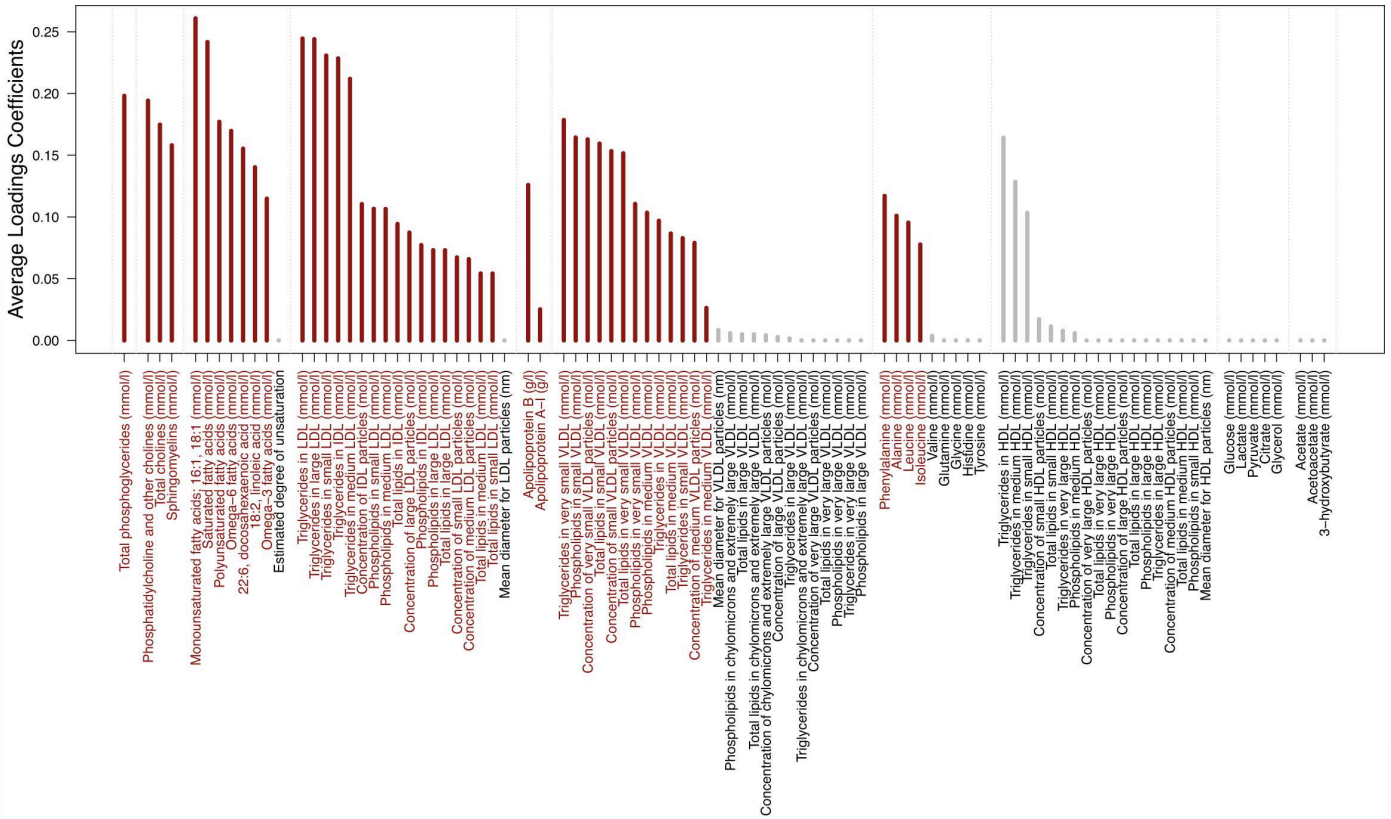
Research Questions:can we identify metabolic markers of the BHS?

- can we use prior information on the metabolites to improve interpretability → grouping
- can we select features within the groups → sparsity through penalisation



⇒ Some irrelevant features are discarded within (N=6) selected groups (VLDL, Apo AI, degree of unsaturation)

OMICs and composite scores: the NFBC cohort

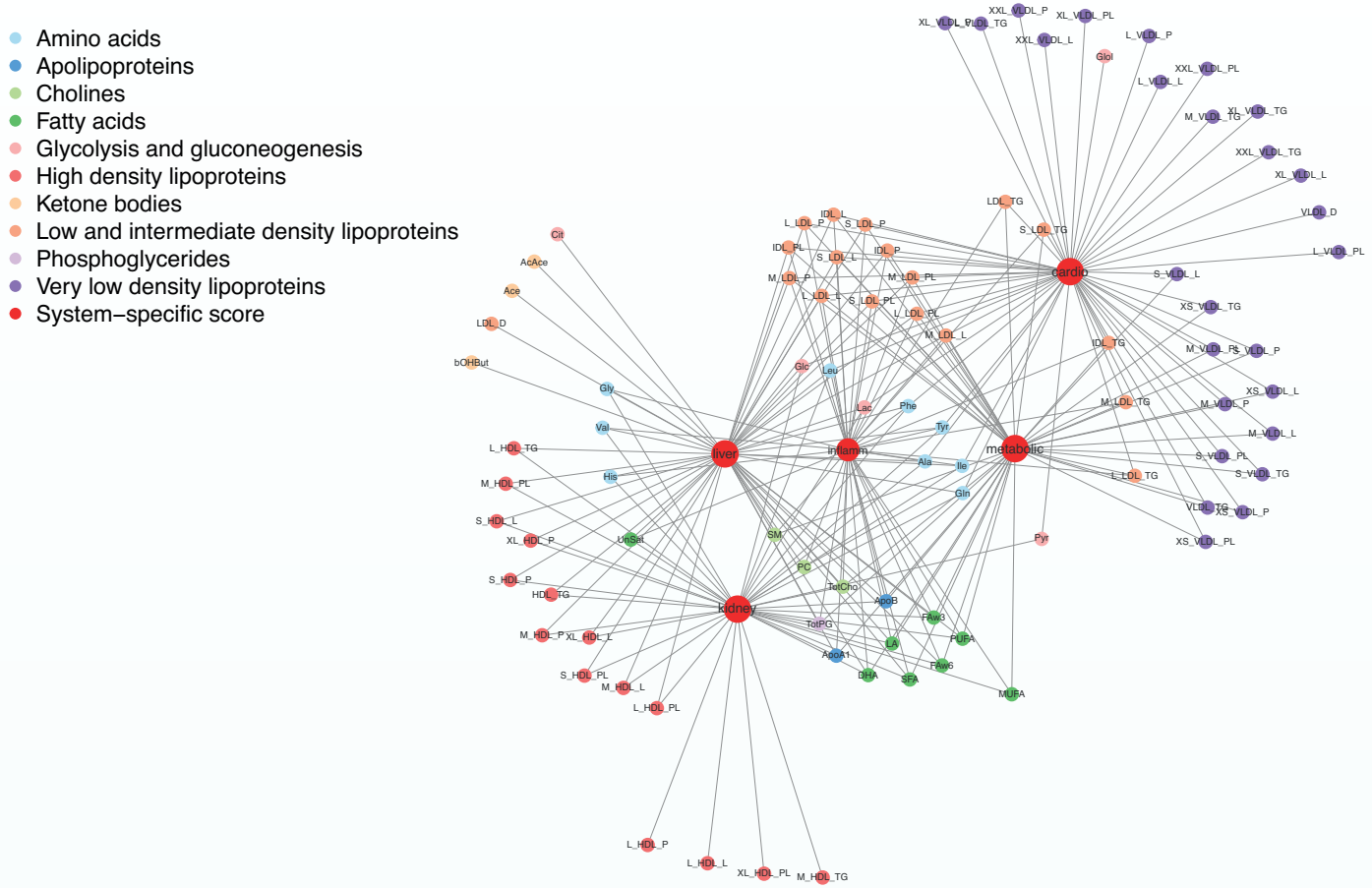


- 43 selected metabolites (from 7/10 groups)
- The group of cholesterol measurements appears redundant

$\Rightarrow$ is there a sub score differential?

OMICs and composite scores: the NFBC cohort

Question: can we identify sub-score specific patterns (sgPLS for subscores)



- Clear and functionally-relevant system specific relationships

⇒ current extensions: model scores as multivariate outcomes & multi-OMICs

Scores, risk factors and CVD prediction: UK Biobank

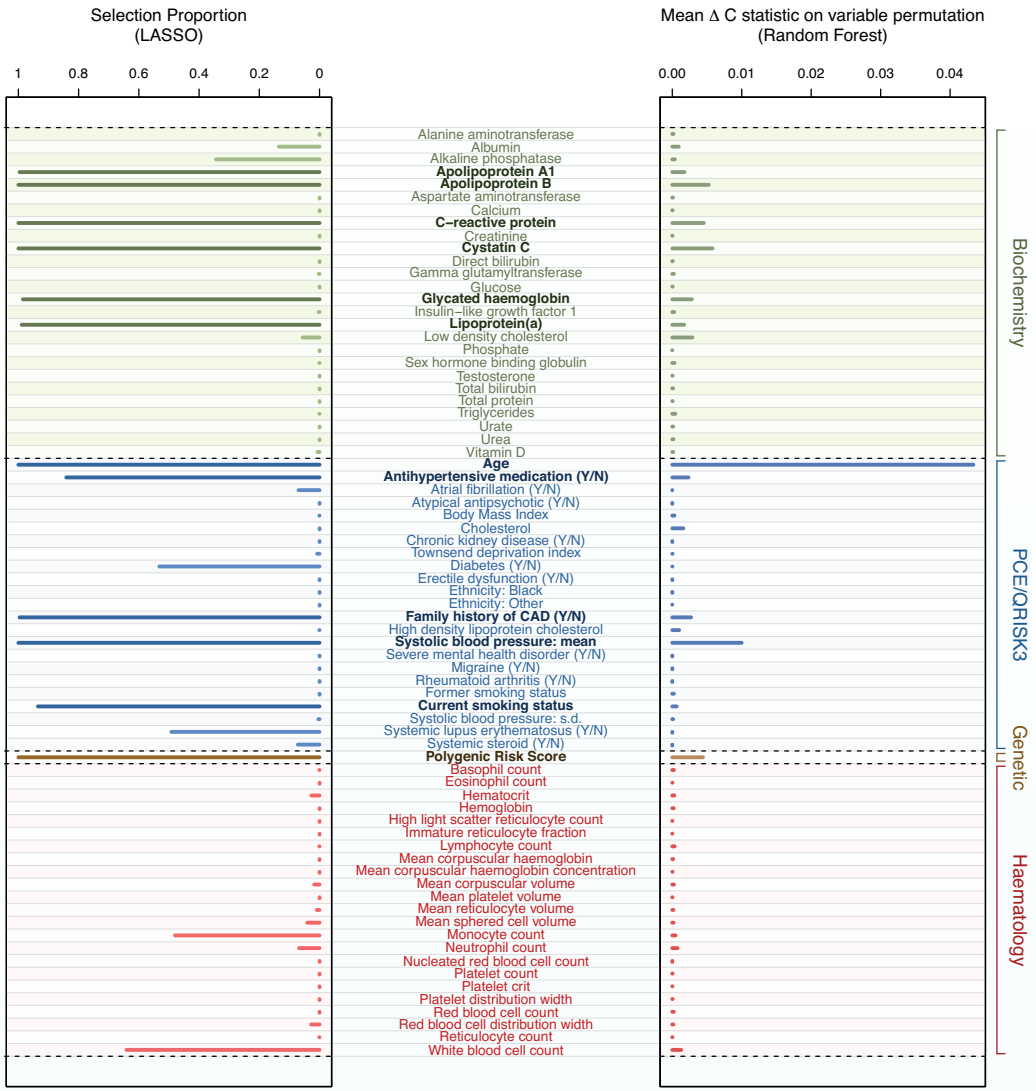
Background: Polygenic Genetic Scores incrementally improves CVD prediction (Elliott J *et al.*, 2020; JAMA)

- The Pooled Cohort Equation (PCE) isand established score for risk of CVD (C-statistics: 0.76)
- Including a bespoke and re-calibrated PRS only increases the C-statistics by 0.02

Question: can other factors (including biochemistry) improve prediction over established scores? (Elliott J *et al.*; Submitted)

- Include biochemistry biomarkers in the prediction model
- Using stability selection and Random Forrest, identify and evaluate the relative importance of the selected predictors among:
  1. Variables included in the PCE or QRISK3 algorithms (N=21)
  2. Genetic information summarised by the PRS
  3. Biochemistry biomarkers (N=26)
  4. Haematology data (N=23)

Scores, risk factors and CVD prediction: UK Biobank

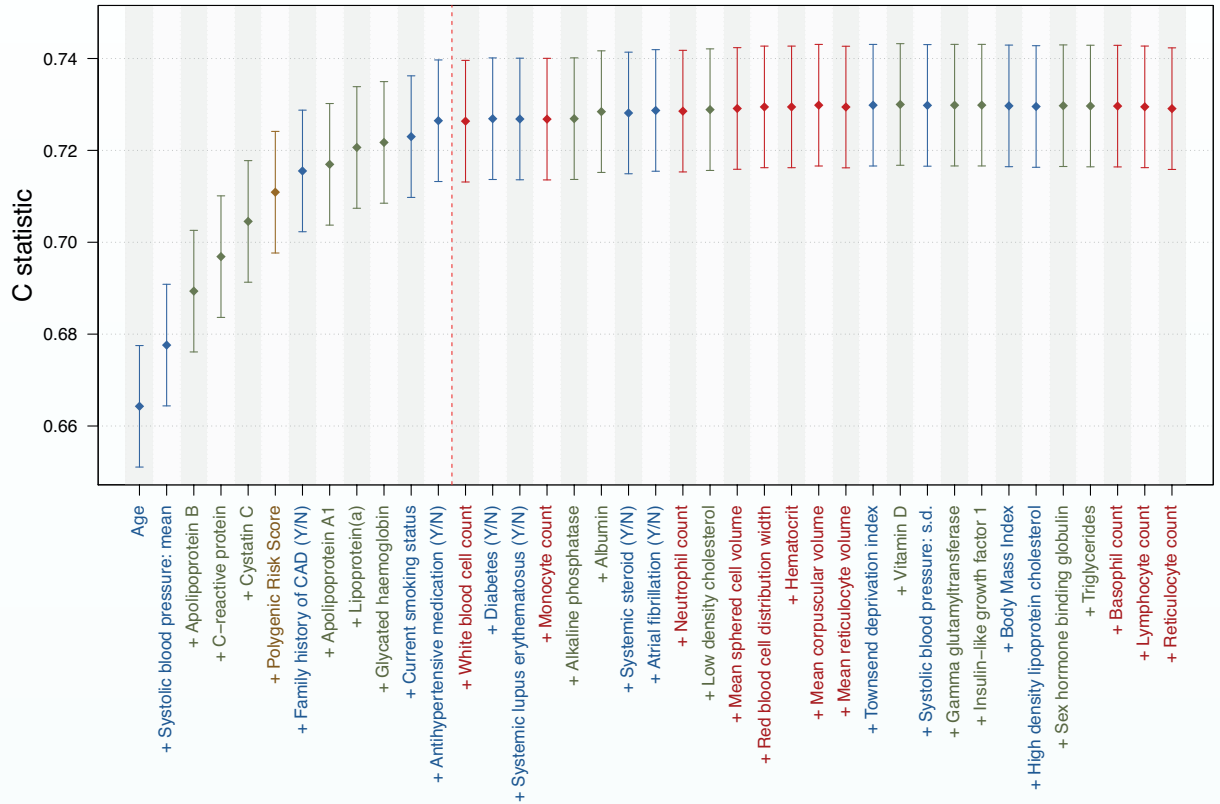


$\Rightarrow$ Identification of N=12 variables with selection proportion ≥ 0.8 in Men

$\Rightarrow$ Consistent results with the Random Forest

Scores, risk factors and CVD prediction: UK Biobank

A series of models sequentially including variables (ranked importance) are fitted in the training set and C statistics in the test set



$\Rightarrow$ Very limited increase in C statistic when including more than the calibrated number of variables

$\Rightarrow$ However: biochemistry adds to the prediction

Scores, risk factors and CVD prediction: UK Biobank

Evaluation of prediction performances: The C statistics in the test set

	Men (1,885/34,694)		Women (1,163/53,712)		Full population (3,048/88,406)
	#	C stat (95% CI)	#	C stat (95% CI)	C stat (95% CI)
PCE		0.732 [0.721-0.742]		0.684 [0.671-0.698]	0.713 [0.696-0.730]
Stability selection	12	0.726 [0.713-0.740]	11	0.745 [0.728-0.762]	0.762 [0.752-0.773]
Random Forest	16	0.727 [0.714-0.740]	13	0.739 [0.722-0.756]	0.760 [0.749-0.770]

$\Rightarrow$ Increase in C statistic of 0.05 for the full population when including other covariates from the selected groups

$\Rightarrow$ of this only up to 2% can be attributed to PRS

$\Rightarrow$ need to include other factors including environmental and other molecular data

$\Rightarrow$ identify other data sources with (multi-) OMICs data $\Rightarrow$ need to focus on finer outcomes

Ongoing Questions:

Summary from the BHS analyses; Composite scores are

- capturing (biological, social) gradients in the population
- explanatory of incident conditions and mortality
- complementary to established factors

Ensuing research questions:

- Can we refine the scores such that they include Exposome features?

  $\Rightarrow$ Clustering approaches

- Are and how much are (constituents of) scores complementing established risk factors?

  $\Rightarrow$ variable selection and prediction models including Expotyes &/or their constituents

- Can we elucidate causal links between scores, their drivers and health outcomes?

  $\Rightarrow$ Mediation & Causal models

Acknowledgments

Imperial College London

- B Bodinier
- D Vuckovic
- V Zuber

Inserm
La science pour la santé
From science to health

- C Delpierre
- M Kelly-Irving
- R Castagné

Universiteit Utrecht

- R Vermeulen
- L Portengen
- J Vlaanderen

- Imperial Team: P Elliott, P Vineis, A Berlanga, S Bowden, S Dagnino, A Pengelly, M Whittaker, M Hedges, J Elliott, D Petrovic, N De Toro Eadie, T Wright, K Asamoah

- EU Financial Support:

Lifepath HEALTHY AGEING FOR ALL

EXPANSE

longITools HEALTH & ENVIRONMENT DYNAMICS