

Supplementary Material

Causal attribution fractions for epidemiological studies, applied to a UK Biobank study of smoking and BMI

A.J. Webster

December 24, 2021

Sensitivity analysis

As discussed in the main text, participants with hospital reports of prior cancers other than non-melanoma skin cancers were excluded from the main study, but self-reported cancers or other prior diseases were not. A sensitivity analysis compared the paper's main results with a second analysis that excluded participants who had reported any cancer other than non-melanoma skin cancer, or any serious cardiovascular disease from heart disease, stroke, arterial or pulmonary embolisms, or subarachnoid haemorrhage. As discussed in the main text, depending on the disease and exposure, the extra exclusions can in principle either weaken or strengthen associations.

The tables are equivalent to the main text, but with values calculated from the sensitivity analysis, allowing estimates to be compared.

The figures are from equivalent analyses to those in the main text. The main differences are driven by differences in the number of statistically significant associations with BMI or smoking, that led to small differences in the numbers of diseases that are summarised in the figures.

Disease	Sex	N	N_{A_f}	A_f	Old rank A_f
E66 Obesity	F	273	268	0.98	1
J44.1 Chronic obstructive pulmonary disease with acute exacerbation, unspecified	F	166	152	0.92	2
J44.0 Chronic obstructive pulmonary disease with acute lower respiratory infection	F	345	306	0.89	3
J44.0 Chronic obstructive pulmonary disease with acute lower respiratory infection	M	337	298	0.88	4
J44.1 Chronic obstructive pulmonary disease with acute exacerbation, unspecified	M	209	184	0.88	5
C34 Malignant neoplasm of bronchus and lung	M	849	687	0.81	6
I70 Atherosclerosis	M	139	100	0.72	7
C34 Malignant neoplasm of bronchus and lung	F	848	609	0.72	8
E11 Non-insulin-dependent diabetes mellitus	M	174	97	0.56	9
I71 Aortic aneurysm and dissection	M	361	194	0.54	10
R04.2 Haemoptysis	M	279	146	0.52	11
C15 Malignant neoplasm of oesophagus	M	419	195	0.47	12
R91 Abnormal findings on diagnostic imaging of lung	M	309	134	0.43	13
C67 Malignant neoplasm of bladder	M	921	365	0.40	14
R91 Abnormal findings on diagnostic imaging of lung	F	323	130	0.40	15
I50 Heart failure	F	239	115	0.48	16
K42 Umbilical hernia	F	282	104	0.37	17
G47.3 Sleep apnoea	F	352	122	0.35	18
J84 Other interstitial pulmonary diseases	F	161	56	0.35	19
I50 Heart failure	M	305	100	0.33	20
R04.2 Haemoptysis	F	222	76	0.34	21
J10 Influenza due to identified influenza virus	F	181	49	0.27	22
M13 Other arthritis	M	204	67	0.33	23
R29.6 Tendency to fall, not elsewhere classified	M	147	61	0.41	24
A41 Other septicaemia	F	812	240	0.30	25
J18 Pneumonia, organism unspecified	M	2505	790	0.32	26
C22 Malignant neoplasm of liver and intrahepatic bile ducts	M	144	40	0.28	27
L72.0 Epidermal cyst	M	401	121	0.30	28
J18 Pneumonia, organism unspecified	F	2400	705	0.29	29
G47.3 Sleep apnoea	M	688	216	0.31	30
N17 Acute renal failure	F	282	73	0.26	31
G44.2 Tension-type headache	F	138	41	0.30	32
K43 Ventral hernia	F	342	92	0.27	33
G62 Other polyneuropathies	M	148	44	0.30	34
C16 Malignant neoplasm of stomach	M	229	47	0.21	35
B37 Candidiasis	M	145	44	0.30	36
R06.0 Dyspnoea	M	555	128	0.23	37
I26 Pulmonary embolism	F	756	221	0.29	38
R63.4 Abnormal weight loss	F	441	121	0.27	39
J84 Other interstitial pulmonary diseases	M	208	57	0.27	40
J22 Unspecified acute lower respiratory infection	F	1263	283	0.22	41
E87.1 Hypo-osmolality and hyponatraemia	M	193	43	0.22	42
I25.9 Chronic ischaemic heart disease, unspecified	M	231	65	0.28	43
M48 Other spondylopathies	F	525	140	0.27	44
K62.1 Rectal polyp	M	1047	248	0.24	45
N17 Acute renal failure	M	434	72	0.17	46
H02.4 Ptosis of eyelid	M	225	55	0.25	47
C90 Multiple myeloma and malignant plasma cell neoplasms	F	229	58	0.25	48
M47 Spondylosis	M	293	72	0.25	49
L60 Nail disorders	F	205	47	0.23	50
R13 Dysphagia	M	880	180	0.21	51
J90 Pleural effusion, not elsewhere classified	M	443	96	0.22	52
J22 Unspecified acute lower respiratory infection	M	1117	187	0.17	53
L03 Cellulitis	M	1766	362	0.20	54
M81 Osteoporosis without pathological fracture	F	849	181	0.21	55
K92.0 Haematemesis	M	126	22	0.18	56
L03 Cellulitis	F	1413	292	0.21	57
K25 Gastric ulcer	M	300	54	0.18	58
I64 Stroke, not specified as haemorrhage or infarction	M	158	28	0.18	59

Table 1: Table 1 of the main text is recalculated for the sensitivity analysis using Eq. 19. Colours: $A_f \geq 0.5$ (red), $0.5 > A_f \geq 0.35$ (orange), $0.35 > A_f \geq 0.2$ (yellow). Sex: diseases in males (M) or females (F), N : total cases, N_{A_f} : attributed cases.

Disease	Sex	N	N_{A_f}	A_f	Old rank	A_f
E66 Obesity	F	273	267	0.98	1	
E11 Non-insulin-dependent diabetes mellitus	M	174	79	0.45	2	
K42 Umbilical hernia	F	282	111	0.39	3	
G47.3 Sleep apnoea	F	352	107	0.30	4	
I50 Heart failure	M	305	83	0.27	5	
G62 Other polyneuropathies	M	148	41	0.27	6	
M13 Other arthritis	M	204	55	0.27	7	
I50 Heart failure	F	239	81	0.34	8	
I26 Pulmonary embolism	F	756	193	0.26	9	
G47.3 Sleep apnoea	M	688	158	0.23	10	
J22 Unspecified acute lower respiratory infection	F	1263	260	0.21	11	
I25.9 Chronic ischaemic heart disease, unspecified	M	231	52	0.22	13	
M81 Osteoporosis without pathological fracture	F	849	168	0.20	14	

Table 2: BMI only: Table 4 of the main text is recalculated for the sensitivity analysis using Eq. 19. Sex indicates diseases in males (M) or females (F), N are total cases, N_{A_f} are cases attributed to BMI, A_f is the attributable fraction for deviations from the mid-tertile of BMI. For obesity, unsurprisingly $A_f \simeq 1$. Reporting errors may have prevented $A_f = 1$ for obesity.

Disease	Sex	N	N_{A_f}	A_f	Old rank	A_f
J44.1 Chronic obstructive pulmonary disease with acute exacerbation, unspecified	F	166	150	0.90	1	
J44.0 Chronic obstructive pulmonary disease with acute lower respiratory infection	F	345	301	0.87	2	
J44.0 Chronic obstructive pulmonary disease with acute lower respiratory infection	M	337	288	0.86	3	
J44.1 Chronic obstructive pulmonary disease with acute exacerbation, unspecified	M	209	180	0.86	4	
C34 Malignant neoplasm of bronchus and lung	M	849	668	0.79	5	
I70 Atherosclerosis	M	139	102	0.73	6	
C34 Malignant neoplasm of bronchus and lung	F	848	599	0.71	7	
I71 Aortic aneurysm and dissection	M	361	200	0.55	8	
R04.2 Haemoptysis	M	279	114	0.41	9	
C15 Malignant neoplasm of oesophagus	M	419	177	0.42	10	
R91 Abnormal findings on diagnostic imaging of lung	M	309	121	0.39	11	
C67 Malignant neoplasm of bladder	M	921	329	0.36	12	
R91 Abnormal findings on diagnostic imaging of lung	F	323	110	0.34	13	
J84 Other interstitial pulmonary diseases	F	161	47	0.29	14	
B37 Candidiasis	M	145	44	0.30	15	
J84 Other interstitial pulmonary diseases	M	208	62	0.30	16	
J90 Pleural effusion, not elsewhere classified	M	443	110	0.25	17	
R04.2 Haemoptysis	F	222	55	0.25	18	
G56.2 Lesion of ulnar nerve	F	200	44	0.22	19	
K92.0 Haematemesis	M	126	25	0.20	20	
C22 Malignant neoplasm of liver and intrahepatic bile ducts	M	144	33	0.23	21	
R29.6 Tendency to fall, not elsewhere classified	M	147	37	0.25	22	
R06.0 Dyspnoea	M	555	110	0.20	23	
J18 Pneumonia, organism unspecified	M	2505	539	0.22	24	
C16 Malignant neoplasm of stomach	M	229	44	0.19	25	
H02.0 Entropion and trichiasis of eyelid	M	290	59	0.20	26	
K62.1 Rectal polyp	M	1047	209	0.20	27	
G44.2 Tension-type headache	F	138	24	0.17	28	

Table 3: Smoking only: Table 4 of the main text is recalculated for the sensitivity analysis using Eq. 19. Sex indicates diseases in males (M) or females (F), N are total cases, N_{A_f} are cases attributed to smoking, A_f is the attributable fraction.

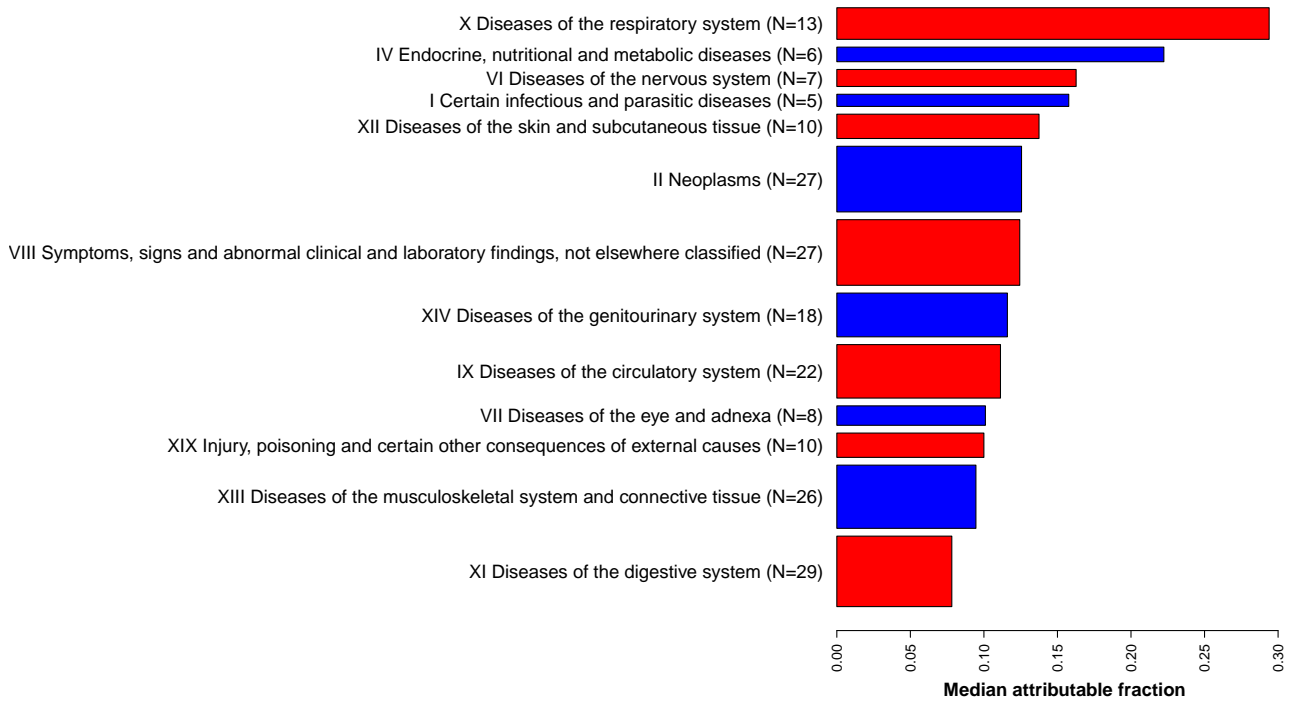


Figure 1: Median attributable fractions for each ICD-10 chapter with at least 5 diseases where $A_f > 0.2$. Bar widths are proportional to the number of diseases in each chapter.

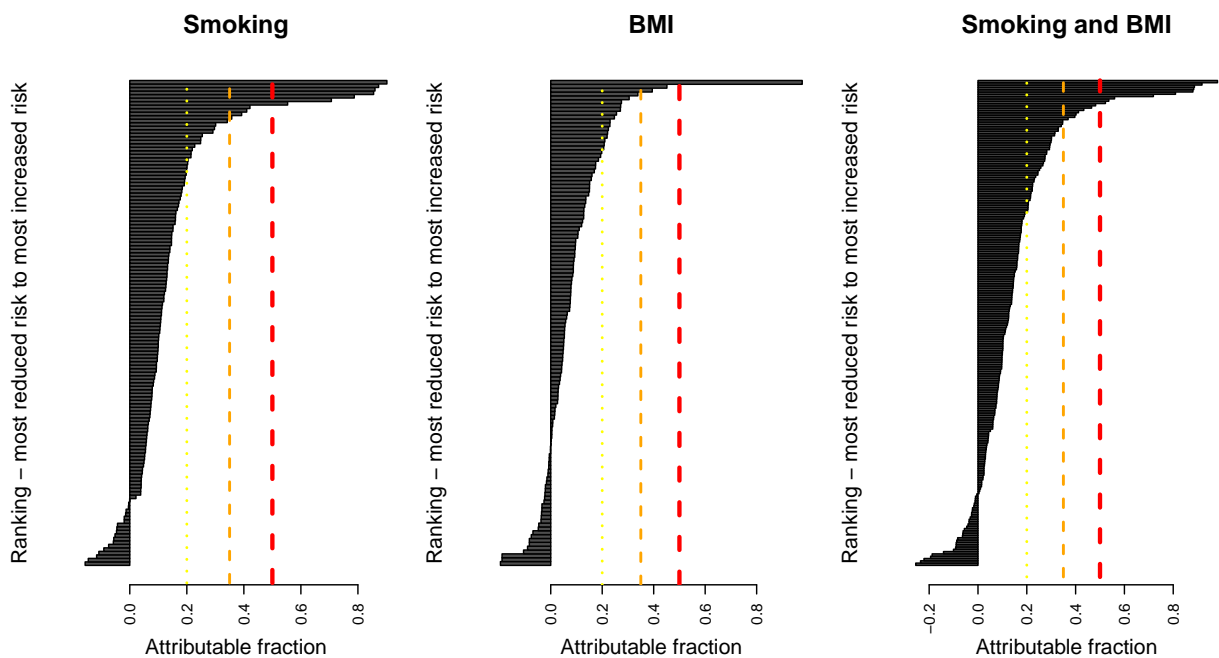
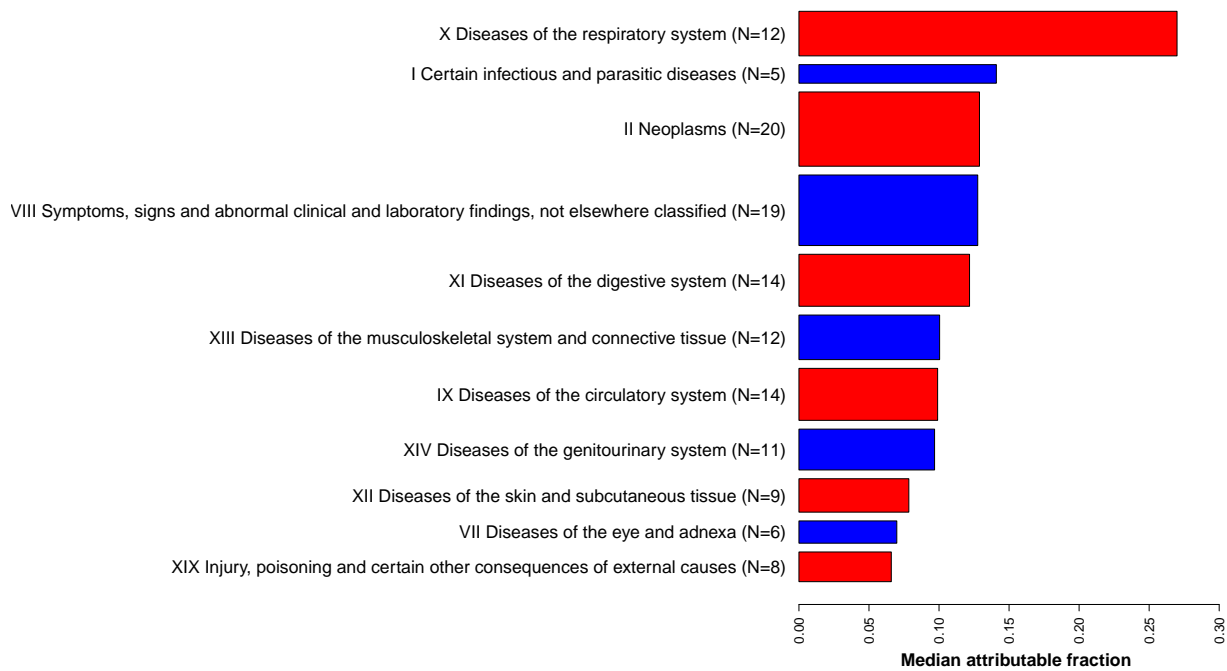
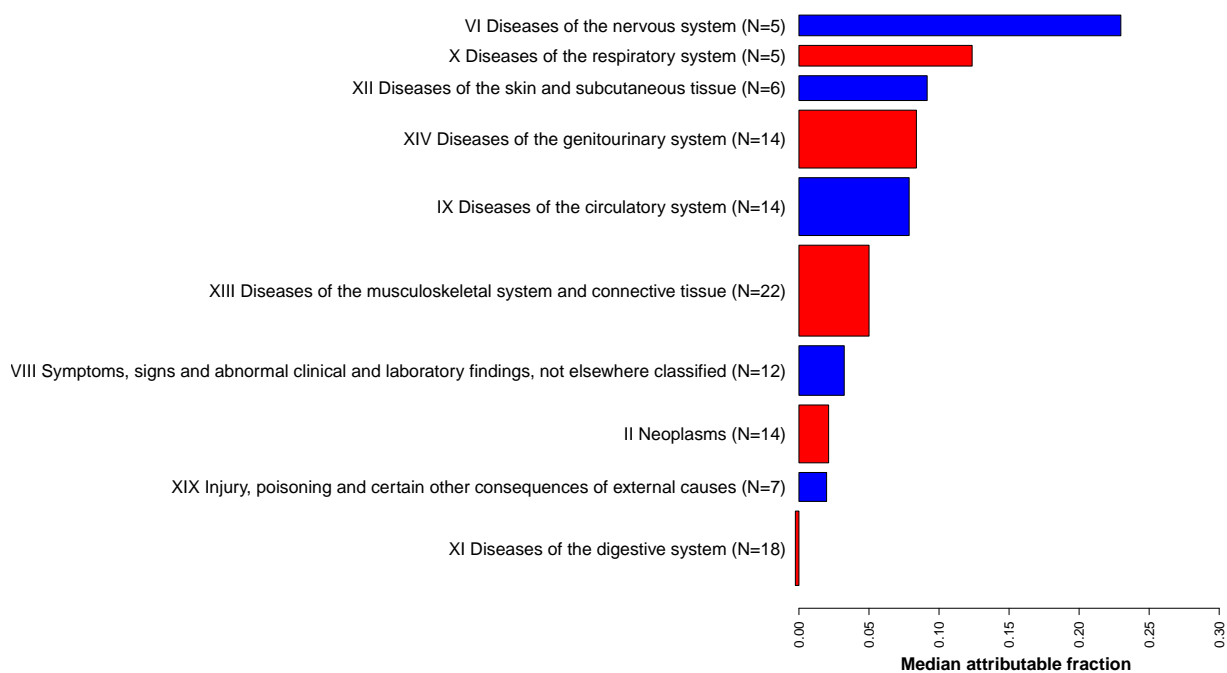


Figure 2: For diseases with a statistically significant association with smoking or BMI after an FDR multiple-testing adjustment, attributable fractions A_f were calculated with Eq. 19. Lines indicate $A_f = 0.5$ (red), $A_f = 0.35$ (orange), $A_f = 0.2$ (yellow). $A_f < 0$ indicates a protective association.



Attribution fractions for smoking (existing UK Biobank population, versus if all never smoked)



Attribution fractions for BMI (existing UK Biobank population, versus if all were mid-tertile BMI)

Figure 3: Median attributable fractions for each ICD-10 chapter with at least 5 diseases where $A_f > 0.2$. Bar widths are proportional to the number of diseases in each chapter.