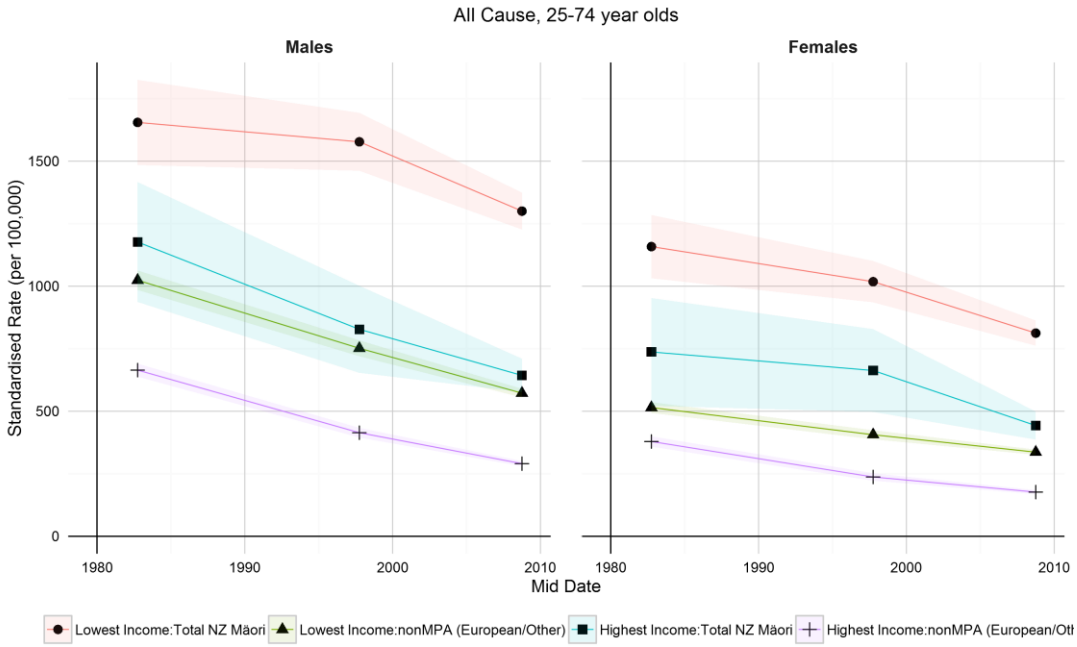
Supplementary material

Contents

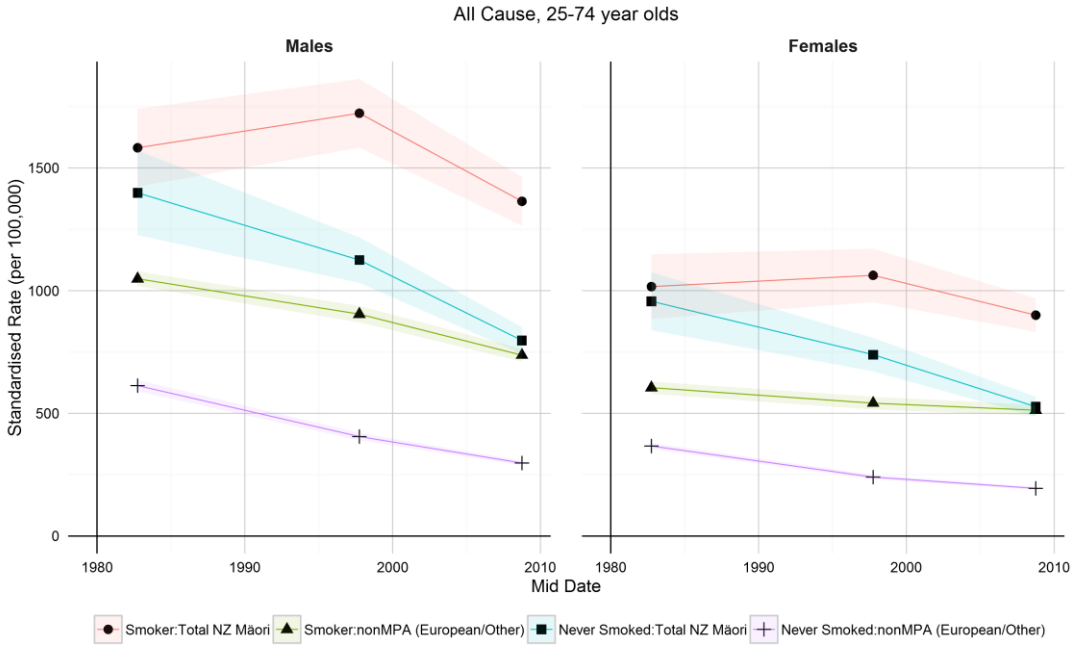
eFigures

eFigure 1: Standardized mortality rates for 25-74 year old all-cause mortality by sex by ethnicity:

a) By income (high and low tertile; middle income tertile suppressed)

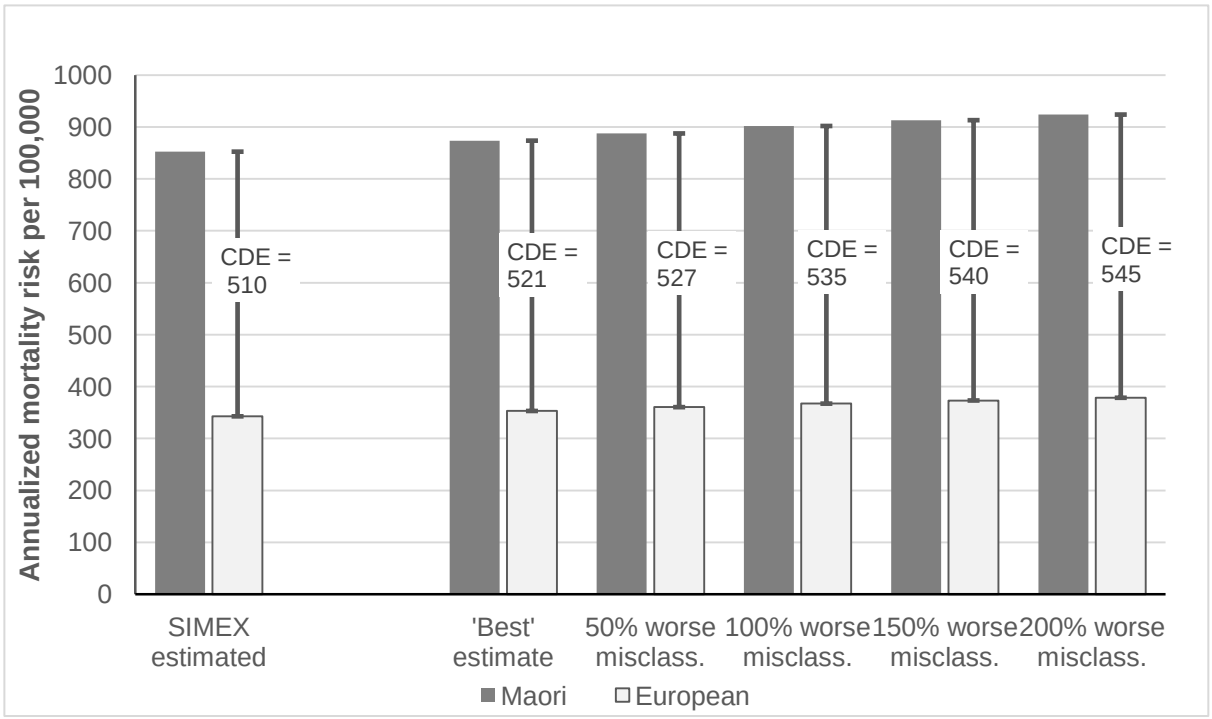


b) By smoking status (never and current; ex-smoker suppressed)

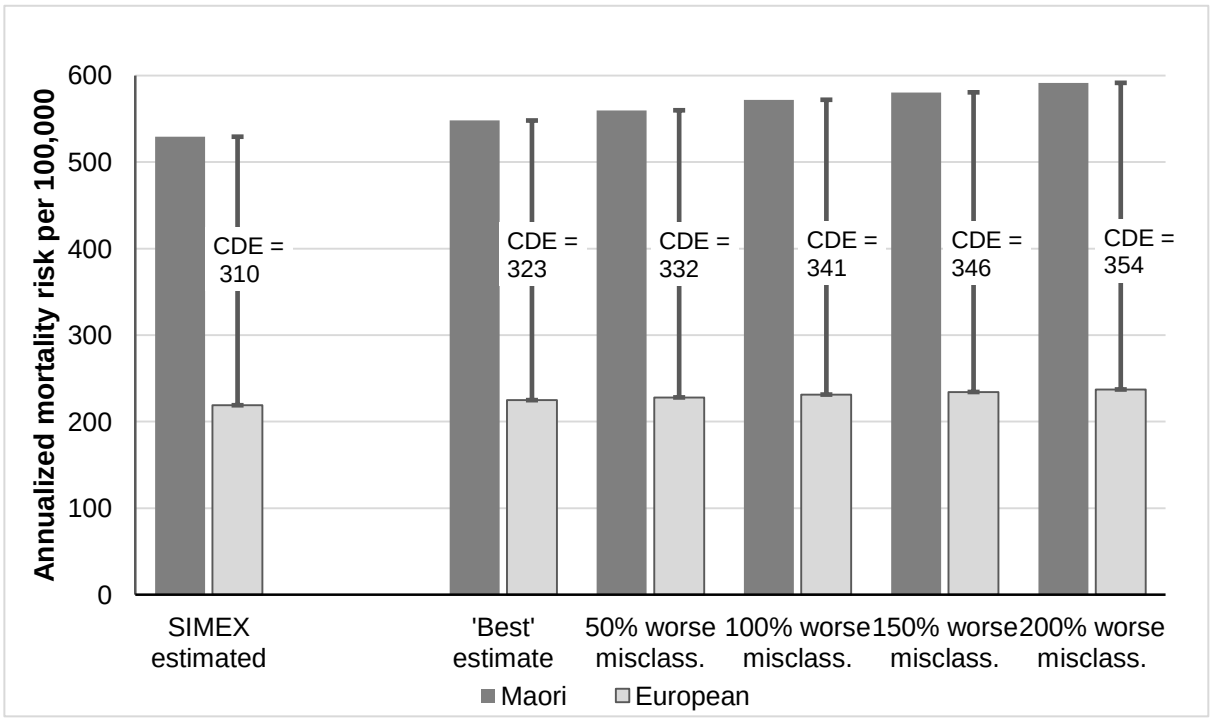


eFigure 2: SIMEX correction of smoking CDEs for plausible misclassification of the smoking variable in the census for the 2006-11 analyses. 'Best' estimate is that based on observed data with assumed amount of misclassification of smoking; SIMEX is that extrapolated to possible value with no misclassification of smoking

a. Males



b. Females



eTables

eTable 1: Standardized rate ratios (SRRs) and rate differences (SRDs) for all-cause mortality (per 100,000 population) comparing low to high income groups, by sex by ethnic strata

	Males					Females			
	Low income	High income	SRR	SRD		Low income	High income	SRR	SRD
Age standardized, whole cohort of Māori and Europeans									
1981-84	1095.8	684.1	1.60	411.7		567.6	392.7	1.45	174.8
1996-99	860.5	431.6	1.99	428.9		472.2	253.0	1.87	219.2
2006-11	658.5	311.3	2.12	347.2		392.5	191.1	2.05	201.5
Age standardized, restricted to Māori									
1981-84	1653.2	1146.6	1.44	506.6		1163.4	746.2	1.56	417.2
1996-99	1597.3	848.8	1.88	748.5		1028.2	649.1	1.58	379.1
2006-11	1322.2	657.0	2.01	665.2		818.5	448.4	1.83	370.1
Age standardized, restricted to Europeans									
1981-84	1028.4	668.4	1.54	359.9		517.3	380.3	1.36	137.0
1996-99	752.6	416.8	1.81	335.9		403.7	236.0	1.71	167.7
2006-11	575.7	294.0	1.96	281.7		339.1	177.6	1.91	161.4

Data restricted to observations with complete sex, age, ethnicity, smoking and income data

eTable 2: Standardized rate ratios (SRRs) and rate differences (SRDs) for all-cause mortality (per 100,000 population) comparing current to never smoker groups, by sex by ethnic strata

	Males					Females			
	Current	Never	SRR	SRD		Low income	High income	SRR	SRD
Age standardized, whole cohort of Māori and Europeans									
1981-84	1091.7	650.1	1.68	441.7		635.9	380.7	1.67	255.2
1996-99	1000.5	446.2	2.24	554.3		603.9	263.2	2.29	340.7
2006-11	818.7	328.3	2.49	490.3		574.0	205.5	2.79	368.6
Age standardized, restricted to Māori									
1981-84	1545.7	1432.9	1.08	112.8		1029.4	908.0	1.13	121.3
1996-99	1713.9	1060.9	1.62	653.0		999.2	730.8	1.37	268.4
2006-11	1351.2	798.1	1.69	553.2		849.5	487.5	1.74	362.0
Age standardized, restricted to Europeans									
1981-84	1054.9	609.8	1.73	445.1		598.7	364.7	1.64	234.0
1996-99	911.1	403.6	2.26	507.5		540.4	237.7	2.27	302.7
2006-11	744.2	297.8	2.50	446.4		517.8	189.6	2.73	328.2

Data restricted to observations with complete sex, age, ethnicity, smoking and income data

ETable 3: Table of regression coefficients for all 12 a priori logistic regression prediction equations, for the complete case dataset

Sex:	Males						Females					
Cohort:	1981-84		1996-99		2006-11		1981-84		1996-99		2006-11	
Model including:	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking
Parameter												
Intercept	-3.07	-3.4	-3.6	-3.9	-3.16	-3.53	-3.77	-4.07	-4.12	-4.44	-3.71	-4.04
25-29 yrs	-2.50	-2.21	-2.13	-1.90	-2.32	-2.00	-3.09	-2.88	-2.62	-2.18	-2.86	-2.52
30-34 yrs	-2.49	-2.1	-2.12	-1.96	-2.1	-1.71	-2.57	-2.19	-2.43	-2.19	-2.43	-2.25
35-39 yrs	-2.17	-1.99	-1.86	-1.84	-1.94	-1.62	-2.10	-1.82	-2.14	-1.94	-1.86	-1.61
40-44 yrs	-1.72	-1.69	-1.71	-1.70	-1.59	-1.37	-1.54	-1.23	-1.63	-1.50	-1.52	-1.38
45-49 yrs	-1.21	-1.10	-1.20	-1.16	-1.21	-1.15	-1.03	-0.897	-1.03	-0.958	-1.03	-0.89
50-54 yrs	-0.656	-0.591	-0.751	-0.731	-0.759	-0.647	-0.701	-0.589	-0.575	-0.513	-0.595	-0.593
55-59 yrs	-0.138	-0.132	-0.237	-0.22	-0.368	-0.302	-0.300	-0.28	-0.271	-0.124	-0.276	-0.184
65-69 yrs	0.628	0.705	0.742	0.743	0.56	0.563	0.462	0.553	0.513	0.484	0.515	0.543
70-74 yrs	1.09	1.10	1.20	1.25	1.03	1.13	0.939	1.10	0.945	1.01	0.945	0.94
Māori	0.119	0.358	0.682	0.941	0.669	0.734	0.752	0.879	0.804	0.765	0.641	0.779
Log Income (centered)	-0.0804	-0.0667	-0.110	-0.0915	-0.139	-0.118	-0.0974	-0.0878	-0.139	-0.124	-0.150	-0.128
NZ Deprivation Deciles (centered)	0.0505	0.0446	0.0585	0.0481	0.0666	0.053	0.0365	0.0314	0.0568	0.047	0.0625	0.0496
Unemployed	0.190	0.152	0.416	0.340	0.449	0.393	0.0939	0.0667	0.0826	0.049	0.379	0.344
Out of labor force	0.787	0.784	0.895	0.856	1.00	0.974	0.507	0.518	0.622	0.617	0.763	0.761
School Qualifications	-0.0679	-0.0635	-0.0653	-0.0591	-0.120	-0.0959	-0.0931	-0.105	-0.0807	-0.0689	-0.127	-0.0968
Post School Qualifications	-0.166	-0.133	-0.124	-0.0873	-0.205	-0.146	-0.223	-0.224	-0.183	-0.163	-0.259	-0.202
65+ yrs and not in work	-0.424	-0.432	-0.491	-0.483	-0.582	-0.571	-0.00247	-0.00846	-0.121	-0.133	-0.313	-0.31
Ex Smoker	NA	0.312	NA	0.337	NA	0.407	NA	0.499	NA	0.48	NA	0.368
Current Smoker	NA	0.561	NA	0.747	NA	0.883	NA	0.683	NA	0.916	NA	1.02
25-29 yrs * Māori	0.656	0.728	-0.290	-0.484	-0.351	-0.612	-0.151	-0.195	-0.403	0.02	-0.39	-0.112
30-34 yrs * Māori	0.504	0.771	0.0070	-0.466	-0.212	-0.359	-0.033	-0.0978	-0.594	-0.634	-0.381	-0.534

Sex:	Males						Females					
Cohort:	1981-84		1996-99		2006-11		1981-84		1996-99		2006-11	
Model including:	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking
Parameter												
35-39 yrs * Māori	0.697	-0.0714	-0.379	-0.557	-0.0586	-0.207	-0.472	-1.05	-0.332	-0.298	-0.401	-0.621
40-44 yrs * Māori	0.548	0.436	0.221	0.321	-0.0754	0.0762	-0.160	-0.391	-0.461	-0.104	-0.219	-0.455
45-49 yrs * Māori	0.706	0.827	0.0107	-0.0531	-0.0501	0.00829	-0.0945	-0.126	-0.109	0.157	-0.116	-0.386
50-54 yrs * Māori	0.434	0.561	0.0263	-0.0048	-0.02	0.0821	-0.103	-0.322	-0.227	-0.168	-0.0529	0.044
55-59 yrs * Māori	0.24	0.183	0.013	-0.253	0.0496	-0.038	0.0537	0.133	-0.171	-0.0862	0.0358	-0.253
65-69 yrs * Māori	0.248	-0.051	-0.136	-0.198	-0.0691	0.0847	-1.60	-1.64	0.298	0.452	-0.0491	-0.284
70-74 yrs * Māori	0.096	0.480	-0.318	-0.668	-0.133	-0.190	-1.70	-1.67	0.264	0.334	0.0298	-0.0275
Log income (centered) * Māori	0.133	0.118	0.0215	0.0138	0.0299	0.0118	0.00815	-0.0047	0.119	0.102	0.0602	0.0389
NZ Deprivation Deciles (centered) * Māori	0.0117	0.0171	0.0215	0.0283	0.0194	0.0298	0.0326	0.0383	0.0371	0.0458	0.040	0.0464
Unemployed * Māori	0.145	0.178	0.0111	0.0615	-0.172	-0.138	0.324	0.346	0.247	0.278	-0.482	-0.468
Out of labor force * Māori	0.465	0.466	0.0256	0.0517	0.0198	0.0414	-0.0779	-0.0971	0.216	0.214	0.0803	0.0756
School Qualifications * Māori	-0.306	-0.323	-0.119	-0.113	-0.0277	-0.0338	-0.0547	-0.0332	0.0115	0.0162	0.0302	0.0232
Post School Qualifications * Māori	-0.195	-0.239	-0.101	-0.113	-0.0459	-0.0764	0.270	0.270	0.020	0.0086	-0.0118	-0.0373
65+ yrs and not in work * Māori	-0.492	-0.505	0.027	0.0181	0.0309	-0.00106	1.53	1.52	-0.598	-0.594	0.0155	-0.00125
Ex Smoker * Māori	NA	-0.0911	NA	-0.588	NA	-0.116	NA	-0.204	NA	0.0983	NA	-0.311
Current Smoker * Māori	NA	-0.510	NA	-0.337	NA	-0.237	NA	-0.391	NA	-0.351	NA	-0.531
25-29 yrs * Ex Smoker	NA	-0.847	NA	-0.355	NA	-0.477	NA	-0.521	NA	-1.30	NA	-0.646
30-34 yrs * Ex Smoker	NA	-0.527	NA	-0.664	NA	-0.687	NA	-0.833	NA	-0.360	NA	-0.431
35-39 yrs * Ex Smoker	NA	-0.318	NA	-0.0744	NA	-0.644	NA	-0.334	NA	-0.507	NA	-0.435
40-44 yrs * Ex Smoker	NA	-0.103	NA	-0.0633	NA	-0.244	NA	-0.506	NA	-0.286	NA	-0.159
45-49 yrs * Ex Smoker	NA	-0.238	NA	-0.243	NA	-0.183	NA	-0.271	NA	-0.226	NA	-0.253
50-54 yrs * Ex Smoker	NA	-0.239	NA	-0.092	NA	-0.295	NA	-0.238	NA	-0.152	NA	-0.0468
55-59 yrs * Ex Smoker	NA	-0.1	NA	-0.128	NA	-0.135	NA	-0.0244	NA	-0.203	NA	-0.09

Sex:	Males						Females					
Cohort:	1981-84		1996-99		2006-11		1981-84		1996-99		2006-11	
Model including:	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking
Parameter												
65-69 yrs * Ex Smoker	NA	-0.0279	NA	0.0227	NA	0.0247	NA	-0.0783	NA	0.168	NA	0.0814
70-74 yrs * Ex Smoker	NA	0.0562	NA	-0.032	NA	-0.0635	NA	-0.147	NA	0.0805	NA	0.235
25-29 yrs * Current Smoker	NA	-0.307	NA	-0.523	NA	-0.745	NA	-0.4	NA	-0.962	NA	-0.975
30-34 yrs * Current Smoker	NA	-0.633	NA	-0.229	NA	-0.878	NA	-0.778	NA	-0.784	NA	-0.530
35-39 yrs * Current Smoker	NA	-0.224	NA	-0.0826	NA	-0.653	NA	-0.708	NA	-0.525	NA	-0.768
40-44 yrs * Current Smoker	NA	0.0186	NA	-0.0758	NA	-0.564	NA	-0.711	NA	-0.453	NA	-0.586
45-49 yrs * Current Smoker	NA	-0.0941	NA	-0.0422	NA	-0.131	NA	-0.276	NA	-0.294	NA	-0.516
50-54 yrs * Current Smoker	NA	-0.0203	NA	-0.098	NA	-0.153	NA	-0.255	NA	-0.293	NA	-0.192
55-59 yrs * Current Smoker	NA	0.031	NA	0.00859	NA	-0.104	NA	-0.102	NA	-0.436	NA	-0.367
65-69 yrs * Current Smoker	NA	-0.116	NA	-0.00192	NA	0.0953	NA	-0.117	NA	0.0325	NA	0.0278
70-74 yrs * Current Smoker	NA	-0.00083	NA	0.0121	NA	0.0486	NA	-0.216	NA	-0.0929	NA	0.174
25-29 yrs * Ex Smoker * Māori	NA	0.208	NA	0.472	NA	-0.0789	NA	-0.688	NA	-0.505	NA	-0.771
30-34 yrs * Ex Smoker * Māori	NA	-1.41	NA	1.16	NA	0.427	NA	0.749	NA	-0.565	NA	0.748
35-39 yrs * Ex Smoker * Māori	NA	0.617	NA	-0.567	NA	-0.010	NA	0.979	NA	0.296	NA	0.258
40-44 yrs * Ex Smoker * Māori	NA	0.0336	NA	0.287	NA	-0.190	NA	0.521	NA	-0.594	NA	0.541
45-49 yrs * Ex Smoker * Māori	NA	-0.00665	NA	0.512	NA	0.163	NA	0.425	NA	-0.477	NA	0.506
50-54 yrs * Ex Smoker * Māori	NA	-0.0909	NA	0.437	NA	0.206	NA	0.486	NA	0.0742	NA	0.154
55-59 yrs * Ex Smoker * Māori	NA	-0.184	NA	0.710	NA	0.302	NA	0.0252	NA	-0.344	NA	0.556
65-69 yrs * Ex Smoker * Māori	NA	0.0675	NA	0.198	NA	-0.152	NA	0.095	NA	-0.282	NA	0.574
70-74 yrs * Ex Smoker * Māori	NA	-0.669	NA	0.853	NA	0.189	NA	0.209	NA	-0.00278	NA	0.184
25-29 yrs * Current Smoker * Māori	NA	-0.102	NA	0.188	NA	0.617	NA	0.253	NA	-0.285	NA	-0.0218
30-34 yrs * Current Smoker * Māori	NA	-0.114	NA	0.418	NA	0.351	NA	0.16	NA	0.532	NA	0.217
35-39 yrs * Current Smoker * Māori	NA	1.10	NA	0.248	NA	0.360	NA	0.902	NA	0.020	NA	0.602

Sex:	Males						Females					
Cohort:	1981-84		1996-99		2006-11		1981-84		1996-99		2006-11	
Model including:	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking	SEP	SEP & smoking
Parameter												
40-44 yrs * Current Smoker * Māori	NA	0.203	NA	-0.442	NA	-0.17	NA	0.5	NA	-0.277	NA	0.452
45-49 yrs * Current Smoker * Māori	NA	-0.248	NA	-0.208	NA	-0.264	NA	-0.0415	NA	-0.19	NA	0.488
50-54 yrs * Current Smoker * Māori	NA	-0.186	NA	-0.205	NA	-0.421	NA	0.338	NA	-0.0348	NA	-0.175
55-59 yrs * Current Smoker * Māori	NA	0.244	NA	0.115	NA	-0.0693	NA	-0.159	NA	0.164	NA	0.436
65-69 yrs * Current Smoker * Māori	NA	0.692	NA	0.108	NA	-0.283	NA	0.0478	NA	-0.169	NA	0.107
70-74 yrs * Current Smoker * Māori	NA	-0.581	NA	0.103	NA	-0.0495	NA	-0.488	NA	-0.26	NA	-0.0713

Small area deprivation, NZDep, was treated as a continuous variable with integer values for each decile of deprivation. To allow for the different meaning of labor force status post-retirement (usually 65 year of age in NZ across the 1980s to 2000s), we additionally include an interaction or dummy variable for those employed older than 65 compared to everyone else.

ETable 4: Mortality risks, risk differences, risk ratios and other causal mediation estimates, by age-group (25-44, 45-64 and 65-74 years of age)

Age-group and outcome	Population and Effect Sizes (predicted)	Males			Females		
		1981	1996	2006	1981	1996	2006
25 – 44 year olds							
Mortality Risks (per 100,000)	Māori	400 (336, 465)	338 (295, 383)	270 (241, 298)	219 (178, 265)	158 (134, 182)	132 (117, 147)
	Europeans	163 (153, 174)	134 (125, 144)	116 (109, 123)	97.2 (89, 106)	68 (62.2, 73.8)	64.6 (60.4, 69.1)
	Māori with Euro SEP (CW)	304 (253, 362)	230 (199, 264)	189 (168, 211)	172 (134, 217)	108 (90, 128)	90.8 (80.2, 102)
	Māori with Euro SEP + Tob (CW)	303 (248, 359)	225 (194, 258)	186 (164, 208)	169 (126, 220)	112 (91.2, 135)	88 (75.6, 101)
	Māori Never Smoking (counterfactual)	400 (289, 528)	305 (243, 377)	255 (211, 301)	198 (118, 287)	177 (131, 229)	117 (89.8, 145)
	Euro Never Smoking (counterfactual)	144 (128, 161)	109 (97.6, 122)	111 (102, 121)	97.3 (86, 110)	63.1 (55.4, 71)	58.9 (53.3, 64.6)
Risk Differences (per 100,000)	Total Effect	236 (172, 301)	204 (161, 249)	154 (124, 183)	122 (78.6, 169)	89.4 (65.7, 115)	67.1 (52.2, 82.3)
	NDE (SEP mediation)	140 (86.8, 198)	96.3 (63.2, 131)	73.2 (50.4, 96.3)	74.9 (35.8, 120)	40.1 (21.2, 62)	26.1 (14.6, 38.3)
	NDE (SEP + Tob mediation)	139 (84.3, 198)	90.4 (58.7, 125)	69.5 (47, 93.6)	72.2 (27.7, 123)	43.9 (22.6, 67.7)	23.4 (10.4, 36.8)
	NIE (SEP mediation)	95.8 (62.6, 130)	107 (88.7, 129)	79.9 (69.2, 91.4)	46.6 (26.2, 67.4)	49.2 (39.1, 60.1)	40.8 (34.6, 47.9)
	NIE (SEP + Tob mediation)	96.5 (58.6, 135)	113 (90.8, 136)	83.4 (70.4, 96.8)	49.3 (19, 76.9)	45.5 (31.4, 59.4)	43.9 (35, 52.6)
	CDE (Tobacco free counterfactual)	256 (141, 386)	195 (131, 270)	143 (97.6, 192)	101 (19.4, 191)	114 (66.2, 169)	58.2 (30.2, 86.9)
% Mediated (natural effects)	% Mediated (SEP)	40.9 (27.5, 54.5)	52.8 (45, 62)	52.2 (45.6, 60)	38.6 (21.6, 59.3)	55.2 (44.3, 69.5)	61 (51.5, 72.7)
	% Mediated (SEP + Tob)	41.2 (26.5, 55.9)	55.6 (46.2, 65.9)	54.4 (46.8, 63.9)	40.9 (16.5, 69.6)	51.3 (35.8, 68.1)	65.3 (52.7, 81.3)
	Mediation Change (SEP to SEP + Tob)	0.33 (-6.44, 7.51)	2.61 (-2.34, 7.62)	2.29 (-2.41, 7.02)	2.21 (-15.5, 20.1)	-3.98 (-16.2, 6.84)	4.43 (-4.71, 14.5)
Counterfactual changes had nobody smoked (controlled effects)	% Reduction in Māori risk	-0.164 (-25.7, 24.4)	9.26 (-8.3, 26.8)	5.43 (-8.46, 18.4)	9.38 (-25, 40.2)	-12.1 (-40.7, 12.7)	11 (-6.54, 29.3)
	% Reduction in Euro risk	11.6 (3.5, 20.2)	18.3 (11.1, 25)	3.94 (-1.49, 9.59)	-0.005 (-8.71, 8.34)	7.1 (-1.38, 15.5)	8.8 (2.63, 14.9)
	% Eradicated (Tobacco free counterfactual)	-8.58 (-52.8, 34.7)	3.52 (-27.2, 32)	6.76 (-19.2, 30.3)	16.7 (-45.9, 80.5)	-27.1 (-79.9, 17.9)	13.1 (-22.8, 50.8)
Risk Ratios	Total Effect	2.45 (2.04, 2.87)	2.52 (2.17, 2.89)	2.32 (2.05, 2.62)	2.25 (1.78, 2.77)	2.31 (1.95, 2.75)	2.04 (1.79, 2.32)
	NDE (SEP mediation)	1.86 (1.52, 2.24)	1.72 (1.46, 2)	1.63 (1.43, 1.86)	1.77 (1.36, 2.26)	1.59 (1.3, 1.94)	1.41 (1.22, 1.61)
	NDE (SEP + Tob mediation)	1.86 (1.51, 2.24)	1.68 (1.43, 1.96)	1.6 (1.4, 1.83)	1.74 (1.28, 2.29)	1.65 (1.32, 2.04)	1.36 (1.16, 1.58)
	NIE (SEP mediation)	1.32 (1.19, 1.45)	1.47 (1.39, 1.55)	1.42 (1.37, 1.48)	1.27 (1.15, 1.43)	1.46 (1.35, 1.56)	1.45 (1.38, 1.53)
	NIE (SEP + Tob mediation)	1.32 (1.18, 1.48)	1.5 (1.4, 1.62)	1.45 (1.37, 1.53)	1.29 (1.1, 1.55)	1.41 (1.26, 1.57)	1.5 (1.38, 1.63)
	CDE (Tobacco free counterfactual)	2.77 (1.94, 3.75)	2.79 (2.15, 3.54)	2.28 (1.84, 2.78)	2.04 (1.2, 3)	2.81 (2.03, 3.82)	2 (1.49, 2.51)

	Population and Effect Sizes (predicted)	Males			Females		
		1981	1996	2006	1981	1996	2006
45 – 64 year olds							
Mortality Risks (per 100,000)	Māori	2197 (1971, 2416)	1808 (1686, 1930)	1351 (1279, 1428)	1533 (1348, 1721)	1317 (1212, 1427)	907 (850, 966)
	Europeans	1089 (1058, 1118)	676 (654, 697)	512 (499, 526)	616 (593, 640)	433 (418, 449)	342 (332, 352)
	Māori with Euro SEP (CW)	1696 (1485, 1937)	1254 (1143, 1365)	941 (877, 1008)	1248 (1080, 1432)	975 (884, 1075)	643 (591, 697)
	Māori with Euro SEP + Tob (CW)	1699 (1486, 1930)	1224 (1116, 1336)	916 (853, 983)	1207 (1036, 1397)	923 (835, 1021)	599 (548, 653)
	Māori Never Smoking (counterfactual)	2169 (1784, 2577)	1596 (1402, 1806)	1085 (973, 1203)	1324 (1051, 1609)	1078 (928, 1239)	718 (628, 818)
	Euro Never Smoking (counterfactual)	818 (769, 872)	506 (476, 535)	384 (367, 400)	483 (457, 512)	338 (320, 358)	265 (253, 277)
Risk Differences (per 100,000)	Total Effect	1110 (882, 1330)	1132 (1008, 1259)	839 (766, 918)	916 (728, 1102)	883 (781, 993)	565 (506, 625)
	NDE (SEP mediation)	608 (394, 848)	578 (467, 689)	429 (364, 495)	632 (465, 820)	543 (449, 643)	301 (248, 357)
	NDE (SEP + Tob mediation)	612 (392, 848)	549 (440, 660)	405 (341, 473)	592 (420, 786)	491 (399, 588)	257 (204, 311)
	NIE (SEP mediation)	500 (354, 644)	553 (479, 630)	411 (363, 456)	282 (173, 399)	341 (272, 407)	264 (230, 299)
	NIE (SEP + Tob mediation)	497 (347, 646)	581 (504, 660)	434 (385, 481)	321 (206, 445)	392 (322, 463)	308 (269, 346)
	CDE (Tobacco free counterfactual)	1351 (962, 1762)	1090 (894, 1301)	702 (588, 823)	839 (572, 1125)	740 (587, 904)	454 (361, 552)
% Mediated (natural effects)	% Mediated (SEP)	45.2 (31.8, 59.3)	48.9 (42.9, 55.4)	48.9 (44, 53.9)	31.1 (19, 42.4)	38.7 (31.7, 45.3)	46.8 (40.9, 53)
	% Mediated (SEP + Tob)	44.7 (30.9, 60.2)	51.4 (45.1, 58.1)	51.7 (46.5, 57)	35.2 (23, 47.4)	44.6 (37.2, 51.3)	54.5 (47.8, 61.2)
	Mediation Change (SEP to SEP + Tob)	-0.22 (-4.29, 3.48)	2.44 (0.69, 4.35)	2.81 (-1.19, 4.39)	4.23 (1.30, 10.1)	5.79 (2.31, 9.13)	7.71 (4.09, 11.1)
Counterfactual changes had nobody smoked (controlled effects)	% Reduction in Māori risk	1.16 (-14.2, 16.4)	11.5 (2.13, 20.4)	19.6 (12.7, 26.7)	13.3 (-1.47, 28)	18.3 (7.41, 27.7)	20.8 (11.4, 29.7)
	% Reduction in Euro risk	24.8 (20.4, 28.7)	25.1 (21.6, 28.9)	25.1 (22.5, 27.8)	21.5 (18.2, 24.7)	21.8 (18.4, 25)	22.4 (19.9, 25)
	% Eradicated (Tobacco free counterfactual)	-22 (-55.3, 8.88)	3.57 (-11.7, 18.2)	16.4 (4.5, 27.8)	8.04 (-17.6, 33.1)	16.7 (0.607, 30.9)	19.8 (4.64, 34.5)
Risk Ratios	Total Effect	2.02 (1.81, 2.23)	2.67 (2.48, 2.88)	2.64 (2.48, 2.81)	2.48 (2.17, 2.81)	3.04 (2.78, 3.32)	2.65 (2.47, 2.85)
	NDE (SEP mediation)	1.56 (1.36, 1.78)	1.85 (1.69, 2.03)	1.84 (1.71, 1.98)	2.03 (1.74, 2.34)	2.25 (2.03, 2.5)	1.88 (1.72, 2.06)
	NDE (SEP + Tob mediation)	1.56 (1.36, 1.78)	1.81 (1.65, 1.99)	1.79 (1.66, 1.93)	1.96 (1.68, 2.29)	2.13 (1.91, 2.37)	1.75 (1.59, 1.93)
	NIE (SEP mediation)	1.3 (1.2, 1.41)	1.44 (1.37, 1.52)	1.44 (1.38, 1.5)	1.23 (1.13, 1.33)	1.35 (1.27, 1.43)	1.41 (1.34, 1.48)
	NIE (SEP + Tob mediation)	1.29 (1.19, 1.42)	1.47 (1.4, 1.56)	1.47 (1.41, 1.54)	1.26 (1.16, 1.39)	1.42 (1.33, 1.52)	1.51 (1.43, 1.6)
	CDE (Tobacco free counterfactual)	2.65 (2.16, 3.19)	3.16 (2.74, 3.62)	2.83 (2.51, 3.17)	2.73 (2.17, 3.35)	3.19 (2.7, 3.7)	2.71 (2.36, 3.11)

	Population and Effect Sizes (predicted)	Males			Females		
		1981	1996	2006	1981	1996	2006
65 – 74 year olds							
Mortality Risks (per 100,000)	Māori	5858 (5093, 6617)	5511 (5059, 6002)	4273 (4013, 4533)	4639 (3917, 5330)	3995 (3574, 4388)	3117 (2899, 3345)
	Europeans	4332 (4232, 4434)	3065 (2992, 3134)	2133 (2091, 2178)	2392 (2320, 2462)	1724 (1674, 1774)	1350 (1314, 1386)
	Māori with Euro SEP (CW)	5029 (4263, 5782)	4537 (4115, 5024)	3455 (3217, 3713)	4024 (3329, 4711)	3297 (2928, 3691)	2554 (2342, 2765)
	Māori with Euro SEP + Tob (CW)	4929 (4174, 5756)	4575 (4133, 5056)	3434 (3195, 3692)	3969 (3266, 4679)	3238 (2864, 3629)	2396 (2189, 2608)
	Māori Never Smoking (counterfactual)	5809 (4545, 7097)	4498 (3848, 5222)	3571 (3170, 3958)	4349 (3403, 5376)	3242 (2754, 3731)	2334 (2032, 2642)
	Euro Never Smoking (counterfactual)	3338 (3163, 3519)	2319 (2212, 2426)	1608 (1552, 1669)	2036 (1950, 2120)	1276 (1219, 1334)	1009 (971, 1050)
Risk Differences (per 100,000)	Total Effect	1528 (743, 2283)	2443 (2003, 2937)	2136 (1871, 2402)	2241 (1530, 2934)	2267 (1848, 2665)	1765 (1546, 1999)
	NDE (SEP mediation)	698 (-63.5, 1449)	1477 (1051, 1953)	1320 (1082, 1585)	1636 (938, 2333)	1575 (1203, 1964)	1202 (981, 1427)
	NDE (SEP + Tob mediation)	611 (-158, 1431)	1511 (1067, 1995)	1305 (1065, 1568)	1579 (867, 2293)	1516 (1140, 1904)	1050 (832, 1262)
	NIE (SEP mediation)	832 (512, 1168)	966 (767, 1180)	813 (681, 947)	609 (332, 878)	686 (505, 869)	562 (453, 676)
	NIE (SEP + Tob mediation)	878 (480, 1248)	923 (695, 1157)	823 (686, 961)	633 (276, 952)	736 (544, 941)	711 (588, 840)
	CDE (Tobacco free counterfactual)	2482 (1211, 3809)	2188 (1518, 2886)	1963 (1547, 2353)	2316 (1353, 3331)	1964 (1467, 2456)	1325 (1023, 1641)
% Mediated (natural effects)	% Mediated (SEP)	54.8 (30.3, 107)	39.7 (30.6, 49.7)	37.9 (31.5, 44.7)	27.4 (15.1, 41.9)	30.5 (22.2, 38.8)	31.8 (25.8, 38.6)
	% Mediated (SEP + Tob)	59.1 (29.8, 120)	38 (27.8, 48.3)	38.6 (32.1, 45.6)	28.7 (12.8, 46.9)	32.9 (24.1, 41.9)	40.5 (33.5, 48.1)
	Mediation Change (SEP to SEP + Tob)	4.25 (-10.4, 22.3)	-1.79 (-6.97, 3.35)	0.60 (-1.45, 2.84)	1.52 (-8.76, 11.2)	2.33 (-1.38, 6.25)	8.53 (4.95, 12.4)
Counterfactual changes had nobody smoked (controlled effects)	% Reduction in Māori risk	0.263 (-17.8, 18.7)	17.9 (8.27, 28.1)	16.2 (9.01, 23.6)	5.62 (-10.1, 20.6)	18.5 (9.3, 27.6)	24.9 (17.3, 32.4)
	% Reduction in Euro risk	22.9 (19.3, 26.6)	24.3 (21.3, 27.3)	24.5 (22.1, 26.8)	14.8 (12.4, 17.2)	25.9 (23.5, 28.2)	25.1 (23, 27.4)
	% Eradicated (Tobacco free counterfactual)	-62.9 (-176, 9.22)	10.3 (-12.5, 34.4)	8.05 (-6.78, 23.6)	-3.38 (-37.7, 28.8)	13.1 (-3.4, 29.9)	25.0 (10.9, 38.5)
Risk Ratios	Total Effect	1.35 (1.17, 1.53)	1.80 (1.65, 1.96)	2.00 (1.88, 2.13)	1.94 (1.64, 2.24)	2.31 (2.08, 2.56)	2.31 (2.13, 2.49)
	NDE (SEP mediation)	1.16 (0.985, 1.34)	1.48 (1.34, 1.64)	1.62 (1.51, 1.75)	1.68 (1.39, 1.98)	1.91 (1.7, 2.14)	1.89 (1.73, 2.06)
	NDE (SEP + Tob mediation)	1.14 (0.964, 1.33)	1.49 (1.35, 1.65)	1.61 (1.5, 1.74)	1.66 (1.36, 1.97)	1.88 (1.66, 2.11)	1.78 (1.61, 1.94)
	NIE (SEP mediation)	1.17 (1.09, 1.25)	1.21 (1.16, 1.27)	1.24 (1.19, 1.28)	1.15 (1.08, 1.23)	1.21 (1.14, 1.27)	1.22 (1.17, 1.27)
	NIE (SEP + Tob mediation)	1.18 (1.09, 1.27)	1.2 (1.15, 1.26)	1.24 (1.19, 1.29)	1.16 (1.06, 1.26)	1.23 (1.16, 1.3)	1.3 (1.24, 1.36)
	CDE (Tobacco free counterfactual)	1.74 (1.36, 2.14)	1.94 (1.65, 2.25)	2.22 (1.96, 2.47)	2.13 (1.67, 2.65)	2.54 (2.14, 2.95)	2.31 (2, 2.63)

eAppendix: R code

Functions for mediation point estimates

```
# Packages needed ------------------------------------------------------

library (dplyr)
library (haven)
library (sas7bdat)
library (readr)
library (broom)
library (reshape)
library (boot)

#### Load functions and models ----------------------------------------

source ([file path for data functions – see bottom of document])
source ([file path for model specifications – see bottom of document])

#### Read data into R and initial prep ---------------------------------

read_data <- function (file_title){

  root <- "[insert file path here]"
  file_title <- file_title
  path <- paste(root, file_title, sep = "")
  mort_data_master <- read_csv(path)

  mort_data <- mort_data_master%>%
               filter (EqIncJen != "NA",
                       AgeC_yrs < 75 & AgeC_yrs >=25,
                       TotMaori %in% c("1","5"),
                       SmkStat %in% c("1", "2", "3"),
                       !is.na(NZDep10),
                       HQual3 != "9")

  mort_data$EqIncJen <- ifelse (mort_data$EqIncJen < 1000, 1000, mort_data$EqIncJen)

  mort_data <- mort_data %>%
               mutate (EqIncJen_log = log(EqIncJen))%>%
               filter (EqIncJen_log > 0 & EqIncJen_log != "NaN")%>%
               mutate(EqIncJen_log_cent = EqIncJen_log - mean(EqIncJen_log),
                      NZDep10_cent = NZDep10 - 5,
                      Age_group = sapply(AgeC_yrs, age_5),
                      Age_group_broad = sapply(AgeC_yrs, age_broad),
                      Retired = sapply(AgeC_yrs, retired),
                      WHO_stand = sapply(AgeC_yrs, WHO_stand_fun),
                      Ret_lfs = ifelse (AgeC_yrs > 64 & LabSt3 >1, 1, 0))%>%
                 droplevels.data.frame(.)

  # ENSURING VARIABLES ARE NUMERIC WHERE APPROPRIATE #

  mort_data$NZDep10<- as.numeric(mort_data$NZDep10)
```

```
  mort_data$NZDep10_cent<- as.numeric(mort_data$NZDep10_cent)
  mort_data$WHO_stand <- as.numeric(mort_data$WHO_stand)

  # ENSURING VARIABLES ARE CATEGORICAL WHERE APPROPRIATE #

  mort_data$Retired <- as.factor(mort_data$Retired)
  mort_data$Sex <- as.factor(mort_data$Sex)
  mort_data$Link <- as.factor(mort_data$Link)
  mort_data$CarAccess <- as.factor(mort_data$CarAccess)
  mort_data$HQual3 <- as.factor(mort_data$HQual3)
  mort_data$TotMaori <- as.factor(mort_data$TotMaori)
  mort_data$LabSt3 <- as.factor(mort_data$LabSt3)
  mort_data$SmkStat <- as.factor(mort_data$SmkStat)
  mort_data$Age_group <- as.factor(mort_data$Age_group)
  mort_data$Age_group_broad <- as.factor(mort_data$Age_group_broad)

  mort_data$TotMaori <- ifelse (mort_data$TotMaori == "1", 1, ifelse (mort_data$TotMaori ==
"5", 0, NA))
  mort_data$Age_group <- relevel(mort_data$Age_group, ref = "6064")

  mort_data$SmkStat <- ifelse (mort_data$SmkStat == "3", 1, ifelse (mort_data$SmkStat ==
"1", 3, 2))
  mort_data$SmkStat <- as.factor(mort_data$SmkStat)

mort_data

}


#### Splitting cohorts into separate sexes, extra prep and calculating weights -------------
----------------------------

sex_prep <- function (cohort, sex){
  #look at code below and assess whether can be moved up to previous function
 sex_mort_data <- filter (cohort,
                          Sex == sex,
                          EqIncJen != "NA",
                          EqIncJen_log > 0 & EqIncJen_log != "NaN")%>%
            droplevels.data.frame(.)

  sex_mort_data <- sex_mort_data %>%
                    mutate(WHO_prop = WHO_stand/0.5413,
                           total_gender = n())%>%
                    group_by(TotMaori)%>%
                    mutate(eth_sum = n(),
                           p_eth = eth_sum/total_gender)

 sex_mort_data <- sex_mort_data %>%
                    group_by(Age_group,TotMaori)%>%
                    mutate(age_eth_sum = n())%>%
                    mutate (p_age_eth = (age_eth_sum/total_gender)/p_eth)

 sex_mort_data <- sex_mort_data %>%
                    mutate(weights = WHO_prop/p_age_eth)
 sex_mort_data

}
```

```
#### Cross world function for natural direct and indirect effects -------------------------
----------------------------


cw_calcs <- function (cohort_sex, sex_sep_mod, sex_tob_mod, mod_set_name) {

  cw_cohort_sex <- cohort_sex

  # Natural effects for SEP
  cw_cohort_sex$cross_ref <- cw_cohort_sex$TotMaori
  cw_cohort_sex$TotMaori <- ifelse(cw_cohort_sex$TotMaori == "0", 1, 1)

  cw_pred <- predict.glm(sex_sep_mod, type = "response", newdata = cw_cohort_sex)
  cw_pred <- as.data.frame(cw_pred)
  #sep_pred <- as.data.frame(sex_sep_mod$fitted.values) # new code
  sep_pred <- predict.glm(sex_sep_mod, type = "response", newdata = cohort_sex)
  sep_pred <- as.data.frame(sep_pred)

  cw_cohort_sex <- bind_cols(cw_cohort_sex, cw_pred)
  cw_cohort_sex <- bind_cols(cw_cohort_sex, sep_pred)

  # Natural effects for SEP + Smoking
  cw_pred_smo <- predict.glm(sex_tob_mod, type = "response", newdata = cw_cohort_sex)
  cw_pred_smo <- as.data.frame(cw_pred_smo)
  #smo_pred <- as.data.frame(sex_tob_mod$fitted.values)
  smo_pred <- predict.glm(sex_tob_mod, type = "response", newdata = cohort_sex)
  smo_pred <- as.data.frame(smo_pred)


  cw_cohort_sex <- bind_cols(cw_cohort_sex, cw_pred_smo)
  cw_cohort_sex <- bind_cols(cw_cohort_sex, smo_pred)

  # Controlled effects ie "never smoker"
  never_cohort_sex <- cohort_sex
  never_cohort_sex$SmkStat <- ifelse(never_cohort_sex$SmkStat ==
"3"|never_cohort_sex$SmkStat == "2", 1,1)
  never_cohort_sex$SmkStat <- as.factor(never_cohort_sex$SmkStat)

  never_pred <- predict.glm(sex_tob_mod, type = "response", newdata = never_cohort_sex)
  never_pred <- as.data.frame(never_pred)

  cw_cohort_sex <- bind_cols(cw_cohort_sex, never_pred)

  cw_cohort_sex <- cw_cohort_sex %>%
                    mutate (model_set = mod_set_name)


  # Observed risk calculation
  cw_cohort_sex$mort <- ifelse (cw_cohort_sex$Link == "1", 1, 0)

  cw_cohort_sex <- cw_cohort_sex

  risks <- cw_cohort_sex %>%
                    mutate (comb_weight = W_AgEthAdj*weights)%>%
```

```
                    group_by(CenYear, Sex, cross_ref, model_set)%>%
                    summarise(obs_risk = weighted.mean(mort, w = comb_weight),
                              pred_sep_risk = weighted.mean (sep_pred, w = weights), #new
                              cw_sep_risk = weighted.mean(cw_pred, w = weights),
                              pred_smo_risk = weighted.mean(smo_pred, w = weights), #new
                              cw_smo_risk = weighted.mean(cw_pred_smo, w = weights),
                              never_risk = weighted.mean (never_pred, w = weights))%>%
                    mutate(Age_group_broad = "2574",
                           Date_run = Sys.Date())



  risks_age <- cw_cohort_sex %>%
                    mutate (comb_weight = W_AgEthAdj*weights)%>%
                    group_by(CenYear,Sex,cross_ref, model_set, Age_group_broad)%>%
                    summarise(obs_risk = weighted.mean(mort, w = comb_weight),
                              pred_sep_risk = weighted.mean(sep_pred, w = weights),
                              cw_sep_risk = weighted.mean(cw_pred, w = weights),
                              pred_smo_risk = weighted.mean(smo_pred, w = weights),
                              cw_smo_risk = weighted.mean(cw_pred_smo, w = weights),
                              never_risk = weighted.mean (never_pred, w = weights))%>%
                    mutate(Date_run = Sys.Date())

  tab_risks <- bind_rows(risks, risks_age)

  tab_risks


}




############ Code for running complete analysis including the model fit element
####################

run_mediation <- function(cohort, glm_sep_title = NULL, glm_tob_title = NULL,
                          glm_sep_spec = NULL, glm_tob_spec = NULL, model_name) {


data_name <- read_data (file_title = paste ("mort", cohort, ".csv", sep = ""))

males_data <- sex_prep(data_name, sex = "1")
females_data <- sex_prep(data_name, sex = "2")



if (!is.null(glm_sep_spec)& !is.null(glm_tob_spec)){
  males_glm_sep <- glm(as.formula(glm_sep_spec),
                       family = binomial(link = "logit"), data = males_data, weights =
W_AgEthAdj)

  females_glm_sep <- glm(as.formula(glm_sep_spec),
                       family = binomial(link = "logit"), data = females_data, weights =
W_AgEthAdj)
```

```
  males_glm_tob <- glm(as.formula(glm_tob_spec),
                       family = binomial(link = "logit"), data = males_data, weights =
W_AgEthAdj)

  females_glm_tob <- glm(as.formula(glm_tob_spec),
                         family = binomial(link = "logit"), data = females_data, weights =
W_AgEthAdj)



  males_cw <- cw_calcs (cohort_sex = males_data,
                        sex_sep_mod = males_glm_sep,
                        sex_tob_mod =  males_glm_tob,
                        mod_set_name = model_name)

  females_cw <- cw_calcs (cohort_sex = females_data,
                          sex_sep_mod = females_glm_sep,
                          sex_tob_mod = females_glm_tob, #replace 06 with 2006
                          mod_set_name = model_name)

}


if (!is.null(glm_sep_title)& !is.null(glm_tob_title)){

root <- [insert file path]

males_glm_sep <- load(file = paste0 (root,"males_", glm_sep_title,".R"))
females_glm_sep <- load(file = paste0 (root,"females_", glm_sep_title,".R"))

males_glm_tob <- load(file = paste0 (root,"males_", glm_tob_title,".R"))
females_glm_tob <- load(file = paste0 (root,"females_", glm_tob_title,".R"))

males_cw <- cw_calcs (cohort_sex = males_data,
                      sex_sep_mod = get(males_glm_sep),
                      sex_tob_mod =  get(males_glm_tob),
                      mod_set_name = model_name)

females_cw <- cw_calcs (cohort_sex = females_data,
                        sex_sep_mod =  get(females_glm_sep),
                        sex_tob_mod = get(females_glm_tob), #replace 06 with 2006
                        mod_set_name = model_name)


}

coef_1 <- tidy (males_glm_sep)%>%
            mutate (Model_fam = model_name,
                    Model = "SEP",
                    Sex = 1,
                    CenYear = cohort,
                    Date_run = Sys.Date())
coef_2 <- tidy (males_glm_tob)%>%
            mutate (Model_fam = model_name,
                    Model = "Tob",
                    Sex = 1,
```

```
                    CenYear = cohort,
                    Date_run = Sys.Date())

coef_3 <- tidy (females_glm_sep)%>%
            mutate (Model_fam = model_name,
                    Model = "SEP",
                    Sex = 2,
                    CenYear = cohort,
                    Date_run = Sys.Date())

coef_4 <- tidy (females_glm_tob)%>%
            mutate (Model_fam = model_name,
                    Model = "Tob",
                    Sex = 2,
                    CenYear = cohort,
                    Date_run = Sys.Date())

fit_stats1 <- glance (males_glm_sep)%>%
                mutate (Model_fam = model_name,
                        Model = "SEP",
                        Sex = 1,
                        CenYear = cohort,
                        Date_run = Sys.Date())

fit_stats2 <- glance (males_glm_tob)%>%
                mutate (Model_fam = model_name,
                        Model = "Tob",
                        Sex = 1,
                        CenYear = cohort,
                        Date_run = Sys.Date())

fit_stats3 <- glance (females_glm_sep)%>%
                mutate (Model_fam = model_name,
                        Model = "SEP",
                        Sex = 2,
                        CenYear = cohort,
                        Date_run = Sys.Date())

fit_stats4 <- glance (females_glm_tob)%>%
                mutate (Model_fam = model_name,
                        Model = "Tob",
                        Sex = 2,
                        CenYear = cohort,
                        Date_run = Sys.Date())

fit_results <- bind_rows(fit_stats1, fit_stats2, fit_stats3, fit_stats4)
coeff_results <- bind_rows(coef_1, coef_2, coef_3, coef_4)
risk_results <- bind_rows(males_cw, females_cw)
return(list(coeff_results, fit_results, risk_results))

}
```

Bootstrap function

```
# Bootstrap ----------------------------------
```

```
boot_cw <- function (sex_cohort, indices, glm_sep_spec, glm_tob_spec, mod_set_name) {


  sep_glm <- glm(as.formula(glm_sep_spec),
                 family = binomial(link = "logit"), data = sex_cohort [indices, ], weights
= W_AgEthAdj)

  tob_glm <- glm(as.formula(glm_tob_spec),
                family = binomial(link = "logit"), data = sex_cohort[indices, ], weights =
W_AgEthAdj)

  cw_data <- sex_cohort[indices, ]
  cw_data$cross_ref <- cw_data$TotMaori
  cw_data$TotMaori <- ifelse (cw_data$TotMaori == "1",1,1)

  smo_data <- sex_cohort[indices, ]
  smo_data$SmkStat <- ifelse(smo_data$SmkStat == "3"|smo_data$SmkStat == "2", 1, 1)
  smo_data$SmkStat <- as.factor(smo_data$SmkStat)

  sep_pred <- predict.glm(sep_glm, type = "response")
  smo_pred <- predict.glm(tob_glm, type = "response")

  cw_pred <- predict.glm(sep_glm, newdata = cw_data, type = "response")
  cw_pred_smo <- predict.glm(tob_glm, newdata = cw_data, type = "response")
  never_pred <- predict.glm(tob_glm, newdata = smo_data, type = "response")

  cw_data$sep_pred <- sep_pred
  cw_data$smo_pred <- smo_pred
  cw_data$cw_pred <- cw_pred
  cw_data$cw_pred_smo <- cw_pred_smo
  cw_data$never_pred <- never_pred
  cw_data$mort <- ifelse(cw_data$Link == "1", 1, 0)

  risk <- cw_data %>%
            mutate(comb_weights = weights*W_AgEthAdj)%>%
            group_by(CenYear, Sex, cross_ref)%>%
            summarise (obs_risk = weighted.mean(mort, w = comb_weights),
                       pred_sep_risk = weighted.mean(sep_pred, w = weights),
                       cw_sep_risk = weighted.mean(cw_pred, w = weights),
                       pred_smo_risk = weighted.mean(smo_pred, w = weights),
                       cw_smo_risk = weighted.mean(cw_pred_smo, w = weights),
                       never_risk = weighted.mean (never_pred, w = weights))%>%
            mutate(Age_group_broad = "2574",
                   model_set = mod_set_name,
                   Date_run = Sys.Date())


  risk_age <- cw_data %>%
                mutate(comb_weights = weights*W_AgEthAdj)%>%
                group_by(CenYear, Sex, cross_ref, Age_group_broad)%>% #check age group
variable
                summarise (obs_risk = weighted.mean(mort, w = comb_weights),
                           pred_sep_risk = weighted.mean(sep_pred, w = weights),
                           cw_sep_risk = weighted.mean(cw_pred, w = weights),
                           pred_smo_risk = weighted.mean(smo_pred, w = weights),
                           cw_smo_risk = weighted.mean(cw_pred_smo, w = weights),
                           never_risk = weighted.mean (never_pred, w = weights))%>%
```

```
                mutate(model_set = mod_set_name,
                       Date_run = Sys.Date())%>%
    bind_rows (.,risk)

  results <- as.matrix(risk_age)
  return(results)

}
```

Functions to help with data prep

```
# Generic functions for data manipulation


### WRITING AGE FUNCTION TO RECODE DATA INTO 5 YEAR AGE GROUPS ###


age_5 <- function (x){
  if (x >= 0 & x <= 4) {
    return ("04")
  } else if (x >= 5 & x <= 9){
    return ("59")
  } else if (x >= 10 & x <= 14){
    return ("1014")
  } else if (x >= 15 & x <= 19){
    return ("1519")
  } else if (x >= 20 & x <= 24){
    return ("2024")
  } else if (x >= 25 & x <= 29){
    return ("2529")
  } else if (x >= 30 & x <= 34){
    return ("3034")
  } else if (x >= 35 & x <= 39){
    return ("3539")
  } else if (x >= 40 & x <= 44){
    return ("4044")
  } else if (x >= 45 & x <= 49){
    return ("4549")
  } else if (x >= 50 & x <= 54){
    return ("5054")
  } else if (x >= 55 & x <= 59){
    return ("5559")
  } else if (x >= 60 & x <= 64){
    return ("6064")
  } else if (x >= 65 & x <= 69){
    return ("6569")
  } else if (x >= 70 & x <= 74){
    return ("7074")
  } else if (x > 74){
    return ("7400")
  } else
    return ("NA")
```

```
}

### FUNCTION TO PUT CASES INTO BROADER AGE GROUPS ###

age_broad <- function (x){

  if (x > 0 & x <= 14){
    return ("014")
  } else if (x >= 15 & x <= 24){
    return ("1524")
  } else if (x >= 25 & x <= 44){
    return ("2544")
  } else if (x >= 45 & x <= 64){
    return ("4564")
  } else if (x >= 65 & x<= 74) {
    return ("6574")
  } else if (x > 74){
    return ("7400")
  } else
    return ("NA")
}




### INDICATOR FUNCTION FOR 65-74 YEAR OLDS ###

retired <- function (x) {

  if (x >= 65 & x <= 74){
    return ("2")
  } else if (x > 75){
    return ("3")
  } else
    return ("1")
}


### FUNCTION TO PRINT WHO STANDARD PROPORTIONS ###

WHO_stand_fun <- function (x) {
  if (x >= 0 & x <= 4) {
    return ("0.0886")
  } else if (x >= 5 & x <= 9){
    return ("0.0869")
  } else if (x >= 10 & x <= 14){
    return ("0.0860")
  } else if (x >= 15 & x <= 19){
    return ("0.0847")
  } else if (x >= 20 & x <= 24){
    return ("0.0822")
  } else if (x >= 25 & x <= 29){
    return ("0.0793")
  } else if (x >= 30 & x <= 34){
    return ("0.0761")
  } else if (x >= 35 & x <= 39){
    return ("0.0715")
```

```
  } else if (x >= 40 & x <= 44){
    return ("0.0659")
  } else if (x >= 45 & x <= 49){
    return ("0.0604")
  } else if (x >= 50 & x <= 54){
    return ("0.0537")
  } else if (x >= 55 & x <= 59){
    return ("0.0455")
  } else if (x >= 60 & x <= 64){
    return ("0.0372")
  } else if (x >= 65 & x <= 69){
    return ("0.0296")
  } else if (x >= 70 & x <= 74){
    return ("0.0221")
  } else if (x > 74){
    return ("0.0306")
  } else
    return ("NA")
}
```

Models

```
fam_p_int_sep <- "Link ~ Age_group + TotMaori + EqIncJen_log_cent + NZDep10_cent +
                         LabSt3 + HQual3 + Ret_lfs + Age_group*TotMaori +
                         TotMaori*EqIncJen_log_cent + TotMaori*NZDep10_cent +
                         TotMaori*LabSt3 + TotMaori*HQual3 + TotMaori*Ret_lfs"

fam_p_int_tob <- "Link ~ Age_group + TotMaori + EqIncJen_log_cent + NZDep10_cent +
                         LabSt3 + HQual3 + Ret_lfs + Age_group*TotMaori +
                         TotMaori*EqIncJen_log_cent + TotMaori*NZDep10_cent +
                         TotMaori*LabSt3 + TotMaori*HQual3 + TotMaori*Ret_lfs +
                         SmkStat + SmkStat*TotMaori + SmkStat*Age_group +
                         SmkStat*Age_group*TotMaori"
```

Code to run mediation functions

```
## Prior interactions results ------------------------------------------------------


# Risks, model fit and coefficients
p_int_1981 <- run_mediation(cohort = 1981,
                            glm_sep_title = NULL,
                            glm_tob_title = NULL,
                            glm_sep_spec = fam_p_int_sep,
                            glm_tob_spec = fam_p_int_tob,
                            model_name = "Prior Interaction")

p_int_1996 <- run_mediation(cohort = 1996,
                            glm_sep_title = NULL,
                            glm_tob_title = NULL,
                            glm_sep_spec = fam_p_int_sep,
                            glm_tob_spec = fam_p_int_tob,
```

```
                              model_name = "Prior Interaction")

p_int_2006 <- run_mediation(cohort = 2006,
                              glm_sep_title = NULL,
                              glm_tob_title = NULL,
                              glm_sep_spec = fam_p_int_sep,
                              glm_tob_spec = fam_p_int_tob,
                              model_name = "Prior Interaction")
```

eAppendix: Weights

Two weights were used in analyses: a weight to adjust for linkage bias in regression prediction, and a weight to adjust (or standardize) for confounding by age.

Weighting for linkage bias

The anonymous and probabilistic linkage of census and mortality records meant that not all decedents could be linked back to their previous census record. The main reason for this non-linkage was moving residence between census night and death, as the geocode was an important linking variable. Errors in other linkage variables (day of birth, month of birth, year of birth, ethnicity, sex, country of birth) lowered linkage rates as well.[1] Fortunately, the mortality file itself included some demographic data (sex, age, ethnicity, region) and one socio-economic factor (neighborhood deprivation, NZDep), allowing us to calculate the probabilities of linkage by these cross-classified strata. To allow for incomplete linkage and mitigate against misclassification bias of the mortality outcome (including differential, by ethnicity and socio-economic factors), in the main cohort analyses each linked census-mortality record was weighted by the inverse probability of linkage (e.g. if 20 out of 25 European male decedents aged 45-49 living in the middle quintile of deprivation in the north of NZ were linked, then each of these records was weighted by 25/20 = 1.25; hereafter called linkage weights). Conversely, census records not linked to a mortality record were slightly down-weighted (i.e. weights just less than 1.0) so that the weighted sum of census respondents in each sex by age by socio-demographic was the same as that actually observed.

Weighting for confounding

The weighted multiple mediator approach uses weighting to adjust for confounding. For our analyses, by sex, age was the only confounder. We also wanted to ensure that the same weighting was used across all six census by sex cohorts. We therefore elected to use weighting equivalent to the WHO World Standard [2], with an added advantage that the risks reported in this study are comparable with existing studies WHO age-standardization weights. We first describe the usual weighting method (p. 122-5 of [3], [4]), then our variant WHO standardization method, and conclude with an example to demonstrate both methods, and their similarities and dissimilarities.

For $E[Y_{a*M*}]$ the usual approach is to calculate stabilized inverse-probability-of-exposure weights:

$$\frac{P(A = a*)}{P(A = a* | C = c_i)} \quad (1)$$

where P(A=a*) is the overall proportion of census respondents who are European (for the population restricted to Europeans and Māori, within each census cohort), and P(A=a*|C=c_i) denotes the proportion European conditional on (or within each stratum of) confounder age (analyses were undertaken separately by sex).

We instead standardized age according to the WHO direct standard, using the weight generated in equation 2 applied to each observation:

$$\frac{\frac{w_i}{\sum_{j=min}^{j=max} w_j}}{P(C = c_i | A = a*)} \quad (2)$$

where c_i is the five-year age-group i under consideration (e.g. 60-64 year olds), a^* denotes Europeans, w_i is the WHO World Standard weight for age-group i (e.g. 0.0372 for 60-64 year olds [2]), and $\sum_{j=min}^{j=max} w_j$ is the sum of WHO World Standard weights for the study population at hand (e.g. 0.5413 for 25-74 years). The numerator of equation 2 is the WHO World Standard weights, scaled for the ages included in the study. Thus, 0.0372/0.5413 = 0.0687 of 25-74 year olds are age 60-64 years. The denominator of equation 2 is the proportion of the study population, a^*, that is in each age group. Thus, the weight rebalances the cohort using weights applied to individuals, to give the same standardized risks has each age group risk been calculated first, then weighted by the WHO World Standard weights.

Similarly, $E[Y_{aM}]$ is the weighted average risk of mortality among Māori, with weights:

$$\frac{\frac{w_i}{\sum_{j=min}^{j=max} w_j}}{P(C = c_i \,|A = a)} \tag{3}$$

To demonstrate these two weightings, the table below presents hypothetical data for two populations, a and a*, with different age structures (flat in population a, young in population a^*) and mortality risks in a^* that are roughly double a within each age strata. The strong difference in age structure generates confounding when comparing the crude mortality risks (RR = 0.80 for a^* compared to a).

The middle panel of the table shows the workings for the usual weighting method, and the right panel for our WHO Standard Population weighting method. Both methods achieve the same weights within age strata across the two populations, and a total pseudo-population (i.e. weighted

sum of actual population) that equals the actual starting population. And both methods return a RR that is close to 2.0; it is not exactly 2.0 as the within age strata RRs randomly vary around 2.0.

In the actual analyses in this paper, these weights were assigned to individuals and used to generate weighted mean values of the expected mortality risks for 25-74 year olds combined. For narrower age groups (i.e. 25-44, 45-64 and 65-74 year olds), we simply altered equation 3 to be for this smaller age range.

1. Blakely T, Salmond C. Probabilistic record linkage and a method to calculate the positive predictive value. *International Journal of Epidemiology* 2002;**31**:1246 - 1252.

2. Ahmad O, Boschi-Pinto C, Lopez A, Murray C, Lozano R, Inoue M. Age standardization of rates: a new WHO standard. Geneva: World Health Organization, 2000.

3. VanderWeele TJ. *Explanation in Causal Inference: Methods for Mediation and Interaction* Oxford University Press, 2015.

4. VanderWeele TJ, Vansteelandt S. Mediation Analysis with Multiple Mediators. Epidemiol Method 2013;2(1):95-115.

Table of two hypothetical populations, under the usual and our WHO World Standard variant weighting methods

						Standard weighting method (eqn 1 above)					WHO age-standardization weighting method (eqn 2&3)			
	Confounder (C) = age group	**N**	**Deaths**	**Mortality risk**		**Weights**	**Pseudo-population**	**Proportion pseudo-population**	**Weighted deaths**		**Weights**	**Pseudo-population**	**Proportion pseudo-population**	**Weighted deaths**
Group 1 (a)	25-29	10000	5	0.0005		1.162	11618.1	0.116	5.8		1.465	14649.9	0.146	7.3
	30-34	10000	10	0.0010		1.103	11033.6	0.110	11.0		1.406	14058.7	0.141	14.1
	35-39	10000	20	0.0020		1.057	10566.0	0.106	21.1		1.321	13208.9	0.132	26.4
	40-44	10000	40	0.0040		1.019	10192.0	0.102	40.8		1.217	12174.4	0.122	48.7
	45-49	10000	80	0.0080		0.989	9892.7	0.099	79.1		1.116	11158.3	0.112	89.3
	50-54	10000	160	0.0160		0.965	9653.3	0.097	154.5		0.992	9920.6	0.099	158.7
	55-59	10000	320	0.0320		0.946	9461.8	0.095	302.8		0.841	8405.7	0.084	269.0
	60-64	10000	640	0.0640		0.931	9308.5	0.093	595.7		0.687	6872.3	0.069	439.8
	65-69	10000	1280	0.1280		0.919	9186.0	0.092	1175.8		0.547	5468.3	0.055	699.9
	70-74	10000	2560	0.2560		0.909	9087.9	0.091	2326.5		0.408	4082.8	0.041	1045.2
	Sum	100000	5115				100000	1.000	4713.2			100000	1.000	2798.4
	Mortality risk			0.0512					0.0471					0.0280
Group 2 (a*)	25-29	3361	3.9	0.0012		0.519	1742.7	0.116	2.0		0.654	2197.5	0.146	2.5
	30-34	2689	5.3	0.0020		0.616	1655.0	0.110	3.3		0.784	2108.8	0.141	4.2
	35-39	2151	7.7	0.0036		0.737	1584.9	0.106	5.7		0.921	1981.3	0.132	7.1
	40-44	1721	13.3	0.0077		0.888	1528.8	0.102	11.8		1.061	1826.2	0.122	14.1
	45-49	1377	25.9	0.0188		1.078	1483.9	0.099	28.0		1.216	1673.7	0.112	31.5
	50-54	1101	38.8	0.0352		1.315	1448.0	0.097	51.0		1.351	1488.1	0.099	52.4
	55-59	881	56.2	0.0638		1.611	1419.3	0.095	90.5		1.431	1260.9	0.084	80.4
	60-64	705	109.1	0.1548		1.981	1396.3	0.093	216.1		1.463	1030.9	0.069	159.6
	65-69	564	134.3	0.2382		2.444	1377.9	0.092	328.1		1.455	820.2	0.055	195.3
	70-74	451	217.6	0.4824		3.022	1363.2	0.091	657.6		1.358	612.4	0.041	295.4
	Sum	15000	612				15000	1.000	1394.2			15000	1.000	842.6
	Mortality risk			0.0408					0.0929					0.0562
a* versus a			RR =	0.80				RR =	1.97				RR =	2.01