

Supplementary Table 1. Concordance of smoking status as per latest UK Biobank record and latest TPP record, brackets in cells correspond to the row percentage (n=109,641 with smoking data recorded in both datasets). Median time between latest record and study period was 4.5 years (IQR: 0.9-13.1 years).

Biobank smoking data	TPP smoking data				
	Never-smoker	Ex-smoker	Light smoker	Moderate smoker	Heavy smoker
Never-smoker	34,645 (75.96)	10,898 (23.90)	39 (0.09)	17 (0.04)	<10
Ex-smoker	3,170 (5.77)	51,547 (93.89)	97 (0.18)	58 (0.11)	28 (0.05)
Light smoker	68 (3.42)	1,549 (77.96)	282 (14.19)	80 (4.03)	<10
Moderate smoker	86 (2.10)	3,038 (74.15)	304 (7.42)	562 (13.72)	107 (2.61)
Heavy smoker	60 (1.97)	2,110 (69.18)	112 (3.67)	372 (12.20)	396 (12.98)

Supplementary Table 2. Concordance of smoking status as per latest UK Biobank record and latest EMIS record, brackets in cells correspond to the row percentage (n=229,583 with smoking data recorded in both datasets). Median time between latest record and study period was 1.1 years (IQR: 0.4-3.4 years).

Biobank smoking data	EMIS smoking data				
	Never-smoker	Ex-smoker	Light smoker	Moderate smoker	Heavy smoker
Never-smoker	129,776 (97.21)	3,705 (2.78)	14 (0.01)	<10 (0.01)	<10 (<0.01)
Ex-smoker	21,941 (26.56)	60,309 (72.99)	164 (0.20)	124 (0.15)	84 (0.10)
Light smoker	258 (8.42)	2,083 (68.01)	573 (18.71)	141 (4.60)	<10 (0.26)
Moderate smoker	349 (5.97)	3,762 (64.32)	580 (9.92)	997 (17.05)	161 (2.75)
Heavy smoker	202 (4.45)	2,791 (61.44)	200 (4.40)	707 (15.56)	643 (14.15)

Supplementary Table 3. Output of maximally adjusted sample selection logistic regression model (likelihood to receive a SARS-CoV-2 test), which was used to generate inverse probability weights for the maximally adjusted observational analysis model for effect of smoking on confirmed infection.

Parameter	(tested / not tested)	Odds ratio (95% confidence interval)
<i>Smoking status</i>		
Never smoker (reference)	(6,565/226,930)	1
Former smoker	(5,135/139,482)	1.16 (1.12 to 1.20)
Light smoker	(200/5,745)	1.14 (0.96 to 1.35)
Moderate smoker	(471/11,421)	1.35 (1.19 to 1.55)
Heavy smoker	(457/8,977)	1.54 (1.32 to 1.79)
<i>Sex</i>		
Female (reference)	(6,841/224,115)	1
Male	(6,457/180,349)	1.06 (1.02 to 1.10)
<i>Age category</i>		
40 – 49 years	(33/1,160)	0.96 (0.68 to 1.37)
50 – 59 years (reference)	(2,644/93,494)	1
60 – 69 years	(1,625/135,812)	0.92 (0.88 to 0.97)
70 – 79 years	(6,527/166,822)	1.29 (1.23 to 1.35)
80 + years	(469/7,176)	2.09 (1.89 to 2.31)
<i>Ethnicity</i>		
White (reference)	(12,274/378,752)	1
Mixed race	(91/2,510)	1.12 (0.90 to 1.38)
Asian or Asian British	(273/7,083)	1.10 (0.97 to 1.24)
Chinese	(34/1,375)	0.85 (0.60 to 1.20)
Other Asian	(53/1,600)	0.95 (0.72 to 1.25)
Black or Black British	(325/7,175)	1.29 (1.15 to 1.45)
Other	(168/2,041)	1.24 (1.06 to 1.45)
<i>Townsend deprivation quintile</i>		
1 (reference)	(5,543/186,697)	1
2	(2,740/90,549)	1.0 (0.96 to 1.06)
3	(2,021/58,677)	1.13 (1.07 to 1.19)
4	(1,963/47,637)	1.31 (1.24 to 1.39)
5 (least affluent)	(1,031/20,904)	1.48 (1.38 to 1.59)
<i>Comorbidities</i>		
Bronchiectasis	(45/499)	1.75 (1.27 to 2.40)
Chronic liver disease	(100/1,096)	2.22 (1.80 to 2.74)
Cystic fibrosis	(<10 / <10)	3.61 (0.39 to 33.5)
Diabetes mellitus	(1,262/20,392)	1.58 (1.47 to 1.67)
Interstitial lung disease	(30/251)	2.54 (1.73 to 3.75)
Asthma	(154/2,212)	1.65 (1.39 to 1.95)
Atrial fibrillation	(525/7,555)	1.62 (1.48 to 1.78)
Chronic obstructive pulmonary disease	(291/3,048)	1.96 (1.72 to 2.22)
Chronic kidney disease		<i>Not estimable</i>
Stage 2		
Stage 3	(<10/21)	4.61 (1.66 to 12.79)
Stage 4	(<10/18)	5.09 (1.63 to 15.92)
Stage 5 / End-stage renal failure	(48/197)	5.54 (3.99 to 7.69)
Unspecified	(58/343)	3.71 (2.78 to 4.94)
Congestive cardiac failure	(151/1,166)	1.92 (1.60 to 2.29)
Hypertension	(75/1,115)	1.54 (1.21 to 1.96)
Ischaemic heart disease	(1,463/23,197)	1.54 (1.45 to 1.64)
Lung cancer	(61/634)	2.04 (1.57 to 2.69)

Supplementary Table 4. Results from multivariable logistic regression models examining effect of smoking on outcomes of interest with serial adjustment guided by directed acyclic graphs. CI = confidence interval, COVID-19 = novel coronavirus disease 2019, BMI = body mass index, A = interstitial lung disease, cystic fibrosis, bronchiectasis, chronic liver disease, diabetes (type 1 or 2), B = lung cancer, asthma, chronic obstructive pulmonary disease, hypertension, ischaemic heart disease, congestive cardiac failure, chronic kidney disease, atrial fibrillation.

Smoking category	(cases / controls)	Odds ratio (95% CI) adjusted for age and sex	Odds ratio (95% CI) adjusting for age, sex, ethnicity, deprivation and non-smoking related comorbidity ^A	Odds ratio (95% CI) adjusting for age, sex, ethnicity, deprivation, non- smoking related comorbidity ^A , smoking-related comorbidity ^B and BMI
<i>Confirmed SARS-CoV-2 infection (1,649 / 419,444)</i>				
Never smoker	849 / 251,126	1 (reference)	1 (reference)	1 (reference)
Former smoker	717 / 154,708	1.34 (1.21 to 1.48)	1.31 (1.19 to 1.45)	1.26 (1.13 to 1.40)
Light smoker (1-9/day)	18 / 3,923	1.28 (0.80 to 2.04)	1.09 (0.68 to 1.74)	1.09 (0.67 to 1.77)
Moderate smoker (10-19/day)	26 / 5,753	1.27 (0.86 to 1.88)	1.12 (0.75 to 1.65)	1.13 (0.76 to 1.69)
Heavy smoker (20+/day)	13 / 3,934	0.90 (0.52 to 1.56)	0.79 (0.46 to 1.38)	0.75 (0.42 to 1.33)
Never smoker	849 / 251,126	1 (reference)	1 (reference)	1 (reference)
Former smoker	717 / 154,708	1.34 (1.21 to 1.48)	1.31 (1.19 to 1.45)	1.26 (1.13 to 1.40)
Current smoker	57 / 13,610	1.16 (0.89 to 1.52)	1.01 (0.77 to 1.33)	1.00 (0.76 to 1.32)
<i>Confirmed SARS-CoV-2 infection (weighted by inverse probability of receiving a test)</i>				
Never smoker	849 / 251,126	1 (reference)	1 (reference)	1 (reference)
Former smoker	717 / 154,708	1.07 (0.97 to 1.19)	1.10 (0.98 to 1.22)	1.12 (1.00 to 1.25)
Light smoker (1-9/day)	18 / 3,923	1.01 (0.63 to 1.64)	1.09 (0.66 to 1.79)	1.11 (0.67 to 1.83)
Moderate smoker (10-19/day)	26 / 5,753	0.82 (0.55 to 1.23)	0.81 (0.54 to 1.23)	0.86 (0.57 to 1.30)
Heavy smoker (20+/day)	13 / 3,934	0.50 (0.29 to 0.89)	0.54 (0.30 to 0.97)	0.55 (0.31 to 0.99)
Never smoker	849 / 251,126	1 (reference)	1 (reference)	1 (reference)
Former smoker	717 / 154,708	1.07 (0.97 to 1.19)	1.10 (0.98 to 1.22)	1.12 (1.00 to 1.24)
Current smoker	57 / 13,610	0.76 (0.57 to 1.00)	0.78 (0.59 to 1.04)	0.82 (0.61 to 1.09)

<i>Hospitalisation related to COVID-19 (968 / 419,444)</i>							
Never smoker	440 / 251,126	1	(reference)	1	(reference)	1	(reference)
Former smoker	457 / 154,708	1.44	(1.26 to 1.64)	1.40	(1.22 to 1.60)	1.31	(1.14 to 1.50)
Light smoker (1-9/day)	12 / 3,923	1.82	(1.03 to 3.24)	1.43	(0.80 to 2.54)	1.31	(0.69 to 2.46)
Moderate smoker (10-19/day)	25 / 5,753	2.59	(1.72 to 3.87)	2.13	(1.41 to 3.19)	2.26	(1.50 to 3.41)
Heavy smoker (20+/day)	14 / 3,934	1.97	(1.16 to 3.37)	1.64	(0.96 to 2.80)	1.65	(0.96 to 2.84)
Never smoker	440 / 251,126	1	(reference)	1	(reference)	1	(reference)
Former smoker	457 / 154,708	1.44	(1.26 to 1.64)	1.40	(1.22 to 1.60)	1.31	(1.14 to 1.50)
Current smoker	51 / 13,610	2.19	(1.63 to 2.92)	1.78	(1.32 to 2.39)	1.80	(1.26 to 2.29)
<i>Death related to COVID-19 (444 / 419,444)</i>							
Never smoker	159 / 251,126	1	(reference)	1	(reference)	1	(reference)
Former smoker	223 / 154,708	1.76	(1.43 to 2.16)	1.67	(1.36 to 2.06)	1.60	(1.29 to 1.99)
Light smoker (1-9/day)	<10 / 3,923	3.35	(1.57 to 7.16)	2.72	(1.27 to 5.82)	2.14	(0.87 to 5.24)
Moderate smoker (10-19/day)	20 / 5,753	6.65	(4.17 to 10.62)	5.54	(3.45 to 8.89)	5.91	(3.66 to 9.54)
Heavy smoker (20+/day)	16 / 3,934	7.44	(4.42 to 12.49)	6.09	(3.60 to 10.31)	6.11	(3.59 to 10.42)
Never smoker	159 / 251,126	1	(reference)	1	(reference)	1	(reference)
Former smoker	223 / 154,708	1.76	(1.43 to 2.16)	1.67	(1.36 to 2.06)	1.60	(1.29 to 1.97)
Current smoker	43 / 13,610	5.93	(4.22 to 8.33)	4.87	(3.44 to 6.88)	4.89	(3.41 to 7.00)

Supplementary Table 5. Results from models using smoking status derived from UK Biobank data only. The latest of UK Biobank follow-up assessment and baseline assessment were used. This sensitivity analysis was used to ascertain the possible impact of reverse causation on our results. Compared to using the most recent smoking status available, the results using UK Biobank data only (measured at median 10.6 years prior to 1st Jan 2020) preserved the direction of the risks, but with odds ratios of lower magnitude.

CI = confidence interval, COVID-19 = novel coronavirus disease 2019, BMI = body mass index, A = interstitial lung disease, cystic fibrosis, bronchiectasis, chronic liver disease, diabetes (type 1 or 2), B = lung cancer, asthma, chronic obstructive pulmonary disease, hypertension, ischaemic heart disease, congestive cardiac failure, chronic kidney disease, atrial fibrillation.

Smoking category	(cases/controls)	Odds ratio (95% CI) adjusted for age and sex	Odds ratio (95% CI) adjusting for demographics and non- smoking related comorbidity ^A	Odds ratio (95% CI) adjusting for demographics, non- smoking related comorbidity ^A , smoking- related comorbidity ^B and BMI
<i>Confirmed SARS-CoV-2 infection</i>		(1,567/403,934)		
Never smoker	(806/232,822)	1 (reference)	1 (reference)	1 (reference)
Former smoker	(645/143,976)	1.28 (1.15 to 1.43)	1.28 (1.15 to 1.43)	1.24 (1.11 to 1.39)
Light smoker (1-9/day)	(23/5,926)	1.09 (0.72 to 1.66)	0.97 (0.64 to 1.47)	0.99 (0.65 to 1.52)
Moderate smoker (10-19/day)	(53/11,829)	1.27 (0.96 to 1.67)	1.12 (0.85 to 1.49)	1.13 (0.85 to 1.50)
Heavy smoker (20+/day)	(40/9,381)	1.17 (0.85 to 1.61)	1.03 (0.75 to 1.42)	0.98 (0.71 to 1.37)
Never smoker	(806/232,822)	1 (reference)	1 (reference)	1 (reference)
Former smoker	(645/143,976)	1.28 (1.15 to 1.43)	1.28 (1.15 to 1.43)	1.24 (1.11 to 1.39)
Current smoker	(116/27,136)	1.19 (0.98 to 1.45)	1.06 (0.87 to 1.29)	1.04 (0.85 to 1.28)
<i>Hospitalisation related to COVID-19</i>		(909/403,934)		
Never smoker	(420/232,822)	1 (reference)	1 (reference)	1 (reference)
Former smoker	(402/143,976)	1.33 (1.16 to 1.53)	1.32 (1.15 to 1.52)	1.22 (1.05 to 1.41)
Light smoker (1-9/day)	(14/5,926)	1.44 (0.84 to 2.45)	1.21 (0.71 to 2.07)	1.24 (0.71 to 2.17)
Moderate smoker (10-19/day)	(35/11,829)	1.72 (1.21 to 2.42)	1.44 (1.01 to 2.04)	1.47 (1.03 to 2.08)

Heavy smoker (20+/day)	(38/9,381)	2.20	(1.57 to 3.07)	1.83	(1.30 to 2.57)	1.79	(1.27 to 2.52)
Never smoker	(420/232,822)	1	(reference)	1	(reference)	1	(reference)
Former smoker	(402/143,976)	1.33	(1.16 to 1.53)	1.32	(1.15 to 1.52)	1.22	(1.05 to 1.41)
Current smoker	(87/27,136)	1.83	(1.45 to 2.31)	1.53	(1.21 to 1.94)	1.54	(1.22 to 1.96)
<i>Death related to COVID-19</i>		(419/403,934)					
Never smoker	(166/232,822)	1	(reference)	1	(reference)	1	(reference)
Former smoker	(206/143,976)	1.53	(1.25 to 1.89)	1.49	(1.20 to 1.83)	1.41	(1.14 to 1.74)
Light smoker (1-9/day)	(<10/5,926)	2.04	(0.96 to 4.35)	1.76	(0.82 to 3.76)	1.38	(0.57 to 3.38)
Moderate smoker (10-19/day)	(22/11,829)	3.01	(1.93 to 4.71)	2.47	(1.58 to 3.88)	2.52	(1.60 to 3.97)
Heavy smoker (20+/day)	(18/9,381)	2.97	(1.82 to 4.83)	2.36	(1.43 to 3.86)	2.33	(1.41 to 3.86)
Never smoker	(166/232,822)	1	(reference)	1	(reference)	1	(reference)
Former smoker	(206/143,976)	1.53	(1.25 to 1.89)	1.49	(1.20 to 1.83)	1.41	(1.14 to 1.74)
Current smoker	(47/27,136)	2.79	(2.02 to 3.87)	2.29	(1.64 to 3.12)	2.23	(1.58 to 3.15)

Supplementary Table 6. Two-sample Mendelian randomisation results for the effects of smoking behaviours on COVID-19 susceptibility and severity (281,105 White British participants)

Exposure	Outcome	MR method	OR (95% CI)	P-value
Smoking initiation	Confirmed infection	IVW	1.45 (1.10-1.91)	0.01
		MR-Egger	1.30 (0.41-4.08)	0.66
		Weighted median	1.53 (1.02-2.28)	0.04
		MR-PRESSO	1.45 (1.10-1.91)	0.01
	Hospitalisation	IVW	1.60 (1.13-2.27)	0.01
		MR-Egger	1.69 (0.40-7.26)	0.48
		Weighted median	1.62 (0.96-2.72)	0.07
		MR-PRESSO	1.60 (1.14-2.26)	0.01
	Death	IVW	1.35 (0.82-2.22)	0.23
		MR-Egger	0.35 (0.04-2.82)	0.33
		Weighted median	1.54 (0.73-3.23)	0.25
		MR-PRESSO	1.35 (0.82-2.22)	0.23
Smoking heaviness (Ever smoked)	Confirmed infection	IVW	2.51 (1.20-5.24)	0.01
		MR-Egger	3.40 (0.90-12.87)	0.07
		Weighted median	4.25 (1.33-13.57)	0.01
		MR-PRESSO	2.51 (1.28-4.89)	0.01
	Hospitalisation	IVW	5.09 (2.04-12.67)	<.001
		MR-Egger	7.14 (1.37-37.14)	0.02
		Weighted median	6.57 (1.54-28.06)	0.01
		MR-PRESSO	5.09 (2.29-11.31)	<.001
	Death	IVW	10.02 (2.53-39.73)	0.001
		MR-Egger estimate	17.68 (1.42-219.98)	0.03
		Weighted median	21.68 (2.77-169.42)	0.003
		MR-PRESSO	10.02 (2.53-39.73)	0.002

Supplementary Table 7. Tests of heterogeneity, directional pleiotropy, directionality, and regression dilution statistics (n = 281,105)

Exposure	Outcome	Method	Estimate	P-value
Smoking initiation	Confirmed infection	Q statistic	398.10	0.22
		Egger intercept	0.00	0.84
		Rucker's Q difference	0.04	0.84
		MR-PRESSO global test	--	0.24
		Steiger directionality test	--	<.001
		I^2_{GX}	0.98	
	Hospitalisation	Q statistic	369.51	0.60
		Egger intercept	-0.00	0.94
		Rucker's Q difference	0.01	0.94
		MR-PRESSO global test	--	0.62
		Steiger directionality test	--	<.001
		I^2_{GX}	0.98	
	Death	Q statistic	376.28	0.50
		Egger intercept	0.03	0.19
		Rucker's Q difference	1.70	0.19
		MR-PRESSO global test	--	0.51
		Steiger directionality test	--	<.001
		I^2_{GX}	0.98	
Smoking heaviness (Ever smoked)	Confirmed infection	Q statistic	44.35	0.82
		Egger intercept	-0.01	0.59
		Rucker's Q difference	0.29	0.59
		MR-PRESSO global test	--	0.79
		Steiger directionality test	--	<.001
		I^2_{GX}	0.99	
	Hospitalisation	Q statistic	41.35	0.90
		Egger intercept	-0.01	0.63
		Rucker's Q difference	0.23	0.63
		MR-PRESSO global test	--	0.88
		Steiger directionality test	--	<.001
		I^2_{GX}	0.99	
	Death	Q statistic	66.28	0.12
		Egger intercept	-0.02	0.60
		Rucker's Q difference	0.35	0.56
		MR-PRESSO global test	--	0.12
		Steiger directionality test	--	<.001
		I^2_{GX}	0.99	

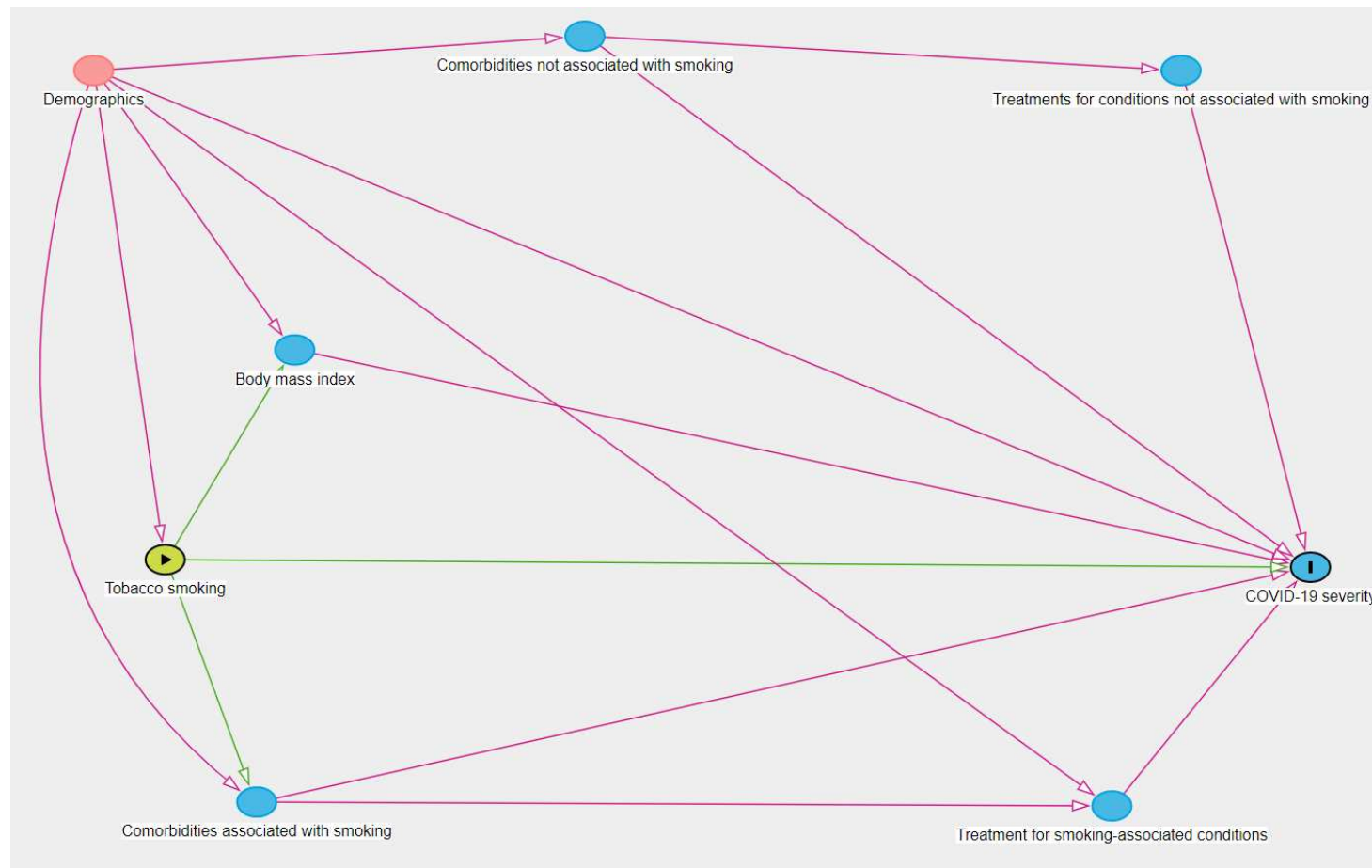
Supplementary Table 8. Sensitivity analyses of smoking behaviours on COVID-19 susceptibility and severity using SNP-exposure estimates that exclude UK Biobank and 23andMe

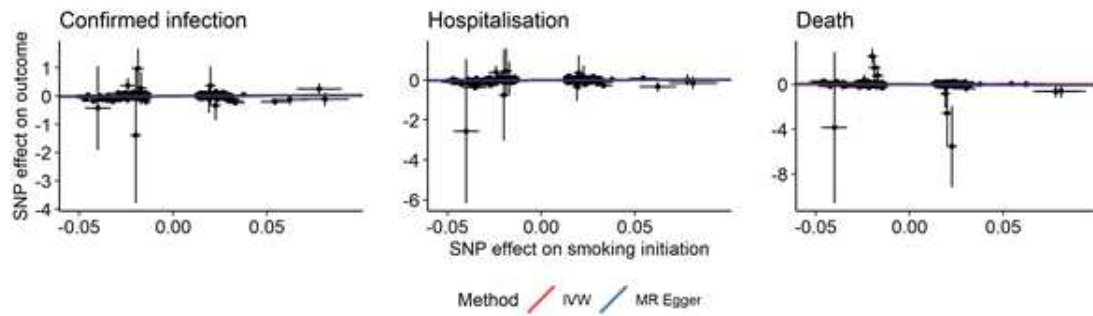
Exposure	Outcome	MR method	OR (95% CI)	P-value
Smoking initiation	Confirmed infection	IVW	2.10 (1.06-4.15)	0.03
		MR-Egger	0.51 (0.09-3.00)	0.45
		Weighted median	2.00 (0.73-5.44)	0.18
		MR-PRESSO	2.10 (1.06-4.15)	0.03
	Hospitalisation	IVW	2.76 (1.17-6.50)	0.02
		MR-Egger	0.84 (0.09-8.05)	0.88
		Weighted median	2.37 (0.65-8.63)	0.19
		MR-PRESSO	2.76 (1.18-6.47)	0.02
	Death	IVW	2.04 (0.60-6.96)	0.26
		MR-Egger	0.37 (0.01-9.49)	0.55
		Weighted median	2.02 (0.32-12.66)	0.45
		MR-PRESSO	2.04 (0.60-6.95)	0.26
Smoking heaviness (Ever smoked)	Confirmed infection	IVW	2.45 (1.44-4.17)	<.001
		MR-Egger	3.65 (1.60-8.28)	0.001
		Weighted median	3.44 (1.64-7.22)	0.001
		MR-PRESSO	2.45 (1.56-3.85)	<.001
	Hospitalisation	IVW	4.20 (2.18-8.08)	<.001
		MR-Egger	6.08 (2.21-16.73)	<.001
		Weighted median	6.85 (2.67-17.56)	<.001
		MR-PRESSO	4.20 (2.48-7.12)	<.001
	Death	IVW	7.60 (2.99-19.31)	<.001
		MR-Egger estimate	17.80 (4.29-73.89)	<.001
		Weighted median	13.26 (3.71-47.46)	<.001
		MR-PRESSO	7.60 (2.99-19.31)	<.001

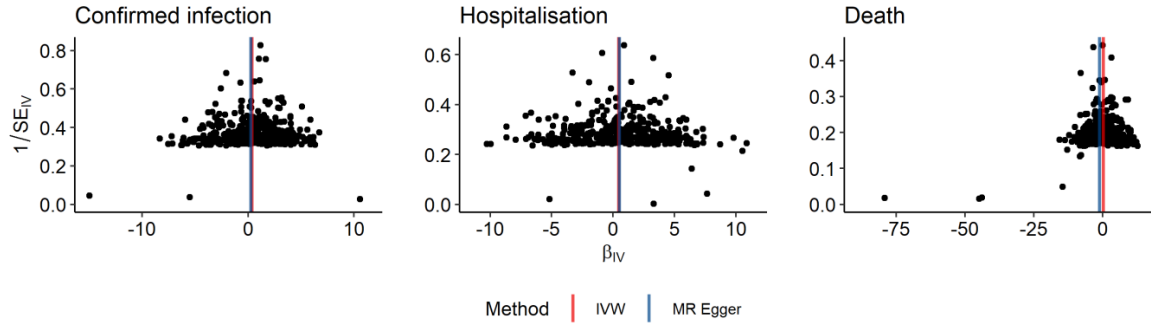
Supplementary Table 9. Sensitivity analyses of smoking behaviours on COVID-19 susceptibility and severity in 341,471 unrelated participants, unrestricted by ancestry.

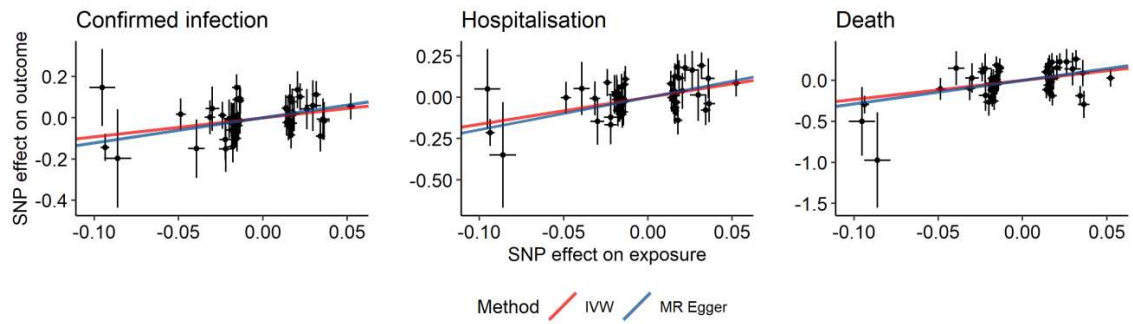
Exposure	Outcome	MR method	OR (95% CI)	P-value
Smoking initiation	Confirmed infection	IVW	1.29 (1.02-1.63)	0.03
		MR-Egger	1.31 (0.48-3.53)	0.60
		Weighted median	1.64 (1.15-2.32)	0.01
		MR-PRESSO	2.51 (1.28-4.89)	0.01
	Hospitalisation	IVW	1.43 (1.05-1.94)	0.02
		MR-Egger	2.27 (0.62-8.29)	0.22
		Weighted median	1.48 (0.94-2.33)	0.09
		MR-PRESSO	1.43 (1.06-1.92)	0.02
	Death	IVW	1.23 (0.79-1.92)	0.37
		MR-Egger	0.48 (0.07-3.16)	0.44
		Weighted median	1.53 (0.79-2.95)	0.21
		MR-PRESSO	1.23 (0.80-1.90)	0.35
Smoking heaviness (Ever smoked)	Confirmed infection	IVW	2.06 (1.06-4.02)	0.03
		MR-Egger	2.51 (0.75-8.41)	0.14
		Weighted median	2.51 (0.85-7.42)	0.10
		MR-PRESSO	2.06 (1.15-3.70)	0.02
	Hospitalisation	IVW	3.81 (1.68-8.64)	0.00
		MR-Egger	4.01 (0.91-17.65)	0.07
		Weighted median	4.38 (1.17-16.34)	0.03
		MR-PRESSO	3.81 (1.96-7.42)	<.001
	Death	IVW	10.35 (3.06-34.97)	<.001
		MR-Egger estimate	27.13 (2.99-246.25)	0.00
		Weighted median	14.93 (2.34-95.19)	0.00
		MR-PRESSO	10.35 (3.06-34.97)	<.001

Supplementary Figure 1. Directed acyclic graph ('causal diagram') used to inform the adjustment strategy in the logistic regression models for the observational analyses. Adjustment for non-smoking-related conditions and demographic factors (age, sex, ethnicity and deprivation) may identify the 'total effect' of smoking on COVID-19 severity, hence, the first adjustment set. Body mass index and smoking-related conditions are on indirect causal pathways (coloured green) from the exposure (smoking) and the outcomes of interest, and therefore adjustment for them may enable isolation of the direct causal effect.



Supplementary Figure 2. Scatter plots for smoking initiation (378 variants)

Supplementary Figure 3. Funnel plots for smoking initiation (378 variants)

Supplementary Figure 4. Scatter plots for smoking heaviness in ever smokers (55 variants)

Supplementary Figure 5. Funnel plots for smoking heaviness in ever smokers (55 variants)