Risk factors for ongoing symptomatic COVID-19 and post-COVID syndrome in the population: analyses of 10 longitudinal studies and electronic health records in the UK

List of Supplementary Tables

Supplementary Table S1. Details of each study

Supplementary Table S2. Ethics and data access statements for each study

Supplementary Table S3. Descriptives of analytic sample (self-reported COVID-19 symptoms) and non-analytic sample (no self-reported COVID-19 symptoms)

Supplementary Table S4. Length of time of symptoms (using self-report direct measures) by COVID-19 status (self-report)

Supplementary Table S5. Length of time of symptoms by COVID-19 status (swab/saliva and antibody test)

Supplementary Table S6. Length of time of symptoms (using sum of individual symptoms) by COVID-19 status (BIB; TwinsUK)

List of Supplementary Figures

Supplementary figure 1: Categorical age associations with symptoms 4+ week in a sub-set of the longitudinal studies

Supplementary figure 2: Categorical age associations with symptoms 12+ week in a sub-set of the longitudinal studies

Supplementary figure 3: Full meta-analysis results for sociodemographic characteristics with symptoms for 4+ weeks in the longitudinal studies

Supplementary figure 4: Full meta-analysis results for health traits with symptoms for 4+ weeks in the longitudinal studies

Supplementary figure 5: Full meta-analysis results for sociodemographic characteristics with symptoms for 12+ weeks in the longitudinal studies

Supplementary figure 6: Full meta-analysis results for health traits with symptoms for 12+ weeks in the longitudinal studies

Supplementary figure 7: Secondary meta-analysis results for sociodemographic characteristics with symptoms for 4+ weeks in the longitudinal studies, including inverse probability weights for COVID-19 risk

Supplementary figure 8: Secondary meta-analysis results for health traits with symptoms for 4+ weeks in the longitudinal studies, including inverse probability weights for COVID-19 risk

Supplementary figure 9: Secondary meta-analysis results for sociodemographic characteristics with symptoms for 12+ weeks in the longitudinal studies, including inverse probability weights for COVID-19 risk

Supplementary figure 10: Secondary meta-analysis results for health traits with symptoms for 12+ weeks in the longitudinal studies, including inverse probability weights for COVID-19 risk

Supplementary figure 11: Sub-group meta-analysis results for sociodemographic characteristics with symptoms for 4+ weeks in individuals with COVID-19 status confirmed by a positive PCR test and/or serology result

Supplementary figure 12: Sub-group meta-analysis results for health traits with symptoms for 4+ weeks in individuals with COVID-19 status confirmed by a positive PCR test and/or serology result

Supplementary figure 13: Sub-group meta-analysis results for sociodemographic characteristics with symptoms for 12+ weeks in individuals with COVID-19 status confirmed by a positive PCR test and/or serology result

Supplementary figure 14: Sub-group meta-analysis results for health traits with symptoms for 12+ weeks in individuals with COVID-19 status confirmed by a positive PCR test and/or serology result

List of Supplementary Information

Supplementary Note 1: Information governance and ethics for the OpenSAFELY platform

Supplementary Note 2: Detail of method to derive long COVID by monthly symptom reporting

Supplementary Note 3: Detail of method to derive inverse probability weights (IPW)

Supplementary Table S1.	Details	of each study
-------------------------	---------	---------------

pplementary rable 31. Details of each	study				
Study Population	Design and Sample Frame	2020 Age Range	Pre- pandemic Survey	Details of Covid surveys (response rate)	Analytic N
ge Homogenous Cohorts					
MCS: Millennium Cohort Study(1,2)	Cohort of UK children born between Sept 2000 and Jan 2002 with regular follow-up surveys from birth.	18-20	2018	Spring 2021 survey response with the issued sample: 33.1%	1055
ALSPAC (G1): Avon Longitudinal Study of Parents and Children- Generation 1(3)	Cohort of children born in the South-West of England between April 1991 and Dec 1992, with regular follow-up surveys from birth. (original young people)	27-29	2017-2018	Three questionnaires: April (19%), June (17.4%), December (26.4%)	668
NS: Next Steps, formerly known as Longitudinal Study of Young People in England(1,4)	Sample recruited via secondary schools in England at around age 13 with regular follow-up surveys thereafter.	29-31	2015	Spring 2021 survey response with the issued sample: 34.3%	848
BCS70: British Cohort Study 1970(1,5)	Cohort of all children born in Great Britain (i.e. England, Wales & Scotland) in one week in 1970, with regular follow-up surveys from birth.	50	2016	Spring 2021 survey response with the issued sample: 45.4%	889
NCDS: National Child Development Study(1,6)	Cohort of all children born in Great Britain (i.e. England, Wales & Scotland) in one week in 1958, with regular follow-up surveys from birth.	62	2013	Spring 2021 survey response with the issued sample: 58.5%	709
ge Heterogeneous Studies					
BIB: Born in Bradford(7,8)	Birth cohort recruiting pregnant women and their children between 2007 and 2011	28-55	2016-2020	Two surveys: April-Jun (30.7%) & Oct-Nov (39.9%)	110
USOC: Understanding Society: the UK Household Longitudinal Survey(9)	A nationally representative longitudinal household panel study, based on a clustered-stratified probability sample of UK households, with all adults aged 16+ in chosen households surveyed annually.	16-96	2018-2019	Seven surveys (full/partial interview): April 2020 (42.0%); May (35.1%); Jun (33.5%); July (32.6%); Sep (30.6%), Nov (28.6%), Jan 2021 (28.5%)	1033
GS: Generation Scotland: the Scottish Family Health Study(10)	A family-structured, population-based Scottish cohort, with participants aged 18-99 recruited between 2006-2011	27-100	2006-2011	Three surveys: April-Jun 2020 (21.3%); Jul-Aug 2020 (15.4%); Feb 2021 (14.3%)	335
ALSPAC(G0): Avon Longitudinal Study of Parents and Children- Generation 0(11)	Parents of the ALSPAC(G1) cohort described above, treated as a separate age-heterogenous study population. (original parents)	45-81	2011-2013	Three questionnaires: April (12.4%), June (12.2%), December (14.3%)	446
TWINSUK: the UK Adult Twin Registry(12,13)	A cohort of UK volunteer adult twins (55% monozygotic and 43% dizygotic) who were sampled between 18-101 years of age.	22-96	2017-2018	Three surveys: April (64.3%), July (77.6%) & November (76.1%)	806

Supplementary Table S2. Ethics and data access statements for each study

NCDS, BCS70,	The most recent sweeps of the NCDS, BCS70, Next Steps and MCS have all been granted ethical approval by the National Health Service (NHS) Research Ethics Committee and all participants
NS and MCS	have given informed consent. Data for NCDS (SN 6137), BCS70 (SN 8547), Next Steps (SN 5545), MCS (SN 8682) and all four COVID-19 surveys (SN 8658) are available through the UK
	Data Service.
ALSPAC	Ethical approval was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. The study website contains details of all the data that is available
	through a fully searchable data dictionary and variable search tool: http://www.bristol.ac.uk/alspac/researchers/our-data. ALSPAC data is available to researchers through an online proposal
	system. Information regarding access can be found on the ALSPAC website (http://www.bristol.ac.uk/media-library/sites/alspac/documents/researchers/data-
	access/ALSPAC_Access_Policy.pdf).
BIB	Ethical approval for Born in Bradford was granted by the National Health Service Health Research Authority Yorkshire and the Humber (Bradford Leeds) Research Ethics Committee (reference:
	16/YH/0320). Data from the various BiB family studies are available to researchers; see the study website for information on how to access data (https://borninbradford.nhs.uk/research/how-to-
	access-data/).
USOC	The University of Essex Ethics Committee has approved all data collection for the Understanding Society main study and COVID-19 waves. No additional ethical approval was necessary for
	this secondary data analysis. All data are available through the UK Data Service (SN 6614 and SN 8644).
GS	Generation Scotland obtained ethical approval from the East of Scotland Committee on Medical Research Ethics (on behalf of the National Health Service). Reference number 20/ES/0021.
	Access to data is approved by the Generation Scotland Access Committee. See https://www.ed.ac.uk/generation-scotland/for-researchers/access or email access@generationscotland.org for
	further details.
TWINSUK	All wave of TwinsUK have received ethical approval associated with TwinsUK Biobank (19/NW/0187), TwinsUK (EC04/015) or Healthy Ageing Twin Study (H.A.T.S) (07/H0802/84) studies
	from NHS Research Ethics Committees at the Department of Twin Research and Genetic Epidemiology, King's College London. The TwinsUK Resource Executive Committee (TREC) oversees
	management, data sharing and collaborations involving the TwinsUK registry (for further details see https://twinsuk.ac.uk/resources-for-researchers/access-our-data/).

Supplementary Table S3. D	Descriptives of anal	lytic sample (se	elf-reported C	OVID-19 sy	mptoms) and	the sample ex	cluded from t	he analysis (n	o self-reporte	d COVID-19	symptoms)									
	Analysis sample	MCS Excluded	Analysis sample	Excluded	Analysis sample	Next Steps Excluded	Analysis sample	BIB Excluded	Analysis sample	Usoc Excluded	Analysis sample	TwinsUK Excluded	Analysis sample	GS Excluded	A Analysis sample	Excluded	Analysis sample	BCS70 Excluded	Analysis sample	NCDS Excluded
Sample size	1055	3,293	668	3446	848	3,317	110		1033	9267	806	4610	335	2819	446	3890	889	4,815	709	6,063
Sumple Size	1000	5,275	000	5110	010	5,517	110	100							110	5670		1,010	107	0,000
Age, mean years (SD)	19.9 (0.3)	19.9 (0.3)	28.4	28.4	31.0 (0.3)	31.0 (0.3)	40.7 (5.9)	41.2 (5.7)	48.5 (14.8)	56.4 (16.1)	52.7 (15.85)	59.9 (15.78)	55.9 (10.6)	60.9 (11.5)	58.3 (4.4)	59.5 (4.8)	51 *	51 *	63 *	63 *
Female sex, %	61.8	60.3	63.8	66.6	61.8	62.3	96.4	94.3	65.3	57.5	88	87.8	64.2	63.7	67.9	70.8	57	57.7	54.9	53.4
Ethnicity, %																				
White	81.7	83.3	95.5	95.8	67.7	73.4	44.5	47.3	85.1	89.9	96.2	97.1	96.1	98.3	98.4	98.0	84	85.5	92	91.2
Na akisa shai asia sisa	18.2	16.2	4.5	4.1	20	24.0	50.9	49.8	12.2	9.7	2.7	2.0	1.5	0.6	1.4	1.0)	2	2.1	2.7	
Non-white ethnic minority	18.2	16.2	4.5	4.1	30	24.9			13.2	8.7	3.7	2.9	1.5	0.6	1.4	1.9)	3	2.1	2.7	1.4
Missing	0.1	0.6	0	0.1	2.4	1.8	4.6	2.9	1.7	1.4	0.1	0.2	2.4	1.1	0.2	0.1	12.9	12.5	5.4	9.4
Education, %																				
Degree	46.8	39.3	50.6	51.4	49.7	43.3	10	9.6	48.4	39.5	49.9	45.5	48.7	45.1	23.8	26.6	42.4	42.3	40.1	41.3
No degree	47.6	54.8	22.3	24.2	42.2	46.7	74.5	76	41.5	46.7	27.8	41.1	49.3	53.0	68.8	65.5	49.9	50.9	58.5	56.4
Missing	5.6	5.9	27.1	24.4	11.1	10	15.5	14.3	10.1	13.8	22.3	13.4	2.1	1.9	7.4	7.8	7.7	6.9	1.4	2.4
IMD quintile, %																				
	1 15.5	18.7	35	34.1	23.7	18.5	45.5	46.3	-	-	7.4	6.3	10.1	6.0	37.7	39.6	11.3	9.1	9.5	8.5
	2 15.2	16.5	24.3	23.6	18.5	19.6	30	24.4	-	-	14.6	12.9	11.9	9.7	27.4	24.9	15.1	13.4	15.1	13.8
	3 16.9	18.8	15.7	16.7	16.8	18.2	11.8	12.1	-	-	21.1	20.8	15.8	15.1	12.8	14.6	19.2	18.2	17.2	19.7
	4 22	19.3	11.8	11.7	17.3	16.9	8.2	8.4	-	-	26.9	26.2	22.7	28.2	9.2	8.8	18.2	21.8	22.4	22.4
	5 27.7	23.4	6.9	6.4	12.5	16.7	0.9	3.1	-	-	29.5	33	39.4	41.0	4.3	3.5	23.9	24.7	25.3	26.5
Missing	2.8	3.3	6.3	7.4	11.2	10.1	3.6	5.7	-	-	0.4	0.8			8.7	8.5	12.4	12.8	10.6	9.1
Occupational class, %																				
Managerial, Admin, Professional	NA	NA	18	14.9	NA	NA	23.6	28.3	38.9	31.1	-	-	52.8	47.3	12.8	11.5	NA	NA	NA	NA
Intermediate	NA	NA	41.9	39.2	NA	NA	32.7	27.7	16.6	16.4	-	-	17.9	17.5	29.2	35.8	NA	NA	NA	NA
Manual/Routine	NA	NA	25.6	31.6	NA	NA	19.1	16.8	21.3	17.4			11.0	8.2	42.6	38.1	NA	NA	NA	NA
Not in employment	NA	NA	0.3	0.5	NA	NA			20.5	33.6	-	-	0.0	0	1.1	0.7	NA	NA	NA	NA
Missing	NA	NA	14.2	13.8	NA	NA	24.5	27.3	2.7	1.5	-		18.2	27.0	14.4	14	NA	NA	NA	NA
Country, %																				
England	70.7	66.4	100	100	97.6	96.5	100	100	83.8	79.9	92.7	91.4	1.2	0.4	100	100	86.6	82.7	86.5	82.1
Scotland	8.8	13.3			0.6	0.6			6	9.5	3.2	4.5	98.8	99.5			6.4	8.5	6.4	8.8
Wales	12.9	11.1			1.1	0.7			6.7	6	3	3.1					5	5	5.4	5.2

	7.1	9.6			0.1	0.1			2.5	1.6	0.1	0.1	0.0	0			0	0.2	0.2	0.1
Northern Ireland Missing	7.1 0.5	8.6 0.6			0.1	0.1			3.5	4.6	0.1	0.1	0.0	0			0	0.2	0.3	0.1
	0.5	0.0			0.0	2			0	0	1						2	0	1.0	5.0
Pre-pandemic mental health, mean scale score (SD)	NA	NA	6.5 (6.2)	6.9 (6.4)	NA	NA	4.3 (5.4)	3.0 (3.8)	12.6 (6.2)	10.8 (5.2)	8.01 (6.13)	7.36 (5.97)	11.8 (6.3)	11.0 (5.4)	7.2 (5.6)	6.3 (5.3)	NA	NA	NA	NA
Pre-pandemic mental health categories, %												(5.5.7)								
Yes	15.4	15.8	14.7	17.9	22.2	22.9	10	5.7	28.1	15.8	3.7	51.5	11.6	8.8	13.9	11.3	15	14.3	11.7	11.6
No	79.2	77.6	55.7	54.4	65.1	65.4	67.3	73	69.3	82.6	40.9	3.3	72.8	71.2	65	68.8	16.7	68.4	78.8	78.3
Missing	5.4	6.6	29.6	27.7	12.7	11.7	22.7	21.3	2.6	1.6	55.3	42.5	15.5	20.0	21.1	19.9	16.3	17.3	9.5	10.1
Self-reported health, %																				ļ
Excellent	NA	NA	16.6	14.7	23.9	21.9	5.5	5.7	8.9	9.8	15.5	13.8			16.6	18	16.2	16.4	11.4	13.9
Very Good	NA	NA	36.8	35.7	35.3	36.7	21.8	24.4	32.7	37.9	23.7	26.3			20.2	27.6	30.6	32.6	33.3	34.4
Good	NA	NA	19.6	23.8	20.8	21.7	30.9	35.9	32.5	32.6	16.00%	18.4			30.9	27	26.1	24.2	29.1	29.2
Fair	NA	NA	4.3	5.1	60.5	6.9	13.6	10.2	15.7	13.7	4.8	5.9			7.4	5.5	10.9	10.6	12.7	11.5
Poor	NA	NA	1.3	1.5	1.3	1.8	5.5	2.5	6.1	3.9	1	0.8			2.9	1.4	3.8	3.7	3.8	3.7
Missing	NA	NA	21.3	19.1	12.3	11	22.7	21.3	4.1	2.1	39	34.8			22	20.6	12.4	12.5	9.7	7.4
Pre-pandemic BMI, mean kg/m2 (SD)	23.3 (4.7)	23.2 (4.7)	25.1 (5.3)	24.8 (5.1)	25.4 (5.8)	25.2 (5.3)	26.8 (5.3)	26.0 (5.5)	-	-	26.2 (4.91)	26.7 (5.17)	26.9 (5.6)	26.5 (5.0)	27.2 (4.9)	26.6 (4.8)	28.7 (5.9)	28.2 (5.3)	27.6 (5.0)	27.1 (5.1)
Pre-pandemic BMI categories, %																				
>18.5	4.6	6.9	1.8	2.1	1.5	2.4	1	1.6	-	-	1.4	1.1	1.5	0.7	0.2	0.7	0	0.3	0.5	0.5
18.5 - 24.9999	63.5	59.6	41	39.0	46.9	47.9	31.8	40.4	-	-	19.7	26.1	32.5	34.1	2	26.5	22.8	24.3	26.8	31.9
25 - 29.9999	15.5	16.6	16.5	16.3	21.7	20.9	21.8	22.7	-	-	16.5	20.4	33.7	28.7	25.3	25	29.7	29.3	35.3	35.6
30+	8.7	8.4	10.3	8.9	13.6	14	20.9	16.4	-	-	8.9	10.8	16.4	16.3	13.9	12.8	27.3	25.7	22.7	19.4
Missing	7.7	8.6	30.4	33.7	16.3	14.8	24.5	18.9	-	-	53.5	41.6	15.8	20.3	38.6	35	20.1	20.4	14.7	12.6
Diabetes, %																				
No	NA	NA	67.8	66.5	NA	NA	NA	NA	93	92	65.1	64.9	82.4	78.5	76.9	77.2	84.5	84.6	83.4	87.7
Yes	NA	NA	0.3	0.4	NA	NA	5.5	3.9	7	8	1.7	3.1	2.1	1.5	1.1	2.3	3.2	2.9	6.8	4.9
Missing	NA	NA	31.9	33.1	NA	NA	94.5	96.1	-	-	33.1	31.9	15.5	20.0	22	20.5	12.4	12.5	9.9	7.5
Hypertension, %																				ļ
No	NA	NA	67.2	65.6	NA	NA	NA	NA	79.7	74.1	60.9	57.6	77.3	69.4	65.5	70.3	NA	NA	69	73.6
Yes	NA	NA	0.8	1.2	NA	NA	12.7	5.9	20.3	25.9	12.5	19.7	7.2	10.6	15.3	13.1	NA	NA	21.3	19.9
Missing	NA	NA	32.0	33.2	NA	NA	87.3	94.1	-	-	26.6	22.7	15.5	20.0	19.3	16.6	NA	NA	9.7	7.5
High cholesterol, %																				
No	NA	NA	67.1	63.5	NA	NA	Na	NA	-	-	58.4	57.6			52	54.9	NA	NA	NA	NA
Yes	NA	NA	0	0.1	NA	NA	1.8	1.8	-	-	15.1	23.2			6.5	7.6	NA	NA	NA	NA
Missing	NA	NA	32.9	36.4	NA	NA	98.2	98.2	-	-	26.4	22.2			41.5	37.5	NA	NA	NA	NA
Asthma, %	74.7	76.2	40.4	49.7	NA	NA	NA	NA	78.9	85.3	7.3	10.6	70.7	72	65.7	68	78	77.3	80.5	81.8
No			49.4																	
Yes	10.9	10.2	18	17.1	NA	NA	10.9	12.3	21.1	14.7	7.2	7.9	13.7	8.1	11.7	11.2	9.7	10.2	9.6	10.7

Missing	14.4	13.6	32.6	33.2	NA	NA	89.1	87.7	-	-	85.5	81.6	15.5	20.0	22.7	20.8	12.4	12.5	9.9	7.5
Current smoker, %																				
No	43.2	48.7	49.4	49.7	64.4	61.2	84.5	83.2	91.5	92.2	85.9	88.5	93.4	94.6	57.2	61.1	72.4	69	79	77.8
Yes	11.8	8.4	18	17.1	8	10.1	5.5	6.6	8.4	7.8	13.8	10.7	5.7	4.8	28.3	26.2	7.3	10	4.5	7.5
Missing	45	42.9	32.63	33.2	27.6	28.7	10	10.2	0.1	0.0	0.4	0.8	0.9	0.6	14.6	12.7	20.3	21	16.5	14.7

Sources: MCS (Millennium Cohort Study); ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); NS (Next Steps); BCS 70 (1970 British Cohort Study), NCDS (National Child Development Study); USoc (Understanding Society); GS (Generation Scotland: the Scottish Family Health Study); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC); BiB (Born in Bradford). Unweighted data. Note. * SD values for age are approximately zero for these cohorts. All participants in NCDS and BCS70 were born in the same week in 1958 and 1970, respectively

Supplementary Table S4. Length of time of symptoms (using self-report direct measures) by COVID-19 status (confirmed infection vs self-report).

	1					1
	Mean age	COVID-19 ascertainment	N with symptom duration data	Duration of symptoms, N (%)		
				Acute (0-4 weeks)	Ongoing symptomatic COVID-19 (4-12 weeks)	Post COVID- 19 syndrome (12+ weeks)
	19.9	Confirmed	552	532 (96.4)	15 (2.7)	5 (0.9)
MCS	19.9	Suspected	503	478 (95.0)	17 (3.4)	8 (1.6)
ALSPAC G1	28.4	Confirmed	187	144 (77.0)	25 (13.4)	18 (9.6)
ALSPAC GI	20.4	Suspected	481	375 (78.0)	15 (22.0)	34 (7.1)
	31	Confirmed	400	360 (90.0)	26 (6.5)	14 (3.5)
Next Steps	51	Suspected	448	413 (92.2)	25 (5.6)	10 (2.2)
	51	Confirmed	386	316 (81.9)	50 (13.0)	20 (5.2)
BCS70	51	Suspected	503	441 (87.7)	34 (6.8)	28 (5.6)
	52.7	Confirmed	377	261 (69.2)	66 (17.57)	50 (13.3)
TwinsUK	32.1	Suspected	429	318 (74.1)	80 (18.6)	31 (7.2)
	55.9	Confirmed	83	51 (61.4)	19 (22.9)	13 (15.7)
GS	55.9	Suspected	252	173 (67.7)	35 (13.9)	44 (17.5)
	58.3	Confirmed	95	73 (76.8)	17 (17.9)	5 (5.3)
ALSPAC G0	56.5	Suspected	351	229 (65.2)	51 (14.5)	71 (20.2)
	63	Confirmed	313	248 (79.2)	49 (15.7)	16 (5.1)
NCDS	03	Suspected	396	330 (83.3)	48 (12.1)	18 (4.6)
	40.7	Confirmed	34	22 (64.7)	8 (23.5)	4 (11.8)
BiB	40.7	Suspected	76	18 (23.7)	17 (22.4)	41 (53.9)

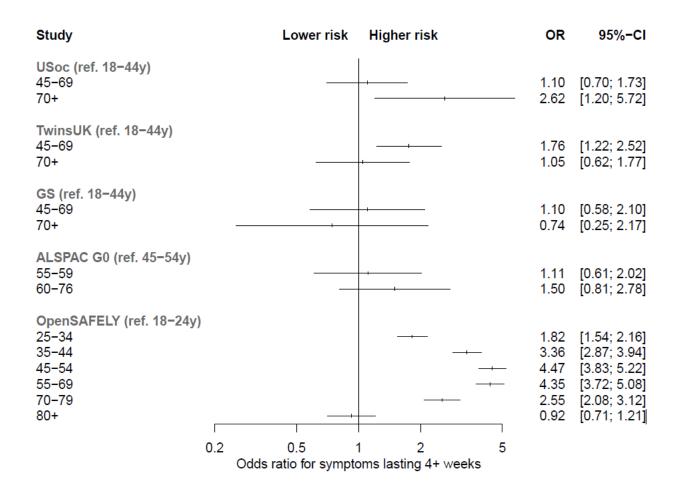
Note. *Confirmed*: tested positive by PCR, antigen and/or antibody test; Suspected: strong personal suspicion and/or medical advice of COVID-19, but no test result to confirm. Sources: MCS (Millennium Cohort Study); ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); NS (Next Steps); BCS 70 (1970 British Cohort Study), NCDS (National Child Development Study); USoc (Understanding Society); GS (Generation Scotland: The Scottish Family Health Study); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC); BiB (Born in Bradford).

		N with symptom	Duration of symptoms, N		
	COVID-19 ascertainment	duration data	(%)		
			Acute (0-4 weeks)	Ongoing symptomatic COVID-19 (4-12 weeks)	Post COVID- 19 syndrome (12 + weeks)
ALSPAC G1	Symptomatic test positive (antibody or PCR) and self- diagnosed	193 (100)	150 (78)	21 (11)	22 (11)
ALSI AC GI	Symptomatic test negative (antibody or PCR) and self- diagnosed	282 (100)	215 (78)	45 (16)	22 (8)
	Symptomatic test positive (antibody or PCR) and self- diagnosed	377(100)	261 (69.2)	66 (17.57)	50 (13.3)
TwinsUK	Symptomatic test negative (antibody or PCR) and self- diagnosed	429 (100)	318 (74.1)	80 (18.6)	31 (7.2)
	Symptomatic test positive (antibody or PCR) and self- diagnosed	57(100)	44 (77.2)	5 (8.8)	8 (14)
ALSPAC G0	Symptomatic test negative (antibody or PCR) and self- diagnosed	202 (100)	131 (64.9)	39 (19.3)	32 (15.8)
	Symptomatic test positive (antibody or PCR) and self- diagnosed	35 (100)	23 (65.7)	8 (22.9)	4 (11.4)
BiB	Symptomatic test negative (antibody or PCR) and self- diagnosed	42 (100)	6 (14.3)	9 (21.4)	27 (64.3)

Sources: ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC); BiB (Born in Bradford).

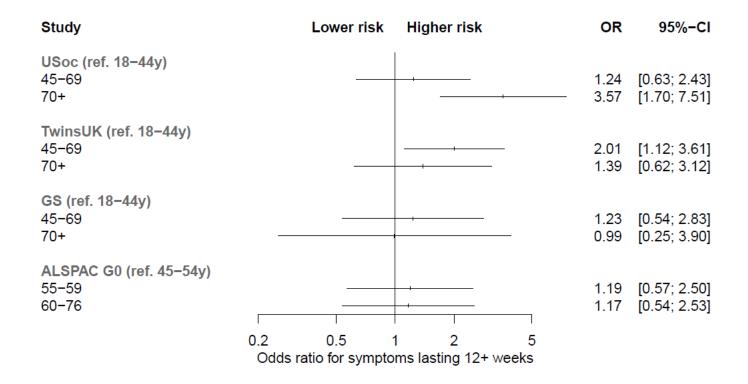
	COVID-19 ascertainment	N with symptom duration data	Duration of symptoms, N (%)		
			Acute (0-4 weeks)	Ongoing symptomatic COVID-19 (4-12 weeks)	Post COVID- 19 syndrome (12+ weeks)
	Not Confirmed	4611	820 (17.8)	1006 (21.8)	1328 (28.8)
	Confirmed	363	87 (24)	104 (28.7)	139 (38.3)
TwinsUK	Suspected	696	185 (26.6)	142 (20.4)	296 (42.5)
	Confirmed	34	22 (64.7)	8 (23.5)	4 (11.8)
BiB	Suspected	76	18 (23.7)	17 (22.4)	41 (53.9)

Supplementary figure 1: Categorical age associations with symptoms 4+ week in the sub-set of the longitudinal studies that are heterogeneous in age and EHRs from OpenSAFELY



Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate.

Sources: MCS (Millennium Cohort Study); ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); NS (Next Steps); BCS 70 (1970 British Cohort Study), NCDS (National Child Development Study); USoc (Understanding Society); GS (Generation Scotland: The Scottish Family Health Study); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). Usoc N = 1033, TwinsUK N = 806, GS N = 335, ALSPAC G0 N = 446, OpenSAFELY N = 4,189



Sources: USoc (Understanding Society); GS (Generation Scotland: the Scottish Family Health Study); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Usoc N = 1033, TwinsUK N = 806, GS N = 335, ALSPAC G0 N = 446.

	Ν	Lower risk Higher risk	OR	95%-CI
Female sex (ref. male)				
MCS	1029		5.84	[2.21; 15.48]
NS	741		1.54	[0.67; 3.53]
BCS70	886		1.92	[1.00; 3.70]
NCDS	708		2.36	[1.29; 4.33]
USoc	926	-* <u>-</u> _	0.78	[0.46; 1.33]
GS	335	-	1.56	[0.95; 2.56]
TwinsUK ALSPAC G0	803 429		1.12 1.29	[0.69; 1.82]
ALSPAC GU	668		1.29	[0.81; 2.04] [1.12; 2.49]
Common effect model	000		1.49	[1.24; 1.79]
Random effects model		\sim	1.56	[1.18; 2.07]
Heterogeneity: $I^2 = 56\%$, $\tau^2 = 0.0986$	6, p = 0.020			,
, , , , , , , , , , , , , , , , , , ,	, ,			
Other ethnicity (ref. white)				
BCS70	774 —		0.18	[0.04; 0.86]
NCDS	671		2.11	[0.56; 7.93]
MCS	1028		1.20	[0.40; 3.58]
TwinsUK	802		0.53	[0.19; 1.44]
USoc	923		0.50	[0.19; 1.31]
NS	741		0.82	[0.34; 2.01]
ALSPAC G1	668		1.24	[0.54; 2.87]
Common effect model Random effects model			0.80	[0.54; 1.19]
Heterogeneity: $I^2 = 31\%$, $\tau^2 = 0.0478$	$s_{\rm p} = 0.188$		0.80	[0.52; 1.22]
Heterogeneity: $r = 51\%$, $t = 0.04\%$	σ, <i>ρ</i> = 0.100			
No higher education (ref. degre	e attained)			
NS	740		1.18	[0.52; 2.70]
BCS70	821		1.24	[0.69; 2.22]
NCDS	699		1.22	[0.69; 2.18]
ALSPAC G0	407		1.04	[0.62; 1.74]
USoc	843		0.98	[0.60; 1.62]
ALSPAC G1	487		0.64	[0.39; 1.05]
GS TwinsUK	328 623		1.05 0.81	[0.65; 1.68]
Common effect model	023		0.01	[0.56; 1.18] [0.80; 1.14]
Random effects model			0.95	[0.80; 1.14]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0$.639	_		[
Index of multiple deprivation (p	,			
ALSPAC G0	390	+	1.00	[0.83; 1.21]
NS	739	+	1.00	[0.84; 1.19]
MCS	999	Ť	1.00	[0.86; 1.16]
ALSPAC G1	626	Ť	1.06	[0.92; 1.22]
TwinsUK	800	Ť	1.04	[0.92; 1.17]
NCDS BCS70	636 780		1.00 0.99	[0.90; 1.12] [0.90; 1.09]
Common effect model	100	\$	1.01	[0.96; 1.06]
Random effects model			1.01	[0.96; 1.06]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0$.991			[0.000, 1.000]
Current smoking (ref. non-sm	oking)			
NS	562		0.29	[0.07; 1.25]
MCS	567		1.19	[0.38; 3.77]
GS	332		1.76	[0.67; 4.59]
NCDS USoc	590 926		0.92	[0.37; 2.32] [0.47; 2.26]
BCS70	926 707		1.03 0.92	[0.47; 2.26]
ALSPAC G1	540		0.92	[0.39; 1.42]
ALSPAC GO	370		1.02	[0.63; 1.63]
Common effect model	0,0	\rightarrow	0.95	[0.73; 1.25]
Random effects model		\diamond	0.95	[0.73; 1.25]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = < 0.000$	1, <i>p</i> = 0.670		-	- / *
·				
		0.1 0.5 1 2 10		
		Odds ratio for symptoms lasting 4+ weeks		

Sources: MCS (Millennium Cohort Study); ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); NS (Next Steps); BCS 70 (1970 British Cohort Study), NCDS (National Child Development Study); USoc (Understanding Society); GS (Generation Scotland); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). Estimates from fixed and random-effect meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Associations adjusted for continuous age and sex, except where sex is the risk factor (only adjusted for age).

Supplementary figure 4: Full meta-analysis results for health factors with symptoms for 4+ weeks in the longitudinal studies

	N	Lower risk Higher risk	OR	95%-C
Poor overall health (ref. go				
NS	730		2.61	[0.69; 9.83
ALSPAC G1	526		1.23	[0.58; 2.63
TwinsUK	489		0.67	[0.33; 1.36
BCS70	780		2.01	[1.00; 4.04
NCDS	642		1.70	[0.88; 3.26
ALSPAC G0	338		2.84	[1.49; 5.40
USoc	924		1.63	
	924			[1.01; 2.65
Common effect model		\diamond		[1.25; 2.09
Random effects model		\diamond	1.61	[1.14; 2.29
Heterogeneity: $I^2 = 42\%$, $\tau^2 = 0$.	.0927, <i>p</i> = 0.113			
Psychological distress (ref	. no distress)			
MCS	974		1.20	[0.42; 3.38
NS	726		1.10	[0.45; 2.72
TwinsUK	357		1.51	[0.68; 3.36
BCS70	745		1.52	[0.69; 3.33
NCDS	644		1.16	[0.55; 2.43
GS	283		2.62	[1.30; 5.27
ALSPAC G0	343		2.07	[1.16; 3.70
USoc	925		0.95	[0.57; 1.58
ALSPAC G1	470		1.55	[0.93; 2.57
Common effect model		\diamond	1.46	[1.17; 1.83
Random effects model		\diamond	1.46	[1.14; 1.87
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0.0$	190, p = 0.455			
BMI (per 5 kg/m^2)				
NS	696		1.31	[0.89; 1.94
BCS70	710	<u> </u>	1.33	[0.94; 1.89
MCS	949	_ _	1.33	[0.94; 1.03
NCDS	607		0.95	
				[0.72; 1.25
ALSPAC G0	264		0.88	[0.67; 1.15
ALSPAC G1	465	м	1.23	[1.01; 1.50
GS	335		1.07	[0.89; 1.29
TwinsUK	803		0.88	[0.74; 1.04
Common effect model		¢	1.05	[0.96; 1.14
Random effects model		\$	1.07	[0.94; 1.20
Heterogeneity: $I^2 = 48\%$, $\tau^2 = 0$.0138, <i>p</i> = 0.064			
Overweight or obese (ref. E	MI-25)			
Overweight of obese (ref. E NS	696		2.19	[0.91; 5.29
	949			
MCS			0.88	[0.38; 2.02
BCS70	710		1.53	[0.67; 3.48
NCDS	607		0.85	[0.43; 1.71
ALSPAC G0	263		0.89	[0.51; 1.56
GS	277		1.19	[0.69; 2.03
TwinsUK	372		1.25	[0.79; 1.98
ALSPAC G1	453		1.72	[1.10; 2.68
Common effect model		\diamond	1.25	[1.02; 1.55
Random effects model		\diamond	1.25	[1.01; 1.55
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0.0$	i030, p = 0.428			
Diabetes				
ALSPAC G0	338		1.52	[0.24; 9.62
GS	283			[0.79; 20.34
BCS70	780			[1.23; 21.72
TwinsUK	536		0.62	
		_		[0.16; 2.39
NCDS	641		0.39	[0.14; 1.07
USoc	926		2.08	[0.95; 4.58
Common effect model			1.39	[0.86; 2.26
Random effects model	6640 0.020		1.48	[0.64; 3.45
Heterogeneity: $I^2 = 63\%$, $\tau^2 = 0$.	0040, p = 0.020			
Hypertension				
GS USoc	283 926		0.93 1.18	[0.36; 2.40 [0.60; 2.30
NCDS	642		1.10	[0.78; 2.87
ALSPAC G0	345		1.49	
		L.		[0.59; 1.83
TwinsUK	589		1.21	[0.75; 1.98
Common effect model				[0.90; 1.56
Random effects model Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, j	n = 0.916	~	1.18	[0.90; 1.56
$for (0, \tau) = 0.00, \tau = 0, \tau$	2 - 0.010			
High cholesterol				
ALSPAC G0	254	_	2.66	[1.18; 5.96
TwinsUK	590	- <u>m</u> -	1.06	[0.66; 1.68
Common effect model				[0.89; 1.99
Random effects model			1.58	[0.64; 3.86
Heterogeneity: $I^2 = 73\%$, $\tau^2 = 0.00$.3112, <i>p</i> = 0.053			
Asthma				
MCS	882		2.98	[0.58; 15.32
TwinsUK	114		1.60	[0.71; 3.62
NCDS	641		1.80	[0.89; 3.66
GS	283		1.22	[0.62; 2.38
BCS70	780		1.46	[0.76; 2.82
ALSPAC G0	335		1.35	[0.72; 2.54
ALSPAC G1	450		1.39	[0.86; 2.25
USoc	926	-+	1.01	[0.68; 1.50
Common effect model		\diamond	1.31	[1.06; 1.62
Random effects model		×		[1.06; 1.62
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$,	n = 0.793		1.31	1.00, 1.02
neterogeneity. $r = 0\%$, $\tau = 0$,	y = 0.185			
	0	1 0.5 1 2 10		
	0. C	1 0.5 1 2 10 Odds ratio for symptoms lasting 4+ weeks		

Sources: MCS (Millennium Cohort Study); ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); NS (Next Steps); BCS 70 (1970 British Cohort Study), NCDS (National Child Development Study); USoc (Understanding Society); GS (Generation Scotland); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). The reference category for 'Diabetes', 'Hypertension', 'High Cholesterol', and 'Asthma' is the absence of condition. Estimates from fixed and random-effect meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Associations adjusted for age and sex.

	Ν	Lower risk Higher risk	OR	95%-CI
Female sex (ref. male)		I		
MCS	1029		1.55	[0.35; 6.83]
NS	741		1.37	[0.32; 5.83]
BCS70	886		1.67	[0.56; 4.98]
TwinsUK	803		1.75	[0.73; 4.20]
NCDS	708		3.32	[1.39; 7.94]
ALSPAC G1	668		2.24	[1.13; 4.44]
GS	335		1.73	[0.91; 3.30]
USoc	926	— <u> </u>	0.89	[0.48; 1.63]
ALSPAC G0	429		1.44	[0.80; 2.61]
Common effect model		\diamond	1.60	[1.23; 2.09]
Random effects model		\diamond	1.62	[1.21; 2.17]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0.0$	0284, p = 0.488			
Other ethnicity (ref. white)				
NCDS	671	+	0.32	[0.04; 2.65]
MCS	1028		4.07	[0.77; 21.42]
TwinsUK	802		0.65	[0.14; 3.04]
NS	741 —		0.08	[0.02; 0.33]
ALSPAC G1	668		1.28	[0.37; 4.39]
USoc	923		0.25	[0.16; 0.38]
Common effect model		\diamond	0.32	[0.22; 0.47]
Random effects model			0.50	[0.17; 1.48]
Heterogeneity: $I^2 = 75\%$, $\tau^2 = 1$.2896, <i>p</i> = 0.001			
No bishes advastice (set				
No higher education (ref. d	•		0.60	10 15: 0 461
NS	740		0.60	[0.15; 2.46]
BCS70	821		0.98	[0.35; 2.79]
NCDS	699		0.50	[0.20; 1.28]
ALSPAC G1	487		0.42	[0.18; 0.96]
ALSPAC G0	407		0.68	[0.36; 1.27]
GS	328		0.67	[0.37; 1.21]
USoc	843		1.52	[0.85; 2.73]
TwinsUK	623		0.59	[0.34; 1.03]
Common effect model		\diamond	0.73	[0.57; 0.94]
Random effects model		\diamond	0.72	[0.52; 0.99]
Heterogeneity: $I^2 = 27\%$, $\tau^2 = 0$	0.0702, p = 0.215			
Index of multiple deprivation	on (per one decile)			
NS	739		1.29	[0.94; 1.78]
ALSPAC G1	626			[0.63; 1.04]
ALSPAC G0	390	<u> </u>	1.06	[0.85; 1.33]
BCS70	780	+	1.02	[0.84; 1.25]
TwinsUK	800	+	1.03	[0.86; 1.24]
NCDS	636		1.16	[0.98; 1.38]
MCS	999	+	0.94	[0.81; 1.09]
Common effect model		\$	1.02	[0.95; 1.10]
Random effects model		\$		[0.94; 1.12]
Heterogeneity: $I^2 = 33\%$, $\tau^2 = 0$	0.0040, p = 0.174			
Current smoking (ref. non-			0.00	[4 0 4, 7 4 70]
MCS	567	*	- 8.80	[1.04; 74.79]
NS	562	*	0.64	[0.08; 5.10]
GS	332		0.91	[0.24; 3.45]
NCDS	590	*	1.89	[0.54; 6.69]
BCS70	707		0.75	[0.22; 2.63]
ALSPAC G1	540		1.05	[0.42; 2.58]
USoc	926		1.44	[0.60; 3.42]
ALSPAC G0	370		0.82	[0.44; 1.51]
Common effect model		\frown	1.08	[0.75; 1.55]
Random effects model		\Diamond	1.08	[0.75; 1.55]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = < 0$	0.0001, <i>p</i> = 0.506			
	<u>_</u>	0.1 0.5 1 2 10 odds ratio for symptoms lasting 12+ weeks		
	C			

Sources: MCS (Millennium Cohort Study); ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); NS (Next Steps); BCS 70 (1970 British Cohort Study), NCDS (National Child Development Study); USoc (Understanding Society); GS (Generation Scotland); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Associations adjusted for age and sex, where relevant.

	N	Lower risk Higher risk	OR	95%-CI
Poor overall health (ref. goo	od health)	1		
NS	730		0.14	[0.01; 1.41]
BCS70	780	.	1.40	[0.36; 5.45]
ALSPAC G1	526		0.87	[0.26; 2.97]
TwinsUK	489		0.60	
				[0.20; 1.79]
NCDS	642		2.44	[0.99; 6.02]
USoc	924		1.58	[0.76; 3.31]
ALSPAC G0	338		3.43	[1.70; 6.95]
Common effect model		\diamond	1.66	[1.15; 2.40]
Random effects model				[0.78; 2.51]
Heterogeneity: $I^2 = 56\%$, $\tau^2 = 0$.	.3133, p = 0.034	—		[0110] 2101]
Developing the first sector of the				
Psychological distress (ref.			4.00	
BCS70	745		1.93	[0.42; 8.96]
MCS	974		5.67	[1.28; 25.11]
NS	726		1.01	[0.23; 4.35]
TwinsUK	357		0.78	[0.21; 2.84]
NCDS	644		2.39	[0.75; 7.56]
ALSPAC G1	470		0.89	
				[0.38; 2.12]
GS	283		3.77	[1.75; 8.14]
ALSPAC G0	343		1.27	[0.62; 2.60]
USoc	925		1.20	[0.62; 2.32]
Common effect model		\diamond		[1.15; 2.17]
Random effects model		Š		
	4469	~	1.61	[1.05; 2.46]
Heterogeneity: $I^2 = 36\%$, $\tau^2 = 0$.	1468, $p = 0.129$			
BMI (per 5 kg/m^2)				
MCS	949		0.21	[0.07; 0.62]
BCS70	710		1.45	[0.67; 3.14]
NCDS	607		0.87	[0.55; 1.39]
NS	696	+	1.02	[0.69; 1.51]
ALSPAC G0	264	-+	0.67	[0.47; 0.97]
ALSPAC G1	465	- <u>h</u> -	1.11	[0.84; 1.46]
TwinsUK	803		0.87	[0.68; 1.13]
GS	335	<u> </u>	1.16	[0.94; 1.44]
Common effect model		Ŷ		[0.87; 1.10]
Random effects model		4	0.94	[0.78; 1.12]
Heterogeneity: $I^2 = 59\%$, $\tau^2 = 0$.	.0275, p = 0.016			-
Overweight or obese (ref. B	MI-25)			
NS	696		2.21	[0.38; 12.79]
MCS	949		0.08	
BCS70				[0.02; 0.41]
	710		0.88	[0.22; 3.53]
NCDS	607		0.58	[0.24; 1.39]
TwinsUK	372		1.29	[0.63; 2.62]
GS	277		1.28	[0.64; 2.55]
ALSPAC G0	263		0.62	[0.31; 1.22]
ALSPAC G1	453		1.50	[0.78; 2.89]
Common effect model		\$		[0.70; 1.28]
Random effects model Heterogeneity: $I^2 = 56\%$, $\tau^2 = 0$.	2450 - 0.020		0.88	[0.55; 1.40]
neterogeneity. $I = 56\%$, $\tau = 0$.	2159, p = 0.026			
Diabetes				
ALSPAC G0	338		1.68	[0.17; 16.15]
NCDS	641			[0.04; 2.40]
GS	283			[0.49; 18.98]
TwinsUK	536		1.28	[0.25; 6.52]
BCS70	780		1.08	[0.24; 4.77]
USoc	926		1.25	[0.44; 3.57]
Common effect model			1.22	[0.64; 2.30]
Random effects model				[0.64; 2.30]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, μ	p = 0.728			
Hypertension	a			
GS	283		2.02	[0.72; 5.68]
NCDS	642		1.11	[0.45; 2.74]
USoc	926		0.77	[0.33; 1.80]
ALSPAC G0	345		0.77	[0.36; 1.61]
		- L_		
TwinsUK	589		1.23	[0.61; 2.47]
Common effect model		~		[0.73; 1.52]
Random effects model			1.05	[0.73; 1.52]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, μ	p = 0.565			
High cholesterol				
	241		3.67	[1 00: e Fe1
ALSPAC GO	241		2.67	[1.09; 6.56]
TwinsUK	590		0.97	[0.49; 1.92]
Common effect model				[0.81; 2.42]
Random effects model			1.54	[0.57; 4.14]
Heterogeneity: $I^2 = 68\%$, $\tau^2 = 0$.	.3492, <i>p</i> = 0.078			-
Aothma				
Asthma MCS	882		0.57	[0.09; 3.79]
	641			
NCDS			1.97	[0.67; 5.73]
NCDS	780		1.53	[0.56; 4.18]
BCS70			1.30	[0.48; 3.53]
BCS70	114		1.21	[0.51; 2.86]
BCS70 TwinsUK	114 283			
BCS70 TwinsUK GS	283			
BCS70 TwinsUK GS ALSPAC G0	283 335		0.92	[0.40; 2.10]
BCS70 TwinsUK GS ALSPAC G0 ALSPAC G1	283 335 450		0.92 1.00	[0.40; 2.10] [0.47; 2.13]
BCS70 TwinsUK GS ALSPAC G0 ALSPAC G1	283 335		0.92 1.00 1.01	[0.40; 2.10] [0.47; 2.13] [0.49; 2.07]
BCS70 TwinsUK GS ALSPAC G0 ALSPAC G1 USoc	283 335 450		0.92 1.00 1.01	[0.40; 2.10] [0.47; 2.13] [0.49; 2.07]
BCS70 TwinsUK GS ALSPAC G0 ALSPAC G1 USoc Common effect model	283 335 450		0.92 1.00 1.01 1.14	[0.40; 2.10] [0.47; 2.13] [0.49; 2.07] [0.83; 1.57]
BCS70 TwinsUK GS ALSPAC G0 ALSPAC G1 USoc Common effect model Random effects model	283 335 450 926		0.92 1.00 1.01 1.14	[0.40; 2.10] [0.47; 2.13] [0.49; 2.07]
BCS70 TwinsUK GS ALSPAC G0 ALSPAC G1 USoc Common effect model	283 335 450 926		0.92 1.00 1.01 1.14	[0.40; 2.10] [0.47; 2.13] [0.49; 2.07] [0.83; 1.57]
BCS70 TwinsUK GS ALSPAC G0 ALSPAC G1 USoc Common effect model Random effects model	283 335 450 926		0.92 1.00 1.01 1.14	[0.40; 2.10] [0.47; 2.13] [0.49; 2.07] [0.83; 1.57]
BCS70 TwinsUK GS ALSPAC G0 ALSPAC G1 USoc Common effect model Random effects model	283 335 450 926 p = 0.933		0.92 1.00 1.01 1.14	[0.40; 2.10] [0.47; 2.13] [0.49; 2.07] [0.83; 1.57]

Sources: MCS (Millennium Cohort Study); ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); NS (Next Steps); BCS 70 (1970 British Cohort Study), NCDS (National Child Development Study); USoc (Understanding Society); GS (Generation Scotland); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). The reference category for 'Diabetes', 'Hypertension', 'High Cholesterol', and 'Asthma' is the absence of condition. Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Associations adjusted for age and sex.

Supplementary figure 7: Secondary meta-analysis results for sociodemographic characteristics with symptoms for 4+ weeks in the longitudinal studies, including inverse probability weights for COVID-19 risk

	Ν	Lower risk Higher risk	OR	95%-CI
Female sex (ref. male)				
MCS	1028		5.02	[1.89; 13.33]
NS	741		1.55	[0.68; 3.55]
BCS70	886		1.85	[0.97; 3.53]
GS	335		1.72	[0.93; 3.19]
NCDS	708		2.13	[1.20; 3.77]
TwinsUK	803		1.33	[0.79; 2.25]
ALSPAC G0	429		1.29	[0.80; 2.06]
USoc	926			[0.60; 1.45]
ALSPAC G1	668			[1.11; 2.47]
Common effect model Random effects model		\$		[1.26; 1.82] [1.24; 1.96]
Heterogeneity: $I^2 = 39\%$, $\tau^2 = 0$	0.0381, <i>p</i> = 0.108			
Other ethnicity (ref. white)			0.40	10.04.0.051
BCS70 NCDS	774 — 671		0.18	[0.04; 0.85]
MCS	1027		2.18 1.08	[0.57; 8.28]
USoc	923		0.61	[0.37; 3.19] [0.22; 1.68]
TwinsUK	802		0.76	[0.28; 2.08]
NS	741		0.81	[0.33; 2.01]
ALSPAC G1	668			[0.54; 2.89]
Common effect model	000			[0.59; 1.29]
Random effects model			0.87	[0.59; 1.29]
Heterogeneity: $I^2 = 17\%$, $\tau^2 = -17\%$	< 0.0001, p = 0.304			. , .
No higher education (ref. on NS	degree attained) 740		1.12	[0.50; 2.50]
BCS70	821		1.20	[0.67; 2.16]
NCDS	699		1.18	[0.67; 2.08]
GS	328		1.01	[0.58; 1.75]
ALSPAC G0	407		1.02	[0.60; 1.71]
ALSPAC G1	487		0.64	[0.39; 1.05]
USoc	843		0.93	[0.57; 1.52]
TwinsUK	623		0.77	[0.53; 1.13]
Common effect model		<	0.92	[0.76; 1.10]
Random effects model Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$,	p = 0.677		0.92	[0.76; 1.10]
Index of multiple deprivati	on (per one decile)			
ALSPAC G0	390		1.01	[0.84; 1.23]
NS	739	+	1.01	[0.86; 1.19]
ALSPAC G1	626	7	1.06	[0.91; 1.23]
MCS	998	±	1.01	[0.88; 1.16]
TwinsUK	800	土	1.04	[0.91; 1.19]
BCS70 NCDS	780 636		0.98	[0.88; 1.09]
Common effect model	030	7		[0.93; 1.14] [0.97; 1.07]
Random effects model		Ľ		[0.97; 1.07]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$,	<i>p</i> = 0.991		1.02	[0.07, 1.07]
Current smoking (ref. non				
NS	562		0.30	[0.07; 1.33]
GS	332		1.34	[0.38; 4.73]
MCS	566		1.08	[0.34; 3.42]
NCDS RCS70	590		0.85	[0.34; 2.15]
BCS70 USoc	707 926		0.97 0.95	[0.46; 2.04] [0.48; 1.88]
ALSPAC G1	920 540			[0.48, 1.88]
ALSPAC G0	370	<u> </u>	1.05	[0.65; 1.69]
Common effect model	010		0.91	[0.70; 1.20]
Random effects model		7	0.91	[0.70; 1.20]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$,	p = 0.845]	0.01	True a second
	-			
		0.1 0.5 1 2 10		
		Odds ratio for symptoms lasting 4+ weeks		

Sources: MCS (Millennium Cohort Study); ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); NS (Next Steps); BCS 70 (1970 British Cohort Study), NCDS (National Child Development Study); USoc (Understanding Society); GS (Generation Scotland); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Associations adjusted for age and sex.

Supplementary figure 8: Secondary meta-analysis results for health factors with symptoms for 4+ weeks in the longitudinal studies, including inverse probability weights for COVID-19 risk

	N	Lower risk	Higher risk	OR	95%-CI
Poor overall health (ref. good h	nealth)				
NS	730		+ +	1.93	[0.53; 6.97]
ALSPAC G1	526			1.20	[0.57; 2.52]
BCS70	780			2.10	[1.03; 4.31]
TwinsUK	489			0.86	[0.44; 1.66]
ALSPAC G0	338			3.02	[1.57; 5.81]
NCDS	642			1.70	[0.91; 3.17]
USoc	924			1.49	[0.93; 2.39]
Common effect model			\diamond	1.60	[1.25; 2.05]
Random effects model			\diamond	1.61	[1.18; 2.20]
Heterogeneity: $I^2 = 28\%$, $\tau^2 = 0.054$	0, <i>p</i> = 0.212				
Psychological distress (ref. no	distress)				
MCS	973			1.34	[0.48; 3.79]
GS	283			2.28	[0.89; 5.85]
TwinsUK	357			1.34	[0.55; 3.27]
NS	726			1.11	[0.46; 2.70]
BCS70	745	_		1.59	[0.68; 3.73]
NCDS	644		x	1.12	[0.55; 2.28]
ALSPAC G0	343			2.06	[1.19; 3.57]
USoc	926		<u>+</u>	1.01	[0.60; 1.68]
ALSPAC G1	470			1.54	[0.93; 2.54]
Common effect model			\diamond		[1.13; 1.79]
Random effects model			à		[1.13; 1.79]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0$	1 720		_	1.42	[1.15, 1.75]
Herefogeneity. $T = 0.06$, $\tau = 0$, $p = 0.05$	5.720				
BMI (per 5 kg/m^2)					
	603			1.00	10 90: 1 971
NS	693			1.29	[0.89; 1.87]
BCS70	710			1.30	[0.92; 1.84]
MCS	948		+	1.30	[0.94; 1.78]
ALSPAC G0	264	-		0.89	[0.68; 1.17]
NCDS	607	-	+	0.99	[0.76; 1.29]
GS	335		-	1.12	[0.88; 1.43]
ALSPAC G1	465		-	1.23	[1.01; 1.49]
TwinsUK	803		÷	0.97	[0.83; 1.12]
Common effect model			¢	1.08	[0.99; 1.17]
Random effects model			¢	1.09	[0.98; 1.21]
Heterogeneity: $I^2 = 26\%$, $\tau^2 = 0.006$	3, p = 0.225				
Overweight or obese (ref. BMI-	<25)				
NS	696	-		1.93	[0.78; 4.77]
BCS70	710	_		1.45	[0.62; 3.37]
MCS	948			1.14	[0.50; 2.58]
NCDS	607			0.97	[0.49; 1.90]
GS	277	_		1.16	[0.62; 2.16]
ALSPAC G0	264			0.91	[0.51; 1.63]
TwinsUK	372			1.68	[1.01; 2.78]
ALSPAC G1	465			1.63	[1.05; 2.54]
Common effect model	400			1.34	[1.08; 1.67]
Random effects model			ě	1.34	[1.08; 1.67]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0$	1645		–	1.54	[1.00, 1.07]
Heterogeneity. $T = 0\%$, $\tau = 0$, $p = 0$	0.040				
Diabetes					
GS	283			4.21	[0.35; 51.04]
ALSPAC G0	338				
				1.85	
BCS70	780	_			[1.52; 24.91]
TwinsUK	536			0.69	[0.25; 1.91]
NCDS	641		1_	0.38	[0.15; 1.01]
USoc	926			1.54	[0.67; 3.51]
Common effect model		<	\sim	1.14	[0.71; 1.83]
Random effects model		~		1.37	[0.58; 3.22]
Heterogeneity: $I^2 = 62\%, \tau^2 = 0.684$	4, <i>p</i> = 0.021				
Hypertension					
GS	283			0.98	[0.31; 3.06]
NCDS	642	-	1	1.42	[0.76; 2.68]
ALSPAC G0	345		-	1.09	[0.60; 1.96]
USoc	926	_	H	1.19	[0.69; 2.06]
TwinsUK	589			1.49	[0.95; 2.33]
Common effect model			\diamond	1.29	[0.99; 1.68]
Random effects model			\diamond	1.29	[0.99; 1.68]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0$	0.892				• · ·
High cholesterol					
ALSPAC G0	254			2.71	[1.20; 6.10]
TwinsUK	590		-	1.33	[0.86; 2.06]
Common effect model			\sim		[1.06; 2.29]
Random effects model			\sim	1.74	[0.89; 3.41]
Heterogeneity: $I^2 = 56\%$, $\tau^2 = 0.139$	5. p = 0.133				Press, entry
	o, p 0.100				
Asthma					
MCS	881	_	+ +	3.77	[0.72; 19.66]
TwinsUK	114		L	2.63	[1.03; 6.71]
GS	283	_		1.43	[0.63; 3.22]
NCDS	641		L	1.43	
			· ·		[0.86; 3.60]
BCS70	780	-		1.48	[0.76; 2.87]
ALSPAC G0	335	_		1.29	[0.67; 2.51]
ALSPAC G1	450		1."	1.38	[0.85; 2.24]
USoc	926	-		1.03	[0.69; 1.55]
Common effect model				1.37	[1.10; 1.72]
Random effects model			\diamond	1.37	[1.10; 1.72]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0$	0.568		ļ		
			1 I I		
		0.1 0.5			
		Odds ratio for sympt	oms lasting 4+ weeks	6	

Sources: MCS (Millennium Cohort Study); ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); NS (Next Steps); BCS 70 (1970 British Cohort Study), NCDS (National Child Development Study); USoc (Understanding Society); GS (Generation Scotland); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). The reference category for 'Diabetes', 'Hypertension', 'High Cholesterol', and 'Asthma' is the absence of condition. Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Associations adjusted for age and sex.

Supplementary figure 9: Secondary meta-analysis results for sociodemographic characteristics with symptoms for 12+ weeks in the longitudinal studies, including inverse probability weights for COVID-19 risk

	Ν	Lower risk Higher risk	OR	95%-0
Female sex (ref. male)		1		
NS	741		1.32	[0.30; 5.8
MCS	1028		1.48	0.35; 6.2
BCS70	886		1.69	[0.58; 4.9
NCDS	708		2.95	[1.26; 6.9
GS	335		1.53	[0.69; 3.3
TwinsUK	803		1.43	[0.68; 3.0
ALSPAC G1	668		2.25	[1.13; 4.4
USoc	926			[0.54; 1.9
ALSPAC G0	429	Ĭ.		[0.76; 2.5
	429			
Common effect model				[1.20; 2.0
Random effects model Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, μ	- 0 722	♦	1.50	[1.20; 2.0
$\frac{1}{1000} = \frac{1}{1000} = 1$	7 - 0.725			
Other ethnicity (ref. white)				
NCDS	671 -	•	0.32	[0.04; 2.6
MCS	1027	+		[0.78; 19.6
NS	741 —		0.08	[0.02; 0.3
TwinsUK	802	*	1.17	[0.29; 4.6
ALSPAC G1	668			[0.39; 4.4
USoc	923			0.22; 0.5
Common effect model		\diamond		[0.29; 0.6
Random effects model			0.60	[0.21; 1.7
Heterogeneity: $I^2 = 74\%$, $\tau^2 = 1.5$	2614, <i>p</i> = 0.002			·,
No higher education (ref. de	aree attained)			
NS	740		0.71	[0.17; 2.9
BCS70	821		0.94	
NCDS				[0.33; 2.6
	699		0.52	[0.21; 1.2
ALSPAC G1	487		0.42	[0.18; 0.9
GS	328		0.59	[0.29; 1.2
USoc	843			[0.58; 2.1
ALSPAC G0	407		0.71	[0.37; 1.3
TwinsUK	623		0.45	[0.24; 0.8
Common effect model		\diamond	0.64	[0.48; 0.8
Random effects model		\diamond	0.64	[0.48; 0.8
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0.0$	113, <i>p</i> = 0.526			
Index of multiple deprivatio	n (per one decile)			
NS	739	+	1.26	[0.92; 1.7
ALSPAC G1	626		0.81	0.63; 1.0
ALSPAC G0	390		1.11	0.88; 1.4
BCS70	780	_ _		[0.79; 1.2
TwinsUK	800	<u>_</u>		[0.86; 1.2
NCDS	636		1.17	•
MCS	998			[0.86; 1.1
Common effect model	330	Ţ		
Random effects model		Ľ		[0.96; 1.1 [0.95; 1.1
Heterogeneity: $I^2 = 26\%$, $\tau^2 = 0.1$	0024, p = 0.234	Ť	1.04	[0.95, 1.1
Current smoking (ref. non-			0.70	10 40 - 0.0
NS	562		0.78	[0.10; 6.3
MCS	566	1 *		[0.93; 56.3
GS	332			[0.22; 5.9
NCDS	590			[0.49; 6.2
BCS70	707			[0.23; 2.7
ALSPAC G1	540		1.03	[0.42; 2.5
USoc	926			0.52; 2.6
ALSPAC G0	370			[0.47; 1.6
Common effect model		\diamond		[0.75; 1.5
Random effects model		↓		[0.75; 1.5
Heterogeneity: $I^2 = 0\%$, $\tau^2 = < 0$.0001, p = 0.702			
	-	0.1 0.5 1 2 10		
	0	dds ratio for symptoms lasting 12+ weeks		

Sources: MCS (Millennium Cohort Study); ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); NS (Next Steps); BCS 70 (1970 British Cohort Study), NCDS (National Child Development Study); USoc (Understanding Society); GS (Generation Scotland); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Associations adjusted for age and sex.

Supplementary figure 10: Secondary meta-analysis results for health traits with symptoms for 12+ weeks in the longitudinal studies, including inverse probability weights for COVID-19 risk

	Ν	Lower risk Higher risk	OR	95%-CI
Poor overall health (ref. god	od health)			
NS	730 —	• • • • • • • • • • • • • • • • • • • •	0.12	[0.01; 1.15]
BCS70	780		1.66	[0.39; 7.08]
ALSPAC G1	526		0.86	[0.26; 2.91]
TwinsUK	489		0.96	[0.37; 2.48]
NCDS USoc	642 924		2.58 1.16	[1.04; 6.41] [0.55; 2.41]
ALSPAC G0	338	· · · · · · · · · · · · · · · · · · ·	3.92	[1.91; 8.06]
Common effect model			1.65	[1.14; 2.38]
Random effects model		\sim	1.44	[0.79; 2.60]
Heterogeneity: $I^2 = 60\%$, $\tau^2 = 0$.	3342, p = 0.021			
Psychological distress (ref.			0.00	10 40 40 401
BCS70 NS	745 726		2.30 0.99	[0.43; 12.48] [0.20; 4.83]
MCS	973			[1.51; 23.14]
NCDS	644		2.38	[0.76; 7.45]
GS	283	n	3.33	[1.18; 9.37]
ALSPAC G1	470		0.88	[0.37; 2.06]
ALSPAC G0	343		1.49	[0.76; 2.95]
USoc	926		1.48	[0.77; 2.84]
Common effect model		\diamond	1.69	[1.21; 2.36]
Random effects model	0.1.17 0.000	\diamond	1.70	[1.20; 2.42]
Heterogeneity: $I^2 = 16\%$, $\tau^2 = 0$.	0147, p = 0.302			
BMI (per 5 kg/m^2)				
MCS	948		0.26	[0.10; 0.74]
BCS70	710		1.31	[0.58; 2.97]
NCDS	607		0.88	[0.56; 1.40]
NS	693		0.95	[0.60; 1.48]
ALSPAC G0	264		0.67	[0.47; 0.95]
GS	335	-	1.17	[0.88; 1.55]
ALSPAC G1	465	1	1.11	[0.86; 1.43]
TwinsUK Common offect model	803	Ť	0.97	[0.78; 1.22] [0.86; 1.10]
Common effect model Random effects model		Ĭ.	0.97 0.95	[0.80; 1.10]
Heterogeneity: $I^2 = 50\%$, $\tau^2 = 0$.	$0.194 \ p = 0.052$	Ť	0.95	[0.00, 1.12]
	6101, p 6.002			
Overweight or obese (ref. E	SMI<25)			
NS	696		1.78	
MCS	948 —		0.11	[0.02; 0.53]
BCS70	710		0.78	[0.19; 3.20]
NCDS GS	607 277		0.64 1.24	[0.26; 1.58] [0.55; 2.78]
TwinsUK	372		1.24	[0.62; 3.05]
ALSPAC G0	264		0.56	[0.28; 1.14]
ALSPAC G1	465		1.47	[0.77; 2.82]
Common effect model		<	0.92	[0.67; 1.27]
Random effects model		\diamond	0.87	[0.56; 1.36]
Heterogeneity: $I^2 = 48\%$, $\tau^2 = 0$.	1645, <i>p</i> = 0.060			
Diabetes				
GS	283		3.02	[0.18; 51.79]
ALSPAC G0	338			[0.16; 15.26]
NCDS	641		0.31	[0.04; 2.49]
BCS70	780		1.20	[0.28; 5.17]
TwinsUK	536		1.03	[0.26; 4.05]
USoc	926		1.05	[0.46; 2.38]
Common effect model		\rightarrow	1.04	[0.59; 1.85]
Random effects model Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, I	- 0.950		1.04	[0.59; 1.85]
Heterogeneity. $T = 0\%$, $\tau = 0$, p	0 = 0.052			
Hypertension				
GS	283		2.11	[0.61; 7.26]
NCDS	642		1.07	[0.43; 2.65]
ALSPAC G0	345		0.94	[0.42; 2.11]
USoc	926		0.66	[0.30; 1.47]
TwinsUK	589		1.62	[0.86; 3.08]
Common effect model			1.15	[0.80; 1.66]
Random effects model Heterogeneity: $I^2 = 3\%$, $\tau^2 = 0.0$	198 n = 0.387	\leftarrow	1.14	[0.77; 1.69]
Therefore the state of the sta	150, p = 0.307			
High cholesterol				
ALSPAC GO	241		2.66	[1.07; 6.60]
TwinsUK	590	+	1.53	[0.81; 2.87]
Common effect model		\sim	1.83	[1.09; 3.07]
Random effects model	- 0.220	\diamond	1.83	[1.09; 3.07]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, I	9 = 0.328			
Asthma				
MCS	881		0.71	[0.11; 4.39]
TwinsUK	114		2.20	[0.72; 6.76]
NCDS	641		1.88	[0.63; 5.55]
BCS70	780		1.47	[0.54; 4.00]
GS	283		1.75	[0.65; 4.71]
ALSPAC G0	335		1.01	[0.43; 2.38]
ALSPAC G1	450	— <u>+</u>	0.98	[0.46; 2.09]
USoc	926		0.97	[0.49; 1.93]
Common effect model Random effects model			1.23	[0.89; 1.72]
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, I	n = 0.826	\sim	1.23	[0.89; 1.72]
10:0:09010ity. r = 070, t = 0, j	- 0.020			
		0.1 0.5 1 2 10		
	0	dds ratio for symptoms lasting 12+ weeks		

Odds ratio for symptoms lasting 12+ weeks

Sources: MCS (Millennium Cohort Study); ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); NS (Next Steps); BCS 70 (1970 British Cohort Study), NCDS (National Child Development Study); USoc (Understanding Society); GS (Generation Scotland); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). The reference category for 'Diabetes', 'Hypertension', 'High Cholesterol', and 'Asthma' is the absence of condition. Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Associations adjusted for age and sex.

Supplementary figure 11: Sub-group meta-analysis results for sociodemographic characteristics with symptoms for 4+ weeks in individuals with COVID-19 status confirmed by a positive PCR test and/or serology result

	Ν	Lower risk	Higher risk	OR	95%-CI
Female sex (ref. male) ALSPAC G0 ALSPAC G1 TwinsUK Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, $p = 10\%$	57 193 377 0.791			2.12 1.82 1.84	[0.18; 6.27] [0.88; 5.14] [0.86; 3.86] [1.07; 3.16] [1.07; 3.16]
Other ethnicity (ref. white) ALSPAC G1 TwinsUK Common effect model Random effects model Heterogeneity: $I^2 = 2\%$, $\tau^2 = 0.0110$	193 377 , <i>p</i> = 0.312			0.47 0.75	[0.32; 5.02] [0.12; 1.77] [0.29; 1.96] [0.29; 1.98]
No higher education (ref. degr ALSPAC G0 ALSPAC G1 TwinsUK Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, $p = 0$	55 150 297		 	0.41 0.57 0.50	[0.03; 1.44] [0.15; 1.16] [0.33; 0.99] [0.31; 0.80] [0.31; 0.80]
Index of multiple deprivation (ALSPAC G0 ALSPAC G1 TwinsUK Common effect model Random effects model Heterogeneity: $l^2 = 1\%$, $\tau^2 = < 0.00$	53 183 376		 	1.01 0.96 1.00	[0.85; 2.43] [0.74; 1.36] [0.80; 1.15] [0.86; 1.16] [0.86; 1.16]
Current smoking (ref. non–sm ALSPAC G0 ALSPAC G1 Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, $p = 1$	51 168 0.331	0.1 0.5 Odds ratio for sympto	1 2 10 2 10	0.53 0.80	[0.31; 6.83] [0.15; 1.94] [0.30; 2.17] [0.30; 2.17]

Sources: ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Adjusted for age, sex and ethnicity, where relevant. Some samples were omitted from specific risk factor long COVID analyses, where sample sizes were too small to contribute, e.g., ALSPAC G0 in the ethnicity meta-analysis.

Supplementary figure 12: Sub-group meta-analysis results for health traits with symptoms for 4+ weeks in individuals with COVID-19 status confirmed by a positive PCR test and/or serology result

	N	Lower risk Higher risk	OR	95%-CI
Poor overall health (ref. g ALSPAC G1 ALSPAC G0 TwinsUK Common effect model Random effects model Heterogeneity: 1 ² = 51%, τ ² =	159 - 50 226		0.57 0.93	[0.04; 3.13] [0.72; 31.26] [0.14; 2.26] [0.35; 2.48] [0.23; 4.23]
Psychological distress (r ALSPAC G0 TwinsUK ALSPAC G1 Common effect model Random effects model Heterogeneity: <i>I</i> ² = 0%, τ ² = 0	51 165 143		2.39 1.35 1.63	[0.09; 13.06] [0.70; 8.14] [0.50; 3.62] [0.78; 3.39] [0.78; 3.39]
BMI (per 5 kg/m^2) ALSPAC G0 ALSPAC G1 TwinsUK Common effect model Random effects model Heterogeneity: I ² = 0%, τ ² = 0	27 146 175 0, <i>p</i> = 0.588		0.90 0.93	[0.70; 1.63]
Overweight or obese (ref. ALSPAC G0 ALSPAC G1 TwinsUK Common effect model Random effects model Heterogenelty: <i>I</i> ² = 0%, τ ² = 0	27 144 175		1.11 1.13 1.10	[0.11; 6.24] [0.51; 2.43] [0.57; 2.21] [0.67; 1.80] [0.67; 1.80]
Hypertension ALSPAC G0 TwinsUK Common effect model Random effects model Heterogeneity: I ² = 47%, τ ² =	49 – 275 0.6633, p = 0.168		2.05 1.74	[0.04; 3.69] [0.97; 4.33] [0.86; 3.54] [0.29; 5.60]
High cholesterol ALSPAC G0 TwinsUK Common effect model Random effects model Heterogeneity: / ² = 15%, τ ² =	42 276 0.1054, p = 0.279		0.66	[0.30; 15.71] [0.28; 1.53] [0.36; 1.71] [0.33; 2.15]
Asthma TwinsUK ALSPAC G1 Common effect model Random effects model Heterogeneity: J ² = 36%, τ ² =	139 140 0.1382, p = 0.211	0.1 0.5 1 2 10 Odds ratio for symptoms lasting 4+ weeks	1.06 0.70	[0.16; 1.20] [0.41; 2.74] [0.35; 1.39] [0.30; 1.64]

Sources: ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). The reference category for 'Diabetes', 'Hypertension', 'High Cholesterol', and 'Asthma' is the absence of condition. Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Adjusted for age, sex and ethnicity. Some samples were omitted from specific risk factor long COVID analyses, where sample sizes were too small to contribute, e.g., ALSPAC G0 in the asthma meta-analysis.

Supplementary figure 13: Sub-group meta-analysis results for sociodemographic characteristics with symptoms for 12+ weeks in individuals with COVID-19 status confirmed by a positive PCR test and/or serology result

	Ν	Lower risk	Higher risk	OR	95%-CI
Female sex (ref. male) TwinsUK ALSPAC G1 Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, p	377 193 9 = 0.666	-	*		
Other ethnicity (ref. white) TwinsUK ALSPAC G1 Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, p	377 193 9 = 0.696			0.39 0.72 0.54 0.54	[0.04; 3.75] [0.09; 5.93] [0.11; 2.53] [0.11; 2.53]
No higher education (ref. de ALSPAC G0 ALSPAC G1 TwinsUK Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, p	46 150 — 297		-	0.24 0.14 0.35 0.30 0.30	[0.03; 2.01] [0.02; 1.09] [0.16; 0.79] [0.15; 0.61] [0.15; 0.61]
Index of multiple deprivation ALSPAC G0 ALSPAC G1 TwinsUK Common effect model Random effects model Heterogeneity: $I^2 = 65\%$, $\tau^2 = 0.5\%$	35 183 376	e) 		1.85 0.71 0.96 0.96 1.02	[0.97; 3.54] [0.45; 1.12] [0.80; 1.15] [0.82; 1.13] [0.65; 1.61]
Current smoking (ref. non-s ALSPAC G0 ALSPAC G1 Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$, p	42 168	0.1 0.5 Odds ratio for sympto	1 2 10 ms lasting 12+ weeks	0.82 0.78 0.79 0.79	[0.12; 5.84] [0.17; 3.66] [0.24; 2.68] [0.24; 2.68]

Sources: ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Adjusted for age, sex and ethnicity, where relevant. Some samples were omitted from specific risk factor long COVID analyses, where sample sizes were too small to contribute, e.g., ALSPAC G0 in the ethnicity meta-analysis.

Supplementary figure 14: Sub-group meta-analysis results for health traits with symptoms for 12+ weeks in individuals with COVID-19 status confirmed by a positive PCR test and/or serology result

	Ν	Lower risk Higher risk	OR	95%-CI
Poor overall health (ref. go ALSPAC G0 TwinsUK Common effect model Random effects model Heterogeneity: I^2 = 72%, τ^2 = 2	42 226			[1.29; 78.50] [0.15; 4.08] [0.59; 7.65] [0.22; 31.34]
Psychological distress (ref ALSPAC G0 TwinsUK ALSPAC G1 Common effect model Random effects model Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$,	42 165 104		1.63 0.47 1.47 1.17 1.17	[0.14; 19.18] [0.05; 4.76] [0.36; 6.02] [0.40; 3.45] [0.40; 3.45]
BMI (per 5 kg/m^2) ALSPAC G0 ALSPAC G1 TwinsUK Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$,	23 146 175 <i>p</i> = 0.692		0.68 1.08 0.96 0.97 0.97	[0.27; 1.73] [0.64; 1.82] [0.61; 1.52] [0.70; 1.33] [0.70; 1.33]
Overweight or obese (ref. I ALSPAC G0 TwinsUK ALSPAC G1 Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$,	23 175 144		0.77 1.81 1.48 1.52 1.52	[0.08; 7.10] [0.66; 4.92] [0.56; 3.86] [0.79; 2.95] [0.79; 2.95]
Hypertension ALSPAC G0 TwinsUK Common effect model Random effects model Heterogeneity: $/^2 = 0\%$, $\tau^2 = 0$,	42 275 p = 0.662		0.76 1.35 1.24 1.24	[0.07; 7.95] [0.51; 3.56] [0.50; 3.04] [0.50; 3.04]
High cholesterol ALSPAC G0 TwinsUK Common effect model Random effects model Heterogeneity: $l^2 = 52\%$, $\tau^2 = 0$	42 276 0.7946, <i>p</i> = 0.148			[0.53; 34.21] [0.25; 2.22] [0.41; 2.86] [0.27; 7.26]
Asthma ALSPAC G1 TwinsUK Common effect model Random effects model Heterogeneity: $l^2 = 0\%$, $\tau^2 = 0$,	140 139 p = 0.741		0.63 0.85 0.75 0.75	
		Odds ratio for symptoms lasting 12+ weeks		

Sources: ALSPAC G1 (Children of the Avon Longitudinal Study of Parents and Children); TwinsUK (UK Adult Twin Registry); ALSPAC G0 (parents of ALSPAC). The reference category for 'Diabetes', 'Hypertension', 'High Cholesterol', and 'Asthma' is the absence of condition. Estimates from fixed and random effects meta-analyses are presented as odds ratios (OR) and 95% confidence intervals (CIs) as appropriate. Adjusted for age, sex, and ethnicity. Some samples were omitted from specific risk factor long COVID analyses, where sample sizes were too small to contribute, e.g., ALSPAC G0 in the asthma meta-analysis.

Supplementary Note 1: Information governance and ethics for the OpenSAFELY platform

NHS England is the data controller; TPP is the data processor; and the key researchers on OpenSAFELY are acting on behalf of NHS England. OpenSAFELY is hosted within the TPP environment which is accredited to the ISO 27001 information security standard and is NHS IG Toolkit compliant; 1, 2 patient data are pseudonymised for analysis and linkage using industry standard cryptographic hashing techniques; all pseudonymised datasets transmitted for linkage onto OpenSAFELY are encrypted; access to the platform is via a virtual private network (VPN) connection, restricted to a small group of researchers who hold contracts with NHS England and only access the platform to initiate database queries and statistical models. Pseudonymised structured data include demographics, medications prescribed from primary care, diagnoses, and laboratory measures. No free text data are included. All database activity is logged; only aggregate statistical outputs leave the platform environment following best practice for anonymisation of results such as statistical disclosure control for low cell counts.3 The OpenSAFELY research platform adheres to the obligations of the UK General Data Protection Regulation (GDPR) and the Data Protection Act 2018. In March 2020, the Secretary of State for Health and Social Care used powers under the UK Health Service (Control of Patient Information) Regulations 2002 (COPI) to require organisations to process confidential patient information for the purposes of protecting public health, providing healthcare services to the public and monitoring and managing the COVID-19 outbreak and incidents of exposure; this sets aside the requirement for patient consent.4 Taken together, these provide the legal bases to link patient datasets on the OpenSAFELY platform. GP practices, from which the primary care data are obtained, are required to share relevant health information to support the public health response to the pandemic and have been informed of the OpenSAFELY analytics platform. This study was approved by

1 NHS Digital. Data Security and Protection Toolkit. 2020. https://digital.nhs.uk/data-and-information/looking-after-information/data-securityand-information-governance/data-security-and-protection-toolkit (accessed Aug 11, 2020).

2 NHS Digital. BETA - Data Security Standards. 2020. https://digital.nhs.uk/about-nhs-digital/our-work/nhs-digital-data-and-technologystandards/framework/beta---data-security-standards (accessed Aug 11, 2020).

3 NHS Digital. ISB1523: Anonymisation Standard for Publishing Health and Social Care Data. 2020. https://digital.nhs.uk/data-andinformation/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-andcollections/isb1523-anonymisation-standard-for-publishing-health-and-social-care-data (accessed Aug 11, 2020).

4 Secretary of State for Health-UK Government. Coronavirus (COVID-19): notification to organisations to share information. 2020. https://www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information (accessed Aug 11, 2020).

Supplementary Note 2: Detail of method to derive long COVID by monthly symptom reporting

Born in Bradford

BiB study members who self-reported COVID-19 were asked to report whether any particular symptoms were present during March-September 2020. Specifically, participants were presented with 27 symptoms (i.e. "decrease in appetite", "nausea and/or vomiting", "diarrhoea", "abdominal pain/tummy ache", "sore eyes", "loss of sense of smell or taste", "sore throat", "hoarse voice", "headache", "dizziness", "new persistent cough", "tightness in the chest", "chest pain", "shortness of breath", "fever", "chills", "difficulty sleeping", "felt more tired than normal", "severe fatigue", "numbness or tingling somewhere in the body", "feeling of heaviness in arms or legs", "achy muscles", "raised, red, itchy areas on the skin", "sudden swelling of the face or lips") and asked to tick whether the symptoms were present for the months March – September 2020. Only 24 of the 27 symptoms were included for consideration as three symptoms ("runny nose", "sneezing", and "blocked nose") were considered non-specific.

They were also asked: (1) whether they had symptoms in the past week (yes; no); (2) whether they think they have had COVID-19 (Yes, confirmed by a positive test; Yes, suspected by a doctor but not tested; Yes, my own suspicions; No) which was re-categorized to a binary response of "yes" if they responded affirmatively and "no" if they did not; (3) if yes to (2), when were you told / when did you think you had COVID-19 (write-in 'reported positive' date), and (4) whether they had a positive result from a swab (polymerase chain reaction [PCR]) test (yes; no; don't know) or an antibody test (yes; no; don't know) which was re-categorized to a binary variable with "yes" indicating a positive result from a swab or antibody test, "no" indicating a negative result from a swab or antibody test, and those responding "don't know" coded as missing.

Data were used to derive symptom length categories above by summing included symptoms present for 0-4 weeks; 4-12 weeks or 12+ weeks. Participants were coded to the 0–4-week category if they selected symptoms in the same month as the reported positive date or the participant reported symptoms in the past week and it falls within the same month as positive date (for example, when the reported positive date is after the last symptom month of September and they reported symptoms in the past week).

Participants were coded to the 4–12-week category if they selected 2-3 symptom months. A 3-month duration was allowed only if one of the months is within the same month as the reported positive date.

Participants were coded to the 12+ week category if they selected 3+ symptom months. A 3-month duration was allowed only if the reported positive date was not one of the selected months.

The final analytical sample of long COVID participants (n=110) included only those who responded affirmatively to have had COVID-19 whether confirmed by a positive test, was suspected by doctor but not tested, or from their own judgment.

TwinsUK

In TwinsUK, all study members were asked to report whether they had experienced particular symptoms (33 in total) between February and November 2020. Specifically, twins were presented with 33 symptoms (i.e. "cold or flu symptoms", "decrease in appetite", "nausea and/or vomiting", "diarrhoea", "abdominal pain/stomach ache", "runny nose", "sneezing", "blocked nose", "unusual eye soreness or discomfort", "loss of sense of smell", "loss of sense of taste", "sore or painful throat", "hoarse voice", "headache", "dizziness, light-headedness or vertigo", "shortness of breath or trouble breathing affecting normal activities", "new persistent cough", "tightness in the chest", "chest pain", "racing heart or palpitations", "fever", "chills (feeling too cold)", "difficulty sleeping", "felt more tired than normal", "severe fatigue", "numbness or tingling somewhere in the body", "feeling of heaviness in arms or legs", "strong muscle pains or aches", "shaking or difficulty while walking", "phlegm production/chesty cough", "raised, red, itchy welts on the skin or sudden swelling of the face or lips", "red/purple sores or blisters on feet", "confusion, disorientation or drowsiness") and were asked to report whether they had experienced the listed symptom during the months of February-March 2020; April-May 2020; June-July 2020 (July questionnaire) and / or July-August 2020; September-October 2020 (November questionnaire).

Only 28 of the 33 symptoms were included for consideration as five symptoms ("runny nose", "sneezing", "blocked nose", "shaking or difficulty while walking" and "phlegm production/chesty cough") were considered non-specific.

Similarly, to BiB, data were used to derive the symptom length categories above through summing whether any of the included symptoms (i.e. were present for 0-4 weeks; 4-12 weeks or 12+ weeks, at any point in time over the specified period. This was performed for people who had had COVID-19 and those who had not (confirmed by negative antibody testing).

Supplementary Note 3: Detail of method to derive inverse probability weights (IPW)

Self-reported COVID-19 status was regressed on each exposure to assess whether COVID-19 was associated with each socio-demographic or pre-pandemic health risk factor. To determine what variables to include across LS, observed associations were meta-analysed to identify consistent predictors of COVID-19 self-report status. To avoid missingness on IPWs, covariates included in each model were imputed using multiple imputation by chained equations (MICE) and IPWs were derived across multiple imputed data sets. All statistical analyses on the LS were performed in Stata version 16 or R (release 3.6.0 or later).

Covariates included in all longitudinal studies

- Sex
- Age
- Ethnicity
- Mental Health Score
- BMI

Additional covariates included in some longitudinal studies

- Asthma (NCDS)
- Smoking (NCDS; BCS70; MCS; NS)

Supplementary References:

- Brown M, Goodman A, Peters A, Ploubidis GB, Sanchez A, Silverwood R, et al. COVID-19 Survey in Five National Longitudinal Studies: Wave 1, 2 and 3. User Guide (Version 3). UCL Cent Longitud Stud MRC Unit Lifelong Heal Ageing London, UK [Internet]. 2020;(June):1–62. Available from: https://cls.ucl.ac.uk/wpcontent/uploads/2021/01/UCL-Cohorts-COVID-19-Survey-user-guide.pdf
- 2. Joshi HE, Fitzsimons E. The UK Millennium Cohort Study: the making of a multi- purpose resource for social science and policy in the UK. Longit Life Course Stud. 2016;7(4):409–30.
- 3. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort profile: The 'Children of the 90s'-The index offspring of the avon longitudinal study of parents and children. Int J Epidemiol. 2013;42(1):111–27.
- 4. Calderwood L, Sanchez C. Next Steps (formerly known as the Longitudinal Study of Young People in England). 2016;2–4.
- 5. Elliott J, Shepherd P. Cohort profile: 1970 British Birth Cohort (BCS70). Int J Epidemiol. 2006;35(4):836–43.
- 6. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). Int J Epidemiol. 2006;35(1):34–41.
- 7. Wright J, Small N, Raynor P, Tuffnell D, Bhopal R, Cameron N, et al. Cohort profile: The born in bradford multiethnic family cohort study. Int J Epidemiol. 2013;42(4):978–91.
- 8. Dickerson J, Bird PK, McEachan RRC, Pickett KE, Waiblinger D, Uphoff E, et al. Born in Bradford's Better Start: An experimental birth cohort study to evaluate the impact of early life interventions. BMC Public Health [Internet]. 2016;16(1):1–14. Available from: http://dx.doi.org/10.1186/s12889-016-3318-0
- 9. Institute for Social and Economic Research. Understanding Society COVID-19 User Guide. Colchester; 2021.
- Smith BH, Campbell A, Linksted P, Fitzpatrick B, Jackson C, Kerr SM, et al. Cohort profile: Generation scotland: Scottish family health study (GS: SFHS). The study, its participants and their potential for genetic research on health and illness. Int J Epidemiol. 2013;42(3):689–700.
- 11. Fraser A, Macdonald-wallis C, Tilling K, Boyd A, Golding J, Smith GD, et al. Cohort Profile : The Avon Longitudinal Study of Parents and Children : ALSPAC mothers cohort. 2013;(April 2012):97–110.
- 12. Verdi S, Abbasian G, Bowyer RCE, Lachance G, Yarand D, Christofidou P, et al. TwinsUK: The UK Adult Twin Registry Update. Twin Res Hum Genet. 2019;(May 2007):1–7.
- 13. Suthahar A, Sharma P, Hart D, García MP, Horsfall R, Bowyer RCE, et al. TwinsUK COVID-19 personal experience questionnaire (CoPE): wave 1 data capture April-May 2020 [version 1; peer review : awaiting peer review]. 2021;(May 2020):1–10.