The effect of social distancing on the reproduction number and number of contacts in the UK from a social contact survey
Report 4

Authors: Christopher Jarvis, Amy Gimma, Kevin Van Zandvoort, Petra Klepac, and John Edmunds on behalf of the LSHTM COVID-19 Modelling Team

Date: 23rd April 2020

Main conclusion

We estimate that R_0 has remained unchanged and is still below one. Potential regional differences need to be explored further.

Aims

1. Assess temporal changes in contact patterns and the reproduction number of COVID-19 under social distance measures in the UK.
2. Assess regional differences in the number of contacts for different settings across the UK.
3. Assess characteristics of individuals for contacts outside of the house.

Change since last report:

1. We have now weighted the POLYMOD data for weekend- and weekdays.
2. Contact ages are recorded in groups. We previously used the mean of the age-group as the estimated age but now sample uniformly between the minimum and maximum age reported for the contact.
3. Change to use the reciprocity of contacts of adult-child contacts instead of imputing from POLYMOD.
4. Changed age bands to be 0-4, 5-12, 13-17 instead of 0-17 to be consistent with the BBC Pandemic study.
5. Decreased the number of bootstrap samples from 5,000 to 2,000.
6. We now model contact counts using a Negative Binomial rather than a Poisson distribution, which better reflects heterogeneity in individual number of contacts.

Methods

CoMix is a behavioural survey, with a study sample recruited to be broadly representative of the UK adult (18+) population. It was launched on 24th of March 2020 and this analysis includes data collected up to the 21st of April. Data is collected weekly, using two different panels who are interviewed using the same questionnaire in alternate weeks. Both panels started with a sample size of 1,816 (Panel A) and 1,560 (Panel B) individuals, Panel A has now decreased to 1,326 and interim data for Panel B has 1,079 participants. Participants recorded direct, face-to-face

contacts that they made on the previous day, specifying certain characteristics for each contact including the age and sex of the contact, whether contact was physical (skin-to-skin contact), and where contact occurred (e.g. at home, work, while undertaking leisure activities, etc). Further details have been published elsewhere.[1] The contact survey is based on the POLYMOD contact survey, which is used as a baseline for social mixing in the UK under normal conditions.[2]

Change in contact patterns over time

We calculated the average number of contacts in different settings for each of the four weeks of the survey.

In contrast to previous reports we are now weighting the POLYMOD data for weekend- and weekdays. We previously used the mean of the age-group as the estimated age but now sample uniformly between the minimum and maximum age reported for the contact, as we do not record exact ages for contacts. We use the reciprocity of contacts of adult-child contacts instead of imputing from POLYMOD. We also changed the age bands for under 18s to be 0-4, 5-12, 13-17 instead of 0-17 to be consistent with the BBC Pandemic study. We impute child-child contacts using the POLYMOD UK data, setting school-contacts to 0 and adjusting contact in other settings (e.g. home) as observed for adults. Further details of the approach can be found here.[1,3]

We assume that R_0 prior to physical distancing measures were in place follows a normal distribution with a mean of 2.6 and sd of 0.54. We then apply a scaling factor of the ratio of dominant eigenvalues between CoMix and Polymod contact matrices to estimate R_0 under the observed contacts patterns in our study following the approach found in Wallinga et al.[4]. This assumes that all other elements of the Next Generation Matrix remain constant, such as transmissibility by age group, which may not be the case.

Uncertainty in the estimates of reduction in R_0 is obtained using 2,000 bootstrap samples of the CoMix and POLYMOD contacts matrices, and applying these ratios to 2,000 sampled values of R_0.

We repeated this process separately for each of the four weeks of data collection so far and present estimates of R_0 for two scenarios: if transmission would be driven by all direct conversational (face-to-face) contacts, and if it would be driven by physical conversational contacts.

Regional differences in contacts for different settings across the UK.

We assessed regional differences in the total number of contacts. We used generalised additive models (GAM) to calculate the relative and absolute difference in number of contacts, for all contacts, contacts within the home, and contacts outside of the home. Relative differences were

calculated using an individual-level GAM assuming contacts followed a negative binomial distribution (modelled using a log link function), with smoothed terms for age and household size, fixed terms for region, gender, week of survey, and survey panel, and a random effect for participants. Absolute differences were calculated using a similar individual-level GAM assuming a negative binomial distribution, but modelled using an identity link function, with a random effect for participants, and a fixed term for week of survey and region. The model for absolute differences was simplified (fewer predictor variables) due to additional complexities of fitting a GAM with an identity link function and negative binomial distribution. We then assessed the relationship between mean contacts by region and estimated effective reproductive number from the Epiforecast report using a generalized linear model.

Results

Between the 24th of March and the 21st of April we collected 5,781 observations from 3,376 individual participants, with information on 14,712 contacts. These consist of two full rounds of survey for panel A and panel B, and a third, partial, round for Panel A.

Change in contact patterns and estimation of reproduction number

Table 1 gives the reported number of contacts made by the participants and the overall estimate of the reproduction number for the fours weeks of the survey (also shown in Figure 1).

For week 4 (Between 16th April and 21st April) we estimated R_0 to be 0.44 (95% CI 0.25 to 0.63) if transmission is driven by all direct contacts and 0.39 (95% CI 0.22 to 0.56) if transmission is driven by physical contacts only. There was a small suggestion of a reduction in average number of daily contacts per person from 2.72 in week 1 to 2.32 in week 3 for all contacts, though interquartile ranges remain similar for all 4 weeks. The ranges of R_0 are consistent for all four weeks though the point estimates show some decrease although caution should be taken not to over-interpret this as there are changes in participants between the weeks. The average number of physical contacts and R_0 for physical contacts only are near identical across the four weeks. The majority of contacts continue to occur within homes and other indoor settings, most participants report few contacts at work, with some outliers reaching over 50 contacts.

We originally estimated R_0 to be 0.61 (0.36, 0.86) in the first week of the survey but have since made multiple changes to our approach, we are exploring what aspect resulted in this reduction.

From Week 1 to Week 3 we lost 490 individuals in panel A, while 481 participants from Panel B have not yet completed this week's survey. The mean number of contacts of repeat respondents was 2.53, while it was 2.60 for individuals who have only completed the survey once. We will continue to monitor loss to follow-up and potential differences between individuals who remain in or drop out of the study.

Table 1. Numbers of participants, reported contacts and reproduction numbers by week.
Numbers of participants (N) in each panel, their number of contacts reported and the estimate of the reproduction number, R_0 by week and type of contact.

Week	Panel	Dates	N	Contact Type	Contacts	Mean (IQR)	R_0 mean (95% CI)
1	A	24/03 to 01/04	1816	All	4941	2.72 (1 to 4)	0.53 (0.33 to 0.75)
			1816	Physical	1522	0.84 (0 to 1)	0.39 (0.23 to 0.56)
2	B	02/04 to 10/04	1560	All	4002	2.57 (1 to 3)	0.48 (0.28 to 0.69)
			1560	Physical	1266	0.81 (0 to 1)	0.37 (0.21 to 0.53)
3	A	07/04 to 15/04	1326	All	3263	2.46 (1 to 3)	0.49 (0.29 to 0.68)
			1326	Physical	1119	0.84 (0 to 1)	0.39 (0.22 to 0.55)
4	B	16/04 to 21/04	1079	All	2506	2.32 (1 to 3)	0.44 (0.25 to 0.63)
			1079	Physical	908	0.84 (0 to 1)	0.39 (0.22 to 0.55)

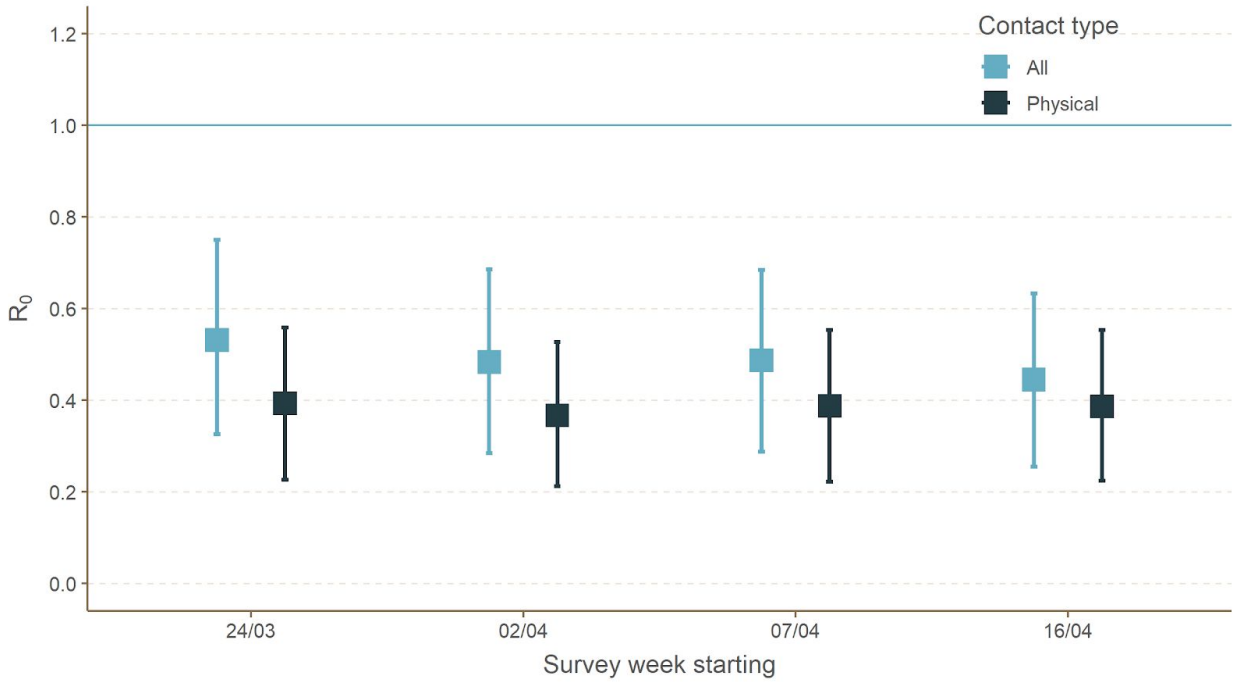

Figure 1: Weekly R_0 estimates.

Regional differences in contacts for different settings across the UK.

Figure 2 displays the relative and absolute difference between the average number of contacts in all regions compared to London adjusting for participant's age and household size. All regions were higher compared to London although the North East, Northern Ireland, East of England and Scotland were still consistent with Greater London for relative and absolute difference. All other regions had a higher average number of contacts with Yorkshire and Humbers, the South West and Wales having the largest relative and absolute differences. The absolute differences between regions are small with all regions within 0.6 contacts of London for each setting. However, stratifying the analysis by contact within and outside of the house suggests that these increases are heavily driven by differences in contacts outside of the house as opposed to within the house.

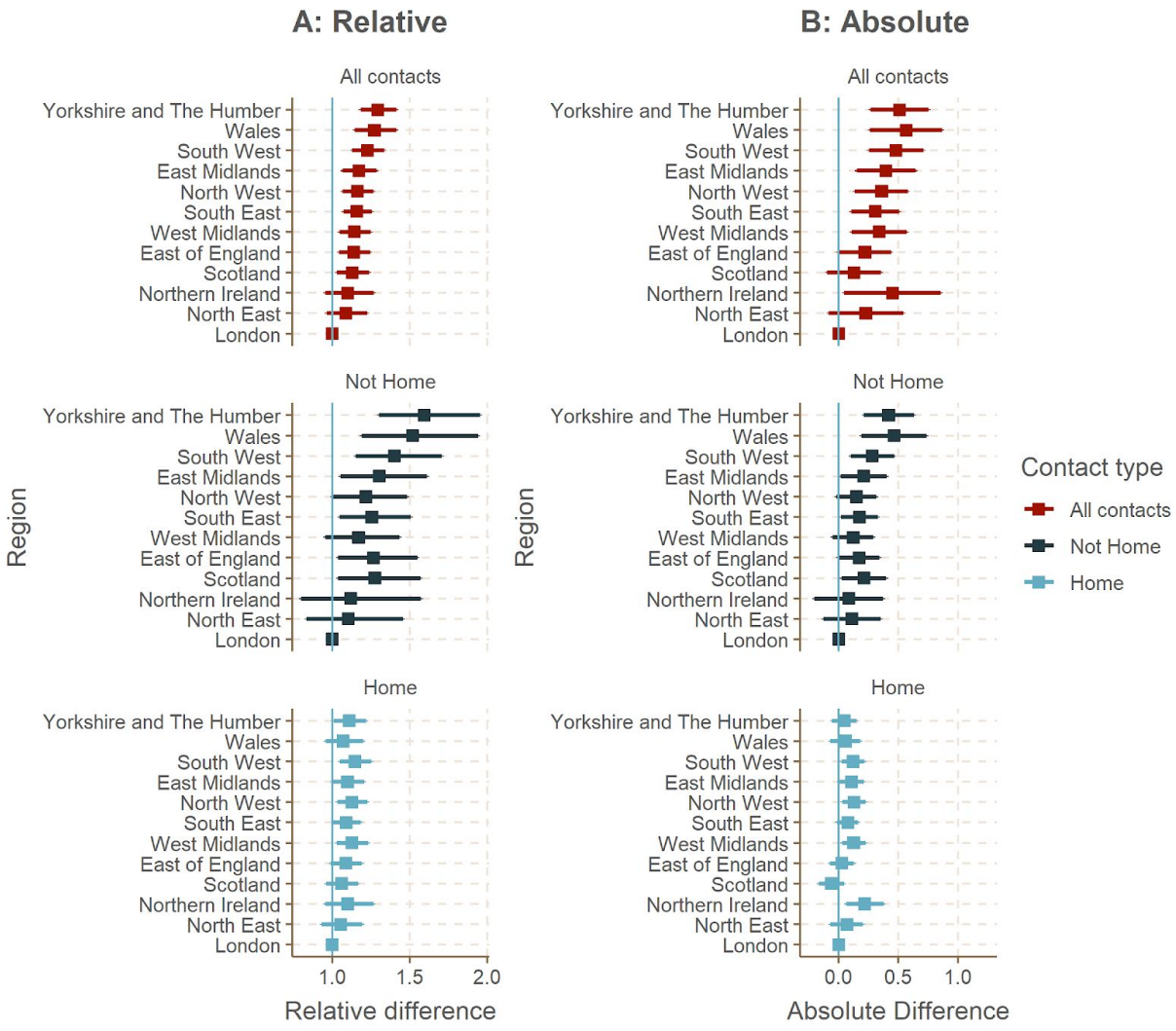

Figure 2: Comparison of number of contacts by setting comparing all regions to London.
A: The relative difference in number of contacts for overall, contacts at home, and contacts not at home

house. B: The absolute difference in number of contacts for overall, within the house, and outside of the house. Relative differences were calculated using an individual-level GAM assuming contacts followed a negative binomial distribution (modelled using a log link function), with smoothed terms for age and household size, fixed terms for region, gender, week of survey, and survey panel, and a random effect for participants. Absolute differences were calculated using a similar individual-level GAM assuming a negative binomial distribution, but modelled using an identity link function, with a random effect for participants, and a fixed term for week of survey and region. The model was simplified for the absolute difference due to extra complexities of fitting a GAM with an identity link function with a negative binomial distribution.

As our sample size was not large enough to directly explore R_0 by region, we explored the relationship between the reported mean number of contacts in the CoMix survey and the effective reproduction number as estimated by Abbot et al.,[5] shown in Figure 3. As Abbot et al. estimates a daily reproduction number, we took the mean value of their regional estimates since the first day of the CoMix survey and plotted its relationship with the mean number of contacts reported across all weeks of the CoMix survey. While there is an indication of a correlation between the two estimates, there is large variability between them. We will investigate this relationship in more detail, using different statistical techniques.

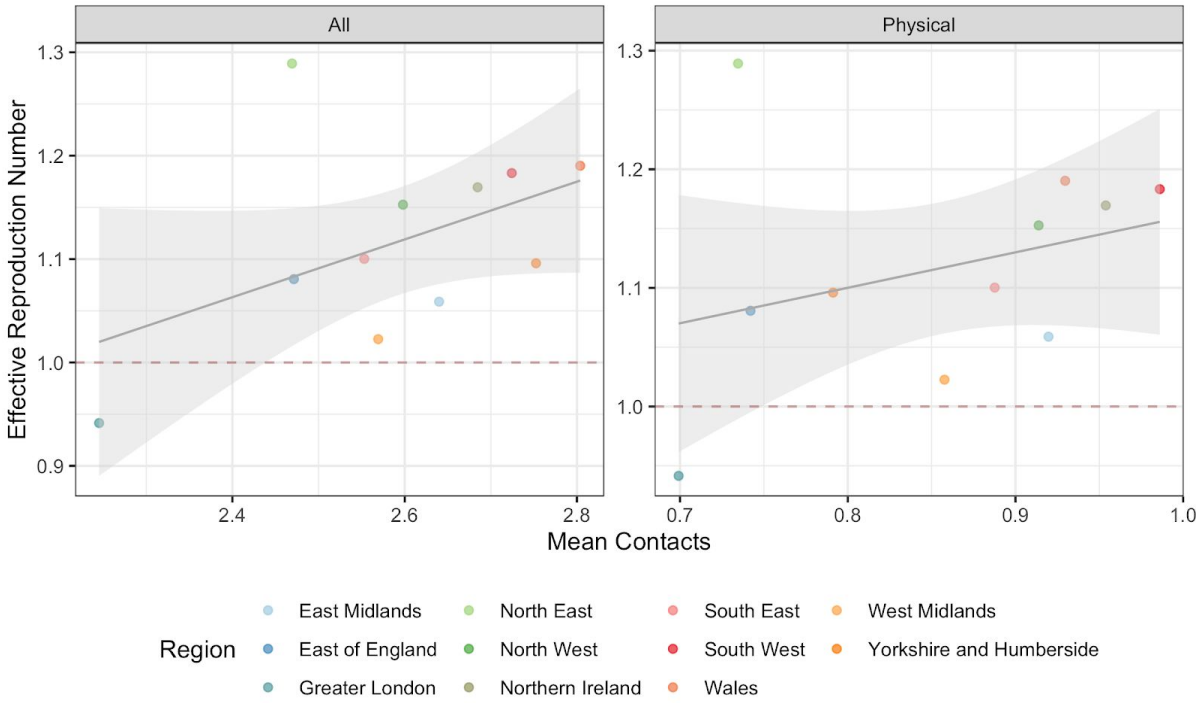

Figure 3. Regional comparison of the relationship between a model-based estimate of the effective reproduction number and the mean number of contacts

Characteristics of individuals for contacts outside of the house

We explored the reported mean number of contacts outside of the home for observations in week 3 and 4 of the survey. The mean number of contacts outside of the home was 0.87 (IQR 0 to 1) for all participants. Table S1 shows the mean number of contacts outside the home reported in different groups within our survey for week 3 and 4. Notably, individuals reporting not having their work closed report having twice the mean number of contacts (0.78; 0 to 1) compared to those who reported their workplace has not been closed (1.49; 0 to 2). Among individuals reporting at least one contact within the workplace, and from which profession we had at least 10 individuals in our dataset, the five professions with the highest mean number of contacts within the workplace are i. nursing and midwifery professionals (7.9 contacts), ii. personal care and related workers (7.5 contacts), iii. teaching professionals (7.4 contacts), iv. customer service clerks (5.6 contacts), and v. health professionals (except nursing; 5.5 contacts).

Discussion

The fourth week of the CoMix survey corresponds to the fourth week of the lockdown in the UK. There is no evidence of any changes in behaviour over these four weeks, and so we estimate that the reproduction number has remained unchanged and is still below one (assuming that it was 2.6 on average before physical distancing interventions).

We noted some regional differences in behaviour, particularly regarding contacts outside the home. Participants from London report fewer contacts than elsewhere, particularly work contacts, whereas participants from Wales, the South West and Yorkshire and Humberside report greater numbers of contacts. Although differences are small in absolute numbers, they are large relative to each other. Further work is required to understand what might be driving these differences. There is a slight suggestion of correlation between the mean number of contacts per region and estimates of the reproduction number from Abbott et al.[5] Latest regional nowcasts (from the 12th April, Abbott et al.[5]), which are based on the epidemiological data (and therefore lagged by 2-3 weeks) suggests that the reproduction number in London is somewhat lower compared to other regions (0.6; CI 0.5-0.7) while the South West has the highest regional reproduction number in England (0.8; CI 0.6-1), in line with our behavioural findings. Across all regions in the UK, Abott et al now estimate the highest regional reproduction number in Northern Ireland (1.3; 1 - 1.6), but this estimate is hard to compare to our CoMix survey due to the small number of respondents in Northern Ireland.

We found that individuals whose workplace was not closed have twice the number of contacts outside the household compared to those whose workplace was closed. Professions with the highest mean number of direct workplace contacts were mostly those of key workers.

References

1 Jarvis CI, Van Zandvoort K, Gimma A, *et al.* Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. Epidemiology. 2020; published online April 3. DOI:10.1101/2020.03.31.20049023.

2 Mossong J, Hens N, Jit M, *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med* 2008; **5**: e74.

3 Klepac P, Kucharski AJ, Conlan AJK, *et al.* Contacts in context: large-scale setting-specific social mixing matrices from the BBC Pandemic project. Epidemiology. 2020; published online Feb 19. DOI:10.1101/2020.02.16.20023754.

4 Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am J Epidemiol* 2006; **164**: 936–44.

5. Epiforecasts. National and Subnational estimates for the United Kingdom [Internet]. CMMID; 2020 Apr [cited 2020 Apr 22]. (Temporal variation in transmission during the COVID-19 outbreak). Available from: https://github.com/epiforecasts/covid-regional/tree/master/united-kingdom/regional-summary

Appendix

Table S1: Breakdown of contacts outside the home by different characteristics for Week 3 and 4 combined.

Category	Values	N	Mean contacts	IQR	Min and Max
Gender	Female	1225	0.89	0 to 1	0 to 30
	Male	1172	0.87	0 to 1	0 to 20
Age	[18,30)	234	0.79	0 to 1	0 to 20
	[30,40)	352	0.73	0 to 1	0 to 24
	[40,50)	414	0.77	0 to 1	0 to 11
	[50,60)	508	1.18	0 to 1	0 to 30
	[60,70)	545	0.93	0 to 1	0 to 16
	[70,120)	352	0.7	0 to 1	0 to 11
Social group	A - Upper middle class	123	0.67	0 to 1	0 to 24
	B - Middle class	649	0.81	0 to 1	0 to 20
	C1 - Lower middle class	825	0.85	0 to 1	0 to 16
	C2 - Skilled working class	378	0.98	0 to 1	0 to 15
	D - Working class	362	1.12	0 to 1	0 to 30
	E - Lower level of subsistence	68	0.4	0 to 1	0 to 2
Income	Under £10,000	120	0.57	0 to 1	0 to 4
	£10,000 - £19,999	460	0.88	0 to 1	0 to 16
	£20,000 - £34,999	703	0.94	0 to 1	0 to 20
	£35,000 - £54,999	502	1.07	0 to 1	0 to 30
	£55,000 - £99,999	295	0.78	0 to 1	0 to 16
	£100,000 or more	75	0.77	0 to 0	0 to 24
Employment status	Employed full-time (34 hours or more)	857	1.09	0 to 1	0 to 30
	Employed part-time (less than 34 hours)	375	1.02	0 to 1	0 to 16
	Full-time parent, homemaker	88	0.61	0 to 1	0 to 5
	Long-term sick or disabled	89	0.34	0 to 1	0 to 3
	Retired	628	0.7	0 to 1	0 to 5
	Self employed	194	0.97	0 to 1	0 to 11
	Student/Pupil	56	0.43	0 to 1	0 to 4
	Unemployed and not looking for a job	42	0.52	0 to 1	0 to 3
	Unemployed but looking for a job	76	0.45	0 to 1	0 to 4
Work closed	Yes	725	0.78	0 to 1	0 to 24
	No	554	1.49	0 to 2	0 to 30
Limit work at least one day in the previous seven	Yes	801	0.76	0 to 1	0 to 24
	No	477	1.56	0 to 2	0 to 30

Table S1 (continued)

Category	Values	N	Mean contacts	IQR	Min and Max
High risk for Covid	Yes	776	0.69	0 to 1	0 to 30
	No	1567	0.96	0 to 1	0 to 24
Covid would be serious for me	Strongly agree	704	0.67	0 to 1	0 to 30
	Tend to agree	746	0.92	0 to 1	0 to 20
	Neither agree nor disagree	451	0.89	0 to 1	0 to 10
	Tend to disagree	271	1.15	0 to 1	0 to 16
	Strongly disagree	58	1.33	0 to 1.75	0 to 24
Likely to get Covid	Strongly agree	94	1.15	0 to 1.75	0 to 20
	Tend to agree	423	1.06	0 to 1	0 to 30
	Neither agree nor disagree	958	0.85	0 to 1	0 to 16
	Tend to disagree	541	0.82	0 to 1	0 to 11
	Strongly disagree	109	0.72	0 to 1	0 to 24
Asked to quarantine	No	2215	0.9	0 to 1	0 to 30
	Yes	170	0.62	0 to 0.75	0 to 24
Quarantine one day in the previous seven	No	2015	0.93	0 to 1	0 to 30
	Yes	361	0.63	0 to 1	0 to 24
Asked to isolate	No	2027	0.93	0 to 1	0 to 30
	Yes	354	0.61	0 to 1	0 to 24
Isolated one day in the previous seven	No	1647	1.01	0 to 1	0 to 30
	Yes	738	0.6	0 to 1	0 to 24