Social contacts during the COVID-19 pandemic in 2020 and 2021 in The Netherlands (CoMix survey)

Jantien A. Backer[1,*], Laurens Bogaardt[1], more authors[?], and Jacco Wallinga[1,2]

[1]National Institute for Public Health and the Environment, Bilthoven, the Netherlands
[2]Leiden University Medical Center, Leiden, the Netherlands
[*]jantien.backer@rivm.nl

ABSTRACT

The COVID-19 pandemic in 2020 and 2021 was for a large part mitigated by reducing contacts in the general population. To track how these contacts changed over the course of the pandemic, social contacts were measured in The Netherlands every two weeks, as part of the European CoMix survey. The survey included 1659 participants from April to August 2020 and 2514 participants from December 2020 to September 2021.

Over the course of the study period, drop-out rates differed between age groups, and a fatigue effect occurred where participants tend to report fewer contacts the longer they participate. We classified the number of reported contacts in activity levels, and analysed these as ordered categorical variables using a generalised additive model. After correcting for age, vaccination and risk status, and participation fatigue, activity levels increased during both study periods, coinciding with relaxation of restrictions. Multi-country longitudinal contact surveys such as CoMix are a useful tool to track how restrictive measures and compliance affect social mixing, provided they can be deployed early and consistently over time.

Introduction

From early 2020 onwards, COVID-19 grew into a global pandemic within a few months affecting nearly all countries worldwide. SARS-CoV-2 virus that causes COVID-19 is transmitted through contacts between persons. Reducing transmission can therefore be achieved in two ways. First, by lowering the transmission probability per contact, by maintaining a safe distance from each other or using protective equipment such as a face mask. And second, by reducing the number of contacts per person, by direct or indirect measures. Direct measures such as closing of schools, theaters and restaurants, make having contacts at that location impossible. Indirect measures such as recommendations for teleworking or inviting fewer people at home rely on compliance.

Over the course of the pandemic, the stringency of measures to reduce the number of contacts varied over time and differed by country. To measure the effect of both direct and indirect measures, the CoMix survey was set up in April 2020 in Belgium, the UK, and the Netherlands, later followed by 19 other European countries[1]. In this contact study, participants were asked to regularly report their contact behaviour, i.e. the number of persons they had contact with in certain locations, as well as some characteristics of the contacted persons such as age and sex.

Conversational ('face-to-face') contacts as measured in contact surveys were found to be a good proxy for transmission of respiratory diseases, proving their worth for infectious disease modelling[2]. The most famous contact survey is the POLYMOD study[3] that served as a blueprint for many others[4]. These pre-COVID-19 studies have in common that they aimed to measure 'normal' contact behaviour. The unprecedented measures that were taken during the pandemic gave a unique opportunity to determine contact behaviour under restricted conditions. Many countries set up contact surveys to learn about the effect of measures on contact behaviour[5–12]. What sets the CoMix survey apart is that it uses a standardised questionnaire in multiple countries and is repeated at regular intervals.

Our aim is to describe the contact behaviour of the Dutch general population, using the CoMix data that was collected in the Netherlands during the survey periods in 2020 and 2021. We will focus on the number of contacts per participant (later aggregated in activity levels). The model results are compared to four rounds of the Pienter Corona (PiCo) survey[8,13] that contains a simpler contact survey and is conducted at greater intervals in a representative sample of the Dutch population.

Methods

Participant recruitment and survey design

This contact survey was part of the CoMix study that was conducted in several EU countries and the UK[1]. Participants were recruited by the market research company Ipsos-MORI, and filled out the contact survey online. The study was carried out in two survey periods: 8 survey rounds from 16 April to 5 August 2020, referred to as the 2020 series, and 20 survey rounds from 23 December 2020 to 22 September 2021, referred to as the 2021 series. At the start of the 2020 series 1500 participants of 18 years and older were recruited, reflecting the age distribution of the Dutch population. An overall drop-out rate of 10% per round was allowed, but these rates differed per age group leading to very few participants in the youngest age group of 18-24 in the last rounds of the 2020 series (Fig 1). For this reason, the younger age groups were oversampled at the start of the 2021 series with 1200 participants of 18 years and older. In addition, 300 children between 0 and 17 years of age were included by asking adult participants to fill out the contact survey for one of their children. After 10 rounds, the age cohorts were supplemented to include 1500 participants in total again in May 2021.

The contact survey was repeated every two weeks. In each round participants were asked about characteristics that can change over time, such as risk perception, risk status and vaccination status against COVID-19. High risk participants were people with chronic respiratory disease, chronic heart disorder, chronic kidney disease, diabetes, reduced resistance to infection (due to illness or medication), morbid obesity (BMI over 40), and pregnant women, i.e. reflecting the medical indication for influenza vaccination. Missing data in the risk status were completed with the most common risk status reported over the survey rounds by the participant. The risk status of participants can change by survey round. The fraction of high risk participants is calculated per age group using only the participants with unambiguous risk status, and compared to the population fraction with a medical indication for influenza vaccination[14]. Participants are considered to be vaccinated after they have reported to have had at least one vaccination dose. The fraction of vaccinated participants is compared to the vaccination coverage of one vaccine dose in the general population over time[15].

In the main part of the survey, participants report their contact behaviour of the previous day. This consists of the persons they had conversational or physical contact with and their characteristics, such as age group and gender, as well as duration and location of the contact. From the third survey round onwards in late May 2020, it was also possible to report group contacts, where the total number of contacted persons in broad age groups (0-17, 18-64, 65+) at work, school or another location could be filled out, instead of individually. More details on the CoMix study and questionnaire can be found in previous publications[1,16,17].

Analysis of number of contacts

The number of contacts per participant is studied separately for three age groups (0-17, 18-64, 65+) and two series (2020, 2021). Only contacts with non-household members are included, as these are most likely to change over the course of the study period. Twelve occasions were excluded where participants reported more than 1000 contacts in a single round, and 58 participants were excluded that participated in four or more rounds but did not report any contact.

Previous research has shown that contacts often follow a power-law distribution, where most people have relatively few contacts per day while some have many[18]. The number of contacts was visually checked for such powerlaw-like behaviour by plotting the complementary cumulative distribution function of the number of contacts on a loglog scale. A linear relation indicates a powerlaw, and as a consequence the distribution of number of contacts lacks clearly defined moments. Depending on the powerlaw coefficient, even the mean and variance of the distribution may lack meaning[19], which precludes the use of parametric distributions. One option to overcome this, is to cap the number of contacts to a lower maximum, such as 50, and assume the number of contacts follows a negative binomial distribution. Instead, we define activity levels of 0, 1, 2, 3-4, 5-9, $\geq$ 10 contacts per participant per round and analyse these using ordinal regression. The levels are chosen in such a way that the activity levels higher than 0 contain similarly sized fractions of reported number of contacts.

The activity levels are outcome variables of a generalised additive model with fixed effects for participant vaccination status, participant risk status, weekends and (school)holidays, cubic splines for calendar time, participant age, participant round (i.e. the n^{th} time a participant participated), and random intercepts for participant id:

$$Pr(\text{activity} \leq l, X) = \Phi[\theta_l - (\beta_{\text{id}} + \beta_0 + \beta_1 * s(\text{date}) + \beta_2 * s(\text{part\_round}) + \beta_3 * s(\text{part\_age}) + \beta_4 * \text{part\_vacc} + \beta_5 * \text{part\_risk} + \beta_6 * \text{weekend} + \beta_7 * \text{holiday})],$$

where Φ is the logistic function, θ_l are the set of thresholds identifying the boundaries between the activity levels l, β_{id} are the random intercepts centered about zero and the $s()$ function indicates that splines are used. The fits are assessed by explained deviance and by comparing the predicted activity levels to the observations. Next, the fitted results are used in a synthetic population that has the size of the study population, but reflects the general population. It is constructed by sampling age, risk status and vaccination status at each day from the general population[20]. Participant id's are sampled with replacement from the study population, and participant round is set to 1 for each day during the study period. 200 parameter sets are sampled from the estimated model parameters, and combined with 200 synthetic populations, to capture both the uncertainty of the model parameters and the uncertainty caused by the sample size of the study population. The predicted activity levels of the synthetic population are compared to the activity levels measured in four Pienter Corona (PiCo) surveys.

All analyses are performed in R version 4.1.3[21] using the *mgcv* package[22] with family *ocat* for ordered categorical variables. Code is available at REFgithub, and data is published at Zenodo[23] in the standardised format of *socialcontactdata.org*.

Results

Study population

In total, 1659 participants were included in the 2020 series, and 2514 participants in the 2021 series. Participants were equally distributed over men and women in 2020, while female participants were slightly overrepresented in 2021. Household sizes of 2 persons were most common (Tab. 1). Each series started with a target population of 1500 participants, distributed over different age groups according to the Dutch population over 18 years of age in the 2020 series, and according to a preferential sampling scheme in the 2021 series (Tab S1, Fig S1). The number of participants declined after recruitment. In 2020 the drop-out rate was on average 4% per round in the oldest age group and 14% in the youngest age group; in 2021 drop out rates were more comparable for all age groups (Fig. 1).

The 2021 series coincided with the roll-out of the COVID-19 vaccination campaign. Health care workers, persons with a high medical risk and care home residents were vaccinated with priority, while the non-risk groups were vaccinated from elderly to young. Most persons were vaccinated with either BionTech/Pfizer, Moderna or Janssen vaccine by the Municipal Health Services, except for 60-64 year olds who were vaccinated at an early stage with AstraZeneca by their general practitioner. The vaccination coverage of the study population increases over time largely in agreement with the general population (Fig. 2). The participants in the age groups of 18-34 years achieve a higher vaccination coverage at an earlier stage than the general population. The participants in the age group of 65 years and older lag behind the general population. The fraction of high risk adult participants is higher than in the general population.

Contact behaviour

The distribution of the number of contacts is plotted on a loglog scale, showing that the adult and elderly age groups in both series exhibit powerlaw-like behaviour (Fig. 3). Powerlaw coefficients for these four groups range from 1.1 to 1.5, indicating that the mean is defined but the variance is not. This confirms the choice to estimate activity levels, rather than a parametric distribution.

Five separate analyses are carried out for the different series and age group combinations. Of the fixed effects only the weekends significantly reduces the activity levels in all analyses (Tab. S2). Activity levels in general increase over time during the two study periods, indicating that the underlying number of contacts increases (Fig. S2). In all analyses the activity levels decrease with participant round. This means that participants tend to report fewer contacts the more often they participate. This is especially the case for the first few rounds, after which the effect stabilises. For age within the age group, no clear trend is observed.

The deviance explained by this model ranges from 14% to 21% (Tab. S2). Using these model results the activity levels are predicted for the study population and compared to the data (Fig. 4). With the exception of some rounds, notably the first round in 2020 and the fourth round in 2021, the predicted activity levels agree well with the observed activity levels.

Next, the model results are applied to the general population over the full course of the study period (Fig. 5). The saw-tooth pattern is caused by the weekend effect, which is largest for the 0-17 age group. Over time activity levels increase in all age groups in both series, indicated by the decrease of the predicted lines in Fig. 5. As a comparison, the results of the representative PiCo study[8] are plotted, of which rounds 1, 2, 4 and 5 fell within the CoMix study periods. Participants of the PiCo study reported higher activity levels than CoMix participants. The discrepancy is largest for the 0-17 age group, whereas the 65+ age group all but agrees with the PiCo study results.

Discussion

In this paper we analysed the Dutch data of the European CoMix survey to describe contact behaviour in the general population during two study periods. Activity levels increased over the course of each study period corresponding to the lifting of measures. In 2020, the lockdown that was imposed on March 16th, was lifted stepwise on May 11th (opening of schools), June 1st (outside bars and restaurants open a.o.) and July 1st (inside bars and restaurants open a.o.). The 2021 study period started with an even stricter lockdown including an evening curfew, that was relaxed from March 23rd onwards while vaccination coverages were increasing.

The independent contact surveys of the PiCo studies show a similar increasing trend in activity levels in the general population as our results, but the activity levels reported in PiCo are higher than in the CoMix data. The two studies differed in how the participants report their contacts. In the CoMix study contacts were reported individually with a lot of detail on location, duration, distance from contact, protection measures, etc. Also for the group contacts at work, school, or other places additional details were requested. In the PiCo study a participant reported the total number of contacts in specific age classes, which made reporting many contacts less cumbersome. This may partly explain the higher activity levels found in the PiCo study.

A counterintuitive observation in the 2021 series, is that while restrictions were lifted the study population reported fewer contacts. This is caused by the strong fatigue effect where participants tend to report fewer contacts when they participate in more rounds. This fatigue effect is also seen in the CoMix data of other countries, but the fatigue effect in the Netherlands is found to be extreme compared to these[24]. A complicating factor in interpreting the fatigue effect is that the participant rounds are not equally distributed over the study period. Most participants have their first round in survey round 1 in 2020, and survey rounds 1 and 11 in 2021, due to the study design. In an alternative design as was used in the UK, the study population is supplemented to reach the target size in every survey round, which leads to a better distibution of the participant rounds over the survey rounds[17]. Another possible adjustment to the study design is to decrease the frequency of the survey rounds, which may lead to a smaller fatigue effect.

In order to say anything about the general population, the fatigue effect needed to be corrected for, and differences between the study population and the general population had to be taken into account. Particularly the age distribution of the study participants varied over the survey rounds, because drop-out rates differed between age groups. Also vaccination and risk status of the participants did not always reflect the vaccination and risk status in the general population. The participants in the age groups of 18-34 years achieved a higher vaccination coverage at an earlier stage than the general population. This could indicate that these groups contained relatively many health care workers, who were vaccinated with priority and may have a higher willingness to get vaccinated. The participants in the age group of 65 years and older lagged behind the general population, because the study population probably did not contain many care home residents who were vaccinated with priority. The fraction of high risk adult participants was higher than in the general population. The risk status was reported by participants themselves and compared to the fraction of the population that is invited for the annual influenza vaccination. Although the high risk definition is the same for both populations, participants may have assessed their risk status differently than their GP would. This is supported by the finding that the high risk participants are mainly overrepresented in 2020 (Tab. 1) when risks may have been perceived to be higher than in 2021. This would explain the discrepancy, but it can not be excluded that they were really high risk participants.

Our analysis showed that contact data collected in the CoMix survey can be used to describe trends in contact behaviour in the general population. As such it can be useful in tracking the effect of control measures that are aimed at contact reduction[17,25,26], and serve as input for infectious disease models[27–29]. Moreover, the data contains many additional details which facilitates in-depth analyses, such as the effect of risk perception[30] and pregnancy[31]. These results can help guide policy in future waves of COVID-19 or other emerging respiratory diseases and emphasize that contact surveys such as the CoMix study are indispensable for providing a quantitative basis to public health policy.

References

1. Verelst, F. *et al.* SOCRATES-CoMix: a platform for timely and open-source contact mixing data during and in between COVID-19 surges and interventions in over 20 European countries. *BMC Med* **19**, 254, DOI: http://dx.doi.org/10.1186/s12916-021-02133-y (2021).
2. 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.* **164**, 936–944, DOI: http://dx.doi.org/10.1093/aje/kwj317 (2006).
3. Mossong, J. *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74, DOI: http://dx.doi.org/10.1371/journal.pmed.0050074 (2008).
4. Hoang, T. *et al.* A Systematic Review of Social Contact Surveys to Inform Transmission Models of Close-contact Infections. *Epidemiology* **30**, 723–736, DOI: http://dx.doi.org/10.1097/EDE.0000000000001047 (2019).
5. Zhang, J. *et al.* Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* **368**, 1481–1486, DOI: http://dx.doi.org/10.1126/science.abb8001 (2020).
6. Latsuzbaia, A., Herold, M., Bertemes, J. P. & Mossong, J. Evolving social contact patterns during the COVID-19 crisis in Luxembourg. *PLoS One* **15**, e0237128, DOI: http://dx.doi.org/10.1371/journal.pone.0237128 (2020).
7. Quaife, M. *et al.* The impact of COVID-19 control measures on social contacts and transmission in Kenyan informal settlements. *BMC Med* **18**, 316, DOI: http://dx.doi.org/10.1186/s12916-020-01779-4 (2020).
8. Backer, J. A. *et al.* Impact of physical distancing measures against COVID-19 on contacts and mixing patterns: repeated cross-sectional surveys, the Netherlands, 2016-17, April 2020 and June 2020. *Euro Surveill* **26**, DOI: http://dx.doi.org/10.2807/1560-7917.ES.2021.26.8.2000994 (2021).

9. Bosetti, P. *et al.* Lockdown impact on age-specific contact patterns and behaviours, France, April 2020. *Euro Surveill* **26**, DOI: http://dx.doi.org/10.2807/1560-7917.ES.2021.26.48.2001636 (2021).

10. Feehan, D. M. & Mahmud, A. S. Quantifying population contact patterns in the United States during the COVID-19 pandemic. *Nat Commun* **12**, 893, DOI: https://dx.doi.org/10.1038/s41467-021-20990-2 (2021).

11. Tomori, D. V. *et al.* Individual social contact data and population mobility data as early markers of SARS-CoV-2 transmission dynamics during the first wave in Germany-an analysis based on the COVIMOD study. *BMC Med* **19**, 271, DOI: http://dx.doi.org/10.3201/doi:10.1186/s12916-021-02139-6 (2021).

12. Drolet, M. *et al.* Time trends in social contacts before and during the COVID-19 pandemic: the CONNECT study. *BMC Public Heal.* **22**, 1032, DOI: https://dx.doi.org/10.1186/s12889-022-13402-7 (2022).

13. Vos, E. R. A. *et al.* Associations Between Measures of Social Distancing and Severe Acute Respiratory Syndrome Coronavirus 2 Seropositivity: A Nationwide Population-based Study in the Netherlands. *Clin Infect Dis* **73**, 2318–2321, DOI: http://dx.doi.org/10.1093/cid/ciab264 (2021).

14. Hooiveld, M., Heins, M., Hendriksen, J. & Korevaar, J. Aantal mensen met medische indicatie voor vaccinatie tegen COVID-19 (Number of people with medical indication for vaccination against COVID-19). https://www.nivel.nl/sites/default/files/bestanden/1003980.pdf (2021). [Online; accessed 8 August 2022].

15. RIVM. Cumulatieve opkomst tenminste één COVID-19 vaccinatie naar geboortejaar en week (Cumulative coverage of at least one COVID-19 vaccination per birth year and calendar week). https://www.rivm.nl/covid-19-vaccinatie/cijfers-vaccinatieprogramma (2022). [Online; accessed 14 April 2022].

16. Jarvis, C. I. *et al.* Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC Med* **18**, 124, DOI: http://dx.doi.org/10.1186/s12916-020-01597-8 (2020).

17. Gimma, A. *et al.* Changes in social contacts in England during the COVID-19 pandemic between March 2020 and March 2021 as measured by the CoMix survey: A repeated cross-sectional study. *PLoS Med* **19**, e1003907, DOI: https://doi.org/10.1371/journal.pmed.1003907 (2022).

18. Liljeros, F., Edling, C. R., Amaral, L. A., Stanley, H. E. & Aberg, Y. The web of human sexual contacts. *Nature* **411**, 907–908, DOI: http://dx.doi.org/10.1038/35082140 (2001).

19. Tagore, S. *Epidemic Models: Their Spread, Analysis and Invasions in Scale-Free Networks*, 1–25 (Springer International Publishing, Cham, 2015).

20. Statistics Netherlands. Bevolking; geslacht, leeftijd en burgerlijke staat, 1 januari 2021. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/7461bev/table (2021). [Online; accessed 1 July 2021].

21. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2022).

22. Wood, S. *Generalized Additive Models: An Introduction with R* (Chapman and Hall/CRC, 2017), 2 edn.

23. Bogaardt, L. & Backer, J. A. CoMix social contact data (The Netherlands). http://zenodourl/ (2022). [Online; accessed 19 July 2022].

24. Wong, K. L. M. *et al.* Social contact patterns during the COVID-19 pandemic in 21 European countries – evidence from a two-year study. *medRxiv* DOI: http://dx.doi.org/10.1101/2022.07.25.22277998 (2022).

25. Coletti, P. *et al.* CoMix: comparing mixing patterns in the Belgian population during and after lockdown. *Sci Rep* **10**, 21885, DOI: https://dx.doi.org/10.1038/s41598-020-78540-7 (2020).

26. Jarvis, C. I. *et al.* The impact of local and national restrictions in response to COVID-19 on social contacts in England: a longitudinal natural experiment. *BMC Med* **19**, 52, DOI: http://dx.doi.org/10.1186/s12916-021-01924-7 (2021).

27. Coletti, P. *et al.* A data-driven metapopulation model for the Belgian COVID-19 epidemic: assessing the impact of lockdown and exit strategies. *BMC Infect Dis* **21**, 503, DOI: https://dx.doi.org/10.1186/s12879-021-06092-w (2021).

28. Munday, J. D. *et al.* Estimating the impact of reopening schools on the reproduction number of SARS-CoV-2 in England, using weekly contact survey data. *BMC Med* **19**, 233, DOI: http://dx.doi.org/10.1186/s12916-021-02107-0 (2021).

29. Franco, N. *et al.* Inferring age-specific differences in susceptibility to and infectiousness upon SARS-CoV-2 infection based on Belgian social contact data. *PLoS Comput. Biol* **18**, e1009965, DOI: https://doi.org/10.1371/journal.pcbi.1009965 (2022).

30. Wambua, J. *et al.* The influence of risk perceptions on close contact frequency during the SARS-CoV-2 pandemic. *Sci Rep* **12**, 5192, DOI: http://dx.doi.org/10.1038/s41598-022-09037-8 (2022).

31. Wong, K. L. M. *et al.* Pregnancy during COVID-19: social contact patterns and vaccine coverage of pregnant women from CoMix in 19 European countries. *medRxiv* DOI: http://dx.doi.org/10.1101/2022.06.01.22275775 (2022).

Acknowledgements

The authors would like to thank all participants of the CoMix survey. This project has received funding from the European Union's Horizon 2020 research and innovation programme - project EpiPose (grant agreement number 101003688).

ook Netherlands Ministry of Health, Welfare and Sport noemen?

Author contributions statement

J.B. managed the data collection, L.B. curated the data, J.B. and L.B. analysed the data, [more authors?], J.W. set up the study. All authors reviewed the manuscript.

Additional information

Competing interests The authors declare no competing interests.
Ethics statement ?

year		2020	2020	2020	2021	2021	2021
population		study	study	general[a]	study	study	general
		n	%	%	n	%	%
total		1659			2514		
age group	0-11				240	10	12
age group	12-17				290	12	7
age group	18-24	84	5	11	511	20	9
age group	25-34	236	14	16	378	15	13
age group	35-44	257	15	15	254	10	12
age group	45-54	337	20	18	271	11	14
age group	55-64	312	19	17	230	9	14
age group	65+	433	26	24	340	14	20
gender	Female	814	49	51	1383	55	50
gender	Male	845	51	49	1131	45	50
high risk	mixed	359	[b]		385	[b]	
high risk	no	744	58	78	1720	81	81
high risk	yes	540	42	22	400	19	19
high risk	NA	16	[b]		9	[b]	
household size	1	409	25	17[c]	573	23	17
household size	2	716	43	31	766	30	31
household size	3	241	15	17	444	18	17
household size	4	203	12	23	496	20	23
household size	5+	90	5	12	235	9	12

Table 1. Number of participants in the 2020 and 2021 series by participant characteristics.
[a] Characteristics for 2020 general population based on 18+ population to match study population.
[b] Percentage of risk status in study population only calculated for participant with unambiguous risk status to allow comparison with general population.
[c] Expected household size distribution per Dutch citizen, based on Dutch households consisting of one (38%), two (33%), three (12%), four (12%) and five or more (5%) persons[8]

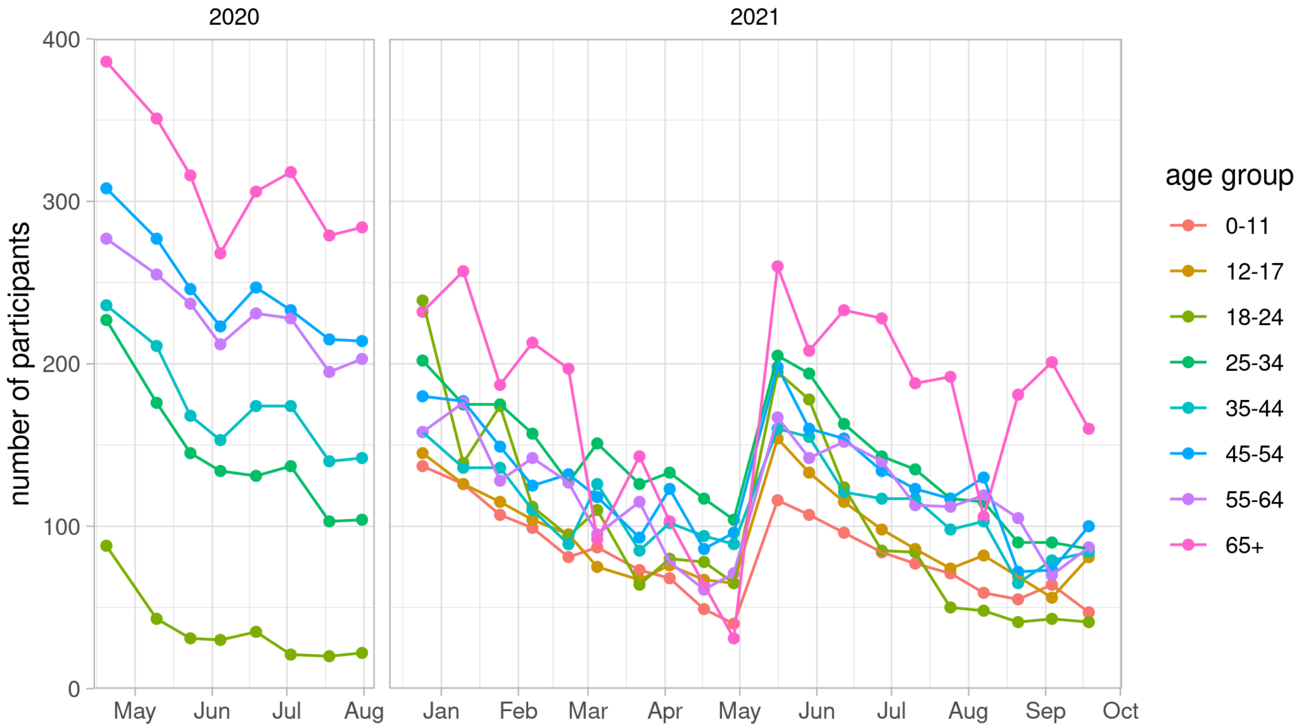

Figure 1. Number of participants included per survey round in the 2020 and 2021 series, by age group. After 10 survey rounds in May 2021, the study population was supplemented to meet the target numbers of the first survey round again.

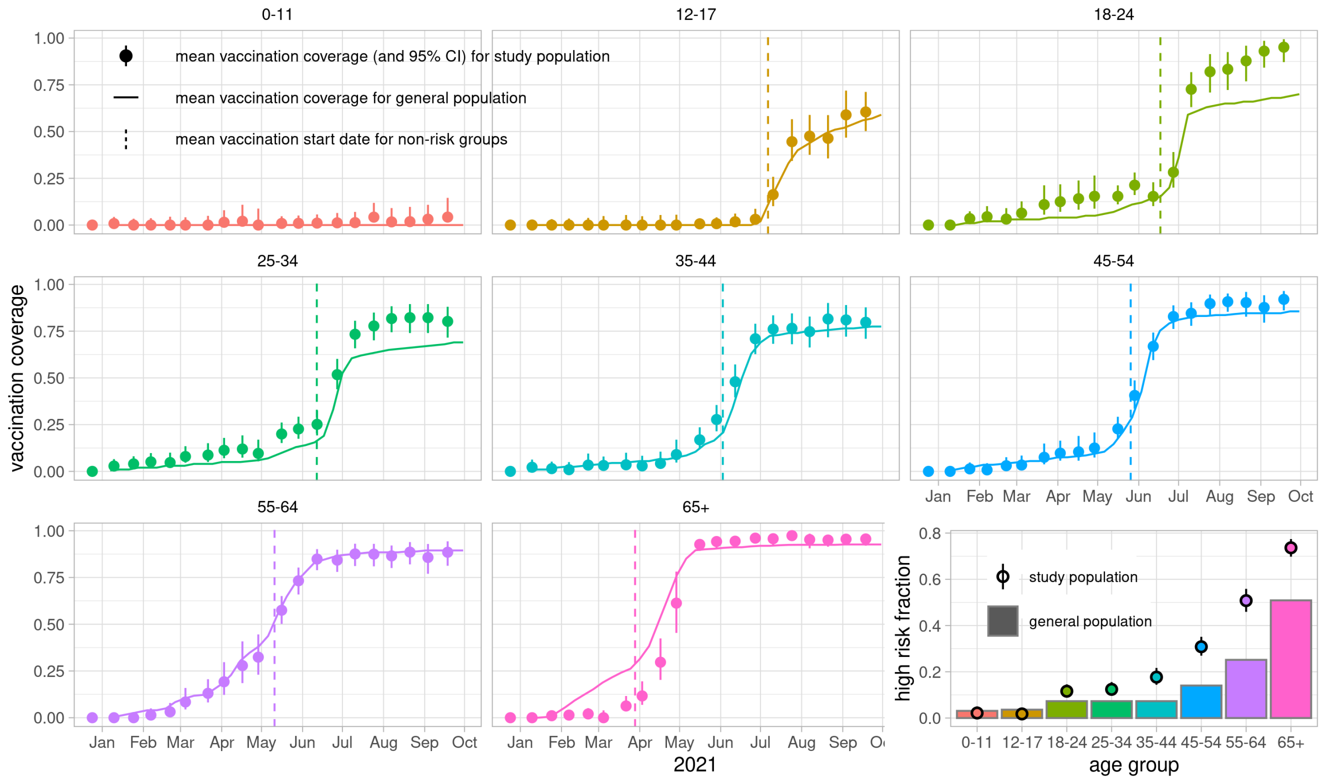

Figure 2. Vaccination and risk status of study and general population. First eight panels show the vaccination coverage in 2021 per age group of the study population (points with 95% confidence interval) and of the general population[15] (line). The dashed line denotes the average date when non-risk groups were invited for vaccination. The dashed line for the 55-64 age group is in fact the average for 55-59 year olds, as 60-64 year olds were invited by their general practitioner. The last panel shows the fraction of high risk participants of the 2020 and 2021 series combined (points with 95% confidence interval), and of the general population[14] (bars).

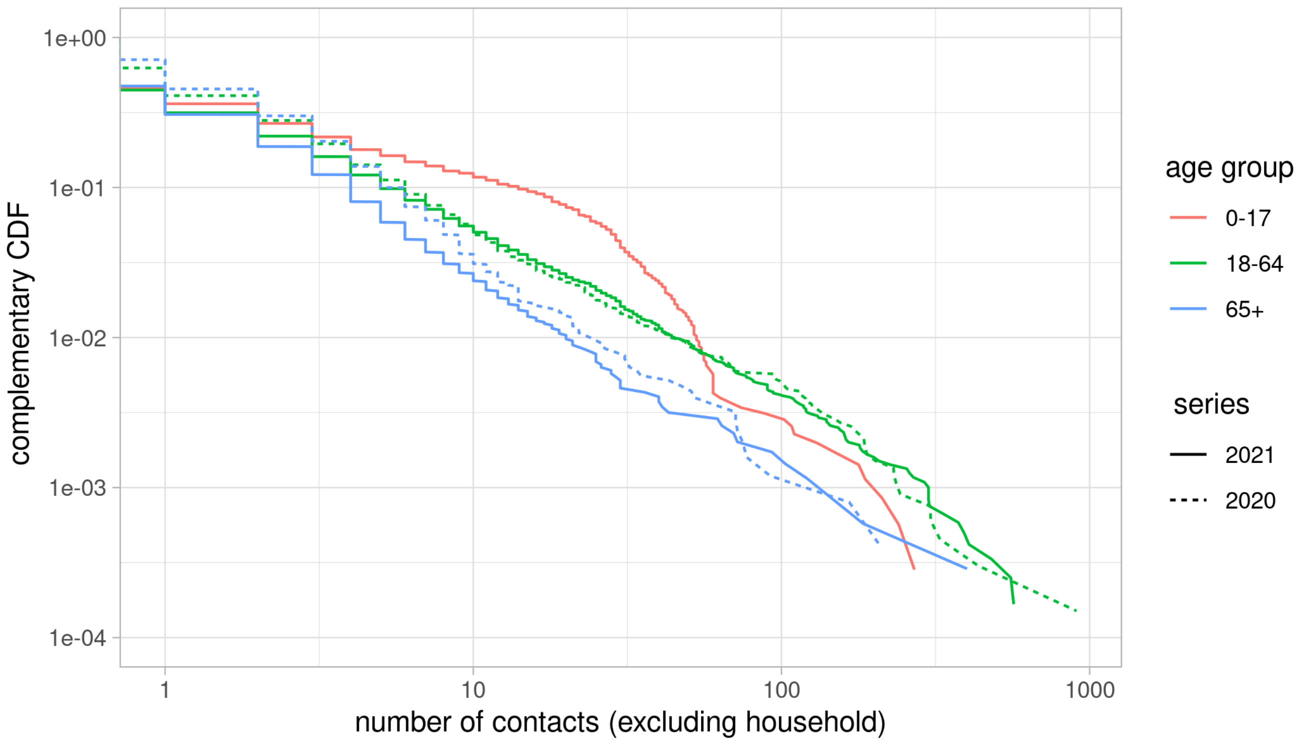

Figure 3. Distribution of number of contacts per participant per round excluding household contacts, by series (line type) and age group (color), with complementary cumulative distribution function on y-axis.

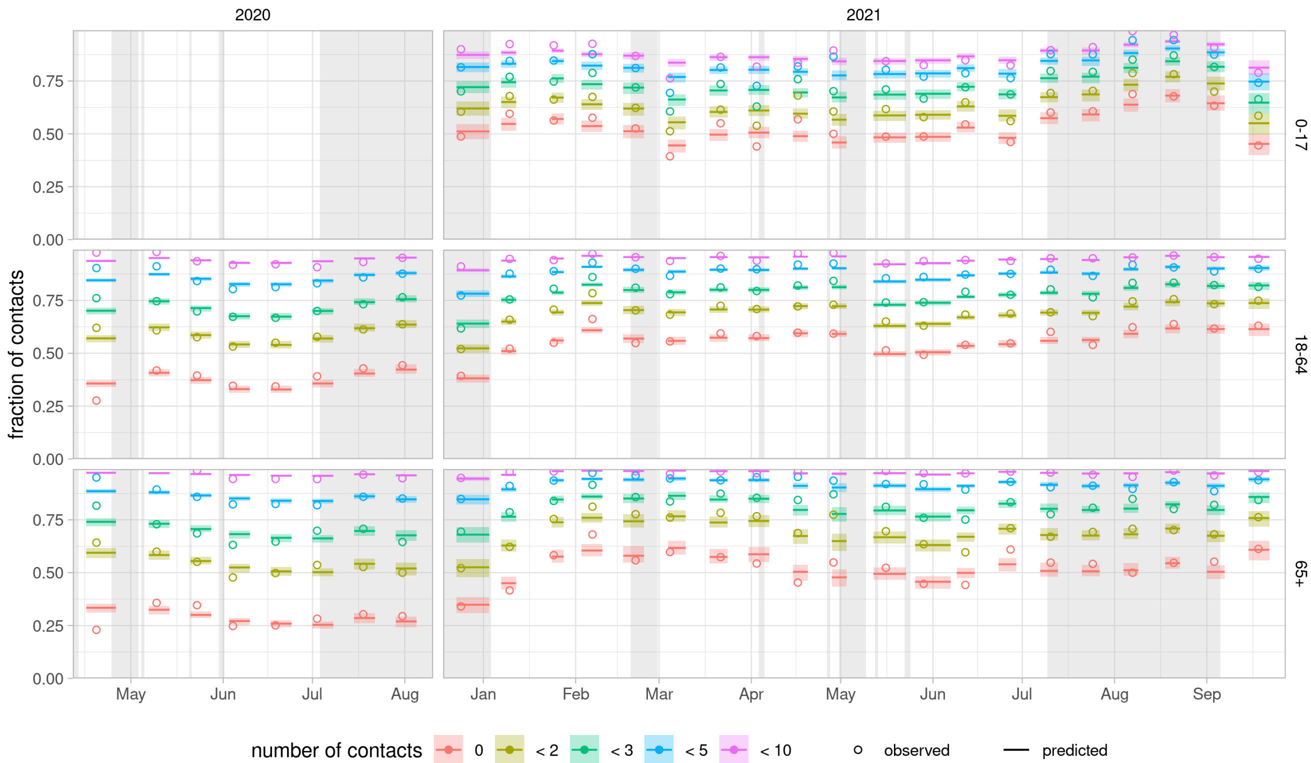

Figure 4. Predicted and observed activity levels over time by series (columns) and age group (rows). Activity levels are shown as the fraction of participants that report less than a certain number of contacts outside the household. With five limits (0, <2, <3, <5 and <10) six activity levels are defined, e.g. the fraction between the limits of <3 and <5 is the activity level that represents 3 or 4 contacts per participant. The model predictions per round are shown by the median (lines) and 95% prediction interval (shaded), from the first to last participation date. The observed activity levels per round (open circles) are placed at the mean participation date. Holidays and school holiday periods are shaded in grey.

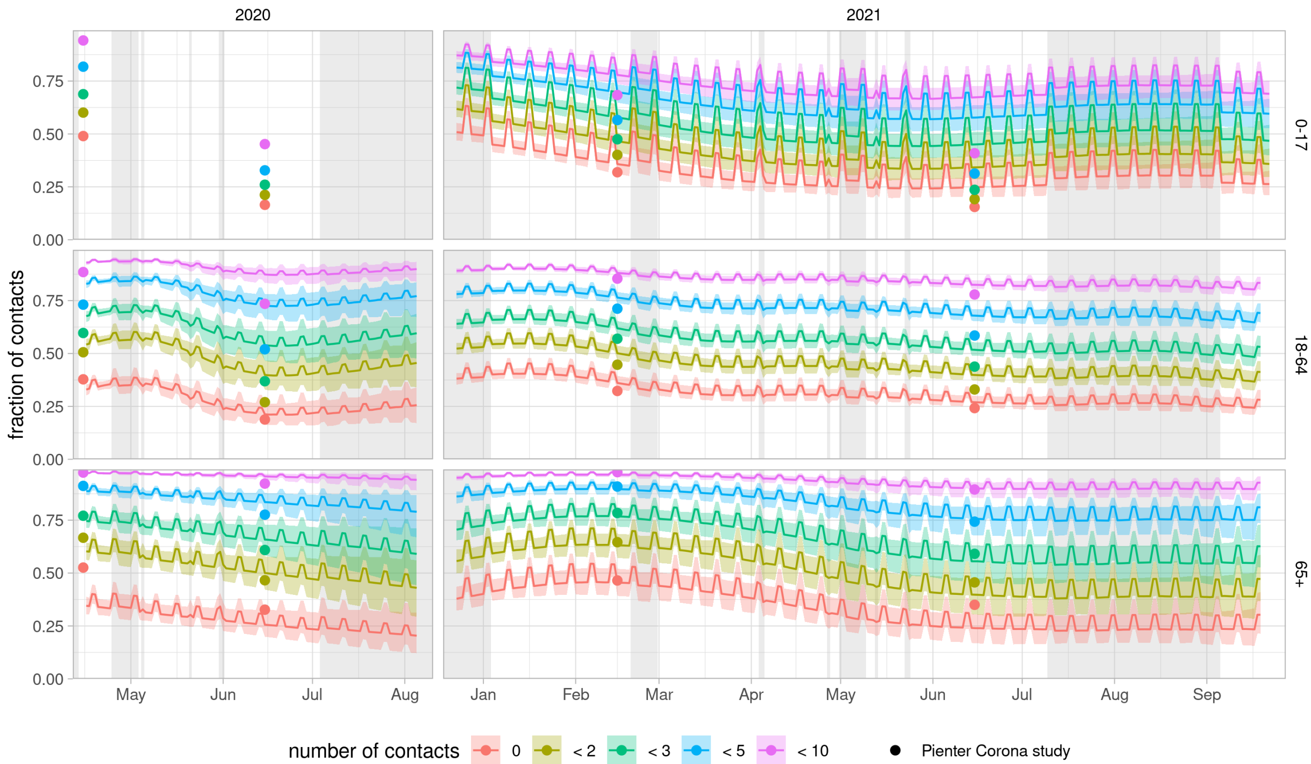

Figure 5. Predicted activity levels over time by series (columns) and age group (rows) for the general population. Activity levels are shown as the fraction of the population that has less than a certain number of contacts outside the household. With five limits (0, <2, <3, <5 and <10) six activity levels are defined, e.g. the fraction between the limits of <3 and <5 is the activity level that represents 3 or 4 contacts per person. The model predictions are shown by the median (lines) and 95% prediction interval (shaded). The activity levels observed in the contact survey of the independent PiCo study are shown for comparison per round at the mean participation date (points). Holidays and school holiday periods are shaded in grey.

Supplementary information

year	round	0-11	12-17	18-24	25-34	35-44	45-54	55-64	65+	total
2020	target	0	0	162	235	225	277	249	352	1500
2020	1	0	0	88	227	236	308	277	386	1522
2020	2	0	0	43	176	211	277	255	351	1313
2020	3	0	0	31	145	168	246	237	316	1143
2020	4	0	0	30	134	153	223	212	268	1020
2020	5	0	0	35	131	174	247	231	306	1124
2020	6	0	0	21	137	174	233	228	318	1111
2020	7	0	0	20	103	140	215	195	279	952
2020	8	0	0	22	104	142	214	203	284	969
2021	target	150	150	248	206	163	184	161	238	1500
2021	1	137	145	239	202	158	180	158	232	1451
2021	2	126	126	139	175	136	177	176	257	1312
2021	3	107	115	174	175	136	149	128	187	1171
2021	4	99	104	112	157	110	125	142	213	1062
2021	5	81	95	94	127	89	132	127	197	942
2021	6	87	75	110	151	126	118	95	92	854
2021	7	73	67	64	126	85	93	115	143	766
2021	8	68	76	80	133	102	123	78	103	763
2021	9	49	67	78	117	94	86	61	64	616
2021	10	40	65	65	104	89	96	71	31	561
2021	11	116	154	195	205	160	198	167	260	1455
2021	12	107	133	178	194	155	160	142	208	1277
2021	13	96	115	124	163	121	154	152	233	1158
2021	14	84	98	85	143	117	134	140	228	1029
2021	15	77	86	84	135	117	123	113	188	923
2021	16	71	74	50	117	98	117	112	192	831
2021	17	59	82	48	115	103	130	119	106	762
2021	18	55	69	41	90	65	72	105	181	678
2021	19	64	56	43	90	79	73	70	201	676
2021	20	47	81	41	86	84	100	87	160	686

Table S1. Number of participants per survey round of study series in 2020 and 2021 in eight age groups. The target number of participants was to be achieved in round 1 in 2020, and in rounds 1 and 11 in 2021.

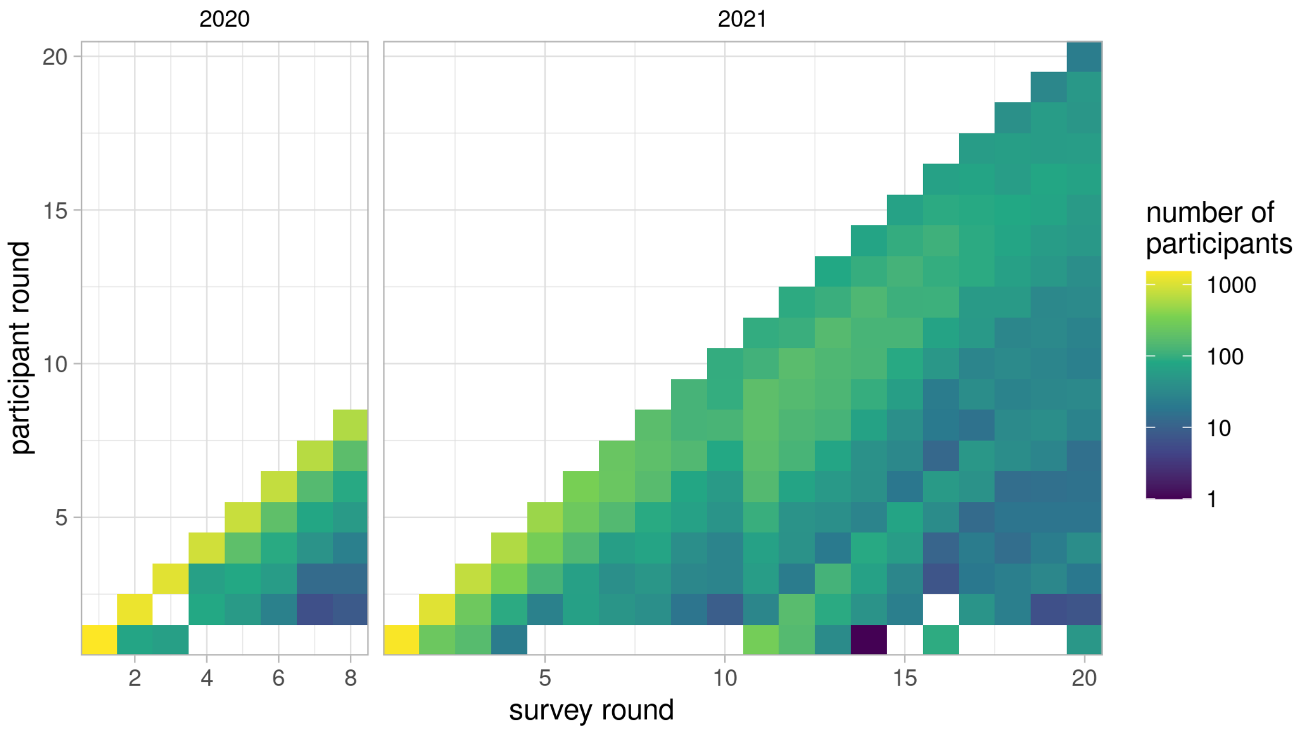

Figure S1. Number of participants by survey round and participant round (i.e. the n^{th} time a participant participated). Series 2020 (left) consisted of 8 rounds, series 2021 (right) consisted of 20 rounds. Most participants participated for the first time the start of the series, but some also started during the study period. The panel was supplemented to the target population in round 11 of series 2021.

series	age_group	Variable	Estimate	Std. Error	t value	Pr(>|t|)
2020	18-64	(Intercept)	-0.24	0.06	-4.15	0.00
2020	18-64	part_vaccTRUE	0.00	0.00		
2020	18-64	part_elevated_riskTRUE	-0.03	0.08	-0.38	0.70
2020	18-64	weekendTRUE	-0.21	0.06	-3.65	0.00
2020	18-64	holidayTRUE	0.07	0.10	0.69	0.49
2020	18-64	Deviance explained	0.16			
2020	65+	(Intercept)	0.03	0.13	0.27	0.79
2020	65+	part_vaccTRUE	0.00	0.00		
2020	65+	part_elevated_riskTRUE	0.25	0.12	2.05	0.04
2020	65+	weekendTRUE	-0.32	0.08	-3.89	0.00
2020	65+	holidayTRUE	-0.08	0.11	-0.70	0.48
2020	65+	Deviance explained	0.14			
2021	0-17	(Intercept)	-0.82	0.08	-10.02	0.00
2021	0-17	part_vaccTRUE	0.12	0.19	0.63	0.53
2021	0-17	part_elevated_riskTRUE	-0.14	0.28	-0.52	0.60
2021	0-17	weekendTRUE	-0.67	0.08	-8.35	0.00
2021	0-17	holidayTRUE	-0.20	0.10	-2.02	0.04
2021	0-17	Deviance explained	0.17			
2021	18-64	(Intercept)	-1.04	0.06	-17.68	0.00
2021	18-64	part_vaccTRUE	0.09	0.07	1.24	0.22
2021	18-64	part_elevated_riskTRUE	-0.09	0.08	-1.20	0.23
2021	18-64	weekendTRUE	-0.27	0.05	-5.95	0.00
2021	18-64	holidayTRUE	0.06	0.06	0.94	0.35
2021	18-64	Deviance explained	0.21			
2021	65+	(Intercept)	-0.86	0.15	-5.74	0.00
2021	65+	part_vaccTRUE	0.07	0.19	0.38	0.70
2021	65+	part_elevated_riskTRUE	-0.09	0.11	-0.81	0.42
2021	65+	weekendTRUE	-0.45	0.08	-5.69	0.00
2021	65+	holidayTRUE	0.04	0.10	0.40	0.69
2021	65+	Deviance explained	0.16			

Table S2. Results for fixed effects of generalised additive model, by series (2020, 2021) and age group (0-17, 16-64, 65+). Included fixed effects are participant vaccination status (part_vacc), participant risk status (part_elevated_risk), weekend and holiday. Holidays include general holidays and school holidays. The last row of each section refers to the explained deviance.

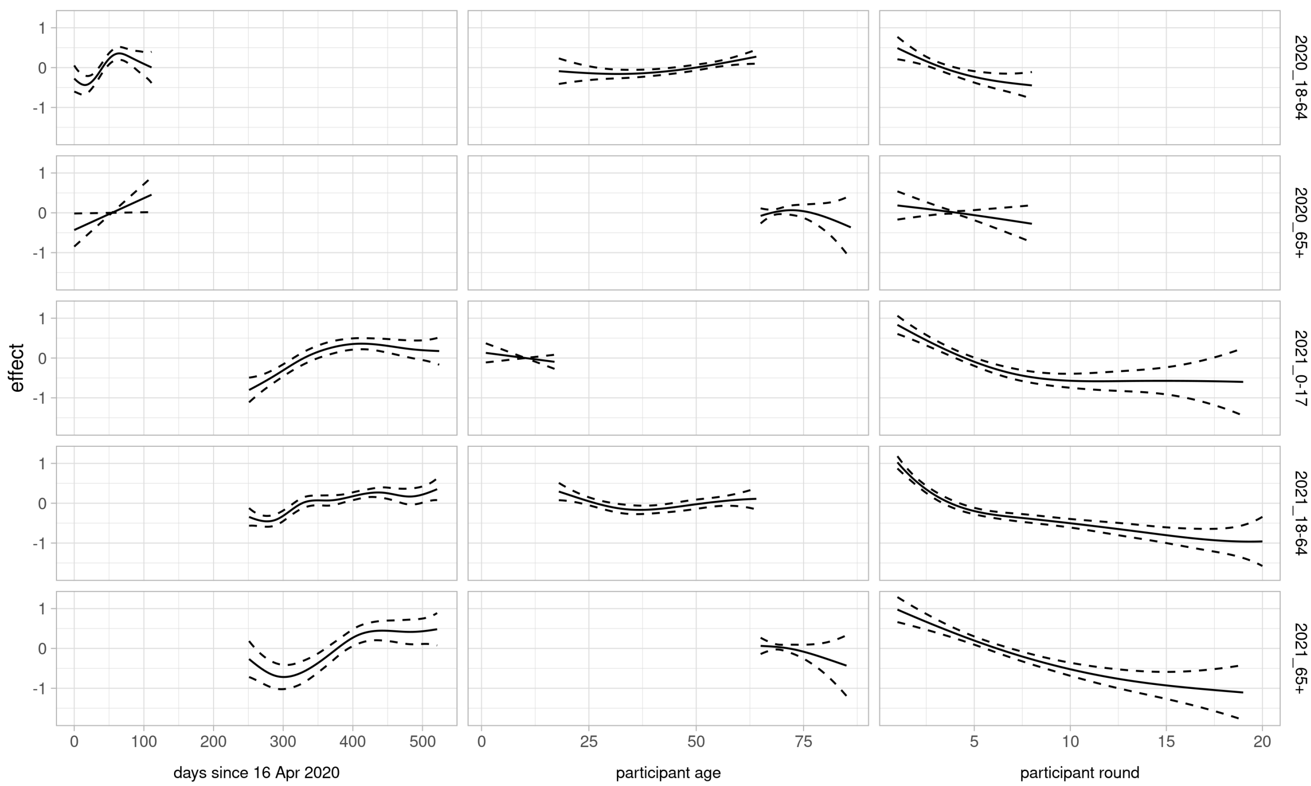

Figure S2. Results for non-fixed effects of generalised additive model, by series (2020, 2021) and age group (0-17, 16-64, 65+). Calendar time (expressed as days since first survey), participant age, and participant round (i.e. the n^{th} time a participant participated) are included as cubic splines. Shown are the fitted splines (solid lines) $\pm$ the standard error (dashed lines).