


BMJ Open Differential impact of physical distancing strategies on social contacts relevant for the spread of SARS-CoV-2: evidence from a cross-national online survey, March–April 2020

Emanuele Del Fava ¹, Jorge Cimentada,¹ Daniela Perrotta,¹ André Grow,¹ Francesco Rampazzo,² Sofia Gil-Clavel,¹ Emilio Zagheni¹

To cite: Del Fava E, Cimentada J, Perrotta D, *et al*. Differential impact of physical distancing strategies on social contacts relevant for the spread of SARS-CoV-2: evidence from a cross-national online survey, March–April 2020. *BMJ Open* 2021;**11**:e050651. doi:10.1136/bmjopen-2021-050651

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-050651>).

Received 25 February 2021
Accepted 07 October 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Laboratory of Digital and Computational Demography, Max Planck Institute for Demographic Research, Rostock, Germany

²Said Business School, Leverhulme Centre for Demographic Science, and Nuffield College, University of Oxford, Oxford, UK

Correspondence to

Dr Emanuele Del Fava;
delfava@demogr.mpg.de

ABSTRACT

Objectives We investigate changes in social contact patterns following the gradual introduction of non-pharmaceutical interventions and their implications for infection transmission in the early phase of the pandemic.

Design, setting and participants We conducted an online survey based on targeted Facebook advertising campaigns across eight countries (Belgium, France, Germany, Italy, the Netherlands, Spain, UK and USA), achieving a sample of 51 233 questionnaires in the period 13 March–12 April 2020. Poststratification weights based on census information were produced to correct for selection bias.

Outcome measures Participants provided data on social contact numbers, adoption of protective behaviours and perceived level of threat. These data were combined to derive a weekly index of infection transmission, the net reproduction number R_t .

Results Evidence from the USA and UK showed that the number of daily contacts mainly decreased after governments issued the first physical distancing guidelines. In mid-April, daily social contact numbers had decreased between 61% in Germany and 87% in Italy with respect to pre-COVID-19 levels, mostly due to a contraction in contacts outside the home. Such reductions, which were uniform across age groups, were compatible with an R_t equal or smaller than one in all countries, except Germany. This indicates lower levels of infection transmission, especially in a period of gradual increase in the adoption rate of the face mask outside the home.

Conclusions We provided a comparable set of statistics on social contact patterns during the COVID-19 pandemic for eight high-income countries, disaggregated by week and other demographic factors, which could be leveraged by the scientific community for developing more realistic epidemic models of COVID-19.

INTRODUCTION

The COVID-19 epidemic has shown the importance of implementing non-pharmaceutical interventions (NPIs), like physical distancing, to contain the spread of infectious diseases

Strengths and limitations of this study

- ⇒ The use of targeted Facebook advertisements to rapidly recruit survey respondents made it possible to collect data on health behaviours in a timely manner and at a lower cost than expected with traditional surveys during public health emergencies.
- ⇒ The use of the same survey instrument in a cross-country perspective enabled us to collect daily data on social contact numbers and adoption of protective behaviours in a consistent manner in different epidemic contexts.
- ⇒ The stratification of the Facebook advertising campaigns by sex, age group and region of residence, along with the construction of poststratification weights based on the population census data of the same variables, helped to reduce the non-representativeness bias in participant recruitment.
- ⇒ Despite the use of poststratification weights, we could not entirely adjust for different behaviours or social inequalities among Facebook users, since only a limited number of demographic variables could be included in the ads stratification.
- ⇒ Although data on the characteristics of social contacts (eg, their age) were not collected, we proposed a method to derive age-specific contact matrices using predictions of the number of social contacts by age group and estimate the net reproduction number R_t .

and to avoid overburdening healthcare systems. As the strain of the virus that causes the disease (SARS-CoV-2) spreads from person to person through infected secretions, like saliva or respiratory droplets from the nose or mouth, disease transmission mainly occurs via close social contacts, generally when susceptible individuals breathe in droplets from infected ones, although there is increasing evidence for airborne transmission through aerosols.¹

In the early phase of the pandemic, in the absence of a vaccine and effective antiviral treatments, the public health measures introduced to contain the disease were largely NPIs aimed at reducing the number of social contacts, for example, travel restrictions, school closures, the cancellation of large gatherings, the promotion of physical distancing practices and, in some countries, nationwide lockdowns. Quantifying changes in social contact patterns in response to these interventions is key to: (1) improve our understanding of the determinants of disease transmission, (2) enable researchers to design more realistic epidemic models and, ultimately (3) assess the effectiveness of public health measures designed to contain the disease and prevent new waves of infection.^{2 3}

In this study, we leverage new opportunities for data collection, enabled by the digitalisation of our lives, to assess changes in social contact patterns across countries and over time and to examine their implications for the spread of SARS-CoV-2. We estimate the net reproduction number, R_t , an index that directly depends on social contact patterns. This is based on a cross-national online survey, the COVID-19 Health Behaviour Survey (CHBS), for which we recruited participants via targeted Facebook advertisements stratified by sex, age and region of residence, during the early phase of the pandemic.^{4 5} One goal of the survey was to evaluate the extent to which social contact numbers changed across countries and over time, especially in comparison with those observed prior to the pandemic, using comparable metrics.^{3 6} Following the literature on the relationship between social contacts and their impact on the reproduction number for close-contact infectious diseases,^{2 7 8} and consistently with other social contact surveys conducted in Asia,⁹ Europe,¹⁰⁻¹⁴ Africa¹⁵ and North America^{16 17} to assess changes in contact patterns and their impact on infection transmission during the early phase of the pandemic, we quantified participants' social contacts, adoption of protective behaviour and perceived threat and estimated the relative impact of changes in social contact numbers on the net reproduction rate R_t , taking also the effect of other behavioural changes into account.

METHODS

Study design

We designed the CHBS as a cross-country, cross-sectional and voluntary opt-in online survey to collect key information on people's health and behaviour in eight high-income countries: Belgium, France, Germany, Italy, the Netherlands, Spain, the UK and the USA. The first wave of data collection ran from 13 March to 12 April 2020, a key period during which the global pandemic was well underway, even though at different stages across countries.⁴ To quickly reach large numbers of survey participants across several countries, recruitment occurred via targeted advertisements implemented through the Facebook Ads Manager, a tool usually employed by advertisers to create marketing campaigns for their products. The

ads were stratified by sex, age group and region of residence, in order to ensure that a minimum number of respondents could be reached in each stratum.^{18 19} Post-stratification weights by sex, age group and region of residence were used in all analyses. These were obtained by dividing the true population proportion in each stratum (based on nationally representative census data^{20 21}) by the sample proportion from the same stratum. The effect of applying these weights to the sample is shown in the online supplemental tables S1-S8. Participation was anonymous, voluntary and open to people aged 18 years or older who gave their informed consent. This study was funded with support from the Max Planck Institute for Demographic Research, which is part of the Max Planck Society.

We asked respondents to report the number of social contacts (defined as any interaction involving either physical contact or a conversation of three or more words in the physical presence of another person) on the day prior to the survey in four different settings, ie, at home, at school, at work and in the general community (eg, during commuting or leisure activities) (online supplemental figure S1). We calculated the overall number of daily contacts per person as the sum of contacts in the four settings, after removing participants who reported more contacts than the 90% quantile of the contact distribution for the respective setting. This choice was motivated by the need to remove low-quality responses with implausibly high contact numbers. Our results were robust to the choice of the threshold, as we also considered a fixed cut-off point equal to 29 contacts, which has been used in a seminal social contact study³ and which resulted in slightly higher average contact numbers (online supplemental tables S9-S16). We also asked participants to report, among others, their perceived level of threat (Q13: 'What level of threat do you think the coronavirus poses to... you personally? ... your family?') and their adoption of any protective behaviours that might have an impact on disease transmissibility²² (Q18: 'Which of the following actions, if any, have you already taken to protect yourself from the coronavirus?... washed hands more often?... worn a face mask?... avoided social activities (eg, meeting friends)?').

Data analysis

We modelled the number of daily contacts reported by participants using negative binomial regression to account for possible overdispersion in the contact data.²³ The models were fitted independently by country, setting of contact and contact threshold, and adjusted for both sociodemographic characteristics and behavioural responses of respondents. The first group accounted for sex, age group, region of residence, education level, working status, household size (continuous and centred around the mean), day type (whether weekday or weekend) and being foreign born. The second group included a dummy variable for each protective behaviour and a threat perception index, built by averaging the

threat to oneself (much higher for older adults than for the 18–24 years age group⁴) and to the family. For each fitted model, we predicted the mean number of contacts, overall and by age group, adjusted at the mean values of the three behavioural variables and the feeling of threat.

Next, we assessed the implications of the changes in social contact numbers for the net reproduction number (R_t , ie, the average number of new infections generated by an infectious person at time t in a population, either partially or fully susceptible, considering the current interventions and the potential spontaneous behavioural change in response to the risk of infection^{24 25}), estimated as the dominant eigenvalue of the next generation matrix N , which provides information on the numbers of newly infected individuals by age group at a given time.²⁶ Under the ‘social contact hypothesis’,^{2 27} we have the relationship $N = DqC$, where C denotes the matrix containing the average number of contacts between age groups, q is the infection transmissibility and D is the length of the infectious period. We used the proportional relationship between N and C , ie, $N = \frac{R_t}{\rho(C)} C$, where $\rho(C)$ denotes the dominant eigenvalue of C , to assess changes in R_t due to changes in social contact numbers.^{11 14 17} We compared our contact data with those collected prior to the epidemic, either from the POLYMOD study (Belgium, Germany, Italy, the Netherlands and the UK),^{3 28} or from the Comes-F study (France),^{6 29} and we derived age-specific social contact matrices C compatible with the demographic structure of each country at the beginning of 2020.³⁰

We assumed that: (1) the basic reproduction number R_0 of the infection was consistent with social mixing as measured prior to COVID-19, under the assumption of absence of immunity in the population and of any behavioural change and interventions; (2) the parameters D and q did not change over time; (3) the initial contact matrices of Italy and the UK could be a reasonable replacement of contact patterns of Spain and the USA, because of cultural similarity, after projecting them on the respective demographic structures.

We obtained weekly contact matrices C_w^{cov} by multiplying the normalised pre-COVID-19 contact matrices C^{pre} with the predicted contact numbers by age group and week (see more details on matrix construction in the online supplemental material). We finally computed the weekly values of R_t^w as $R_0 f_R^w$, where the weekly reduction factor f_R^w was given by the ratio between the dominant eigenvalue of the weekly contact matrix and that of the pre-COVID-19 matrix, that is, $f_R^w = \rho(C_w^{cov}) / \rho(C^{pre})$.

Our estimates of R_t^w accounted for three sources of uncertainty, namely (1) the baseline basic reproduction number R_0 which, according to a meta-analysis of the literature, was assumed to be normally distributed with mean 2.6 and SD 0.54, resulting in the 95% CI (1.54 to 3.66)^{14 17}; (2) the weekly predictions of the daily overall number of contacts, assumed to be normally distributed with mean and SD coming from our predictions; and (3)

the pre-COVID-19 contact matrices, which were derived by applying non-parametric bootstrap to the original data. For each source of uncertainty, 5000 replicates were taken.

All data preparation and visualisation were performed in R software,³¹ while the negative binomial regression models were fitted in Stata V.16.

Patient and public involvement

Respondents were not involved in the design of the study. However, as the survey was disseminated through Facebook ads, users could comment on the ads, leave posts on the study page and review the page itself. To maintain trust with both current and prospective participants, we engaged with user comments and provided additional information if requested.⁵

RESULTS

From an initial sample of 60 048 respondents, we obtained a final sample of 51 233 individuals who: (1) completed the questionnaire between 13 March and 12 April 2020, (2) provided the information on sex, age group and region of residence, necessary for the computation of the poststratification weights and (3) reported knowing at least something about the COVID-19 outbreak at the time of their participation.

Individual social contacts decreased during the study period in all surveyed countries, plateauing at the lowest levels between calendar week 2020–13 (23–29 March) and calendar week 2020–14 (30 March–5 April) under both contact number thresholds (figure 1 and online

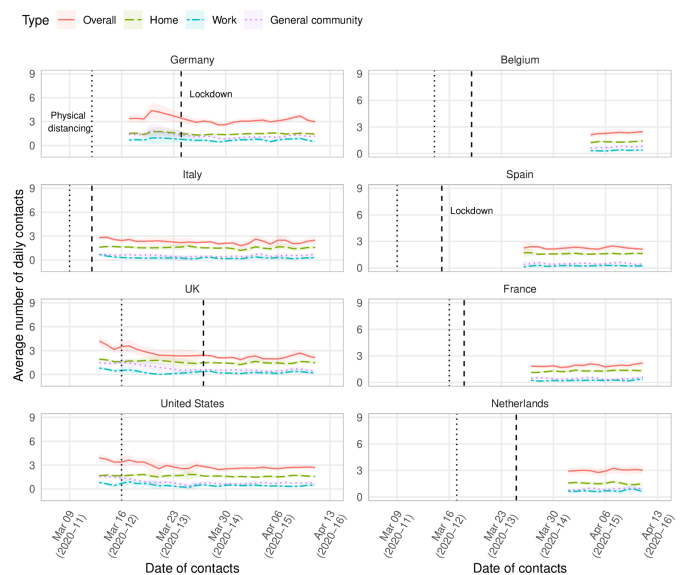


Figure 1 Mean overall number of daily social contacts (with 95% CI) smoothed by a simple 2-day moving average, by country and study day, March–April 2020. The dotted line corresponds to the date in which the physical distancing guidelines were introduced at the national level; the dashed line corresponds to the date in which the lockdown, regardless of being full or partial, was ordered.

supplemental figure S2). The decrease was such that around 30% of participants in each country reported having less than one contact per day during the study period (online supplemental figure S3).

Results for the USA and the UK showed that the overall numbers of daily social contacts declined considerably after governments issued physical distancing guidelines, mainly occurring during week 2020–11 (9–15 March), rather than after more stringent lockdown measures were announced. Comparing trends before and after the announcement of physical distancing measures, both occurring between week 2020–11 and week 2020–12,^{32 33} we estimated that, under the 90% quantile threshold, the average number of daily contacts decreased from 3.67 (8.81 under the ≤ 29 threshold) in week 2020–11 to 2.82 (4.81) in week 2020–13 in the USA, and from 3.43 (8.77) to 2.40 (3.85) in the UK. Then, in week 2020–14, after the implementation of lockdown measures, contact numbers decreased even further in the UK—to 2.17 (3.07)—and to a smaller extent in the USA—to 2.67 (4.42)—where contact numbers had essentially plateaued since the beginning of week 2020–13 (table 1 and online supplemental table S17). Also in the other countries, where physical distancing guidelines were introduced before the beginning of our study, we found that the average number of contacts had declined to their lowest levels in week 2020–14, ranging from 2.94 and 2.97 in Germany and the Netherlands, respectively, to 1.93 and 2.17 in France and Italy, respectively (table 1). These results show how reductions in social contacts differed between those countries that eventually applied a full lockdown (people were required to stay at home and not allowed to go out if not necessary), such as Belgium, France, Italy, Spain and the UK, and those that opted for a partial lockdown (a stricter version of the physical distancing guidelines was imposed, but people had the freedom to go out), such as Germany and the Netherlands.

The comparison between the pre-pandemic contact data and those in week 2020–15 showed that the largest reduction occurred in Italy, at 87% under the 90% quantile threshold (81% under the ≤ 29 threshold), though substantial declines also occurred in the other countries, ranging from 61% (32%) in Germany to a level between 78% and 82% (63% and 70%) in Belgium, France, the Netherlands and the UK. This variation was mostly due to a decrease of contact numbers in the general community and, to a lower extent, at work. Indeed, while the former setting was the main contributor to the overall number of daily contacts in the pre-COVID-19 period, at the end of the study period (week 2020–15) and under the 90% quantile threshold, home contacts had become the largest component, representing up to 50% (in Germany and the Netherlands) and around 70% (in Italy, Spain and the UK) of all the contacts (online supplemental figure S4); this result was slightly attenuated under the ≤ 29 threshold, as in Germany and the Netherlands, both the work and the general community represented more important contributors than the household (online

supplemental figure S5), likely due to the smaller average household size in these two countries (online supplemental figure S6).

The analysis of the determinants of the contact numbers showed a pattern of cross-country heterogeneity, with a dependence on sociodemographic variables such as the household size and being in paid work—all positively associated—and, to a lower extent, being male, and reporting data during the weekend (online supplemental figures S7 and S8). In terms of behavioural factors, at the introduction of the physical distancing guidelines in week 2020–11 followed an increase in the rate of adoption of protective behaviour, especially wearing a face mask and avoiding social activities (figure 2). This is particularly evident in the UK, the USA and Germany, where we included participants in a period of increasing restrictions. What strikes the most is the increased use of face coverings in all countries, except for Belgium and the Netherlands (stable at 10%).⁴ The largest increases were reported in Italy (from 40% in week 2020–11 to 80% in week 2020–15) and in the USA (from 10% in week 2020–11 to 60% in week 2020–15), while in the UK, it reached 20% by the same date. Between 80% and 90% of participants in every country avoided social activities, with evidence of a sharp increase over time between week 2020–11 and week 2020–13 in the UK and, to a lesser extent, the USA. Little to no variation was instead observed for handwashing and the perceived threat. As concerns the association between these other health behaviours and the contact numbers, both avoiding social activities and the feeling of threat were generally associated with a lower number of contacts, while wearing a face mask was often positively associated (online supplemental figure S7). Age-specific contact numbers over time were quite heterogeneous across settings. In general, people aged 18–24 years reported higher contact numbers at home (representing the main source of contacts), while people aged 45–64 years and older adults reported higher contacts in the general community (online supplemental figure S9–S12 for the 90% quantile threshold and online supplemental figures S13–S16 for the ≤ 29 threshold).

Finally, the weekly estimates of the net reproduction number R_t^w generally decreased from the baseline value of the basic reproduction number R_0 under pre-COVID-19 social mixing, although with large heterogeneity among countries (figure 3). In Italy, Spain, France and Belgium, the reduction in contact numbers under the 90% quantile threshold led to decreases between 75% and 90% of the baseline R_0 , and hence to a R_t^w with 95% CI fully below one, as early as in week 2020–11 in Italy, while in the UK and in the USA, such evidence did not appear until week 2020–13. In Germany, the decline in the overall contact numbers led to a rather small decrease in the R_t^w , which reached its minimum value, 1.06 (0.63–1.51), in week 2020–13. Under the ≤ 29 threshold, we found that the higher number of contacts entailed reductions of smaller entity in both social contacts and in the R_t^w , which was below 1 for Italy and Spain, and equal to 1 for Belgium,

Table 1 Model-predicted mean number of daily contacts per person (with SEs) compared with prepandemic model predictions, by country, setting and week, March–April 2020

Country	Setting	Prior to		Week 2020–11		Week 2020–12		Week 2020–13		Week 2020–14		Week 2020–15	
		COVID-19		(9–15 March)		(16–22 March)		(23–29 March)		(30 March–5 April)		(6–12 April)	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Belgium	Home	3.15	0.15							1.33	0.05	1.35	0.03
	Community	5.34	0.3							0.64	0.06	0.74	0.03
	Work	3.05	0.36							0.35	0.07	0.36	0.04
	Overall	11.78	0.48							2.29	0.11	2.38	0.06
France	Home	2.7	0.07					1.28	0.03	1.31	0.03	1.25	0.03
	Community	4.59	0.13					0.5	0.04	0.45	0.03	0.41	0.03
	Work	2.37	0.12					0.24	0.04	0.23	0.03	0.26	0.04
	Overall	10.3	0.21					1.96	0.07	1.93	0.05	1.89	0.06
Germany	Home	2.73	0.1			1.52	0.04	1.44	0.03	1.45	0.02	1.46	0.02
	Community	2.97	0.14			1.33	0.07	1.13	0.05	0.93	0.03	1.15	0.04
	Work	1.82	0.17			0.76	0.1	0.56	0.05	0.71	0.05	0.68	0.05
	Overall	8.03	0.27			3.29	0.12	3.1	0.09	2.94	0.06	3.15	0.06
Italy	Home	3.56	0.19	1.67	0.03	1.63	0.03	1.57	0.03	1.5	0.04	1.6	0.05
	Community	7.91	0.36	0.71	0.04	0.6	0.03	0.47	0.03	0.48	0.03	0.56	0.05
	Work	5.28	0.53	0.51	0.04	0.36	0.04	0.25	0.03	0.21	0.03	0.3	0.06
	Overall	18.19	0.61	2.9	0.08	2.47	0.06	2.23	0.06	2.17	0.06	2.36	0.09
Netherlands	Home	3	0.23							1.57	0.04	1.56	0.04
	Community	7.13	0.76							0.85	0.04	0.91	0.04
	Work	4.97	1.08							0.67	0.06	0.74	0.07
	Overall	15.03	1.2							2.97	0.09	3.11	0.1
Spain	Home							1.63	0.04	1.61	0.03	1.65	0.03
	Community							0.66	0.05	0.45	0.03	0.49	0.03
	Work							0.25	0.05	0.19	0.03	0.26	0.03
	Overall							2.47	0.08	2.25	0.05	2.3	0.06
UK	Home	3.54	0.14	1.65	0.06	1.62	0.05	1.51	0.03	1.52	0.03	1.56	0.04
	Community	4.07	0.21	1.27	0.09	1.3	0.1	0.62	0.03	0.51	0.03	0.53	0.04
	Work	2.89	0.39	0.65	0.12	0.51	0.12	0.37	0.04	0.2	0.04	0.2	0.04
	Overall	10.58	0.35	3.43	0.17	3.18	0.15	2.4	0.05	2.17	0.06	2.28	0.07
USA	Home			1.66	0.05	1.69	0.03	1.66	0.02	1.56	0.04	1.69	0.03
	Community			1.5	0.07	1.07	0.05	0.8	0.03	0.7	0.03	0.72	0.03
	Work			0.64	0.08	0.55	0.06	0.47	0.04	0.45	0.05	0.41	0.05
	Overall			3.67	0.13	3.17	0.08	2.82	0.06	2.67	0.05	2.72	0.06

The 'overall' category encompasses contacts reported in all four surveyed settings, ie, home, school, work and general community.

France, Netherlands and, from week 13, the UK and the USA (online supplemental figure S17).

DISCUSSION

This paper provided comparative estimates of the daily number of social contacts, overall and by setting of contact, that people had during the early phase of the COVID-19 pandemic, between 13 March (week 2020–11) and 12 April (week 2020–15), in Belgium, France, Germany, Italy,

the Netherlands, Spain, the UK and the USA, disaggregated by week. Compared with the pre-COVID-19 period, overall contact numbers had substantively decreased, although not uniformly across all countries. Reductions were smaller than the 86% observed in Wuhan (China) in February 2020,⁹ as they ranged from 32% in Germany to 81% in Italy, under the ≤ 29 contacts threshold. Our estimates were consistent with those from other independent studies conducted at the same time in Europe and the

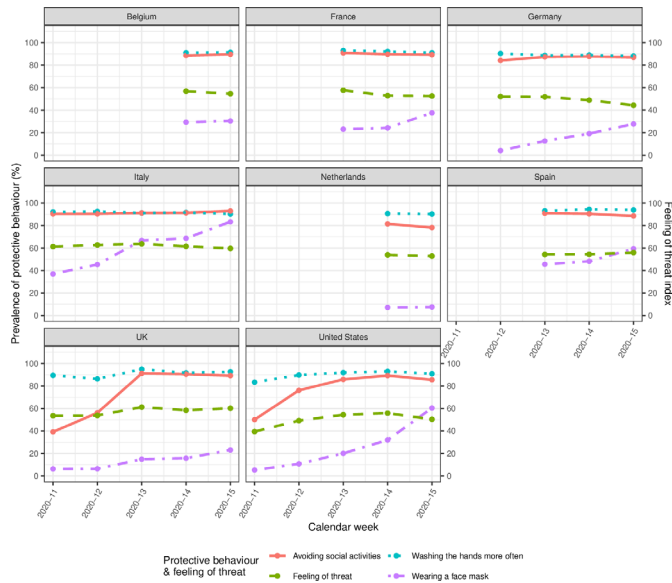


Figure 2 Prevalence of participants' uptake of different behavioural measures to protect themselves from the COVID-19 and feeling of threat (averaging the feeling of threat to oneself with that to the family, after rescaling to a 0–100 scale), by country and week.

USA.^{10–14 17} The unique aspect of our study is that it is the only one to use a consistent approach to collect comparative data across several countries and on a daily basis over an extended period of time.

Three results are especially relevant to characterise the changes in social contact patterns and their implications for infection transmission. First, our findings suggest that the countries with the highest numbers of social contacts were those that implemented only partial lockdowns (or stricter versions of the physical distancing guidelines), whereas those that implemented full lockdowns

experienced much lower contact numbers. In this latter group, compared with the prepandemic period, the household replaced the general community as the main source of social contacts, while in Germany and the Netherlands the general community and the work remained the settings contributing the most to the overall contact patterns.

Second, results for the UK and the USA showed that the numbers of contacts declined sharply after governments issued physical distancing guidelines, even before the introduction of lockdown measures, showing how spontaneous individual responses may also have played a role in the rapid adoption of protective behaviour. This observation is consistent with the strong association between high perception risk and the wearing of face masks among vulnerable people,³⁴ as well as with mobile phone data made available by Google, illustrating the change in human mobility compared with January and February 2020.^{35 36}

Third, in almost all countries, we found that the reduction in contact numbers, while keeping fixed other epidemiological factors, resulted in a reduction of the net reproduction number R_t below 1 and was therefore associated with a lower infection transmission, especially in a period in which the adoption rate of the face mask outside the home was still low, although gradually increasing. The exception was Germany, where the estimates of R_t remained consistently above 1 during the study period. Independent estimates of R_t at the beginning of April in Germany, that is, 1.1 with 95% CI 0.9 to 1.4, calculated by the Robert Koch Institute using incidence data, were consistent with our estimates under the 90% quantile threshold.³⁷ Our results can be explained by the relatively low social contact numbers in Germany during the pre-COVID-19 period (eight contacts per

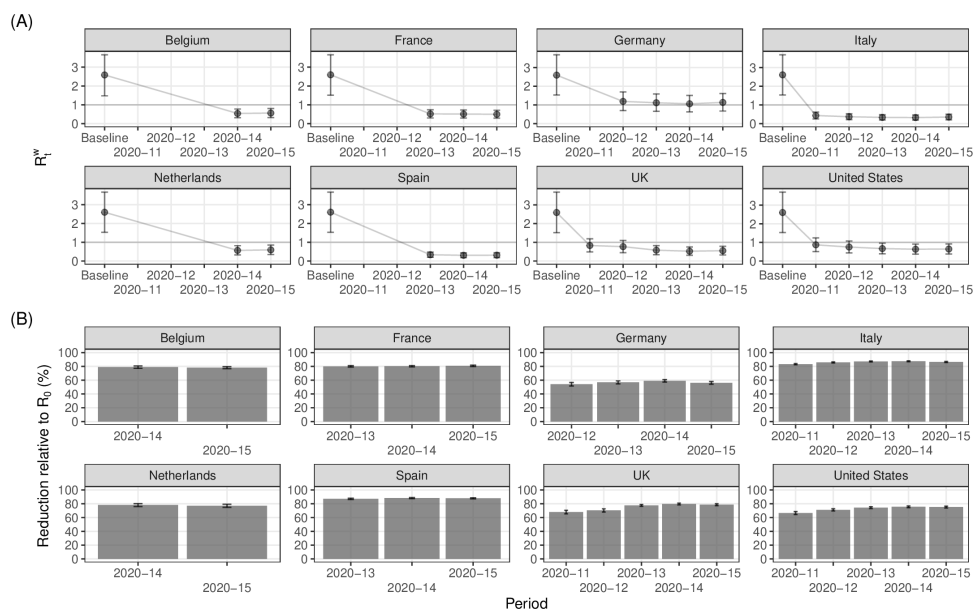


Figure 3 Absolute change in the weekly net reproduction number R_t^w with respect to the basic reproduction number R_0 at baseline (A) and percent reduction with respect to R_0 (B), by country and week. The 95% CIs are based on 5000 replicates.

day, on average) and the smaller reductions in social contact numbers during the study period. Nonetheless, our estimates should not be interpreted as evidence that the infection spread was exceptionally high in Germany despite physical distancing interventions. Changes in the R_t are brought about by changes in social contact patterns and in infection transmissibility, which might be affected by other health behaviours (eg, physical distancing in public spaces, increased hand washing or sanitising and compliance with mask wearing^{33 38}) or environmental factors (eg, surface disinfection and indoor ventilation). Although our sample provided data on some of the behavioural factors (ie, hand washing, face mask wearing and avoiding social activities)—which we included in the estimation of the contact numbers—we could not explicitly account for their impact, as well as for the impact of other important factors for which we did not collect data, in the calculation of R_t , which may be seen as a limitation of this study.

The study has other limitations that should be considered when interpreting the results. First, we collected data using a sample of Facebook users. Although such nonprobability samples are somewhat less accurate than probability samples in obtaining statistics that are generalisable to the whole population, with the appropriate statistical adjustments, such as those we made, they offer a good approximation to results that could be obtained from probabilistic samples, at least at the country level. Furthermore, samples of Facebook users can be gathered rapidly and therefore offer data that would be difficult to collect otherwise during a pandemic. Indeed, to the best of our knowledge, our study is the only cross-country survey on social contacts that exists for the unique period of time considered. Finally, our data cover a large population at minimal costs compared with more traditional surveys.⁵

Second, our approach of recruiting survey respondents through targeted online advertisements may aggravate issues of self-selection bias. Indeed, our participants might differ to some extent from the general population in terms of their sociability patterns (due to their Facebook use) and their concerns about health-related issues (as they opted in to participate in the survey). We tackled this issue by stratifying the campaigns and applying poststratification weights using factors that are linked to both survey participation and the outcomes of interest^{39 40} and ensuring that only respondents who were shown the Facebook ads and clicked on the link could participate in the study (thereby avoiding biases in the sample that we cannot adjust for). However, poststratification weights cannot correct entirely for different behaviours or social inequalities (eg, education level, ethnicity, type of profession or urban/rural residence) among Facebook users, since such variables cannot be included in the ads stratification. Further improvements, which go beyond of the scope of this work, may come from using a multilevel regression and poststratification approach, which has shown to work well in practice when making inference

from non-probability samples, especially when having at disposal data from census or a true probability sample on the demographic strata of interest in the population.⁴¹

Third, several survey participants reported extremely high contact numbers, perhaps because they had little motivation to provide accurate counts. This is a general problem in social contact surveys. Typically, these surveys either include a maximum number of contacts that respondents can report, or researchers assess ex-post the threshold level for the right tail of the contact distribution of contacts. We followed the latter approach and removed respondents reporting implausible values. We found that our results are robust to a choice of cut-off values consistent with those used in standard surveys in the field. Future contact surveys may consider including attention checks items to reduce potential issues with outliers.

Finally, the calculation of the net reproduction number hinges on two factors characterised by large uncertainty, namely, the initial estimate of the basic reproduction number R_0 and the age-specific social contact matrix. For R_0 , we employed a common value for all parameters coming from a meta-analysis based on data from China, South Korea and Italy; although this estimate was not specific for each surveyed country, the attached uncertainty made it consistent with estimates used in other studies^{11 14 17} and allowed us to disentangle the impact of the reduction in contact numbers from that of other epidemiological factors. For the contact matrices, as we did not collect the age of the individuals contacted by respondents like other single-country studies did,^{10–12 14 17} we scaled down the prepandemic matrices by age-specific factors derived from the data collected in our study. Although our matrices are not based on the observed age mixing patterns during the early phase of the pandemic, we believe that the results on the reduced infection transmission entailed by the reduced contact numbers are still valid and consistent with the results of the other studies.

In conclusion, we provided an international comparable set of statistics on social contact patterns during the COVID-19 pandemic, disaggregated by country, setting and week (<https://osf.io/esdmg/>). As these estimates offer a more grounded alternative to the theoretical assumptions that have often been used,^{38 42 43} we believe that the scientific community can draw on this information for developing more realistic epidemic models of COVID-19.

Correction notice This article has been corrected since it first published. Supplementary file has been updated.

Twitter Emanuele Del Fava @emanudelfava

Acknowledgements We would like to thank all participants who took the time to voluntarily complete our survey; the staff at the Max Planck Institute for Demographic Research who contributed to the realisation of this project, including the administrative department, the IT department and the press department; and all our colleagues who took the time to review the translation and the implementation of the survey (Maarten J Bijlsma, Mathias Lerch, Marília R Nepomuceno, Tom Theile, Nicolas Todd and Joshua Wilde). We would also like to thank two anonymous

referees and the editor of the journal, whose valuable comments allowed us to greatly improve the quality of the manuscript.

Contributors Conceptualisation: EDF and EZ; data curation: AG and JC; formal analysis and investigation: EDF and JC; project administration: DP and AG; writing – original draft preparation: EDF, JC and EZ; writing – review and editing: all authors; funding acquisition: EZ; resources: DP, AG and FR and SG-C; supervision: EZ; guarantor: EZ.

Funding This study was funded with support from the Max Planck Institute for Demographic Research (no award/grant number received), which is part of the Max Planck Society.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Ethical approval for the study was obtained from the Ethics Council of the Max Planck Society (Application No: 2020_07). Participation in this study was voluntary, anonymous and open to people who were at least 18 years old. Electronic informed consent was obtained from all participants who actively opted in to participate in the online survey, enabling the collection, storage and processing of their answers.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. The aggregated contact data that underlie the analyses reported in the paper are provided as weighted averages of daily contact numbers by threshold, country, calendar week, and setting, both overall and stratified by region of residence, age group and sex. They are available in a public, open access repository at the following link: <https://osf.io/esdmng/>

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Emanuele Del Fava <http://orcid.org/0000-0001-7352-8525>

REFERENCES

- Greenhalgh T, Jimenez JL, Prather KA, *et al*. Ten scientific reasons in support of airborne transmission of SARS-CoV-2. *Lancet* 2021;397:1603–5.
- 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.
- 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.
- Perrotta D, Grow A, Rampazzo F, *et al*. Behaviours and attitudes in response to the COVID-19 pandemic: insights from a cross-national Facebook survey. *EPJ Data Sci* 2021;10:1–13.
- Grow A, Perrotta D, Del Fava E, *et al*. Addressing public health emergencies via Facebook surveys: advantages, challenges, and practical considerations. *J Med Internet Res* 2020;22:e20653.
- Béraud G, Kazmierczak S, Beutels P, *et al*. The French connection: the first large population-based contact survey in France relevant for the spread of infectious diseases. *PLoS One* 2015;10:e0133203.
- Melegaro A, Jit M, Gay N, *et al*. What types of contacts are important for the spread of infections?: Using contact survey data to explore European mixing patterns. *Epidemics* 2011;3:143–51.
- Brooks-Pollock E, Read JM, McLean AR, *et al*. Mapping social distancing measures to the reproduction number for COVID-19. *Philos Trans R Soc Lond B Biol Sci* 2021;376:20200276.
- Zhang J, Litvinova M, Liang Y, *et al*. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* 2020;368:1481–6.
- Backer JA, Mollema L, Vos ERA, *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. *Eurosurveillance* 2021;26:2000994.
- Coletti P, Wambua J, Gimma A, *et al*. CoMix: comparing mixing patterns in the Belgian population during and after lockdown. *Sci Rep* 2020;10:21885.
- Latsuzbaia A, Herold M, Bertemes J-P, *et al*. Evolving social contact patterns during the COVID-19 crisis in Luxembourg. *PLoS One* 2020;15:e0237128.
- Sypsa V, Roussos S, Paraskevis D, *et al*. Effects of social distancing measures during the first epidemic wave of severe acute respiratory syndrome infection, Greece. *Emerg Infect Dis* 2021;27:452–62.
- 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. *BMC Med* 2020;18:124.
- Quaife M, van Zandvoort K, Gimma A, *et al*. The impact of COVID-19 control measures on social contacts and transmission in Kenyan informal settlements. *BMC Med* 2020;18:316.
- Dorélien AM, Simon A, Hagge S, *et al*. Minnesota social contacts and mixing patterns survey with implications for modelling of infectious disease transmission and control. *Surv Pract* 2020;13:1–12.
- Feehan DM, Mahmud AS. Quantifying population contact patterns in the United States during the COVID-19 pandemic. *Nat Commun* 2021;12:893.
- Pötzschke S, Braun M. Migrant sampling using Facebook advertisements: a case study of Polish migrants in four European countries. *Soc Sci Comput Rev* 2016;35:633–53.
- Zhang B, Mildenerberger M, Howe PD. Quota sampling using Facebook advertisements. *Polit Sci Res Methods* 2018;8:1–7.
- Eurostat. Eurostat regional yearbook: 2019 edition, 2019. Available: http://publications.europa.eu/publication/manifesteration_identifler/PUB_KSHA19001ENN [Accessed 15 May 2020].
- US Census Bureau. 2018 population estimates by age, sex, race and Hispanic origin. U. S. census Bur, 2019. Available: <https://www.census.gov/newsroom/press-kits/2019/detailed-estimates.html> [Accessed 15 May 2020].
- Teslya A, Pham TM, Godijk NG, *et al*. Impact of self-imposed prevention measures and short-term government-imposed social distancing on mitigating and delaying a COVID-19 epidemic: a modelling study. *PLoS Med* 2020;17:e1003166.
- Hens N, Goeyvaerts N, Aerts M, *et al*. Mining social mixing patterns for infectious disease models based on a two-day population survey in Belgium. *BMC Infect Dis* 2009;9:5.
- Liu Q-H, Ajelli M, Aleta A, *et al*. Measurability of the epidemic reproduction number in data-driven contact networks. *Proc Natl Acad Sci U S A* 2018;115:12680–5.
- The Royal Society. Reproduction number (R) and growth rate (R) of the COVID-19 epidemic in the UK, 2020. Available: <https://royalsociety.org/news/2020/09/set-c-covid-r-rate/>
- Diekmann O, Heesterbeek JA, Metz JA. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *J Math Biol* 1990;28:365–82.
- Goeyvaerts N, Hens N, Ogunjimi B, *et al*. Estimating infectious disease parameters from data on social contacts and serological status. *J R Stat Soc Ser C Appl Stat* 2010;59:255–77.
- Mossong J, Hens N, Jit M. *Data from: POLYMOD social contact data*. Zenodo, 2020.
- Béraud G, Kazmierczak S, Beutels P. *Data from: social contact data for France*. Zenodo, 2020.
- Arregui S, Aleta A, Sanz J, *et al*. Projecting social contact matrices to different demographic structures. *PLoS Comput Biol* 2018;14:e1006638.
- R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2021. <https://www.R-project.org/>
- Mervosh S, Lu D, Swales V. See which states and cities have told residents to stay at home. N. Y. times, 2020. Available: <https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html> [Accessed 25 Sep 2020].
- Flaxman S, Mishra S, Gandy A, *et al*. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 2020;584:257–61.
- Perrotta D, Del Fava E, Zagheni E. The role of vulnerability in mediating the relationship between threat perception and the use of face masks in the context of COVID-19. *Workshop Proceedings of*

- the 15th International AAAI Conference on Web and Social Media, 2021.
- 35 Google LLC. Google COVID-19 community mobility reports. Available: <https://www.google.com/covid19/mobility/> [Accessed 22 Jun 2020].
 - 36 Basellini U, Alburez-Gutierrez D, Del Fava E, *et al*. Linking excess mortality to mobility data during the first wave of COVID-19 in England and Wales. *SSM Popul Health* 2021;14:100799.
 - 37 RKI. RKI - Coronavirus SARS-CoV-2 - Situation report - 9 Apr 2020. Robert Koch Institute, 2020. Available: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/2020-04-09-en.pdf?__blob=publicationFile [Accessed 15 May 2020].
 - 38 Dehning J, Zierenberg J, Spitzner FP, *et al*. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* 2020;369. doi:10.1126/science.abb9789. [Epub ahead of print: 10 07 2020].
 - 39 Keiding N, Louis TA. Perils and potentials of self-selected entry to epidemiological studies and surveys. *J R Stat Soc Ser A Stat Soc* 2016;179:319–76.
 - 40 Mercer AW, Kreuter F, Keeter S. Theory and practice in Nonprobability surveys: parallels between causal inference and survey inference. *Public Opin Q* 2017;81:250–71.
 - 41 Wang W, Rothschild D, Goel S, *et al*. Forecasting elections with non-representative polls. *Int J Forecast* 2015;31:980–91.
 - 42 Gatto M, Bertuzzo E, Mari L, *et al*. Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures. *Proc Natl Acad Sci U S A* 2020;117:10484–91.
 - 43 Salje H, Tran Kiem C, Lefrancq N, *et al*. Estimating the burden of SARS-CoV-2 in France. *Science* 2020;369:208–11.