RESEARCH ARTICLE SUMMARY

CORONAVIRUS

Drivers of epidemic dynamics in real time from daily digital COVID-19 measurements

Michelle Kendall*†, Luca Ferretti*†, Chris Wymant, Daphne Tsallis, James Petrie, Andrea Di Francia, Francesco Di Lauro, Lucie Abeler-Dörner, Harrison Manley, Jasmina Panovska-Griffiths, Alice Ledda, Xavier Didelot, Christophe Fraser*

INTRODUCTION: Monitoring is the first step in responding to epidemics. A key metric is the time-varying reproduction number $R(t)$—the average number of secondary infections arising from an infected individual at time t. However, many more aspects of the underlying transmission processes can be measured, for example, the contexts in which more transmission occurs. This could inform more targeted interventions, helping to minimize negative socioeconomic impacts.

RATIONALE: Digital contact tracing was introduced during the COVID-19 pandemic to reduce viral transmission: Apps recorded proximity between users and notified them later if they had a risky exposure to an infected case. As a by-product, this process generated data on the link between exposure and transmission. The potential of such data to improve epidemic monitoring and understanding has received little attention. We explored this using data from the National Health Service (NHS) COVID-19 app for England and Wales, which was used by 6 million to 18 million people (aged 16 years and older) over 55 months.

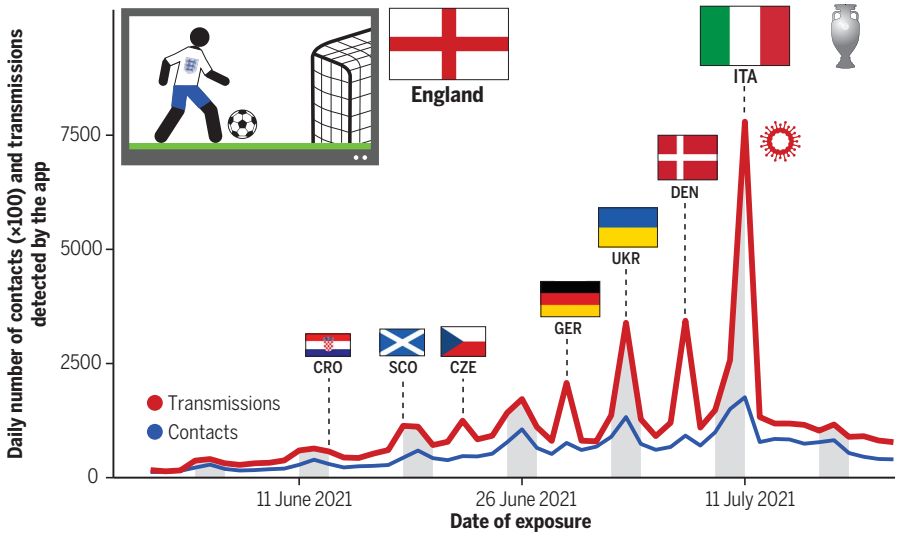
RESULTS: We calculated an indicator of $R(t)$, called $R_{app}(t)$, which identified epidemic changepoints consistently with other estimates of $R(t)$ and was available at least 5 days earlier. It offered the capability of separating the individual roles of contact rates and the probability of reported infection among notified contacts. We measured contact rates as the mean number of exposure notifications triggered for each index case, normalized by app usage. Contact rates changed markedly with restrictions and during Christmas holidays; the probability of reported infection changed markedly with vaccination and the arrival of the Omicron variant. Before Omicron, changes in contact rates contributed more than transmission risk to the daily variation in $R_{app}(t)$, whereas changes in $R_{app}(t)$ for Omicron were almost entirely explained by changes in transmission risk.

We defined transmissions as risky exposure events recorded by the app for which the exposed individual reported testing positive sh afterward, after subtracting the backgro risk of infection from undetected exposures. True infection levels in app users are underestimated because of underascertainment in voluntary testing and reporting. We classified transmissions into approximate settings based on their duration and frequency and saw pronounced variability over days of the week. For example, household transmissions peaked on Sundays, with 1.3 times as many as on Fridays; those from fleeting exposures (<30 min) peaked on Saturdays, with 2.6 times as many as on Mondays.

As one example of the app's capability to provide epidemiological insights with daily resolution into contact and transmission patterns, we analyzed the 2021 Christmas season. In the weeks before Christmas Day, contact rates were elevated and there were local peaks in transmissions from 1-day exposures on Saturdays. On Christmas Day, contact rates decreased and transmissions shifted toward household settings, before shifting back to briefer exposures for Boxing Day and (nonhousehold) recurring contacts on New Year's Eve. These findings are in line with known seasonal patterns of shopping and socializing. We found similar results for Christmas 2022.

As a second example, we analyzed the Euro football tournament in 2021. Days with England matches showed sharp spikes in transmissions (see the figure), owing to both increased contact rates and increased probability of reported transmission. We saw analogous patterns for Wales, though based on fewer app users. The day of the final match (England versus Italy) recorded the highest number of daily transmissions for 2021.

CONCLUSION: In addition to its primary purpose as an intervention to reduce transmission, digital contact tracing for respiratory pathogens can provide insights into epidemic dynamics with unprecedented time resolution and rapid availability, facilitating responsive decision-making. We shared results with the UK government and public health authorities weekly or at peak times daily. Despite collecting minimal anonymized data from a user base potentially not representative of the population, the app-based indicator tracked changes in $R(t)$ that were observed in other data streams. We should continue developing the programmatic infrastructure and our analytic capabilities for digital contact tracing to be ready for the next pandemic. ■



Example of app-based epidemic monitoring in England during the Euro 2020 football tournament. The highest peak in contacts (176,000) and COVID-19 transmissions (7800) occurred on the day of the final. The lines show the estimated daily number of app-detected contacts (blue, in units of a hundred individuals) and transmissions (red), along with dates of matches involving the England team, opponents [identified by flags and Fédération Internationale de Football Association (FIFA) country codes], and weekends (shaded gray).

The list of author affiliations is available in the full article online.
*Corresponding author. Email: michelle.kendall@warwick.ac.uk (M.K.); luca.ferretti@ndm.ox.ac.uk (L.F.); christophe.fraser@ndm.ox.ac.uk (C.F.)
†These authors contributed equally to this work.
Cite this article as M. Kendall *et al.*, *Science* **385**, eadm8103 (2024). DOI: 10.1126/science.adm8103

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adm8103

RESEARCH ARTICLE

CORONAVIRUS

Drivers of epidemic dynamics in real time from daily digital COVID-19 measurements

Michelle Kendall[1]*†, Luca Ferretti[2,3]*†, Chris Wymant[2,3], Daphne Tsallis[4], James Petrie[2,3], Andrea Di Francia[5], Francesco Di Lauro[2,3], Lucie Abeler-Dörner[2,3], Harrison Manley[5], Jasmina Panovska-Griffiths[2,3,5], Alice Ledda[5], Xavier Didelot[1,6], Christophe Fraser[2,3]*

Understanding the drivers of respiratory pathogen spread is challenging, particularly in a timely manner during an ongoing epidemic. In this work, we present insights that we obtained using daily data from the National Health Service COVID-19 app for England and Wales and that we shared with health authorities in almost real time. Our indicator of the reproduction number $R(t)$ was available days earlier than other estimates, with an innovative capability to decompose $R(t)$ into contact rates and probabilities of infection. When Omicron arrived, the main epidemic driver switched from contacts to transmissibility. We separated contacts and transmissions by day of exposure and setting and found pronounced variability over days of the week and during Christmas holidays and events. For example, during the Euro football tournament in 2021, days with England matches showed sharp spikes in exposures and transmissibility. Digital contact-tracing technologies can help control epidemics not only by directly preventing transmissions but also by enabling rapid analysis at scale and with unprecedented resolution.

Monitoring is the first step in responding to epidemics, guiding all decisions that follow. Early in the COVID-19 pandemic, the director-general of the World Health Organization, Tedros Adhanom Ghebreyesus, urged countries to test widely, arguing, "You cannot fight a fire blindfolded. And we cannot stop this pandemic if we don't know who is infected" (*1*). Most countries heeded this advice and developed testing infrastructure at scale. Testing is the most basic ingredient of epidemic monitoring—counting the number of people affected at a given time.

The time-varying reproduction number $R(t)$ characterizes the dynamics of an epidemic. There is some variation in proposed definitions of the measure, but the core idea is to capture the average number of secondary infections that arise from an infected individual at time t over the course of their whole infectious period (*2*–*5*). Using a model for epidemic dynamics, $R(t)$ can be inferred from observations reported at different times, including not only case counts but also survey data, hospital admissions, and deaths (*6*–*15*). $R(t)$ is a key metric for evaluating past and predicting future impacts of interventions and is widely used to assess the recent, present, or future state of the epidemic.

Epidemics are driven by complex patterns of pathogen transmission in dynamically interacting populations. The complexity of the system implies the possibility of measuring many more aspects of the transmission process than case counts—at least in theory. In practice, the usually rapid transmission of respiratory pathogens in close-contact settings makes it difficult to obtain measurements at scale. Better characterization of complex epidemic systems has the power to inform public health decisions, particularly when the results are rapidly available. More detailed understanding of the changing drivers of an epidemic and the impact of interventions allows future interventions to be more targeted and tailored, helping to minimize negative socioeconomic impacts.

Digital contact tracing was proposed early in the COVID-19 pandemic as a tool to reduce transmission (*16*–*18*). Bluetooth-based contact-tracing apps were introduced in many countries (*19*, *20*), and their varying degree of success in reducing transmission has been evaluated (*21*–*26*). Little attention has been paid to the possibility of using digital contact tracing to improve epidemic monitoring, using the minimized and anonymized data that were generated to ensure the correct functioning of the app (*27*, *28*). In this work, we demonstrate this potential using data from the National Health Service (NHS) COVID-19 app for England and Wales (henceforth, "the app"). The dataset produced as a by-product of the app's contact-tracing function was large and epidemiologically informative, which is attributable to the app being active on 6 million to 18 million devices (12 to 37% of the population aged 16 years and older, assuming one device per person) at all times over its 55 months of operation. We show that these data enabled rapid epidemic monitoring with an unprecedented degree of granularity, providing singular insights into the drivers of epidemic dynamics.

Data from the NHS COVID-19 app

We analyzed different sources of data available through the app covering early 2021 to early 2023. One source was analytics data, described previously (*22*). This comprised a small amount of anonymous data sent to a central server each day by each installation of the app on a mobile device. It indicated, for example, whether the user received an exposure notification today or recently, whether they reported a positive test result today, and their self-declared postcode district. No data on age was available through the app, aside from users self-reporting being aged 16 years or older, which was an eligibility requirement for using the app. Thus, all results that we report here are for individuals aged 16 years and older.

In the UK, there was never an obligation to take a test after an exposure notification, and recommendations to do so were only in place from 16 August 2021 to 24 February 2022. Therefore, all our results concerning infections among notified app users are underestimates of the true amount of infection, which was due to underascertainment when considering only voluntarily reported positive tests. Infections may also have been caused by an interaction that was not detected by the app, particularly when background prevalence was high.

Disaggregating reproduction number dynamics into contact rates and infection probability

From app data, we calculated a contact rate (CR), defined as the mean number of contacts notified per positive test reported through the app on day t, normalized by the proportion of the population using the app and the proportion of test-positive users that consented to contact tracing that day. This measure estimates the average number of people (with or without the app) who came into close contact with a test-positive app user in a time window of a few days before they tested positive (see Materials and methods). The contact rate CR (Fig. 1A) captured similar trends to Google Mobility (*29*) and the CoMix study (*30*) (figs. S1 and S2). Figure 1B shows the proportion of notified app users who reported through the app that they tested positive shortly after

[1]Department of Statistics, University of Warwick, Coventry CV4 7AL, UK. [2]Pandemic Sciences Institute, Nuffield Department for Medicine, University of Oxford, Old Road Campus, Oxford OX3 7DQ, UK. [3]Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Medicine, Old Road Campus, University of Oxford, Oxford OX3 7LF, UK. [4]Zühlke Engineering Ltd., 80 Great Eastern Street, London EC2A 3JL, UK. [5]UK Health Security Agency, Nobel House, 17 Smith Square, London SW1P 3JR, UK. [6]School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK.
*Corresponding author. Email: michelle.kendall@warwick.ac.uk (M.K.); luca.ferretti@ndm.ox.ac.uk (L.F.); christophe.fraser@ndm.ox.ac.uk (C.F.)
†These authors contributed equally to this work.

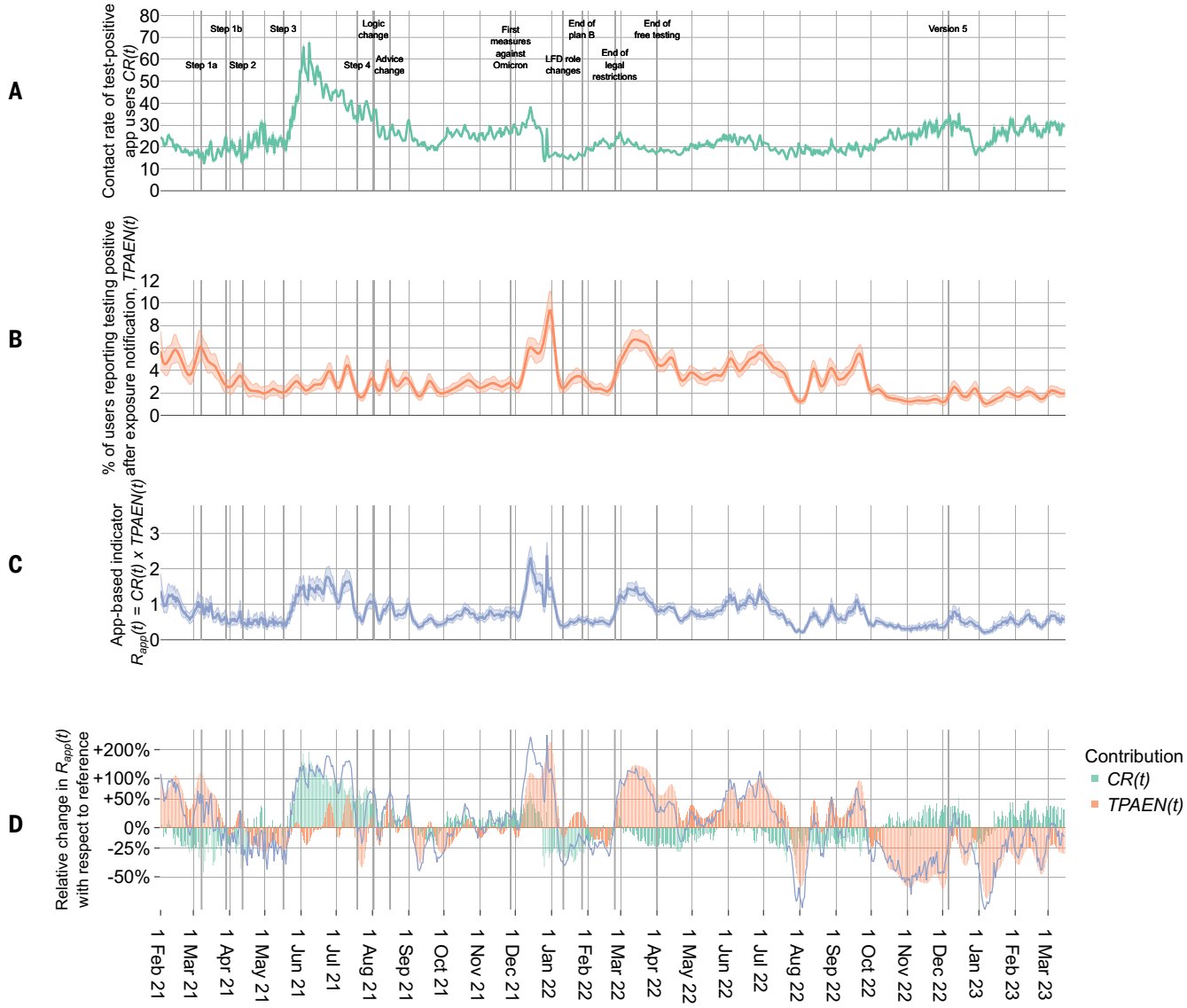


Fig. 1. Contribution of contact rates of test-positive app users and the probability of reported infection to the reproduction number. (**A**) Contact rate *CR*. (**B**) *TPAEN*. (**C**) The app-based indicator, $R_{app}(t) = CR \times TPAEN$. (**D**) The relative change in $R_{app}(t)$ and the respective contributions of *CR* and *TPAEN*. The *y* axis is log scaled such that the contributions sum precisely to the changes in $R_{app}(t)$. Annotated vertical lines show key policy changes and app updates (see timeline in supplementary text section 5). In (A) to (C), the line shows the mean and shading shows the 95% credible interval.

exposure notification (*TPAEN*) (*23*). This metric is similar to the secondary attack rate but is an underestimate of the proportion of notified users who were truly infected due to underascertainment of cases. Nevertheless, it captures the expected dynamics of transmissibility, including a decrease during vaccine rollout in the first half of 2021 and an increase with the arrival of the Omicron variant in December 2021. The product of *CR* and *TPAEN*, which we denote $R_{app}(t)$, is an indicator for $R(t)$ (Fig. 1C): It estimates the average number of contacts who reported an infection after being exposed to a test-positive app user. These exposures must occur in a short time window before the user reports a positive test (at which point, their app temporarily pauses contact tracing), making $R_{app}(t)$ an underestimate of $R(t)$.

Unlike existing $R(t)$ estimators from time series of cases, $R_{app}(t)$ allows changes in transmission rates over time to be attributed to changes in contact rates (*CR*) or changes in probability of reported infection in those contacts (measured by *TPAEN*) or both (Fig. 1D). The rapid increase in $R_{app}(t)$ in May 2021 was driven more by an increase in contact rates than by an increase in the probability of reported infection. A road map of policy changes for lifting social restrictions was presented in spring 2021 (*31*, *32*) (see the timeline in supplementary text section 5); the biggest increases in contact rates occurred after stage 3, with stage 4 ("Freedom Day") causing little change in contact rates.

Some new viral variants had increased transmissibility and thus increased $R(t)$ (*33*, *34*). The arrival of the Delta variant caused surprisingly little increase in the probability of reported infection, likely owing to the contemporaneous vaccination program. The arrival of the Omicron variant (*35*) in November 2021, however, caused an increase in $R_{app}(t)$ that was almost entirely driven by higher transmissibility. Before Omicron appeared at the end of November 2021, *TPAEN* alone explained 30% of the daily variability in $R_{app}(t)$ and contact rates alone explained 46% (as assessed

by the squared log-scaled correlation coefficient). Daily variations in *TPAEN* and *CR* were anticorrelated in this period [log-scaled Pearson correlation coefficient (r) = −0.25]. With the arrival of Omicron, the driver of the epidemic shifted from contacts to transmissibility (fig. S3). From December 2021 onward, *TPAEN* and *CR* were still anticorrelated (r = −0.37), but the daily variation in $R_{app}(t)$ was almost entirely explained by *TPAEN* (83% versus 0.2% by contact rates).

Figure 2 shows publicly available estimates of $R(t)$ together with the app-based indicator $R_{app}(t)$. We see broad agreement around major change points across the different measures, suggesting that the sample of app users was sufficiently large and diverse to provide a reliable signal of epidemic dynamics at the population level (despite our expectation that app users are not a representative sample of the general population; the app did not collect any demographic data about its users to test this). For example, these measures all agree on an increase in $R(t)$ around the beginning of December 2021. Some variation in timing is explained by differences such as the handling of lags in reporting and the precise definitions of the measures, such as whether $R(t)$ is defined for someone who was infected, infectious, or testing positive on day t. Figure S4 shows the measures that were combined to provide the official estimates of $R(t)$, published initially by the Scientific Advisory Group for

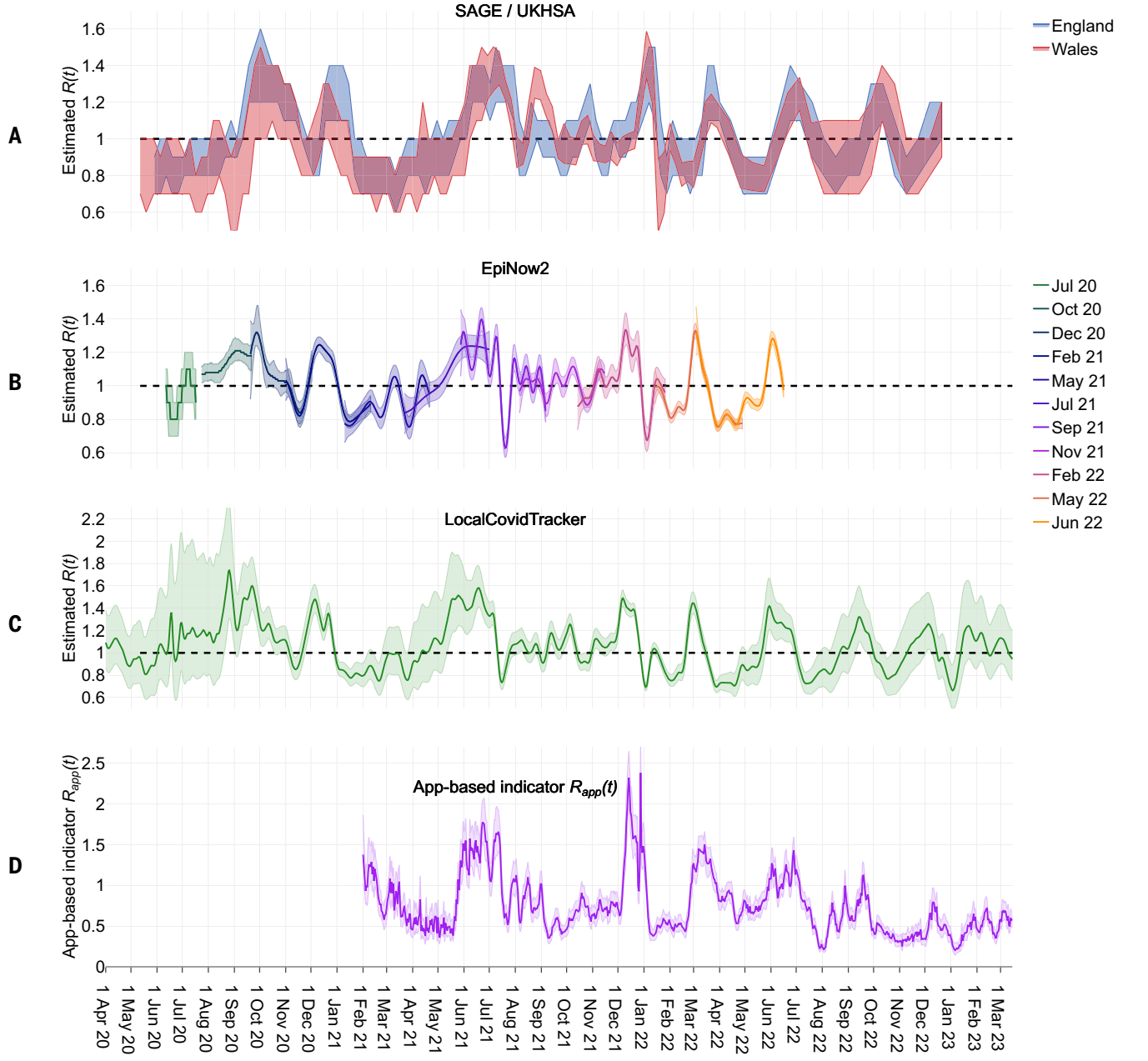


Fig. 2. Estimates of the time-varying *R(t)*. (**A**) Estimated $R(t)$ for England (*15*) and Wales (*55*) as published by SAGE, and later UKHSA, using an ensemble of methods (*56*). (**B**) Estimated $R(t)$ for the UK using the EpiNow2 method (*10*). Estimates were published (*48*) almost daily for the previous 12 weeks; plotted are snapshots of the measure as published approximately every 10 weeks, as indicated by the colors shown in the legend. (**C**) Estimated $R(t)$ for England and Wales combined using the LocalCovidTracker method (*11*, *57*). (**D**) The app-based indicator $R_{app}(t)$. In (A) to (C), the dashed horizontal lines show $R(t)$ = 1. Shading represents credibility intervals.

Emergencies (SAGE) and later by the UK Health Security Agency (UKHSA).

For many applications, specifying the exact date of an $R(t)$ change point is less important than simply identifying that there has been a change, and doing so rapidly. Estimates of $R_{app}(t)$ were available with a 6-day lag from the date of a positive test being reported in the app. The lags for other estimates varied according to data availability and publication schedules; we consider December 2021 as an example. An increase in $R(t)$ became apparent on 17, 16, 15, and 10 December for SAGE/UKHSA (*15*), EpiNow2 (*10*), LocalCovidTracker (*11*, *36*), and the app, respectively. Thus, the app-based indicator $R_{app}(t)$ alerted us to a change point around 5 days earlier than these publicly available estimators (figs. S5 and S6). App delays are shown in figs. S7 to S9.

Disaggregating *CR* and *TPAEN* across Wales and the nine regions of England (Fig. 3), we

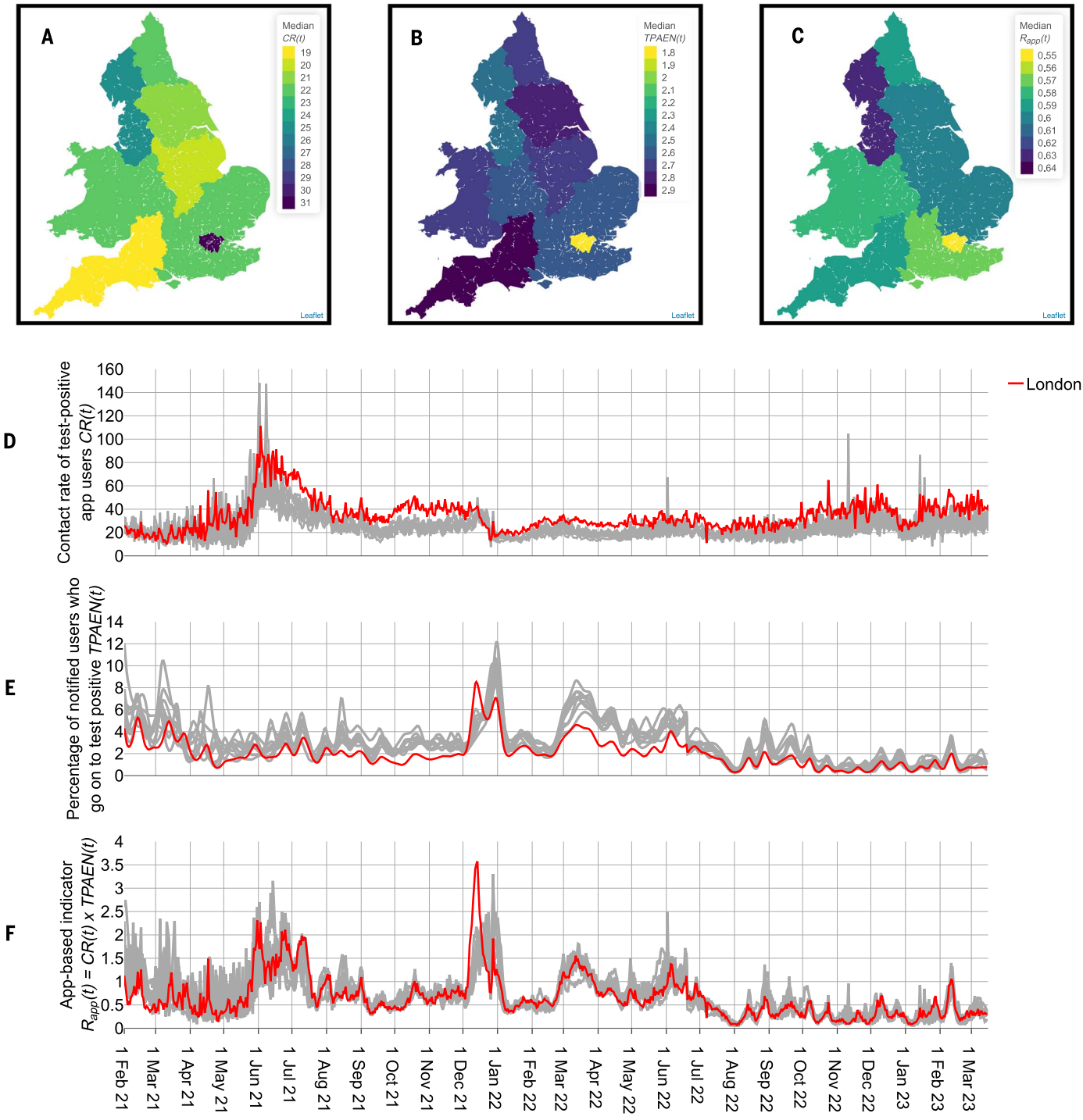


Fig. 3. Subnational variation in *CR*, *TPAEN*, and $R_{app}(t)$ between 1 February 2021 and 15 March 2023. (**A** to **C**) Median *CR*, median *TPAEN*, and median $R_{app}(t)$, respectively, for each region of England and Wales. (**D** to **F**) Time series of *CR*, *TPAEN*, and $R_{app}(t)$, respectively, for each region of England and Wales, with the London region shown in red.

saw some regional variation, which was especially stark for London (red in Fig. 3, D to F). From 1 February 2021 to 15 March 2023, London's median *CR* was 32, compared with 20 to 26 for other regions, and its median *TPAEN* was 1.8, compared with 2.5 to 2.9 for other regions (Fig. 3, A and B). These differences were not driven by only a few key moments or events—they were present at most times throughout the period of study (Fig. 3, D and E). London's high *CR* and low *TPAEN* combined to give a moderate $R_{app}(t)$ estimate (Fig. 3C), which was similar to those of other regions at most times (Fig. 3F). We observed other regional variations, for example, the median $R_{app}(t)$ for South West England was comparable to that for much of East of England but consisted of lower contact rates and higher *TPAEN*.

Data on exposure events

The analyses above were all based on the daily analytics data recorded by the app. An additional source of data came in the form of event data, which recorded risky exposure events. We described these data previously (*37*). Briefly, when two app users spent time with their phones in close proximity, anonymized data about this interaction was stored for a 2-week period on their respective phones. If one of the users recorded a positive test in the app during this period, and if the timing and proximity of the interaction crossed a risk threshold, then an exposure notification was triggered for the other user, and some limited data about the exposure was sent to a central server.

We define a contact event to be an interaction where two app users' phones were sufficiently close for long enough to trigger an exposure notification after one of them reported a positive test. A contact is an individual who received a notification. We define an infection event detected by the app to be a positive test result reported via the app after a risky exposure notification, during an interval beginning with notification and ending 14 days after the last exposure. For such events, we estimated and subtracted the number attributed to background transmissions not detected by the app, leaving those transmission events that we attributed to the recorded exposure (*37*). The fraction of contact events that result in transmission events defines the probability of reported transmission. For each contact, we calculated their cumulative risk score by summing the risk score calculated by the app, based on proximity and duration, for each separate 30-min window covering the whole exposure event. Previously, we showed this measure to be the best single predictor of probability of reported transmission, as detected by the app (*37*).

For contact and transmission events with an exposure that spanned multiple days, we assigned a single representative day probabilistically in proportion to the total duration of exposure on each day. Normally, case data are indexed only by the date of reporting, which has a variable delay from date of infection, introducing a smoothing effect on any daily variation in infections (fig. S10). Here, having actual dates of exposure provided sharper resolution into the underlying transmission patterns, as demonstrated in the next three sections.

The app data do not contain any information about the location of a contact event, but the total duration and frequency of exposure between contact and index case can be used to approximately classify them into settings. We previously defined these as household contacts (exposure on at least 1 day for at least 8 hours), recurring contacts (exposure on more than 1 day for less than 8 hours each day), single-day contacts (between 30 min and 8 hours), and fleeting contacts (less than 30 min) (*37*). Table S1 summarizes the data in these four settings.

Disaggregating contact and transmission events by day and setting

Figure 4A shows the daily total number of notified contacts, indexed by their date of exposure and colored by their mean cumulative risk score. This risk score is closely related to the actual risk of transmission (*37*). The steep increase in contacts in spring 2021 reflects the aforementioned roadmap for lifting restrictions. In particular, a set of policy changes on 17 May 2021 led to increased mobility and app use (*23*, *38*), coinciding with the arrival of the Delta variant to the UK (*36*) and increasing levels of virus circulation (*39*). Nevertheless, the mean of the cumulative risk score each day was relatively low in spring 2021, which was likely due to the remaining restrictions reducing the duration and/or the proximity of contacts.

The number of daily contacts reached a peak of 176,000 on 11 July 2021, the date of the final match of the (postponed) Euro 2020 tournament. This was followed by a decrease in contacts that was due to a decrease in cases, a decrease in app use (fig. S13), and probably some effects of summer holiday travel and other behavioral changes as restrictions eased further. Indeed, cumulative risk scores were typically higher after July 2021, corresponding to longer and/or closer exposures. In late 2021, there was another wave of contacts that was partially due to the arrival of the Omicron variant in November 2021 (*36*). The daily number of exposure notifications received by app users closely tracked the number of contact events, with a delay of 3 to 6 days (figs. S7 and S10).

The app detected an estimated 271,000 transmission events in the period 1 April 2021 to 21 February 2022. Figure 4B shows that these events occurred in two main waves, as expected from the daily numbers and riskiness of contact events (Fig. 4A). Clear weekly patterns are visible and are explored further in Fig. 6.

In Fig. 5, we show the proportions of different types of contact and transmission events varying over time. Using median values for this period (1 April 2021 to 21 February 2022), household contacts accounted for only 6% of contacts but 39% of transmission events, whereas fleeting interactions accounted for 48% of contacts but only 12% of transmission events. During the lifting of social restrictions in April and May 2021, there was a gradual shift away from household contacts toward briefer contact events, with 8 to 19% of contacts from households in the first week of April compared with 2 to 4% by the last week of May. Because of the low number of transmission events in these months (Fig. 4B), the fractions of transmission events (Fig. 5B) are noisy. Late November and December 2021 saw the Christmas shopping season and the arrival of the Omicron variant; the fraction of contacts from each setting remained fairly stable (always with strong day-of-the-week effects), though with an increasing burden of transmissions from briefer contact events. Of the transmission events, 14 to 23% were from fleeting contacts in the first week of November, compared with 26 to 41% in December (excluding Christmas Day and Boxing Day, which are analyzed in more detail in the next section). Figure S14 shows how these results vary with geographical region, and fig. S15 shows the variability in the app's cumulative risk score by setting. Although absolute numbers of fleeting contacts were nine times higher than those of household contacts, cumulative risk scores of household contacts were 120 times higher than those of fleeting contacts.

Figure 6B shows day-of-the-week effects when disaggregating the transmissions by setting. Household transmissions were highest on Sundays, with 1.3 times as many transmissions as on Fridays. Fleeting transmissions peaked on Saturdays, at 2.6 times the level of transmissions on Mondays (Fig. 6B). The notable exception to these patterns was Christmas Day 2021, which was a Saturday.

Dynamics during the Christmas season

Socially, the Christmas season in England and Wales typically sees extensive shopping, work celebrations, and family reunions, finally followed by a quiet January. Similar patterns occur for Christmas and other celebrations in much of the world. The epidemiological impact of this is understudied.

Figure 7A shows that the contact rate *CR* increased during the 2021 pre-Christmas season and then decreased markedly around Christmas Day and through January 2022 to values at or below the mean value of 20 measured during lockdown during February to April 2021. We see a similar pattern for Christmas 2022 but with a faster recovery in January 2023. It is possible that caution and soft social restrictions

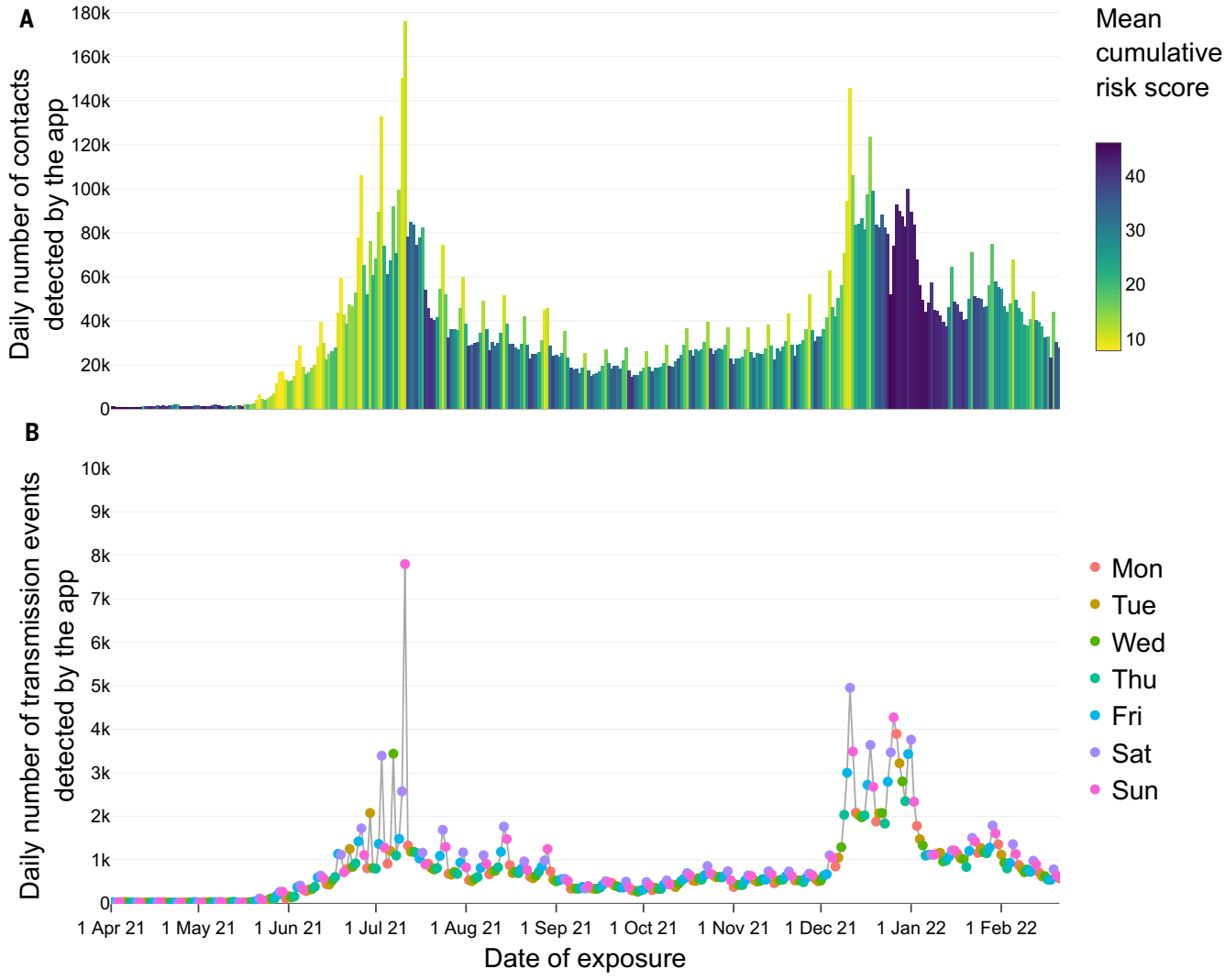


Fig. 4. Daily numbers of contacts (risky exposures) and estimated transmission events detected by the app, indexed by their date of exposure. (**A**) Contacts, colored by the mean of their cumulative risk scores. (**B**) Transmission events, colored by day of the week. Infection events are shown in fig. S11.

after the announcement of the arrival of the Omicron variant in late November 2021 (*32*), together with the high numbers of people self-isolating in January 2022, extended and intensified typical behavioral changes associated with the holiday season (annual leave, staying home in cold weather, etc.). Note that these two Christmas seasons differed in viral prevalence, testing behaviors, and app use (see the timeline in supplementary text section 5).

Figure 7B shows the fraction of contacts between app users in each setting for the 6-week period from 1 December 2021 to 11 January 2022 (top panel) and from 1 December 2022 to 11 January 2023 (bottom panel). Again, there were similar patterns across the 2 years, although the latter is more noisy owing to lower case numbers and app use (fig. S13). Christmas Day, and to a lesser extent the following few days, saw a shift toward household and recurring contacts and away from single-day and fleeting contacts. The median proportion of contacts classified as household for this 6-week period in 2021–2022 and 2022–2023 was 7 and 6%, respectively, but it was 18 and 16% on Christmas Day. Recurring contacts increased from a median of 15 to 23% on Christmas Day 2021 and from 12 to 27% in 2022.

Figure 7C shows the transmission events during the 6-week period from 1 December 2021 to 11 January 2022. Fifteen percent of these occurred on Christmas Day and the two preceding Saturdays, likely related to social and shopping habits. Although Christmas and Boxing Day 2021 had fewer contacts than in the preceding weeks (Fig. 7A), there was a local peak in transmission events on those days, particularly in households, which reached their overall peak of 1500 on Christmas (Figs. 6 and 7C). By contrast, transmission events from single-day or fleeting contacts were notably lower on Christmas Day than would be expected from their trends. There were 160 transmission events on Christmas Day from fleeting contacts, which was 80 to 83% lower than on the Saturdays either side. The shift on Sunday 26 December 2021 toward more brief encounters for contact and transmission events likely captures a propensity for more social mixing and the reopening of many shops for popular Boxing Day sales. A similar pattern can be observed for 2022, despite limitations in detecting transmissions beyond mid-2022 (fig. S16).

There was a notable peak in transmissions from recurring contacts on Friday 31 December 2021, a day when most workplaces were open and many social gatherings took place. Social gatherings on this day in Belgium were implicated in increased infection events (*40*). On Saturday 1 January 2022 (a Bank Holiday, when workplaces are typically shut), there was another local peak in household transmissions. The precise timings of the Christmas holiday

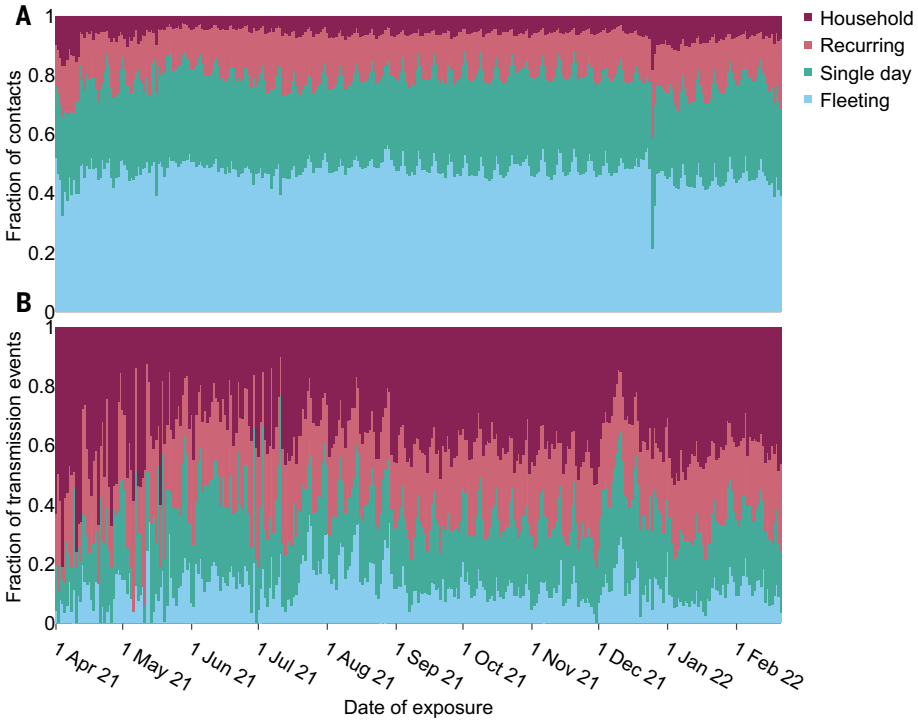


Fig. 5. Relative daily contribution of different settings. Daily fractions of contacts (**A**) and transmission events (**B**) detected by the app, classified by setting: household, recurring, single-day, or fleeting.

period varied across workplaces and schools, but it is common for schools to be closed for 2 to 3 weeks. Correspondingly, contact and transmission patterns returned to more normal levels within the first 2 weeks of January 2022.

Impact of the Euro 2020 football tournament

During the 2020 Union of European Football Associations (UEFA) European men's football championship, held from 11 June to 11 July 2021, there was a large increase in transmissions in the UK (*39*, *41*). The tournament saw the best performance by the England men's football team since 1966. It occurred several months into the roadmap for lifting restrictions and led to an increase in gatherings in pubs and homes. The app data reveal how the epidemic was shaped by this specific event.

Days on which the English or Welsh team played showed strong peaks in transmissions for app users registered as residing in England or Wales, respectively (Fig. 8A). This often disrupted the typical pattern of transmissions peaking on Saturdays (Fig. 4B). The peak for users residing in England was on Sunday 11 July 2021 when England played Italy in the tournament final: Transmission events were 6.1 and 8.7 times higher that day than on the previous and following Sundays, respectively. Of the 57,100 transmission events detected by the app during the 31 days of the tournament, 7800 (18%) of them occurred on the day of the final match. Overall, days with England matches recorded 1.1 to 6.6 times higher numbers of daily transmissions compared with a baseline rolling 15-day median, and the excess of transmissions on these match days accounted for 29% of all transmissions in the period. Only 4% of app users registered as residing in Wales, giving smaller numbers of transmission events (1.8% of those shown) but with similar patterns.

Match days had increased contact rates and increased probability of transmission and infection (Fig. 8B and fig. S17). For counts of contact and transmission events, as well as the probability of a transmission being reported through the app per contact, we calculated trend-corrected values by subtracting the overall trend and day of the week effects (see supplementary text). Increased contact rates and increased probability of transmission combined to cause the considerable peaks in transmissions.

The peaks in infections on match days were driven by an increasing burden from single-day and fleeting contacts. For England, 48% of the transmission events detected during match days were from single-day contacts and 21% from fleeting contacts, compared with 27 and 12%, respectively, for days without England matches (Fig. 5). This indicates that individuals who would not regularly come into contact were mixing more than usual at a level relevant for risk of infection (fig. S18).

The disaggregation by day and country strongly supports the hypothesis that the observed peaks in transmissions in this period were driven by social behaviors surrounding watching the matches (e.g., in homes, in pubs, at large public screenings, and in the Wembley stadium). A large burden of additional infections from Euro match days was also reported in a study across 12 countries (*42*).

Discussion

We analyzed anonymized data gathered by the NHS COVID-19 app in England and Wales in 2021–2023 and found that the app provided rich insights into the COVID-19 epidemic in addition to its primary role of digital contact tracing. Highly interesting among our results is an indicator $R_{app}(t)$ of the time-varying $R(t)$ that is disaggregated into multiplicative contributions from contact rates and from the probability of reported infection in those contacts (which we measured by *TPAEN*). This allowed us to see when changes in transmission levels could be attributed to changes in contact rates, such as at relaxation of restrictions, or to changes in probability of infection per contact, such as with the arrival of the Omicron variant. Furthermore, existing $R(t)$ measures are inferences drawn using a model of epidemic dynamics to link different times, commonly the renewal equation (*4*). $R_{app}(t)$, by contrast, requires no such model and is thus a more direct measure of per-person transmission.

$R_{app}(t)$ provided a reliable signal of epidemic change points and their direction, though not of the exact magnitude of $R(t)$. We underascertained infections among notified app users, relying on voluntary case reporting through the app; we also only considered transmissions from index cases before reporting of a positive test (at which point the app stopped recording exposures until after the set isolation period). Both points caused underestimation of $R(t)$ by a factor that changed with testing policies, app version, and isolation behavior. We found that when $R(t)$ from other sources was increasing and greater than 1, the corresponding upward trend was typically exaggerated in $R_{app}(t)$, which is apparent in a comparison with high-quality REACT-1 estimates (*43*) (fig. S19). This was because the background risk of infection from individuals not using the app increased during times of high prevalence, artificially driving up *TPAEN* in the calculation. This exaggeration was helpful for rapidly spotting an increase; for careful historic analysis of the impacts of interventions, slower, more accurate estimates of $R(t)$ are more appropriate (*5*). The 6-day lag to a stable estimate of $R_{app}(t)$ was consistent throughout the app's lifetime. As the pandemic progressed, publicly available estimates of $R(t)$ ceased or were published less often, in part because case reporting on the national dashboard was reduced from its original daily frequency (*44*).

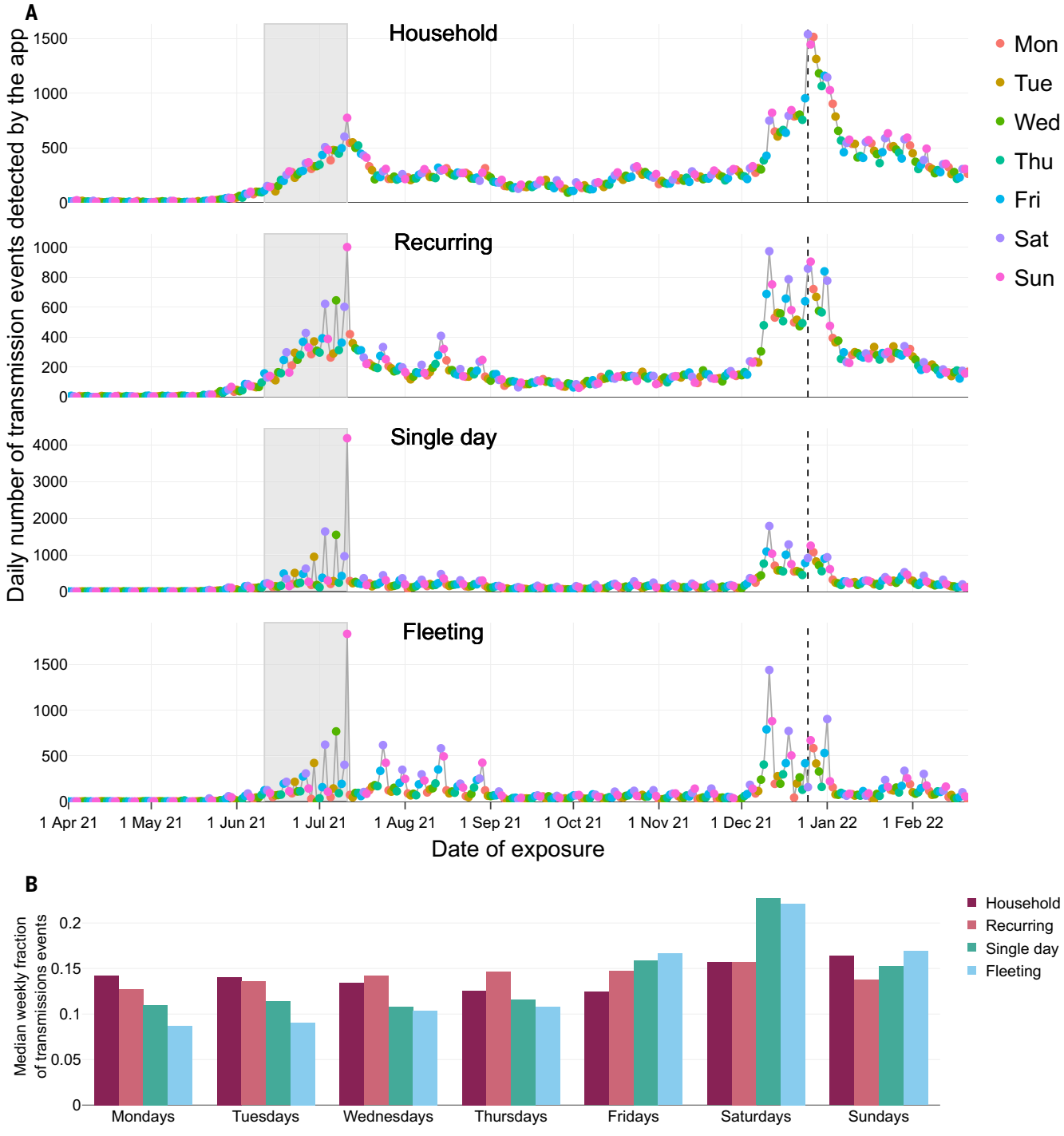


Fig. 6. Transmission events disaggregated by setting and day of the week. (**A**) The estimated number of transmission events detected by the app, disaggregated by setting, indexed by date of exposure, and colored by day of the week. The period of the Euro 2020 tournament is shaded gray; Christmas Day is annotated with a dashed line. Infection events are shown in fig. S12. (**B**) For each setting, the median fraction of weekly transmission events from exposures each day of the week. Each setting is normalized separately to sum to 1 over the 7 days.

Even with the best approach, there are intrinsic unavoidable delays in estimating $R(t)$. No method will ever determine contact rates of cases before the cases are reported, and no method will ever measure either secondary attack rates or $R(t)$ before the secondary infections are detected. For the COVID-19 epidemic in England and Wales, the delay for the app to detect exposures to confirmed cases was 3 to 5 days by polymerase chain reaction (PCR) and 2 to 3 days by rapid testing, and the delay from transmission to detection of secondary cases was 4 to 8 days by PCR and 3 to 6 days by rapid testing (fig. S7). Hence, $R(t)$ could not be estimated earlier than a week after the actual time of transmission.

We analyzed the app's daily analytics data in almost real time during the epidemic. We provided updates of many of the results presented here to the UK government and public health authorities with weekly frequency, and

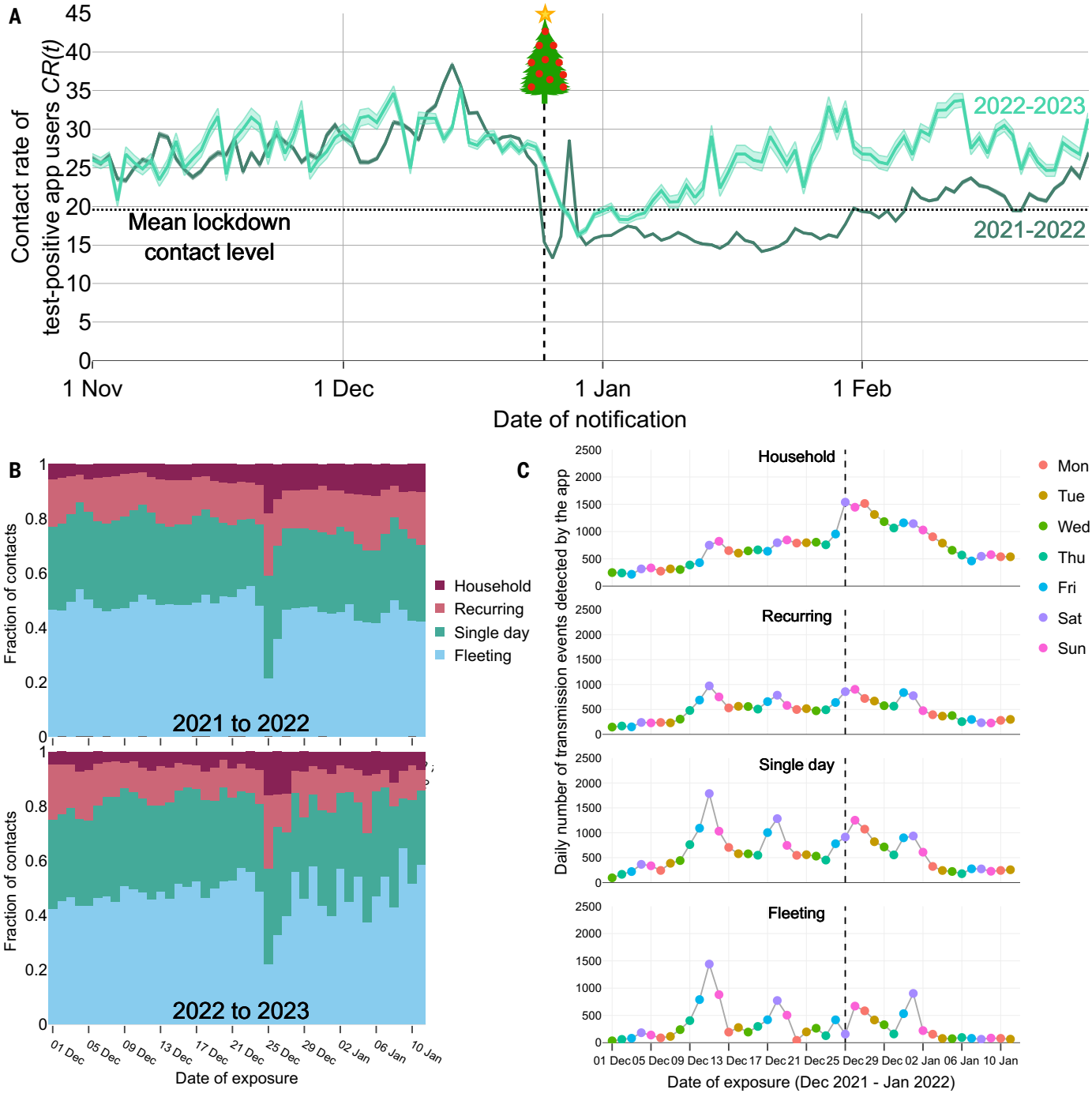


Fig. 7. Contact and transmission event patterns during Christmas holiday periods. Christmas Day is indicated by a vertical dashed line. (**A**) Contact rate *CR* for 1 November 2021 to 28 February 2022 and for 1 November 2022 to 28 February 2023. Shading around the lines shows the 95% confidence interval (this is narrow for most dates). The dotted line shows the mean level of contacts during lockdown during February to April 2021 for comparison. (**B**) Fraction of contacts by setting for 1 December 2021 to 11 January 2022 (top) and 1 December 2022 to 11 January 2023 (bottom). (**C**) Daily number of transmission events detected by the app, disaggregated by setting, for 1 December 2021 to 11 January 2022. Colors show days of the week.

at peak times daily, until the app's decommissioning in April 2023. Because they were based on a new data source, our reports were strongly caveated and were always used alongside other more established sources such as case counts and hospital admissions (fig. S20). However, in hindsight, signals from app data were notably reliable, even across major change points in testing policy. The app also enabled more localized insights, which were helpful for local decision-making, as tested on one NHS Integrated Care System (a collection of 10 lower-tier local authorities). At the level of single local authorities, we found that estimates of $R_{app}(t)$ were too noisy and inaccurate. The results obtained using events data are reported here for the first time; in future apps, such analyses could be conducted in real time.

The Euro 2020 tournament saw a large increase in case numbers; we resolved cases

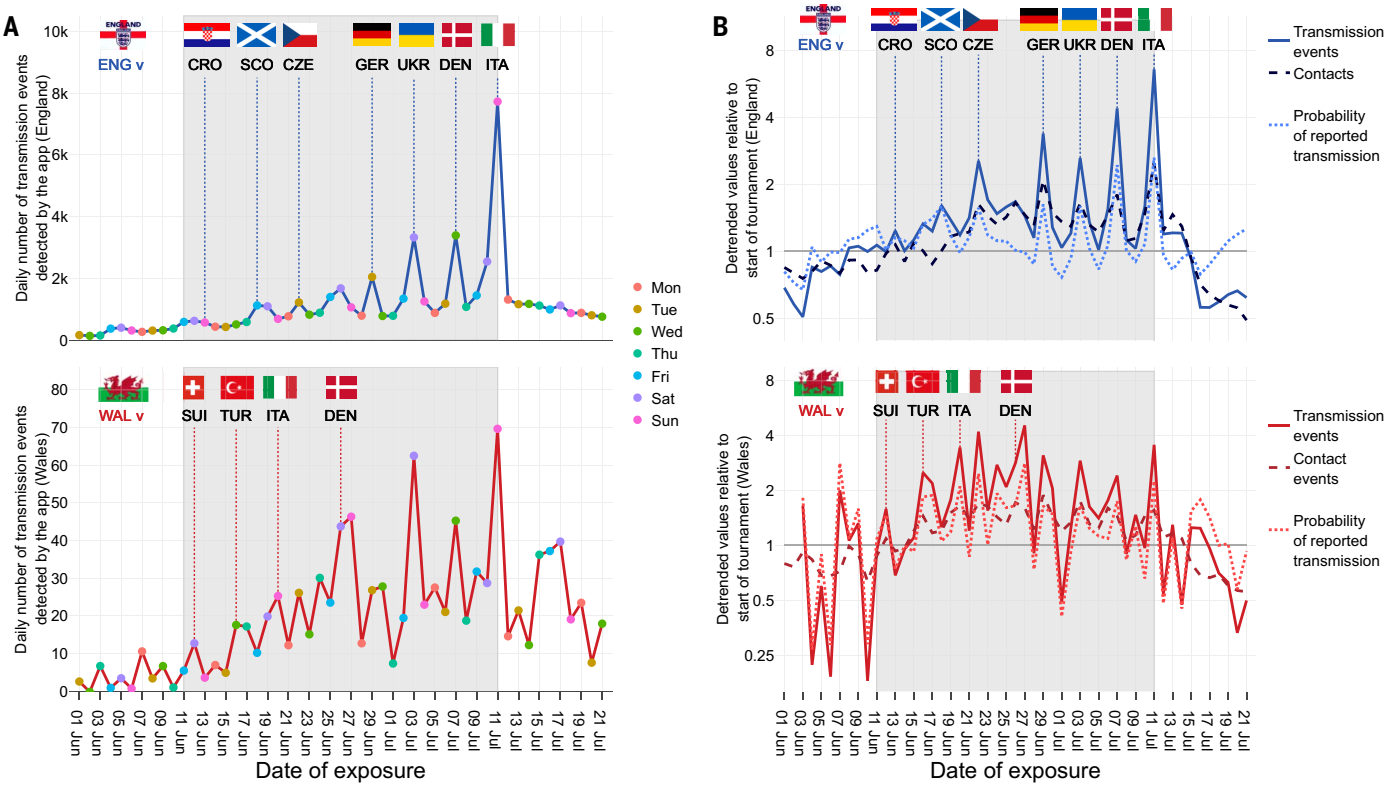


Fig. 8. Epidemiological indicators in England and Wales during the Euro 2020 tournament (11 June to 11 July 2021). The period 11 June to 11 July 2021 is shaded gray. Dashed vertical lines show match days for England and Wales, with flags and Fédération Internationale de Football Association (FIFA) country codes (*58*) over these lines showing the nation against whom they played on that day. (**A**) Daily number of transmission events colored by day of the week. Note the different y-axis scales between England and Wales, which reflect that 96% of app users registered an English postcode. (**B**) Trend-corrected values of transmission events, contact events, and the probability of reported transmission shown on a log scale. The trend was inferred through linear regression by date and day of the week, excluding match days.

reported through the app down to daily resolution in the time of risky exposure and found clear and strong peaks on the days of England matches for users in England, and likewise for Wales. We did not perform a causal analysis—estimating the difference between observations and a modeled counterfactual epidemic with no tournament—but the strength and resolution of the peaks, on a fairly smooth background, is strongly supportive of the match days having an impact. Transmissions on match days came mainly from single-day and fleeting contacts, indicating that individuals who would not regularly come into contact were mixing more than usual at a risk level relevant for infection. We speculate that this contributed to viral spread outside of typical contact networks. Compared with usual analyses of the dates on which cases were reported, our analysis of the dates on which cases were actually exposed to the virus enabled the high temporal resolution necessary to consider specific individual days such as those during Euro 2020, as well as Christmas and New Year's Eve.

The most prominent peaks in COVID-19 transmissions correspond to days associated with nationwide decentralized gatherings. Although small gatherings are seen as less risky than big events because of their smaller scale and reduced chance for superspreading, a high number of small decentralized gatherings involving a sizable fraction of the population may have a proportionally large impact on epidemic dynamics. Future studies should be devoted to large-scale monitoring and epidemiological understanding of these events.

Limitations to our analyses largely arose from the minimal, anonymized nature of the app data. No demographic data were collected, making it impossible to check the representativeness of app users as a (large) sample of the population and then to correct for biases. Data on age would have been particularly helpful, but this was not available, aside from users self-reporting being aged 16 or older. Thus, all results presented here concern only those aged 16 and older; other studies are needed for the contact and transmission patterns in those younger than 16 years and more generally for age-specific patterns. Also, our geographical data (postcode district of residence) were self-declared only once by users when they installed the app. These data were more likely to reflect location earlier in the life of the app, when users were less likely to have changed address since app installation and when heavier restrictions limited travel. It was particularly unlikely to be representative for users who traveled and/or changed address frequently, for example, commuters or students studying at a university in a different region from their home address. In our regional analysis of contact rates (Fig. 3), we implicitly assumed that the user testing positive and the user(s) receiving the resulting notifications were registered in the same region. These biases do not seem to have a major impact on our analyses, as illustrated by the reliability of our indicator of $R(t)$ and by the robustness of the results with respect to geographical reweighting of the data to correct for spatially variable sampling biases (fig. S21). Another example of the limitations related to minimal data is that we could calculate the average value each day for the number of notifications per test-positive app user, but not the dispersion around this average, related to superspreading. Finally, our estimates for the number of transmissions assume that the relative background risk of infection for notified individuals compared with other app users is

approximately constant in time. This assumption may not hold for episodic decentralized gatherings because both the background risk and the chance of being notified are higher for individuals attending these gatherings; this could lead to an overestimation of the number of daily transmissions at these events.

The lack of contextual data from the app made our metric of contact rates less rich than existing metrics of mixing and mobility provided by the CoMix study (*30*) and Google Mobility (*29*). These measures included data such as user age, vaccination status, whether contacts were inside or outside, and the type of venue visited (e.g., park, retail, etc.). The representativeness of app users is unknown, although some insights are available from survey data (*38*). Daily sample sizes derived from app data are likely to be considerably smaller than those of Google Mobility (composed of Google account holders with location history turned on) but larger than those of the CoMix study, which had approximately 1000 participants in each round. A distinctive epidemiological quality of our metric is that it is specifically for a subpopulation of individuals who tested positive for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), specifically for the period when they were infectious. Furthermore, it was available quickly, with a 1-day lag from the date of notification. Google Mobility reports were published regularly using data up to 2 to 3 days before the date of publication; CoMix reports were published weekly or fortnightly, using data from up to 7 to 9 days before the date of publication.

We have demonstrated that digital contact tracing for SARS-CoV-2 can provide rich insights into epidemic dynamics with unprecedented time resolution, in addition to its primary purpose of reducing transmission. When decisions must be made quickly, evidence must be available quickly, and digital contact-tracing technologies have strong potential to support this real-time aspect of public health. In the future, similar apps could inform policy decisions with data on epidemic drivers, the impacts of interventions, population behaviors, regional variations, and near-term forecasting. However, there is an urgent need to learn lessons for improvement while individual and institutional memories are fresh. Digital tools are likely to play an even larger role in the next pandemic, given their unparalleled scalability and the insight they can provide for precision public health. Frameworks for their implementation should be agreed on in advance with stakeholders, based on robust evaluation of their potential impacts across public health and wider society. We should continue developing the associated infrastructure, regulatory framework, and analytic capabilities in preparation for future epidemics and pandemics.

Materials and methods

The app used the Google Apple Exposure Notification framework (*45*) and maintained a strong focus on user privacy throughout its operation (*46*). Counterbalancing this, the UK Information Commissioner required enough data to be collected for monitoring and evaluation (within an overarching framework of data minimalism). For this reason, the reconstruction of contact and infection events is inferential and complex. Throughout the app's operation, there were two streams of data collected: analytics data and event data. We described these data in detail previously (*22*, *23*, *37*); here, we provide a brief summary and more detail about the additional methods used in this manuscript.

We described the app's daily analytics data primarily in (*22*) and recapped in (*23*, *37*). The data were composed of "'analytics packets," which were sent to a central server once each day by each correctly functioning installation of the app on a mobile device that had access to the internet. We defined the total number of analytics packets each day to be the number of active users that day. Each packet indicated whether on that day the user was notified, had been recently notified (specifically, whether they were in a recommended isolation period or the 14 days after it), and/or registered a positive test result. Also included were up to four individual characteristics: device model, operating system, self-reported postcode district, and, in some later app versions, self-reported lower-tier local authority.

We described the app's event data in (*37*). These data were composed of "event packets" sent to the central server whenever an app user was notified of a risky exposure to a confirmed case. Each event packet summarized an exposure window of 30 min or less spent in close contact between the notified user and the associated index case, only if it had a risk score over the threshold for notification. If a pair of users spent more than 30 min in close contact, this was recorded as separate 30-min exposure windows. The event packets included information on proximity, duration, and date of exposure, as well as the same four individual characteristics as in the analytics packets. If a notified individual reported a positive test in the app during the observation interval, which started with their notification and ended 14 days after the exposure, the same event packets that were sent when the individual was notified were sent once more to the central server with a flag indicating that this was the report-positive stage not the notification stage. This allowed us to see which risky exposure events were followed by the notified contact reporting a positive test.

As described in (*37*), only 59% of event packets could be grouped together such that each group represented the set of all risky exposure windows recorded by the app for a given notified contact. This is because event packets did not contain a unique identifier for the app user who sent them. We filtered the analytics data to instances where only one app user with a distinct combination of individual characteristics was notified on a given day. In those instances, all event packets sent on that day with those individual characteristics could be assumed to come from one app user, and they were grouped together to give a combined description of a single exposure split over multiple 30-min windows. In (*37*), we only used analytics data to perform that filtering; here, because we also present results from analytics data, we scaled the counts we obtained in events data for consistency with corresponding counts in the analytics data. Explicitly, the number of contact events on a given day is determined from the number of analytics packets reporting notification that day, and because the number of groups we reconstructed from event packets that day was less than this (by a factor of about 60% on average over all days), we scaled up the number of groups to match it.

The availability of app data varied over its lifetime. The app was launched on 24 September 2020 and decommissioned on 27 April 2023 (*47*). The decommissioning was announced on 28 March 2023, and there was a corresponding decrease in active app use and engagement. We could conduct detailed analysis of analytics data from 1 February 2021 until 15 March 2023 and of joint analytics and event data from 1 April 2021 to 21 February 2022. Full details of the challenges posed by the data outside of those time windows are given in (*23*) and (*37*), respectively. They are due to app updates and changes in app use, test availability, and viral prevalence. In the supplementary materials, we also present some noisier results from exposure data later in 2022.

We previously (*23*) presented the daily number of notifications per positive test reported in the app. In this work, to develop this into a measure of contact rates, we divided by app uptake (number of active users divided by total population) and by the proportion of users who, on reporting a positive test via the app, then consented to contact tracing. (Consent was never automatic, and test-positive app users who did not consent did not trigger any notifications.) Thus, we calculated

$$CR = \text{notifications}(t)/[\text{positives}(t) \times \text{uptake}(t) \times \text{proportion consenting to tracing}(t)]$$

and the corresponding Poisson confidence intervals based on a likelihood ratio test. The app-based indicator $R_{\text{app}}(t)$ is then

$$R_{app}(t) = CR \times TPAEN$$

where we modeled *TPAEN* as a natural cubic spline with weekly knots, calculated by combining the number of daily notifications, the

delay between notification and reported positive tests, and the number of daily positive cases reported after notification, as detailed in (*23*).

To calculate the contributions of changes in *CR* and *TPAEN* to changes in $R_{app}(t)$ (Fig. 1D) we took the sum of the logarithms of the changes in each of the component measures, which describes $\log[R(t)]$ up to a constant:

$$\log[R_{app}(t)] = \log[CR/\text{median}(CR)] + \log[TPAEN/\text{median}(TPAEN)] + \text{constant}$$

When comparing the app-based indicator $R_{app}(t)$ to other measures of $R(t)$, we selected measures for which values were made publicly available for a period of at least a year. We plotted each from their first to last available dates. For the EpiNow2 estimates, each publication on GitHub provided estimates for the previous 12 weeks (*48*). We used the git history to extract estimates as published every 10 weeks, to the nearest available publication date.

For event data, we followed the standard approach to remove background infections that assumes that the background risk for an individual is proportional to the local rate of positive tests reported through the app in the same period (*37*) and refined it to account for potential day-to-day heterogeneities in transmissibility. In particular, we implemented a log-binomial model as in (*37*), but here we also assumed a linear relation between transmission risk and cumulative risk score, a result shown empirically in (*37*) for exposures of cumulative duration of less than 3 hours. The log-binomial model for individual i in lower-tier local authority x_i notified on date n_i and with date of a randomly selected exposure window t_i has the form

$$TPAEN_i = 1 - \exp\{\beta \times \log[1 - \text{app reported positive test rate}(x_i, n_i)] - \tau(t_i) \times \text{cumulative risk score}_i\}$$

where $\tau(t_i)$ is a time-dependent transmissibility factor (see supplementary text section 2).

We estimated the background risk coefficient β separately for fleeting contacts (exposed during a single 30-min risky window) and for all contacts spanning two to six risky exposure windows. We then used these two β values to estimate the background infection risk for contacts with one window or at least two windows, respectively, calculating individual *i*'s background risk as the above expression with cumulative risk score set to zero. Individuals reporting a positive test contributed 1 minus their background risk to the total count of transmission events; individuals not reporting a positive test contributed minus their background risk. In this way, our total counts for transmission events are less than our total counts for reported infections by the amount that we attributed to the background risk, leaving an amount that we attribute to the exposures recorded by the app. In case of negative count for transmission events for some category, we set the final estimate for that category to zero. The result is illustrated in fig. S22 for the Euro 2020 period.

For contact events and transmission events, we used a single date to be representative of the exposure, for example, to show counts of these events by date. For exposures that took place over multiple days, we assigned a date stochastically with a probability of choosing each day equal to the fraction of exposure windows that were on that day (a close proxy for the fraction of the overall exposure duration that was on that day). When summing many transmission events, the counts for representative dates become the likely dates of transmission. For example, if there were many transmission events, all with 60% of their exposure duration on day 1 and 40% on day 2, 60% of them would have day 1 assigned and 40% of them day 2. Though the assignment of day is stochastic at the individual level, the counts for many individuals (as present in this dataset) reflect our expectation at the population level of when the transmissions occurred (under a model where risk remains proportional to exposure duration, i.e., neglecting risk saturation for very long exposures). Statistics for the duration of different types of exposures are presented in fig. S23 and table S1.

We conducted our analyses in R (*49*), with particular use of the packages data.table (*50*), leaflet (*51*), plotly (*52*), RAthena (*53*), and tidyverse (*54*).

REFERENCES AND NOTES

1. World Health Organization, WHO Director-General's opening remarks at the media briefing on COVID-19 - 16 March 2020 (2020); https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—16-march-2020.
2. R. M. Anderson, R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ. Press, 1991). doi: 10.1093/oso/9780198545996.001.0001
3. J. Wallinga, P. Teunis, Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am. J. Epidemiol.* **160**, 509–516 (2004). doi: 10.1093/aje/kwh255; pmid: 15353409
4. C. Fraser, Estimating individual and household reproduction numbers in an emerging epidemic. *PLOS ONE* **2**, e758 (2007). doi: 10.1371/journal.pone.0000758; pmid: 17712406
5. K. M. Gostic *et al.*, Practical considerations for measuring the effective reproductive number, Rt. *PLOS Comput. Biol.* **16**, e1008409 (2020). doi: 10.1371/journal.pcbi.1008409; pmid: 33301457
6. H. Vöhringer, M. Sinnott, T. Sanderson, R. Amato, I. Martincorena, D. Kwiatkowski, J. C. Barrett, M. Gerstung, GenomicSurveillance. Github (2021); https://github.com/sagar87/genomicsurveillance.
7. C. C. Kerr *et al.*, Covasim: An agent-based model of COVID-19 dynamics and interventions. *PLOS Comput. Biol.* **17**, e1009149 (2021). doi: 10.1371/journal.pcbi.1009149; pmid: 34310589
8. C. E. Overton *et al.*, EpiBeds: Data informed modelling of the COVID-19 hospital burden in England. *PLOS Comput. Biol.* **18**, e1010406 (2022). doi: 10.1371/journal.pcbi.1010406; pmid: 36067224
9. J. A. Scott, A. Gandy, S. Mishra, J. Unwin, S. Flaxman, S. Bhatt, epidemia: Modeling of Epidemics using Hierarchical Bayesian Models. R package version 1.0.0 (2020); https://imperialcollegelondon.github.io/epidemia/index.html.
10. S. Abbott *et al.*, Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. *Wellcome Open Res.* **5**, 112 (2020). doi: 10.12688/wellcomeopenres.16006.1
11. M. Kendall *et al.*, LocalCovidTracker, Github (2020); https://github.com/BDI-pathogens/LocalCovidTracker [accessed 10 October 2023].
12. OxfordCSML group, UK Local Covid Map; https://localcovid.info/ [accessed 10 October 2023].
13. R. Hinch *et al.*, OpenABM-Covid19-An agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing. *PLOS Comput. Biol.* **17**, e1009146 (2021). doi: 10.1371/journal.pcbi.1009146; pmid: 34252083
14. MRC Biostatistics Unit, Nowcasting and Forecasting of the COVID-19 Pandemic; https://hprubse.nihr.ac.uk/research/covid-19-research/nowcasting-and-forecasting-of-the-covid-19-pandemic/ [accessed 10 October 2023].
15. UK Government, The R value and growth rate (2020); https://www.gov.uk/guidance/the-r-value-and-growth-rate#full-publication-update-history [accessed 10 October 2023].
16. L. Ferretti *et al.*, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, eabb6936 (2020). doi: 10.1126/science.abb6936; pmid: 32234805
17. M. E. Kretzschmar *et al.*, Impact of delays on effectiveness of contact tracing strategies for COVID-19: A modelling study. *Lancet Public Health* **5**, e452–e459 (2020). doi: 10.1016/S2468-2667(20)30157-2; pmid: 32682487
18. A. J. Kucharski *et al.*, Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: A mathematical modelling study. *Lancet Infect. Dis.* **20**, 1151–1160 (2020). doi: 10.1016/S1473-3099(20)30457-6; pmid: 32559451
19. LibertiesEU, COVID-19 contact tracing apps in the EU (2021); https://www.liberties.eu/en/stories/trackerhub1-mainpage/43437.
20. C. Nebeker *et al.*, Digital exposure notification tools: A global landscape analysis. *PLOS Digit. Health* **2**, e0000287 (2023). doi: 10.1371/journal.pdig.0000287; pmid: 37656671
21. K. Jenniskens *et al.*, Effectiveness of contact tracing apps for SARS-CoV-2: An updated systematic review. *F1000 Res.* **11**, 515 (2022). doi: 10.12688/f1000research.110668.1
22. C. Wymant *et al.*, The epidemiological impact of the NHS COVID-19 app. *Nature* **594**, 408–412 (2021). doi: 10.1038/s41586-021-03606-z; pmid: 33979832
23. M. Kendall *et al.*, Epidemiological impacts of the NHS COVID-19 app in England and Wales throughout its first year. *Nat. Commun.* **14**, 858 (2023). doi: 10.1038/s41467-023-36495-z; pmid: 36813770
24. F. Vogt, B. Haire, L. Selvey, A. L. Katelaris, J. Kaldor, Effectiveness evaluation of digital contact tracing for COVID-19 in New South Wales, Australia. *Lancet Public Health* **7**, e250–e258 (2022). doi: 10.1016/S2468-2667(22)00010-X; pmid: 35131045
25. M. Salathé *et al.*, Early evidence of effectiveness of digital contact tracing for SARS-CoV-2 in Switzerland. *Swiss Med. Wkly.* **150**, w20457 (2020). doi: 10.4414/smw.2020.20457; pmid: 33327003
26. T. Ballouz *et al.*, Individual-level evaluation of the exposure notification cascade in the SwissCovid Digital Proximity Tracing app: Observational study. *JMIR Public Health Surveill.* **8**, e35653 (2022). doi: 10.2196/35653; pmid: 35476726
27. E. Aronoff-Spencer *et al.*, Exposure notification system activity as a leading indicator for SARS-COV-2 caseload forecasting. *PLOS ONE* **18**, e0287368 (2023). doi: 10.1371/journal.pone.0287368; pmid: 37594936
28. J. Mellor *et al.*, Understanding the leading indicators of hospital admissions from COVID-19 across successive waves in the UK. arXiv.2303.12037 [stat.AP] (2023).
29. Google, COVID-19 community mobility reports; https://www.google.com/covid19/mobility/ [accessed 10 October 2023].
30. C. I. Jarvis *et al.*, CoMix study – Social contact survey in the UK. CMMID Repository (2020); https://cmmid.github.io/topics/covid19/comix-reports.html.
31. UK Government, COVID-19 response – Spring 2021 (summary) (2021); https://www.gov.uk/government/publications/covid-19-response-spring-2021/covid-19-response-spring-2021-summary.
32. Wikipedia, Timeline of the COVID-19 pandemic in England (2021); https://en.wikipedia.org/wiki/Timeline_of_the_COVID-19_pandemic_in_England_(2021) [accessed 10 October 2023].
33. P. V. Markov *et al.*, The evolution of SARS-CoV-2. *Nat. Rev. Microbiol.* **21**, 361–379 (2023). doi: 10.1038/s41579-023-00878-2; pmid: 37020110
34. E. Volz, Fitness, growth and transmissibility of SARS-CoV-2 genetic variants. *Nat. Rev. Genet.* **24**, 724–734 (2023). doi: 10.1038/s41576-023-00610-z; pmid: 37328556

35. F. P. Lyngse *et al*., Household transmission of the SARS-CoV-2 Omicron variant in Denmark. *Nat. Commun.* **13**, 5573 (2022). doi: 10.1038/s41467-022-33328-3; pmid: 36151099
36. UK Government, COVID-19 variants: Genomically confirmed case numbers (2021); https://www.gov.uk/government/publications/covid-19-variants-genomically-confirmed-case-numbers#full-publication-update-history [accessed 10 October 2023].
37. L. Ferretti *et al*., Digital measurement of SARS-CoV-2 transmission risk from 7 million contacts. *Nature* **626**, 145–150 (2024). doi: 10.1038/s41586-023-06952-2; pmid: 38122820
38. M. Al-Haboubi, J. Exley, K. Allel, B. Erens, N. Mays, One year of digital contact tracing: Who was more likely to install the NHS COVID-19 app? Results from a tracker survey in England and Wales. *Digit. Health* **9**, 1–18 (2023). doi: 10.1177/20552076231159449
39. UK Health Security Agency, UKHSA data dashboard; https://ukhsa-dashboard.data.gov.uk/ [accessed 10 October 2023].
40. C. Geenen *et al*., Unravelling the effect of New Year's Eve celebrations on SARS-CoV-2 transmission. *Sci. Rep.* **13**, 22195 (2023). doi: 10.1038/s41598-023-49678-x; pmid: 38097713
41. Office for National Statistics, Coronavirus (COVID-19) Infection Survey, UK statistical bulletins (2020); https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/ conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/previousReleases [accessed 10 October 2023].
42. J. Dehning *et al*., Impact of the Euro 2020 championship on the spread of COVID-19. *Nat. Commun.* **14**, 122 (2023). doi: 10.1038/s41467-022-35512-x; pmid: 36653337
43. O. Eales *et al*., Trends in SARS-CoV-2 infection prevalence during England's roadmap out of lockdown, January to July 2021. *PLOS Comput. Biol.* **18**, e1010724 (2022). doi: 10.1371/journal.pcbi.1010724; pmid: 36417468
44. UK Government, What's new: Coronavirus (COVID-19) in the UK; https://ukhsa-dashboard.data.gov.uk/whats-new [accessed 10 October 2023].
45. Google, Apple Google Exposure Notifications: Using technology to help public health authorities fight COVID 19 (2020); https://www.google.com/covid19/exposurenotifications/ [accessed 10 October 2023].
46. UK Health Security Agency, NHS COVID-19 app: Your data and privacy (2022); https://www.gov.uk/guidance/nhs-covid-19-app-your-data-and-privacy [accessed 10 October 2023].
47. UK Health Security Agency, NHS COVID-19 app statistics (2022); https://www.gov.uk/government/publications/nhs-covid-19-app-statistics [accessed 10 October 2023].
48. London School of Hygiene & Tropical Medicine, epiforecasts/covid-rt-estimates. Github (2020); https://github.com/epiforecasts/covid-rt-estimates/blob/master/national/cases/summary/rt.csv [accessed 10 October 2023].
49. R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing (2021); https://www.R-project.org/.
50. M. Dowle, A. Srinivasan, data.table: Extension of 'data.frame.' R package version 1.14.8 (2021); https://CRAN.R-project.org/package=data.table [accessed 10 October 2023].
51. J. Cheng, B. Karambelkar, Y. Xie, leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library. R package version 2.2.0 (2022); https://CRAN.R-project.org/package=leaflet [accessed 10 October 2023].
52. C. Sievert, Interactive web-based data visualization with R, plotly, and shiny. R package version 4.10.2 (2020); https://cran.r-project.org/web/packages/plotly/index.html [accessed 10 October 2023].
53. D. Jones, RAthena: Connect to 'AWS Athena' using 'Boto3' ('DBI' Interface). R package version 2.6.1 (2022); https://github.com/DyfanJones/RAthena [accessed 10 October 2023].
54. H. Wickham *et al*., Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019). doi: 10.21105/joss.01686
55. StatsWales, Reproduction (R) number (2020); https://statswales.gov.wales/Catalogue/Health-and-Social-Care/coronavirus-covid-19/reproduction-r-number [accessed 10 October 2023].
56. H. Manley *et al*., Combining models to generate a consensus effective reproduction number R for the COVID-19 epidemic status in England. medRxiv 2023.02.27.23286501 [Preprint] (2023); doi: 10.1101/2023.02.27.23286501
57. M. Kendall, UK recent R estimate. Github (2021); https://github.com/MichelleKendall/UK_recent_R_estimate [accessed 10 October 2023].
58. Wikipedia, List of FIFA country codes; https://en.wikipedia.org/wiki/List_of_FIFA_country_codes [accessed 10 October 2023].
59. UK Government, Accessing UKHSA protected data: How to apply for access to protected data from UKHSA; https://www.gov.uk/government/publications/accessing-ukhsa-protected-data.
60. M. Kendall, MichelleKendall/drivers_of_epi_dynamics_from_nhs_covid19_app_measurements: Initial release (v1.0.0). Zenodo (2022); https://doi.org/10.5281/zenodo.11625147.
61. M. Kendall, ukhsa-collaboration/drivers_of_epi_dynamics_from_nhs_covid19_app_measurements. Github (2024); https://github.com/ukhsa-collaboration/drivers_epidemic_dynamics_from_nhs_covid19_app.

ACKNOWLEDGMENTS

We are grateful for the help and support from teams across UKHSA, Zühlke, and previously at NHS Test and Trace. In particular, we thank the NHS COVID-19 app data and analytics team for their invaluable support with data access, management, and analytics. **Funding:** This work was funded by a Li Ka Shing Foundation award and research grant funding from the UK Department of Health and Social Care (DHSC), both to C.F., and by the National Institute for Health and Care Research to the Health Protection Research Unit in Genomics and Enabling Data, grant number NIHR200892, to M.K. and X.D. C.W. was funded by the Wellcome Trust (collaboration award 206298/Z/17/Z, ARTIC-Network). **Author contributions:** All authors contributed to the conceptualizing, writing, and reviewing of this manuscript. M.K., L.F., C.W., J.P., and C.F. developed the methods. M.K., L.F., C.W., D.T., and C.F. performed the analyses. **Competing interests:** The views expressed in this article are those of the author(s) and are not necessarily those of UK Health Security Agency (UKHSA) or the Department of Health and Social Care (DHSC). M.K. has a data-sharing agreement with UKHSA. D.T. was an employee of Zühlke, which provided consultancy to UKHSA. C.W., L.F., and C.F. were named researchers on a grant from DHSC to Oxford University. A.D.F., A.L., H.M., and J.P.-G. were employees of UKHSA. J.P. has an equity interest in WeHealth Solutions PBC, which distributed an exposure notification app to Arizona and Bermuda. F.D.L, L.A.-D., and X.D. declare no competing interests. **Data and materials availability:** Data access is managed by UKHSA, which will make available on request (*59*) the anonymized personal data needed to replicate the key results. Access is controlled for privacy reasons (dataaccess@ukhsa.gov.uk). Code to replicate the analysis is available on Zenodo (*60*) and GitHub (*61*). **License information:** Copyright © 2024 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse. In the interest of rapid dissemination of results with immediate public health relevance and because this research was funded in whole or in part by the Wellcome Trust (206298/Z/17/Z), a cOAlition S organization, the author will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adm8103
Supplementary Text
Figs. S1 to S23
Table S1
Reference (*62*)
MDAR Reproducibility Checklist
Data S1 to S3

Submitted 8 November 2023; accepted 19 June 2024
Published online 11 July 2024
10.1126/science.adm8103