

# Drivers of SARS-CoV-2 testing behaviour: a modelling study using nationwide testing data in England

Received: 29 October 2022

Accepted: 30 March 2023

Published online: 14 April 2023

 Check for updatesYounjung Kim<sup>1</sup>, Christl A. Donnelly<sup>2,3,4</sup> & Pierre Nouvellet<sup>1,4</sup> 

During the COVID-19 pandemic, national testing programmes were conducted worldwide on unprecedented scales. While testing behaviour is generally recognised as dynamic and complex, current literature demonstrating and quantifying such relationships is scarce, despite its importance for infectious disease surveillance and control. Here, we characterise the impacts of SARS-CoV-2 transmission, disease susceptibility/severity, risk perception, and public health measures on SARS-CoV-2 PCR testing behaviour in England over 20 months of the pandemic, by linking testing trends to underlying epidemic trends and contextual meta-data within a systematic conceptual framework. The best-fitting model describing SARS-CoV-2 PCR testing behaviour explained close to 80% of the total deviance in NHS test data. Testing behaviour showed complex associations with factors reflecting transmission level, disease susceptibility/severity (e.g. age, dominant variant, and vaccination), public health measures (e.g. testing strategies and lockdown), and associated changes in risk perception, varying throughout the pandemic and differing between infected and non-infected people.

During the COVID-19 pandemic, national SARS-CoV-2 testing programmes were conducted on an unprecedented scale across the world<sup>1</sup>. They played a significant role in monitoring and controlling SARS-CoV-2 transmission and mitigating its impact on public health. In most countries, contrasting with most testing for other infectious diseases, the majority of SARS-CoV-2 tests were conducted on a voluntary basis, making the interpretation of testing data challenging<sup>2–4</sup>. This highlights the importance of understanding factors driving testing behaviour<sup>5,6</sup> in SARS-CoV-2 surveillance and control.

Currently, most studies interested in SARS-CoV-2 testing behaviour attempted either to characterise the willingness to be tested based on cross-sectional survey data<sup>7–9</sup> or explored factors associated with actual SARS-CoV-2 testing behaviour based on longitudinal participatory survey data<sup>10,11</sup>. While providing a wealth of information at the individual level, as the authors acknowledged, such studies often suffered from selection bias, as survey participation is likely correlated

to actual testing behaviour. Additionally, none of the studies accounted for fast-changing epidemiological contexts, such as transmission level, variant emergence, and vaccine rollout among other public health measures, although they could potentially bias study results when associated with variables of interest. In fact, throughout the pandemic, according to SARS-CoV-2 surveillance studies, most observed variability in positivity rate has been attributed to increased viral transmission within certain demographic groups<sup>12,13</sup>, which might have also led to heterogeneities in testing behaviour over time. Potentially due to this issue, the studies based on longitudinal survey data covered only a limited period (i.e. around 5 and 2 months, respectively)<sup>10,11</sup>. This led to a static view of factors influencing testing behaviour, although it is likely to have changed dynamically throughout the pandemic.

The above limitations suggest the need for a new methodological approach to better understand testing behaviour for COVID-19 or

<sup>1</sup>Department of Evolution, Behaviour, and Environment, School of Life Sciences, University of Sussex, Brighton, UK. <sup>2</sup>Department of Statistics, University of Oxford, Oxford, UK. <sup>3</sup>Pandemic Sciences Institute, University of Oxford, Oxford, UK. <sup>4</sup>MRC Centre for Global Infectious Disease Analysis and Abdul Latif Jameel Institute for Disease and Emergency Analytics, Imperial College London, London, UK. ✉ e-mail: [pierre.nouvellet@sussex.ac.uk](mailto:pierre.nouvellet@sussex.ac.uk)

other diseases requiring public health responses at the regional, national, or global level. Importantly, this methodological approach would require a conceptual framework that systematically describes mechanisms underpinning testing behaviour and disentangles their interrelationships while accounting for the underlying epidemic trends. Therefore, we propose a conceptual framework for SARS-CoV-2 testing behaviour upon which this study is built (Fig. 1). In this framework, the level of viral transmission is the primary driver for SARS-CoV-2 testing with higher incidence leading to more testing<sup>14</sup>. Then, at a given epidemic point, individuals' testing probability would be mainly determined by a combined effect of the following three mechanisms: (i) disease susceptibility/severity<sup>5,7,15–20</sup>, (ii) risk perception<sup>5,7,15–17</sup>, and (iii) public health measures<sup>21</sup>. Our conceptual framework proposes that those three mechanisms are interrelated with one another (Fig. 1a) and could act differentially on people's testing behaviour depending on their infection status (Fig. 1b). First, disease susceptibility/severity from SARS-CoV-2 infection would influence testing propensity only among infected people, with experiencing more severe symptoms increasing testing propensity. Second, public health measures, such as testing capacity, testing strategies, and social distancing, would affect testing propensity equally or differentially among infected and non-infected people depending on their target population. Finally, based on various factors, including awareness of disease susceptibility/severity and public health measures in place, people would perceive the risk of SARS-CoV-2 infection and its consequences at different levels. This risk perception would equally impact the testing propensity of infected and non-infected people, considering that most people take a SARS-CoV-2 test without knowing their infection status at the time of testing.

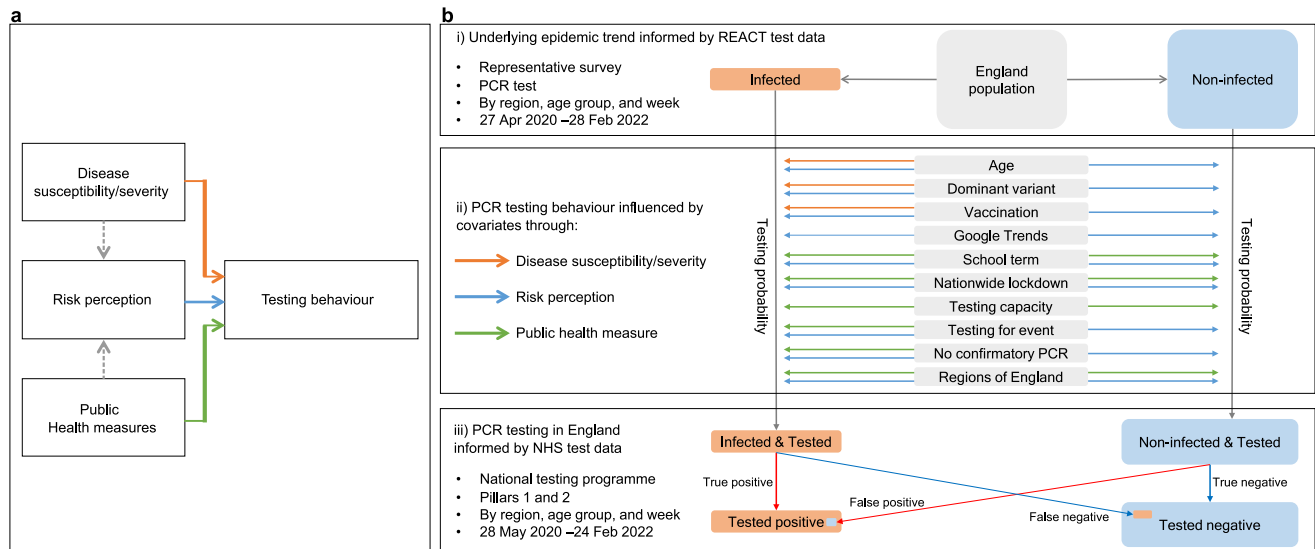
Here, based on the above conceptual framework for SARS-CoV-2 testing, we aim to understand factors influencing age-specific, spatial, and temporal trends in SARS-CoV-2 testing in England between 28 May 2020 and 24 February 2022. More specifically, using publicly available population-level data for England disaggregated by regions and age groups over 92 weeks, our study seeks to link PCR test results from the National Health Service (NHS) Test and Trace<sup>22</sup> (hereafter NHS test

data) to testing behaviour within dynamic models. Our models characterise testing behaviour separately for infected (defined as SARS-CoV-2 shedding as determined by PCR) and non-infected (defined as no SARS-CoV-2 shedding as determined by PCR) people based on age- and region-specific level of SARS-CoV-2 circulation from a continuously running representative PCR survey (i.e. REal-time Assessment of Community Transmission (REACT) study, hereafter REACT test data)<sup>23</sup>. This modelling approach allows us to explicitly describe the process of seeking and receiving a test in relation to various contextual meta-data related to disease susceptibility/severity, risk perception, and public health measures, accounting for underlying epidemic trends.

## Results

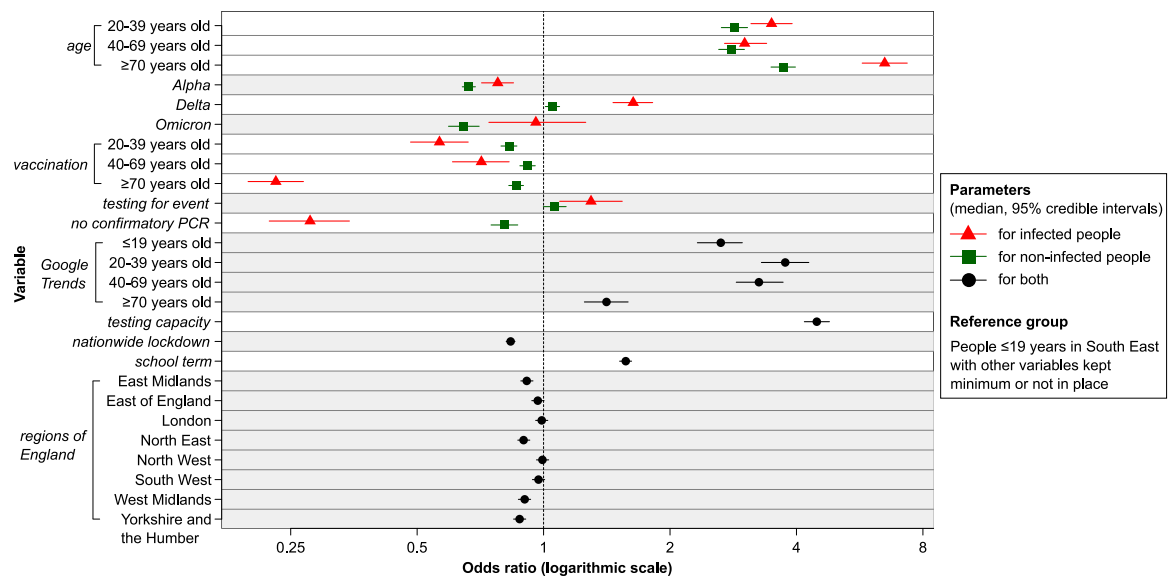
### SARS-CoV-2 PCR testing and underlying epidemic trends

NHS test data for England gave weekly PCR test results in 9 regions and 4 age groups over 92 weeks, resulting in 3,312 data points. The weekly number of PCR tests in England varied greatly throughout the epidemic, ranging from 392,873 (18 to 24 Jun 2020) to 3,835,758 (16 to 22 Dec 2021). Of these, the proportion of positive tests also showed large variability, ranging from 0.01 (29 April to 5 May 2021) to 0.31 (30 Dec 2021 to 5 Jan 2022). On the other hand, REACT test data provided weekly PCR test results from random samples of the population of England generated by 18 REACT rounds over 74 weeks. The weekly number of PCR tests ranged from 10 to 105,698, with a median of 23,867 (with the very low numbers per week arising from individuals returning tests taken outside of the intended time interval). The weekly percentage of positive PCR test results ranged from 0 to 4.6%, with a median of 0.4%. The latter ranged from 0 to 15.3%, with a median of 0.3%, when stratified by age and region. Based on the mean prevalence of PCR swab positivity obtained by fitting a generalised additive model (GAM) to REACT test data (hereafter REACT GAM fit) (Fig. S1), we estimated the number of infected and non-infected people in each region and age group for the start date of each week defined in the NHS test data. In England, the estimated weekly number of infected people fluctuated significantly, ranging from 30,763 (17 Jul 2020) to 2,529,972 (13 Jan 2022). Reflecting a relatively low level of infection prevalence at



**Fig. 1 | Conceptual framework for SARS-CoV-2 testing behaviour.** **a** The relationships between mechanisms proposed to influence testing behaviour. The conceptual framework proposes that testing behaviour is influenced by the combined effect of (i) disease susceptibility/severity, (ii) risk perception, and (iii) public health measures (solid lines). Disease susceptibility/severity and public health measures are also proposed to influence risk perception (dotted lines), in addition to their direct impact on testing behaviour (solid lines). However, the indirect impact (i.e. dotted lines) could not be formally disentangled within our

statistical framework. **b** Organisation of data and covariates under the conceptual framework. In models built on this framework, the population was divided into infected or non-infected people for each region of England, age group, and week, based on REACT test data (i). Then, infected or non-infected people were hypothesised to have taken a SARS-CoV-2 PCR test, depending on factors related to disease susceptibility/severity, risk perception, and public health measures (ii). Finally, the outcome of PCR test results was fitted to NHS test data by dynamic models (iii).



**Fig. 2 | Odds ratios for taking a SARS-CoV-2 PCR test from the best-fitting model.** The x-axis shows odds ratios on a logarithmic scale, and the y-axis shows variables included in the best-fitting model. Points and horizontal lines correspond to median and 95% credible intervals, respectively, estimated for the infected (red and triangle), non-infected (green and square), or both (black and circles). Infected (or non-infected) people aged  $\leq 19$  years (age group 1) in South East, with other variables kept minimum or not in place, represented the reference group for

parameters estimated for infected (or non-infected) people. Parameters estimated for both infected and non-infected people had the same reference group with no differentiation by infection status. For testing capacity and Google Trends, odds ratios comparing the maximum and minimum values are shown. The models in this study were fitted to weekly NHS test data for England, which ranged from 392,873 (18 to 24 Jun 2020) to 3,835,758 (16 to 22 Dec 2021) tests over 92 weeks. See UK Health Security Agency<sup>22</sup> for NHS test data.

the population level, the estimated weekly number of non-infected people remained relatively constant, fluctuating by less than 4.7% (from 54,020,166 to 56,519,375). The number of infected people estimated by the REACT GAM fit was larger than the observed number of positive tests in NHS test data, except for 33 out of 3,312 data points (1.0%). Most of those data points originated from the early period of Delta variant circulation between 17 June and 19 Aug 2021 ( $n = 23$ , 69.7%), notably in people aged 20 to 39 years ( $n = 16$ , 48.5%), and the 59-day interval between REACT rounds 13 and 14 that was much longer than the average between-round interval (median: 17 days, 1<sup>st</sup>–3<sup>rd</sup> quantiles: 13 to 19 days).

### Associations with testing behaviour

Testing behaviour was associated with various factors related to the following three mechanisms allowed for by our conceptual framework: disease susceptibility/severity, risk perception, and public health measures (Figs. 1 and 2 and Tables S1 and S2).

First, the relationship between testing probability and variables linked to disease susceptibility/severity, i.e. age, dominance of the Alpha, Delta, and Omicron variants, and vaccination, typically followed the same direction for infected and non-infected people, but varied significantly in magnitude, likely reflecting the added impact of increased disease susceptibility/severity on testing behaviour among those infected (blue and orange arrows in Fig. 1).

In the best-fitting model, the probability of testing appeared to increase markedly with age in both infected and non-infected people (Fig. 2). The increase in testing probability was particularly substantial among infected people, being even more pronounced in people aged  $\geq 70$  years (OR: 6.49, 95%CrI: 5.73–7.34) compared with non-infected people (OR: 3.72, 95%CrI: 3.48–3.98), highlighting the impact of both disease susceptibility/severity and that of risk perception. The transmission of particular SARS-CoV-2 Variants of Concern<sup>24</sup> was also associated with significant changes in testing behaviour. In particular, predominant circulation of the Delta variant was associated with increased testing, particularly among infected people (OR linked to  $\geq 50\%$  of circulating SARS-CoV-2 being Delta variant: 1.63, 95%CrI:

1.46–1.82). Contrastingly, predominant circulation of the Alpha and Omicron variant was associated with a decrease in testing, which was more pronounced among non-infected people (OR: 0.66, 95%CrI: 0.64–0.69 for the Alpha variant; OR: 0.65, 95%CrI: 0.59–0.70 for the Omicron variant).

Vaccination was associated with a decrease in testing behaviour among both infected and non-infected people and across age groups. The most substantial decrease in testing probability was among infected people, particularly in those aged  $\geq 70$  years (OR linked to  $\geq 50\%$  of the population receiving a second dose of vaccination: 0.23, 95%CrI: 0.20–0.27).

While the influence of variables described above was estimated differentially among infected and non-infected people by altering disease susceptibility/severity and risk perception, the following set of variables was allowed to affect equally or differentially infected and non-infected people through changes in risk perception (blue arrows in Fig. 1b) and/or public health measures (green arrows in Fig. 1b). COVID-19 social amplification measured by Google Trends as a proxy (search keyword: ‘covid’, normalised range: 0–100) appeared to play an essential role in driving testing behaviour, and the impact was least pronounced for people aged  $\geq 70$  years. The increase from the minimum (=14) to maximum (=100) Google Trends value was estimated to increase the odds of testing most in people aged 20 and 39 by 3.76 (95%CrI: 3.30–4.28) and least in those aged  $\geq 70$  by 1.41 (95%CrI: 1.25–1.59). The lateral flow device (LFD) testing requirement for attending large events appeared to increase PCR testing probability, with a significant increase among infected people (OR: 1.09–1.54). In contrast, PCR testing probability decreased when a confirmatory PCR test for people with a positive LFD test result was no longer required, especially among infected people (OR: 0.28, 95%CrI: 0.22–0.34). While both testing capacity and in-person school attendance were also associated with increased testing, the nationwide lockdown was associated with decreased testing. Finally, compared with South East, the probability of testing appeared to have been lower in other regions, including East Midlands, North East, Yorkshire and the Humber, and West Midlands.

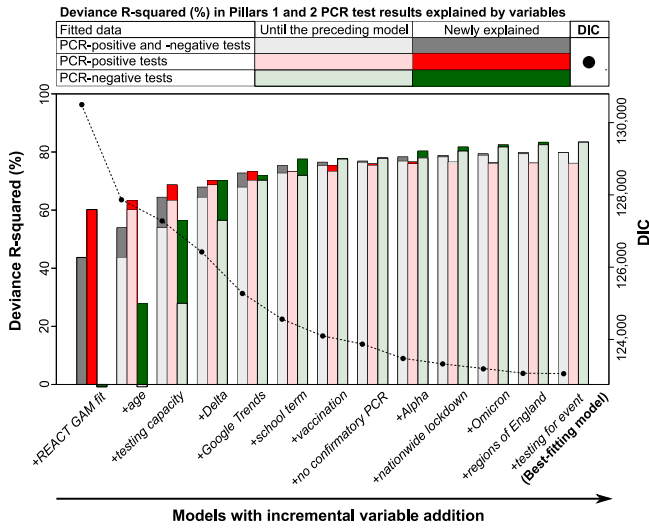
**Contribution to explaining variability in NHS test data**

All variables included in the best-fitting model contributed to reducing the deviance information criterion (DIC) by more than 5 units (Fig. 3). The best-fitting model explained 79.8% of the total deviance in NHS test data, with marginally better performance for explaining the

variability in the number of negative tests (83.4% explained) than positive tests (76.1% explained) (Fig. 3). For both positive and negative tests, the predicted numbers of PCR tests fitted reasonably well with the observed numbers over time by age group (Fig. 4 and S2) and region (Figs. S3 and S4).

Most of the variability in the number of positive tests was explained by the underlying epidemic trends (i.e. REACT GAM fit explained 60.2% of the deviance). To further explain variability in the number of positive tests, other important variables included testing capacity, age, and Google Trends, but their contributions to the overall deviance explained were relatively small (5.3%, 3.2%, and 3.0%, respectively). In contrast, explaining variability in the number of negative tests required accounting for more factors, notably age, testing capacity, and dominance of the Delta variant (28.7%, 28.6%, and 13.8%, respectively).

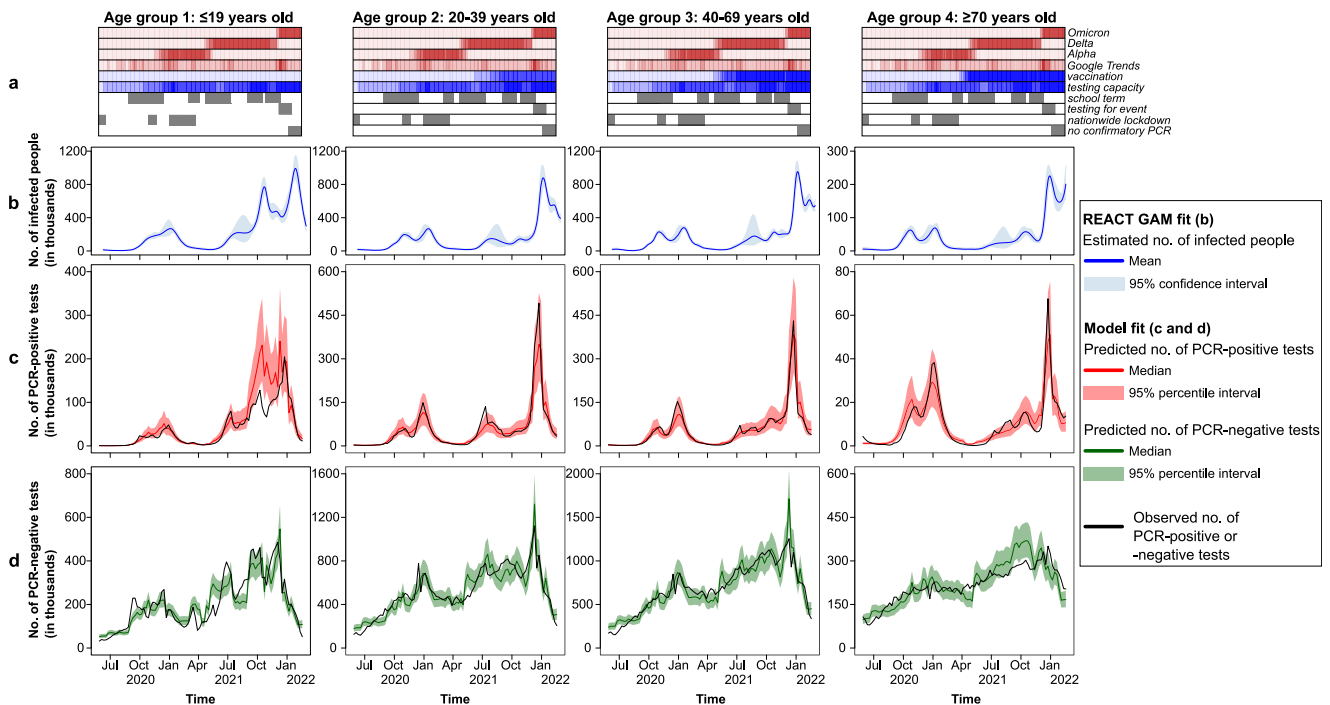
After accounting for variables in the best-fitting model, the probability of testing showed significant fluctuations throughout the epidemic (Fig. S5). Among infected people, the probability of testing gradually increased in all age groups, with intermittent drops. Notably, a sudden drop in testing probability was observed in people aged  $\geq 70$  when the population received a second dose of vaccination, whereas, in younger age groups, the largest drop was observed after the peak of the Omicron wave. Among those non-infected, the probability of testing also gradually increased over time until it decreased sharply after the peak of the Omicron wave in all age groups.



**Fig. 3 | Deviance R-squared (%) explained during a forward stepwise selection procedure.** Variables were added incrementally to the baseline model with only REACT GAM fit (most left) until the best-fitting model (most right). For each model, bars represent the total deviance R-squared explained by variables added in the preceding steps and the newly added variable (left y-axis). Different bar colours represent the deviance R-squared for PCR-positive tests (red), PCR-negative tests (green), or both (grey), with darker portions representing the deviance R-squared additionally explained by the newly added variable. Points represent the deviance information criterion (DIC) of each model (right y-axis).

**Sensitivity analysis**

Models with an alternative age classification in NHS test data and those with no age group weighting for the REACT GAM fit and vaccination did not substantially change the interpretation of the best-fitting model results (Fig. S6). More specifically, after including younger people in age group 4 (i.e. aged  $\geq 60$  years, instead of  $\geq 70$  years) (alternative model A), the effect of age and vaccination on testing behaviour decreased in age group 4, particularly among infected



**Fig. 4 | Predictive posterior check of the best-fitting model by age group and over time.** Row a shows the temporal trend of variables included in the best-fitting model, and row b shows the temporal trend of the number of infected (mean, 95% confidence intervals) estimated by REACT GAM fit. Row c shows the predicted (red

lines: median, reddish shades: 95% percentile intervals) and observed (black lines) numbers of Pillar 1 and Pillar 2 PCR-positive test results. Row d shows the predicted (green lines: median, greenish shades: 95% percentile intervals) and observed (black lines) numbers of Pillar 1 and Pillar 2 PCR-negative test results.



people, whereas the effect of Google Trends became more substantial in age group 4. Those associations remained significant with alternative model A. No age group weighting (alternative model B) resulted in no significant association with vaccination in age group 1, linked to having  $\leq 17$ , instead of  $\leq 19$  years old in the age group. There were no other significant changes with alternative model B.

## Discussion

Our study results show complex dynamics of SARS-CoV-2 PCR testing behaviour, varying during 20 months of the pandemic in England and differing between infected and non-infected people, in association with factors hypothesised under our conceptual framework (Fig. 1).

One of the core hypotheses in our conceptual framework was that risk perception influenced the testing behaviour of both infected and non-infected people. Then, in addition to risk perception, factors related to disease susceptibility/severity, including age, SARS-CoV-2 virulence, and immune status, directly affected the testing behaviour of infected people by causing different levels of symptoms. Under this hypothesis, for a variable to affect testing behaviour by changing both disease susceptibility/severity and risk perception, the observed difference in the magnitude of associations between infected and non-infected people would reflect the relative contribution of disease susceptibility/severity and risk perception to testing behaviour, although, to quantify such an effect, one would need to know the precise form of the relationship of testing behaviour with those two mechanisms.

Indeed, considerable differences were observed for several variables between infected and non-infected people. The associations between testing behaviour and age in infected people, particularly among those aged  $\geq 70$  years, suggests that age-dependent susceptibility/severity to COVID-19<sup>25,26</sup> served as the primary determinant of testing participation. Similar age group-specific patterns in non-infected people would indicate that different age groups formed different levels of risk perception based on their perceived risk of contracting COVID-19 and experiencing severe illness upon infection. In addition, people aged  $\geq 70$  years could have been more likely to take a test than those in younger age groups, if they had COVID-19-like symptoms, but not actually from SARS-CoV-2 infection, more frequently. Furthermore, people aged  $\geq 70$  years would have been prescribed a PCR test more often, for example, when admitted to hospitals or living in care homes. This would have further increased testing participation in this age group.

Our findings also suggest that people significantly changed their testing behaviour during the circulation of different dominant variants, potentially through changes in risk perception, in addition to heterogeneities in the variants' virulence actually realised among those infected. For Delta variant circulation, a considerable increase in testing probability among infected people is likely to reflect the variant's greater virulence<sup>18,19,27</sup>. Among non-infected people, the increase in testing probability may suggest that the Delta variant's potential to cause severe disease increased testing participation by heightening risk perception. The same logic could explain the decrease in testing probability with Omicron variant circulation, given the variant's association with generally milder and less specific symptoms<sup>18,27</sup>. While the association with Omicron variant circulation was significant among non-infected people, it was not among infected people. This was likely influenced by withdrawing a confirmatory PCR testing scheme in the middle of Omicron variant circulation, as it appeared to explain most of the decrease in testing participation in the best-fitting model.

For Alpha variant circulation, the decrease in testing probability was unexpected as the variant has been suggested to be more transmissible and cause more severe disease than pre-Alpha SARS-CoV-2 strains<sup>20,28,29</sup>. One possible explanation is the influence of epidemic trends on risk perception. The Alpha variant became dominant and peaked during the winter holiday season shortly after the first

epidemic wave and the second nationwide lockdown. If risk perception was decreased by these factors, people could have become less willing to receive a test, offsetting the possible impact of the variant's increased transmissibility and virulence. Importantly, this suggests that testing behaviour would not always follow the dynamics of viral circulation and virulence in the presence of other external factors.

With COVID-19 vaccination, the decrease in testing behaviour in both infected and non-infected people suggests that risk perception toward COVID-19 was reduced in the public with the widespread vaccine rollout. A greater decrease among infected people further indicates that vaccine effectiveness against symptomatic disease could have influenced the testing behaviour of those vaccinated upon infection<sup>30–32</sup>. Indeed, among infected people, the largest decrease was observed in those aged  $\geq 70$ , who are most likely to experience symptomatic COVID-19 without vaccination<sup>25,26</sup>. This indicates that a significant fraction of this age group could have benefited from vaccination in avoiding apparent symptoms, thereby having considerably reduced test propensity. Our findings also suggest that efforts to increase testing participation should be maintained to control community transmission during widespread vaccine rollout because, while COVID-19 vaccination is proven to be effective for preventing severe disease<sup>30–32</sup>, its contribution to reducing SARS-CoV-2 transmissibility is becoming increasingly limited with the emergence of new variants<sup>33,34</sup>.

The increase in testing probability with Google Trends for the keyword 'covid' highlights the influence of COVID-19 social amplification on testing behaviour, potentially by affecting risk perception. The increase in testing probability was more pronounced in younger age groups, which might reflect different internet usage patterns between age groups<sup>35</sup>. For example, more frequent and interactive nature of digital media, rather than traditional media, could have affected risk perception and, thus, testing behaviour differently depending on internet usage patterns. However, there could also be other explanations not necessarily related to risk perception. First, compared with those aged  $\geq 70$  years, the testing behaviour in younger age groups could have been more responsive to COVID-19 social amplification, e.g. through increased use of online PCR test appointments, especially when the demand for PCR tests was high<sup>36</sup>. Secondly, if disease susceptibility/severity was a predominant factor determining testing behaviour among those aged  $\geq 70$  years, changes in the level of COVID-19 social amplification would have had limited influence on testing behaviour.

Our findings also suggest that testing behaviour was limited or driven by public health measures. Among them, testing capacity explained a relatively large fraction of the variability in NHS test data, particularly for the number of PCR-negative tests, confirming its importance for the design of the testing programmes. Moreover, its positive association with testing behaviour suggests that a shortage of available tests might have restricted testing participation, as witnessed at the beginning of the pandemic and some periods during the Omicron variant surge<sup>37–39</sup>. However, while increasing testing capacity would identify more infected people, it could diminish optimal use of limited testing resources if the increase in testing became disproportionately higher among non-infected people. This indicates that, if the primary objective of testing is to identify as many infected people as possible with a limited number of available tests, increasing testing capacity should be accompanied by strategies designed to increase the probability of testing infected people and therefore optimise the allocation of testing resources. Such strategies might include targeting high-risk groups, contact tracing, and effective risk communications.

Although public health measures introduced during particular periods showed an association with testing behaviour, they explained a relatively small proportion of the variability in NHS test data. First, during school terms, a confirmed COVID-19 case in a class would trigger testing students in the school, thereby increasing testing

participation. In particular, the introduction of regular testing in schools in March 2021 likely facilitated this process<sup>40</sup>. Furthermore, during school terms, students could have re-evaluated the risk of contracting COVID-19 and therefore changed their testing behaviour. Similarly, requiring an LFD test for attending large events was also shown effective for increasing PCR testing probability, especially among infected people, likely from requiring a confirmatory PCR test on those who tested positive with an LFD test. This, in turn, explains the decrease in testing participation after withdrawing the confirmatory PCR test policy, particularly among infected people. Nationwide lockdowns were also negatively associated with testing participation, possibly reflecting a reduced need for testing from having fewer contacts and a reduced level of risk perception following a rapidly shirking epidemic trend during lockdown periods.

Finally, the probability of testing varied among regions. Interestingly, regions with a significantly lower testing probability—East Midlands, West Midlands, Yorkshire and the Humber, and North East—also have the lowest average household income in England; while the reference region, i.e. South East, has the second highest average household income, close to that of London<sup>41</sup>. This might suggest that socio-economic deprivation hindered testing participation to some extent during the pandemic. Other factors could have also contributed to regional differences in testing behaviour. In fact, despite having a relatively low average household income, North West showed a similar testing probability as the reference region. The relatively high testing probability in North West may have been affected by the voluntary mass LFD testing pilot conducted in Liverpool, one of the largest cities in the region, for people with or without COVID-19 symptoms<sup>42</sup>. Between 6 November 2020 and 30 April 2021, 57% of Liverpool residents took an LFD test, which led to an increase in PCR testing to confirm a positive LFD test result<sup>43</sup>. The awareness of this pilot may have also influenced risk perception and, therefore, PCR testing in the region. In addition, population size could have also impacted testing behaviour if accessibility to testing centres or perceived risk of contracting COVID-19 was associated with population size; South East and North West are ranked 1<sup>st</sup> and 3<sup>rd</sup> by population size, respectively<sup>44</sup>, whereas regions with a significantly lower testing probability have relatively small populations.

Several factors must be considered when interpreting the findings of this study. First, the key mechanisms proposed by our modelling framework, i.e. disease susceptibility/severity, risk perception, and public health measures, would have affected testing behaviour by modifying one another, in addition to their direct impact on testing behaviour. In our models, multiple variables were selected as proxies for each mechanism, and some of those variables were associated with more than one mechanism, depending on data available at the population level. While this modelling approach helped disentangle and explore testing behaviour in relation to various contextual meta-data within our conceptual framework, it limited our ability to evaluate other forms of causal relationships, for example, the indirect impact of disease susceptibility/severity and public health measures on risk perception (dotted lines in Fig. 1a). Thus, care must be taken when making causal inferences from such observational studies. Studies based on individual-level data might be better suited for exploring formally the causal relationships hypothesised here. For example, the probability of testing could be modelled at the individual level, with testing history as the outcome assumed to have various causal relationships at the individual-level (e.g. age and vaccination history) and population-level (e.g. dominant variant in circulation and public health measures in place) factors. However, it must also be noted that such studies face challenges in accounting for selection bias and factors that constantly change over the pandemic, notably the underlying epidemic trends. These complexities highlight the need to explore complementary study designs and results to further understand testing behaviour.

Secondly, there could be other factors associated with testing behaviour which were not explicitly investigated. Notably, data on LFD tests were not analysed as we considered that their association with PCR testing behaviour was explained by other variables included, specifically a confirmatory PCR testing requirement for those testing positive with LFD tests and PCR testing capacity for both infected and non-infected people. However, if compliance with the confirmatory PCR test requirement and the availability of PCR and LFD tests changed over time, the trend in LFD tests could have affected PCR testing behaviour differently during the study period. This relationship highlights the need for future research into the dynamics between LFD and PCR testing behaviour to optimise surveillance and control strategies with those two test types<sup>4</sup>. In addition, age- or region-specific associations with testing behaviour were not explored for some of the variables, mainly due to data unavailability at corresponding levels<sup>45–47</sup>. For example, we did not use age-specific parameters for the dominant variants to avoid correlations with parameters for vaccination; the vaccine rollout and Delta variant emergence occurred concurrently at the national level. Also, data for PCR testing capacity and Google Trends were not available at the regional level, although these variables could have varied between regions depending on regional epidemic situations. These data gaps suggest that future research based on our conceptual framework could be strengthened further by employing age- and region-specific data, should they become available.

Finally, the underlying epidemic trends were informed by the mean prevalence of SARS-CoV-2 swab positivity, i.e. REACT GAM fit, without accounting for its uncertainty. Thus, under- or over-estimation of the true prevalence could have affected model estimates by under- or over-estimating the number of infected and non-infected people. This was apparent during the early Delta variant circulation in people aged 20–39 years when the observed number of positive tests in NHS test data was greater than the estimated number of infected people by REACT GAM fit. In fact, the interval between REACT rounds during this period was much longer than other periods, suggesting that REACT GAM fit could have under-estimated the prevalence of SARS-CoV-2 swab positivity during this period, potentially by missing increasing epidemic trends when tests were not conducted between REACT study rounds. However, the overall impact of the underestimate is unlikely to be significant, considering that those observations comprised only 1.0% of the data points.

Our findings suggest how testing strategies could be optimised for different objectives, notably monitoring cases, controlling transmission, and clinical management. For monitoring cases, varying levels of disease susceptibility/severity and risk perception need to be considered when interpreting epidemic trends, as transmission among people expected to experience asymptomatic or mild infection could be underestimated in voluntary test data. For the transmission control objective, changes in risk perception require continuous assessments, and efforts should be made to maintain risk awareness and, thus, test participation. This will be particularly important when a new variant emerges with different virulence and transmissibility, as shown by this study during Omicron variant circulation, or when public health measures are newly introduced, with a potential impact on risk perception. For clinical management, the public needs to be well informed about the risk profile of disease susceptibility/severity in order to increase testing frequency among people who would need treatments upon infection. Finally, it must be noted that different testing objectives cannot be considered alone, as efforts made for one testing objective could have synergistic or counterproductive effects on other objectives. This need for balance highlights how testing strategies should be designed and updated constantly based on the relative importance of different testing objectives under given epidemic contexts.

In conclusion, our study highlights the usefulness of representative infection survey data, combined with voluntary testing and other

meta-data, for understanding factors underpinning the complex dynamics of testing behaviour. While the underlying epidemic trends and disease susceptibility/severity were found to drive these dynamics, changes in risk perception and public health measures were also associated with testing behaviour. These illustrate that testing strategies should be designed based on the combined effect of these mechanisms on testing behaviour and the relative importance of different testing purposes for given epidemic situations.

## Methods

### Data

**NHS testing data.** The weekly number of people tested and the weekly number of people testing positive for COVID-19 by PCR tests via Pillar 1 and Pillar 2 routes were obtained from weekly statistics for NHS Test and Trace (England)<sup>22</sup>. These data represented COVID-19 swab testing in the UK: Pillar 1 targeted people with a clinical need, and health and care workers, and Pillar 2 targeted the wider population. The numbers were aggregated by week, region, and age group (Table 1).

**REACT test data.** REACT test data included PCR test results from random samples of the population of England between 27 April 2020 and 28 February 2022 (REACT rounds 1–18)<sup>23</sup>. The original data included the weekly number of PCR tests and, among them, the weekly number of positive tests in 9 regions and 8 age groups over 74 weeks. Since these original data were prepared using different age group definitions, we re-classified age groups by weighting the total number of tests and the number of positive tests by accounting for age group population sizes. After weighting, a binomial GAM was fitted to estimate the prevalence of SARS-CoV-2 swab positivity in each region and age group, REACT GAM fit, using the *gam* function of the *mgcv* package<sup>48</sup> (see SI for details).

**Covariates included.** The definition and source of covariates are detailed in Table 1.

### Model

Models describing the process of people seeking and receiving a SARS-CoV-2 PCR test were fitted to NHS test data<sup>49</sup>. We allowed infected and non-infected people to have different likelihoods of taking a PCR test, considering that their decision to take a PCR test was likely influenced by several factors, notably the presence and severity of COVID-19 symptoms and recent contacts with those confirmed infected. Therefore, for each region, age group, and week, the study population was divided into infected and non-infected people first. The number of infected people ( $N_{R,A,W}^1$ , i.e. producing a true positive or false negative test result if tested, orange square in the top panel of Fig. 1b) was estimated as a product of the population size and the mean prevalence of SARS-CoV-2 swab positivity from REACT GAM fit. The number of non-infected people ( $N_{R,A,W}^2$ , i.e. producing a true negative or false positive test result if tested, blue square in the top panel of Fig. 1b) was estimated similarly. Then, the expected numbers of positive and negative PCR test results were modelled with the following structure:

$$n_{R,A,W}^p = N_{R,A,W}^1 \theta_{R,A,W}^1 \varphi + N_{R,A,W}^2 \theta_{R,A,W}^2 (1 - \omega) \quad (1)$$

$$n_{R,A,W}^n = N_{R,A,W}^2 \theta_{R,A,W}^2 \omega + N_{R,A,W}^1 \theta_{R,A,W}^1 (1 - \varphi) \quad (2)$$

For region,  $R$ , age group  $A$ , and week,  $W$ ,  $n_{R,A,W}^p$  and  $n_{R,A,W}^n$  represent the expected number of positive and negative PCR test results, respectively, to be compared with respective NHS test results during parameter estimation based on their likelihoods (see below).

The expected number of positive test results was modelled as the sum of true and false positive test results, and in the same way for the expected number of negative test results. For the expected number of

positive tests,  $n_{R,A,W}^p$ , the first and second terms of the right-hand side of (1) represent the expected numbers of people with true and false positive test results, respectively (red arrows in the bottom panel of Fig. 1b). The expected number of negative tests,  $n_{R,A,W}^n$ , was modelled in the similar manner (2) (blue arrows in the bottom panel of Fig. 1b). We assumed 95.0% and 99.9% as PCR test sensitivity ( $\varphi$ ) and specificity ( $\omega$ ), respectively<sup>50</sup>.

The probability of seeking and conducting a test was allowed to differ between infected and non-infected people,  $\theta_{R,A,W}^1$  and  $\theta_{R,A,W}^2$  respectively, within a logistic regression framework (Fig. 1b):

$$\text{logit}(\theta_{R,A,W}^1) = \alpha^1 + \sum_i \beta_i^1 X_{i,R,A,W} + \sum_j \beta_j^3 X_{j,R,A,W} \quad (3)$$

$$\text{logit}(\theta_{R,A,W}^2) = \alpha^2 + \sum_i \beta_i^2 X_{i,R,A,W} + \sum_j \beta_j^3 X_{j,R,A,W} \quad (4)$$

An intercept was estimated separately for infected and non-infected people (noted as  $\alpha^1$  and  $\alpha^2$ ), assuming that they had different baseline testing probabilities. The youngest age group (i.e.  $\leq 19$  years old) was chosen as the reference to present the association of age with testing behaviour in order of age and also considering its larger number of positive tests than the oldest age group (i.e.  $\geq 70$  years old). South East was chosen as the reference given its largest population among regions of England.

$X_{R,A,W}$  represent covariates for a given region, age group, and week which were allowed to influence testing probability (see Table 1). Some variables were hypothesised to affect testing behaviour differently for infected and non-infected people (noted as  $\beta^1$  and  $\beta^2$ , respectively). For example, while some variables may increase the testing behaviour of both infected and non-infected people by changing risk perception (blue arrows in the middle panel of Fig. 1b), they may also have an additional impact on the testing behaviour of infected people by changing the nature of the infection itself, thereby leading to greater disease susceptibility/severity (orange arrows in the middle panel of Fig. 1b).

Therefore, regression coefficients for variables likely linked to disease susceptibility/severity, i.e. *age*, *Alpha*, *Delta*, *Omicron*, and *vaccination*, were estimated separately for infected and non-infected people. Similarly, coefficients associated with *testing for event* and *no confirmatory PCR* were also estimated separately as those public health measures had different implications for taking PCR tests following positive or negative LFD test results.

In contrast, coefficients associated with variables not directly linked to disease susceptibility/severity, including *Google Trends*, *testing capacity*, *school term*, *nationwide lockdown*, and *regions of England*, were assumed the same for infected and non-infected people, i.e. prior to testing, those factors would equally affect infected and non-infected people through changes in public health measures and risk perception (noted as  $\beta^3$ ).

In addition to *age*, the impacts of *vaccination* and *Google Trends* were also allowed to be age-specific for all age groups, and *school term* was assumed to affect age group 1 (i.e. those aged 0–19) only.

Note that age classifications in REACT test and vaccination data were slightly different from that of NHS test data. In this study, age classes were defined following age categories in NHS test data, and covariates were re-grouped into these age classes by weighting their values by the population size of each age group (See SI for details, including sensitivity analyses). A Metropolis-Hastings Markov chain Monte Carlo (MCMC) algorithm<sup>51</sup> was implemented in R 4.1.2<sup>52</sup> to estimate parameter based on optimising the following likelihood

**Table 1 | Use, definition, and source of variables**

Variable	Category	Note	Source
<b>Outcome where models were fitted</b>			
NHS PCR test data		• Weekly numbers of positive and negative PCR test results via NHS Test and Trace Pillar 1 and Pillar 2 routes over 92 weeks between 28 <sup>th</sup> May 2020 and 24 <sup>th</sup> February 2022.	22
<b>Explanatory</b>			
<b>1. To account the underlying epidemic trends</b>			
REACT GAM fit		• Mean prevalence of SARS-CoV-2 swab positivity estimated from REACT test data • Used to divide the study population into infected and non-infected people. Thus, no parameter estimation was made for this variable.	23
<b>2. To assess the process of seeking and receiving a PCR test</b>			
age	age group 1	• ≤19 years old (Reference group)	
	age group 2	• 20–39 years old	
	age group 3	• 40–69 years old	
	age group 4	• ≥70 years old	
Alpha	< 50% of circulation	• Other periods	55
	≥ 50% of circulation	• 10 Dec 2020 to 13 May 2021	
Delta	< 50% of circulation	• Other periods	56
	≥ 50% of circulation	• 13 May 2021 to 09 Dec 2021	
Omicron	< 50% of circulation	• Other periods	56
	≥ 50% of circulation	• 09 Dec 2021 to 24 Feb 2022	
vaccination	< 50% of population	• Cumulative proportion of people with the second vaccine dose • Age group-specific coefficients were used. ◦ From 12 Aug 2021 in age group 2 ◦ From 27 May 2021 in age group 3 ◦ From 15 Apr 2021 in age group 4 ◦ Age group 1 was not considered as vaccination did not reach 50%	57
	≥ 50% of population		
testing for event	not required	• 1 if a lateral flow device test was required for entering nightclubs and large events, 0 otherwise • 15 Dec 2021 to 27 Jan 2022	58
	required		
no confirmatory PCR	confirmatory PCR required	• The requirement for a confirmatory PCR on people with positive LFD test results • Required from 11 Jan 2022 • PCR testing was still required for people with symptoms until the end of the study period.	59
	confirmatory PCR not required		
testing capacity	Continuous	• Daily no. of PCR testing capacity averaged for each week	57
Google Trends	Continuous	• Google Trends index for search keyword ‘covid’ in the United Kingdom • Age group-specific coefficients were used. • The same value was assumed for all age groups.	60
school term	school holiday	• Applied only to age group 1 • In 2021, students returned to schools in March due to the third nationwide lockdown	61
	school term		
regions of England	South East (reference group)		
	East Midlands		
	East of England		
	London		
	North East		
	North West		
	South West		
	West Midlands		
nationwide lockdown	no lockdown	• 1st lockdown: 23 Mar 2020-23 Jun 2020 • 2nd lockdown: 5 Nov 2020-2 Dec 2020 • 3rd lockdown: 6 Jan 2021-29 Mar 2021 • As 3rd lockdown was gradually relaxed, we set that it ended when the ‘Stay at home’ order was abolished <sup>21</sup> , assuming the impact on testing behaviour was most substantial by this order.	21
	In lockdown		

function,  $L$ :

$$L = \prod_{R,A,W} \Pr\left(k_{R,A,W}^p; n_{R,A,W}^p, \varnothing^p\right) \Pr\left(k_{R,A,W}^n; n_{R,A,W}^n, \varnothing^n\right) \quad (5)$$

Here, the likelihood of observing the number of NHS positive test results,  $k_{R,A,W}^p$ , (or negative test results,  $k_{R,A,W}^n$ ) was assumed to follow a negative binomial distribution with mean  $n_{R,A,W}^p$  (or  $n_{R,A,W}^n$ ) and over-dispersion parameter  $\varnothing^p$  (or  $\varnothing^n$ ). Uniform prior distributions were

used, except for  $\varnothing^p$  and  $\varnothing^n$ , which were constrained to be positive. After discarding burn-in posterior estimate, 200,000 posterior samples were retained and convergence was assessed by visual inspection of MCMC trace plots and Gelman-Rubin convergence diagnostic (<1.01)<sup>53</sup>.

Variables were added through a manual forward-stepwise selection procedure. In the first cycle, all univariate models were considered in turn, and the model with the lowest DIC was retained, provided that the decrease in deviance was >5 units compared to the null,



intercept-only, model. The same process was then repeated for the remaining variables in the next cycles until no other variables resulted in >5 DIC unit difference, leading to the best-fitting model. Posterior predictive check was performed for a model selected in each cycle by comparing the observed and simulated number of PCR-positive and -negative test results for each region, age group and week.

To assess how much variability in NHS test data was explained when variables were added incrementally, the Deviance *R*-squared was estimated for each selected model separately for the positive and negative test results and for both, accounting for overdispersed count data<sup>54</sup>. Additionally,  $\theta^1$  and  $\theta^2$  were computed accounting for all variables in the best-fitting model, to assess the trends in testing behaviour during the pandemic.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

This study used publicly available population-level data. Data sources are presented in Table 1.

### Code availability

Model codes are available from <https://github.com/KimYounjung/COVIDTestingBehaviour>.

### References

- Hasell, J. et al. A cross-country database of COVID-19 testing. *Sci. Data* **7**, 345 (2020).
- UK Health Security Agency. [WITHDRAWN] COVID-19: People With COVID-19 And Their Contacts. <https://www.gov.uk/government/publications/covid-19-people-with-covid-19-and-their-contacts/covid-19-people-with-covid-19-and-their-contacts> (2022).
- World Health Organization. Recommendations for National SARS-CoV-2 Testing Strategies And Diagnostic Capacities (WHO, 2021) <https://www.who.int/publications/i/item/WHO-2019-nCoV-lab-testing-2021.1-eng>.
- WHO Regional Office for Europe and Stockholm. Operational Considerations For Respiratory Virus Surveillance in Europe (European Centre for Disease Prevention and Control, 2022) <https://www.ecdc.europa.eu/en/publications-data/operational-considerations-respiratory-virus-surveillance-europe>.
- Ferrer, R. & Klein, W. M. Risk perceptions and health behavior. *Curr. Opin. Psychol.* **5**, 85–89 (2015).
- Schneider, C. R. et al. COVID-19 risk perception: a longitudinal analysis of its predictors and associations with health protective behaviours in the United Kingdom. *J. Risk Res.* **24**, 294–313 (2021).
- Vandrevala, T., Montague, A., Terry, P. & Fielder, M. D. Willingness of the UK public to volunteer for testing in relation to the COVID-19 pandemic. *BMC Public Health* **22**, 565 (2022).
- Smith, L. E. et al. Adherence to the test, trace, and isolate system in the UK: results from 37 nationally representative surveys. *BMJ* **372**, n608 (2021).
- Jaspal, R. & Breakwell, G. M. Social support, perceived risk and the likelihood of COVID-19 testing and vaccination: cross-sectional data from the United Kingdom. *Curr. Psychol.* **41**, 492–504 (2022).
- McDonald, S. A. et al. Testing behaviour and positivity for SARS-CoV-2 infection: insights from web-based participatory surveillance in the Netherlands. *BMJ Open* **11**, e056077 (2021).
- Allen, W. E. et al. Population-scale longitudinal mapping of COVID-19 symptoms, behaviour and testing. *Nat. Hum. Behav.* **4**, 972–982 (2020).
- Elliott, P. et al. Exponential growth, high prevalence of SARS-CoV-2, and vaccine effectiveness associated with the Delta variant. *Science* **374**, eabl9551 (2021).
- Elliott, P. et al. Twin peaks: The Omicron SARS-CoV-2 BA.1 and BA.2 epidemics in England. *Science* **376**, eabq4411 (2022).
- Beauté, J., Adlhoch, C., Bundle, N., Melidou, A. & Spiteri, G. Testing indicators to monitor the COVID-19 pandemic. *Lancet Infect. Dis.* **21**, 1344–1345 (2021).
- Wise, T., Zbozinek, T. D., Michelini, G., Hagan, C. C. & Mobbs, D. Changes in risk perception and self-reported protective behaviour during the first week of the COVID-19 pandemic in the United States. *R. Soc. Open Sci.* **7**, 200742 (2020).
- Moss, R., Zarebski, A. E., Carlson, S. J. & McCaw, J. M. Accounting for healthcare-seeking behaviours and testing practices in real-time influenza forecasts. *Trop. Med. Infect. Dis.* **4**, 12 (2019).
- Ibuka, Y., Chapman, G. B., Meyers, L. A., Li, M. & Galvani, A. P. The dynamics of risk perceptions and precautionary behavior in response to 2009 (H1N1) pandemic influenza. *BMC Infect. Dis.* **10**, 296 (2010).
- Nyberg, T. et al. Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study. *Lancet* **399**, 1303–1312 (2022).
- Twohig, K. A. et al. Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. *Lancet Infect. Dis.* **22**, 35–42 (2022).
- Scientific Advisory Group for Emergencies. NERVTAG: Update Note On B.1.1.7 Severity. <https://www.gov.uk/government/publications/nervtag-update-note-on-b117-severity-11-february-2021> (2021).
- The Institute for Government. Timeline of UK Government Coronavirus Lockdowns And Measures, March 2020 to December 2021. <https://www.instituteforgovernment.org.uk/data-visualisation/timeline-coronavirus-lockdowns> (2021).
- UK Health Security Agency. Weekly Statistics For NHS Test and Trace (England): 3 to 9 March 2022. <https://www.gov.uk/government/collections/nhs-test-and-trace-statistics-england-weekly-reports> (2022).
- Riley, S. et al. REal-time Assessment of Community Transmission (REACT) of SARS-CoV-2 virus: study protocol. *Wellcome Open Res.* **5**, 200 (2020).
- World Health Organization. Tracking SARS-CoV-2 Variants. <https://www.who.int/activities/tracking-SARS-CoV-2-variants> (2022).
- Davies, N. G. et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* **26**, 1205–1211 (2020).
- Romero Starke, K. et al. The isolated effect of age on the risk of COVID-19 severe outcomes: a systematic review with meta-analysis. *BMJ Global Health* **6**, e006434 (2021).
- Menni, C. et al. Symptom prevalence, duration, and risk of hospital admission in individuals infected with SARS-CoV-2 during periods of omicron and delta variant dominance: a prospective observational study from the ZOE COVID Study. *Lancet* **399**, 1618–1624 (2022).
- Volz, E. et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature* **593**, 266–269 (2021).
- Davies, N. G. et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **372**, eabg3055 (2021).
- Lopez Bernal, J. et al. Effectiveness of covid-19 vaccines against the B.1.617.2 (Delta) variant. *N. Eng. J. Med.* **385**, 585–594 (2021).
- Lopez Bernal, J. et al. Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. *BMJ* **373**, n1088 (2021).
- Moline, H. L. et al. Effectiveness of COVID-19 vaccines in preventing hospitalization among adults aged ≥65 years—COVID-NET, 13 States, February–April 2021. *MMWR Morb. Mortal Wkly Rep.* **70**, 1088–1093 (2021).
- Franco-Paredes, C. Transmissibility of SARS-CoV-2 among fully vaccinated individuals. *Lancet Infect. Dis.* **22**, 16 (2022).

34. Ferguson, N., et al. *Growth, Population Distribution And Immune Escape of the Omicron in England* (Imperial College London, 2021) <https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2021-12-16-COVID19-Report-49.pdf>.
35. Johnson, J. *Internet Activities Performed in Great Britain 2020*, By Age (Statista, 2020) <https://www.statista.com/statistics/275805/internet-activities-performed-in-great-britain/>.
36. Green, M. A. et al. Evaluating social and spatial inequalities of large scale rapid lateral flow SARS-CoV-2 antigen testing in COVID-19 management: An observational study of Liverpool, UK (November 2020 to January 2021). *Lancet Reg. Health Eur.* <https://doi.org/10.1016/j.lanepe.2021.100107> (2021).
37. Paun, C. *Tension Over Shortage of Coronavirus Tests Across Europe* (Politico, 2020) <https://www.politico.eu/article/europe-testing-question-coronavirus-crisis/#:~:text=The%20lack%20of%20access%20across,laboratory%20testing%20materials%20and%20capacity.>
38. Iacobucci, G. Covid-19: Lack of capacity led to halting of community testing in March, admits deputy chief medical officer. *BMJ* **369**, m1845 (2020).
39. Barnes, O., Burn-Murdoch, J. & Hughes, L. *Why Are There Covid Test Shortages In The UK?* (Financial Times, 2020) <https://www.ft.com/content/25c9aa67-2b7e-4b8d-a25b-0b7b7ac48452>.
40. Department of Education. *Coronavirus (COVID-19) Asymptomatic Testing in Schools and Colleges.* <https://www.gov.uk/government/publications/coronavirus-covid-19-asymptomatic-testing-in-schools-and-colleges/coronavirus-covid-19-asymptomatic-testing-in-schools-and-colleges> (2021).
41. Office For National Statistics. *What Are The Regional Differences In Income And Productivity?* <https://www.ons.gov.uk/visualisations/dvc1370/> (2021).
42. Zhang, X. et al. Impact of community asymptomatic rapid antigen testing on covid-19 related hospital admissions: synthetic control study. *BMJ* **379**, e071374, <https://doi.org/10.1136/bmj-2022-071374> (2022).
43. Department of Health and Social Care. *Liverpool Coronavirus (COVID-19) Community Testing Pilot: Full Evaluation Report Summary.* <https://www.gov.uk/government/publications/liverpool-coronavirus-covid-19-community-testing-pilot-full-evaluation-report-summary> (2021).
44. Office For National Statistics. *Mid-Year Population Estimates June 2021.* <https://www.gov.wales/mid-year-estimates-population-2021> (2022).
45. Fisman, D. N. & Tuite, A. R. Age-specific changes in virulence associated with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants of concern. *Clin. Infect. Dis.* **75**, e69–e75 (2022).
46. O'Driscoll, M. et al. Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature* **590**, 140–145 (2021).
47. Sah P, et al. Asymptomatic SARS-CoV-2 infection: a systematic review and meta-analysis. *Proc. Natl Acad. Sci. USA* **118**, e2109229118 (2021).
48. Wood, S. N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B* **73**, 3–36 (2011).
49. Department of Health and Social Care. *Coronavirus (COVID-19): Scaling Up Our Testing Programmes.* <https://www.gov.uk/government/publications/coronavirus-covid-19-scaling-up-testing-programmes> (2020).
50. Kara Steel RY. *Coronavirus (COVID-19) Infection Survey: Methods And Further Information* (Statistics, 2022) <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/methodologies/covid19infectionsurveyipilotmethodsandfurtherinformation>.
51. Gilks, W. R., Richardson, S. & Spiegelhalter, D. J. *Markov Chain Monte Carlo in Practice* (Chapman & Hall, 1998).
52. R Core Team. *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. See <https://www.R-project.org/>. (2022).
53. Gelman, A. et al. *Bayesian Data Analysis.* (CRC Press, Boca Raton, 2014).
54. Cameron, A. C. & Windmeijer, F. A. G. R-squared measures for count data regression models with applications to health-care utilization. *J. Bus. Econ. Stat.* **14**, 209–220 (1996).
55. Sonabend, R. et al. Non-pharmaceutical interventions, vaccination, and the SARS-CoV-2 delta variant in England: a mathematical modelling study. *Lancet* **398**, 1825–1835 (2021).
56. UK Health Security Agency. *Investigation of SARS-CoV-2 Variants: Technical Briefings.* <https://www.gov.uk/government/publications/investigation-of-sars-cov-2-variants-technical-briefings> (2022).
57. UK Health Security Agency. *Coronavirus (COVID-19) in the UK Dashboard.* <https://coronavirus.data.gov.uk/> (2022).
58. Phillips A. *COVID-19 Passes For Nightclubs And Large Events Come Into Force in England as Booster Jabs Open To All Adults* (Skynews, 2021) <https://news.sky.com/story/covid-19-passes-for-nightclubs-and-large-events-come-into-force-in-england-as-booster-jabs-open-to-all-adults-12496084>.
59. UK Health Security Agency. *Confirmatory PCR Tests to be Temporarily Suspended For Positive Lateral Flow Test Results.* <https://www.gov.uk/government/news/confirmatory-pcr-tests-to-be-temporarily-suspended-for-positive-lateral-flow-test-results> (2022).
60. Google. *Google Trends.* <https://trends.google.com/home> (2022).
61. Brighton and Hove City Council. *School Term And Holiday Dates.* <https://www.brighton-hove.gov.uk/schools-and-learning/brighton-hove-schools/school-term-and-holiday-dates> (2022).

## Acknowledgements

P.N. and Y.K. acknowledge support from the BBSRC, through the ERA-NET ICRAD program (BB/V019945/1). C.A.D. acknowledges support from the MRC Centre for Global Infectious Disease Analysis (MR/R015600/1) and the NIHR Health Protection Research Unit in Emerging and Zoonotic Infections (NIHR200907). C.A.D. is an investigator at the Pandemic Sciences Institute, University of Oxford.

## Author contributions

Y.K., P.N., and C.A.D. conceived the study. Y.K. and P.N. developed the conceptual framework and performed the analyses. C.A.D. reviewed and advised on the conceptual framework and the analyses. Y.K. wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-37813-1>.

**Correspondence** and requests for materials should be addressed to Pierre Nouvellet.

**Peer review information** *Nature Communications* thanks Julien Beauté and Iain Buchan for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023