Early View

Original research article

Broadening symptom criteria improves early case identification in SARS-CoV-2 contacts

Hamish Houston, Seran Hakki, Timesh D Pillay, Kieran Madon, Nieves Derqui-Fernandez, Aleksandra Koycheva, Anika Singanayagam, Joe Fenn, Rhia Kundu, Emily Conibear, Robert Varro, Jessica Cutajar, Valerie Quinn, Lulu Wang, Janakan S Narean, Mica R Tolosa-Wright, Jack Barnett, Onn Min Kon, Richard Tedder, Graham Taylor, Maria Zambon, Neil Ferguson, Jake Dunning, Jonathan J Deeks, Ajit Lalvani

Please cite this article as: Houston H, Hakki S, Pillay TD, *et al.* Broadening symptom criteria improves early case identification in SARS-CoV-2 contacts. *Eur Respir J* 2021; in press (<https://doi.org/10.1183/13993003.02308-2021>).

This manuscript has recently been accepted for publication in the *European Respiratory Journal*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJ online.

Title:

Broadening symptom criteria improves early case identification in SARS-CoV-2 contacts.

Authors:

Hamish Houston^{1*}, Seran Hakki^{1*}, Timesh D Pillay¹, Kieran Madon¹, Nieves Derqui-Fernandez¹, Aleksandra Koycheva¹, Anika Singanayagam², Joe Fenn¹, Rhia Kundu¹, Emily Conibear¹, Robert Varro¹, Jessica Cutajar¹, Valerie Quinn¹, Lulu Wang¹, Janakan S Narean¹, Mica R Tolosa-Wright¹, Jack Barnett¹, Onn Min Kon^{1,3}, Richard Tedder⁴, Graham Taylor⁵, Maria Zambon², Neil Ferguson⁶, Jake Dunning^{2,7}, Jonathan J Deeks^{8**}, Ajit Lalvani^{1**}

* HH & SH contributed equally

** JDe & AL contributed equally

Affiliations:

1. NIHR Health Protection Research Unit in Respiratory Infections, Imperial College London, London, UK
2. National Infection Service, Public Health England, London, UK
3. Tuberculosis Service, Imperial College Healthcare NHS Trust, London, UK
4. Molecular Diagnostic's Unit, Imperial College London, London, UK
5. Section of Virology, Department of Infectious Disease, Imperial College London, London, UK.
6. Department of Infectious Disease Epidemiology, Faculty of Medicine, Imperial College London, St Mary's Campus, Norfolk Place, London, UK.
7. NIHR Health Protection Research Unit in Emerging and Zoonotic Infections, University of Oxford, Oxford, UK
8. Test Evaluation Research Group, Institute of Applied Health Research, University of Birmingham, Birmingham, UK

Corresponding Author:

Ajit Lalvani

NIHR Health Protection Research Unit in Respiratory Infections, National Heart and Lung Institute, Imperial College London

a.lalvani@imperial.ac.uk

Word Count:

Abstract: 249; Main Text: 3000

KEYWORDS:

Symptom; case definition; case identification; COVID-19; SARS-CoV-2; community testing

TAKE HOME MESSAGE:

Data from two prospective longitudinal community-based cohorts show broadened symptom criteria identify more COVID-19 infections and reduce time-to-detection with minimal reductions in specificity, providing greater opportunities to prevent transmission.

ABSTRACT

Introduction

The success of case isolation and contact tracing for the control of SARS-CoV-2 transmission depends on the accuracy and speed of case identification. We assessed whether inclusion of additional symptoms alongside three canonical symptoms (CS) - fever; cough; loss or change in smell or taste – could improve case definitions and accelerate case identification in SARS-CoV-2 contacts.

Methods

Two prospective longitudinal London-based cohorts of community SARS-CoV-2 contacts, recruited within 5 days of exposure, provided independent training and test datasets. Infected and uninfected contacts completed daily symptom diaries from the earliest possible time-points. Diagnostic information gained by adding symptoms to the CS was quantified using likelihood ratios and AUC-ROC. Improvements in sensitivity and time-to-detection were compared to penalties in terms of specificity and number-needed-to-test.

Results

Of 529 contacts within two cohorts, 164 (31%) developed PCR-confirmed infection and 365 (69%) remained uninfected. In the training dataset (n=168), 29% of infected contacts did not report the CS. Four symptoms (sore throat, muscle aches, headache and appetite loss) were identified as early-predictors (EP) which added diagnostic value to the CS. The broadened symptom criterion “ ≥ 1 of the CS, or ≥ 2 of the EP” identified PCR-positive contacts in the test dataset on average 2 days earlier after exposure (p=0.07) than “ ≥ 1 of the CS”, with only modest reduction in specificity (5.7%).

Conclusions

Broadening symptom criteria to include individuals with at least 2 of muscle aches, headache, appetite loss and sore throat identifies more infections and reduces time-to-detection, providing greater opportunities to prevent SARS-CoV-2 transmission.

INTRODUCTION:

SARS-CoV-2 transmission is not entirely prevented by current vaccines.[1] As vaccination coverage increases, blanket isolation rules for SARS-CoV-2 contacts (i.e. test-trace-isolate) become less acceptable to society. Rapid identification and isolation of contacts who become infected is an increasingly important alternative strategy for prevention and containment.[2] Effectiveness depends crucially on how quickly such cases are detected and initiate self-isolation,[3] because individuals are most infectious early in the course of infection.[4, 5] However, case definitions must also be sufficiently specific to avoid overwhelming testing capacity.[6, 7]

There is considerable international heterogeneity in policy for COVID-19 community testing within the general population (S1). Most criteria include fever, cough, and loss or change in smell or taste (hereafter referred to as the canonical symptoms [CS]) alongside a range of other symptoms. Some countries are currently considering altering their case definitions.[13] There is thus an urgent need for empirical data to identify whether additional symptoms (we call early-predictors [EPs]) can augment the CS within community case definitions.

Surprisingly, empirical longitudinal data from recently exposed SARS-CoV-2 contacts are scarce. Recent large-scale cross-sectional studies of community testing data support adding more symptoms to the CS. [7, 14] Descriptive longitudinal retrospective studies of SARS-CoV-2 infections also exist.[15, 16] However, a high-resolution longitudinal evaluation of symptom combinations for differentiating infected SARS-CoV-2 contacts from exposed but uninfected controls has not, to our knowledge, been performed before.

Using data from two prospective longitudinal cohorts of SARS-CoV-2 contacts, we aimed to establish definitively whether broadening symptom criteria beyond the CS can accelerate and improve case detection without weakening specificity. Rapid recruitment following clearly defined exposure enabled optimal symptom criteria to be identified. Daily contemporaneously recorded symptom diaries ensured symptom onset times were recorded and time-savings measured with maximum precision. Through direct study of relevant community-based cohorts we provide generalisable evidence-based criteria for effective case definitions to rapidly identify and isolate infectious cases.

METHODS:

Recruitment and Study Procedures

INSTINCT (Integrated Network for Surveillance, Trials and Investigations into COVID-19 Transmission) and ATACCC (Assessment of Transmission And Contagiousness of COVID-19 in Contacts) were two community-based cohort studies in which contacts of COVID-19 cases in Greater London were identified and recruited from 10th May 2020 through 31st March 2021.

Index cases, or contacts identified by the UK contact tracing system (NHS Test and Trace [NTAT]), were referred from Public Health England (PHE). Initially, referrals were also received from the Royal College of General Practitioners Research and Surveillance Centre (RCGP-RSC) network. Contacts referred within five days of their index case symptom onset (ISO), who provided valid informed consent were enrolled within our recruitment capacity until the end of the 2nd pandemic wave in the UK. Ethics approval was granted by the Health Research Authority (Research Ethics Committee reference: 20/NW/0231).

In INSTINCT, household-contacts living with their index cases were enrolled at home by research nurses (day 0) and visited again on days 7, 14 and 27. Date of ISO was recorded at enrolment and served as a proxy for exposure. Combined nose and throat swabs (CNTS) for RT-PCR testing and blood samples for serology were taken by research nurses at each visit and an additional CNTS by participants on day 4. Samples were processed at the Molecular Diagnostics Unit, Imperial College London. Antibody (IgM and IgG) to SARS-CoV-2 receptor binding domain (anti-RBD) was measured using a two-step double antigen binding assay (DABA) with recombinant S1 antigen on the solid-phase and labelled recombinant RBD as detector in the fluid-phase.[16] In ATACCC, household and non-household-contacts (i.e. not residing with their index) were enrolled. Dates of ISO (household-contacts) or exposure event (non-household-contacts) were provided by NTAT. After nurse-delivered training, participants self-sampled CNTS daily for 14 consecutive days. SARS-CoV-2 RT-PCR testing was performed at the Virus Reference Department, PHE Colindale.

At enrolment, demographic information was collected and participants recorded the onset date of prior symptoms. After enrolment, participants completed a daily symptom diary which assessed 20 symptoms (S2). Loss or change in smell or taste was recorded as one item (hereafter referred to as anosmia).

Definitions and Reference Standards

INSTINCT data were used as the training dataset. 'Current infection' was set as the target condition and a rigorous composite reference standard was constructed to establish its presence or absence with maximum accuracy.[17] Contacts were assigned to the 'infected' group if they were PCR-positive at day 0, 4 or 7. Contacts were assigned to the 'uninfected' group if they were PCR-negative

and had undetectable SARS-CoV-2 antibodies at all time-points. Participants were excluded if they had no serology results or were PCR-negative at all time-points but had detectable SARS-CoV-2 antibodies at study day 0, 7 or 27.

ATACCC data were used as the test dataset. In this cohort, daily PCR results were available but serological testing was not performed routinely. Contacts were assigned to the 'PCR-positive' group if they had a positive PCR result by 7 days after enrolment and PCR-negative if all results were negative. Participants who became PCR-positive after study day 7 or had no PCR results were excluded from the analysis. Participants with only one positive PCR result with a high Ct value (>28) were excluded to minimise false-positives caused by recent rather than current infection.

In both cohorts, participants were made aware of their PCR results as they became available. Participants with missing ISO or exposure dates were excluded from analyses requiring these data. The study flowchart (Figure 1) depicts participant numbers included in each analysis.

[Figure 1]

Statistical Analyses

Time-to-event Analysis

We used time-to-event analysis to describe the onset of COVID-19-related symptoms relative to ISO in INSTINCT (Figure 1 cohort A). Briefly, we used symptoms reported by 'uninfected' contacts to define baseline time-dependent hazards, and the difference between 'infected' and 'uninfected' contacts to define COVID-19-related hazards for each symptom (S3 for detailed methodology). Symptoms with a probability of occurring due to COVID-19 of $>15\%$ by 10 days post-ISO were selected as candidate symptoms for further evaluation.

Spiegelhalter Knill-Jones Models

We aimed to quantify any additional diagnostic value gained by adding each of the candidate symptoms to the CS using likelihood ratios (LRs) estimated for individual symptoms within combinations of symptoms. The Spiegelhalter Knill-Jones (SKJ) method was used rather than the independence Bayes approach in order to adjust for dependency caused by symptom co-occurrence.[18] This method is summarised in S4, having been described in detail previously.[18–22] Symptoms were considered as a series of binary tests based on their occurrence by each study day (e.g. fever by day 3 would be regarded as positive if fever had been reported on study day 2). Persistent cough and productive cough were combined into a single cough variable. We compared models using the CS to those with an additional symptom. AUC-ROC allows evaluation of model

discrimination in training and test datasets. Candidate symptoms with useful LR_s after adjustment for dependency with the CS and whose addition improved AUC-ROC across multiple early time-points were considered 'early-predictors' (EP).

Evaluating Simple Case Definitions

To assess real-world impact through readily applicable case definitions, each of the EP were added to the CS individually and together as a list requiring more than 1 to be positive by using the words "at least". Diagnostic performance was assessed against the serial PCR reference standard in the test dataset (cohort D, Figure 1) at each day post-exposure. We used time-to-event analysis to measure how quickly broadened case definitions would identify PCR-positive individuals and log-rank tests to make comparisons with the CS. Finally, we quantified the prevalence-dependent trade-off between true-positives and false-positives by calculating the NNT: the number of false-positives for every true-positive plus 1.

Software

Statistical analyses were performed in Stata version 17.0 (StataCorp, LLC, College Station TX) and R (R Core Team, 2020).

RESULTS:

Patient cohorts:

53011 referrals were received via three recruitment pathways (S5). Of 529 contacts within two cohorts, 164 (31%) developed PCR-confirmed infection and 365 (69%) remained uninfected. S6 shows demographic details for INSTINCT and ATACCC. Sex, ethnicity and BMI were similar between cohorts. Participants were slightly older in ATACCC than in INSTINCT (median 38 vs 34 years, $p < 0.001$).

Participants were enrolled a median 3 days (IQR 2-4) post-ISO in INSTINCT, 4 days (IQR 4-5) post-ISO in household-contacts in ATACCC and 5 days (IQR 4-6) post-exposure-event in non-household-contacts. In INSTINCT 22/168 (13.1%) contacts were linked to index cases identified through the RCGP-RSC network with suspected, but not confirmed COVID-19 and were included to avoid selection bias. Symptom diaries were completed in INSTINCT for median 26 days (IQR 7-27) by

‘infected’ and 7 days (IQR 0-27) by ‘uninfected’ contacts, and in ATACCC for 13 days (IQR 0-20) by ‘PCR-positives’ and 7 days (IQR 0-14) by ‘PCR-negatives’.

Four participants with confirmed infection required hospitalisation. Whilst more than 90% of ‘infected’ contacts in INSTINCT (68/73, 93.1%) reported at least one of the 20 symptoms by day 7, over a quarter (21/73, 28.8%) did not report fever, cough or anosmia by day 7.

Sequence of onset of COVID-19 related symptoms:

Time-to-event analysis of symptom onset following exposure (S7-8) showed that fever preceded anosmia and persistent cough preceded productive cough. Sore throat and rhinitis occurred early, and breathlessness later. Fatigue was commonly reported by ‘uninfected’ contacts.

Thirteen symptoms had a probability of occurring due to COVID-19 of >15% by 10 days post-ISO (fever, persistent cough, productive cough, anosmia, headache, muscle aches, sore throat, rhinitis, appetite loss, breathlessness, diarrhoea, nausea and abdominal pain). 9 of these 13 symptoms are not included in the CS and were denoted candidate symptoms in further analyses. Other than the CS: rhinitis, sore throat, headache, muscle aches, and appetite loss had the largest cumulative COVID-19-related hazards.

Additional diagnostic value of candidate symptoms

Raw counts of participants who had reported each symptom by each study day in the training cohort are presented in S9. Used alone, cough, rhinitis, headache and muscle aches were the most sensitive symptoms whilst nausea and abdominal pain were insensitive (S10). Anosmia, fever, and appetite loss were highly specific symptoms.

The crude LRs (S11) show that any of the symptoms will affect post-test odds when they are used alone. However, when used in combination with other symptoms, their LRs after adjustment using the SKJ approach (Figure 2, Table 1, S12) were all less extreme than their crude LRs, indicating considerable dependency between symptoms.

[Figure 2]

[Table 1]

When cough was used in combination with fever and anosmia it’s adjusted LRs were closer to 1 than those of anosmia or fever (Figure 2A, S12). This was most likely due to the higher specificity of

anosmia and fever (S10). When combined with the CS the presence or absence of nausea did not independently affect post-test odds, its adjusted LRs lying close to 1 (Figure 2). Breathlessness was more common in the 'infected' group. However, whilst breathlessness was reported without fever, anosmia or cough by 'uninfected' contacts, this was rare in the 'infected', explaining why its adjusted positive LRs are below 1 and negative LRs above 1.

In training and test datasets, AUC-ROCs increased with study day reflecting improved discrimination afforded by greater accumulation of symptoms by later study days in 'infected' contacts (Figure 3). AUC-ROC was often greater in test data than training data, likely reflecting the longer median time to recruitment in ATACCC.

[Figure 3]

Between study days 0 to 3, the addition of headache, sore throat, muscle aches and appetite loss to the CS yielded the greatest improvements in AUC-ROC in the test dataset. When combined with the CS, appetite loss, headache, sore throat and muscle aches all consistently had positive adjusted LRs above 1 and negative adjusted LRs below 1 showing that both their presence and their absence added to the CS's ability to discriminate between the infected and uninfected. These symptoms were therefore considered 'early-predictors' (EP).

Evaluating Simple Case Definitions

Each of the four EP were combined individually with the CS using an "or" operator, as well as together using "or" and "at least" operators (Box 1).

[Box 1]

The addition of any symptom to the CS using an "or" operator increased sensitivity (Figure 4A) whilst reducing specificity (Figure 4B). The addition of appetite loss produced the smallest changes compared to the CS.

[Figure 4]

The CS identified 50% of PCR-positives by 6 days post exposure (Figure 5, Table 2, S14). Adding headache yields the greatest increase in sensitivity (Figure 4A, S13A) and would identify PCR-positives on average 2 days earlier ($p=0.02$), but causes the largest reduction in specificity (15.2% at 5 days post-exposure; Figure 4B, S13A). In contrast, *CS or sore throat* only reduced specificity by 5.7% at 5 days and identified PCR-positive cases earlier than the CS, by 1 day on average. This

change was not statistically significant given the small number of PCR-positive participants in the cohort ($p=0.1$, $n=91$).

[Figure 5]

[Table 2]

When all 4 EP are added to the CS, if all 4 are required, there is very little difference to the CS. In contrast, the case definition ≥ 1 of the CS, or ≥ 1 of the EP would increase sensitivity and identify PCR-positive cases a median 2 days earlier than the CS ($p=0.002$). However, the corresponding reduction in specificity by 5 days post-exposure (19.7%, Figure 4B, S13B) would lead 25% of PCR-negative individuals to be inappropriately identified (Table 2, Figure 5B). ≥ 1 of the CS, or ≥ 2 of the EP identified PCR-positive cases a median 2 days earlier than the CS ($p=0.07$) with a reduction in specificity of only 5.7% at 5 days post-exposure. This reduction is smaller than that caused by moving from the CS to various other international case definitions (S16). None of the EP were dispensable from this proposed criterion (S17).

The number of individuals identified in order to yield a single PCR-positive case, the NNT, increases rapidly immediately after exposure, reflecting an initial accumulation of false-positives because no one has yet developed symptoms actually caused by infection (Figure 4C). NNT plateaus around 4-5 days following exposure, reflecting the incubation period. At 25.6% prevalence, ≥ 1 of the CS, or ≥ 2 of the EP had a NNT at 5 days post-exposure of 1.78 compared to 1.61 for the CS, indicating 17 additional individuals identified for every 100 infected individuals identified.

DISCUSSION:

This, to our knowledge, is the first study to use daily symptom data prospectively collected from recently exposed infected and uninfected SARS-CoV-2 contacts to evaluate the diagnostic performance of symptom combinations for detecting infection.

Using this definitive study design, we found that 29% of individuals with PCR-confirmed COVID-19 did not report any of the CS, but 93% reported at least one symptom from a broader list of 20. We identified 4 EP symptoms (sore throat, headache, muscle aches, and appetite loss) providing additional early predictive power for identifying SARS-CoV-2-infected contacts. The case definition ≥ 1 of the CS, or ≥ 2 of the EP identified PCR-positive contacts 2 days earlier after exposure than the

CS alone ($p=0.07$). This time-saving is critical given that shortening the delay from infectiousness to self-isolation from 2.6 to 1.2 days has been estimated to reduce transmission by 47%.^[3] Moreover, the proportion of “symptomatic” infections and time-to-symptom onset are critical parameters in studies modelling effectiveness of testing and isolation strategies for contacts.^[24]

Consistent with previous studies, headache and sore throat were sensitive symptoms,^[7] which occurred early in the course of infection,^[15] and were prevalent in our relatively young participants.^[24] Importance of these symptoms will increase as vaccination of older age groups increases the proportion of infections occurring in the young. In agreement with the Real-time Assessment of Community Transmission-1 (REACT-1) study, we found that headache, muscle aches and appetite loss improved discrimination within statistical prediction models.^[14] We add a crucial evaluation of readily applicable case definitions. We observed that both the structure of symptom criteria (e.g. use of the Boolean operator “at least”) and time-from-exposure had a considerable effect on diagnostic performance.

The SKJ approach enabled another important new observation. Although an important indicator of disease severity,^[26] breathlessness was not a useful additional symptom for identifying early and mild infections because a hierarchy of symptoms exists. Breathlessness is unlikely to occur due to COVID-19 without prior fever, cough or anosmia and its inclusion reduces specificity.

Further strengths include day-by-day measurement of diagnostic performance following exposure and prospective data collection which mitigates recall bias. The rigorous reference standard employed in our training cohort maximised accuracy for the target condition and ensured only the most useful symptoms were taken forward to the test data. Neither serology nor PCR have 100% sensitivity for SARS-CoV-2 infection.^[27] [Click or tap here to enter text.](#) Using both serology and PCR at multiple time-points to define the absence of infection we minimised false-negatives. False-positive PCR results caused by recent rather than current infection were likely less common in our longitudinal study of recently exposed contacts than in studies involving random community-sampling.^[14]

Limitations include modest sample size, largely White British population, minor differences between training and test cohorts and the potential for tick-box and behavioural biases. Study participants were usually highly motivated and attentiveness to mild symptoms (e.g. rhinitis) may have been increased by awareness of exposure, frequent study visits and co-residence with other participants. Contacts could not be blinded to their PCR results or those of their index.

Since we studied community-based COVID-19 contacts identified through NTAT, our findings are very likely generalizable. As large-scale cross-sectional data replicate our findings in smaller-scale daily-resolution longitudinal data, the combined evidence-base is now sufficient to influence policy. Broadening symptom criteria for use in the general population would likely identify more infections and reduce time-to-detection, reducing transmission. We propose that symptom criteria within case definitions to prompt symptomatic isolation and testing of SARS-CoV-2 contacts should include headache, sore throat, muscle aches and appetite loss as well as the canonical symptoms to optimise sensitivity. Two of these additional symptoms should be required to maximise specificity.

As highly vaccinated regions transition to lower COVID-19 incidence, investment in RT-PCR testing capacity will make such broader case definitions feasible. As societies develop alternatives to test-trace-isolate, application of evidence-based symptom criteria alongside judicious testing will be critical for early discrimination of infected and uninfected contacts. Accordingly, our findings should inform development of evidence-based national testing policies in many parts of the world now and in subsequent phases of the pandemic.

REPORTING STANDARDS:

Results are presented in accordance with STARD (Standards for Reporting of Diagnostic Accuracy Studies) and STROBE (Strengthening Reporting of Observational Studies in Epidemiology) guidelines.

FUNDING STATEMENT:

This work is supported by a Department of Health and Social Care COVID-19 Fighting Fund award (ATACCC), by the NIHR Health Protection Research Unit in Respiratory Infections at Imperial College London in partnership with Public Health England (award NIHR200927) and by MRC/UKRI; nCoV: Serological detection of past SARS-CoV-2 infection by non-invasive sampling for field epidemiology and quantitative antibody detection.

AUTHOR CONTRIBUTIONS:

HH, SH & AL made substantial contribution to the conception of the work. TP, RK, RV, JF, HH, SH and AL contributed to study design and data collection methodology. HH, SH, AL, NF & JDe made substantial contribution to the design of the analyses. RV, MT-W, JB, VQ, JC, JDu, EC, LW, JN, JF, RK, SH, ND-F, AK, GT & MZ delivered and supported the administrative and technical study logistics. SH, ND-F, AK, & KM carried out the initial data acquisition. RT provided the serology data. GT & MZ provided the RT-PCR data. HH, KM & JDe analysed the data. Statistical analyses were checked by JDe. HH, SH, KM, JDe, AS, JDu, OMK & AL contributed to data interpretation. HH & SH drafted the manuscript. AL was the Principal Investigator for INSTINCT. AL and JDu were joint Principal Investigators for ATACCC. All authors contributed to revising the manuscript critically for important intellectual content, approved the final manuscript and are accountable for all aspects of the work.

COMPETING INTERESTS STATEMENT:

No competing interests were declared by any of the study authors.

ACKNOWLEDGEMENTS:

We would like to thank all the participants who were involved in the study and the support of our study research nurses Kristel Timcang, Sarah Hammett, Eimear McDermott, Constanta Luca, Sam Bremang Jada Samuel. We thank Myra McClure and Ellie Parker on behalf of the Molecular Diagnostic's Unit, Imperial College London. We thank Joanna Ellis and the Virus Reference Department Staff, and Lucy Moss crop, Carolina Rosadas and the Molecular Diagnostic's Unit staff for providing SARS-CoV-2 PCR data. We thank PHE England staff for facilitating recruitment into the study. We thank Michael Whitfield, Mohammed Essoussi, Chitra Tejpai, Guilia Miserocchi, Harriet Catchpole and Anjeli Ketkar for conducting data entry and quality control, Holly Grey and Megan

Davies for providing further logistical support and Charlotte Williams and Michael Whitfield for performing a literature search of the SARS-CoV-2 symptom case definitions used worldwide.

REFERENCES:

1. Singanayagam A, Hakki S, Dunning J, Madon KJ, Crone MA, Koycheva A, Derqui-Fernandez N, Barnett JL, Whitfield MG, Varro R, Charlett A, Kundu R, Fenn J, Cutajar J, Quinn V, Conibear E, Barclay W, Freemont PS, Taylor GP, Ahmad S, Zambon M, Ferguson NM, Lalvani A, Badhan A, Dustan S, Tejpal C, Ketkar A v, Narean JS, Hammett S, McDermott E, et al. Community transmission and viral load kinetics of the SARS-CoV-2 delta (B.1.617.2) variant in vaccinated and unvaccinated individuals in the UK: a prospective, longitudinal, cohort study. *The Lancet Infectious Diseases* [Internet] 2021; Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1473309921006484>.
2. World Health Organization. Contact tracing in the context of COVID-19. World Health Organization; 1 Feb 2021 [cited 5 July 2021]. Available: <https://www.who.int/publications/i/item/contact-tracing-in-the-context-of-covid-19>.
3. Kucharski AJ, Klepac P, Conlan AJK, Kissler SM, Tang ML, Fry H, Gog JR, Edmunds WJ, Emery JC, Medley G, Munday JD, Russell TW, Leclerc QJ, Diamond C, Procter SR, Gimma A, Sun FY, Gibbs HP, Rosello A, van Zandvoort K, Hué S, Meakin SR, Deol AK, Knight G, Jombart T, Foss AM, Bosse NI, Atkins KE, Quilty BJ, Lowe R, 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. *The Lancet Infectious Diseases* Lancet Publishing Group; 2020; 20: 1151–1160.
4. Bullard J, Dust K, Funk D, Strong JE, Alexander D, Garnett L, Boodman C, Bello A, Hedley A, Schiffman Z, Doan K, Bastien N, Li Y, van Caeseele PG, Poliquin G. Predicting infectious severe acute respiratory syndrome coronavirus 2 from diagnostic samples. *Clinical Infectious Diseases* Oxford University Press; 2020; 71: 2663–2666.
5. Salvatore PP, Dawson P, Wadhwa A, Rabold EM, Buono S, Dietrich EA, Reses HE, Vuong J, Pawloski L, Dasu T, Bhattacharyya S, Pevzner E, Hall AJ, Tate JE, Kirking HL. Epidemiological Correlates of Polymerase Chain Reaction Cycle Threshold Values in the Detection of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). *Clinical Infectious Diseases* Oxford University Press (OUP); 2021; 72: e761–e767.
6. Fragaszy E, Shrotri M, Geismar C, Aryee A, Beale S, Braithwaite I, Byrne T, Lam W, Fong E, Gibbs J, Hardelid P, Kovar J, Lamos V, Nastouli E, Navaratnam MD, Nguyen V, Patel P, Aldridge RW, Hayward A, Wijlaars L, Spyer M, Killingley B, Cox I, Mckendry RA, Cheng T, Liu Y, Johnson AM, Gilson R. Symptom profiles and accuracy of clinical definitions for COVID-19 in the community. Results of the Virus Watch community cohort. on behalf of Virus Watch Collaborative*. *MedRxiv [pre-print]* [Internet] 2021; Available from: <https://doi.org/10.1101/2021.05.14.21257229>.
7. Antonelli M, Capdevila J, Chaudhari A, Granerod J, Canas LS, Graham MS, Klaser K, Modat M, Molteni E, Murray B, Sudre CH, Davies R, May A, Nguyen LH, Drew DA, Joshi A, Chan AT, Cramer JP, Spector T, Wolf J, Ourselin S, Steves CJ, Loeliger AE. Optimal symptom combinations to aid COVID-19 case identification: Analysis from a

community-based, prospective, observational cohort. *Journal of Infection* W.B. Saunders Ltd; 2021; 82: 384–390.

8. World Health Organisation. WHO COVID-19 Case definition. [cited July 4, 2021]. Available: https://www.who.int/publications-detail-redirect/WHO-2019-nCoV-Surveillance_Case_Definition-2020.2. .
9. Centers for Disease Control and Prevention. Coronavirus disease 2019 (covid-19) 2020 interim case definition. 2020. [cited 5 July 2021]. Available: <https://wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-covid-19/case-definition/2020/08/05/>. .
10. Gouvernement de la République française. Tester- Alerter- Protéger (2021) [cited 15th July 2021] Available: <https://www.gouvernement.fr/info-coronavirus/tests-et-depistage>. . .
11. Ministerio de Sanidad and Instituto de Salud Carlos III. Estrategia de detección precoz, vigilancia y control de covid-19. (2021) [cited 14th July 2021] Available: https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/documentos/COVID19_Estrategia_vigilancia_y_control_e_indicadores.pdf. .
12. UK Government. Get a free PCR test to check if you have coronavirus [cited 13 July, 2021], Available: <https://www.gov.uk/get-coronavirus-test>. .
13. Crozier A, Dunning J, Rajan S, Semple MG, Buchan IE. Could expanding the covid-19 case definition improve the UK's pandemic response? *BMJ* [Internet] 2021; : n1625 Available from: <https://www.bmj.com/lookup/doi/10.1136/bmj.n1625>.
14. Elliott J, Whitaker M, Bodinier B, Eales O, Riley S, Ward H, Cooke G, Darzi A, Chadeau-Hyam M, Elliott P. Predictive symptoms for COVID-19 in the community: REACT-1 study of over 1 million people. *PLOS Medicine* Public Library of Science (PLOS); 2021; 18: e1003777.
15. Wiegele PN, Kabar I, Kerschke L, Froemmel C, Hüsing-Kabar A, Schmidt H, Vorona E, Vollenberg R, Tepasse PR. Symptom diary-based analysis of disease course among patients with mild coronavirus disease, Germany, 2020. *Emerging Infectious Diseases* Centers for Disease Control and Prevention (CDC); 2021; 27: 1353–1361.
16. Khan SM, Farland L v, Austhof E, Bell ML, Catalfamo CJ, Chen Z, Cordova-Marks F, Ernst KC, Garcia-Filion P, Heslin KM, Hoskinson J, Jehn ML, Joseph ECS, Kelley CP, Klimentidis Y, Carroll SR, Kohler LN, Pogreba-Brown K, Jacobs ET. Symptoms of COVID-19 in a population-based cohort study. *medRxiv [pre-print]* [Internet] 2021; Available from: <https://doi.org/10.1101/2021.03.20.21254040>.
17. Rosadas C, Randell P, Khan M, McClure M, Tedder R. Testing for responses to the wrong SARS-CoV-2 antigen? *Lancet* NLM (Medline); 2020; 396: 1291–1292.
18. Medicines and Healthcare products Regulatory Agency (B) (MHRA). Target product profile: In Vitro Diagnostic (IVD) self-tests for the detection of SARS-CoV-2 in people

without symptoms. [cited 14 July 2021] Available:

<https://www.gov.uk/government/publications/how-tests-and-testing-kits-for-coronavirus-covid-19-work>. .

19. Chan SF, Deeks JJ, Macaskill P, Irwig L. Three methods to construct predictive models using logistic regression and likelihood ratios to facilitate adjustment for pretest probability give similar results. *Journal of Clinical Epidemiology* 2008; 61: 52–63.
20. Spiegelhalter DJ, Knill-Jones RP. Statistical and Knowledge-Based Approaches to Clinical Decision-Support Systems, with an Application in Gastroenterology. *Journal of the Royal Statistical Society. Series A (General)* 1984; 147: 35–77.
21. Spiegelhalter DJ. Probabilistic prediction in patient management and clinical trials. *Statistics in Medicine* 1986; 5: 421–433.
22. R P Knill-Jones. Diagnostic systems as an aid to clinical decision making. *BMJ* 1987; 295: 1392–1396.
23. Seymour DG, Green M, Vaz FG. Making better decisions: Construction of clinical scoring systems by the Spiegelhalter-Knill-Jones approach. *British Medical Journal* 1990; 300: 223–226.
24. Quilty BJ, Clifford S, Hellewell J, Russell TW, Kucharski AJ, Flasche S, Edmunds WJ, Atkins KE, Foss AM, Waterlow NR, Abbas K, Lowe R, Pearson CAB, Funk S, Rosello A, Knight GM, Bosse NI, Procter SR, Gore-Langton GR, Showering A, Munday JD, Sherratt K, Jombart T, Nightingale ES, Liu Y, Jarvis CI, Medley G, Brady O, Gibbs HP, Simons D, et al. Quarantine and testing strategies in contact tracing for SARS-CoV-2: a modelling study. *The Lancet Public Health* Elsevier Ltd; 2021; 6: e175–e183.
25. Swann O v., Holden KA, Turtle L, Pollock L, Fairfield CJ, Drake TM, Seth S, Egan C, Hardwick HE, Halpin S, Girvan M, Donohue C, Pritchard M, Patel LB, Ladhani S, Sigfrid L, Sinha IP, Olliaro PL, Nguyen-Van-Tam JS, Horby PW, Merson L, Carson G, Dunning J, Openshaw PJM, Baillie JK, Harrison EM, Docherty AB, Semple MG. Clinical characteristics of children and young people admitted to hospital with covid-19 in United Kingdom: Prospective multicentre observational cohort study. *The BMJ* BMJ Publishing Group; 2020; 370.
26. Borges do Nascimento IJB, von Groote TC, O’Mathúna DP, Abdulazeem HM, Henderson C, Jayarajah U, Weerasekara I, Pericic TP, Gerald Klapproth HE, Puljak L, Cacic N, Zakarija-Grkovic I, Meirelles Guimarães SM, Atallah AN, Bragazzi NL, Marcolino MS, Marusic A, Jeroncic A. Clinical, laboratory and radiological characteristics and outcomes of novel coronavirus (SARS-CoV-2) infection in humans: A systematic review and series of meta-analyses. *PLoS ONE* Public Library of Science; 2020; 15.
27. Long QX, Tang XJ, Shi QL, Li Q, Deng HJ, Yuan J, Hu JL, Xu W, Zhang Y, Lv FJ, Su K, Zhang F, Gong J, Wu B, Liu XM, Li JJ, Qiu JF, Chen J, Huang AL. Clinical and immunological

assessment of asymptomatic SARS-CoV-2 infections. *Nature Medicine* Nature Research; 2020; 26: 1200–1204.

TABLES & BOXES:

Table 1: Adjusted Likelihood Ratios for individual symptoms when canonical symptoms are used in combination with an additional candidate symptom.

LRs for individual symptoms in combination with other symptoms were calculated using the Spiegelhalter Knill-Jones method which adjusts for the dependency caused by symptom co-occurrence. Adjusted LRs for study days 0-7 are represented graphically in Figure 2 and values for study days 0, 2 & 4 are presented here. For any combination of symptoms by a particular time-point, the adjusted LRs can be multiplied together to estimate effect on posttest odds. For example, on study day 0, a contact with fever and headache but without cough or anosmia has $3.22 \times 0.94 \times 0.84 \times 1.37 = 3.90$ times the odds of being infected compared to a contact where symptoms are unknown (see S4 for worked example). LR = likelihood ratio, CS = canonical symptoms

Candidate added to CS	Study Day	Symptom	Adjusted Positive LR	Adjusted Negative LR	Candidate added to CS	Study Day	Symptom	Adjusted Positive LR	Adjusted Negative LR	Candidate added to CS	Study Day	Symptom	Adjusted Positive LR	Adjusted Negative LR
abdominal pain	0	fever	3.54	0.90	appetite loss	0	fever	3.68	0.90	breathlessness	0	fever	3.10	0.91
		cough	1.48	0.90			cough	1.29	0.94			cough	2.12	0.82
		anosmia	10.73	0.84			anosmia	9.46	0.85			anosmia	11.94	0.83
		abdominal pain	1.19	0.98			appetite loss	1.96	0.92			breathlessness	0.31	1.09
abdominal pain	2	fever	3.04	0.80	appetite loss	2	fever	3.00	0.80	breathlessness	2	fever	3.69	0.76
		cough	1.62	0.82			cough	1.44	0.86			cough	1.88	0.77
		anosmia	4.42	0.81			anosmia	4.39	0.81			anosmia	5.80	0.78
		abdominal pain	1.58	0.94			appetite loss	2.22	0.84			breathlessness	0.52	1.11
abdominal pain	4	fever	3.83	0.71	appetite loss	4	fever	3.52	0.73	breathlessness	4	fever	5.48	0.65
		cough	1.33	0.87			cough	1.22	0.90			cough	1.56	0.80
		anosmia	7.35	0.69			anosmia	6.17	0.71			anosmia	11.10	0.64
		abdominal pain	1.04	0.99			appetite loss	2.01	0.83			breathlessness	0.34	1.31
diarrhoea	0	fever	4.08	0.89	headache	0	fever	3.22	0.91	muscle aches	0	fever	3.64	0.90
		cough	1.26	0.94			cough	1.27	0.94			cough	1.29	0.94
		anosmia	11.15	0.84			anosmia	10.29	0.84			anosmia	10.70	0.84
		diarrhoea	2.88	0.93			headache	1.37	0.89			muscle aches	1.37	0.91
diarrhoea	2	fever	2.93	0.80	headache	2	fever	2.69	0.82	muscle aches	2	fever	2.83	0.81

		cough	1.38	0.88			cough	1.44	0.86			cough	1.35	0.88
		anosmia	5.43	0.79			anosmia	4.29	0.81			anosmia	4.75	0.80
		diarrhoea	2.81	0.88			headache	1.33	0.79			muscle aches	1.76	0.72
diarrhoea	4	fever	3.82	0.72	headache	4	fever	3.43	0.73	muscle aches	4	fever	3.36	0.74
		cough	1.31	0.87			cough	1.20	0.91			cough	1.21	0.91
		anosmia	7.35	0.69			anosmia	6.61	0.70			anosmia	7.05	0.69
		diarrhoea	1.08	0.98			headache	1.30	0.76			muscle aches	1.37	0.80
nausea	0	fever	3.54	0.90	rhinitis	0	fever	2.75	0.92	sore throat	0	fever	2.75	0.92
		cough	1.53	0.89			cough	1.28	0.94			cough	1.28	0.94
		anosmia	10.86	0.84			anosmia	6.17	0.87			anosmia	6.17	0.87
		nausea	1.02	1.00			rhinitis	2.57	0.80			sore throat	2.57	0.80
nausea	2	fever	3.01	0.80	rhinitis	2	fever	2.95	0.80	sore throat	2	fever	2.95	0.80
		cough	1.56	0.83			cough	1.40	0.87			cough	1.40	0.87
		anosmia	5.08	0.79			anosmia	3.07	0.85			anosmia	3.07	0.85
		nausea	1.30	0.97			rhinitis	1.99	0.69			sore throat	1.99	0.69
nausea	4	fever	3.81	0.72	rhinitis	4	fever	3.33	0.74	sore throat	4	fever	3.33	0.74
		cough	1.31	0.87			cough	1.06	0.97			cough	1.06	0.97
		anosmia	7.36	0.69			anosmia	5.58	0.72			anosmia	5.58	0.72
		nausea	1.08	0.99			rhinitis	2.29	0.58			sore throat	2.29	0.58

Box 1: Construction of easily comprehensible case definitions using Boolean operators Case definitions can contain Boolean logical operators such as: <i>AND</i> , <i>OR</i> and <i>at least</i> (see supplementary box 1 for examples in current use)	
Symptom groups:	Referred to in text as:
<i>fever, cough, anosmia</i>	Canonical Symptoms (CS)
<i>headache, sore throat, muscle aches, appetite loss</i>	Early-predictors (EP)
The ‘early-predictors’ were each combined individually with the CS using an “OR” operator.	
<i>fever OR cough OR anosmia OR headache</i>	CS or headache
<i>fever OR cough OR anosmia OR sore throat</i>	CS or sore throat
<i>fever OR cough OR anosmia OR muscle aches</i>	CS or muscle aches
<i>fever OR cough OR anosmia OR appetite loss</i>	CS or appetite loss
The ‘early-predictors’ were all combined together with the CS using “at least” and “OR” operators.	
<i>fever OR cough OR anosmia OR AT LEAST 1 OF headache, sore throat, muscle aches and appetite loss</i>	≥1 of the CS, or ≥1 of the EP
<i>fever OR cough OR anosmia OR AT LEAST 1 OF headache, sore throat, muscle aches and appetite loss</i>	≥1 of the CS, or ≥2 of the EP
<i>fever OR cough OR anosmia OR AT LEAST 1 OF headache, sore throat, muscle aches and appetite loss</i>	≥1 of the CS, or ≥3 of the EP
<i>fever OR cough OR anosmia OR AT LEAST 1 OF headache, sore throat, muscle aches and appetite loss</i>	≥1 of the CS, or ≥4 of the EP

Table 2: Time until positive identification by simple case definitions including the canonical symptoms (CS) and broadened symptom criteria. Contacts were assigned to the ‘PCR-positive’ group if they had a positive PCR result by 7 days after enrolment and PCR-negative if all results were negative. The ‘early-predictors’ (sore throat, headache, muscle aches, appetite loss) were each combined individually with the canonical symptoms (CS) using an “OR” operator and all were added together using “at least” and “OR” operators (as described in Box 1).

Case definition	PCR-positive			PCR-negative		
	Time until 25% identified (days post exposure)	Time until 50% identified (days post exposure)	Time until 75% identified (days post exposure)	Time until 25% identified (days post exposure)	Time until 50% identified (days post exposure)	Time until 75% identified (days post exposure)
fever OR cough OR anosmia (CS)	2	6	10	never	never	never
CS OR sore throat	2	5	8	never	never	never
CS OR headache	2	4	7	6	never	never
CS OR muscle aches	2	5	9	never	never	never
CS OR appetite loss	2	6	9	never	never	never
CS OR at least 1 of the Early-predictors	2	4	7	5	never	never
CS OR at least 2 of the Early-predictors	2	4	8	never	never	never
CS OR at least 3 of the Early-predictors	2	5	9	never	never	never
CS OR at least 4 of the Early-predictors	2	6	9	never	never	never

Figure 1: Flow diagram for the inclusion and exclusion of INSTINCT and ATACCC study participants for each analysis. Cohort A was used for a time-to-event analysis describing symptom onset in time post index symptom onset (ISO). Cohorts B and C were used to create Spiegelhalter Knill-Jones models for each study day. Cohort D was used to evaluate the performance of simple case definitions at different time-points following exposure.

383 participants were recruited within INSTINCT, of which 138 were indexes and were excluded. 6/245 contacts were excluded because of missing PCR data or enrolment symptom data and 1 was excluded as they became PCR-positive after study day 7. 73/238 contacts were PCR-positive and assigned to the 'infected' group. 43/165 PCR-negative contacts were seropositive at study day 0 or 7 (possible prior infection or vaccination) and 2/165 seroconverted at day 27, (possible separate exposure event) and were excluded. 0/168 PCR-negative participants seroconverted at study day 14. 25/165 PCR-negative participants had no serology data available and were excluded. 21/168 contacts were excluded from cohort A due to missing ISO date, and 1 was excluded as they reported symptoms several days prior to ISO.

411 contacts were recruited in ATACCC. 31 were excluded because of missing PCR or enrolment symptom data and 1 was excluded as they became PCR-positive after study day 7. 15 were excluded because they were PCR-positive at only one time-point with a high Ct value. 91 of the remaining 361 contacts had at least one positive PCR result by study day 7. 6/361 were excluded from cohort D because of missing exposure date.

Figure 2: Adjusted likelihood ratios for individual symptoms within symptom combinations. Symptoms were considered as a series of binary tests based on their occurrence by each study day. Spiegelhalter Knill-Jones models were created using (A) 3 predictors (fever, cough, anosmia) to evaluate the diagnostic performance of the canonical symptoms (CS), and (B) 4 predictors to evaluate the effect of adding one of the 9 candidate variables to the CS (fever, cough, anosmia, candidate). Models were created for the day of enrolment and each of the first 7 study days. Positive and negative LR for the presence or absence of each symptom by each study day were calculated (S11) and then adjusted for dependency with the other predictors within the model to measure the independent predictive value of each symptom within the symptom combination. See S4 for full description of the SKJ method and a worked example. Adjusted LR for study days 0, 2 & 4 are presented in Table 1. Adjusted positive LR are shown on the left and adjusted negative LR on the right. In each plot the horizontal line drawn is drawn at 1. LR lying above the line increase post-test odds and those below the line reduce post-test odds. Bootstrap CIs for adjusted LR could not be calculated because some bootstrap iterations resulted in samples with singularities. LR = Likelihood Ratio, CI = Confidence Interval

Figure 3: Direct Comparison of Area under the Receiver Operating Characteristic Curve in Training and Test Datasets Symptoms were considered as a series of binary tests based on their occurrence by each study day. A series of Spiegelhalter Knill-Jones models were created, one for each study day, using 3 predictors (fever, cough, anosmia) to evaluate the canonical symptoms (CS) (black points). Nine further series of models were created using 4 predictors (fever, cough, anosmia, candidate) to evaluate the effect of adding one of the 9 candidate variables to the CS at each study day (coloured points). Model predictions were evaluated in training and test datasets by calculating area under the receiver operating characteristic curves. AUC Train = Area under the receiver operating characteristic curve in training dataset (cohort B, Figure 1). AUC Test = Area under the receiver operating characteristic curve in test dataset (cohort C, Figure 1). The hard-line marks AUC Test for the CS model. The dotted-line marks AUC Train for the CS. Models where the addition of a candidate symptom yielded better predictions in the training dataset lie to the right of the dotted-line and models where better predictions were yielded in the test dataset lie above the hard-line.

Figure 4: Sensitivity, specificity and number-needed-to-test for the canonical symptoms and broadened symptom criteria. The 'early-predictors' (EP): sore throat, headache, muscle aches, appetite loss, were each combined individually with the canonical symptoms (CS): fever, cough and anosmia, using an "OR" operator and all were added together using "at least" and "OR" operators (as described in Box 1). Sensitivity [A], specificity [B] and number-needed-to-test [C] were calculated for each symptom criterion by day post exposure (ISO for household-contacts) against a serial PCR reference standard. Full results are given in S13. Number-needed-to-test is calculated by dividing the number of false-positives by the number of truepositives and adding 1. Rarely, symptoms were reported at enrolment without an onset date. We imputed onset dates for these symptoms by assuming the median number of days pre-enrolment (maximum 2 participants [0.55%] for rhinitis).

Figure 5: Kaplan-Meier plots showing time from exposure until positive identification by the canonical symptoms and broadened symptom criteria. The ‘early-predictors’ (sore throat, headache, muscle aches, appetite loss) were each combined individually with the canonical symptoms (CS) using an “OR” operator and all were added together using “at least” and “OR” operators (as described in Box 1). The proportion of PCR-positive (left) and PCR-negative (right) participants who were positively identified by each case definition by each day following exposure (ISO for household-contacts) is shown using a Kaplan-Meier plot. The plot for the CS is shown in black. Life-tables are presented in S14 and median time-to-diagnosis in Table 2. Rarely, symptoms were reported at enrolment without an onset date. We imputed onset dates for these symptoms by assuming the median number of days pre-enrolment (maximum 2 participants [0.55%] for rhinitis).

SUPPLEMENTARY MATERIALS:

S1: International Variation in Symptom Criteria within Case Definitions for suspected COVID-19 within the general community population (as of 16th July 2021):

World Health Organisation	1) Fever AND Cough
	2) AT LEAST 3 OF: Fever Cough Fatigue Headache Myalgia Sore throat Coryza Anorexia/nausea/vomiting Dyspnoea Diarrhoea Altered mental status
	3) Loss or change in smell or taste AND Anorexia/Nausea/Vomiting in the absence of any other identified cause
United States (Center for Disease Control)	1) Cough OR Dyspnoea OR Shortness of breath OR Loss or change in smell or taste
	2) AT LEAST 2 OF: Fever Chills Rigors Myalgia Headache Sore throat Diarrhoea Fatigue Coryza Nausea/vomiting
France	Dyspnoea OR Cough OR Fever OR Loss or change in smell or taste Fatigue OR Sore throat OR Headache OR Myalgia
Spain	Dyspnoea OR Cough OR Fever OR Loss or change in smell or taste Diarrhoea OR Sore throat OR Headache OR Myalgia
United Kingdom	Cough OR Fever OR Loss or change in smell or taste

Exact wordings of criteria can be found at: World Health Organization,[7] United States (Centers for Disease Control),[8] France,[9] Spain,[10] United Kingdom.[11]

SUPPLEMENTARY METHODS:

S2: Precise Wording of Symptom Items within Symptom Diary. The order in which items were asked about within the symptom diary is given in the right-hand column. The names for each symptom item used in the main text are given in the column entitled *Variable name*.

Type	Variable Name	Wording within symptom diary	Order within symptom diary
		In the last 24 hours, have you...	
Canonical Symptom	fever	felt feverish?	1
	persistent cough	had a persistent cough lasting for more than 4 hours?	2
	productive cough	coughed up phlegm or mucous?	3
	anosmia	had loss or change of smell or taste?	17
Lower Respiratory	breathlessness	felt unusually dyspnoea?	5
	wheeze	had a wheeze?	20
	cough blood	coughed blood?	4
	chest pain	had chest pain?	11
Upper Respiratory	rhinitis	had a runny or blocked nose?	14
	sore throat	had a sore throat?	13
	conjunctivitis	had red or irritated eyes?	16
Gastrointestinal	abdominal pain	had unusual tummy pains?	19
	appetite loss	had loss of appetite?	18
	nausea	felt sick or vomited?	7
	diarrhoea	had diarrhoea?	10
Systemic	headache	had a headache?	12
	muscle aches	had unusual muscle aches?	6
	fatigue	felt tired or generally unwell?	8
	confusion	been confused?	9
	rash	had a rash?	15

S3: Describing symptom onset using time-to-event analysis

We aimed to identify symptoms with a high probability of occurring due to COVID-19 by 10 days post index symptom onset. Index symptom onset (ISO) was defined as the onset date of the first of any of the 20 symptoms. We used time-to-event analysis to describe the onset of COVID-19 related symptoms by comparing symptom onset in ‘infected’ and ‘uninfected’ contacts in cohort A (Figure 1).

Participants were excluded from this analysis if the date of index symptom onset was unknown, or if the contact reported symptoms which occurred more than 3 days prior to index symptom onset. Rarely, symptoms were reported at enrolment without an onset date. We imputed onset dates for these symptoms using the median number of days pre-enrolment (maximum 9 participants [5.4%] for muscle aches). Participants with incomplete symptom diaries were right-censored on the day their symptom diary ended.

We used symptoms reported by ‘uninfected’ contacts to define a baseline time-dependent hazard and calculated the difference in hazards between ‘infected’ and ‘uninfected’ contacts to represent the COVID-19-related time-dependent hazard of onset of each symptom.

$$h(t) = h^+(t) - h^-(t)$$

Cumulative COVID-19-related hazards $H(t)$ were calculated by adding up the daily $h(t)$ up to time t . The probability of an ‘infected’ individual experiencing a symptom as a result of their infection by a particular time t , was estimated by calculating 1 minus the survival probability $S(t)$:

$$1 - S(t) = 1 - (\exp[-H(t)]).$$

We selected symptoms where this probability was more than 15% at $t = 10$ days as candidate symptoms for further evaluation. Bootstrap 90% confidence intervals were calculated for $h(t)$ and $H(t)$.

S4: Spiegelhalter Knill-Jones Method

Likelihood ratios (LRs) have clinical value as they quantify how much a diagnostic test result (or symptom) will raise or lower the pre-test probability of the target condition. Bayes's theorem (by which LRs are calculated) assumes that diagnostic tests act independently. However, symptoms often co-occur and this dependency must be accounted for to avoid overoptimistic evaluations of diagnostic performance. Spiegelhalter and Knill-Jones combined Bayes's theorem with logistic regression to produce LRs which are adjusted for dependency.

We used Spiegelhalter Knill-Jones models to study the predictive power of symptom combinations and measure the diagnostic usefulness of each symptom within the combinations. The analysis was performed using Cohort B as the training dataset and Cohort C as the test dataset (Figure 1).

Symptoms were considered as diagnostic tests based on their occurrence by each study day for the day of recruitment (day 0) and the first 7 days thereafter. Diagnostic performance is assessed compared to a composite serial PCR and serology reference standard in the training dataset and a serial PCR reference standard in the test dataset (described in the main text).

A series of 8 Spiegelhalter Knill-Jones models were created, one for each study day, using 3 predictors (fever, cough, anosmia) to evaluate the canonical symptoms (CS). Nine further sets of 8 models were created using 4 predictors (fever, cough, anosmia, and the candidate symptom) to evaluate the effect of adding one of the 9 candidate variables to the 3 CS at each study day. Adjusted LRs were calculated for each of the canonical symptoms (fever, cough and anosmia) when used together, and for each candidate symptom when added to the canonical symptoms.

In brief, rather than including sets of (0,1) indicator variables for each symptom in a logistic regression, the indicator variables take the value of the crude positive and negative likelihood ratios LR^{crude} in a logistic regression equation of the form:

$$\log(posttest\ odds) = \log(pretest\ odds) + \gamma_1 \log LR_1^{crude} + \gamma_2 \log LR_2^{crude}$$

using $\log(pretest\ odds)$ as an offset term and not including an intercept term. Missing symptom data are entered as having log likelihood ratios of zero (making the assumption that a missing value is not in itself diagnostic). The parameter γ_i is called the shrinkage factor for symptom i . A value of γ_i of 1 indicates no dependency between symptoms in the model. A value of zero indicates that the symptom adds no additional diagnostic information as there is total dependency with the other symptoms. Adjusted LRs for each symptom when used in combination with the other symptoms in the model are calculated using the shrinkage factor and the crude LR for that symptom when used as on its own.

The formula for calculating adjusted LRs is given by

$$LR_i^{adj} = \exp(\gamma_i \times \log(LR_i^{crude})).$$

We calculated the area under the receiver operating characteristic curve (AUC-ROC) for each predictive model in the training dataset and test dataset (cohort C, Figure 1) and compared to the AUC-ROCs for predictive models containing only fever, cough and anosmia. Candidate symptoms whose addition yielded consistent improvements in AUC-ROC across multiple early time-points were considered 'early predictors' and were evaluated within simple diagnostic decision tools.

Worked example of a Spiegelhalter Knill-Jones models.

A Spiegelhalter Knill-Jones model using fever, cough, anosmia and headache as predictor variables at study day 0 is presented as a worked example to illustrate how the adjusted likelihood ratios (LRs) are calculated and interpreted. The full output of every model is available on request.

A 2x2 table is constructed for each symptom (raw counts from Table 1 in main text):

	infected	uninfected		
	fever			
symptom present	13	2	Total (TP + FP)	15
symptom absent	60	93	Total (TN + FN)	153
	cough			
symptom present	25	10	Total (TP + FP)	35
symptom absent	48	85	Total (TN + FN)	133
	Anosmia			
symptom present	15	1	Total (TP + FP)	16
symptom absent	58	94	Total (TN + FN)	152
	Headache			
symptom present	29	16	Total (TP + FP)	45
symptom absent	44	79	Total (TN + FN)	123
	Total (TP + FN)	Total (TN + FP)		
	73	95		

Crude positive and negative likelihood ratios (LRs) (supplementary figure 2) are calculated for each symptom i using the equations:

$$\text{positive } LR_i^{\text{crude}} = \frac{\left(\frac{TP}{TP+FN}\right)}{1-\left(\frac{TN}{TN+FP}\right)} \quad \text{negative } LR_i^{\text{crude}} = \frac{1-\left(\frac{TP}{TP+FN}\right)}{\left(\frac{TN}{TN+FP}\right)}$$

	Crude Positive LR	Crude Negative LR
Fever	8.46	0.84
Cough	3.25	0.73
Anosmia	19.52	0.80
Headache	2.36	0.72

The log of the odds of infection at study day 1 is used as the offset term:

$$\log\left(\frac{73}{95}\right) = \log(0.77)$$

Instead of 1 and 0, the indicator variables take the value of the crude positive and negative likelihood ratios for each of the CS with missing data entered as having log likelihood ratios of zero. The model fitted for the example is:

$$\log(\text{posttest odds}) = \log(0.77) + 0.547\log(LR_{\text{fever}}^{\text{crude}}) + 0.204\log(LR_{\text{cough}}^{\text{crude}}) + 0.785\log(LR_{\text{anosmia}}^{\text{crude}}) + 0.366\log(LR_{\text{headache}}^{\text{crude}})$$

The shrinkage factor γ_i is used to calculate the adjusted LRs (Figure 2) for symptom i . For example, the positive adjusted LR for fever when used in combination with cough, anosmia & sore throat on day 0 is calculated as follows:

$$\text{positive } LR_{\text{fever}}^{\text{adj}} = \exp(y_{\text{fever}} \times \log(\text{positive } LR_{\text{fever}}^{\text{crude}})) = \exp(0.547 \times \log(8.46)) = 3.22$$

Note how the magnitude of the LR has been shrunk to account for the symptom's dependency with other symptoms in the model.

Symptom	Crude Positive LR	Crude Negative LR	Shrinkage Factor	Adjusted Positive LR	Adjusted Negative LR
Fever	8.46	0.84	0.547	3.22	0.91
Cough	3.25	0.73	0.204	1.27	0.94
Anosmia	19.52	0.80	0.785	10.29	0.84
Headache	2.36	0.72	0.366	1.37	0.89

Estimates of overall LR for combinations of test results are obtained by multiplying these values. On study day 0, a contact with fever and headache but without cough or anosmia has $3.22 \times 0.94 \times 0.84 \times 1.37 = 3.90$ times the odds of being infected compared to a contact where symptoms are unknown.

We then used AUC-ROC to evaluate model discrimination in training and test datasets (cohorts B & C, Figure 1) and compare models using the CS to those with an additional symptom (Figure3).

SUPPLEMENTARY RESULTS:

S5: Recruitment cascade.

Early in the study, contacts were referred to our study through the Royal College of General Practitioners Research and Surveillance Centre. Later in the study, contacts were referred either directly through National Test And Trace or, in the case of most household-exposed contacts, through their index case, which PHE had identified. In total, 53011 referrals were received via these three pathways. If contacts were identified via PHE, PHE would initially ascertain whether participants consented to being contacted to take part in a research study. Consenting Contacts would then be contacted by Imperial, of which a subset would agree to take part. PHE and Imperial strove to contact all eligible contacts, but were limited by staff resources. The potentially eligible contacts PHE and Imperial attempted to call were non-biased in that the staff systematically worked down their list of participants, which was organised at random. Where referrals concerned index cases, multiple contacts may have been recruited following a single referral.

S6: Baseline Demographics for INSTINCT and ATACCC cohorts. Demographics are shown for cohorts B and C in Figure 1. Cohorts were compared using chi-squared tests for proportions and Wilcoxon rank sum to compare non-normally distributed populations.

BMI = Body Mass Index, IQR = Inter-quartile range

		INSTINCT	ATACCC	p-value
N		168	361	
Contact Type	Household (n/N, %)	168/168, 100%	227/361, 62.9%	X
	Non-household (n/N, %)	0/168, 0%	134/361, 37.1%	X
Age	(median, IQR)	34, 21 – 47	38, 28 – 52	<0.001
Sex	Female (n/N, %)	79/168, 47.0%	182/361, 50.4%	0.33
	Male (n/N, %)	88/168, 52.4%	166/361, 46.0%	
BMI	(median, IQR)	25.2, 22.8 – 27.4	25.2, 22.3 – 29.1	0.93
Ethnicity	White (n/N, %)	138/168, 82.1%	299/361, 82.8%	0.93
	Non-white (n/N, %)	25/168, 14.9%	51/361, 14.1%	
Smoking History	Current smoker (n/N, %)	20/168, 11.9%	57/361, 15.8%	0.67
	Vape/E-cigarettes (n/N, %)	4/168, 2.4%	8/361, 2.2%	
	Ex-smoker (n/N, %)	17/168, 10.1%	30/361, 8.3%	
	Non-smoker (n/N, %)	118/168, 70.2%	257/361, 71.2%	
Number of Comorbidities	None (n/N, %)	141/166, 84.9% (2 missing)	235/353, 66.6% (8 missing)	0.002
	One or more (n/N, %)	25/166, 15.1% (2 missing)	118/353, 33.4% (8 missing)	

S7: Cumulative COVID-19 related hazards for the onset of each symptom over time in days post index symptom onset. Cumulative COVID-19 related hazards $H(t)$ were calculated by adding up the daily hazards $h(t)$ up to time t . As $H(t)$ is conditional on not having experienced the symptom before t , for a population it may sum over 1. The probability of an 'infected' individual experiencing a symptom as a result of their infection by a particular time t , was estimated by the formula: $1 - S(t) = 1 - (\exp[-H(t)])$. Symptoms which had $1 - S(t)$ of more than 0.15 at $t = 10\text{days}$, corresponding to $H(10\text{days})$ of more than $-\ln(0.85) = 0.163$ were selected as candidate symptoms for evaluation. ISO = Index Symptom Onset.

S8: Life table for time-to-event analysis describing time from index any symptom onset to onset of individual symptoms in the contact. This analysis was performed using Cohort A (Figure 1). For each 2 day time interval, the number of participants at risk of onset of each symptom at the start of each time period is given. The numbers of participants reporting symptom onset, and participants whose study diary ended within the time interval are given in brackets. At enrolment, some contacts reported symptom onset prior to any symptom onset in the index. These events are shown in time periods -4 to -2 and -2 to 0 (days post index symptom onset). As prospective study diaries always started after index symptom onset, the number of participants whose symptom diary ended within these time intervals is not applicable (NA). Rarely, symptoms were reported at enrolment without an onset date. We imputed onset dates for these symptoms using the median number of days pre-enrolment (maximum 9 participants [5.4%] for muscle aches). URT = Upper Respiratory Tract, LRT = Lower Respiratory Tract, GI = Gastrointestinal

Infection Status	Group	Symptom	Time Period (Days post index symptom onset)							
			-4 to -2	-2 to 0	0 to 2	2 to 4	4 to 6	6 to 8	8 to 10	10 to 12
			Number at risk at start of time period (Symptom First reported, Last Day of Symptom Diary)							
Current Infection	Canonical Symptoms	fever		65 (2,NA)	63 (6,0)	57 (10,0)	47 (7,0)	40 (6,1)	33 (1,1)	31 (0,4)
		persistent cough		65 (1,NA)	64 (1,0)	63 (4,3)	56 (8,3)	45 (4,3)	38 (3,4)	31 (3,1)
		productive cough	65 (1,NA)	64 (2,NA)	62 (6,1)	55 (5,0)	50 (4,0)	46 (8,0)	38 (1,0)	37 (2,6)
		anosmia	65 (1,NA)	64 (1,NA)	63 (5,1)	57 (4,0)	53 (5,0)	48 (9,0)	39 (4,1)	34 (1,5)
	URT	rhinitis	65 (2,NA)	63 (2,NA)	61 (10,1)	50 (12,0)	38 (8,0)	30 (8,0)	22 (3,1)	18 (1,2)
		sore throat	65 (1,NA)	64 (0,NA)	64 (8,3)	53 (5,0)	48 (4,4)	40 (4,0)	36 (0,5)	31 (0,6)
		conjunctivitis			65 (0,1)	64 (2,0)	62 (4,2)	56 (1,2)	53 (2,3)	48 (1,9)
	LRT	dyspnoea	65 (2,NA)	63 (0,NA)	63 (1,1)	61 (2,0)	59 (7,3)	49 (2,1)	46 (1,2)	43 (1,8)
		wheeze	65 (1,NA)	64 (1,NA)	63 (0,1)	62 (3,0)	59 (3,0)	56 (2,2)	52 (1,4)	47 (0,9)
		chest pain	65 (1,1)	63 (0,NA)	63 (4,0)	59 (4,0)	55 (4,0)	51 (0,2)	49 (1,3)	45 (0,9)
		cough blood			65 (0,1)	64 (0,0)	64 (0,0)	64 (0,3)	61 (0,4)	57 (1,10)
	GI	diarrhoea	65 (1,NA)	64 (0,NA)	64 (2,4)	57 (1,1)	55 (6,0)	49 (0,1)	48 (2,3)	43 (2,8)
		appetite loss	65 (1,1)	63 (1,NA)	62 (2,4)	56 (4,3)	49 (6,1)	42 (4,2)	36 (0,2)	34 (0,9)
		abdominal pain	65 (1,NA)	64 (0,NA)	64 (4,1)	59 (5,0)	54 (3,0)	51 (0,1)	50 (1,4)	45 (1,7)
		nausea	65 (0,1)	64 (0,NA)	64 (2,2)	60 (5,0)	55 (1,3)	51 (2,2)	47 (3,2)	42 (2,8)
	Systemic	muscle aches	65 (2,NA)	63 (0,NA)	63 (4,7)	52 (6,7)	39 (4,4)	31 (3,0)	28 (1,1)	26 (1,6)
		headache	65 (1,NA)	64 (3,2)	59 (4,7)	48 (6,8)	34 (4,6)	24 (5,0)	19 (1,1)	17 (0,1)
		fatigue	65 (2,2)	61 (0,NA)	61 (2,11)	48 (3,11)	34 (2,5)	27 (2,2)	23 (1,2)	20 (1,4)
		confusion			65 (0,1)	64 (2,0)	62 (3,0)	59 (0,2)	57 (1,4)	52 (0,9)
		rash			65 (0,1)	64 (0,0)	64 (0,0)	64 (2,3)	59 (0,3)	56 (1,10)
Uninfected	Canonical Symptoms	fever			81 (1,2)	77 (1,2)	74 (0,14)	60 (1,5)	54 (0,8)	46 (1,10)
		persistent cough		81 (1,NA)	80 (2,3)	75 (2,2)	71 (1,13)	57 (1,5)	51 (0,9)	42 (0,10)
		productive cough			81 (1,2)	77 (3,2)	72 (1,14)	57 (0,5)	52 (0,9)	43 (0,11)
		anosmia			81 (0,3)	78 (1,2)	75 (0,14)	61 (0,5)	56 (1,9)	46 (0,11)
	URT	rhinitis			81 (4,3)	74 (1,2)	71 (5,14)	52 (1,5)	46 (0,9)	37 (0,9)
		sore throat		81 (1,NA)	80 (2,3)	75 (2,2)	71 (4,14)	53 (0,5)	48 (2,8)	38 (0,9)
		conjunctivitis			81 (0,3)	78 (0,2)	76 (1,15)	60 (0,5)	55 (0,8)	47 (0,11)
	LRT	dyspnoea			81 (0,4)	77 (2,2)	73 (2,15)	56 (2,5)	49 (0,6)	43 (0,10)
		wheeze			81 (0,3)	78 (1,2)	75 (0,14)	61 (0,5)	56 (1,9)	46 (0,12)
		chest pain			81 (2,3)	76 (1,2)	73 (2,14)	57 (0,6)	51 (0,8)	43 (0,9)
		cough blood			81 (0,3)	78 (0,2)	76 (0,14)	62 (0,5)	57 (0,9)	48 (0,12)
	GI	diarrhoea			81 (0,3)	78 (1,2)	75 (2,14)	59 (1,5)	53 (1,9)	43 (0,11)
		appetite loss			81 (1,3)	77 (1,2)	74 (1,14)	60 (1,6)	53 (0,8)	45 (1,11)
		abdominal pain			81 (1,3)	77 (1,2)	74 (1,14)	59 (0,5)	54 (0,8)	46 (0,11)
		nausea			81 (1,3)	77 (1,2)	74 (0,15)	59 (0,5)	54 (0,8)	46 (0,11)
	Systemic	muscle aches			81 (5,5)	71 (3,2)	65 (3,15)	51 (3,6)	45 (0,6)	38 (0,8)
		headache			81 (6,4)	71 (3,2)	66 (3,15)	48 (3,6)	39 (0,6)	33 (0,8)
		fatigue			81 (5,3)	73 (6,3)	64 (4,14)	46 (4,5)	37 (0,6)	31 (0,7)
		confusion			81 (0,3)	78 (2,2)	74 (0,14)	60 (2,5)	53 (0,8)	45 (1,11)
		rash			81 (1,3)	77 (0,2)	75 (0,14)	61 (0,5)	56 (1,8)	47 (0,11)

S9: Number of participants who had experienced a particular symptom by study day.

Symptoms were considered as a series of binary tests based on their occurrence by each study day. Contacts were defined as infected or uninfected based on a composite PCR and serology reference standard (see main text). These raw counts were used to calculate diagnostic sensitivity and specificity of individual symptoms (S10) as well as crude likelihood ratios (S11) which were used in the Spiegelhalter Knill-Jones analysis. TP = True-positive, FN = False-negative, TN = True-negative, FP = False-positive.

Infection Status			fever	cough	anosmia	muscle aches	headache	sore throat	rhinitis	breathless	appetite loss	abdominal pain	nausea	diarrhoea
	Day	TP + FN	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP
Infected	0	73	13	25	15	25	29	14	26	5	15	9	7	9
	1	72	20	31	15	34	39	20	37	11	20	12	10	12
	2	72	24	33	21	39	45	25	40	15	23	14	14	15
	3	71	28	34	26	40	47	27	43	18	26	14	14	15
	4	71	29	39	28	42	49	29	46	20	28	14	16	18
	5	69	28	38	28	41	48	30	48	20	28	14	16	20
	6	65	29	40	28	38	46	30	46	20	27	14	16	22
	7	57	25	33	25	35	39	26	40	16	23	12	15	18
Uninfected	Day	TN + FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP	FP
	0	95	2	10	1	14	16	5	7	6	4	5	4	3
	1	62	3	9	2	11	16	7	10	5	4	4	4	3
	2	62	4	10	2	13	18	10	11	5	5	4	4	3
	3	61	4	9	2	15	19	10	11	7	5	5	4	5
	4	61	4	10	2	15	20	11	11	8	5	5	4	6
	5	58	4	10	2	14	19	11	11	8	5	5	4	6
	6	57	3	10	2	13	18	11	11	7	4	4	3	6
	7	55	2	9	2	12	17	12	11	6	5	3	2	5

S10: Sensitivity and specificity, of individual symptoms by study day. Symptoms were considered as a series of binary tests based on their presence by each study day. Raw counts are given in S9. Diagnostic sensitivity [A] and specificity [B]) were calculated compared to a composite PCR and serology reference standard (described in the main text) in cohort B (Figure 1). The Clopper-Pearson interval was used to calculate 95% confidence intervals.[21]

S11: Positive and Negative Crude Likelihood Ratios for individual symptoms. Symptoms were considered as a series of binary tests based on their presence by each study day. Raw counts are given in S9. Crude LR were calculated compared to a composite PCR and serology reference standard (described in the main text) in cohort B (Figure 1). The Koopman method was used to calculate 95% confidence intervals.[1] LR = likelihood ratio.

S12: Adjusted Likelihood Ratios for the canonical symptoms when used in combination. LR

= likelihood ratio.

Study Day	Symptom	Adjusted Positive LR	Adjusted Negative LR	Study Day	Symptom	Adjusted Positive LR	Adjusted Negative LR
0	fever	3.55	0.9	4	fever	3.87	0.71
	cough	1.53	0.89		cough	1.33	0.86
	anosmia	10.84	0.84		anosmia	7.38	0.69
1	fever	3.53	0.82	5	fever	3.87	0.71
	cough	1.83	0.8		cough	1.31	0.87
	anosmia	3.36	0.88		anosmia	7.87	0.67
2	fever	3.18	0.79	6	fever	5.67	0.65
	cough	1.63	0.81		cough	1.36	0.83
	anosmia	4.97	0.8		anosmia	8.06	0.64
3	fever	3.68	0.73	7	fever	6.83	0.66
	cough	1.48	0.85		cough	1.31	0.86
	anosmia	6.53	0.72		anosmia	7.04	0.65

S13 (overleaf): Sensitivity, specificity and Number-Needed-to-Test by day since exposure

for broadened case definitions. Participants were excluded from this analysis if the date of exposure (index symptom onset for household contacts) was unknown. The 4 ‘early predictor’ symptoms (muscle aches, headache, sore throat, appetite loss) were combined individually with the canonical symptoms (CS) using an “OR” operator, for example *fever OR cough OR anosmia OR headache* (5A), as well as together in a list using an “at least” operator, for example *fever OR cough OR anosmia OR more than 2 of muscle aches, headache, sore throat, appetite loss*. Diagnostic performance was evaluated for each case definition against a serial PCR reference standard in the test dataset. The increase in TP and FP compared to the CS (*fever OR cough OR anosmia*) was measured. Number needed to test was calculated (number of false positives for every true positive plus 1). The Clopper-Pearson interval was used to calculate confidence intervals for proportions.[2] The data in this table are visualised in figure 4 in the main text. TP = true positive, FN = false negative, TN = true negative, FP = false positive, c.f. = compared to, Pos LR = positive likelihood ratio, Neg LR = Negative likelihood ratio, upper.ci = upper 95% confidence interval, lower.ci = lower 95% confidence interval.

S13A: “OR rule” decision tools

Decision Tool	Days post exposure	TP	FN	Increase in TP	Sensitivity	upper.ci	lower.ci	TN	FP	Increase in FP	Specificity	upper.ci	lower.ci	NNT
fever OR cough OR anosmia	0	17	74	NA	0.19	0.28	0.11	260	4	NA	0.98	1.00	0.96	1.24
	1	24	67	NA	0.26	0.37	0.18	256	8	NA	0.97	0.99	0.94	1.33
	2	32	59	NA	0.35	0.46	0.25	251	13	NA	0.95	0.97	0.92	1.41
	3	38	53	NA	0.42	0.53	0.32	243	21	NA	0.92	0.95	0.88	1.55
	4	42	49	NA	0.46	0.57	0.36	237	27	NA	0.90	0.93	0.85	1.64
	5	49	42	NA	0.54	0.64	0.43	234	30	NA	0.89	0.92	0.84	1.61
	6	55	36	NA	0.60	0.71	0.50	233	31	NA	0.88	0.92	0.84	1.56
	7	58	33	NA	0.64	0.74	0.53	231	33	NA	0.88	0.91	0.83	1.57
	8	62	29	NA	0.68	0.78	0.58	230	34	NA	0.87	0.91	0.82	1.55
	9	63	28	NA	0.69	0.78	0.59	230	34	NA	0.87	0.91	0.82	1.54
	10	64	27	NA	0.70	0.79	0.60	230	34	NA	0.87	0.91	0.82	1.53
fever OR cough OR anosmia OR muscle_aches	0	18	73	1	0.20	0.29	0.12	258	6	2	0.98	0.99	0.95	1.33
	1	27	64	3	0.30	0.40	0.21	253	11	3	0.96	0.98	0.93	1.41
	2	37	54	5	0.41	0.51	0.30	247	17	4	0.94	0.96	0.90	1.46
	3	44	47	6	0.48	0.59	0.38	237	27	6	0.90	0.93	0.85	1.61
	4	48	43	6	0.53	0.63	0.42	230	34	7	0.87	0.91	0.82	1.71
	5	54	37	5	0.59	0.70	0.49	226	38	8	0.86	0.90	0.81	1.70
	6	59	32	4	0.65	0.75	0.54	222	42	11	0.84	0.88	0.79	1.71
	7	61	30	3	0.67	0.77	0.56	220	44	11	0.83	0.88	0.78	1.72
	8	65	26	3	0.71	0.80	0.61	219	45	11	0.83	0.87	0.78	1.69
	9	66	25	3	0.73	0.81	0.62	219	45	11	0.83	0.87	0.78	1.68
	10	66	25	2	0.73	0.81	0.62	219	45	11	0.83	0.87	0.78	1.68
fever OR cough OR anosmia OR headache	0	22	69	5	0.24	0.34	0.16	254	10	6	0.96	0.98	0.93	1.45
	1	33	58	9	0.36	0.47	0.26	243	21	13	0.92	0.95	0.88	1.64
	2	42	49	10	0.46	0.57	0.36	229	35	22	0.87	0.91	0.82	1.83
	3	49	42	11	0.54	0.64	0.43	215	49	28	0.81	0.86	0.76	2.00
	4	57	34	15	0.63	0.73	0.52	205	59	32	0.78	0.83	0.72	2.04
	5	61	30	12	0.67	0.77	0.56	194	70	40	0.73	0.79	0.68	2.15
	6	66	25	11	0.73	0.81	0.62	189	75	44	0.72	0.77	0.66	2.14
	7	70	21	12	0.77	0.85	0.67	187	77	44	0.71	0.76	0.65	2.10
	8	73	18	11	0.80	0.88	0.71	185	79	45	0.70	0.76	0.64	2.08
	9	74	17	11	0.81	0.89	0.72	184	80	46	0.70	0.75	0.64	2.08
	10	74	17	10	0.81	0.89	0.72	184	80	46	0.70	0.75	0.64	2.08
fever OR cough OR anosmia OR sore_throat	0	20	71	3	0.22	0.32	0.14	257	7	3	0.97	0.99	0.95	1.35
	1	28	63	4	0.31	0.41	0.22	252	12	4	0.95	0.98	0.92	1.43
	2	37	54	5	0.41	0.51	0.30	246	18	5	0.93	0.96	0.89	1.49
	3	43	48	5	0.47	0.58	0.37	236	28	7	0.89	0.93	0.85	1.65
	4	50	41	8	0.55	0.65	0.44	229	35	8	0.87	0.91	0.82	1.70
	5	59	32	10	0.65	0.75	0.54	219	45	15	0.83	0.87	0.78	1.76
	6	64	27	9	0.70	0.79	0.60	216	48	17	0.82	0.86	0.77	1.75
	7	69	22	11	0.76	0.84	0.66	212	52	19	0.80	0.85	0.75	1.75
	8	71	20	9	0.78	0.86	0.68	210	54	20	0.80	0.84	0.74	1.76
	9	71	20	8	0.78	0.86	0.68	208	56	22	0.79	0.84	0.73	1.79
	10	71	20	7	0.78	0.86	0.68	208	56	22	0.79	0.84	0.73	1.79
fever OR cough OR anosmia OR appetite_loss	0	18	73	1	0.20	0.29	0.12	259	5	1	0.98	0.99	0.96	1.28
	1	26	65	2	0.29	0.39	0.20	254	10	2	0.96	0.98	0.93	1.38
	2	34	57	2	0.37	0.48	0.27	249	15	2	0.94	0.97	0.91	1.44
	3	41	50	3	0.45	0.56	0.35	240	24	3	0.91	0.94	0.87	1.59
	4	45	46	3	0.49	0.60	0.39	232	32	5	0.88	0.92	0.83	1.71
	5	51	40	2	0.56	0.66	0.45	227	37	7	0.86	0.90	0.81	1.73
	6	57	34	2	0.63	0.73	0.52	225	39	8	0.85	0.89	0.80	1.68
	7	60	31	2	0.66	0.76	0.55	223	41	8	0.84	0.89	0.80	1.68
	8	65	26	3	0.71	0.80	0.61	222	42	8	0.84	0.88	0.79	1.65
	9	66	25	3	0.73	0.81	0.62	222	42	8	0.84	0.88	0.79	1.64
	10	66	25	2	0.73	0.81	0.62	222	42	8	0.84	0.88	0.79	1.64

S13B: “At least” decision tools

Decision Tool	Days post exposure	TP	FN	Increase in TP	Sensitivity	upper.ci	lower.ci	TN	FP	Increase in FP	Specificity	upper.ci	lower.ci	NNT
fever OR cough OR anosmia	0	17	74	NA	0.19	0.28	0.11	260	4	NA	0.98	1.00	0.96	1.24
	1	24	67	NA	0.26	0.37	0.18	256	8	NA	0.97	0.99	0.94	1.33
	2	32	59	NA	0.35	0.46	0.25	251	13	NA	0.95	0.97	0.92	1.41
	3	38	53	NA	0.42	0.53	0.32	243	21	NA	0.92	0.95	0.88	1.55
	4	42	49	NA	0.46	0.57	0.36	237	27	NA	0.90	0.93	0.85	1.64
	5	49	42	NA	0.54	0.64	0.43	234	30	NA	0.89	0.92	0.84	1.61
	6	55	36	NA	0.60	0.71	0.50	233	31	NA	0.88	0.92	0.84	1.56
	7	58	33	NA	0.64	0.74	0.53	231	33	NA	0.88	0.91	0.83	1.57
	8	62	29	NA	0.68	0.78	0.58	230	34	NA	0.87	0.91	0.82	1.55
	9	63	28	NA	0.69	0.78	0.59	230	34	NA	0.87	0.91	0.82	1.54
	10	64	27	NA	0.70	0.79	0.60	230	34	NA	0.87	0.91	0.82	1.53
fever OR cough OR anosmia OR at least 1 of muscle_aches, headache, sore throat and appetite loss	0	22	69	5	0.24	0.34	0.16	251	13	9	0.95	0.97	0.92	1.59
	1	34	57	10	0.37	0.48	0.27	240	24	16	0.91	0.94	0.87	1.71
	2	45	46	13	0.49	0.60	0.39	225	39	26	0.85	0.89	0.80	1.87
	3	51	40	13	0.56	0.66	0.45	209	55	34	0.79	0.84	0.74	2.08
	4	60	31	18	0.66	0.76	0.55	196	68	41	0.74	0.79	0.69	2.13
	5	66	25	17	0.73	0.81	0.62	182	82	52	0.69	0.74	0.63	2.24
	6	70	21	15	0.77	0.85	0.67	176	88	57	0.67	0.72	0.61	2.26
	7	74	17	16	0.81	0.89	0.72	173	91	58	0.66	0.71	0.59	2.23
	8	76	15	14	0.84	0.90	0.74	171	93	59	0.65	0.71	0.59	2.22
	9	76	15	13	0.84	0.90	0.74	169	95	61	0.64	0.70	0.58	2.25
	10	76	15	12	0.84	0.90	0.74	169	95	61	0.64	0.70	0.58	2.25
fever OR cough OR anosmia OR at least 2 of muscle_aches, headache, sore throat and appetite loss	0	21	70	4	0.23	0.33	0.15	258	6	2	0.98	0.99	0.95	1.29
	1	30	61	6	0.33	0.44	0.23	251	13	5	0.95	0.97	0.92	1.43
	2	39	52	7	0.43	0.54	0.33	245	19	6	0.93	0.96	0.89	1.49
	3	47	44	9	0.52	0.62	0.41	234	30	9	0.89	0.92	0.84	1.64
	4	52	39	10	0.57	0.67	0.46	228	36	9	0.86	0.90	0.82	1.69
	5	58	33	9	0.64	0.74	0.53	219	45	15	0.83	0.87	0.78	1.78
	6	63	28	8	0.69	0.78	0.59	213	51	20	0.81	0.85	0.75	1.81
	7	68	23	10	0.75	0.83	0.65	210	54	21	0.80	0.84	0.74	1.79
	8	70	21	8	0.77	0.85	0.67	209	55	21	0.79	0.84	0.74	1.79
	9	71	20	8	0.78	0.86	0.68	209	55	21	0.79	0.84	0.74	1.77
	10	71	20	7	0.78	0.86	0.68	209	55	21	0.79	0.84	0.74	1.77
fever OR cough OR anosmia OR at least 3 of muscle_aches, headache, sore throat and appetite loss	0	18	73	1	0.20	0.29	0.12	259	5	1	0.98	0.99	0.96	1.28
	1	26	65	2	0.29	0.39	0.20	255	9	1	0.97	0.98	0.94	1.35
	2	34	57	2	0.37	0.48	0.27	250	14	1	0.95	0.97	0.91	1.41
	3	41	50	3	0.45	0.56	0.35	242	22	1	0.92	0.95	0.88	1.54
	4	46	45	4	0.51	0.61	0.40	235	29	2	0.89	0.93	0.85	1.63
	5	52	39	3	0.57	0.67	0.46	231	33	3	0.88	0.91	0.83	1.63
	6	58	33	3	0.64	0.74	0.53	230	34	3	0.87	0.91	0.82	1.59
	7	60	31	2	0.66	0.76	0.55	228	36	3	0.86	0.90	0.82	1.60
	8	65	26	3	0.71	0.80	0.61	226	38	4	0.86	0.90	0.81	1.58
	9	66	25	3	0.73	0.81	0.62	225	39	5	0.85	0.89	0.80	1.59
	10	66	25	2	0.73	0.81	0.62	225	39	5	0.85	0.89	0.80	1.59
fever OR cough OR anosmia OR at least 4 of muscle_aches, headache, sore throat and appetite loss	0	17	74	0	0.19	0.28	0.11	260	4	0	0.98	1.00	0.96	1.24
	1	24	67	0	0.26	0.37	0.18	256	8	0	0.97	0.99	0.94	1.33
	2	32	59	0	0.35	0.46	0.25	251	13	0	0.95	0.97	0.92	1.41
	3	38	53	0	0.42	0.53	0.32	243	21	0	0.92	0.95	0.88	1.55
	4	42	49	0	0.46	0.57	0.36	237	27	0	0.90	0.93	0.85	1.64
	5	49	42	0	0.54	0.64	0.43	234	30	0	0.89	0.92	0.84	1.61
	6	55	36	0	0.60	0.71	0.50	233	31	0	0.88	0.92	0.84	1.56
	7	58	33	0	0.64	0.74	0.53	231	33	0	0.88	0.91	0.83	1.57
	8	63	28	1	0.69	0.78	0.59	230	34	0	0.87	0.91	0.82	1.54
	9	64	27	1	0.70	0.79	0.60	230	34	0	0.87	0.91	0.82	1.53
	10	64	27	0	0.70	0.79	0.60	230	34	0	0.87	0.91	0.82	1.53

S14 (overleaf): Time-to-event analysis describing time in days from exposure until participants were positively identified by a case definition. Easily comprehensible case definitions containing the canonical symptoms (CS): fever, cough and anosmia, and the early-predictors (EP): headache, sore throat, muscle aches and appetite loss were constructed as described in Box 1 (main text). The number of participants to whom the case definition was applied at each time-point is given ('at risk'). The numbers of participants identified positively by the case definition ('identified'), and the number of participants whose study diary ended ('lost') are given for each time-point. The median, 25th and 75th percentile of the time till positive identification by each case definition in days post exposure is given for PCR-positive and PCR-negative participants. Where participants were identified at enrolment, this value is 0 representing study day 0. If 25% of participants were not identified by the case definition, the value is "Never". Rarely, symptoms were reported at enrolment without an onset date. We imputed onset dates for these symptoms using the median number of days pre-enrolment (maximum 2 participants [0.55%] for rhinitis). TP = true positive, FN = false negative, TN = true negative, FP = false positive.

A - PCR-positive		Time Period (Days since exposure)											Time to Detection	
		0	1	2	3	4	5	6	7	8	9	10	%	Days
≥1 of the CS...	At Risk (TP + FN)	91	74	67	59	52	44	32	24	21	17	15	25	2
	Identified (TP)	17	7	8	6	4	7	6	3	4	1	1	50	6
	Lost	0	0	0	1	4	5	2	0	0	1	0	70	10
...OR sore throat	At Risk (TP + FN)	91	71	63	54	47	37	23	16	11	9	8	25	2
	Identified (TP)	20	8	9	6	7	9	5	5	2	0	0	50	5
	Lost	0	0	0	1	3	5	2	0	0	1	0	75	8
... OR headache	At Risk (TP + FN)	91	69	58	49	41	30	22	15	11	8	6	25	2
	Identified (TP)	22	11	9	7	8	4	5	4	3	1	0	50	4
	Lost	0	0	0	1	3	4	2	0	0	1	0	75	7
...OR muscle aches	At Risk (TP + FN)	91	73	64	54	46	38	27	20	18	14	13	25	2
	Identified (TP)	18	9	10	7	4	6	5	2	4	1	0	50	5
	Lost	0	0	0	1	4	5	2	0	0	0	0	75	9
...OR appetite loss	At Risk (TP + FN)	91	73	65	57	49	41	30	22	19	14	12	25	2
	Identified (TP)	18	8	8	7	4	6	6	3	5	1	0	50	6
	Lost	0	0	0	1	4	5	2	0	0	1	0	75	9
...or ≥1 of the EP	At Risk (TP + FN)	91	69	57	46	39	27	17	11	7	5	5	25	2
	Identified (TP)	22	12	11	6	9	6	4	4	2	0	0	50	4
	Lost	0	0	0	1	3	4	2	0	0	0	0	75	7
...or ≥2 of the EP	At Risk (TP + FN)	91	70	61	52	43	35	24	17	12	10	8	25	2
	Identified (TP)	21	9	9	8	5	6	5	5	2	1	0	50	4
	Lost	0	0	0	1	3	5	2	0	0	1	0	75	8
...or ≥3 of the EP	At Risk (TP + FN)	91	73	65	57	49	40	29	21	19	14	12	25	2
	Identified (TP)	18	8	8	7	5	6	6	2	5	1	0	50	5
	Lost	0	0	0	1	4	5	2	0	0	1	0	75	9
...or ≥4 of the EP	At Risk (TP + FN)	91	74	67	59	52	44	32	24	21	16	14	25	2
	Identified (TP)	17	7	8	6	4	7	6	3	5	1	0	50	6
	Lost	0	0	0	1	4	5	2	0	0	1	0	75	9

B - PCR-negative		Time Period (Days since exposure)											Time to Detection	
		0	1	2	3	4	5	6	7	8	9	10	%	Days
≥1 of the CS...	At Risk (TN + FP)	264	260	256	251	236	201	165	145	137	134	134	25	never
	Identified (FP)	4	4	5	8	6	3	1	2	1	0	0	50	never
	Lost	0	0	0	7	29	33	19	6	2	0	2	70	never
...OR sore throat	At Risk (TN + FP)	264	257	252	246	229	194	152	131	121	117	115	25	never
	Identified (FP)	7	5	6	10	7	10	3	4	2	2	0	50	never
	Lost	0	0	0	7	28	32	18	6	2	0	2	75	never
... OR headache	At Risk (TN + FP)	264	254	243	229	209	174	132	114	107	103	102	25	6
	Identified (FP)	10	11	14	14	10	11	5	2	2	1	0	50	never
	Lost	0	0	0	6	25	31	13	5	2	0	2	75	never
...OR muscle aches	At Risk (TN + FP)	264	258	253	247	230	195	158	138	131	128	128	25	never
	Identified (FP)	6	5	6	10	7	4	4	2	1	0	0	50	never
	Lost	0	0	0	7	28	33	16	5	2	0	2	75	never
...OR appetite loss	At Risk (TN + FP)	264	259	254	249	233	196	160	140	133	130	130	25	never
	Identified (FP)	5	5	5	9	8	5	2	2	1	0	0	50	never
	Lost	0	0	0	7	29	31	18	5	2	0	2	75	never
...or ≥1 of the EP	At Risk (TN + FP)	264	251	240	225	203	167	124	105	98	94	92	25	5
	Identified (FP)	13	11	15	16	13	14	6	3	2	2	0	50	never
	Lost	0	0	0	6	23	29	13	4	2	0	2	75	never
...or ≥2 of the EP	At Risk (TN + FP)	264	258	251	245	227	192	151	130	122	119	119	25	never
	Identified (FP)	6	7	6	11	6	9	6	3	1	0	0	50	never
	Lost	0	0	0	7	29	32	15	5	2	0	2	75	never
...or ≥3 of the EP	At Risk (TN + FP)	264	259	255	250	235	199	162	143	135	131	130	25	never
	Identified (FP)	5	4	5	8	7	4	1	2	2	1	0	50	never
	Lost	0	0	0	7	29	33	18	6	2	0	2	75	never
...or ≥4 of the EP	At Risk (TN + FP)	264	260	256	251	236	201	165	145	137	134	134	25	never
	Identified (FP)	4	4	5	8	6	3	1	2	1	0	0	50	never
	Lost	0	0	0	7	29	33	19	6	2	0	2	75	never

S16 (overleaf): Comparison of proposed symptom criteria with international case

definitions. The presence or absence of symptoms by each day post-exposure was used to calculate diagnostic sensitivity (A) and specificity (B) for different case definitions (cohort D, figure 1) against a serial PCR reference standard. We compared the canonical symptoms (CS), our proposed symptom criteria (≥ 1 of the CS, or ≥ 2 of the EP) and international case definitions (S1, [3–7]). Where the precise wording of items in our symptom diaries (S2) did not correspond exactly with wording in international case definitions the closest approximation was chosen. In the WHO case definition, anorexia/nausea/vomiting is considered one symptom item whereas in our analysis loss of appetite and nausea/vomiting were considered two separate items, possibly overestimating sensitivity and underestimating specificity. In the US Centers for Disease Control case definition, chills and rigors are included as two separate symptom items as well as fever. In our analysis we considered fever as one variable, possibly underestimating sensitivity and overestimating specificity. The Clopper-Pearson interval was used to calculate 95% confidence intervals.[2]

Kaplan-Meier plots were constructed showing time from exposure (index symptom onset for household-contacts) until positive identification by these symptom criteria in PCR-positive (C) and PCR-negative (D) contacts. The plot for the CS is shown in black. Rarely, symptoms were reported at enrolment without an onset date. We imputed onset dates for these symptoms by assuming the median number of days pre-enrolment (maximum 2 participants [0.55%] for rhinitis).

Legend:

- fever OR cough OR anosmia (CS)
- ▼ CS OR at least 2 of the Early Predictors
- × WHO
- × USA (CDC)
- × France
- × Spain

A

B

C

D

S17: Leave-one-out analysis

Kaplan-Meier plots showing time from exposure until positive identification by different symptom criteria. The proportion of PCR-positive (left) and PCR-negative (right) participants who were positively identified by each case definition by each day following exposure (index symptom onset for household-contacts) is shown using a Kaplan-Meier plot. We compared the performance of the canonical symptoms (fever OR cough OR anosmia; black line) with our proposed symptom criteria (fever OR cough OR anosmia OR AT LEAST 2 OF headache, muscle aches, sore throat, appetite loss; purple line). These data are shown in figure 5 (main text) and life-tables presented in Supplementary Table 7.

We performed a leave-one-out analysis to determine whether each of the ‘early predictor’ symptoms (headache, muscle aches, sore throat, appetite loss) were necessary within the proposed symptom criteria. Four new symptom criteria (each with one of the EPs left out) are applied (red, green, blue and brown lines) and compared to the proposed symptom criteria. HA=Headache, MA=muscle aches, ST= sore throat, AL=appetite loss.

SUPPLEMENTARY REFERENCES:

1. Koopman PAR. Confidence Intervals for the Ratio of Two Binomial Proportions. *Biometrics* [Internet] 1984; 40: 513–517 Available from: <https://www.jstor.org/stable/2531405>.
2. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* [Internet] 1934; 26: 404–413 Available from: <https://academic.oup.com/biomet/article/26/4/404/291538>.
3. UK Government. Get a free PCR test to check if you have coronavirus [cited 13 July, 2021], Available: <https://www.gov.uk/get-coronavirus-test>. .
4. World Health Organisation. WHO COVID-19 Case definition. [cited July 4, 2021]. Available: https://www.who.int/publications-detail-redirect/WHO-2019-nCoV-Surveillance_Case_Definition-2020.2. .
5. Centers for Disease Control and Prevention. Coronavirus disease 2019 (covid-19) 2020 interim case definition. 2020. [cited 5 July 2021]. Available: <https://wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-covid-19/case-definition/2020/08/05/>. .
6. Gouvernement de la République française. Tester- Alerter- Protéger (2021) [cited 15th July 2021] Available: <https://www.gouvernement.fr/info-coronavirus/tests-et-depistage>. . .
7. Ministerio de Sanidad and Instituto de Salud Carlos III. Estrategia de detección precoz, vigilancia y control de covid-19. (2021) [cited 14th July 2021] Available: https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/documentos/COVID19_Estrategia_vigilancia_y_control_e_indicadores.pdf. .