medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Symptom profiles and accuracy of clinical definitions for COVID-19 in the community. Results of the Virus Watch community cohort.

Ellen Fragaszy[1, 2], Madhumita Shrotri[1], Cyril Geismar[1,3], Anna Aryee[4], Sarah Beale[1,3], Isobel Braithwaite[1], Thomas Byrne[1], Wing Lam Erica Fong[1], Jo Gibbs[5], Pia Hardelid[6], Jana Kovar[3], Vasileios Lampos[7], Eleni Nastouli[8,9], Annalan M D Navaratnam[1,3], Vincent Nguyen[1,3], Parth Patel[1], Robert W Aldridge[1], Andrew Hayward[3] on behalf of Virus Watch Collaborative*

* *Linda Wijlaars[6], Moira Spyer[6,9], Ben Killingley[10,11], Ingemar Cox[7], Rachel A McKendry[12], Tao Cheng, Yunzhe Liu[13], Anne M Johnson[5], Richard Gilson[5]*

1. Centre for Public Health Data Science, Institute of Health Informatics, University College London, UK.
2. Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London, UK.
3. Institute of Epidemiology and Health Care, University College London, London, UK.
4. Institute of Health Informatics, University College London, 222 Euston Rd, London NW1 2DA, UK.
5. Institute for Global Health, University College London, London, UK.
6. Department of Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, London, UK
7. Department of Computer Science, University College London, London, UK.
8. Department of Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, London, UK.
9. Francis Crick Institute, London, UK.
10. Health Protection and Influenza Research Group, Division of Epidemiology and Public Health, University of Nottingham School of Medicine, Nottingham, United Kingdom
11. University College London Hospital, London, United Kingdom.
12. London Centre for Nanotechnology and Division of Medicine, London, UCL
13. SpaceTimeLab, Department of Civil, Environmental and Geomatic Engineering, University College London, London, UK.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Abstract

Background: Understanding the symptomatology and accuracy of clinical case definitions for COVID-19 in the community is important for the initiation of Test, Trace and Isolate (TTI) and may, in future, be important for early prescription of antivirals.

Methods: Virus Watch is a large community cohort with prospective daily recording of a wide range of symptoms and self-reporting of swab results (mainly undertaken through the UK TTI System). We compared frequency, severity, timing, and duration of symptoms in test positive and test negative cases. We compared the test performance of the current UK case definition used by TTI (any one of: new continuous cough, high temperature or loss of or change in sense of smell or taste) with a wider definition that also included muscle aches or chills or headache or loss of appetite.

Findings: We included results from 8213 swabbed illnesses, 944 of which tested positive for COVID-19. All symptoms were more common in swab positive than swab negative illnesses and symptoms were also more severe and of longer duration. Common symptoms such as cough, headache, fatigue, muscle aches and loss of appetite occurred early in the course of illness but were also very common in test-negative illnesses. Rarer symptoms such as fever or loss or altered sense of smell or taste were often not present but were markedly more common in swab positive compared to swab negative cases. The current UK definition had a sensitivity and specificity of 81% and 47% respectively for symptomatic COVID-19 compared to 93% and 26% for the broader definition. On average cases met the broader case definition one day earlier than current definition. 1.7-fold more illnesses met the broader definition than the current case definition.

Interpretation: COVID-19 is difficult to distinguish from other respiratory infections and common ailments on the basis of symptoms. Broadening the list of symptoms used to encourage engagement with TTI could moderately increase the number of infections identified and shorten delays but with a large increase in the number of tests needed and in the number of people and contacts who do not have COVID-19 but might need to self-isolate whilst awaiting results.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Introduction

The natural history of COVID-19 infection can range from asymptomatic infection in around 25% of infections[1] to severe or fatal disease at a rate that is highly age dependent[2]. Understanding the natural history of symptomatic COVID-19 in the community is critical to the control of infection because it informs decisions about who should seek testing, whether those with symptoms should self-isolate and whether those in contact with symptomatic people should self-isolate. Understanding the normal course of symptoms is also potentially helpful to patients and clinicians assessing whether care needs to be escalated due to unexpectedly severe or prolonged symptoms. In future, symptom profiles may also trigger early use of antivirals to prevent deterioration and potentially to minimise transmission[3]. Finally understanding symptom profiles is important to inform syndromic surveillance.

A wide range of clinical case definitions for COVID-19 are available utilising different combinations of symptoms to alert individuals to the need for testing, isolation and contact tracing. For example, WHO include the following symptoms in the clinical case definition of a suspected case - Acute onset of fever AND cough; OR Acute onset of ANY THREE OR MORE of the following signs or symptoms: Fever, cough, general weakness/fatigue, headache, myalgia, sore throat, coryza, dyspnoea, anorexia/nausea/vomiting, diarrhoea, altered mental status[4]. European Centers for Disease Control define a possible case based on - at least one of Cough, Fever, Shortness of breath or sudden loss of sudden onset of anosmia, ageusia or dysgeusia[4]; and US Centers for Disease Control use the following clinical criteria - at least two of fever (measured or subjective), chills, rigors, myalgia, headache, sore throat, new olfactory and taste disorder(s) OR at least one of fever (measured or subjective), chills, rigors, myalgia, headache, sore throat, new olfactory and taste disorder(s)[5].

In the UK, the Test Trace and Isolate Community testing programme (TTI) asks individuals to seek testing if they have any of the following symptoms - a new continuous cough, a high temperature, loss of or altered sense of smell or taste[6].

Although it is clear that addition of further symptoms could increase sensitivity, this is at the cost of a loss of specificity and increasing numbers of people who require testing, isolation and contact tracing[7].

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

For example, recent analysis of data from the REACT community survey suggests that adding in loss of appetite, chills, headache or muscle aches could increase the proportion of cases identified from 53% to 75% but at the cost of a 2.7 fold increase in required testing capacity[8]. Altering clinical case definitions will have different implications at different disease prevalences, since positive and negative predictive values of tests depend on disease prevalence as well as sensitivity and specificity. The timing within the course of an illness at which infected people meet the case definition is important since early isolation of cases and contacts reduces transmission.

Prospective community studies where participants record symptoms in near real time are needed to accurately measure COVID-19 symptom profiles with minimal recall bias. Here we describe prospectively recorded symptom profiles (frequency, severity and duration) of illnesses that tested positive for COVID-19 and illnesses that tested negative within a large community cohort study (Virus Watch). We compare the test characteristics (Sensitivity, Specificity, Positive and Negative Predictive Value, timing of meeting the case definition, number of tests needed to identify one case) for the current UK definition and a wider definition proposed following analysis of the REACT study, which adds headache or chills or loss of appetite or muscle aches to the existing UK definition. We explore how symptom profiles and case definition performance vary by age and stage of the pandemic. We will also publish the dataset from which these analyses were conducted allowing readers to explore the implications of different symptom combinations on case definition performance.

Methods

Study design and data collection

The Virus Watch study is an online, prospective, community cohort study following up entire households in England and Wales during the COVID-19 pandemic. As of 04 May 2021, 24,296 households and 50,699 people across England and Wales have joined the study. The full study protocol is published elsewhere[9].

In brief, after registering with the study, Virus Watch participants completed baseline surveys and were then asked to prospectively complete detailed daily symptom diaries recording the

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

presence and severity of any symptoms of acute respiratory and gastrointestinal infections during periods of any illness occurring during follow up. At the end of each week participants were emailed links to a weekly survey where they reported any symptoms from that previous week as well as the dates and outcomes of any COVID-19 swabbing conducted outside of the Virus Watch study (mainly as part of TTI). Within the main cohort there is a nested laboratory sub-cohort of 10,766 participants who have additionally provided study-specific swab specimens tested using COVID-19 PCR from the end of December 2020 onwards.

Symptom data gathered through the weekly survey were grouped into illness episodes and matched to swab results (see appendix for further details). The start date of an illness episode was defined as the first day any symptoms were reported, and the end date was the final day of reported symptoms. A 7-day washout period where no symptoms were reported was used to define the end of one illness episode and the start of a new illness episode. The data presented in this analysis includes illnesses which began between the start of the study through 02 May 2021. Within illness episodes, we investigated a wide range of individual symptoms (see appendix for further definition of symptoms collected) and the following symptom groupings: **UK Case definition** – one or more of the following: cough, measured fever or feeling feverish, loss of, or change to, sense of smell or taste. **Broader case definition** - one or more of the following: cough, measured fever or feeling feverish, loss of, or change to, sense of smell or taste, headache, muscle aches, loss of appetite or chills.

We present simple descriptive analyses of the frequency, severity and duration of symptoms in COVID-19 positive and negative illnesses. We calculate sensitivity and specificity for individual symptoms and, for the two case definitions we also calculate the positive predictive value (PPV), negative predictive value (NPV) number of people meeting the definition within the cohort and the numbers needed to test to identify one positive case (NNT).

We will provide a dataset with age group classified into broad categories and timing of illnesses and tests removed to preserve anonymity. This can be used to assess sensitivity and specificity of different symptom combinations.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Results

Overall, there were 29,083 illnesses (with at least one of the symptoms defined in supplementary Table S1) reported in the cohort. 8,213 illnesses were swabbed and of these 944 (11.5%) tested positive for COVID-19; 436 of these swabs were conducted as part of the study, including 22 positives. The percentage of swabbed illnesses testing positive were highest in young adults aged 16-24 and lowest in children aged 0-15, highest in London and lowest in the South East and South West regions and peaked in December 2020 (Table 1). Supplementary appendix Table S2 shows the proportion of illnesses with swabs for each symptom - those with the symptoms recommended for swabbing by TTI are more likely to be swabbed than those with other symptoms.

Figure 1a shows the proportion of swabbed illnesses that reported each symptom according to whether they tested positive or negative for COVID-19. All of the wide range of reported symptoms were more common in illnesses that tested positive for COVID-19 than in other illnesses. Amongst COVID-19 positive illnesses the 10 most commonly reported symptoms in decreasing order of frequency were: fatigue, headache, cough, muscle ache, loss or change to sense of smell or taste, needing to spend extra time in bed, sore throat, difficulties in undertaking daily activities, feeling feverish and sneezing. The percentage showing each symptom by day of illness is shown in figure 1b and c illustrating both the higher frequency and longer duration of key symptoms in COVID-19 positive and negative cases.

Table S2 in the supplementary appendix shows the sensitivity and specificity of each symptom and the mean and median day of illness on which they are first reported. It can be seen that although constitutional symptoms such as headache, fatigue and muscle aches are common and occur early in the course of illness they are also a common feature of non-COVID illnesses and have low specificity.

Figure 2 and Table S3 (in the supplementary appendix) shows the maximum reported severity for a range of key symptoms in COVID-19 positive and negative illnesses. It can be seen that when symptoms do occur, they are more likely to be severe in COVID-19 illnesses than in test negative illnesses. Figure 3 shows the distribution of the duration of illnesses in COVID-19 positive and negative illnesses. It can be seen that the duration of illness is longer in COVID-19 illnesses than in other illnesses. Table S4 shows how illness duration varies by age and gender

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

in COVID-19 positive and negative illnesses and Figure S1 shows the distribution of duration of illnesses among COVID-19 illnesses by age group. Illnesses tend to be of longer duration in older cases.

Table 2 shows the mean and median day of meeting the current and broader case definition, Sensitivity, Specificity, Positive predictive value (PPV), Negative predictive value (NPV) and numbers needed to test to identify a case (NNT). The numbers meeting the case definition and how many fold higher this is for the broader case definition are shown (multiplication factor). These results are stratified by age and calendar time.

Sensitivity of the current case definition (i.e. the proportion of all those illnesses testing COVID-19 positive who met the definition) was 81% compared to 93% for the broader case definition. Specificity (i.e. the proportion of all those illnesses testing COVID-19 negative who did not meet the case definition) was 47% for the current case definition and was 26% for the broader case definition. Sensitivity and specificity of both case definitions was lower in children age 0-15 than in older age groups. Sensitivity of case definitions remained stable over time. Specificity of the current UK definition appeared lower in August/September (coinciding with a large national outbreak of Rhinovirus in children). The PPV (the proportion of those meeting the clinical case definition who test positive for COVID) was 17% for the current UK case definition and 15% for the broader definition. PPV was substantially lower during August/September when disease rates were low (8% for current UK definition and 7% for the broader definition). NPV (the proportion of those with illnesses not meeting the clinical case definition who test negative for COVID-19) was 95% for the current UK case definition and 96% for the expanded case definition. The number of illnesses meeting the broader case definition was 1.7 fold higher than those meeting the current definition.

Discussion

We characterised the symptom profiles and estimated the accuracy of clinical case definitions for COVID-19 among community cases arising in a large, prospective population-based cohort study based in the UK. All symptoms asked about were more frequently reported and when present were generally more severe and longer lasting in COVID-19 positive illnesses compared to COVID-19 negative illnesses. Individually, cough and some constitutional

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

symptoms including headache, muscle ache and fatigue presented early in illness and had moderate sensitivity and specificity as they were common in both COVID-19 positive and COVID-19 negative illnesses. In contrast, fever and loss or change to smell or taste presented slightly later in illness and had a lower sensitivity but higher specificity as they were not as common in COVID-19 positive illnesses but were even less common in COVID-19 negative illnesses. The combination of symptoms in the current UK TTI case definition had a higher sensitivity (81%) than any individual symptom, with a specificity of 47%. Adding additional symptoms to the case definition can lead to earlier case identification and higher sensitivity but at the cost of specificity and consequently, a substantial increase in the number of illnesses eligible for testing. For example, when we compare the broader case definition to the current UK case definition, cases on average met the case definition one day earlier and there was a moderate increase in sensitivity to 93% but at a much lower specificity of 26%. It would also lead to 1.7 times more illnesses eligible for testing.

Strengths of this work include the prospective daily recording of a wide range of symptoms across a large community cohort and linkage of these to self reported swab results ascertained on a weekly basis. This should maximise the accuracy of symptom data amongst swabbed participants. Our sample is not fully representative of the population of England and Wales but includes participants in every local authority area. There is a moderate overrepresentation of those aged over 65 and an underrepresentation of those in more deprived areas[10]. Whilst we collected information on a very wide range of symptoms testing primarily relied on that conducted through the national TTI programme meaning that those meeting the current case definition are more likely to be tested. This is likely to lead to an overestimation of the sensitivity of the current UK case definition. A further strength is that we will publish the dataset online to enable replication of analyses and for others to explore the test characteristics of various combinations of symptoms for case definitions.

COVID-19 is difficult to distinguish from other respiratory infections or common ailments on the basis of symptoms alone. Also, a high proportion of infections are asymptomatic or have a pre-symptomatic phase when transmission can occur. As such, systems to identify symptomatic cases, test them and isolate them and their contacts can only ever reduce, rather than prevent all transmission. These programmes, when part of a broader programme of Non-Pharmaceutical Interventions, can however contribute to control of infection and may be particularly effective when introduced at very low levels of infection. For example, countries that combined early and strict border controls, intensive testing and rapid introduction of lockdowns

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

have had substantially lower COVID-19 transmission and mortality than other countries such as the UK[11].

The success of testing and isolation programmes is dependent on public understanding and engagement. Data from behavioural surveys in England show only 51% of participants knew the symptoms that testing is recommended for, only 18% sought testing if they had the symptoms and only 42.5% fully adhered to self-isolation. Engagement was lower in younger people and amongst those in financial hardship[12]. Engagement with population level asymptomatic testing using lateral flow testing has also been shown to be low, particularly in socioeconomically disadvantaged areas[13].

Policy makers, considering which symptoms might prompt testing, tracing and isolation need to balance the availability of testing capacity at different stages of the pandemic, the speed with which samples can be taken and results returned, the harms incurred by asking large numbers of people who do not have COVID-19 and their contacts to self-isolate whilst awaiting test results, the consistency and simplicity of public health messaging and the likely public engagement with the system. Alteration of symptom profiles triggering testing and isolation may have less impact than other approaches to increase uptake and engagement with programmes and to ensure timely and effective contact tracing of household and non-household contacts.

The fact that COVID-19 may present as any of a very wide range of symptoms or with no symptoms at all is one of the key challenges in implementing successful TTI systems. Low levels of engagement also limit effectiveness. This emphasises the importance of not placing undue reliance on such systems as a mechanism to allow relaxation of other social distancing measures and the critical importance of protecting populations globally through immunisation.

Ethics:

This study has been approved by the Hampstead NHS Health Research Authority Ethics Committee. Ethics approval number - 20/HRA/2320.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Funding:

The research costs for the study have been supported by the MRC Grant Ref: MC_PC 19070 awarded to UCL on 30 March 2020 and MRC Grant Ref: MR/V028375/1 awarded on 17 August 2020. The study also received $15,000 of Facebook advertising credit to support a pilot social media recruitment campaign on 18th August 2020. This study was supported by the Wellcome Trust through a Wellcome Clinical Research Career Development Fellowship to RA [206602].

Conflicts of interest

ACH serves on the UK New and Emerging Respiratory Virus Threats Advisory Group. AMJ was a Governor of Wellcome Trust from 2011-18 and is Chair of the Committee for Strategic Coordination for Health of the Public Research.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

References

1. Alene M, Yismaw L, Assemie MA, et al. Magnitude of asymptomatic COVID-19 cases throughout the course of infection: A systematic review and meta-analysis. Kwok KO, ed. *PLoS ONE*. 2021;16(3):e0249090. doi:10.1371/journal.pone.0249090

2. Levin AT, Hanage WP, Owusu-Boaitey N, Cochran KB, Walsh SP, Meyerowitz-Katz G. Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *Eur J Epidemiol*. 2020;35(12):1123-1138. doi:10.1007/s10654-020-00698-1

3. Government launches COVID-19 Antivirals Taskforce to roll out innovative home treatments this autumn. GOV.UK. Accessed May 13, 2021. https://www.gov.uk/government/news/government-launches-covid-19-antivirals-taskforce-to-roll-out-innovative-home-treatments-this-autumn

4. WHO COVID-19 Case definition. Accessed May 13, 2021. https://www.who.int/publications-detail-redirect/WHO-2019-nCoV-Surveillance_Case_Definition-2020.2

5. Coronavirus Disease 2019 (COVID-19) | 2020 Interim Case Definition, Approved April 5, 2020. Accessed May 13, 2021. /nndss/conditions/coronavirus-disease-2019-covid-19/case-definition/2020/

6. Get a free PCR test to check if you have coronavirus (COVID-19). GOV.UK. Accessed May 13, 2021. https://www.gov.uk/get-coronavirus-test

7. Antonelli M, Capdevila J, Chaudhari A, et al. Optimal symptom combinations to aid COVID-19 case identification: analysis from a community-based, prospective, observational cohort. *medRxiv*. Published online February 8, 2021:2020.11.23.20237313. doi:10.1101/2020.11.23.20237313

8. Elliott J, Whitaker M, Bodinier B, et al. Symptom reporting in over 1 million people: community detection of COVID-19. *medRxiv*. Published online February 12, 2021:2021.02.10.21251480. doi:10.1101/2021.02.10.21251480

9. Hayward A, Fragaszy E, Kovar J, et al. Risk factors, symptom reporting, healthcare-seeking behaviour and adherence to public health guidance: protocol for Virus Watch, a prospective community cohort study. *medRxiv*. Published online December 16, 2020:2020.12.15.20248254. doi:10.1101/2020.12.15.20248254

10. UCL Virus Watch - help stop the spread of COVID-19. Accessed May 13, 2021. https://ucl-virus-watch.net

11. Oliu-Barton M, Pradelski BSR, Aghion P, et al. SARS-CoV-2 elimination, not mitigation, creates best outcomes for health, the economy, and civil liberties. *The Lancet*. 2021;0(0). doi:10.1016/S0140-6736(21)00978-8

12. Smith LE, Potts HWW, Amlôt R, Fear NT, Michie S, Rubin GJ. Adherence to the test, trace, and isolate system in the UK: results from 37 nationally representative surveys. *BMJ*. 2021;372:n608. doi:10.1136/bmj.n608

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

13. Liverpool COVID-19 community testing pilot: interim evaluation report summary. GOV.UK. Accessed May 13, 2021. https://www.gov.uk/government/publications/liverpool-covid-19-community-testing-pilot-interim-evaluation-report-summary/liverpool-covid-19-community-testing-pilot-interim-evaluation-report-summary

medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Tables and Figures

Table 1. Characteristics of illnesses by demographics and swab outcome

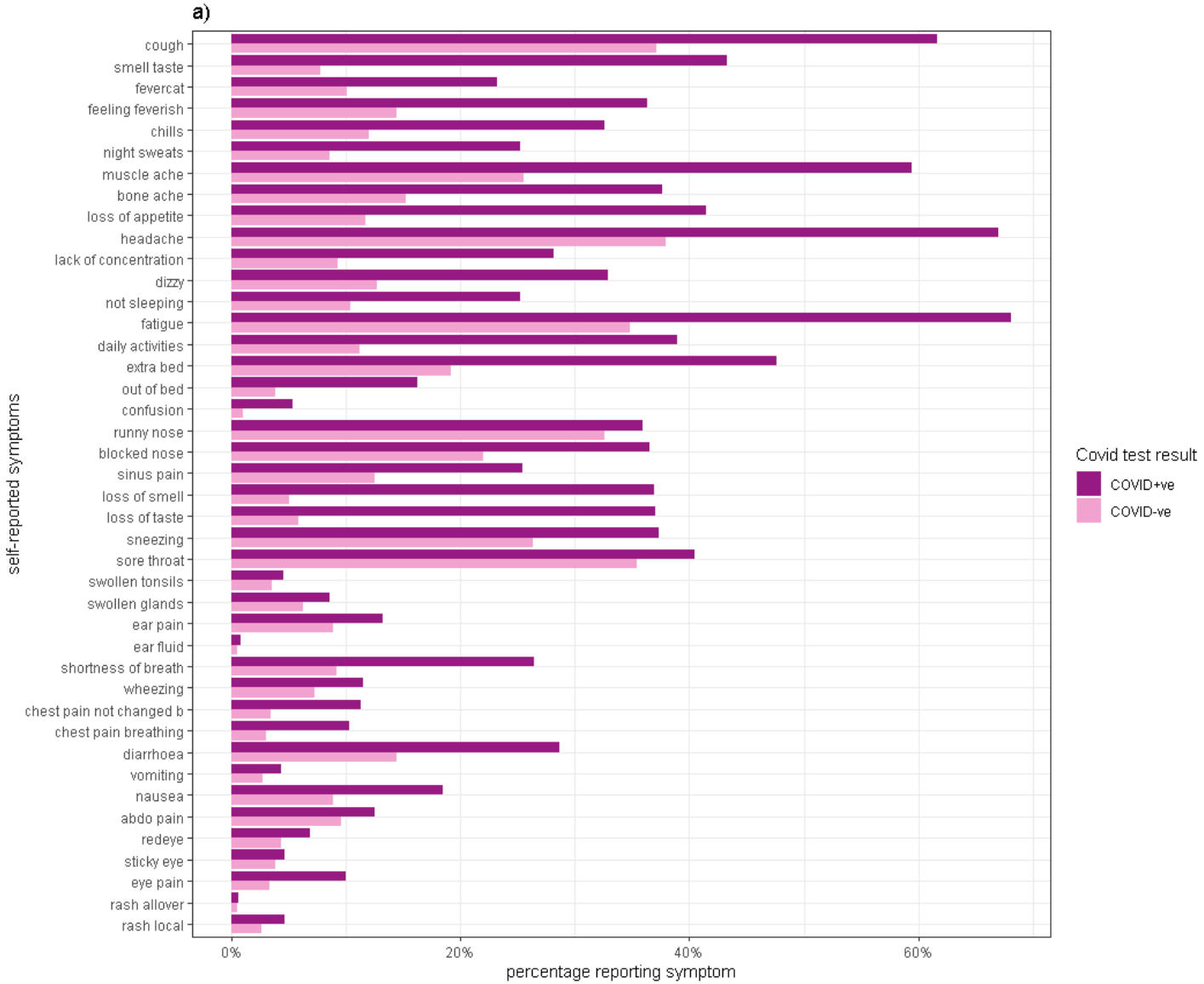
	All illnesses		Swabbed illnesses	Swab+ illnesses	
	N	% of all illnesses (column %)	N	N	% of swabbed illnesses (row %)
Overall	29083	100.0%	8213	944	11.5%
By Age Group*					
0-15	4110	14.3%	1114	76	6.8%
16-24	877	3.0%	300	54	18.0%
25-44	6637	23.0%	2153	270	12.5%
45-64	10350	35.9%	3158	379	12.0%
65+	6859	23.8%	1425	163	11.4%
By Sex**					
Male	9864	35.8%	2604	359	13.8%
Female	17713	64.2%	5197	517	9.9%
By Region***					
East Midlands	2535	9.0%	787	93	11.8%
East of England	4861	17.3%	1332	148	11.1%
London	4039	14.4%	1255	189	15.1%
North East	1505	5.4%	368	43	11.7%
North West	3225	11.5%	981	131	13.4%
South East	5394	19.2%	1520	119	7.8%
South West	2281	8.1%	529	49	9.3%
Wales	746	2.7%	165	19	11.5%
West Midlands	1793	6.4%	521	55	10.6%
Yorkshire and The Humber	1661	5.9%	470	59	12.6%
By Month					
Jun-Aug	1871	6.4%	203	5	2.5%
Sep	5646	19.4%	694	52	7.5%
Oct	3897	13.4%	876	105	12.0%
Nov	3412	11.7%	968	120	12.4%
Dec	4164	14.3%	1647	355	21.6%
Jan	2866	9.9%	1239	215	17.4%
Feb	2499	8.6%	809	54	6.7%
Mar	2740	9.4%	979	29	3.0%
Apr	1902	6.5%	773	9	1.2%
May	86	0.3%	25	0	0.0%

* Age missing for 250 illnesses

** Sex missing for 1506 illnesses

*** Region missing for 1043 illnesses

Figure 1. COVID symptoms**. a,** Self-reported symptoms by swab-confirmed COVID positive and COVID- illnesses. **b-c**, Proportion of COVID+ve illnesses **(b)** and COVID-ve illnesses **(c)** experiencing symptoms on a given day of illness within the first three week of illness. Day 1 represents the onset of symptoms.



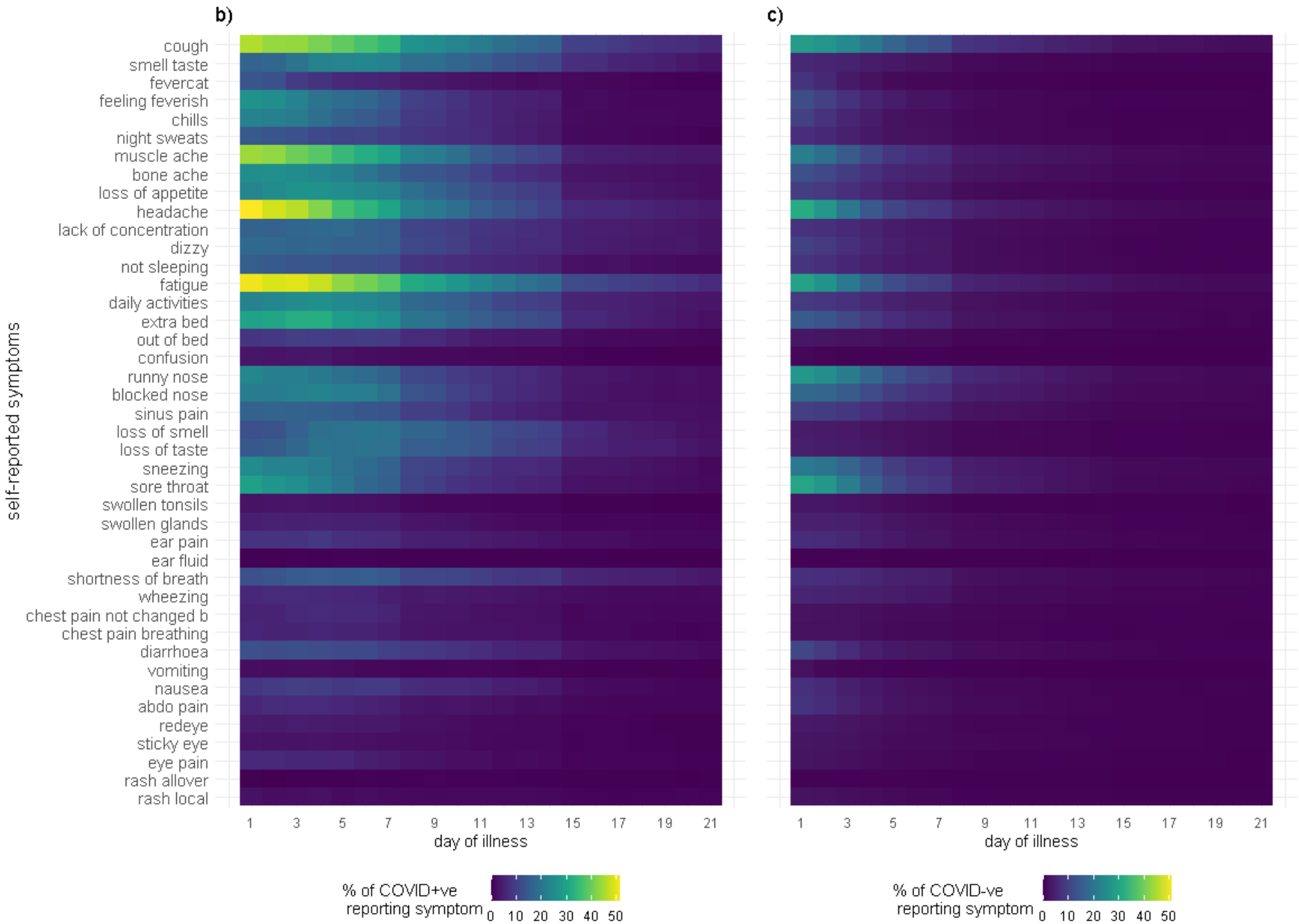
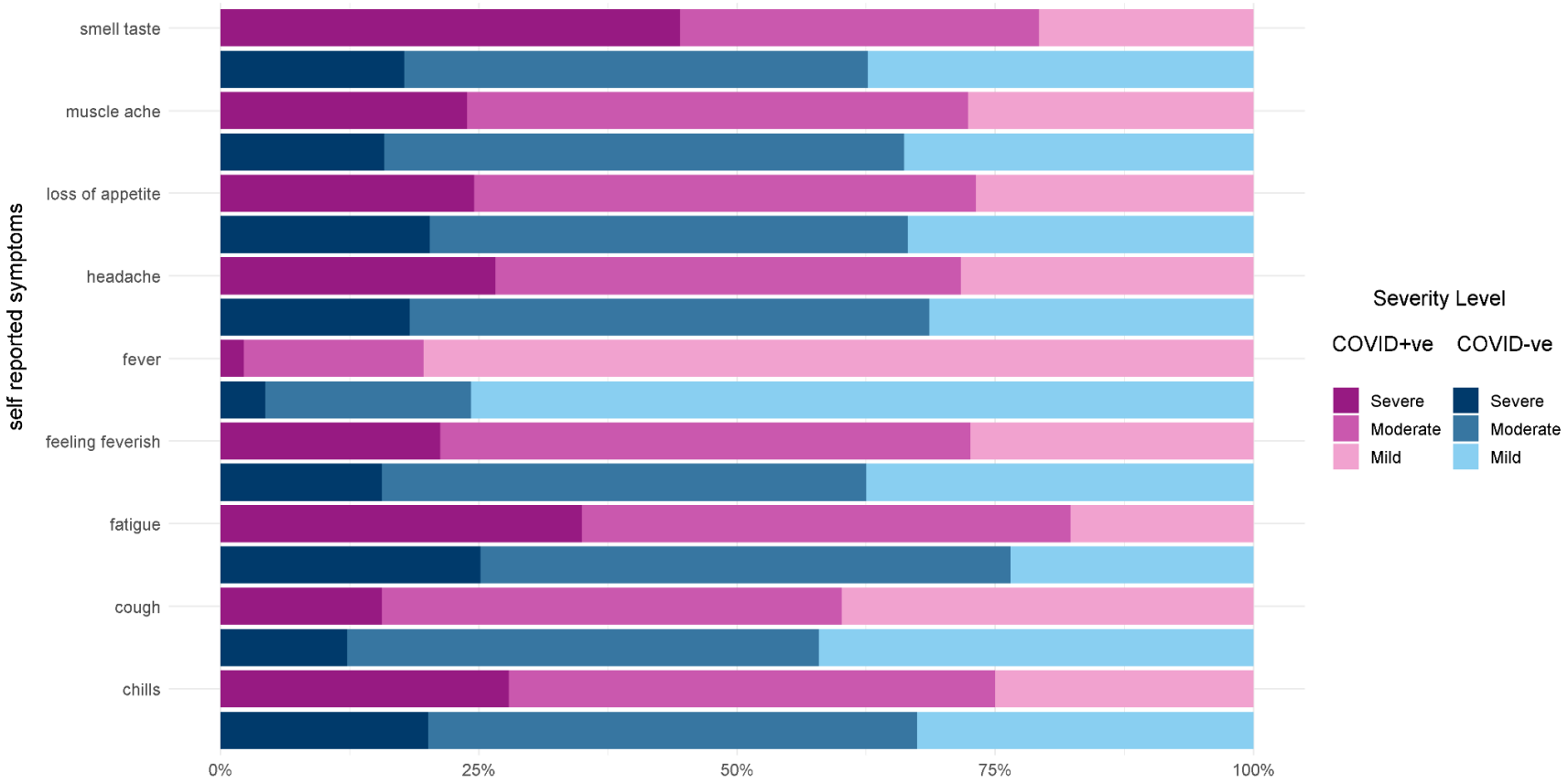

b)
c)
self-reported symptoms
cough
smell taste
fevercat
feeling feverish
chills
night sweats
muscle ache
bone ache
loss of appetite
headache
lack of concentration
dizzy
not sleeping
fatigue
daily activities
extra bed
out of bed
confusion
runny nose
blocked nose
sinus pain
loss of smell
loss of taste
sneezing
sore throat
swollen tonsils
swollen glands
ear pain
ear fluid
shortness of breath
wheezing
chest pain not changed b
chest pain breathing
diarrhoea
vomiting
nausea
abdo pain
redeye
sticky eye
eye pain
rash allover
rash local
1 3 5 7 9 11 13 15 17 19 21
day of illness
% of COVID+ve reporting symptom
0 10 20 30 40 50
% of COVID-ve reporting symptom
0 10 20 30 40 50

Figure 2. Severity of symptoms among swab-confirmed COVID positive and negative illnesses reporting the symptom



medRxiv preprint doi: https://doi.org/10.1101/2021.05.14.21257229; this version posted May 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Figure 3. Distribution of illness duration by COVID swab status

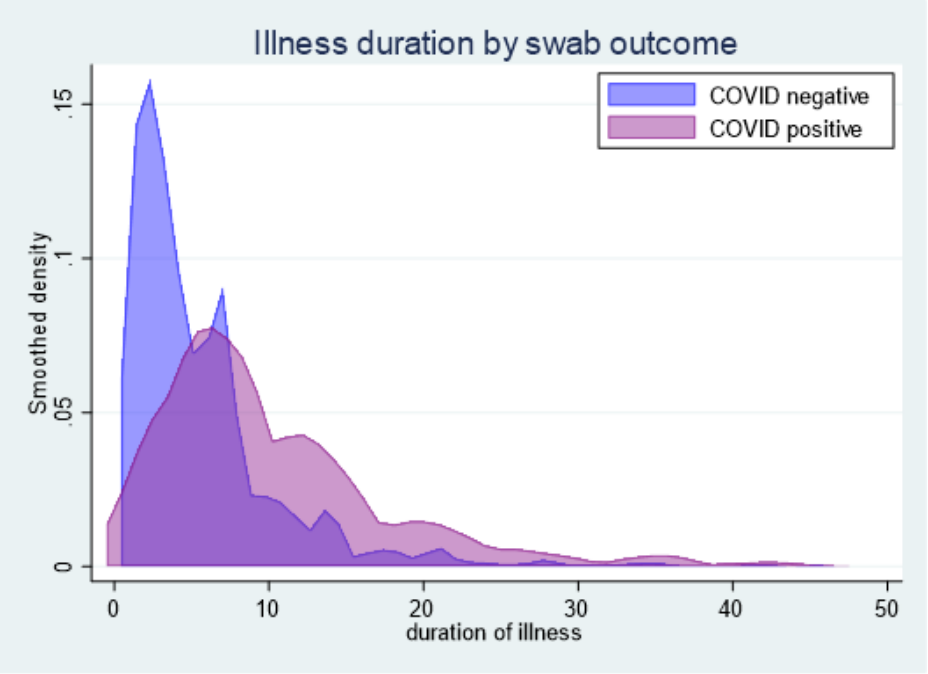


Table 2. Speed of identifying cases, proportion of all community illnesses requiring testing and test characteristics for the current and proposed UK Test and Trace case definitions

Case Definition	Strata		Onset Mean	Onset Median	Sensitivity	Specificity	PPV*	NPV*	NNT*	N illnesses eligible for testing	Multiplication factor
Current Case Definition	**Overall**	Overall	3.315	1	81%	47%	17%	95%	5.8	10134	1.00
	by age group	0-15	2.435	1	61%	28%	6%	91%	16.9	1566	1.00
		16-44	3.079	1	83%	47%	20%	95%	5.0	2682	1.00
		45-64	3.458	1	82%	52%	20%	95%	5.0	3535	1.00
		65+	3.171	1	81%	50%	19%	95%	5.3	2244	1.00
	by time period	Aug-Sep	3.81	1	80%	29%	8%	95%	13.2	2258	1.00
		Oct-Nov	2.783	1	84%	41%	18%	94%	5.6	2434	1.00
		Dec-Jan	2.612	1	82%	46%	28%	91%	3.5	2781	1.00
Proposed Case Definition	**Overall**	Overall	2.312	1	93%	26%	15%	96%	6.7	17154	1.69
	by age group	0-15	2.179	1	81%	21%	7%	94%	14.2	1963	1.25
		16-44	2.305	1	93%	26%	17%	96%	6.0	4552	1.70
		45-64	1.988	1	94%	26%	16%	97%	6.3	6482	1.83
		65+	2.413	1	95%	29%	16%	97%	6.2	4002	1.78
	by time period	Aug-Sep	2.583	1	89%	17%	7%	96%	14.0	3686	1.63
		Oct-Nov	2.474	1	94%	21%	15%	96%	6.5	4144	1.70
		Dec-Jan	1.888	1	94%	27%	25%	94%	4.0	4267	1.53

* PPV and NPV calcualated at a study COVID prevalence of 12.2%