

**Testing decisions in the Pandemic:**

**How do we use imperfect tests for COVID-19 (SARS-CoV-2) to make clinical decisions?**

Katy JL Bell  
Associate Professor  
The University of Sydney  
School of Public Health  
Sydney, New South Wales, Australia

Fiona Stanaway  
Senior Lecturer and Course Coordinator  
The University of Sydney  
School of Public Health  
Sydney, New South Wales, Australia

Les M Irwig  
Professor of Epidemiology  
The University of Sydney  
School of Public Health  
Sydney, New South Wales, Australia

Andrea R Horvath  
Director  
Prince of Wales Hospital and Community Health Services  
Clinical Chemistry and Endocrinology  
Randwick, New South Wales, Australia

Armando Teixeira-Pinto  
Professor  
The University of Sydney  
School of Public Health  
Sydney, New South Wales, Australia

Westmead Millennium Institute for Medical Research  
Centre for Kidney Research  
Westmead, New South Wales, Australia

Clement Loy  
Professor  
The University of Sydney  
School of Public Health  
Sydney, New South Wales, Australia

Director  
Westmead Hospital  
Huntington Disease Service  
Westmead, New South Wales, Australia

Competing interests: No relevant disclosures

Accepted Manuscript

**Abstract**

This is the second article of a series on how to use diagnostic test accuracy evidence. In this article, we explain how diagnostic test accuracy estimates may be applied in clinical practice, using the example of coronavirus disease 2019 (COVID-19). Using clinical scenarios, we guide the reader through the process of applying estimates of diagnostic accuracy to change the pre-test probability of disease, and make clinical decisions about their patient.

Word Count = 1593

Accepted Manuscript

## ***The Medical Journal of Australia – Pre-print – 2 October 2020***

A key recommendation for controlling the coronavirus disease 2019 (COVID-19) pandemic is to “test, test test!”(1). If we were able to test everyone using a test that was both 100% sensitive and specific for the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, we would have no false-positives (wrong identification of people without the infection) and no false-negatives (cases of infection are missed). With this perfect test we could identify, isolate and treat all infected individuals away from uninfected individuals. Contact tracing could identify further individuals in the incubation period for quarantine and testing as needed.

However, as we saw in the first article in this series(2), tests are rarely 100% sensitive and 100% specific, and usually there is a trade-off between the two. Typically, a highly specific test ( $Sp$ ) has few false positives so returning a positive ( $P$ ) result effectively rules in the diagnosis ( $SpPin$ ), but this is traded off against greater risk of false-negatives. In contrast, a highly sensitive test ( $Sn$ ) has few false negatives so returning a negative result ( $N$ ) effectively rules out the diagnosis ( $SnNout$ ), but this is traded off against greater risk of false-positives. Although these are useful ‘rules of thumb’, in order to make clinical decisions we usually need a more defined probability of disease in an individual patient. In this second article, we use the clinical example of COVID-19 to guide the reader through how to apply estimates of diagnostic test accuracy in clinical practice. The process of making a diagnosis can be thought of as an estimation and re-estimation of the probability of disease in a continuous process. When thought of in that way, each item of history and examination, as well as laboratory and imaging tests, are individual tests in themselves, with the results on each increasing or reducing the probability of disease.

For example, suppose a 26 year-old man in Sydney presents with a blocked nose, cough, and fever in late March 2020. To help you determine whether this could be COVID-19 rather than something else (e.g. other viral respiratory pathogen), you might ask him about contact with anyone with COVID-19, shortness of breath, recent travel, and where he lives (some areas have higher risk of community transmission). The answers to these and other questions would give you some kind of pre-test probability for COVID-19. You would then readjust the disease probability after further questions (e.g. loss of smell or taste, duration of symptoms, past medical history), physical examination (if undertaken), and tests results such as SARS-CoV-2 tests, blood tests (e.g. full blood count, C-reactive protein), and chest imaging tests. If we know the “test” accuracy for each of these steps then we can continuously recalculate the probability of disease.

### **Likelihood Ratios**

The conceptual approach outlined above, updating disease probability for a patient once we have new information from a diagnostic test, can be operationalised by calculating likelihood ratios for the test and using Fagan’s nomogram(3), by estimating post-test probability (Figure 1)(4). Likelihood ratios describe the ratio of the probability of a test result in people who truly have the disease, to the probability of the same test result in people who truly do not have the disease. In the simplest case of a dichotomous test (positive or negative for SARS-CoV-2), we can calculate a positive likelihood ratio to decide on implications of a positive test result for our patient, and a negative likelihood ratio for implications of a negative test result.

Box 1 outlines some relevant formulae for these calculations, and interested readers may enjoy reading more about how the Bayes theorem applies to diagnostic tests(4).

When a likelihood ratio is greater than 1, the probability of disease after receiving this test result (post-test probability) is increased from what it was before the test (pre-test probability). The higher the likelihood ratio, the more powerful the test is at shifting the probability of disease upwards from pre- to post-test, and the better the test is at ruling in disease. Conversely when the likelihood ratio is less than 1, then the probability of disease is decreased from what it was before doing the test. The lower the negative likelihood ratio, the more powerful the test is at shifting the probability of disease downwards from pre- to post-test, and the better the test is at ruling out disease.

### **Decision thresholds**

The decision to request a test for a patient needs to take into consideration not just the accuracy of the test, but also how the test results will help us make clinical decisions. To do this, the first step is to define decision thresholds for different types of management. We use tests to move the probability of disease from one side of these decision thresholds to the other. These thresholds may be set by considering the possible impacts that each course of action may have, and by weighing up the benefits and harms of each approach. For example, we may use increasingly higher post-test probability thresholds of SARS-CoV-2 infection to recommend: self-isolation while symptomatic with a respiratory infection, strict quarantine for 14 days, diagnostic testing in people who have contact with vulnerable members of the community (e.g. healthcare workers, aged care workers), diagnostic testing in people who do not have such contact, diagnosis of SARS-CoV-2 infection with strict isolation, and tracing and quarantining/testing all contacts, and diagnosis of COVID-19 with consideration of treatment and/or of enrolling in a trial of experimental treatment (Box 2).

### **Pre-test probability**

The next step after setting decision thresholds is to decide on our patient's pre-test probability. Often it is not something that we consciously think about, even though we may intuitively use it to make clinical decisions. Although we normally rely on this "clinical intuition" – our best guess of probability of disease from our clinical experience - this may be less useful in the setting of a new disease like COVID-19. To explicitly estimate the pre-test probability of a patient, we might use the prevalence of disease in a similar clinical cohort from the literature, or from our local database. We would need to make sure the characteristics of the study cohort used for the prevalence estimate (the denominator of people at risk of infection included in the calculation) matches the clinical cohort for our patient. This should include the specific time and place the estimates are based on, and importantly the criteria for testing used.

### **COVID-19 diagnostic test accuracy**

We now move to finding and applying the evidence on diagnostic test accuracy. After a brief literature search, we identify several test accuracy studies for detecting the SARS-CoV-2 virus with reverse transcriptase polymerase chain reaction (RT-PCR) molecular testing. As we have outlined previously(2), we may use the acronym 'RAM' to decide if a diagnostic accuracy study is likely to be valid and applicable in our own setting: Representative (covers a similar spectrum of patients); Ascertainment (a reference standard is used to verify all test results); Measurement (an accurate reference standard is applied independently and without knowledge of the test result). There is generally a high risk of bias in available diagnostic accuracy studies on SARS-CoV-2/COVID-19 tests (5, 6) as well as important applicability issues(7). Nevertheless, we find a paper that reports on 1014 patients presenting to a one hospital in Wuhan, China, who

underwent both chest computed tomography (CT) and RT-PCR tests for the evaluation of possible COVID-19(8). Using a reference standard that combines all clinical information and all test results available in the paper, we calculate sensitivity, specificity, and likelihood ratios for RT-PCR and chest CT (Table 1; RT-PCR estimates for sensitivity are similar to others' (5, 6, 9)). This retrospective hospital based study may overestimate diagnostic accuracy for our intended test use in the community, however, the study appears to have a lower risk of bias than many others published to date.

### **Post-test probability**

Back to the 26 year-old man in Sydney presenting with a blocked nose, cough, and fever in late March 2020. He tells us that he lives in Bondi, and using available prevalence data, we estimate that his pre-test probability estimate is around 6%. His pre-test probability is above the testing threshold of 2% (general population) and so we collect a nasopharyngeal swab and request an RT-PCR test for the SARS-CoV-2 virus and recommend that he self-isolates at home. Two days later we receive the RT-PCR results, which are negative. He has a positive result for another respiratory virus, however, we are aware that co-infection is possible and that this alone does not rule out COVID-19(10). Using Fagan's nomogram and the negative LR we calculated, we revise the probability of COVID-19 downwards to 3%, which is still above the testing threshold (Blue line in Figure 1). A second negative PCR is needed before the probability of disease is sufficiently low to rule out COVID-19 and stop testing (but he still needs to self isolate while symptomatic). In the scenario where his housemate (a close contact) has COVID-19, more than two negative PCR results are needed to rule out the disease and a chest CT may be helpful (see



scenarios 1C-1G in Table 2). The process of test interpretations in this and three other clinical scenarios are presented in Table 2, along with key concepts illustrated.

### **Conclusion**

Tests for COVID-19, like tests for other diseases, are not perfect and will sometimes miss people who are infected and falsely identify others who are not infected. Despite this, by using likelihood ratios to estimate the post-test probability of disease, we can make the important clinical decisions needed to provide patient-centred care while also preventing spread of the pandemic.

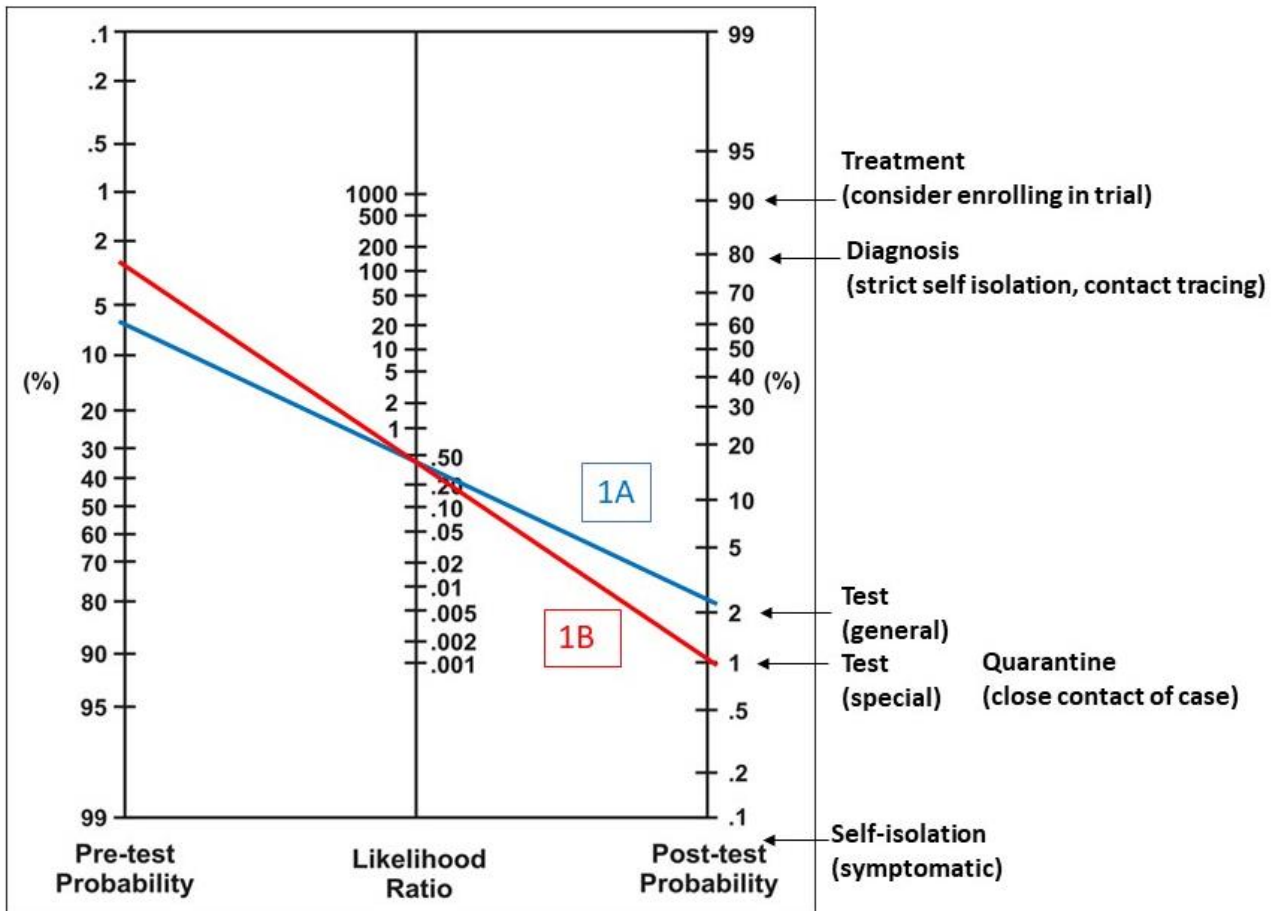
**References**

1. World Health Organisation. WHO Director-General's opening remarks at the media briefing on COVID-19 - 16 March 2020 [press release].  
<https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---16-march-2020> (accessed August 2020).
2. Bell KJ, Macaskill P, Loy C. Test accuracy and potential sources of bias in diagnostic test evaluation. *Med J Aust.* 2020;212:10-13.e1.
3. Fagan T. Nomogram for Bayes's Theorem. *N Engl J Med.* 1975;293:257.
4. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA.* 1994;271:703-7.
5. Watson J, Whiting PF, Brush JE. Interpreting a covid-19 test result. *BMJ.* 2020;369:m1808.
6. Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D, Zambrano-Achig P, del Campo R, Ciapponi A, et al. False-negative results of initial RT-PCR assays for COVID-19: a systematic review. *medRxiv.* 2020; DOI: <https://doi.org/10.1101/2020.04.16.20066787>
7. Bossuyt PM. Testing COVID-19 tests faces methodological challenges. *J Clin Epidemiol.* 2020; DOI: <https://doi.org/10.1016/j.jclinepi.2020.06.037>
8. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology.* 2020;200642.

9. Woloshin S, Patel N, Kesselheim AS. False Negative Tests for SARS-CoV-2 Infection - Challenges and Implications. *N Engl J Med.* 2020; 383:e38.
10. Kim D, Quinn J, Pinsky B, Shah NH, Brown I. Rates of Co-infection Between SARS-CoV-2 and Other Respiratory Pathogens. *JAMA.* 2020; 323:2085-2086.
11. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ.* 2004;329(7458):168.

Accepted Manuscript

Figure 1: Fagan’s Nomogram for COVID-19



## Notes

1. Fagan’s Nomogram modified from Figure in JAMA 1994;271(9):703-707. To use the nomogram, draw a straight line from the pre-test probability for the patient (lefthand side) through the likelihood ratio for the patient’s test result (middle) to arrive at the estimated post-test probability of disease (righthand side).
2. Depicts Scenarios 1 from Table 2: 26 year old man from Bondi presents with blocked nose, cough and fever, pre-test probability 6%.
3. **Blue line (1A)**: First PCR test negative, negative likelihood ratio = 0.41, post test probability =3% (1A from Table 2)
4. **Red line (1B)**: Second PCR test negative, post test probability =1% (1B from Table 2)

**Box 1: Likelihood Ratio definitions**

The Likelihood Ratio (LR) is the probability of a given test result in a patient with the disease, compared to the probability of same result in a patient without the disease(11).

- Positive LR = probability of a positive result among people with disease (sensitivity) / probability of a positive result among people without disease (1-specificity)
- Negative LR = probability of a negative result among people with disease (1-sensitivity) / probability of a negative result among people without disease (specificity)

As well as allowing application of diagnostic test accuracy estimates to a patient in the clinic, likelihood ratios have other advantages over sensitivity and specificity, including allowing for multi-category results (interval likelihood ratios).

**Box 2: Possible Decision Thresholds for SARS-CoV-2/COVID-19<sup>1</sup>**

| <b>Management</b>   | <b>Individual Disease<br/>Probability threshold before<br/>recommendation</b> |
|---|---|
| Self isolate (while symptomatic)  | ≥0%   |
| Quarantine (14 days)  | >1%   |
| Testing threshold - special (healthcare workers, aged care workers)                       | >1%   |
| Testing threshold – general (not in contact with vulnerable)                              | >2%   |
| Diagnostic threshold - strict isolation, trace and test all contacts                      | >80%  |
| Treatment threshold – might include enrolment in trial of experimental COVID-19 treatment | >90%  |

Note:

1. SARS-CoV-2: severe acute respiratory syndrome coronavirus 2; COVID-19: coronavirus disease 2019
2. These decision thresholds are for illustrative purposes only and are not based on evidence of benefits vs harms; readers may choose to use their own.

**Table 1: Sensitivity and Specificity of RT-PCR and chest CT for COVID-19 in 1014 patients suspected of having COVID-19**

|                              | COVID-19+ | COVID-19 - |      | Sens | Spec | LR+   | LR-  |
|------------------------------|-----------|------------|------|------|------|-------|------|
| <b>RT-PCR+ (single test)</b> | 500       | 1          | 501  |      |      |       |      |
| <b>RT-PCR- (single test)</b> | 350       | 163        | 513  |      |      |       |      |
|                              | 850       | 164        | 1014 | 59%  | 99%  | 96.47 | 0.41 |
| <b>Chest CT+</b>             | 830       | 58         | 888  |      |      |       |      |
| <b>Chest CT-</b>             | 20        | 106        | 126  |      |      |       |      |
|                              | 850       | 164        | 1014 | 98%  | 65%  | 2.76  | 0.04 |

**Notes:**

1. RT-PCR: reverse transcriptase polychain reaction; CT: computed tomography; COVID-19: coronavirus disease 2019; Sens: sensitivity; Spec: specificity, LR+: positive likelihood ratio; LR-: negative likelihood ratio; +:positive; -:negative.
2. Tables constructed from data presented in Ai et al(8). We assumed that all 580 patients with positive RT-PCR and positive CT, 20/21 patients with positive RT-PCR and negative CT, and 250/308 patients with negative RT-PCR tests and positive CT had COVID-19. We assumed that the remaining 164 people did not have COVID-19, including one person with positive initial RT-PCR and negative CT, and negative subsequent RT-PCR tests and serial CTs. Similarly, we assumed that 58/308 people with negative PCR and positive CT, and all 105 people with negative RT-PCR and negative CT, did not have COVID-19.
3. RT-PCR results for a single test are calculated as follows. In the report by Ai et al(8), individuals had up to three repeat tests after the initial RT-PCR, and were classified as RT-PCR positive if any of the tests were positive. Among 90 individuals where RT-PCR

was ultimately positive (and at least 4 days between repeated tests), 15 individuals were initially RT-PCR negative. We therefore assumed that 17% (15/90) of the 600 RT-PCR+ve COVID-19 patients would be RT-PCR-ve on a single test. This means that of the 851 COVID-19 patients, 500 would be RT-PCR+ on a single test.

Accepted Manuscript

**Table 2: Hypothetical scenarios of patients in community setting<sup>1</sup>**

| Scenario   |   | pre-test probability | New test result | post-test probability | Decision threshold  | Management <sup>1</sup>                       | Key Concepts  |
|--|---|----------------------|-----------------|-----------------------|---|---|---|
| 1. 26 year old man from Bondi, Sydney presents in late March 2020 with blocked nose, cough, and fever. | 1A. Nil   | 6%                   | RT-PCR -ve      | 3%                    | Self isolate (while symptomatic)<br>Testing threshold – general | Keep testing (go to 1B)                       | <ul style="list-style-type: none"> <li>•In situations of higher pre-test probability may need multiple negative test results before can rule out SARS-CoV-2/COVID-19</li> <li>•Different thresholds for self isolation while symptomatic vs strict quarantine vs ruling out infection vs ruling in infection</li> <li>•RT-PCR may be better at ruling in infection than ruling out (Positive result on highly Specific test: SpPIN)</li> <li>•Chest CT may be better at ruling out infection than ruling in (Negative result on highly Sensitivity test: SnNOUT)</li> </ul> |
|  | 1B. First PCR -ve.  | 3%                   | RT-PCR -ve      | 1%                    | Self isolate (while symptomatic)                                | Stop testing                                  |   |
|  | 1C. Housemate has been found to have COVID-19.                    | 15%                  | RT-PCR -ve      | 7%                    | Quarantine (14 days)<br>Testing threshold – general             | Keep testing (go to 1D), repeat PCR           |   |
|  | 1D. Housemate has been found to have COVID-19. First PCR -ve.     | 7%                   | RT-PCR -ve      | 3%                    | Quarantine (14 days)<br>Testing threshold – general             | Keep testing (go to 1E), move on to CT        |   |
|  | 1E. Housemate has been found to have COVID-19. Two PCR tests -ve. | 3%                   | Chest CT -ve    | 0%                    | Quarantine (14 days)  | Stop testing                                  |   |
|  | 1F. Housemate has been found to have COVID-19. Two PCR tests -ve. | 3%                   | Chest CT +ve    | 8%                    | Quarantine (14 days)<br>Testing threshold – general             | Keep testing (go to 1G), repeat PCR           |   |
|  | 1G. Housemate has been found to have                              | 8%                   | RT-PCR +ve      | 89%                   | Diagnostic threshold  | Strict isolation, trace and test all contacts |   |



**The Medical Journal of Australia – Pre-print – 2 October 2020**

|   | COVID-19. Two PCR tests -ve and CT +ve   |     |            |     |   |   |   |
|---|--|-----|------------|-----|---|---|---|
| 2. 42 year old woman in Brisbane, close contact of a confirmed case in late April 2020 with no symptoms.              | 2A. Nil                                  | 1%  | nil        | 1%  | Quarantine (14 days)                                | Don't test                                    | <ul style="list-style-type: none"> <li>•As all case contacts undergo strict quarantine they may not need to be tested unless pre-test probability is increased (e.g. develop symptoms)</li> <li>•Two negative RT-PCR results in a case contact with symptoms may be needed to stop testing while continuing quarantine which prevents further transmission</li> </ul> |
|   | 2B. Develops sore throat.                | 6%  | RT-PCR -ve | 3%  | Quarantine (14 days)<br>Testing threshold – general | Keep testing (go to 2C), repeat PCR           |   |
|   | 2C. Develops sore throat. First PCR -ve. | 3%  | RT-PCR -ve | 1%  | Quarantine (14 days)                                | Stop testing                                  |   |
|   | 2D. Develops sore throat.                | 6%  | RT-PCR +ve | 86% | Diagnostic threshold                                | Strict isolation, trace and test all contacts |   |
| 3. 56 year old woman in Melbourne, works as nurse in an aged care facility. Has a sore throat and fever in June 2020. | 3A. Nil                                  | 2%  | RT-PCR -ve | 1%  | Self isolate (while symptomatic)                    | Stop testing                                  | <ul style="list-style-type: none"> <li>•We need to use a lower threshold to rule out disease in people who pose higher risk of spreading infection to at risk communities like aged care facilities as the</li> </ul>   |
|   | 3B. Nil                                  | 2%  | RT-PCR +ve | 66% | Testing threshold – special                         | Keep testing (go to 3C), repeat PCR           |   |
|   | 3C. First PCR had +ve result             | 66% | RT-PCR -ve | 45% | Testing threshold – special                         | Keep testing (go to 3D), repeat PCR           |   |

**The Medical Journal of Australia – Pre-print – 2 October 2020**

|   |   |     |              |      |                                  |  |   |
|---|---|-----|--------------|------|----------------------------------|--|---|
|   | 3D. First PCR +ve, 2 <sup>nd</sup> PCR -ve. | 45% | RT-PCR -ve   | 25%  | Testing threshold – special      | Keep testing (go to 3E), move on to CT                                 | consequences of getting this wrong are more severe.<br>•Chest CT may be better at ruling out disease than PCR   |
|   | 3E. PCR results: +ve, -ve, -ve              | 25% | Chest CT -ve | 1%   | Self isolate (while symptomatic) | Stop testing   |   |
| 4. 72 year old man in Melbourne, presents July 2020 with sore throat, fever, myalgia. | 4A. Nil                                     | 4%  | RT-PCR -ve   | 2%   | Self isolate (while symptomatic) | Stop testing   | •We need a higher threshold to consider treatment than to make the diagnosis – especially in the setting of COVID-19 where benefits and harms of experimental treatments are unknown. |
|   | 4B. Nil                                     | 4%  | RT-PCR +ve   | 80%  | Diagnostic threshold             | Strict isolation, trace and test all contacts (go to 4C)               |   |
|   | 4C. First PCR +ve                           | 80% | RT-PCR +ve   | 100% | Treatment threshold              | Consider treatment and/or enrolling in trial of experimental treatment |   |

Notes:

1. RT-PCR: reverse transcriptase polychain reaction; CT: computed tomography; COVID-19: coronavirus disease 2019; -ve: negative; +ve: positive.
2. The suggested management in these scenarios are for illustrative purposes only and should not be taken as clinical guidance recommendations.