In this study, the complicated reasoning and processes inherent in diagnostic testing were analyzed, and a mathematical theory was developed for effectively stopping the transmission of infection in the context of coronavirus disease 2019 (COVID-19). As a result of this work, a new formula was developed for the “boundary condition for contagion containment,” which, based on a horizontal transmission model, gives the lower limit of sensitivity for a diagnostic test to stop the virus spreading. Two parameters are considered in the model: the level of transmission and the effective reproduction number. In example computations, the formula indicated that a one-off polymerase chain reaction-based test with a sensitivity of 85% would not be sufficient to contain highly contagious infections such as the Delta variant of SARS-CoV-2, which would likely require a sensitivity close to 100% for its containment. Furthermore, a cascade judgment system for multiple tests was proposed and examined as a form of triplet test system. This approach can enhance the accuracy of COVID-19 testing up to the minimum level needed to stop the virus spreading. The theory developed in this study will not only contribute as an academic exercise, but also be useful for making evidence-based decisions on public policy for pandemic control. (DOI: 10.2302/kjm.2021-0009-IR) 

Keywords: coronavirus, diagnostic test, sensitivity and specificity, transmission model

Introduction

Coronavirus disease 2019 (COVID-19) was first identified in Wuhan, China, at the end of 2019 and was later declared a pandemic by the World Health Organization (WHO); it has dramatically changed the lifestyles of people around the world. The global crisis has persisted, and Johns Hopkins University reported the number of confirmed cases worldwide to be 186,763,389 as of July 11, 2021, of which 4,029,580 resulted in death (i.e., a crude fatality rate of 2.2%).

Testing for COVID-19 is required not only in the clinical setting for diagnosis and decision-making for treatment, but also in the public health context because it enables the identification of infected individuals and facilitates their isolation to reduce the spread of infection. The comprehensive preventive strategies of “testing and tracing” are just as important as vaccinations and treatments, if not more so.

According to the WHO definition, a confirmed COVID-19 case is a case with the laboratory confirmation of SARS-CoV-2 infection, irrespective of clinical signs and symptoms. The polymerase chain reaction (PCR)-based test is the most accurate for identifying coronavi-
virus infection. Therefore, SARS-CoV-2 infection is best confirmed by a positive PCR test. In clinical settings, however, a rapid antigen test is often utilized as a convenient tool, despite the test accuracy being less than that of a PCR test.

In many countries, it is now recognized that, because of insufficient test accuracy, the preventive strategy of “testing and tracing” did not work effectively to control the surge of infected cases. No test produces results that are 100% accurate, with some false-negative and false-positive results inevitably occurring. Because of its high specificity for identifying COVID-19 RNA, PCR testing is effective in confirming a diagnosis at the level of an individual patient suspected of being infected with COVID-19. However, the risk of false-negative results is a concern in terms of epidemic control from the perspective of public health and nosocomial infection.

To determine the accuracy of a test, such as its sensitivity (1−false-negative rate) and specificity (1−false-positive rate), a “gold standard” is needed that can confirm the presence of the target disease. However, the absence of such a clear-cut “gold standard” for COVID-19 testing makes the evaluation of test accuracy challenging. Certain ranges of variation have been reported in the literature regarding the sensitivity and specificity of the PCR test for COVID-19. High specificity has generally been observed, but with moderate sensitivity. Based on a systematic review, Rodriguez et al. reported sensitivity to be in the range 71%–98%. In addition, Padhye estimated the sensitivity to be 70.7% (95% CI: 66.8–74.9%) and the specificity to be 85.1% (95% CI: 77.4–94.1%), although a caveat about the generalizability of these findings was raised because of the diverse geographical origins of the data. The COVID-19 Infection Survey Studies, Office for National Statistics, UK, suggested that the sensitivity of PCR tests may be somewhere between 85% and 98%, whereas the specificity was likely very close to 100%, with a lowest estimate of 99.92%. Watson et al. employed approximate values of 70% for sensitivity and 95% for specificity, for illustrative purposes, whereas Ontario Public Health reported 85% for sensitivity and 99.99% for specificity. For the illustrative computations performed in the current study, given the wide range of estimates reported in the literature, PCR test sensitivity is assumed to be 85% and specificity to be 99.99%, as per Ontario Public Health’s recommendations.

Assuming a moderate sensitivity and a high specificity for PCR tests, the author aimed, as an academic exercise, to analyze the complicated reasoning and processes inherent in diagnostic testing and to develop a new theory for infection control. The goal of this study was to find the necessary test accuracy and to explore a test system that can effectively stop the spread of infection in the context of COVID-19.

Methods

The issue is how diagnostic testing can decrease the number of infected individuals in a population vulnerable to the spread of the virus. This study gives a theoretical solution and numerical computations for each of the following questions from the perspectives of test accuracy and testing system architecture.

Question 1: What test accuracy is required to stop the virus spreading?

In epidemiological statistics, a transmission model is conventionally used to predict the outbreak of an infectious disease. Consequently, to resolve the issue of how to decrease the number of infected cases, the following horizontal transmission model was employed in this analysis.

Horizontal transmission model

1. Let \( D^+ \) be the number of cases truly infected in a population tested for diagnosis at time \( t \). It should be noted that \( D^+ \) is not identical to the number of cases that test positive, \( T^+ \), in the context of the 2 × 2 table shown in Appendix Fig. 1. If the disease prevalence, \( D^+/N \), is known (where \( N \) is the total number tested), it is conventional to calculate \( D^+ \) given \( T^+ \), a known sensitivity \( S_n \), and a known specificity \( S_p \). If the prevalence is not known in advance, as is often the case in testing, a sophisticated estimation of \( D^+ \) must be conducted based on the 2 × 2 table shown in the Appendix.

2. Those testing positive, mainly consisting of true positives, are isolated and transmit the virus to no other individuals in the population.

3. The diagnostic tests yield false negatives as a result of insufficient test sensitivity, \( S_n \), among the cases of \( D^+ \) truly infected.

4. The false-negative cases secondarily transmit the virus directly to some individuals by the timing of the next test, \( t+1 \).

5. Secondary infections horizontally expand to the next level according to the effective reproduction number, \( R_n \), which is an indicator of infectivity showing how many individuals will be infected directly from one infected person in the actual setting of the current state of a population.

Based on this horizontal transmission model, the required accuracy of a diagnostic test will be analyzed to find the conditions necessary for stopping the transmission of infection.
**Question 2: How should a test system be set up to attain the test accuracy required to stop the virus spreading?**

A solution to Question 1 raises the second question of how the protocols or systems of COVID-19 testing should be organized in practice to attain the test accuracy predicted in theory. To address this question, a system architecture with cascade testing is considered and theoretically analyzed because a cascade combination of multiple tests has the potential to enhance the accuracy of testing beyond that of a one-off PCR test.

**Results**

**Solution to Question 1**

With the horizontal transmission model, the number of false negatives resulting from testing is expressed as $D^+(1–S_n)$ because the false-negative rate is $1–S_n$. Those with a false-negative result retain infectivity to others. Consequently, the number of cases secondarily infected at time $t+1$ can be expressed as $D^+(1–S_n)R_t$. Consequently, the total number of cases that are infectious at time $t+1$ is the sum of the two groups, the primary and secondary infections, as follows:

$$D^+(1–S_n)+D^+(1–S_n)R_t = D^+(1–S_n)(1+R_t)$$

This process is summarized as follows: the infectious population $D^+$ with an effective reproduction number $R_t$ becomes $D^+(1–S_n)(1+R_t)$ as a result of a testing intervention with sensitivity $S_n$. To make the infection rate converge, the total number of infected cases at the secondary level, including false negatives after testing, must be fewer than that before, i.e., the following inequality must be satisfied:

$$D^+(1–S_n)(1+R_t)<D^+$$

From this, the following inequality is further developed:

$$(1–S_n)(1+R_t)<1$$

When it is solved for sensitivity $S_n$, the formula is transformed into:

$$S_n>R_t / (1+R_t)$$

To make the infection rate converge, this inequality indicates that the test sensitivity must be higher than $R_t / (1+R_t)$, a quantity determined solely by $R_t$. Therefore, we may call this inequality “the boundary condition for contagion containment.” Moreover, by substituting $S_n$ with the false-negative rate (FNR), based on the equation $S_n=1–FNR$, the following relationship is derived:

$$FNR<1 / (1+R_t)$$

In other words, for an infectious disease to converge, the false-negative rate must be lower than the reciprocal of the effective reproduction number plus one. For example, if $R_t$ is 2.5 in the case of COVID-19, the reciprocal of 3.5 is 0.286. Therefore, the test sensitivity must be 71.4% or greater (or the false-negative rate must be lower than 28.6%). The assumption of a sensitivity of 85% is greater than the minimum estimate of 71.4%.

Conversely, if the effective reproduction number is considered as a function of sensitivity, for a test with sensitivity of $S_n$ to be able to suppress the scale of infection, the effective reproduction number must not exceed the upper limit as expressed by the following inequality:

$$R_t< S_n / (1–S_n)$$

For example, if the PCR test for COVID-19 has a sensitivity of 85%, the effective reproduction number must be lower than 5.7 $[=0.85 / (1–0.85)]$ for the viral spread to converge. Therefore, in a highly infectious situation where the effective reproduction number might exceed 5.7, the conventional one-off application of PCR tests would be ineffective in containing the spread of the virus. This high reproduction number likely occurs in physical situations characterized by the “three Cs”: closed, crowded, and close contact.

The model of horizontal transmission to the secondary level can be extended to a more general setting of transmission to the k-th level ($k=1, 2, 3, \ldots$). That is, if the secondary infected cases further transmit the virus to the tertiary level with effective reproduction number $R_t$, the number of tertiary infected cases becomes $D^*(1–S_n)R_t^2$. Similarly, consider a cascade of transmissions continued to the k-th level. Then, the total number of cases infected is formulated as follows:

$$D^*(1–S_n)+D^*(1–S_n)R_t+D^*(1–S_n)R_t^2+D^*(1–S_n)R_t^3+\ldots+D^*(1–S_n)R_t^{k–1}=\Sigma D^*(1–S_n) R_t^{i−1}=D^*(1–S_n) (1–R_t^k) / (1–R_t)$$

(Where the notation of $\Sigma X_i$ means summations of $X_i$ for $i=1, 2, 3, \ldots, k$)

This total number of infected cases must be smaller than that before testing. Therefore, for viral transmissions to converge, the following inequality must be satisfied:

$$D^*(1–S_n) (1–R_t^k) / (1–R_t)<D^+$$

When this is solved for sensitivity $S_n$, the result is as follows:

$$S_n>(R_t−R_t^k) / (1–R_t^k)$$

This inequality prescribes the minimum limit of the test sensitivity required to contain the infection by testing, which is independent of $D^+$ in the current model. That is, the test sensitivity $S_n$ must be higher than the quantity $(R_t−R_t^k) / (1–R_t^k)$, a term that is expressed with two parameters, the effective reproduction number $R_t$ and the level of viral transmission reached by time $t+1$.

**Fig. 1** presents an example computation for the “bound-
ary condition for contagion containment" in practical settings regarding combinations of RT and k (RT in the range 1.5–5; k in the range 2–6). In the case of COVID-19 PCR testing, the assumed sensitivity of 85% covers all of the cells at the level of k = 2 and the cell with RT = 1.5 and k = 3. If the sensitivity were to be reduced to 70% in the worst case, only the white cells (RT ≥ 2 and k = 2) are associated with effective containment. This means that a sensitivity of 70% is too low to achieve convergence of the infection.

The Delta variant of SARS-CoV-2 is more contagious than the original; therefore, the numbers in Fig. 1 should be interpreted with caution for that variant. The infectivity of the Delta variant is associated with a twofold increase in the basic reproduction number (WHO suggested RT = 2.5 for the original SARS-CoV-2). Therefore, the effective reproduction number of the Delta variant might be around RT = 5, depending on the practical setting. For such a high value of RT, Fig. 1 suggests that an extremely high sensitivity, close to 1 (i.e., greater than that shown in the black cells) will be required to contain the Delta variant. This implies that a one-off application of the standard PCR test, with a sensitivity of 85%, will by no means be effective in containing the spread of the Delta variant.

**Solution to Question 2**

One of the answers to the question of how a diagnostic test can become more accurate is a system approach that uses “cascade judgment” for multiple testing managed by a combination of repeated tests. Conducting a diagnostic test with a dichotomous outcome, e.g., “positive” or “negative,” is regarded as a Bernoulli trial in which the likelihood of each possible test outcome does not change from trial to trial.

The overall sensitivity of testing can be improved by independent repeated testing. For example, consider triplicate testing in which three rapid antigen tests, assumed to have a sensitivity of 75% (i.e., about 10% lower than that of the PCR test), are separately applied, and if all three outcomes are “negative,” then the final test result is adjudged to be negative. A PCR test with higher sensitivity is conventionally preferred to an antigen test in repeated tests conducted periodically, for example, every 2 or 3 days. Differently from the conventional repeated tests, the outcomes in a triplet test system must be determined on site in a short time because a combination of the three test outcomes determines the final result. Therefore, a rapid antigen test is, due to its rapidity, more practical in this context than a PCR test.

Assuming that the tests are fair and independent, for any infected individual, the probability of a “negative” result in each test is 0.25 (=1−sensitivity of 75%). Thus, the probability of a “negative” outcome in all three trials for an infected individual is given by 0.25 × 0.25 × 0.25 = 0.0156 regarded as an independent trial. If the triple test system defines the judgment of “negative” only as three negative outcomes [i.e., if at least one outcome of the three tests is positive, the judgment is “positive”], we can minimize the likelihood of false negatives with a system false negative rate estimated to be 1.56% (0.0156), which is equivalent to a sensitivity of 98.4% (0.984 = 1 − 0.0156). With this level of sensitivity, the area with gray cells and some black cells in Fig. 1 will be covered.

Based on the above consideration to enhance the system accuracy, Fig. 2 illustrates the cascade judgment for a system of triple testing that employs the parallel application of three rapid antigen tests. A controversial issue of this cascade system is how the final judgment should be made based on the combination of positive and negative results that may be obtained. The judgment would be made with a primary focus on either reducing false positives or reducing false negatives. In the case of the former approach, individuals testing positive in all three tests would be judged as positive, with the rest being judged as
negative. In contrast, in the latter case, individuals testing negative in all three tests would be judged as negative, with the rest being judged as positive. If the diagnostic test is intended to contain the virus, it naturally needs to focus on false negatives, as illustrated in the boundary condition for contagion containment.

It should be noted that false positives will inevitably increase if the focus is on minimizing false negatives. To handle such a trade-off between false positives and false negatives, as shown in the rightmost column in Fig. 2, individuals testing positive in all repeated sequences are finally judged “positive” to minimize the false positives, while keeping false negatives to a minimum. Those with a mixture of positive and negative outcomes will be assigned to an intermediate group of “judgment withheld or quasi-positive” as a third category, neither “positive” nor “negative.”

Practical recommendations on how to manage individuals classified into each of the three categories are as follows: those in the “positive” category should be hospitalized under the national government’s order that regards COVID-19 as a designated infectious disease, with those showing no symptoms being allowed to stay at a designated non-medical facility or at home for follow-up. In this category, the false positives will be drastically reduced because of the extremely high specificity of the cascade system. For those in the intermediate category of “judgment withheld or quasi-positive,” it would be appropriate to hospitalize those who show moderate or more severe symptoms and to allow those with mild or no symptoms to choose between staying at a non-medical facility or at home for follow-up. However, room should be left for the relevant authorities to exercise some flexibility in the judgment in accordance with the stress level of each regional medical system. Those who are judged as being in the third category of “negative” are, of course, not required to undergo any medical follow-up, but should be mindful of preventive measures such as wearing masks since the risk of a false-negative result is not zero.

**Discussion**

Amid the continuing anxiety in many countries about the resurgence of COVID-19 as of July 2021, and in anticipation of the expected arrival of a full-scale wave of Delta-variant coronavirus, this study analyzed the theoretical foundation of diagnostic testing to win the fight against this virus and proved the validity of a formula related to test accuracy.

Using a horizontal transmission model, the “boundary condition for contagion containment” was developed in the form of an inequality: $S_{n} > (R_{t} - R_{k}) / (1 - R_{k})$. This boundary condition indicates the lower limit of sensitivity for a diagnostic test to stop the virus spreading when we consider a horizontal transmission model with widespread transmission to the $k$-th level and the effective reproduction number, $R_{t}$, as an indicator of infectivity. This inequality, despite being a theoretical construct, can be regarded as a “rule of thumb” formula because estimates of $k$ and $R_{t}$ will inevitably be set by empirical analysis or hypothesis.

Some example computations using the formula showed that the sensitivity of 85% assumed for one-off PCR testing would be too low to satisfy the minimum sensitivity required to contain highly contagious infections such as the Delta variant of SARS-CoV-2. Although it is commonly believed that the more tests that are conducted, the more effectively the infection can be contained, the
boundary condition theoretically suggests that such a belief in large-scale testing is not necessarily supported by theory and that the control of infection is an issue of test accuracy, not test quantity.

As a real-world strategy, a cascade judgment system was proposed using multiple tests to enhance the sensitivity of COVID-19 tests to the minimum level at which spread of the virus will be contained. For multiple-test trials, however, it is assumed that the individual tests are mutually independent in statistical terms. Therefore, theoretical questions remain to be answered for the case when individual tests are not mutually independent. If we assume that there is some mutual dependence among individual repeated tests, more complicated analyses must be conducted with more complex modeling.

With the current horizontal transmission model, the minimal test sensitivity necessary for infection control was found to be independent of D\(^+\). However, if we go with a more complex transmission model, D\(^-\) might remain as one of the variables for analysis. In that case, the formulae shown in the Appendix will be useful for estimating D\(^-\) and its related values.

Some aspects remain as work for future studies with more complex modeling. The first issue is to optimize the interval between one test and the next. One of the concerns in repeated testing is that if the inter-test interval is too long it will allow false-negative individuals to transmit infection to others. In the context of the horizontal transmission model used in this study, computing analyses with deeper levels of k would partially address such a concern. However, to find an optimal solution for the test interval is not simply a question of the transmission level, k, since the test sensitivity changes over time. Consequently, in more sophisticated modeling, each clinical stage of individual infected patients must be considered to comply to the time dependency of test sensitivity.

Another limitation of the current transmission model is that the proportion of immunity acquired by vaccination is not considered among those in close contact. As a more realistic model, vaccinated cases must be incorporated for further investigation.

One of the lessons learned in the fight against COVID-19 is that our scientific knowledge and practical resources are still limited in their ability to control the pandemic. To overcome these limitations, further theoretical development in terms of clinical epidemiology and medical technology assessment must be pursued to make the best evidence-based decisions on public policy, thereby avoiding an ad hoc response to the pandemic.

**Appendix: Estimation of D\(^+\) when the prevalence of infection is unknown**

When the numbers of total cases tested and positive cases are given in a public survey in the context of **Appendix Fig. 1**, it is quite difficult to know the number of truly infected cases because the prevalence (i.e., D\(^+\)/N) is unknown due to the unknown D\(^-\). To solve this problem, we revisit the 2 × 2 table for the truth of “disease present” or “disease not present” versus a test result of “positive” or “negative.” Let T\(^\prime\), T\(^\prime\prime\), D\(^+\), and D\(^-\) be, respectively, the numbers of cases testing positive and negative, and the numbers of cases truly infected and not infected. The total number of cases tested is noted as N (i.e., N=T\(^\prime\)+T\(^\prime\prime\) or N=D\(^+\)+D\(^-\)). It is assumed that the test sensitivity, S\(_n\), and the test specificity, S\(_p\), are both known.

Considering the 2 × 2 table, we can establish the following equations to satisfy the consistency of the data in the table, regarding D\(^+\) and D\(^-\) as two unknown variables:

\[
S_n(D^+ + (1 - S_p)D^-) = T^+ \\
(1 - S_n)D^+ + S_pD^- = T^- \\
\]

This system of equations can be mathematically solved, leading to the following solutions:

\[
D^+ = (NS_p - N + T^\prime\prime) / (S_n + S_p - 1) \\
D^- = (NS_n - T^\prime) / (S_n + S_p - 1) \\
\]

These solutions for D\(^+\) and D\(^-\) mean that the numbers of truly infected and not infected cases can be estimated, given S\(_n\), S\(_p\), N, and T\(^\prime\), even though the prevalence (i.e., D\(^+\)/N) is unknown. Once the estimates of D\(^+\) and D\(^-\) are known, the prevalence of infected cases is obtained by the estimate of D\(^+\) divided by N, i.e., the total number of cases tested. Subsequently, the key numbers in the four categories of the two-by-two table are estimated according to the definition of each characteristic, using the estimate of D\(^+\) or D\(^-\) as follows:

\[
\text{True positives, TP} = S_nD^+ = S_n(NS_p - N + T^\prime\prime) / (S_n + S_p - 1) \\
\text{False positives, FP} = (1 - S_p)D^- = (1 - S_p)(NS_n - T^\prime) / (S_n + S_p - 1) \\
\text{False negatives, FN} = (1 - S_n)D^+ = (1 - S_n)(NS_p - N + T^\prime\prime) / (S_n + S_p - 1) \\
\text{True negatives, TN} = S_pD^- = S_p(NS_n - T^\prime) / (S_n + S_p - 1) \\
\]

Consequently, using the TP and TN values determined above, two conditional probabilities, namely, the positive predictive value (PPV) and the negative predictive value (NPV), can be estimated as follows:

\[
\text{PPV} = TP / T^\prime = S_n(NS_p - N + T^\prime\prime) / [(S_n + S_p - 1)T^\prime] \\
\text{NPV} = TN / T^\prime = S_p(NS_n - T^\prime) / [(S_n + S_p - 1)(N - T^\prime\prime)] \\
\]

**Appendix Fig. 1** shows the results of an example calculation that used these equations for the test statistics in Japan as of July 8, 2021, with N = 16,846,353, T\(^\prime\)=811,712, S\(_n\)=85%, and S\(_p\)=99.99%. From this table, we find that the total cumulative estimate of errors is 144,552 (=1589 false positives plus 142,963 false negatives), which amounts
to nearly 17.8% of the cumulative number of individuals who tested positive (811,712 cases). This suggests that such a substantial number of errors should have been taken into consideration to make informed decisions on public health. The predictive values are both high: PPV = 99.8% due to the low infection rate of 5.66% (=D/N = 953,086 / 16,846,353) and NPV = 99.1% due to the high test S of 99.99%. However, it is conjectured that the 142,963 false negatives overlooked by PCR tests might be one of the factors making it difficult to control the epidemic in Japan.

It should be noted that the infection rate of 5.66% estimated in Appendix Fig. 1 is just a test-based prevalence rate, which is subject to selection bias such as cluster testing. This value does not necessarily reflect the prevalence of infection in the community, which likely includes asymptomatic cases.

Acknowledgments

The author is deeply grateful to Professor Hideyuki Saya and Professor Naoki Hasegawa, who moderated my lecture with an invitation to the 100th Symposium, The Keio Medical Society, on November 21, 2020, which became the origin of this article. The author is also honored to be invited to submit this article to The Keio Journal of Medicine by Dean Masayuki Amagai and Editor-in-Chief and Associate Professor Kenjiro Kosaki, Keio University School of Medicine.

Conflicts of Interest

The author has received grants from CRECON Medical Assessment Inc., CMIC Holdings Co. Ltd., Takeda Pharmaceutical Co. Ltd., and Japan Becton Dickinson and Co., outside of the submitted work.

References

1. Coronavirus Resource Center: Johns Hopkins University. https://coronavirus.jhu.edu/ (Accessed 2021–07-12).
2. Day M: Covid-19: Identifying and isolating asymptomatic people helped eliminate virus in Italian village. BMJ 2020; 368: m1165. DOI:10.1136/bmj.m1165, Harvard University. Testing & Tracing. https://www.harvard.edu/coronavirus/testing-tracing/ (Accessed 2021–07-12)
3. WHO: Covid-19 case definitions. https://www.who.int/publications/i/item/WHO-2019-nCoV-Surveillance_Case_Definition-2020.2 (Accessed 2021–07-12).
4. Straus SE, Glasziou P, Richardson WS, Haynes RB. Evidence-Based Medicine: How to Practice and Teach EBMM. 5th edn, Elsevier, 2018.
5. West CP, Montori VM, Sampoorthkumar P: COVID-19 testing: the threat of false-negative results. Mayo Clin Proc 2020; 95: 1127–1129. PMID:32376102, DOI:10.1016/j.mayocp.2020.04.004
6. Arevalo-Rodriguez I, Butrago-Garcia D, Simancas-Racines D, Zambrano-Achig P, Del Campo R, Ciapponi A, Sued O, Martinez-Garcia L, Rutjes AW, Low N, Bossuyt PM, Perez-Molina JA, Zamora J: False-negative results of initial RT-PCR assays for COVID-19: a systematic review. PLoS One 2020; 15: e0242958. PMID:, DOI:10.1101/2020.04.16.20066787
7. Padhye NS: Reconstructed diagnostic sensitivity and specificity of the RT-PCR test for COVID-19. https://www.medrxiv.org/content/10.1101/2020.04.24.20078949v2 (Accessed 2021–07-12).
8. Office for National Statistics: UK: COVID-19 Infection Survey: methods and further information. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/methodologies/covid19infectionsurveypilotmethodsdurtherinformation/text-sensitivity-and-specificity (Accessed 2021–07-12).
9. Watson J, Whiting PF, Brush JE: Interpreting a COVID-19 test result. BMJ 2020; 369: m1808. DOI:10.1136/bmj.m1808
10. Ontario Public Health: COVID-19 Laboratory Testing Q&As. https://www.publichealthontario.ca/-/media/documents/lab/covid-19-lab-testing-faq.pdf?la=en&sc_lang=en&hash=F2F1C5303919FC691AB6B062AF922E42 (Accessed 2021–07-12).
11. Nishiura H, Inaba H: Prediction of infectious disease outbreak with particular emphasis on the statistical issues using transmission model. Proc Inst Math Stat Math 2006; 54: 461–480 (in Japanese; abstract in English).
12. Nature. Delta coronavirus variant: Scientists brace for impact. News 22 June 2021. https://www.nature.com/articles/d41586-021-01696-3 (Accessed 2021–07-12).
## Appendix Fig. 1

Example of the breakdown of test results in a 2 × 2 table. The statistics $N = 16,846,353$ and $T^+ = 811,712$ were reported by the Japanese Ministry of Health Labour and Welfare as of July 8, 2021. Test sensitivity was assumed to be $S_n = 0.85$ and test specificity $S_p = 0.9999$.

|                      | COVID (+) | COVID (-) |
|----------------------|-----------|-----------|
| **Test result**      |           |           |
| +                    | True positives | 810,123 | False positives | 1,589 |
|                      | T^+       | 811,712   |
| -                    | False negatives | 142,963 | True negatives | 15,891,678 |
|                      | T^-       | 16,034,641|
|                      | D^+       | 953,086   |
|                      | D^-       | 15,893,267|
|                      | N         | 16,846,353|

The table shows the breakdown of test results into true positives, false positives, false negatives, and true negatives, with the total number of tests $N = 16,846,353$ and the total number of positive tests $T^+ = 811,712$.