Comparative analyses of all FDA EUA-approved rapid antigen tests and RT-PCR for COVID-19 quarantine and surveillance-based isolation

Chad R. Wells¹, Abhishek Pandey¹, Seyed M. Moghadas², Burton H. Singer³, Gary Krieger⁴,⁵, Richard J.L. Herron⁶, David E. Turner⁷, Justin P. Abshire⁸, Kimberly M. Phillips⁹, A. Michael Donoghue¹⁰, Alison P. Galvani¹¹, and Jeffrey P. Townsend¹¹,¹²,¹³,¹⁴*

¹Center for Infectious Disease Modeling and Analysis (CIDMA), Yale School of Public Health, New Haven, Connecticut, 06520, USA
²Agent-Based Modelling Laboratory, York University, Toronto, Ontario, Canada
³Emerging Pathogens Institute, University of Florida, P.O. Box 100009, Gainesville, FL 32610, USA
⁴NewFields E&E, Boulder, Colorado 80301, USA
⁵Skaggs School of Pharmacy and Pharmaceutical Science, University of Colorado Anschutz Medical Campus, Aurora, Colorado 80045, USA
⁶BP Plc, 1 ST James’s Square, London, England SW1Y 4PD, UK
⁷BP America Inc, 501 Westlake Park Boulevard, Houston, Texas 77079, USA
⁸HSE Specialties, BHP Petroleum, 1500 Post Oak Blvd, Houston, Texas 77056, USA
⁹BHP Petroleum, 1500 Post Oak Blvd, Houston, Texas 77056, USA
¹⁰Group HSE, BHP Group Ltd, 171 Collins Street, Melbourne, Victoria 3000, Australia
¹¹Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06525, USA
¹²Department of Biostatistics, Yale School of Public Health, New Haven, Connecticut 06510, USA
¹³Program in Computational Biology and Bioinformatics, Yale University, New Haven, Connecticut 06511, USA
¹⁴Program in Microbiology, Yale University, New Haven, Connecticut 06511, USA

*Corresponding author: Jeffrey.Townsend@yale.edu
Abstract

Rapid antigen (RA) tests are being increasingly employed for detection of COVID-19 infections in implementations of quarantine and surveillance. We conducted a comparative analysis of quarantine durations, testing frequencies, and false-positive rates for all of the 18 RA tests for which emergency use authorization (EUA) has been given by the FDA and a nasopharyngeal RT-PCR test. For each test, we employed a mathematical model of imminent infections to calculate the effective reproductive number in the context of the test used for quarantine or serial testing strategy. We informed the model with data on test specificity and temporal diagnostic sensitivity, convolved with a data-driven profile of COVID-19 infectiousness across the disease time course. Our results demonstrate that the relative effectiveness of RA and RT-PCR tests in reducing post-quarantine transmission depends on the duration of quarantine and the turnaround time of testing results. When quarantines are shorter than five days, our results suggest that an RA test on entry to and on exit from quarantine would reduce onward transmission more than a single RT-PCR test conducted upon exit. Conducting surveillance via serial RT-PCR testing with a 24-h turnaround time, the testing frequency paired with isolation of positives that is required to suppress the effective reproduction number ($R_E$) below one was found to require a minimum frequency of every six days. RA tests reduce $R_E$ below one when conducted at a minimum frequency that ranges from every six days to every eight days—depending on the type of RA test—with a median of seven days. Our analysis also highlights that the risk of onward transmission during serial testing increases with the delay in obtaining the results. False-positives were found to be more frequent with RA tests, an issue that could be mitigated with clinical and RT-PCR follow-up. Accounting for the specific diagnostic traits of RA tests, they are an important component of the tool set for policy decision-making, and can serve as a viable alternative to RT-PCR in efforts to control the spread of disease.
Introduction

Testing for SARS-CoV-2 infections has played a central role in combating the COVID-19 pandemic. In countries with high vaccination coverage, testing will continue to be important for screening and surveillance, enabling timely isolation of individuals with a positive test and decreasing the risk of further disease spread. For instance, previous studies have shown that a 14-day quarantine with no testing can safely be shortened to a seven-day quarantine if a nasopharyngeal RT-PCR test is conducted on exit from the quarantine, a reduction notified for widespread implementation by the Centers for Disease Control and Prevention (CDC). Complementary analyses have also been conducted to evaluate the optimal frequency of RT-PCR testing in at-risk populations to minimize the probability of an outbreak.

Over the course of the pandemic, the diversity of SARS-CoV-2 tests with regulatory approval has increased immensely. Although the RT-PCR tests remain the gold standard for diagnosis, rapid antigen (RA) tests have aided in scaling up testing capacities worldwide. The fast turnaround time, widespread availability and much lower cost makes RA tests an attractive choice for workplace screening, especially in remote shift team work or offshore environments. The antigen test kits require minimal training and expertise to administer and do not require significant ongoing equipment maintenance and calibration. As a result, many businesses and organizations are shifting to using these tests for screening employees instead of solely relying on the more costly and time-consuming RT-PCR. With access to a variety of RA tests, there could be pronounced changes in the reduction of onward transmission for quarantine and testing strategies that previously had been conducted using RT-PCR-based testing.

Since the emergence of the RA tests for SARS-CoV-2, there has been some dispute surrounding their utility in disease screening and control efforts. One aspect of concern with RA tests pertains to their higher rates of false positives and false negatives. As a tool for workplace screening or community surveillance, frequency of testing has often been indicated as a critical factor in attaining an effective reproduction number \(R_e\) below one. However, increasing the number of tests used in screening and surveillance elevates the probability of false positives. Although false-positive results do
not carry any epidemiological risks, they can lead to undesirable logistical and cost challenges. For example, in an offshore and or remote workplace setting, a false positive could entail medical evacuation via helicopter or other aviation platform. From a workplace risk-analysis perspective, such a false positive is nevertheless less disruptive to operations than false negatives resulting in transmission and a full-scale outbreak.

Since the onset of the COVID-19 pandemic, the large-scale extractive industry sector (i.e., oil, gas, and mining) has developed various strategies for identifying, assessing, controlling, and recovering from the impacts of COVID-19 on their employees and operations to manage their health, safety, and environmental risks. One of the critical control strategies identified early on was systematic quarantine and testing. Within the workforce, quarantine is a significant workforce stressor. To reduce quarantine duration while limiting the probability of an outbreak offshore, a strategy of quarantine and testing with serial offshore testing could be a feasible option. Therefore, data-driven studies that can inform deployment of a considerably short quarantine as well as optimal serial testing frequency for surveillance have significant salience for ensuring operational integrity for private and public sectors.

Over the course of the COVID-19 pandemic, there has been extensive analysis and evaluation of the optimal testing strategies for both RT-PCR and RA tests. The performance of serial RA testing in identifying cases has been evaluated in both the health care and university setting. Screening in these studies was directed towards asymptomatic and symptomatic individuals, and were often conducted during an active COVID-19 outbreak or in a tertiary hospital setting. Outside these settings, screening asymptomatic individuals without known or suspected exposure to SARS-CoV-2 has been proposed and discussed in the literature. For example, remote industrial settings are associated with distinct challenges that differ from either the healthcare or university environment because exposure occurs predominantly within the isolated population but there are logistical constraints to imposing isolation or offering treatment. Therefore, there is currently minimal published evidence of the effectiveness of large-scale serial RA testing studies in mitigating outbreaks within an industrial setting.
Moreover, most of these analyses do not proceed to quantify the degree to which their use would
suppress individual-level transmission in applications of quarantine\(^2\) or serial surveillance testing with
isolation of positives\(^10,11,13,15,16\), and most address a generalized or single RA test\(^5,6,9,11,12,16,18,32,33\) or focus
only on RT-PCR\(^4,8,10,13,15,17\). In contrast, there have been multiple RA tests given regulatory Emergency
Use Authorization (EUA)\(^39\), each with distinct temporal diagnostic sensitivity that influences optimal
testing strategy. Here we construct diagnostic sensitivity curves for 18 RA tests that have received EUA
using data on percent positive agreement with a RT-PCR test and temporal diagnostic sensitivity of a
RT-PCR test. To determine when RA tests can serve as a suitable alternative to the more costly and
laborious RT-PCR tests, we perform the first comprehensive analysis of all 18 RA tests, calculating for
each test i) the probability of post-quarantine transmission for quarantine durations from one to 14 days
with testing on exit or both entry and exit; ii) the extent of onward transmission for testing every day to
every two weeks; and iii) the probability of false-positives during serial testing.

**Results**

We constructed the diagnostic sensitivity curves of 18 RA tests by applying linear logistic
regression and a maximum likelihood approach to data on their percent positive agreement with an
RT-PCR assay, and the diagnostic sensitivity of a RT-PCR test. RA tests were classified into three
categories in terms of their percent positive agreement over time: (i) stable (CareStart anterior nasal
swab, Clip COVID, Liaison anterior nasal and nasopharyngeal swab, Omnia, Vitros, and Sofia),
(ii) gradually declining (LumiraDx anterior nasal swab, SCoV-2, Status COVID+Flu, Simoa, and
BinaxNOW), and (iii) rapidly declining (CareStart nasopharyngeal swab, LumiraDx nasopharyngeal
swab, BD Veritor, Celltrion DiaTrust, Sofia 2 Flu+SARS, and Ellume; Figs. S1–S15; Table S2). Among
the five most commonly used RA tests (i.e., LumiraDx, Sofia, BinaxNOW, BD Veritor, and CareStart)\(^40\),
there were negligible differences in the estimated diagnostic sensitivity of LumiraDx and Sofia tests over
time compared to the sensitivity of RT-PCR tests (Fig. 1). BinaxNOW, BD Veritor, and CareStart
exhibited noticeably lower diagnostic sensitivity than the RT-PCR tests. We found that the diagnostic sensitivity of BD Veritor was higher than BinaxNow and CareStart around the time of symptom onset, but BinaxNow and CareStart exhibit a greater probability of detection earlier and later in the disease time course.

Figure 1: Diagnostic sensitivity of RT-PCR and rapid antigen tests. Specifying an incubation period of 8.29 days (to symptom onset, vertical dashed line), the estimated diagnostic sensitivity of an RT-PCR test based on data from Hellewell et al.\textsuperscript{11} (black stars) and rapid antigen test LumiraDx (blue squares); Sofia (green diamonds); BinaxNOW (yellow triangles); BD Veritor (red circles); and CareStart (purple hexagram) based on percent positive agreement data from Emergency Use Authorization documentation over the course of infection.

Quarantine and testing strategies

Specifying an infectivity profile with an incubation period of 8.29 days\textsuperscript{41} and a basic reproduction number of 2.5\textsuperscript{42}, where 30.8% of infections are asymptomatic\textsuperscript{43}, we quantified the effectiveness of quarantine and testing strategies in reducing post-quarantine transmission by calculating the probability of post-quarantine transmission for individuals entering quarantine randomly in the absence of symptoms, with those who exhibit symptoms isolating upon symptom onset\textsuperscript{4}. For quarantine
durations varying from one to 14 days, we compared the probability of post-quarantine transmission when performing a single RT-PCR test on exit to a RA test on exit or RA tests on both entry and exit.

As a baseline reference for post-quarantine transmission, the CDC has recommended a seven-day quarantine with a RT-PCR test conducted within 48 h of quarantine exit. In our analysis, the RT-PCR test was specified to be conducted 24 h before exit from quarantine, whereas no lag was specified for the RA tests. We found that a probability of post-quarantine transmission equivalent to the CDC-suggested quarantine and testing strategy could be achieved using RA tests with only modest changes to the quarantine duration. In particular, conducting a RA test on entry and exit can shorten quarantine by a day with LumiraDx but does not change quarantine duration for Sofia. For BinaxNow, CareStart and BD Veritor RA tests, the duration of quarantine would require a three- to four-day extension (i.e. to a 10- or 11-day quarantine). Equivalent durations to a seven-day quarantine with an RT-PCR test conducted 24 h before exit ranged from six to 11 days—with a median between six and seven days—when an antigen test was conducted on entry and on exit from quarantine (Table S3, Fig. S1B–S15B).

The zero-day delay in turnaround time for RA test results has the potential to offset their lower sensitivity compared to RT-PCR with a 24-h delay in obtaining results—especially in the context of short quarantine duration. Conducting a RA test on exit for quarantines with a duration of five days or shorter was more effective than an RT-PCR on exit with a 24-h delay (Fig. 2C, Figs. S1B–S15B). For over half of the RA tests, conducting a RA test on both entry and exit produced a lower probability of post-quarantine transmission than RT-PCR on exit for quarantines of 14 days or shorter (Fig. 2D, Figs. S1B–S15B).
Figure 2. Probability of post-quarantine transmission. Specifying a negative-binomial distribution for expected post-quarantine transmission, 30.8% of infections being asymptomatic, a 24-h delay in obtaining RT-PCR test results, a negligible delay in receiving rapid antigen test results, an incubation period of 8.29 days, self-isolation upon symptom onset, and the diagnostic sensitivity curve for RT-PCR based on data from Hellewell et al\textsuperscript{11}, the probability of post-quarantine transmission when conducting an RT-PCR test only on exit (solid line; black circles) and the rapid antigen tests (dashed lines) LumiraDx (blue squares); Sofia (green diamonds); BinaxNOW (yellow triangles); and BD Veritor (red circles) performed (A) on exit and (B) on both entry and exit; and the fraction of the 18 rapid antigen tests whose use conferred a lower probability of post-quarantine transmission than did an RT-PCR test conducted 24 h before exit from quarantine, when the rapid antigen test was conducted (C) on exit and (D) on both entry and exit.

Serial testing

Our calculation of $R_E$ with testing frequencies ranging from every day to once every two weeks (incorporating a 24-h delay for obtaining RT-PCR test results) and isolating positives revealed that an RT-PCR test every six days was required to maintain $R_E < 1$ (Fig. 3A). The frequency of RA testing required to maintain $R_E < 1$ ranged from every six to every eight days (Fig. 3A). Evaluating the importance of turnaround time for RT-PCR tests, we found that any significant delay has a substantial impact on the required frequency of serial testing. With a negligible delay in turnaround time, RT-PCR tests can be performed just once every eight days. A two-day delay in obtaining RT-PCR results increases the frequency of testing to every three days, making RA tests a more practical tool than RT-PCR tests for serial testing without consideration of false positive rates. Turnaround times for RT-PCR are highly
variable both in the USA and internationally with a 12–24 hour delay common in the USA; however, delays of up to five days have been experienced by some companies in some non-USA locations. If the turnaround time of RT-PCR results exceeds three days, even a daily testing strategy is insufficient to achieve $R_E < 1$ (Fig. 3B).

For all testing frequencies examined, we found that RA tests would have a lower effective reproduction number than an RT-PCR test with a 24-h delay (Fig. 3C). All RA tests were found to maintain $R_E < 1$ when conducted with frequency of at least every six days. However, no RA test could achieve $R_E < 1$ when the testing frequency exceeded nine days (Fig. 3D, Table S3, Fig. S19).

Figure 3. Effective reproduction number in the context of different frequencies of serial testing. Specifying 30.8% of infections as asymptomatic, a one-day delay in receiving RT-PCR and rapid antigen test results, an incubation period of 8.29 days, self-isolation upon symptom onset, and a RT-PCR diagnostic sensitivity curve informed by data from Hellewell et al.11, (A) the expected effective reproduction number with serial testing using an RT-PCR test (black) and the rapid antigen tests LumiraDx (blue); Sofia (green); BinaxNOW (yellow); and BD Veritor (red), and (B) for serial testing every day to every 14 days with a zero- to five-day delay (black to light gray) in obtaining the results for an RT-PCR test and isolation of positives in comparison to no testing (solid gray line). (C) The fraction of rapid antigen tests of the 18 tests that had a lower effective reproduction number than a RT-PCR test with a 24-h delay, and (D) that had an effective reproduction number ($R_E$) below one for testing frequencies ranging from every day to every two weeks and isolating positives.

False positives for serial testing
As the frequency of serial testing increases, an additional consideration to the reduction of transmission is the concomitant probability of obtaining false positives. For each type of the test, we calculated the probability of obtaining at least one false positive over a two week period based on the specificity of the test and the frequency of testing (Methods). There is an inverse relationship between the effective reproduction number and the probability of obtaining at least one false positive (Fig. 4 and Fig. S20; \( r = -0.436 \) and \( P < 0.001 \)). For a specified \( R_e \) as a result of the testing strategy, we found that RT-PCR tests yield a lower probability of false-positive results than eight of the 18 RA tests (Fig. S20)—including BD Veritor, BinaxNOW and LumiraDx of the five most commonly used RA tests (Fig. 4). At the frequency of testing required to suppress long-term transmission, we estimated that the probability of one or more false-positives over a two-week period ranged from 0.002–0.063 (Table S3). Eight of the 18 RA tests exhibited a range of elevated false-positive rates (0.013–0.063), whereas the remaining 10 had similar false-positive probabilities to RT-PCR (Fig. S20, Tables S3–S4).

![Figure 4. The effective reproduction number, and risk of one or more false positives, for varying frequencies of serial testing with RT-PCR and rapid antigen tests.](https://doi.org/10.1101/2021.08.23.21262499)
with no delay to five-day delay (black stars gradient) in obtaining test results and the rapid antigen test LumiraDx (blue squares); Sofia (green diamonds); BinaxNOW (yellow triangles); BD Veritor (red circles); and CareStart (purple hexagrams) for testing everyday to every 14 days and isolating positives (small dots: longer time between tests; larger dots: shorter time between tests) and the corresponding probability of at least one false positive over a two-week period (x axis).

Comparing percent positive agreement data sets

The percent positive agreement data used to inform the diagnostic sensitivity for the RA tests were obtained in a more clinical trial setting for companies seeking EUA from the U.S. Food and Drug Administration (i.e., internal) than when they may be applied in a real-world setting by independent investigators (i.e., external). Three of the RA tests (CareStart, Sofia, and BinaxNOW) had data available from studies evaluating their abilities to detect infections within a community. With the availability of these internal and external datasets, we conducted a comparative analysis of these datasets.

Around the time of symptom onset, the diagnostic sensitivity informed by the internal dataset for Sofia and CareStart was greater than that inferred from the external dataset (Fig. S16–S17). In contrast, for BinaxNOW, the diagnostic sensitivity estimated from the external dataset was always greater than that determined from the internal dataset (Fig. S18). We found that the required quarantine duration was at most three days longer under the external dataset as compared to the internal for Sofia, whereas the required quarantine duration under the BinaxNOW external dataset was three days shorter than that for internal dataset, and there was no change for the CareStart test (Table S3). The time between tests was extended by at most two days for the external dataset as compared to the internal for all three tests (Table S3).

Scenario analysis

Compared to the RT-PCR diagnostic sensitivity curve informed by the Hellewell et al. 11 data, an alternative temporal sensitivity curve from Wells et al. 4 yields a higher probability of detecting infection over a longer duration (Fig. S21 vs Fig. 1). Under this alternative diagnostic sensitivity curve, a RT-PCR test on exit from quarantines six days or longer led to a lower probability of post-quarantine transmission than the majority of RA tests (Fig. S22). Compared to a seven-day quarantine with an RT-PCR test
conducted 24 h before exit, the median equivalent quarantine duration with an RA test on both entry and exit was eight days (Table S3) with the BD Veritor test requiring the longest quarantine (12 days). Quarantine durations with testing on entry and exit equivalent to a seven-day quarantine with an RT-PCR test on exit ranged from one day shorter to one day longer (with a median of one day longer) when using the alternative RT-PCR diagnostic sensitivity curve compared to the durations estimated in the baseline analysis.

For the alternative diagnostic sensitivity curve, the minimum testing frequency for RT-PCR test with a 24-h delay in turnaround time of results was every nine days, compared to every six days in our baseline analysis (Fig. 3 vs Fig. S23). The minimum testing frequency for the 18 RA tests ranged from every eight to every 11 days, with a median of every 10 days (Table S3). For the Wells et al. diagnostic sensitivity curve, only 13 of the 18 RA tests could outperform RT-PCR when testing as frequently as every day to every two weeks, compared to all 18 under the baseline diagnostic sensitivity (Fig. 3C vs Fig S23C). Under the alternative diagnostic sensitivity curve, the majority of RA tests can maintain $R_E < 1$ with testing frequency of every 10 days, compared to every seven days for the sensitivity inferred from the Hellewell et al. data used in the baseline analysis (Fig. 3D vs Fig S23D). Examining the minimum testing frequencies to maintain $R_E < 1$, we found that the maximum time interval between tests is extended two to three days for the alternative diagnostic sensitivity curve compared to our baseline results (Table S3).

**Empirical validation of serial RA testing within and subsequent to quarantine for suppression of transmission**

To examine the agreement between a RA test and RT-PCR in a cohort that was prescreened for COVID-19 symptoms, paired testing of BD Veritor and RT-PCR was conducted upon entry to quarantine, and on day three, and day four of a five-day quarantine (Table 1). From this paired testing within a prescreened cohort, the fraction of RA tests in agreement with negative RT-PCR was 2196/2199 (99.86%). This proportion was not significantly different from the fraction of RA tests in agreement with
negative RT-PCR from the EUA documentation for BD Veritor (212/213 = 99.53%; two-tailed Fisher Exact Test $P = 0.318$). The percent positive agreement between the BD Veritor RA test and RT-PCR test declined over the course of quarantine, starting at 25% on entry, to 20% on day three, to 0% agreement by day four (Table 1). The 16 false negatives all exhibited high cycle times ($\geq25$), indicating an early phase of asymptomatic disease and low levels of viral RNA for which RA tests are less sensitive than RT-PCR (Table S5). The three of the four true positive RA tests, there were low cycle times ($\leq19$) indicating a later phase of asymptomatic disease and high levels of viral RNA for which RA tests are frequently sensitive (Table S5).

The BD Veritor test was utilized offshore with testing being done either on days three, six, and nine of being offshore or days two, five, and eight of being offshore after undergoing a 24-h quarantine with an RT-PCR test on entry. A total of 1714 antigen tests were performed for the combined test cohorts. A total of three RT-PCR confirmed cases were identified within the first six days of being offshore (Tables 2 and 3), with all individuals being asymptomatic at the time of the positive RA test result.

Table 1. Paired BD Veritor and RT-PCR tests conducted, results, their agreement, and false positives

<table>
<thead>
<tr>
<th>Quarantine day after entry $^a$</th>
<th>Day 0 $^b$</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired tests conducted</td>
<td>818</td>
<td>726</td>
<td>675</td>
<td>2219</td>
</tr>
<tr>
<td>RT-PCR tests positive</td>
<td>12</td>
<td>5</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>RA tests positive</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>RA test false negative $^c$</td>
<td>9</td>
<td>4</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>(RA $^-$/RT-PCR $^+$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RA test false positives $^c$</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>(RA $^+$/RT-PCR $^-)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RA test true positive $^c$</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>(RA $^+$/RT-PCR $^+$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ BHP onshore quarantine testing, 22 November 2020 to 17 January 2021

$^b$ Test on entry into quarantine

$^c$ Assessed by comparison to RT-PCR.
Table 2. Serial offshore BD Veritor rapid antigen tests, positives, and false positives on days 3, 6, and 9

<table>
<thead>
<tr>
<th>Day after off-shore entry (^a)</th>
<th>Day 3</th>
<th>Day 6</th>
<th>Day 9</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA tests conducted</td>
<td>458</td>
<td>458</td>
<td>457</td>
<td>1373</td>
</tr>
<tr>
<td>RA tests positive</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>RA false positives (^b)</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

\(^a\) BP platform test data, 2 March 2021 to 22 May 2021

\(^b\) Subsequent onshore laboratory RT-PCR was negative

Table 3. Serial offshore BD Veritor rapid antigen tests, positives, and false positives on day 2, 5, and 8

<table>
<thead>
<tr>
<th>Day after offshore entry (^a)</th>
<th>Day 2</th>
<th>Day 5</th>
<th>Day 8</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA tests conducted</td>
<td>124</td>
<td>121</td>
<td>96</td>
<td>341</td>
</tr>
<tr>
<td>RA tests positive</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RA false positives</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^a\) BP platform test data, 5 March 2021 to 24 May 2021

Discussion

Our study demonstrates the utility of RT-PCR and RA tests in suppressing onward transmission of COVID-19 in the contexts of quarantine and serial testing. Examining 18 RA tests that have received EUA, we showed that for quarantines with a duration of five days or shorter, conducting RA tests on entry and exit provides a greater reduction in onward transmission than conducting a single RT-PCR test 24 h prior to quarantine exit. If turnaround time for RT-PCR testing results is delayed by one or more days\(^{45-47}\), RA tests outperform RT-PCR at reducing onward transmission via serial testing and case isolation due to their faster turnaround time. Accordingly, the frequency of serial RT-PCR testing required to suppress outbreaks increases substantially with the associated waiting times in obtaining results. While RA tests can reduce the effective reproduction number below one with equivalent or even in some cases
less-frequent testing than the RT-PCR tests, the probabilities of false-positives are markedly greater, and consideration should be devoted to appropriate follow-up testing of potential cases that are identified and isolated.

For quarantine strategies with an exit test, RT-PCR tests need to be conducted at least a day prior to the end of quarantine and RA tests could be used closer to the end of quarantines due to test turnaround times. As a result of increasing test sensitivity during short quarantines, the majority of RA tests conducted on exit outperformed RT-PCR tests for quarantine durations less than a week. Supplementing an exit RA test with an RA test upon entry to quarantine could allow RA tests to outperform a single RT-PCR test on exit for quarantine durations longer than a week as well. Specifically, more than 50% of RA tests considered produced a lower probability of post-quarantine transmission than an RT-PCR test on exit for quarantine durations up to 14 days, when conducted at both entry and exit from quarantine. Long quarantine durations have minimal practicality for entities who are trying to minimize interruptions to operations. Therefore, RA tests can be suitable alternatives to a single RT-PCR, especially for quarantine durations less than a week.

RT-PCR tests exhibit greater diagnostic sensitivity than RA tests, so it may seem counterintuitive that the RA testing could provide equivalent performance or outperform RT-PCR testing during serial testing. This counterintuitive result arises from the ability of RA tests to furnish results within minutes, leading to swift isolation of individuals prior to substantial transmission. Whereas obtaining results from RT-PCR sampling might be substantially longer than 24 h in high-prevalence or resource-limited settings. Utilizing rapid RT-PCR instrumentation would enable testing with shorter turnaround times. However, significant supply chain and maintenance hurdles for deploying large-scale point of care RT-PCR testing limit their wide applicability and remain infeasible for some remote settings. Unlike in quarantine, delays of testing results from surveillance sampling lead to continued transmission from infected individuals, requiring more frequent testing to prevent outbreaks (i.e., sustaining $R_E < 1$). More frequent testing would then give rise to logistical challenges and increased costs, and could further slow RT-PCR turnaround times, resulting in a positive-feedback loop with regard to delays that would be
extremely challenging to manage and that would exacerbate negative outcomes. Therefore, rapid turnaround times for testing are essential to the utility of a surveillance program intended to suppress the spread of disease.

Some of the challenges and costs arising from serial testing with RT-PCR can be mitigated using RA tests. However, these tests come with an increased chance of producing a false positive. From an operational standpoint, it could be critically problematic to send numerous or essential employees home as a result of false-positive results. The consequences of surveillance frequency in terms of the number of false-positives and extent of transmission are driven by several factors, including: (i) the sensitivity and specificity of the test, (ii) the number of people being tested, (iii) the frequency of testing, and (iv) the background prevalence of disease. The probability of at least one false-positive increases geometrically with the number of people tested, and similarly with the frequency of testing per individual. Given the higher specificity of RT-PCR, a follow-up RT-PCR test can be used for individuals testing positive with an RA test. The challenge to follow-up with RT-PCR testing is not just the increased logistical and economic costs: false-negatives can result in reintroduction into the surveilled population and onward transmission. Therefore, multiple follow-up tests are likely necessary to the investigation of conflicting test results. Accordingly, there are tradeoffs between risks of transmission, number of tests conducted, processing times, and underlying costs from false positives regardless of the test utilized in surveillance.

The effectiveness of quarantine and serial testing strategies is also dependent on the background effective reproduction number, which in turn is influenced by any ongoing disease interventions. We evaluated quarantine and serial testing strategies under the assumption that individuals self-isolate upon symptom onset, resulting in an $R_E$ of 1.6 when a basic reproduction number of 2.5 is specified. Organizations, institutions, and localities will likely specify and enforce distinct disease-control measures, resulting in idiosyncratic $R_E$. Higher $R_E$ or lower levels of self-isolation upon symptom onset will require more stringent testing strategies than indicated here. Nevertheless, the heterogeneity across the diagnostic sensitivity among various tests and their relative performance is independent of the basic reproduction
number. Therefore, we expect the qualitative differences in the performance of RT-PCR tests and RA tests to remain intact despite any differences in $R_E$.

The accuracy of the results presented here is strongly dependent on the quality of data collected. Substantial heterogeneity across RA tests exists in the span of time over which their sensitivity has been assessed. For example, tests such as BinaxNOW, LumiraDx, and Simoa provided percent positive agreement data extending to at least 11 days post symptom onset, whereas other tests such as Clip COVID, SCoV-2, and Sofia 2 Flu + SARS provided results extending only five days beyond symptom onset (Table S1). A limited number of samples at some time points can lead to uncertainty in the longitudinal diagnostic sensitivity of the RA tests. For example, the single data point at the last observed time for the Ellume RA test was the only one that did not agree with RT-PCR. This thresholded disagreement yielded a rapid decline in the logistic regression model for the percent positive agreement. Establishing a minimum sample size that is sufficient to provide statistical power for inference of sensitivity across time-points spanning the disease course would improve sensitivity measurements and facilitate refinement in testing frequencies and necessary quarantines to suppress transmission. As additional information about the diagnostics sensitivity of specific tests over the disease course emerges, more accurate inferences regarding quarantine durations and testing frequencies can be drawn.

Nonetheless, for the results presented here, sensitivity of the test during the late stage of disease matters far less than that during the early stage of disease (due to low infectiousness and limited transmission in the late stage), so the relative uncertainty regarding the late sensitivity of RA tests should only have a limited effect on our findings.

Improper sampling and analysis of samples of these RA tests can influence the rate of false-positives and false-negatives, and would therefore impact the required quarantine duration and testing frequency. Independently evaluated datasets on the testing of COVID-19 infections in the community were available for three of the RA tests, enabling a comparison to the datasets that were provided from the EUA documentation. In two of the three comparisons (Sofia and CareStart), diagnostic sensitivity in independent analyses was lower than in ideal conditions. The estimated quarantine duration
and testing frequency was equivalent or more stringent under the internal dataset compared to the external
dataset for two of the three tests (BinaxNOW and CareStart). These findings provide reassurance that
internal datasets do not deviate substantially from the real-world utility and effectiveness of RA tests for
disease control.

Further assurance of the real-world sensitivity of these tests (and therefore their effectiveness in the
context of quarantine and serial testing and isolation) awaits additional data becoming available reporting
application of RA tests in additional community settings. Our field test data demonstrated that with
prescreening and a 24-h onshore quarantine with an RT-PCR test on entry, serial offshore antigen testing
can be deployed while still maintaining an acceptable risk profile. By pre-screening individuals entering
quarantine, the prevalence of disease that is potentially departing to offshore platforms is reduced. The
extent of prospective secondary infections is diminished further with the combined 24 h quarantine and
RT-PCR testing. With RA testing being more effective for short quarantines, they could also be a suitable
alternative for the entry test for the test and fly protocol. Our analysis for the BD Veritor RA tests
highlights that testing every six days is sufficient to maintain $R_e < 1$. The frequency of testing was every
three days on the offshore platform, although only conducting testing for part of the worker’s time
offshore. Upon identification of a RT-PCR confirmed positive test, contact tracing was conducted to
mitigate additional transmission chains. Therefore, the data from the remote densely populated setting
illustrates that RA tests can be effectively integrated into the layered control measure approach for
mitigating disease outbreaks.

Our analyses of 18 COVID-19 RA tests currently approved for use in the United States provide a
comprehensive comparative assessment with reference to the gold-standard RT-PCR. Our results highlight
the scenarios in which RA tests would serve as a suitable and even beneficial alternative to RT-PCR tests,
incorporating the temporal dynamics that crucially describe their sensitivity and specificity, and
characterize the operationally important outcomes: for quarantine, post-quarantine transmission, and for
surveillance and isolation, suppression of the reproduction number. New tests should similarly be
evaluated by these key measurable outcomes, so that their utility can be compared for strategy
implementation. Such analyses provide crucial insights into the key trade-offs to decisions on the type of
tests to use for both congregate and community settings.

**Methods**

**Infectivity profile**

The infectivity profile was constructed using the R code from He et al.\textsuperscript{55} based on data of 77
contact transmission pairs, specifying a latent period of 2.9 days\textsuperscript{56} and an incubation period of 8.29 days
\textsuperscript{41} assuming exponential increase of virus from the time of infection to the end of the latent period\textsuperscript{4}.
Infected individuals were considered infectious no longer than 20 days after symptom onset\textsuperscript{57–59}.

**Diagnostic sensitivity of the RT-PCR test**

To construct a temporal diagnostic sensitivity curve, we fitted a log-Normal distribution to
nasopharyngeal RT-PCR testing percent-positivity data from Hellewell et al.\textsuperscript{11} using a
maximum-likelihood approach (Supplementary Methods).

**Diagnostic sensitivity of antigen tests**

We fitted a linear logit model to the discrete percent positive agreement data to estimate a
continuous percent positive agreement curve from the time of symptom onset for each RA test.
(Supplementary Methods). Percent positive agreement data were only available for post-symptom onset.
Therefore, we inferred the pre-symptomatic diagnostic sensitivity of the RA tests by constructing a
mapping between the inferred diagnostic sensitivity post symptom onset and the level of infectivity, then
applying that mapping to pre-symptomatic levels of infectivity (Supplementary Methods)\textsuperscript{4}. The
sensitivity of an RA test was calculated as the product of the percent positive agreement curve and the
diagnostic sensitivity of a RT-PCR test at the specified time.
For our baseline results, we examined the five most commonly used RA tests: LumiraDX, Sofia, BinaxNOW, BD Veritor, and CareStart. For the analysis of the LumiraDX and CareStart antigen test, we utilized the percent positive agreement data for the anterior nasal swab, as this method of sampling was used in gathering data for the BD Veritor, BinaxNOW and Sofia antigen tests. Furthermore, the anterior nasal sample can be obtained by a broad range of individuals with less specialized training compared to a nasopharyngeal sample. We also examined both the anterior nasal and nasopharyngeal swab for the LumiraDX and CareStart antigen test (Supplementary Material).

**Probability of post-quarantine transmission**

Accounting for substantial variance in transmission among COVID-19 cases, we specified that secondary cases were negative-binomially distributed:

\[ f(x|k, p) = \frac{\Gamma(k+x)}{\Gamma(k)\Gamma(x+1)} p^k (1 - p)^x , \]

with dispersion parameter \( k = 0.25 \) and \( p = k / (k + R) \)—such that the average number of secondary cases is equal to the expected post-quarantine transmission \( R \). The value for the dispersion parameter is relatively consistent with estimates from other studies. Accordingly, the probability of post-quarantine transmission was calculated by \( 1 - f(0|k, p) \).

**Probability of a false-positive**

With a specificity \( \zeta_i \) for test \( i \) (Table S4) and testing every \( f \) days, we computed the average probability that at least one false-positive occurred over a two-week period:

\[ P_f = \frac{1}{f} \sum_{j=1}^{f} 1 - \zeta_i^{\tau_j} , \]

where \( \tau_j \) is the number of tests to occur in the \( j \)th two-week period since the start of serial testing. For each testing frequency \( f \) (i.e., the time between two consecutive tests), we investigated the sequence of test times \( \{1, 1 + f, 1 + 2f, ..., 1 + 13f\} \) that comprises all the unique testing patterns possible over a two-week period to calculate the average probability.
Paired testing of oil workers in quarantine

The BD Veritor kit was selected by BP/BHP because at the time it was possible to secure a sufficient volume of the test kits and readers for a potential 6-month evaluation. With full informed consent from the offshore oil workers, onshore paired testing was conducted by laboratory-based RT-PCR and RA testing on entry into quarantine (day 0), day three, and day four. The study period spanned 22 November 2020 to 17 January 2021, and corresponded with an observed rise in community transmission across Texas and Louisiana, the overwhelming residence for the majority of the offshore platform workers. Before entry to quarantine, workers had to pass a pre-screening questionnaire that filtered symptomatic individuals and those with recent exposure. During quarantine, a positive RT-PCR led to removal of the individual from the quarantine environment, isolation for 10 days with medical follow-up. Afterward, workers could return to work after two negative RT-PCR tests.

Offshore serial rapid antigen testing

Implementation of a test-and-fly protocol began at the start of the staged vaccination rollout in the USA—2 March 2021 to 22 May 2021. Thus, none of the offshore workers were vaccinated nor likely would be vaccinated for several months. The test-and fly-strategy consisted of initial screening for symptoms, an RT-PCR test on entry, and an approximately 24-h quarantine. The exact quarantine duration was dependent on the commercial laboratory-based RT-PCR turnaround time: after receiving a negative RT-PCR test, the worker was permitted to enter the offshore work environment. An RT-PCR positive individual was removed from the quarantine environment, provided medical follow-up, and advised to conduct a 10-day home-based self-quarantine.

After entering the offshore work environment, the worker underwent serial antigen testing within their first nine of 14 or more days being offshore. Testing of a worker happened either on offshore days three, six, and nine; or on offshore days two, five, and eight. The two patterns were selected based on the results from our data-driven model for the degree of suppression achieved at each frequency of testing, by
the BD Veritor test kit (Fig. S4). All testing was conducted by the platform medic using the BD Veritor kit and reader. Any positive individual was isolated pending helicopter transfer (typically within 12–24 hours) to the established onshore medical facility, whereupon an RT-PCR nasal swab was obtained and sent to a commercial laboratory. A positive antigen test was considered to be a false positive if the follow-up RT-PCR was negative.

The use of the onshore and offshore testing data of the oil platform employees was approved by the Human Participants Review Sub-Committee, York University’s Ethics Review Board (Certificate Number: 2021-003).

Data availability

All data is referenced and provided within the main text or supplementary material.

Author contributions

GK, APG, and JPT conceived and designed the study with contributions from SMM, BS, CRW, and AP. CRW, AP, and JPT developed analytical approaches. CRW wrote computational code with contributions from AP. CRW executed analyses, with guidance from AP and JPT. All authors contributed to interpretation of results. CRW drafted the manuscript with contributions from AP and JPT. All authors contributed to revision of the manuscript and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Acknowledgements

This study was supported by the Notsew Orm Sands Foundation BHP and British Petroleum. SMM acknowledges the support from the Canadian Institutes of Health Research [OV4 – 170643, COVID-19]
Rapid Research] and the Natural Sciences and Engineering Research Council of Canada, Emerging Infectious Disease Modelling, MfPH grant.

References


34. McKay, S. L. et al. Performance Evaluation of Serial SARS-CoV-2 Rapid Antigen Testing During a


Validation of a rapid antigen test as a screening tool for SARS-CoV-2 infection in asymptomatic populations. Sensitivity, specificity and predictive values. EClinicalMedicine 100954 (2021).


672–675 (2020).


64. Adam, D. et al. Clustering and superspreading potential of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in Hong Kong. doi:10.21203/rs.3.rs-29548/v1.


66. Hébert-Dufresne, L., Althouse, B. M., Scarpino, S. V. & Allard, A. Beyond R0: Heterogeneity in