Antigen-based rapid diagnostic testing or alternatives for diagnosis of symptomatic COVID-19: A simulation-based net benefit analysis

Authors: Emily A Kendall,[1*] Nimalan Arinaminpathy,[2] Jilian A Sacks,[3] Yukari C Manabe,[1] Sabine Dittrich,[3] Samuel G Schumacher,[3#] David W Dowdy[4#]

1. Division of Infectious Diseases, Johns Hopkins University School of Medicine, Baltimore
2. Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore
3. MRC Centre for Global Infectious Disease Analysis, Imperial College, London
4. Foundation for Innovative New Diagnostics, Geneva

# Co-senior authors contributed equally

Abstract

Background: SARS-CoV-2 antigen-detection rapid diagnostic tests (Ag-RDT) offer the ability to diagnose COVID-19 rapidly and at low cost; however, lower sensitivity than nucleic acid amplification tests (NAAT) has limited adoption of Ag-RDT in clinical settings.

Methods: We compared Ag-RDT, NAAT, and clinical judgment alone for diagnosing COVID-19 among symptomatic patients. We considered two scenarios: a high-prevalence hospital setting with 24-hour NAAT turnaround, and a lower-prevalence outpatient setting with 3-day NAAT turnaround. We simulated transmission from cases and contacts and relationships between time, viral burden, transmission, and case detection. We used decision curve analysis to compare the net benefit of diagnostic approaches relying on Ag-RDT versus NAAT.

Results: Greater net benefit was achieved with Ag-RDT than NAAT in the outpatient setting, as long as NAAT turnaround time was longer than one day. NAAT was predicted to offer greater net benefit than Ag-RDT in the hospital setting, unless NAAT turnaround times exceeded 2 days. Findings were robust to data-consistent variation in Ag-RDT performance, empiric isolation practices, duration of symptoms, and other model parameters. Both tests provided greater benefit than management based on clinical judgment alone, unless the available interventions carried minimal harm and could be provided at full intensity to all patients in whom COVID-19 diagnosis was considered.

Conclusions: Ag-RDT may provide greater net benefit than NAAT for diagnosis of symptomatic COVID-19 in outpatient settings when NAAT turnaround times are longer than one day. NAAT is likely the optimal testing strategy for hospitalized patients, especially those with prolonged symptoms prior to admission.

Keywords: SARS-CoV-2, Diagnostic Tests, Symptomatic outpatients, Hospitals, Rapid antigen tests, Simulation, Decision curve analysis

Introduction

Accurate diagnosis of COVID-19 can guide clinical management; reduce transmission; and inform appropriate allocation of resources for isolation, contact tracing, and treatment. In many settings – including low- and middle-income countries (LMICs) as well as high-income countries with large outbreaks – efforts to diagnose COVID-19 using nucleic acid amplification tests (NAATs) frequently exceed capacity.[1–3]

Antigen-detection rapid diagnostic tests (Ag-RDTs) are less expensive than NAATs and can be performed in minutes without centralized laboratory infrastructure. Thus, they could facilitate higher volumes of testing and provide rapid results while relieving strains on laboratory capacity. Ag-RDTs are, however, less sensitive than NAATs. Some experts have argued that all SARS-CoV-2 testing must be highly sensitive, [4,5] while others advocate less sensitive testing primarily for community-level surveillance.[6] Recognizing the limited capacity for NAAT in many settings, WHO and national public health agencies have issued target product profiles for Ag-RDTs[7] and interim guidance for their use in select circumstances.[8–10] Meanwhile, a global partnership has begun to manufacture and distribute 120 million Ag-RDTs in LMICs.[11]

As Ag-RDTs become more broadly available, it is important to understand the conditions under which implementing Ag-RDTs would be preferable to relying on NAAT and/or clinician judgment alone. This understanding must balance accuracy (i.e., sensitivity and specificity) with other considerations such as turnaround time and the likely response to a positive (or negative) diagnosis. We sought to aid decision-making by quantifying these tradeoffs for patients presenting with symptomatic COVID-19 in the inpatient and outpatient settings.

Methods

Overview

We developed a simulation model of COVID-19 diagnostic evaluation in which we compared the use of point-of-care Ag-RDT, centralized (e.g., hospital-based) NAAT, and clinical judgment (i.e., without virological testing) among individuals presenting with symptoms suggestive of COVID-19. We parameterized the model using published data (and assumptions, where published data did not exist) regarding SARS-CoV-2 viral dynamics, clinical features of COVID-19, and diagnostic assay performance. The model accounts for the following aspects of disease dynamics and diagnostic testing (Figure 1): (1) time- dependent transmission from index cases and their contacts (Figure 1A); (2) variable timing of clinical presentation; (3) non-uniform distribution of peak viral burden across the population (Figure 1B); (4) correlation of viral burden with both assay sensitivity (Figure 1B, dotted horizontal lines) and infectivity (eAppendix Figure A1); and (5) assay-dependent decline in sensitivity with time since symptom onset (Figure 1C).

Simulated patient populations and testing strategies

We separately simulated two symptomatic patient populations: patients with moderate-to-severe illness in a hospital setting (of whom 40% were assumed to have COVID-19), and mildly symptomatic people in an outpatient setting (assuming that 10% have COVID-19[12]). Other differences between the two settings included a longer turnaround time for NAAT in the outpatient setting (1 day in the hospital, 3 days for outpatients[13,14]), a greater amount of presumptive isolation while awaiting a diagnostic result in the hospital setting, a longer duration of symptoms before presentation in the hospital setting, and a greater value placed on clinical improvement in the hospital setting versus reduction of transmission in the outpatient setting; details in eAppendix. These features were all evaluated in sensitivity analyses.

We also modeled several differences between NAAT and Ag-RDT, based on available data from published studies and preprints. Ag-RDT was assumed to have lower initial sensitivity (85% relative to NAAT at illness onset, Figure 2B and eAppendix Table A1) but faster turnaround time (3 hours, versus 1 to 3 days for NAAT). The modeled decline in assay sensitivity with time since symptom onset was also faster for Ag-RDT than for NAAT (Figure 1C).

In addition to comparing Ag-RDT to NAAT, we compared both virological assays to two clinical (i.e., non-virological) diagnostic approaches: a "treat all" approach that provides the available interventions (such as isolation, treatment, or contact notification and quarantine) to all patients for whom COVID-19 diagnoses were considered, or a "clinician judgment" approach with intervention for those patients whom clinicians deemed most likely to have COVID-19. Clinician judgment was estimated to have 80% sensitivity (comparable, at the time of clinical presentation, to the proportion of patients detectable by NAAT), with corresponding 50% specificity[15] (compared to 99.5% and 99%, respectively, for NAAT and Ag-RDT, eAppendix Table A1). Due to the absence of virological confirmation, we assumed that clinical diagnoses may result in less intense intervention – for example, less stringent isolation, less case reporting or contact notification, or more circumspect clinical management – and therefore reduced benefit or harm of intervention (by 25% in the primary analysis), compared to virological diagnoses made using Ag-RDT or NAAT.

For directness of comparison, we assumed that only a single diagnostic approach would be used in a given setting and patient population, with no further testing or clinical risk stratification after the initial evaluation.

Simulated SARS-CoV-2 transmission dynamics and diagnosis

Our model stochastically simulated the timing of transmission events from index cases and their contacts, relative to the index cases' symptom onset (Figure 1a; details in eAppendix). The model accounted for pre-symptomatic transmission,[16–18] delays between symptom onset and clinical presentation,[19–21] partial isolation while awaiting test results, and more stringent (but imperfect) isolation and contact tracing/quarantine once a COVID-19 diagnosis was made.[22,23]

We assumed that peak upper-airway viral burden coincided with the onset of symptoms and was log-normally distributed in the population (Figure 1b),[24,25] and that both diagnostic sensitivity and infectivity were functions of this viral burden. Specifically, based on observed relationships between NAAT cycle threshold, viral culture, and infectivity,[19,26–28] and similar to other models,[6] we assumed a log-linear relationship between peak viral burden and peak infectivity, above a minimum threshold of 10^3 viral genome copies. We also assumed that for both Ag-RDT and NAAT, the patients whose SARS-CoV-2 infections were detected had higher viral burdens than those not detected by the same assay on the same day of illness (Figure 1b). Thus, since viral burden was correlated with infectivity, cases who could

be diagnosed only by a highly sensitive assay (NAAT) were less infectious, on average, that those who could be diagnosed by a less sensitive assay (Ag-RDT) (eAppendix Figure A1). We also assumed that assay sensitivity declined over time after an initial six-day period of maximal sensitivity; Ag-RDT sensitivity declined on a similar timeline as infectivity, while NAAT sensitivity declined more slowly and thus detected some patients who were no longer infectious (Figure 1c; details in eAppendix).[19,29,30]

Using these assumptions, we simulated a population of 1,000,000 "index" patients, with COVID-19 each with a specified peak viral burden and time to clinical presentation. Our model included transmission events from index patients to their contacts (distributed according to index patients' relative viral burden and thus infectivity) and transmission events from those contacts (assuming no correlation between infectivity of index patients and their contacts). We simulated the result of each diagnostic test based on the patient's peak viral burden, the time (from symptom onset) that the test was performed, and the assay's sensitivity at that point in time (Figure 1C). Based on these simulated diagnostic results and the corresponding intensity of interventions (e.g., isolation before the test, while awaiting test results, and after results returned) that determined the probability that intervention prevented transmission events, we then simulated which transmission events would be averted under each diagnostic strategy.

In addition to averted transmission, we also modeled the clinical impact of interventions as being time-sensitive. Specifically, we assumed that avertible COVID-19 morbidity and mortality in the index patient declined exponentially with time since symptom onset, with a small magnitude and short half-life of 3.5 days in the outpatient setting (reflecting a short window of opportunity for antiviral therapy to avert hospitalization[31]) and a greater magnitude (based on increased average severity of disease) and slightly longer half-life in the hospital setting (details in eAppendix).

Decision curve analysis and net benefit estimation

We compared the value of different diagnostic strategies using an approach rooted in decision curve analysis (DCA), a method commonly used to weigh risks of over- and under-treatment.[32–34] DCA combines the benefits of true-positive diagnoses and the harms of false-positive diagnoses on a single scale of "net benefit," accounting for the prevalence of disease and the strength of preference for a true-positive versus false-positive diagnosis. We modified this approach to account for the importance of making diagnoses not only accurately, but also rapidly, among individuals with the highest

transmission potential, and with a robust response (e.g., directed clinical therapy, stringent isolation, and contact notification/quarantine). We then used net benefit analysis to estimate, and compare on a single scale, the value added by different diagnostic strategies, accounting for both transmission-related and clinical benefit achieved through a given strategy and the harm incurred from intervening on patients without COVID-19.[32]

To accomplish this, we represented the net benefit of true-positive diagnoses as a sum of the reduction in transmission (stochastically simulated across the population) and the clinical benefit (decaying exponentially with time since clinical presentation, as described above).

We thus calculated the net benefit of diagnosis with assay or clinical diagnostic approach *j* as

$$B_j = \left(p \, \frac{g_1(T_{1j}) + (1-g_1)\,(T_{2j}) + c\,\sum_i x_{ij} e^{-d\,t_{ij}}}{(1+c)} - \; (1 - sp_j)(1-p)\,\frac{q}{1-q} \right) l_j \,,$$

where p is the prevalence of COVID-19 among the patients evaluated; g_1 is the relative value of preventing first-generation versus downstream transmission (eAppendix Table A1); T_{1j} and T_{2j} are the proportion of transmission events from index cases and immediate contacts, respectively, that are prevented by approach *j* after accounting for both sensitivity and infectivity in the simulation of 1,000,000 patients described above; c is the expected clinical benefit of clinical intervention at symptom onset, compared to the benefit of preventing all transmission from the same cases; x_{ij} indicates whether case *i* is detected by approach *j*; d is the rate at which potential clinical benefit is lost as time passes; t_{ij} is the time from symptom onset to detection with approach *j*; sp_j is the specificity of assay/approach *j*; $\frac{q}{1-q}$ is the harm of intervening on a false positive, relative to the benefit of preventing all transmission and avertible morbidity and mortality from one true case; and l_j is the reduction in intervention stringency associated with a given diagnostic method (i.e., $l_j \leq 1$ when *j* is non-virological diagnosis).

We present results as decision curves, which plot the net benefit of each diagnostic approach across a range of values of the "threshold probability" q. This threshold probability represents the diagnostic certitude (i.e., probability of a patient having COVID-19) at which intervening at the onset of illness and not intervening would have equal expected benefit.[34] For one-way and two-way sensitivity analyses of other parameters, we hold q fixed at a value reflecting the willingness in current practice to isolate or quarantine individuals with specified probabilities of having or developing COVID-19.

Parameter estimates are described in eAppendix Table A1. All analyses were performed using R version 4.0.2.

Results

Transmission-related benefits of diagnosis

The simulated number and timing of transmission events, under each diagnostic scenario, are shown for the outpatient setting in Figure 2. Specifically, 62% of COVID-19 cases were detected under the Ag-RDT scenario, compared to 80% for both NAAT and clinical judgment (Figure 2A). However, since Ag-RDT detected cases more rapidly than NAAT (Figure 2A) and also detected the most infectious cases (responsible for 77% of future transmission), Ag-RDT was projected to avert more transmission than either NAAT or clinical judgment (Figure 2B). For example, in the primary analysis, Ag-RDT-based diagnosis was estimated to avert 31% of transmission from outpatients with COVID-19 (assigning 1:4 weight to immediate versus downstream transmission; eAppendix), versus 28% for NAAT and 20% for non-virological (clinical) diagnosis alone. By contrast, in the hospital setting – where NAAT turnaround time is shorter and isolation while awaiting test results was assumed to be stringent – NAAT prevented a greater proportion of transmission (34%) than Ag-RDT (30%) (Figure 2D).

Net benefit and decision curves

In the outpatient setting, Ag-RDT provide greater estimated net benefit than NAAT, considering not only reductions in transmission but also clinical benefits and harms of false-positive diagnosis (Figure 3A). This was true for all threshold probabilities below 50% (i.e., if the harm of missing a COVID-19 case is deemed greater than the harm of making a false-positive diagnosis). NAAT was estimated to provide greater net benefit than Ag-RDT if NAAT turnaround time was reduced to 1 day (Figure 3A). The net benefit of Ag-RDT relative to NAAT was lower when delays in diagnosis were less clinically consequential, Ag-RDT turnaround time was long, or adherence to isolation was high while awaiting a test result or low after a positive result (eAppendix, Figure A6) – but Ag-RDT still offered greater net benefit than NAAT in each of these scenarios. NAAT could also outperform Ag-RDT on the basis of its superior specificity, in situations where the harm of a false-positive diagnosis was deemed to be substantially greater than the harm of a missed COVID-19 diagnosis (high threshold probabilities in

Figure 3A). This might occur, for example, with high-risk patients who would be placed into congregate isolation facilities if they tested positive.

In the hospital setting, by contrast, NAAT offered greater net benefit than Ag-RDT under the assumptions of our primary model. In sensitivity analyses, Ag-RDT was estimated to offer greater net benefit at most threshold probabilities if NAAT turnaround time increased to 2 days (Figure 3B), and the benefit of Ag-RDT approached that of NAAT if clinical interventions available after COVID-19 diagnosis were highly time-sensitive.

Clinical diagnostic approaches provided greater net benefit than virological (Ag-RDT or NAAT) testing only if the relative harm of false-positive diagnoses was deemed to be very low (e.g., threshold of <3% in the outpatient setting or <7% in the hospital setting, Figure 4). In comparison, we estimated the empirical threshold probability for COVID-19 diagnosis in many settings to be 10% or higher (eAppendix). Assuming that people diagnosed clinically receive the same benefits (e.g., stringent isolation, antiviral therapy) as people diagnosed virologically may also overestimate the benefit of non-virological diagnosis, especially at low threshold probabilities when some diagnoses are made for people with low probability of having COVID-19. When we considered the possibility that this low diagnostic certainty could result in less intensive intervention (reducing the benefit of each true-positive diagnosis by 25% when made clinically, solid green and black lines in Figure 4B and 4D), virological testing became the preferred diagnostic approach at nearly all threshold probabilities.

These results contrast with those produced using a conventional approach to decision curve analysis,[33,34] in which the same net benefit was assigned to all true-positive diagnoses regardless of the speed of diagnosis or the infectivity of diagnosed cases. Such an analysis strongly favored NAAT in all settings on the basis of its high sensitivity (eAppendix Figure A5), even when Ag-RDT would avert more transmission and lead to better clinical and population-level outcomes.

Discussion

As antigen-detection SARS-CoV-2 rapid diagnostic tests become more widely available, it is important to identify the settings under which such tests offer incremental benefit to NAAT or clinical judgment alone. Using a novel adaptation of net benefit analysis, we demonstrate that Ag-RDT is likely to outperform NAAT under typical outpatient conditions that include >24-hour turnaround times for NAAT, even if the Ag-RDT has considerably lower sensitivity. We also demonstrate conditions under which Ag-

RDT would be favored over NAAT in a hospital setting, including NAAT turnaround-times of 48 hours or more, or use of Ag-RDT assays that achieve sensitivity >95% relative to NAAT during acute illness. We also demonstrate that both NAAT-based and Ag-RDT-based testing are preferable over clinical judgment-based or "treat-all" approaches under reasonable assumptions about the likely intensity of response to virologically unconfirmed diagnoses.

Our novel application of decision curve analysis to COVID-19 demonstrates the importance of accounting for factors beyond sensitivity and specificity when evaluating infectious disease diagnostics. Although NAAT offers higher sensitivity (and, in symptomatic patients, potentially better specificity) than Ag-RDT, the fact that Ag-RDT delivers results more rapidly and identifies the most highly infectious individuals can make it equivalent or superior to NAAT in averting transmission. Incorporating all of these considerations into a single metric of net benefit, our analysis suggests that Ag-RDT is likely to produce better outcomes than NAAT in many outpatient settings with >24 hour NAAT turnaround time (whereas NAAT is likely to be preferred in most inpatient facilities with NAAT capacity that can achieve <48-hour turnaround). Conventional decision curve analysis – based only on sensitivity, specificity, prevalence, and threshold probability – would incorrectly favor the more accurate test (i.e., NAAT) in all settings, even when the timing of diagnosis is critical.

Although any model-based findings are subject to bias, our estimates are likely to be conservative regarding the benefit of adopting Ag-RDT in clinical settings. First, our analysis does not incorporate the lower economic cost of Ag-RDTs relative to NAAT – though these costs may be partially offset by additional follow-up tests for SARS-CoV-2 or alternative diagnoses that may be required after false-negative Ag-RDT results. Second, our primary estimates of Ag-RDT sensitivity in the acute phase (i.e., 85% relative to NAAT) may be conservative.[19] Third, our model does not account for the potential that patients who test falsely negative by Ag-RDT may still receive some degree of isolation or clinical intervention on the basis of high clinical suspicion.

Our model is limited by underlying data availability about SARS-CoV-2 dynamics. In particular, data on the relationship between viral burden and infectivity remain sparse. Our ability to draw conclusions that apply across settings and assays is also limited by varying clinical and public health practice – for instance, the extent to which symptomatic people self-isolate or contact tracing is performed are likely to vary widely across settings. Estimates of threshold probability are also context specific, depending, for example, on the individual and societal economic costs of interventions such as contact investigation and quarantine in each particular setting. Third, certain parameter estimates – such as the importance

of preventing downstream transmission relative to preventing poor clinical outcomes in people currently infected – depend on the future course of the epidemic more broadly. Preventing transmission may be more important in settings with emerging or widespread transmission and less important in settings with resolving epidemics or imminent widespread vaccination. Finally, our analysis focuses on diagnosis of symptomatic individuals and quarantine of their direct contacts, and does not consider the potential role of Ag-RDT in preventing pre-symptomatic or asymptomatic transmission.

In conclusion, this novel application of decision curve analysis demonstrates that for individuals with symptoms suggestive of COVID-19, an Ag-RDT with 85% sensitivity (measured relative to NAAT and during acute symptoms) could offer greater net benefit than either NAAT or clinician-driven diagnosis. This is true particularly in outpatient settings – where NAAT turnaround times are often long, prevalence relatively low, and the benefits of diagnosis dominated by the potential to reduce onward transmission. In most inpatient contexts, NAAT is likely to be preferred, particularly for patients with severe illness presenting after more than a week of symptoms, for whom NAAT may guide clinical intervention while Ag-RDT has a high probability of being falsely negative. Even in hospitalized populations, however, Ag-RDTs could offer greater overall benefit than NAAT-based testing in when NAAT delays are long or Ag-RDT sensitivity at symptom onset surpasses 95% that of NAAT. Thus, in most clinical settings where NAAT testing cannot be consistently performed within 48 hours, Ag-RDT should be strongly considered for diagnosis of patients presenting with COVID-19 symptoms.

References

1. Dan-Nwafor C, Ochu CL, Elimian K, et al. Nigeria's public health response to the COVID-19 pandemic: January to May 2020. *J Glob Health*. 10(2). doi:10.7189/jogh.10.020399

2. Aziz AB, Raqib R, Khan WA, et al. Integrated control of COVID-19 in resource-poor countries. *International Journal of Infectious Diseases*. 2020;101:98-101. doi:10.1016/j.ijid.2020.09.009

3. Giri AK, Rana DR. Charting the challenges behind the testing of COVID-19 in developing countries: Nepal as a case study. *Biosaf Health*. 2020;2(2):53-56. doi:10.1016/j.bsheal.2020.05.002

4. Woloshin S, Patel N, Kesselheim AS. False Negative Tests for SARS-CoV-2 Infection — Challenges and Implications. *New England Journal of Medicine*. 2020;383(6):e38. doi:10.1056/NEJMp2015897

5. West CP, Montori VM, Sampathkumar P. COVID-19 Testing: The Threat of False-Negative Results. *Mayo Clin Proc*. 2020;95(6):1127-1129. doi:10.1016/j.mayocp.2020.04.004

6. Larremore DB, Wilder B, Lester E, et al. Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening. *Sci Adv*. Published online November 20, 2020. doi:10.1126/sciadv.abd5393

7. World Health Organization. *COVID-19 Target Product Profiles for Priority Diagnostics to Support Response to the COVID-19 Pandemic v.1.0.*; 2020. Accessed October 25, 2020. https://www.who.int/publications/m/item/covid-19-target-product-profiles-for-priority-diagnostics-to-support-response-to-the-covid-19-pandemic-v.0.1

8. US Centers for Diseases Control and Prevention. Interim Guidance for Antigen Testing for SARS-CoV-2. COVID-19 (Coronavirus Disease). Published December 5, 2020. Accessed December 15, 2020. https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antigen-tests-guidelines.html

9. European Centre for Disease Prevention and Control. Options for the use of rapid antigen tests for COVID-19 in the EU/EEA and the UK. Published online November 19, 2020. Accessed December 15, 2020. https://www.ecdc.europa.eu/sites/default/files/documents/Options-use-of-rapid-antigen-tests-for-COVID-19.pdf

10. World Health Organization. Antigen-detection in the diagnosis of SARS-CoV-2 infection using rapid immunoassays: Interim guidance. Published online September 11, 2020. Accessed December 15, 2020. https://www.who.int/publications/i/item/antigen-detection-in-the-diagnosis-of-sars-cov-2infection-using-rapid-immunoassays

11. Global partnership to make available 120 million affordable, quality COVID-19 rapid tests for low- and middle-income countries. Accessed October 25, 2020. https://www.who.int/news/item/28-09-2020-global-partnership-to-make-available-120-million-affordable-quality-covid-19-rapid-tests-for-low--and-middle-income-countries

12. The share of COVID-19 tests that are positive. Our World in Data. Accessed November 10, 2020. https://ourworldindata.org/grapher/positive-rate-daily-smoothed

13. Weiner R, Wan W, Hauslohner A. Long delays in getting test results hobble coronavirus response. *Washington Post*. https://www.washingtonpost.com/health/long-delays-in-getting-test-results-hobble-coronavirus-response/2020/07/12/d32f7fa8-c1fe-11ea-b4f6-cb39cd8940fb_story.html. Accessed November 10, 2020.

14. Coronavirus test results must come in 24 hours, says Sage scientist. the Guardian. Published October 18, 2020. Accessed November 10, 2020. http://www.theguardian.com/world/2020/oct/18/coronavirus-test-results-must-come-in-24-hours-says-sage-scientist

15. Sun Y, Koh V, Marimuthu K, et al. Epidemiological and Clinical Predictors of COVID-19. *Clin Infect Dis*. 2020;71(15):786-792. doi:10.1093/cid/ciaa322

16. He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine*. 2020;26(5):672-675. doi:10.1038/s41591-020-0869-5

17. Tindale LC, Stockdale JE, Coombe M, et al. Evidence for transmission of COVID-19 prior to symptom onset. Franco E, Lipsitch M, Lipsitch M, Miller J, Pitzer VE, eds. *eLife*. 2020;9:e57149. doi:10.7554/eLife.57149

18. Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA. Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. *Emerg Infect Dis*. 2020;26(6):1341-1343. doi:10.3201/eid2606.200357

19. Iglⷿi Z, Velzing J, van Beek J, et al. Clinical evaluation of the Roche/SD Biosensor rapid antigen test with symptomatic, non-hospitalized patients in a municipal health service drive-through testing site. *medRxiv*. Published online January 1, 2020:2020.11.18.20234104. doi:10.1101/2020.11.18.20234104

20. Tenforde MW. Symptom Duration and Risk Factors for Delayed Return to Usual Health Among Outpatients with COVID-19 in a Multistate Health Care Systems Network — United States, March–June 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69. doi:10.15585/mmwr.mm6930e1

21. Faes C, Abrams S, Van Beckhoven D, Meyfroidt G, Vlieghe E, Hens N. Time between Symptom Onset, Hospitalisation and Recovery or Death: Statistical Analysis of Belgian COVID-19 Patients. *Int J Environ Res Public Health*. 2020;17(20). doi:10.3390/ijerph17207560

22. Steens A, Blasio BF de, Veneti L, et al. Poor self-reported adherence to COVID-19-related quarantine/isolation requests, Norway, April to July 2020. *Eurosurveillance*. 2020;25(37):2001607. doi:10.2807/1560-7917.ES.2020.25.37.2001607

23. Smith LE, Amlôt R, Lambert H, et al. Factors associated with adherence to self-isolation and lockdown measures in the UK; a cross-sectional survey. *medRxiv*. Published online June 2, 2020:2020.06.01.20119040. doi:10.1101/2020.06.01.20119040

24. Kissler SM, Fauver JR, Mack C, et al. *Viral Dynamics of SARS-CoV-2 Infection and the Predictive Value of Repeat Testing*. Epidemiology; 2020. doi:10.1101/2020.10.21.20217042

25. To KK-W, Tsang OT-Y, Leung W-S, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *The Lancet Infectious Diseases*. 2020;20(5):565-574. doi:10.1016/S1473-3099(20)30196-1

26. Sia SF, Yan L-M, Chin AWH, et al. Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. *Nature*. 2020;583(7818):834-838. doi:10.1038/s41586-020-2342-5

27. Jefferson T, Spencer E, Brassey J, Heneghan C. Viral cultures for COVID-19 infectivity assessment. Systematic review. *medRxiv*. Published online January 1, 2020:2020.08.04.20167932. doi:10.1101/2020.08.04.20167932

28. Singanayagam A, Patel M, Charlett A, et al. Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. *Euro Surveill*. 2020;25(32). doi:10.2807/1560-7917.ES.2020.25.32.2001483

29. Berger A, Ngo Nsoga MT, Perez-Rodriguez FJ, et al. Diagnostic accuracy of two commercial SARS-CoV-2 Antigen-detecting rapid tests at the point of care in community-based testing centers. *medRxiv*. Published online January 1, 2020:2020.11.20.20235341. doi:10.1101/2020.11.20.20235341

30. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction–Based SARS-CoV-2 Tests by Time Since Exposure. *Ann Intern Med*. Published online May 13, 2020. doi:10.7326/M20-1495

31. Chen P, Nirula A, Heller B, et al. SARS-CoV-2 Neutralizing Antibody LY-CoV555 in Outpatients with Covid-19. *New England Journal of Medicine*. 2020;0(0):null. doi:10.1056/NEJMoa2029849

32. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ*. 2016;352:i6. doi:10.1136/bmj.i6

33. Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. *Eur Urol*. 2018;74(6):796-804. doi:10.1016/j.eururo.2018.08.038

34. Fitzgerald M, Saville BR, Lewis RJ. Decision Curve Analysis. *JAMA*. 2015;313(4):409. doi:10.1001/jama.2015.37

Figure 1

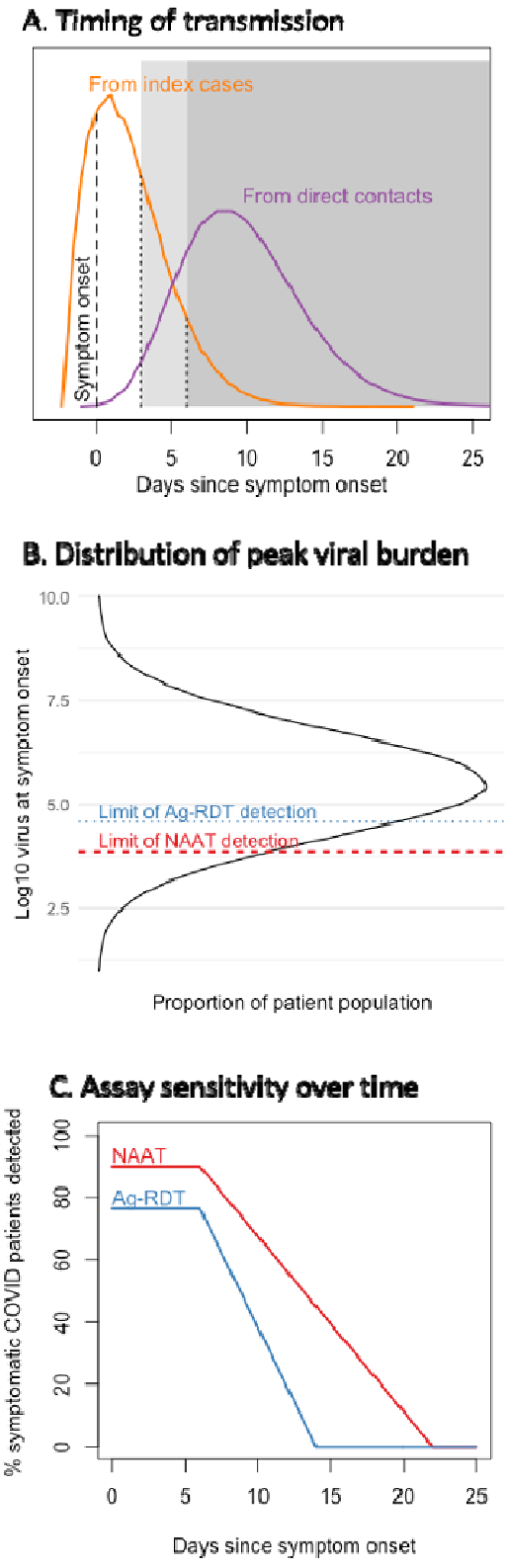

Figure 2

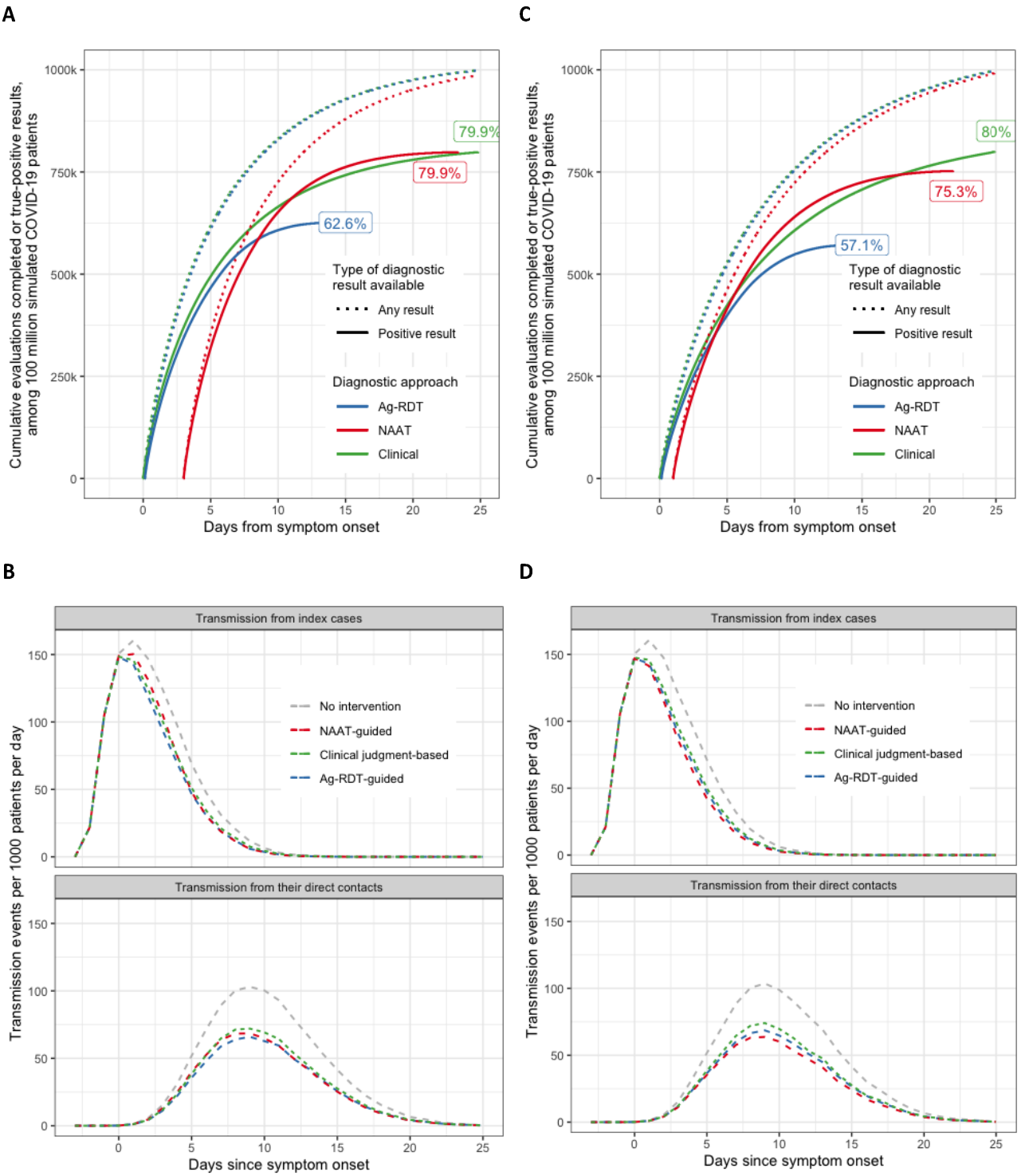

Figure 3

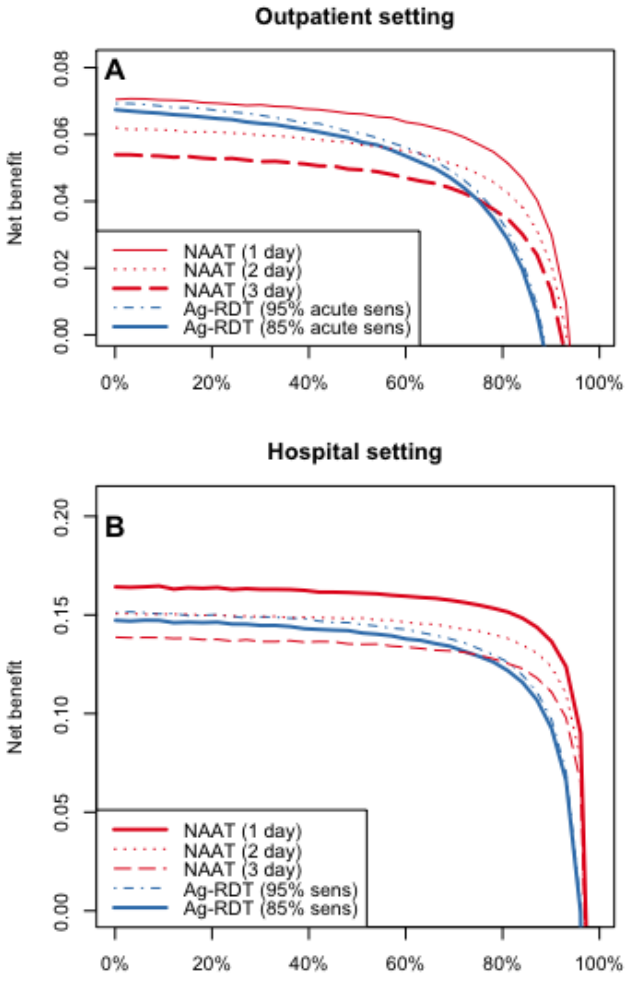

Figure 4

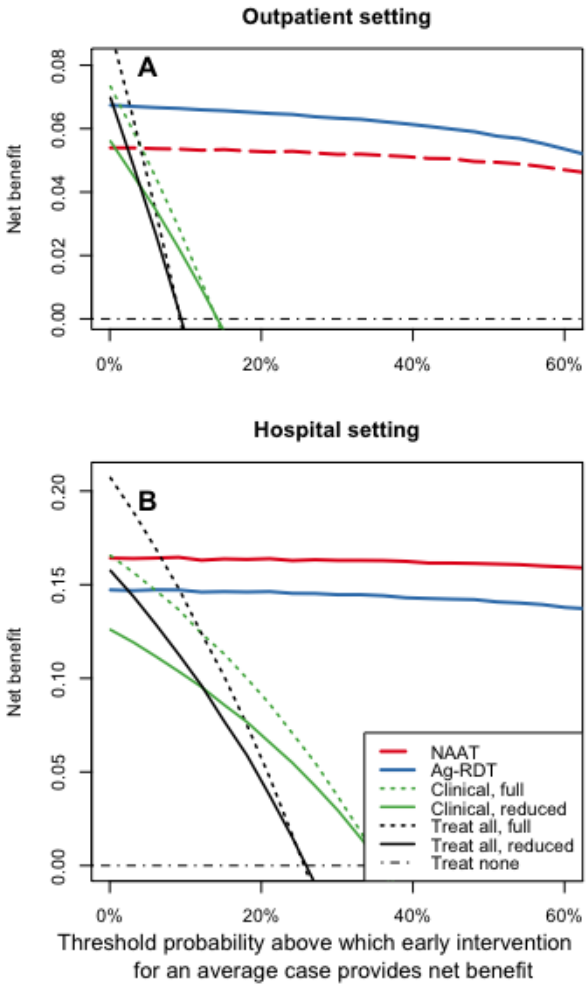

Figure Legends

Figure 1: Key assumptions in simulation of net benefit. Panel A: The timing of transmission events both from cases and from their direct contacts, in absence of intervention, follows a non-uniform distribution relative to the time of symptom onset in the index case. Shading illustrates that different amounts of transmission are preventable after earlier (light shading) versus later (dark shading) diagnosis. **Panel B**: The simulated viral burden in upper respiratory diagnostic specimens at symptom onset is log-normally distributed across a simulated population. This may be conceptualized in terms of genome copies on a quantitative NAAT (units shown on y axis) or antigen copies (assumed to vary in proportion to genome copies during acute illness). **Panel C**: After an initial period of maximal detection, the simulated sensitivity of Ag-RDT declines on a similar timeline as infectivity, and sensitivity of NAAT declines more slowly.

Figure 2: Timing of testing and SARS-CoV-2 transmission events under different diagnostic algorithms in the outpatient setting. Panel A: Test results over time among 1,000,000 COVID-19 patients. Patients present at a median 3 days after symptom onset. Ag-RDT and clinical judgment provide same-day results (and thus have curves shifted to the left) relative to NAAT, but NAAT ultimately detects more true-positive COVID-19 cases than Ag-RDT (80% versus 63%). Clinical judgment may diagnose as many true COVID-19 cases as NAAT, but with lower specificity and no preference for the most infectious cases (not shown). **Panel B:** Quantity and timing of transmission with and without intervention. Much of the potential effect of testing on transmission is diminished by delayed clinical presentation and incomplete implementation of or adherence to isolation and contact quarantine interventions, with the stringency of intervention further reduced when diagnosis is non-virological (based on clinical judgment alone).

Figure 3: Net benefit analysis of COVID-19 diagnostic strategies. The 'threshold probability' is the post-test probability of having COVID-19, at which the benefits of intervening are balanced with the risks (thus a higher threshold probability corresponds to more stringent conditions for diagnosis, and vice versa). For the current analysis, in which the benefit of intervention is time-dependent, the threshold probability is defined for intervention at the onset of illness. In all panels, the most favored test at a given threshold probability is that maximizing net benefit, i.e. corresponding to the uppermost curve. Heavy red and blue curves represent the net benefit achieved by NAAT-driven and Ag-RDT-driven

intervention under the primary assumptions in the outpatient (**panel A**) and hospital (**panel B**) settings. At most threshold probabilities, greater net benefit is achieved with Ag-RDT in the outpatient setting (solid blue curve in panel A) and with NAAT in the hospital setting (dashed red curve in panel B). Additional curves show the effects of varying turnaround times of NAAT (solid = 1 day, dotted = 2 days, dashed = 3 days) and Ag-RDT sensitivity relative to NAAT in acute infection (solid = 85%, dash-dot = 95%).

Figure 4. Comparison of virologic diagnosis (using NAAT or Ag-RDT) to clinical diagnostic approaches (treating all, black, or treating those judged most likely to have COVID-19, green). If clinical diagnoses receive the same full intervention as virologically diagnosed cases (dashed green and black lines), then treating all patients with suspected COVID-19 could be the preferred approach at low threshold probabilities (i.e., at low relative harm of intervening on false positives). If, however, the lower certainty of a clinical diagnosis results in a reduced intensity of intervention and 25% lower net benefit compared to a virologic diagnosis as in our primary analysis (solid green and black lines), then virologic testing provides greater net benefit than non-virologic diagnosis at all or nearly all threshold probabilities.