BIOMEDICINE

REVISTA DE Biología Tropical

https://doi.org/10.15517/rev.biol.trop..v73i1.57971

The performance of mass testing strategies for COVID-19: a case study for Costa Rica

Maikol Solís[1]*; http://orcid.org/0000-0002-8776-6142
Carlos Pasquier-Jaramillo[1]; https://orcid.org/0009-0004-8758-6872
Santiago Nuñez-Corrales[2]; http://orcid.org/0000-0003-4342-6223
Germán Madrigal-Redondo[3]; http://orcid.org/0000-0002-9856-4044
Andrés Gatica-Arias[4]; http://orcid.org/0000-0002-3841-0238

1. Centro de Investigación en Matemática Pura y Aplicada (CIMPA), Escuela de Matemática, Universidad de Costa Rica, San José, Costa Rica; maikol.solis@ucr.ac.cr (*Correspondence); carlos.pasquier@ucr.ac.cr
2. National Center for Supercomputing Applications, University of Illinois Urbana-Champaign, Urbana, Illinois, United States of America; nunezco2@illinois.edu
3. Instituto de Investigaciones Farmacéuticas (INIFAR), Universidad de Costa Rica, San José, Costa Rica; german.madrigal@ucr.ac.cr
4. Escuela de Biología, Universidad de Costa Rica, San José, Costa Rica; andres.gatica@ucr.ac.cr

Received 09-IV-2024. Corrected 28-XI-2024. Accepted 04-III-2025.

ABSTRACT

Introduction: In this article, we derive the behavior of four different mass testing strategies, grounded in guidelines and public health policies issued by the Costa Rican public healthcare system.
Objective: To formally develop the changes of each studied mass testing strategy under different contexts related to people's risk, costs of testing, and accessibility to alternative testing technologies.
Methods: We take over a pre-classifier applied to individuals capable of partitioning suspected individuals into low-risk and high-risk groups. We consider the impact of three testing technologies: RT-qPCR, antigen-based testing, and saliva-based testing (RT-LAMP). When available, we introduced a category of essential workers.
Results: Numerical simulation results confirm that strategies using only RT-qPCR tests cannot achieve sufficient stock capacity to provide efficient detection regardless of prevalence, sensitivity, or specificity. Strategies that harness the power of pooling and RT-LAMP either maximize stock capacity, detection efficiency, or both.
Conclusions: Investing in data quality and classification accuracy can improve the odds of achieving pandemic control and mitigation. Future work will be focused on, based on our findings, constructing representative synthetic data through agent-based modeling and studying the properties of specific pre-classifiers under various scenarios.

Keywords: mass testing; COVID-19 Costa Rica; RT-qPCR; antigen test; RT-LAMP; pooling; detection strategies.

RESUMEN

Desempeño de estrategias de pruebas masivas para COVID-19: un estudio de caso para Costa Rica

Introducción: En este artículo, derivamos el comportamiento de cuatro diferentes estrategias de pruebas masivas, basadas en las directrices y políticas de salud pública emitidas por el sistema de salud pública de Costa Rica.
Objetivos: Desarrollar formalmente los cambios de cada estrategia de pruebas masivas estudiada bajo diferentes contextos relacionados con riesgo de las personas, costos de la prueba y acceso a tecnologías alternativas de pruebas.

Métodos: Asumimos un pre-clasificador aplicado a individuos, capaz de dividir a los sospechosos en grupos de bajo riesgo y alto riesgo. Consideramos el impacto de tres tecnologías de prueba: RT-qPCR, pruebas basadas en antígenos y pruebas de saliva (RT-LAMP). Cuando estuvo disponible, introdujimos una categoría de trabajadores esenciales.

Resultados: Los resultados de simulaciones numéricas confirman que las estrategias que utilizan únicamente pruebas RT-qPCR no pueden lograr una capacidad de existencias suficiente para proporcionar una detección eficiente, independientemente de la prevalencia, sensibilidad o especificidad. Las estrategias que aprovechan el poder tanto del agrupamiento (pooling) como del RT-LAMP maximizan la capacidad de existencias o la eficiencia de detección, o ambos.

Conclusiones: Invertir en la calidad de los datos como en la precisión de la clasificación puede mejorar las probabilidades de lograr el control y la mitigación de la pandemia. El trabajo futuro se concentrará, basándonos en nuestros hallazgos, en construir datos sintéticos representativos a través de modelado basado en agentes y estudiar las propiedades de pre-clasificadores específicos bajo varios escenarios.

Palabras clave: pruebas masivas; COVID-19 Costa Rica; RT-qPCR; pruebas de antígenos; RT-LAMP; agrupamiento; estrategias de detección.

INTRODUCTION

The recent SARS-CoV-2 pandemic has made mass testing strategies a key tool for managing and understanding the trajectory of communicable diseases. Recent studies suggest these strategies help to control outbreaks if they are underpinned by robust estimations of the pandemic's current and future impact. This requires a framework rooted in evidence-based planning, steering clear of potentially misleading metrics, and using complementary, information-driven efforts (Grantz et al., 2021). Implementing effective testing and epidemiological surveillance is hampered by many obstacles in countries around the globe, ranging from structural deficiencies in public health systems (Caliendo et al., 2013) to financial constraints that restrict access to essential technologies (Beaudevin et al., 2021; Yang & Rothman, 2004).

Let us first define asymptomatic as the population infected that will never develop symptoms, while pre-symptomatic patients develop symptoms after the incubation time (He et al., 2020; Tindale et al., 2020). Maximizing resources for a mass testing strategy result in a nonlinear allocation issue with generalized cost functions (Brandeau, 2004). However, the overall problem can change when we introduce pre-symptomatic and asymptomatic patients because of their potential to spread the virus unnoticed. The work by Kırkızlar et al. (2010) found, for the asymptomatic case, the cost-effectiveness of mass testing intervention is equivalent to a Markov Decision Process. This process must include prior data about individual test outcomes and behavioral change induced by an awareness of the disease. Clinical and theoretical studies reveal the importance of controlling the pre-symptomatic and asymptomatic individuals in early stages. Some examples include mass testing with pooling samples (Comess et al., 2022), comparative analysis with isolation components (Du et al., 2021), simulations with rapid saliva-based testing (Núñez-Corrales & Jakobsson, 2020). Longitudinal studies revealed a conservative estimate of 30-45 % asymptomatic cases, including pre-symptomatic (Oran & Topol, 2020; Oran & Topol 2021; Sah et al., 2021).

One of the major flaws in every mass testing strategy is the availability of resources and technology to perform it. Since December 2019, the healthcare systems in Latin America struggled with inadequate and tracking systems mostly due to the infeasibility to perform RT-qPCR to the entire population (Rubinstein, 2025). During high peak waves, global scarcity of reagents forces the laboratories to use alternative options (Avaniss-Aghajani et al., 2020). Despite its accuracy, RT-qPCR was an inadequate solution, given that it requires robust laboratory facilities and trained staff, both of

which are limited. Reporting of results normally takes 2-5 days, depending on the healthcare system capacity to process samples and perform administrative follow-up (Watkins et al., 2021). The cost per test ranges between $ 50 to $ 100 per result (Centers for Medicare & Medicaid Services, 2020).

Antigen-based tests comprise a less expensive alternative strategy to RT-qPCR. Those could cost between $ 30 to $ 50 and can give a result within a maximum of two hours (Wiencek et al., 2020). Detection becomes reliable during the first week after symptoms onset (Mercer & Salit, 2021), and *in vitro* studies show high specificity (> 99 %) but low sensitivity (> 66 % for nucleocapsid, > 85 % spike). In practice, the proportion of false negatives can increase to unacceptable levels when testing occurs after the week in which first symptoms appear. To provide a baseline for antigen-based alternatives, World Human Organization established in 2020 a minimum sensitivity of 80 % and specificity of 97 % compared with RT-qPCR. The Center for Disease Control and Prevention (CDC) published a set of guidelines and good practices along these lines (CDC, 2020), including confirmatory RT-qPCR test when antigen-based alternatives yield inconclusive results. Another technology is the Reverse Transcriptase Loop Mediated Isothermal Amplification (RT-LAMP). Its technology is similar to RT-qPCR given their molecular detection (Österdahl et al., 2020). The protocol can be scaled up because of the use of abundant standard reagents (Saidani et al., 2021) and, when tests are inconclusive, these are repeated at low costs. In relation to the progression of the disease, preliminary research on the sensitivity of RT-LAMP found its value on 85.2 % during the first nine days of infection and 44.4 % afterward, and an average of 60 % for asymptomatic patients (Nagura-Ikeda et al., 2020).

The mentioned technologies can be improved with alternative strategies like pooling or retesting. The pooling technique was first proposed by Dorfman (1943). The scheme divides the total number into different pools and tests each group. The negative groups declare all the individuals as negative. With the positive ones, another round of individual testing allows the detection of the infected individual(s). Performing multiple tests on the low-risk population constitutes an alternative to pooling. This scheme implies weekly or biweekly tests for the same group of individuals. Several studies have shown that frequent testing reduces the positivity rate and the number of sick leaves among workers (Haigh & Gandhi, 2021; Larremore et al., 2021; Plantes et al., 2021; Sandmann et al., 2020).

The review by Prado et al., (2023) about the pandemic situation showed how the RT-qPCR was the primary line of detection during the early phase of the pandemic coupled with contact tracing. In 2020, the lack and efficient diagnostic screening system, the country suffered multiple delays and high costs by tourists and commercial transporters (García-Puerta et al., 2023). In 2021, the Costa Rican Ministry of Health (Ministerio de Salud Costa Rica [MINSA]) introduced regulations for antigen-based testing (MINSA, 2021b). The regulation allowed private healthcare providers to test with this technology. A negative antigen-based test performed by these private providers does not require an RT-qPCR confirmatory test, though. This assumption suggests that negative patients could still be healthy, but a negative test result may not completely rule out the possibility of infection due to the test's low sensitivity. Only much later in their opening, the authorities allowed commercial import of antigen-based tests for the public. Hence, the antigen-based testing was not part of a mass testing strategy in Costa Rica.

This article explores the Costa Rica case on limited testing capabilities with RT-qPCR, scarce testing alternatives, and limited information infrastructure for patient tracing. We examine strategies for maximizing infrastructure effectiveness using multiple test types. The study shows *in silico* the behavior of different population-level strategies under the existence of mechanisms capable of predicting individual risk of contagion. We hypothesize the population's infection stays on *a prior* set of individual

and collective factors that allow predicting the outcome. Therefore, given any available-and possibly anonymized-information, the authorities allocate testing during an emergency.

To this end, we split the population into low and high risk. The high-risk group contains all symptomatic individuals, their epidemiological nexus, healthcare workers and essential workers with frequent viral exposure. The low-risk group consists of individuals whose features prevent COVID-19 like lack of comorbidities, young age, accessibility to health, among others (Zhang et al., 2023). It also represents the largest potential gain for proactive screening of pre-and asymptomatic populations. The work by Escobar et al. (2022) applied a Gradient Boosting Machine (GBM) to clinical and sociodemographic factors to split the population into those groups. They then evaluated pooled testing by computing the efficiency between Dorfman's pooling and matrix pooling strategies, as well as one-stage and two-stage strategies. Reported efficiency gains were significant.

In this work, we study the statistical and mathematical mechanisms behind massive testing strategies when using a classifier to detect subjects at risk before testing. We explore four different strategies. Strategy 1 follows the official guidelines for the high-risk group, ignoring the low-risk one. For the low-risk group, Strategy 2 uses a pooling technique, while Strategy 3 uses a multiple testing scheme. Finally, Strategy 4 combines RT-qPCR, Antigen and RT-LAMP to maximize the benefits of each technology. We formulated a probabilistic model to quantify the costs, detected positives and number of tests per person required in each strategy. To the best of our knowledge, the Costa Rican authorities have not implemented a similar study. Understanding the statistical properties of testing strategies based on separating vulnerable individuals into risk categories can aid in optimizing the efficiency of existing resources for future pandemics.

MATERIAL AND METHODS

In this study, we performed an *in silico* evaluation of the behavior of different massive testing strategies preceded by a patient classification mechanism. This section describes the contextual framework of our work.

Sensitivity, specificity, PPV and NPV: Let D_P be the condition of having the disease (i.e., infected) and D_N the condition of being not infected. The prevalence is estimated by $P(D_P)$ such that $P(D_N) = 1 - P(D_P)$. Let also N be the total population undergo testing. Thus, N x $P(D_P)$ are the true infected and N x $(1 - P(D_P))$ the true healthy people. Denote as R_P^j and R_N^j the results positive and negative of each test, respectively. In addition, let j = *PCR*, *Ag* or *LAMP* denote each available testing technology, RT-qPCR, Antigen or RT-LAMP respectively. We can thus define sensitivity as the proportion of people infected who are correctly identified as positive in the test, or $P(R_P^j \mid D_P)$. Specificity constitutes the proportion of people not infected who are correctly identified as negative in the test, or $P(R_N^j \mid D_N)$. When the prevalence is known, the relationships for testing positive or negative in a test become,

$$P(R_P^j) = P(R_P^j \mid D_P)P(D_P) + (1 - P(R_N^j \mid D_N))(1-P(D_P))P(R_N^j) = 1 - P(R_P^j).$$

Meanwhile, the positive predictive value (PPV) is the probability of being actually positive when infected, or $P(D_P \mid R_P^j)$. In contrast, the negative predicted value (NPV) is the probability of being negative while not having the disease, or $P(D_N \mid R_N^j)$. By virtue of Bayes' theorem,

$$P(D_P \mid R_P^j) = \frac{P(R_P^j \mid D_P)P(D_P)}{P(R_P^j)}, P(D_N \mid R_N^j) = \frac{P(R_N^j \mid D_N)P(D_N)}{P(R_N^j)}$$

In general, sensitivity and specificity are fixed values given testing technology. Nevertheless, NPV and PPV depend on the current prevalence. More generally, PPV increases and NPV decreases as a function of increasing prevalence. Given RT-qPCR testing has high sensitivity and specificity, PPV and NPV values are normally above 90 % regardless of the prevalence.

Antigen-based testing requires special attention due to its low sensitivity (80 %). For a population with 25-50 % prevalence, antigen-based testing will yield a PPV of 90-96 % given that the test is used 5 days after symptom onset. Increasing prevalence, for example, above 36 % yields a decrease in NPV decreases below 90 %, producing many false negatives. In other words, more than 10 % of patients tested were declared as negative when in reality they were infected. The procedure to follow in this case is to collect another sample across individuals whose results were negative results and perform an RT-qPCR confirmatory test.

Prevalence in the range of 1-10 % entails low impact from false negatives (NPV = 98-100 %). However, an unacceptable number of false positives arises when PPV reaches values between 21-75 %. That is, large quantities of healthy people are being declared as infected when they are not, with potentially negative impacts to workforce availability. This can become particularly significant when essential workers are involved. The recommended strategy for this group therefore becomes to apply antigen-based testing with high sensitivity to all the population for an initial screening (CDC, 2020). And to either perform a second round of antigen-based testing (with higher specificity) or an RT-qPCR test, those whose first test was positive.

Mass testing strategies: Pooling and multiple testing: Increasing the effectiveness of mass testing can be achieved through pooling or multiple testing. Pool testing requires three important conditions to work: (a) if all members in a group are negative, then the group yields negative in the pool analysis; (b) a single positive sample within a group makes the group test result positive-further testing is necessary to identify the true positives-and (c) the fraction of expected positive cases is small. Current literature describes two large classes of pooling strategies: adaptive and non-adaptive. Adaptive ones require incremental results to further stratify testing across the population. Non-adaptive methods set the pooling scheme prior to testing, and each group is tested independent of each other (Millioni & Mortarino, 2021). To simplify the estimation process, we used the most commf7on algorithm pooling strategy, the *one-dimensional (1D) protocol* (Dorfman, 1943). This scheme consists of mixing a group of samples, taken in batches. The analysis is then carried out only over these batches. If one batch is positive, then all members must be analyzed individually.

The main limitation with this technique is that it becomes useful only at low prevalence levels. Consider, for example, 100 people divided into groups of ten with only one positive patient (prevalence of 1 %). In this situation, nine of ten groups will be assigned a negative result. The remainder group with one positive case should be tested entirely again. The strategy described above required 20 tests instead of 100. In contrast, if ten people are infected and each group has one positive case in each group (prevalence of 10 %), this pooling strategy results in a total of 110 tests. Other issues include loss of sensitivity due to dilution or possible artifacts introduced by the actual sample collection protocol (Peeling et al., 2021; Watkins et al., 2021).

When pooling schemes are infeasible, multiple testing provides a straightforward solution. Results in Du et al. (2021) show that weekly testing and 2-week periods of isolation work best when transmission rates are high. If transmission rates decrease, then monthly testing and 1-week isolation periods provide the best solution to maintain the economy afloat. Two unpredictable factors make it challenging to translate results into policy. First, the asymptomatic and pre-symptomatic fractions of the population tend to be the most uncertain,

especially when testing strategies are being devised. Knowing how they behave explains the rate of disease spreading. Second, local transmission rates are modulated by multiple factors, including population, density, mitigation policies, and local immunity, where little or no control is possible.

Costa Rican testing guidelines: In the Costa Rican case, the Ministry of Health in MINSA (2021b), MINSA (2022), defined guidelines for antigen-based testing as an alternative to RT-qPCR, depending on whether the patient is tested in public or private health services. The discriminating element is the use of RT-qPCR confirmatory test after an antigen-based test outcome is negative within the public healthcare system. The private system is excluded from required confirmatory testing. These guidelines define a *suspicious patient* (i.e., high risk) when both symptoms (e.g., high fever, cold, loss of sense of smell sense) and a well identified epidemiological nexus (e.g., living with positive individuals, recent travel history) are present. Asymptomatic patients are deemed low risk. Therefore, the underlying principle establishes that high-risk patients must go to the public healthcare system, while the low-risk ones are directed to the private one.

Costa Rican guidelines directly follow CDC recommendations, which distinguish between congregate and community living settings. We note that Costa Rican guidelines have failed to consider prevalence across the population as a significant factor in how they differentiate between public and private health services. The main assumption behind this is that every patient tested in the private service has a low-risk of infection, and that consequently, antigen-based testing is reliable. This may not hold in the complex reality of disease spread of a small size, emerging economy.

High and low risk classification: Any successful mass testing strategy should be able to screen rapidly individuals while controlling as strictly as possible for false negatives and positives. Three elements are reported in this work to achieve this goal: cost-effectiveness, positive rate and number of tests per person. We focus on the sequence of events leading to a confirmatory test, depending on whether the person is symptomatic or not and the current level of prevalence of the disease. We hence propose a set of alternative configurations informed by features of the public-private healthcare system discussed above. Our work includes a two-step strategy for massive testing: classifying the patients into high-risk and low-risk categories, and later applying a suitable adaptive mass testing strategy per group.

The general strategy proceeds as follows:

1. Collect or access patient data in advance corresponding to factors that determine the probability of becoming exposed to COVID-19. Due to privacy reasons or local legislation, the patient data could be confidential. In those cases, we can use aggregated statistics and estimate synthetic models to simulate a usable data table.
2. Predict patient risk categories based using data above. All symptomatic patients or those with an epidemiological nexus are automatically classified as high-risk regardless of prediction outcomes.
3. Select a strategy based on the predicted risk category:
    a. High-risk group: provide antigen-based testing if symptoms started for 5 days or less or provide RT-qPCR testing otherwise. All negative outcomes must be confirmed with RT-qPCR.
    b. Low-risk group:
        i. Use a pooling technique with a pool size of five.
        ii. Perform antigen-based testing across all groups and perform antigen-based confirmatory testing to all members of groups with at least one positive.

The effectiveness of each strategy will depend on the prevalence, sensitivity, specificity, PPV and NPV of each test, as well as on the accuracy of the predictive model. As

mentioned before, we explore only the theoretical properties of such strategies assuming an arbitrary predictive model. Models can be fitted using a wide variety of information (i.e., residence-work location, socioeconomic status, comorbidities, recent travel). Then, we will use the combined characteristics of RT-qPCR and Antigen tests to create a massive strategy for all the population.

For the purposes of this study, define C^{PCR} = \$100 and C^{Ag} = \$50 as the cost of a single RT-qPCR and antigen-based test, respectively. Administrative expenses, fees and other costs were excluded. We assume a total population of N = 1000 individuals. For instance, using RT-qPCR test for the entire population yields $N \times C^{PCR}$ = \$100 000.

We denote a high-risk classification outcome by M_H, and a low-risk one by M_L. We define the classifier's sensitivity as $P(M_H \mid R_P^j)$, which contrasts the prediction against laboratory test results for each testing technology j. This value estimates the proportion of people being classified as high-risk when they have indeed a positive test result. To simplify, we assume the same sensitivity for RT-qPCR and antigen-based tests. We establish the following:

$$P(M_H \mid R_P) = P(M_H \mid R_P^{PCR}) = P(M_H \mid R_P^{Ag})$$

Meanwhile, the specificity $P(M_L \mid R_N^j)$ corresponds to the proportion of people classified as low-risk having a negative result. Again, we assume that both technologies have the same specificity, and we denoted just as $P(M_L \mid R_N)$. For the purpose of our computational study, we explored classifier combinations of sensitivity and specificity at 30 %, 60 % and 90 % for both variables. Knowing the prevalence, we can estimate

$$p_{M_H}^j = P(M_H \mid R_P)P(R_P^j) + (1 - P(M_L \mid R_N))P(R_N^j), p_{M_L}^j = 1 - p_{M_H}^j,$$

the probabilities of being classified as high or low risk depending on the testing technology. To combine both probabilities, we use the logit transformation

$$logit(p) = \log(\frac{p}{1-p})$$

which leads to

$$P(M_H) = logit^{-1}(\frac{logit(p_{M_H}^{PCR}) + logit(p_{M_{qH}}^{Ag})}{2})P(M_L) = 1 - P(M_H). (1)$$

The corresponding values for PPV and NPV are $P(R_P^j \mid M_H)$ and $P(R_N^j \mid M_L)$. By applying Equation (1), this becomes on:

$$P(R_P^j \mid M_H) = \frac{P(M_H \mid R_P^j)P(R_P^j)}{P(M_H)} P(R_N^j \mid M_L) = \frac{P(M_L \mid R_N^j)P(R_N^j)}{P(M_L)}$$

We obtain combined probabilities $P(R_P \mid M_H)$ and $P(R_N \mid M_L)$ via a similar treatment with the logit transformation.

In the context of antigen-based testing, we denote as S_{-5} the event if a patient has less than 5 days since the beginning of symptoms and S_{+5} otherwise. Since neither the high-risk condition nor the result of the test alter the distribution of patient symptoms, we assume that S_{-5} and S_{+5} are independent of M_H, R_P^j or R_N^j. While this assumption may not hold in all the cases, it will not affect the results due to the theoretical nature of this study. To re-estimate the probability correctly when the assumption does not hold, a detailed study of the patients should make results more precise to confirm the hypothesis. In our computational experiments, we set $P(S_{+5})$ with values of 30, 60 and 90 %. We define $P(S_{-5}) = 1 - P(S_{+5})$. We assume a greater proportion of RT-qPCR tests used directly on high-risk patients when (S_{+5}) increases, and the number of antigen-based tests used at the group level increases when (S_{-5}) increases. The following sections define formulas for the overall cost, number of tests per person and number of positive reports of each strategy.

Strategy 1: antigen-based testing: We model this scenario based on the Costa Rican public healthcare guidelines. Fig. 1 depicts the steps involved in this strategy, which adds a new decision layer prior to laboratory testing (blue box). The layer uses a classifier to determine high-risk (red box) or low-risk (green box) individuals. Using this label, the strategy applies a different mechanism to each group. The assumptions for this scenario are:

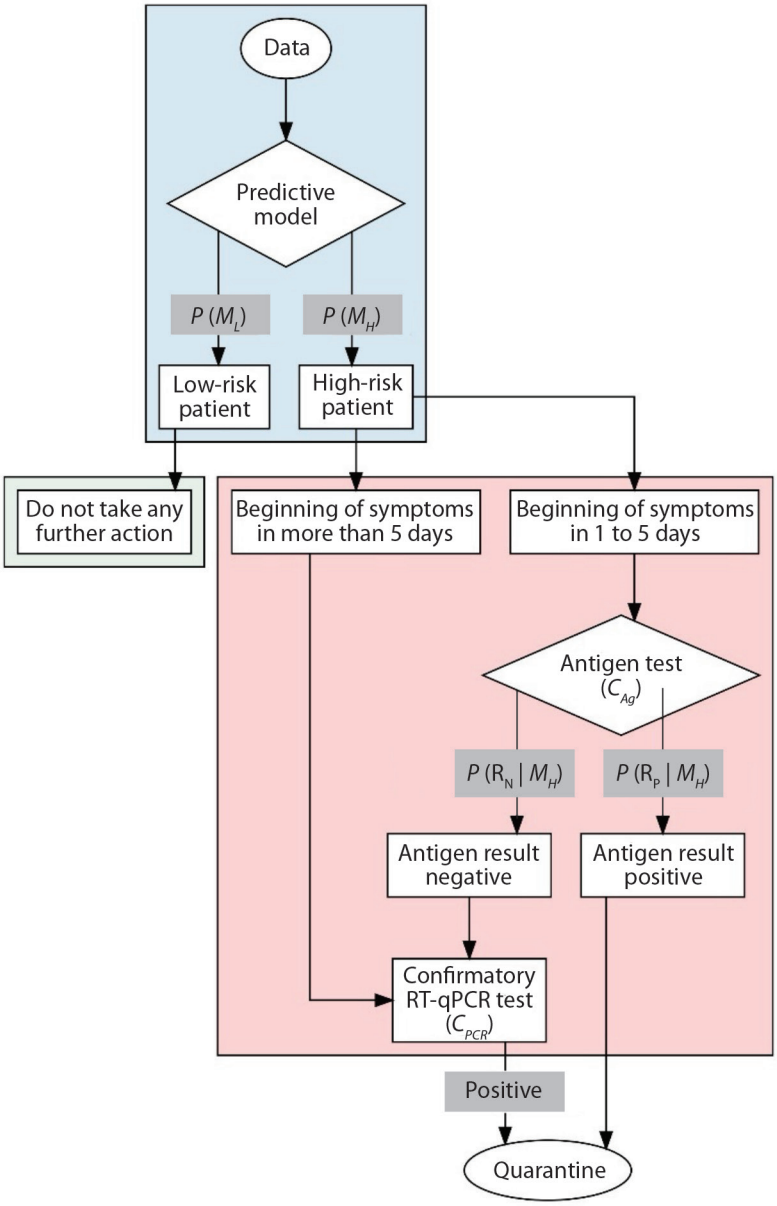

Fig. 1. Strategy 1, mass testing with RT-qPCR and antigen-based technologies.

1. Patients in the low-risk group (M_L) are not tested.
2. Patients in the high-risk group (M_H) are tested according to symptom onset: 1. Patients with less than 5 days since the beginning of symptoms (S_{-5}) undergo antigen-based testing.
   a. If the test is positive ($R_P^{Ag} \mid M_H$), the patient is declared positive. 2. If the test is negative ($R_N^{Ag} \mid M_H$), apply a confirmatory RT-qPCR (C^{PCR}) test.
   b. Apply an RT-qPCR test for patients with more than 5 days after symptom onset (S_{+5}).

To estimate overall costs, given by the number of tests per person and positive captured by the strategy, we define formally each component. First, the population at risk is given by

$$N_{M_H} = N \times P(M_H)$$

Since we only apply tests to high-risk patients, we can establish the number of tests applied for each technology,

$$T_1^{Ag} = N(_{M_H}) \times P(S_{-5}) \quad (2)$$
$$T_1^{PCR} = N(_{M_H}) \times (P(S_{+5}) + P(S_{+5}) P(R_N^{Ag} \vee M_H) \quad (3)$$

and then, the cost for this strategy becomes

$$C_1 = C^{Ag} T_1^{Ag} + C^{PCR} T_1^{PCR}$$

For the number of tests per person, we simply compute

$$T_1^{perperson} = \frac{T_1^{Ag} + T_1^{PCR}}{N_{M_H}}$$

For the number of positive cases reported, we estimate the population which has undergone either antigen-based or RT-qPCR testing and multiply it by the probability of having a positive result in each case. This estimate is

$$P_1^{Reported} = N_{M_H} P(S_{-5}) + P(R_P^{Ag} \mid M_H) + P(R_N^{Ag} \mid M_H) P(R_P^{PCR} \mid M_H) + N_{M_H} P(S_{+5}) P(R_P^{PCR} \mid M_H)$$

Strategy 2: pooling: We maintain all features from Strategy 1 but include a pooling component for the low-risk group (SMF1). The assumptions are

1. The high-risk group follows Strategy 1
2. Pooling is applied to the low-risk group (M_L), with a pool size of 5 samples.

Similarly, as before, we define the high-risk group as N_M and the low-risk one as $N_L = N \times P(M_L)$.

Since we did not modify the number of antigen-based tests, we use the same value T_1^{Ag} as in Equation (2). The RT-qPCR tests applied in this scenario disaggregate into two components. The first one is the same in Equation (3) called here T_1^{PCR}. For the second one, we need to determine the number of tests used in the pooling strategy.

The first element to establish is the prevalence among the low-risk subpopulation. Given the model, we need to estimate those individuals who are expected to be positive given the M_L classification. The negative predictive value of the model is given by $P(R_N \mid M_L)$. Therefore, we define the prevalence in this subgroup as the false omission rate estimated by $p_L = 1 - P(R_N \mid M_L)$.

If no loss of sensitivity occurs in the pooling technique and that the sensitivity of an RT-qPCR test is $P(R_P^{PCR} \mid D_P)$, we estimate the number of positive groups

$$P_2^{groups} = [1 - (1 - P(R_P^{PCR} \mid D_P) p_L)^g] N_{M_L}$$

with given a total test population of size N_{M_L} divided into groups of size g. The total number of tests required are

$$T_2^{Pooling} = g(\frac{1}{N_{M_L}} + P_2^{groups}).$$

Having those elements, we define the total number of RT-qPCR tests applied as

$$T_2^{PCR} = T_1^{PCR} + T_2^{Pooling}$$

and the total cost is, therefore,

$$C_2 = C^{Ag} T_2^{Ag} + C^{PCR} T_2^{PCR}$$

The estimate of the number of tests per person required becomes

$$T_2^{perperson} = \frac{T_2^{Ag} + T_2^{PCR}}{N}.$$

For the number of positive cases reported, we have again two components. First, we have the same number as Strategy 1 for the high-risk population. For the low-risk branch, we need to consider only those groups with positive test outcomes. We estimate the probability that their individual test in the Dorfman scheme attains a positive result. We therefore multiply the prediction outcome for pooling by the

positive predictive value of an RT-qPCR test and by the group size,

$$P_2^{Reported} = P_1^{Reported} + gP_2^{Pooling}\, P(R_P^{PCR} \mid M_L)$$

Strategy 3: consecutive antigen-based testing: Another alternative to increase the efficiency of Strategy 1 entails applying consecutive tests to the low-risk population (SMF2). This requires applying an antigen-based test to all low-risk patients, and in case of a positive result, a second confirmatory test should be performed within the next week or two. The measure is suboptimal due to false positive rates in current antigen-based testing technologies.

The assumptions behind this strategy are:

1. All patients in the high-risk group follows Strategy 1.
2. All patients in the low-risk group (M_L) undergo antigen-based testing.
   a. If the result is negative, we declare the person negative.
   b. If the result is positive, we apply a confirmatory antigen-based test within one or two weeks.

The number of antigen-based tests has two components, due to re-testing. For the high-risk population, we use the same value as Strategy 1, T_1^{Ag}. For the low-risk population, all patients undergo a first round of testing, and positive patients undergo a second one. At the end, the total number of antigen-based tests required during re-testing is

$$T_3^{Retest} = N_{M_L}\,(1 + P(R_P^{Ag} \mid M_L))$$

and the total number of antigen-based tests becomes

$$T_3^{Ag} = T_1^{Ag} + T_3^{Retest}$$

RT-qPCR tests applied are the same as the Scenario 1, T_1^{PCR}. The total cost due to testing for Strategy 3 is

$$C_3 = C^{Ag}\, T_3^{Ag} + C^{PCR}\, T_1^{PCR}.$$

For the number of tests per person, we simply estimate

$$T_3^{perperson} = \frac{T_3^{Ag} + T_3^{PCR}}{N}$$

Finally, the number of positive cases reported divides into two components. First, we have the same number of positive cases as Strategy 1 for the high-risk population. For the low-risk branch, we need to consider only the tests that were positives in the first or second round. This estimate is defined as

$$P_3^{Reported} = P_1^{Reported} + M_L\, P(R_P^{Ag} \mid N_L)^2.$$

Strategy 4: the role of saliva-based testing: Prior strategies model the current state of healthcare guidelines in Costa Rica, anchored in RT-qPCR tests as the main line of defense, which does not scale for mass testing purposes. Antigen-based testing has lower costs, but its low sensitivity makes confirmatory tests of negative results still necessary. An alternative solution is to include saliva-based RT-LAMP testing into the mix, as suggested by a prior study (Segura-Ulate et al., 2022). RT-LAMP and other saliva-based testing technologies reach values above 90 % for sensitivity and above 95 % for specificity and can be adapted quickly to new variants. In addition, the sampling process is inexpensive, requires lower biosafety standards and trained personnel than nasopharyngeal swabs.

The fourth strategy proposed here seeks to overcome the flaws of other technologies by targeting them to appropriate groups based on a data-driven assessment of individual patient risk. We first separate high-risk patients further into essential workers and other high-risk. For essential workers, an RT-qPCR test is mandatory to ensure continuity of services without risking high numbers of false positives or negatives. Other high-risk patients undergo saliva-based RT-LAMP testing, well suited to in particular for high peak waves and massive screening. To capture all positive cases, a confirmatory RT-LAMP should be performed over negative

cases. Finally, the low-risk group is subjected to antigen-based testing at home or in point-of-care (POC) centers. As with Strategy 3, all positive cases must confirm their result with a second test within one or two weeks.

The main assumptions behind this strategy are:

1. Essential workers are tested with RT-qPCR.
2. Patients in the high-risk group are tested with RT-LAMP.
   a. If the result is positive, we declare the person as positive.
   b. If the result is negative, we perform a confirmatory test by RT-LAMP.
3. Patients in the low-risk group () are tested with antigen-based tests.
   a. If the result is negative, we declare the person as negative.
   b. If the result is positive, we apply a confirmatory antigen-based test in one or two weeks.

Where M_E represents the class of essential workers on this Strategy. For simulation purposes, we set the proportion of essential workers at a fixed value of 1 %. The value is a conservative estimate based on the 1.25 % of total healthcare workers in Costa Rica: 2 470 in the Ministry of Health, 62 814 in the public social security from a total population of 5 213 374 inhabitants (Brenes-Camacho et al., 2013; Caja Costarricense de Seguro Social [CCSS], 2021; MINSA, 2021a). Therefore, we estimate the high- and low-risk groups with the remainder of the population,

$$N_{M_H} = (N\text{-}N_{M_E}) \times P(M_H) N_{M_L} = (N\text{-}N_{M_E}) \times P(M_L)$$

The number of tests applied in each case will depend on the technology. For RT-qPCR, we have

$$T_4^{PCR} = N_{M_E}$$

RT-LAMP tests only apply to the high-risk group, with a confirmatory test in case of a negative result,

$$T_4^{LAMP} = N_{M_H} + N_{M_H} P(R_N \mid M_H).$$

For antigen-based tests, the number is equal to that in Strategy 3, . The strategy total costs become

$$C_4 = C^{PCR}\, T_4^{PCR} + C^{LAMP}\, T_4^{LAMP} + C^{Ag}\, T_4^{Ag}$$

and for the number of tests per person, we estimate

$$T_4^{perperson} = \frac{T_4^{PCR} + T_4^{LAMP} + T_4^{Ag}}{N}$$

Finally, the number of positive cases can be decomposed into

$$P_4^{Report} = N_{M_E} P(R_P^{PCR}) + N_{M_H} P(R_P^{LAMP} \mid M_H) (1 + P(R_N^{LAMP} \mid M_H)) + N_{M_L} P(R_P^{Ag} \mid M_L)^2$$

RESULTS

In this section, we compare the strategies above according to their costs (C_i), number of tests per person ($T_i^{perperson}$) and number of positive cases reported ($P_i^{Reported}$) for i = 1, 2, 3. The total population used is N = 1 000 and the prevalence ranges from 0 to 30 %. The cost of an RT-qPCR test is set to $ 100 and an antigen-based test to $ 50. Across all figures, the red dashed line is the cost of applying an RT-qPCR test to each true infected. Formally, it is equal to 1 000 × $ 100 × $P(D_P)$. Those reported as positive correspond to the number of true infected individuals $ 100 × $P(D_P)$. For the number of tests per person, we set to the constant 1 indicating a baseline. Blue lines represent the percent of antigen-based tests used in each strategy according to the proportion of people showing symptoms for less than 5 days. From dark to light blue, we assume proportions of

25, 50 and 75 %. The primary *x* axis represents percent prevalence and the *y* axis varies per target: cost in dollars number of people or tests per person. Secondary axes show the model specificity and sensitivity used in each case. Our code is available in a GitHub repository for reproducibility purposes.

Costs: Computational experiments show that using the pre-classifier reduces the total cost by correctly identifying the high-risk individuals in Strategy 1 (SMG1). As the pre-classifier increases its predictive accuracy, cost decreases to only for those truly infected. Notice that specificity has a greater effect in reducing cost relative to sensitivity. Since this strategy excluded low-risk individuals, false negatives do not contribute to the overall cost. Conversely, false positive cases appear (i.e., false high-risk individuals), the strategy applies an antigen-based test with a confirmatory RT-qPCR in case of negative outcome.

Sensitivity and specificity modulate the effectiveness of the classifier to rebalance the overall cost structure depending on prevalence. Specificity determines the sign of the slope of the resulting curves, while sensitivity determines the percentage of antigen-based tests applied to the population. Proportionally, applying more antigen-based tests becomes more effective at prevalence values higher than 10 % with tests having high specificity (90 %) and medium to high sensitivity (60 %, 90 %).

For Strategy 2, false positive cases represent the largest cost factor (SMG2), similar to Strategy 1. However, individuals misclassified as low-risk individuals do not increase dramatically overall costs, since it becomes a natural an overhead already accounted for in the method. Misclassifying high-risk individuals leads to incorrectly applying Strategy 1 to a healthy individual, or to applying a pooling technique to a group with at least one infected individual.

We observe how the pre-classifier helps to reduce the total cost of identifying correctly the high-risk individuals. When the pre-classifier has high levels of sensitivity and specificity, we achieve outcomes similar to the Strategy 1 with a small overhead due to the cost introduced by pooling. Again, as the model becomes more accurate, this overhead decreases. Sensitivity and specificity play the same role as in Strategy 1.

In Strategy 3 (SMG3), total costs are higher than the Strategy 1 or Strategy 2 due to massive testing with antigen-based technologies for the low-risk group. Even if it is possible to classify correctly most of the population according to their risk, the minimum will be of at least $ 60 000 for each 1 000 individuals.

Finally, Strategy 4 has a similar cost structure compared with the pooling scheme in Strategy 2 (Fig. 2). Using maximally targeted technologies to each type of patient is similar to applying complex (and difficult) techniques like pooling. Given that we use antigen-based testing without the restriction of incubation periods, sensitivity is the only factor affecting the sign of the slope.

Positive Cases Reported: In the case of positive reported, Strategy 1 performs well with a prior classification. Even when ignoring low-risk individuals, we capture almost all true positives when the sensitivity and specificity of the model is 90 %. Sensitivity helps to discard potential true negatives because it has determined correctly the most possible positive cases. When the sensitivity is low, the strategy misses those true positives, who are thus left untested.

Strategy 2 includes the low-risk individuals (SMG4), with an increase in positive reported from the start, decreasing the number of mismatches. Even when the classifier has low sensitivity and specificity, pooling captures the infected individuals identified as low-risk at the expense of higher costs than only using RT-qPCR.

Strategy 3 (SMG5) increases detection of true positives even more regarding Strategy 1, specially at low prevalence contingent on reaching high sensitivity (90 %); the number of false negatives increases at high prevalence below this sensitivity value. A large group of infected individuals are declared as low-risk. Combined

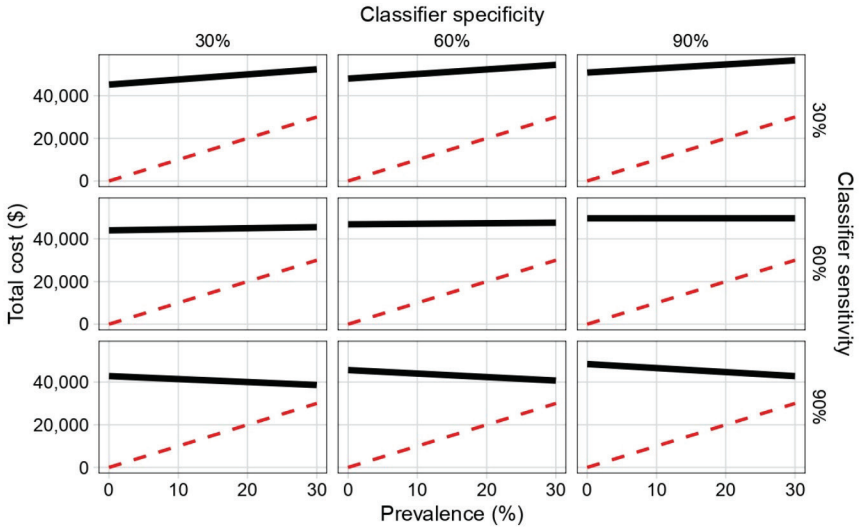

Fig. 2. Total cost according to Strategy 4. Introducing RT-LAMP testing significantly decreases total costs compared with all other strategies.

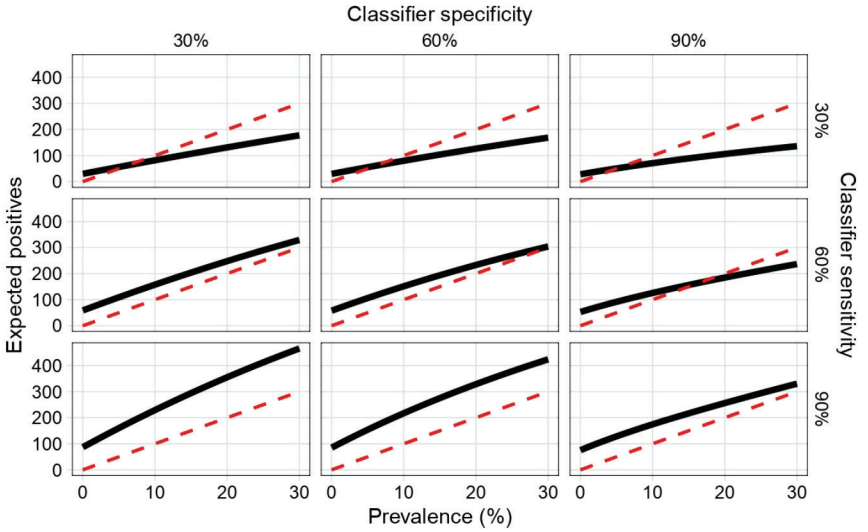

Fig. 3. Number of reported as positive according to Strategy 4. Introducing RT-LAMP testing yields concave curves at all sensitivity and specificity values. Outcomes are qualitatively similar to Strategies 1 and 3.

with the application of antigen-based tests which have lower sensitivity than RT-qPCR ones, the probability of capturing true positives is reduced.

Strategy 4 (Fig. 3) shows a similar pattern as Strategies 1 and 3. We observe that all functions are concave, implying improvements in detection as prevalence increases for sensitivity beyond 60 %. Even when the outcome of the classifier resembles that of Strategies 1 and 3, the robustness of the curves, indicates that RT-LAMP reduces the variability introduced by antigen-based testing.

Finally, we observe that sensitivity below 50 % appears to yield convex curves for the number of positives reported, while curves

corresponding to values above 50 % seem to be all concave for strategies 1-3; this is modulated by the number of antigen-based tests when applicable. This is significant, since it delineates a response function in terms of testing efforts needed at a certain value of prevalence given a current combination of resources. The higher the prevalence, the more likely it is to increase detection of true positives. Similarly, the more antigen-based tests are used, the more likely false negatives will appear. However, it also implies that the impact of RT-LAMP and similar technologies is significant, since even at low sensitivity of the classifier the effort function is concave.

Number of Tests per Person: For Strategy 1 (SMG6), the number of tests per person obtained with computational experiments is as expected. The less accurate the model in identifying high-risk individuals, the larger the number of tests needs to be spent, given the confirmatory mechanism of antigen-based testing against RT-qPCR. When the model is poorly fitted, the strategy spends around 1.2-1.7 tests per person. As the model sensitivity and specificity increases, the curves approach 1 at high prevalence. In all scenarios, the number of tests per person is high (1.2-1.7) at low prevalence, since negatives are the majority, and the strategy must spend two tests to confirm true positives.

When pooling is introduced (SMG7), a 0.5 reduction in average occurs when the model is correctly fitted regarding Strategy 1. Specificity controls the behavior of the curve in terms of convexity and slope. Low specificity increases misclassification of low-risk individuals, increasing the detection of true positives in the pooling technique.

The number of tests per person in Strategy 3 descends linearly as specificity increases (SMG8). Compared against Strategy 1, multiple testing can be reduced if the model is well-fitted. Strategy 2, in contrast, maintains better performance in this aspect. A similar pattern occurs in Strategy 4 (SMG9). However, it is worth noting that the number of tests per person remains relatively constant and close to 1 when the classifier shows high sensitivity and specificity in both Strategies. This is significant, since the resulting curve indicates scalability.

Performance across strategies: To compare the relative performance across different strategies, we establish two new quantities, which we call stock capacity () and detection efficiency (). To do so, we define an amortization index per Strategy $i \in \{1,2,3,4\}$

$$S_i = \left(\frac{T_i^{Total}}{C_i}\right)\left(\frac{N_E + N_H + N_L}{1000}\right),$$

where T_i^{Total} is the total number of tests performed by that Strategy. The left-most factor in S_i represents the buying power of testing per each dollar spent. The right-most factor scales the number to the effectively covered population. This is the case of Strategy 1 where it only considers the high-risk population. For instance, if S_i = 0.01 and the budget is $ 100 000, then the healthcare system can only afford $S_i \times 100\,000 = 100$ tests in total according to each strategy (a mix between RT-qPCR, Antigen and RT-LAMP). Meanwhile, for $i \in \{1,2,3,4\}$ the detection efficiency is

$$E_i = \left(\frac{P_i^{Report}}{C_i}\right)\left(\frac{N_E + N_H + N_L}{1000}\right).$$

We interpret the index as the capacity of each strategy to detect a positive case per each dollar spend. Similar to, S_i the number is scaled to the effective population covered. In the case of a value S_i = 0.001, and plans to spend $ 100 000 in the strategy, we can expect to detect $E_i \times \$\ 100\ 000 = 10$ positive cases. (Fig. 4) shows the values of S_i and E_i across all the strategies. We set here a fixed budget of $ 100 000. The red arrow (or point) represents a base case, with detection of $1\,000 \times P(D_p)$ positive cases spending $\$ (100 \times 1\,000 \times P(D_p))$ using only RT-qPCR tests. Arrows per strategy (i.e., hues of blue) indicate prevalence increase. Comparison of the four strategies regarding the stock capacity versus their detection efficiency. Arrows go from 0 (start) to 30 % (point)

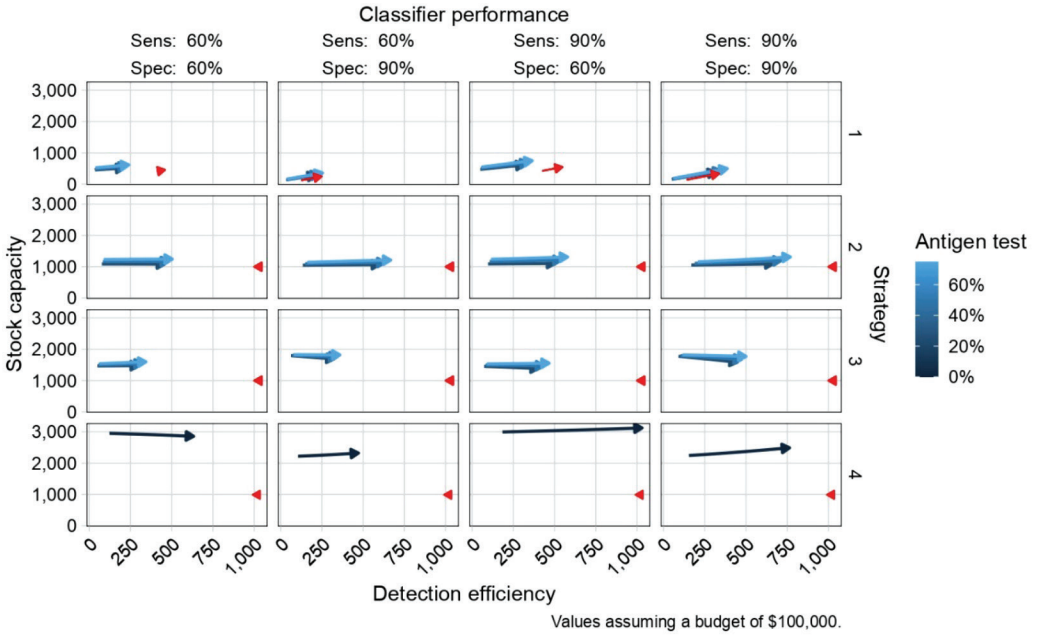

Fig. 4. Comparison of the four strategies with respect to the stock capacity versus their detection efficiency. Arrows go from 0 (start) to 30 % (point). Red arrows (or single triangle) represent the perfect case when there is 1 000 people × P (DP) positives, with a cost of 1 000 people × P (DP) × $ 100 using only RT-qPCR.

prevalence. Red arrows (or single triangle) represent the perfect case when there are positives, with a cost of $ (1 000 × $P(D_p)$ × 100) using only RT-qPCR.

Strategy 1 shows its weakness due to the small capacity to buy tests and overall effectiveness. In other words, there is no difference between using antigen-based testing or one using only RT-qPCR if only patients classified as high-risk are tested in the best-case scenario and significantly deteriorate when the classifier performs poorly. Meanwhile, Strategies 2 and 3 increase their capacities by covering the weak points from Strategy 1. Pooling (Strategy 2) increases the capacity of detection by maintaining the number of tests stable. Retesting (Strategy 3) is inefficient to capture positive cases even when the number of tests is still small. This is explained due to the low sensitivity of antigen-based tests, around 80 %.

Finally, Strategy 4 presents an augmented buying power of tests and detection efficiency. Targeting technologies to specific pre-selected groups appears to be the best strategy to maximize budget impacts across healthcare systems.

DISCUSSION

In this study, we investigated the theoretical impact of four different mass testing strategies for COVID-19 in Costa Rica, incorporating a pre-classification mechanism to improve testing efficiency. By simulating these strategies, we evaluated their overall costs, the number of positive cases detected, and the number of tests required per person. Our pre-classifier, based on machine learning methods, stratifies the population into high-risk and low-risk groups using variables such as social determinants, while maintaining patient privacy and information security. Furthermore, we reformulated the outcomes of each strategy in terms of purchasing power (i.e., stock capacity) and detection efficiency per dollar spent, providing a comprehensive analysis of resource allocation during a pandemic.

Our research gains significance when viewed in the context of previous public health interventions that have leveraged predictive modeling and mass testing. For instance, Jehi et al. (2020) developed a risk prediction model to prioritize COVID-19 testing based on individual patient characteristics, improving testing efficiency and resource allocation. Similarly, Schwab et al. (2020) reviewed clinical predictive models for COVID-19, highlighting the potential of machine learning approaches in enhancing patient triage and clinical decision-making. Our work extends these insights by quantitatively comparing different testing strategies and showing how a well-fitted classifier can significantly reduce costs and increase the detection of positive cases. Finally, Huang et al. (2022) proved that a data-driven testing strategy can detect 89.35 % of positives with only 48.1 7% of the available resources.

The introduction of a predictive model or classifier brings two strategic advantages. First, it can reduce overall costs, time and human efforts. Second, it increases information richness across the testing process. The first advantage relates to the system capacity to choose the best and cheaper technology according to each patient. If the model classifies individuals correctly, testing efforts can be optimized. Furthermore, healthcare systems can cover deficiencies present in one technology with the advantages of another (i.e., scalability), using the probabilistic prediction of the classifier as a triaging device while waiting for laboratory tests to finish and confirm or reject the result. Having more data, and consequently better prediction capabilities, allows clustering individuals into subgroups according to particular features such as their social, demographic or economic indicators and mobility patterns, among others. This information could lead healthcare authorities to adopt more personalized measures to cover certain vulnerable groups.

Our results show that all the strategies become more effective when the classifier -arguably a sophisticated machine learning method-is well-fitted, reaching sensitivity and specificity levels of 60 % or higher. We showed that sensitivity (identification of potential positives) plays a crucial role in reducing costs and increasing confirmation of positives. For the pooling scenario, specificity controls the number of tests per person.

One of the fundamental limitations of achieving a good fit for such models is access to high-quality individual data. The quality of data remains a challenge since the beginning of the pandemic, particularly in developing nations and emerging economies. For instance (Wynants et al., 2020) reviewed more than 126 000 studies related to COVID-19 prognosis prediction which only USA, Brazil and Mexico have formal studies about it. Available data tend to only reflect the reality of people who have undergone testing, and even when that is the case, datasets are biased by the administrative reality -and shortcomings- of the specific healthcare system. Therefore, we can expect a similar systematic bias in the classification process due to the different epidemiological moments across the pandemic. Testing increases during high-peak waves, confirming symptomatic patients and capturing asymptomatic nexus of them. When the pandemic wave passes and minimum cases are reached, the testing strategy tends to focus on confirming symptomatic cases arriving at clinical centers. During these periods, the real number of infected asymptomatic people remains unclear. In addition, overloading of the healthcare services impacts data production, which may be ready for consumption days or weeks later. This requires, as proposed, adjusting the model to correct for administrative and systematic lags.

Another set of limitations corresponds to the choice of potential classification models as well. We mention a non-exhaustive list of classification methods with their respective advantages and disadvantages. The classic logistic regression model is easy to implement, but the implicit assumptions and the inclusion of administrative lags in the data can negatively impact the interpretability of results due to an artificial increase in the number of coefficients; in this situation, a Ridge, or Lasso regularization could reduce their number. Another

option, if the data exhibits non-linearity, is to use a support vector machine (SVM), which can handle situations in which classes are not-linearly separable. The downside here is the computational cost during the training stage, which has to be performed a limited number of times as the pandemic evolves. Tree ensemble approaches are popular, including Random Forest, XGboost, and Gradient Boosting. In practice, these methods perform better than the mentioned classifiers but require fine-tuning of hyperparameters whose interpretation may not be direct. Finally, deep-learning algorithms can be used to fit the classifier at the expense of complexity and interpretability (Escobar et al., 2022).

We envision a series of challenges in the implementation of a classification system such as that described here. The main one is the adoption of machine learning assisted system by clinical and health policy authorities to triage the population before performing laboratory tests. While unforeseen clinical or ethical reasons may hamper the implementation of the model, the aim of this statistical approach is to become a companion instead of a competitor for healthcare providers. The advantage of classification-assisted triaging of patients in clinical contexts has been discussed and demonstrated in literature (Jehi et al., 2020; Schwab et al., 2020). Having some prior information about the possible test result can better prepare clinicians and staff to handle wave peaks efficiently, allocate resources more appropriately, and anticipate critical resource usage and patient mortality counterfactuals. Another challenge is the actual capacity of systems to triage patients. Even with an algorithm ready, further studies are needed about how to integrate it into workflows across medical centers and public health authorities. In the particular case of Costa Rica, the EDUS (Expediente Digital Único en Salud) system can serve as the channel to deliver results from the algorithm to laboratory technicians and physicians. However, creation of a new submodule will require testing, validation and data assurance in compliance with information security standards in the public health service (CCSS). Even if the EDUS system already collects most of the information about patients, the process of anonymizing, handling, securing, and ensuring responsible use of personal information must remain as a top priority. Finally, the attitude of the public around collection of information and its handling constitutes a challenge of uncertain proportions.

Our next step is to fit a classifier using both real and synthetic datasets. The EDUS is the main source of individual data of the Costa Rican public health. When a patient arrives at a medical appointment, physicians register the health status, diagnosis, demographic and related factors of each patient. During the COVID-19 pandemic, the tool was used to track down the symptoms across the population, to provide hot-lines for medical support and to validate the number of vaccines already applied. We believe this information source can be responsibly used further in benefit of all users. Its main advantage is the massive information density and patient coverage. Given the universal healthcare system in Costa Rica, information about a wide range of groups exists regardless of economic status. Another secondary corresponds to the Instituto Costarricense de Investigación y Enseñanza en Nutrición y Salud (INCIENSA: National Institute of Research and Education on Nutrition and Health). At the beginning of the pandemic, INCIENSA collected numerous COVID-19 samples alongside epidemiological and sociodemographic data of infected patients. Even if the diversity in this source is less than that of EDUS, it could be an important source to adjust the model.

Finally, we expect to develop synthetic datasets through simulation. Prior experience with agent-based modeling (Núñez-Corrales & Jakobsson, 2020) showed that it is possible to replicate features of epidemic waves and the effect of public policy measures *in silico*, to then overlay our strategies and determine performance under various scenarios and constraints; other methods exist and will be explored. These datasets can be openly shared across all relevant stakeholders without risking healthcare data leaks, while still

being representative of aggregate statistics of the underlying population.

These strategies may extend beyond COVID-19 to other infectious diseases with similar transmission characteristics, such as influenza and tuberculosis. Targeted testing combined with predictive modeling can improve the efficient use of testing resources and enable timely interventions. However, effectiveness depends on disease-specific factors like incubation periods, transmission rates, and modes of transmission. For diseases with longer incubation periods or different transmission modes (e.g., vector-borne diseases like malaria), adjustments to predictive models and testing protocols may be necessary. Future research should focus on adapting these methodologies to a broader range of pathogens to enhance applicability across various public health contexts.

Ethical statement: the authors declare that they all agree with this publication and made significant contributions; that there is no conflict of interest of any kind; and that we followed all pertinent ethical and legal procedures and requirements. All financial sources are fully and clearly stated in the acknowledgments section. A signed document has been filed in the journal archives.

See supplementary material
a17v73n1-suppl1

REFERENCES

Avaniss-Aghajani, E., Sarkissian, A., Fernando, F., & Avaniss-Aghajani, A. (2020). Validation of the hologic aptima unisex and multitest specimen collection kits used for endocervical and male urethral swab specimens (aptima swabs) for collection of samples from sars-cov-2-infected patients. *Journal of Clinical Microbiology*, *58*(8), e00753-20. https://doi.org/10.1128/JCM.00753-20

Beaudevin, C., Berlivet, L., Boudia, S., Bourgain, C., Cassier, M., Gaudillière, J.-P., & Löwy, I. (2021). 'Test, test, test!': Scarcity, tinkering, and testing policy early in the covid-19 epidemic in france. *Medicine Anthropology Theory*, *8*(2), 13–1 https://doi.org/10.17157/mat.8.2.5116

Brandeau, M. L. (2004). Allocating resources to control infectious diseases. In M. L. Brandeau, F. Sainfort, & W. P. Pierskalla (Eds.), *Operations research and health care: A handbook of methods and applications* (pp. 443–464). Springer. https://doi.org/10.1007/1-4020-8066-2_17

Brenes-Camacho, G., Araya-Umaña, O. M., González-Quesada, M. E., & Méndez-Fonseca, F. (2013). *Estimaciones y proyecciones de población por sexo y edad, 1950-2050*. Instituto Nacional de Estadística y Censos (INEC).

Caliendo, A. M., Gilbert, D. N., Ginocchio, C. C., Hanson, K. E., May, L., Quinn, T. C., Tenover, F. C., Alland, D., Blaschke, A. J., Bonomo, R. A., Carroll, K. C., Ferraro, M. J., Hirschhorn, L. R., Joseph, W. P., Karchmer, T., MacIntyre, A. T., Reller, L. B., & Jackson, A. F. (2013). Better tests, better care: Improved diagnostics for infectious diseases. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *57*(Supplement 3), S139–S170. https://doi.org/10.1093/cid/cit578

Caja Costarricense de Seguro Social. (2021). *Memoria Institucional 2021*. Caja Costarricense de Seguro Social (CCSS).

Center for Disease Control and Prevention (2020). *Interim Guidance for Antigen Testing for SARS-CoV-2*. Centers for Disease Control and Prevention.

Centers for Medicare & Medicaid Services (2020, October 15). *Changes Medicare Payment to Support Faster COVID-19 Diagnostic Testing* [Press Releases]. Centers for Medicare & Medicaid Services.

Comess, S., Wang, H., Holmes, S., & Donnat, C. (2022). Statistical modeling for practical pooled testing during the COVID-19 pandemic. *Statistical Science*, *37*(2), 229–250. https://doi.org/10.1214/22-STS857

Dorfman, R. (1943). The detection of defective members of large populations. *The Annals of Mathematical Statistics*, *14*(4), 436–440. https://doi.org/10.1214/aoms/1177731363

Du, Z., Pandey, A., Bai, Y., Fitzpatrick, M. C., Chinazzi, M., Pastore-Piontti, A. P., Lachmann, M., Vespignani, A., Cowling, B. J., Galvani, A. P., & Meyers, L. A. (2021). Comparative cost-effectiveness of SARS-CoV-2 testing strategies in the USA: A modelling study. *The Lancet Public Health*, *6*(3), e184–e191. https://doi.org/10.1016/S2468-2667(21)00002-5

Escobar, M., Jeanneret, G., Bravo-Sánchez, L., Castillo, A., Gómez, C., Valderrama, D., Roa, M., Martínez, J., Madrid-Wolff, J., Cepeda, M., Guevara-Suarez, M., Sarmiento, O. L., Medaglia, A. L., Forero-Shelton, M., Velasco, M., Pedraza, J. M., Laajaj, R., Restrepo, S., & Arbelaez, P. (2022). Smart pooling: AI-powered COVID-19 informative group testing. *Scientific Reports*, *12*(1), 6519. https://doi.org/10.1038/s41598-022-10128-9

García-Puerta, Y. E., Vásquez-Brenes, P. A., Calvo-Alpízar, J. G., Barboza-Chinchilla, L. A., Sánchez-Peña, F. A., Rivas-Chaves, T., Mery-Valdovinos, G. A., & Pérez-Rosales, M. D. (2023). *Modelos matemáticos y análisis estadísticos implementados para el estudio de Covid19 en Costa Rica*. Organización Panamericana de la Salud.

Grantz, K. H., Lee, E. C., McGowan, L. D., Lee, K. H., Metcalf, C. J. E., Gurley, E. S., & Lessler, J. (2021). Maximizing and evaluating the impact of test-trace-isolate programs: A modeling study. *PLoS Medicine*, *18*(4), e1003585. https://doi.org/10.1371/journal.pmed.1003585

Haigh, K. Z., & Gandhi, M. (2021). COVID-19 Mitigation with appropriate safety measures in an essential workplace: lessons for opening work settings in the United States during COVID-19. *Open Forum Infectious Diseases*, *8*(4), ofab086. https://doi.org/10.1093/ofid/ofab086

He, X., Lau, E. H. Y., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y. C., Wong, J. Y., Guan, Y., Tan, X., Mo, X., Chen, Y., Liao, B., Chen, W., Hu, F., Zhang, Q., Zhong, M., Wu, Y., Zhao, L., … Leung, G. M. (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine*, *26*, 672–675 https://doi.org/10.1038/s41591-020-0869-5

Huang, C., Wang, M., Rafaqat, W., Shabbir, S., Lian, L., Zhang, J., Lo, S., & Song, W. (2022). Data-driven test strategy for COVID-19 using machine learning: A study in Lahore, Pakistan. *Socio-Economic Planning Sciences*, *80*, 101091. https://doi.org/10.1016/j.seps.2021.101091

Jehi, L., Ji, X., Milinovich, A., Erzurum, S., Rubin, B. P., Gordon, S., Young, J. B., & Kattan, M. W. (2020). Individualizing risk prediction for positive coronavirus disease 2019 testing: results from 11 672 patients. *Chest*, *158*(4), 1364–1375. https://doi.org/10.1016/j.chest.2020.05.580

Kırkızlar, E., Faissol, D. M., Griffin, P. M., & Swann, J. L. (2010). Timing of testing and treatment for asymptomatic diseases. *Mathematical Biosciences*, *226*(1), 28–37. https://doi.org/10.1016/j.mbs.2010.03.007

Larremore, D. B., Wilder, B., Lester, E., Shehata, S., Burke, J. M., Hay, J. A., Tambe, M., Mina, M. J., & Parker, R. (2021). Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening. *Science Advances*, *7*(1), eabd5393. https://doi.org/10.1126/sciadv.abd5393

Mercer, T. R., & Salit, M. (2021). Testing at scale during the COVID-19 pandemic. *Nature Reviews Genetics*, *22*(7), 415–426. https://doi.org/10.1038/s41576-021-00360-w

Millioni, R., & Mortarino, C. (2021). Test Groups, not individuals: A review of the pooling approaches for SARS-CoV-2 diagnosis. *Diagnostics*, *11*(1), 68. https://doi.org/10.3390/diagnostics11010068

Ministerio de Salud Costa Rica. (2021a). *Informe de Gestión 2020-2021*. Ministerio de Salud (MINSA).

Ministerio de Salud Costa Rica. (2021b). *LS-SS-012. Lineamientos generales para el uso de pruebas alternativas (antígeno, pruebas moleculares isotérmicas) al estándar de oro (RT- PCR) para el diagnóstico de COVID-19*. Ministerio de Salud (MINSA).

Ministerio de Salud Costa Rica. (2022). *LS-VS-001. Lineamientos Nacionales para la Vigilancia de la enfermedad COVID-19 (versión 26) (LS-VS-001-26)*. Ministerio de Salud (MINSA).

Nagura-Ikeda, M., Imai, K., Tabata, S., Miyoshi, K., Murahara, N., Mizuno, T., Horiuchi, M., Kato, K., Imoto, Y., Iwata, M., Mimura, S., Ito, T., Tamura, K., & Kato, Y. (2020). Clinical evaluation of self-collected saliva by quantitative reverse transcription-PCR (RT-qPCR), direct RT-qPCR, reverse transcription-loop-mediated isothermal amplification, and a rapid antigen test to diagnose COVID-19. *Journal of Clinical Microbiology*, *58*(9), e01438-20. https://doi.org/10.1128/JCM.01438-20

Núñez-Corrales, S., & Jakobsson, E. (2020). The epidemiology workbench: A tool for communities to strategize in response to COVID-19 and other infectious diseases. *medRxiv*, *2020*, 20159798. https://doi.org/10.1101/2020.07.22.20159798

Oran, D. P., & Topol, E. J. (2020). Prevalence of asymptomatic SARS-CoV-2 infection: A narrative review. *Annals of Internal Medicine*, *173*(5), 362–367. https://doi.org/10.7326/M20-3012

Oran, D. P., & Topol, E. J. (2021). The proportion of SARS-CoV-2 infections that are asymptomatic: A systematic review. *Annals of Internal Medicine*, *174*(5), 655–662. https://doi.org/10.7326/M20-6976

Österdahl, M. F., Lee, K. A., Lochlainn, M. N., Wilson, S., Douthwaite, S., Horsfall, R., Sheedy, A., Goldenberg, S. D., Stanley, C. J., Spector, T. D., & Steves, C. J. (2020). Detecting SARS-CoV-2 at point of care: Preliminary data comparing loop-mediated isothermal amplification (LAMP) to polymerase chain reaction (PCR). *BMC Infectious Diseases*, *20*, 783. https://doi.org/10.1186/s12879-020-05484-8

Peeling, R. W., Olliaro, P. L., Boeras, D. I., & Fongwen, N. (2021). Scaling up COVID-19 rapid antigen tests: Promises and challenges. *The Lancet Infectious Diseases*, *21*(9), e290–e295. https://doi.org/10.1016/S1473-3099(21)00048-7

Plantes, P. J., Fragala, M. S., Clarke, C., Goldberg, Z. N., Radcliff, J., & Goldberg, S. E. (2021). Model for mitigation of workplace transmission of COVID-19 through population-based testing and surveillance.

Population Health Management, *24*(Supplement 1), S16–S25. https://doi.org/10.1089/pop.2020.0322

Prado, A. M., Pearson, A. A., Mora-García, C. A., Guha, M., Tanugi-Carresse, A. C., Mullen, L., Bennett, S., & Nuzzo, J. (2023). *COVID-19 response and maintenance of essential health services in Costa Rica*. Exemplars in Global Health.

Rubinstein, A. (2025). How Latin American health care systems will respond to the next crises? Lessons and challenges after the COVID-19 pandemic. *Archives of Medical Research*, *56*(1), 103069. https://doi.org/10.1016/j.arcmed.2024.103069

Sah, P., Fitzpatrick, M. C., Zimmer, C. F., Abdollahi, E., Juden-Kelly, L., Moghadas, S. M., Singer, B. H., & Galvani, A. P. (2021). Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis. *Proceedings of the National Academy of Sciences*, *118*(34), e2109229118. https://doi.org/10.1073/pnas.2109229118

Saidani, M., Kim, H., & Kim, J. (2021). Designing optimal COVID-19 testing stations locally: A discrete event simulation model applied on a university campus. *PLoS ONE*, *16*(6), e0253869. https://doi.org/10.1371/journal.pone.0253869

Sandmann, F. G., White, P. J., Ramsay, M., & Jit, M. (2020). Optimizing benefits of testing key workers for infection with SARS-CoV-2: A mathematical modeling analysis. *Clinical Infectious Diseases*, *71*(12), 3196–3203. https://doi.org/10.1093/cid/ciaa901

Schwab, P., DuMont-Schütte, A., Dietz, B., & Bauer, S. (2020). Clinical predictive models for COVID-19: Systematic study. *Journal of Medical Internet Research*, *22*(10), e21439. https://doi.org/10.2196/21439

Segura-Ulate, I., Bolívar-González, A., Madrigal-Redondo, G., Nuñez-Corrales, S., & Gatica-Arias, A. (2022). Reverse-Transcription Loop-Mediated Isothermal Amplification and alternative protocols for lower cost, large-scale COVID-19 testing: Lessons from an emerging economy. *Revista de Biología Tropical*, *70*(1), 173–189. https://doi.org/10.15517/rev.biol.trop..v70i1.47407

Tindale, L. C., Stockdale, J. E., Coombe, M., Garlock, E. S., Lau, W. Y. V., Saraswat, M., Zhang, L., Chen, D., Wallinga, J., & Colijn, C. (2020). Evidence for transmission of COVID-19 prior to symptom onset. *eLife*, *9*, e57149. https://doi.org/10.7554/eLife.57149

Watkins, A. E., Fenichel, E. P., Weinberger, D. M., Vogels, C. B. F., Brackney, D. E., Casanovas-Massana, A., Campbell, M., Fournier, J., Bermejo, S., Datta, R., Dela Cruz, C. S., Farhadian, S. F., Iwasaki, A., Ko, A. I., Grubaugh, N. D., & Wyllie, A. L. (2021). Increased SARS-CoV-2 testing capacity with pooled saliva samples. *Emerging Infectious Diseases*, *27*(4), 1184–1187 https://doi.org/10.3201/eid2704.204200

Wiencek, J. R., Head, C. L., Sifri, C. D., & Parsons, A. S. (2020). Clinical ordering practices of the SARS-CoV-2 antibody test at a large academic medical center. *Open Forum Infectious Diseases*, *7*(10), ofaa406. https://doi.org/10.1093/ofid/ofaa406

Wynants, L., Calster, B. V., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., Albu, E., Arshi, B., Bellou, V., Bonten, M. M. J., Dahly, D. L., Damen, J. A., Debray, T. P. A., Jong, V. M. T., De Vos, M., Dhiman, P., Ensor, J., Gao, S., Haller, M. C., … van Smeden, M. (2020). Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. *BMJ*, *369*, m1328. https://doi.org/10.1136/bmj.m1328

Yang, S., & Rothman, R. E. (2004). PCR-based diagnostics for infectious diseases: Uses, limitations, and future applications in acute-care settings. *The Lancet Infectious Diseases*, *4*(6), 337–348. https://doi.org/10.1016/S1473-3099(04)01044-8

Zhang, J. J., Dong, X., Liu, G. H., & Gao, Y. (2023). Risk and protective factors for COVID-19 morbidity, severity, and mortality. *Clinical Reviews in Allergy & Immunology*, *64*, 90–107. https://doi.org/10.1007/s12016-022-08921-5