



Application of pooled testing in estimating the prevalence of COVID-19

Pritha Guha¹ · Apratim Guha¹ · Tathagata Bandyopadhyay²

Received: 1 January 2021 / Revised: 27 June 2021 / Accepted: 23 July 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Testing at a mass scale has been widely accepted as an effective way to contain the spread of the SARS-CoV-2 Virus. In the initial stages, the shortage of test kits severely restricted mass-scale testing. Pooled testing was offered as a partial solution to this problem. However, it is a relatively lesser-known fact that pooled testing can also result in significant gains, both in terms of cost savings as well as measurement accuracy, in prevalence estimation surveys. We review here the statistical theory of pooled testing for screening as well as for prevalence estimation. We study the impact of the diagnostic errors, and misspecification of the sensitivity and the specificity on the performances of the pooled as well as individual testing procedures. Our investigation clarifies some of the issues hotly debated in the context of COVID-19 and shows the potential gains for the Indian Council for Medical Research (ICMR) in using a pooled sampling for their upcoming COVID-19 prevalence surveys.

Keywords Pooled testing · Prevalence estimation · Imperfect test · COVID-19 · Sensitivity · Specificity

1 Introduction

The COVID-19 pandemic, caused by the virus SARS-CoV-2, has been devastating the world now for more than a year. It has created catastrophic social, economic, and health consequences for countries around the world. For an emerging economy like India, it has posed a formidable challenge to the policymakers to contain its spread at a level so that the health care infrastructure is not overwhelmed. While the first COVID-19 case

✉ Tathagata Bandyopadhyay
tathagata@iima.ac.in

Pritha Guha
pritha@xlri.ac.in

Apratim Guha
apratim@xlri.ac.in

¹ XLRI, Xavier School of Management, Jamshedpur, India

² Indian Institute of Management, Ahmedabad, India

in India was detected on January 30th, 2020, as of 15th December 2020, the number infected has swelled to nearly 10 million, and more than 140,000 deaths have occurred due to COVID-19. India is right now placed second in the world in terms of the total number of cases, right behind the United States, and is battling the rapid rise in the number of cases and resultant unprecedented economic and social consequences. While the government is desperately looking for an exit route to bring the economy back to normalcy, it is facing considerable challenges in the implementation of social distancing, mass testing, contact tracing, and quarantining which are essential to contain the spread of the virus. Given that a large percentage of COVID-19 cases are asymptomatic and are responsible for spreading the virus (Aguilar et al. 2020; Medical Xpress 2020; Yu et al. 2020), it has been advised by the World Health Organization (WHO) that the most effective way to control the spread of the disease is to test as many people as possible. For example, their guidance document on implementing and adjusting public health and social measures suggests testing at least one person per 1000 population per week (WHO 2020a). This was found to be effective in containing the spread of the virus in the initial phases in countries like South Korea, Singapore, and China. However, the high cost and the short supply of testing kits (Alluri and Pathi 2020) are the main impediments for a country like India in conducting mass testing. In order to ramp up the testing efforts, some experts (Bilder et al. 2020; Hanel et al. 2020; Lakdawalla et al. 2020; Pouwels et al. 2020) recommended the use of pooled testing technique, a technique originally proposed by Dorfman (1943) to increase the speed, and reduce the cost, of screening the US army recruits for syphilis during the Second World War. In April 2020, the Indian Council for Medical Research (ICMR) gave its approval (SNS 2020) and issued detailed guidelines (ICMR 2020) for carrying out pooled testing for screening purposes, which was subsequently followed by the Indian states.

In addition to the attempts made towards the containment of the disease, it is also of vital importance to understand the true spread of the disease. To understand the spread of Covid-19 in India, ICMR has conducted two serosurveys (till December 2020), the first during May–June 2020 (Murhekar et al. 2020), and a second one during September–October 2020. A third survey is being planned for the early part of 2021 (Mudur 2020). Such surveys are extremely important to plan for the future requirement of resources to fight the disease effectively. Moreover, as observed by many (Bendavid et al. 2021; Li et al. 2020), the estimation of the prevalence of the SARS-CoV-2 virus is crucial for an accurate estimation of the fatality rate. The initial estimate of the fatality rate given by WHO (2020b) was scandalously high because it was based only on deaths among symptomatic patients. Later, it was revised downward after taking into consideration the fact that a significant percentage of the infected population is asymptomatic, and hence remains undetected. If 80% of the cases are asymptomatic, then the fatality rate computed only from the symptomatic cases would be five times higher than the true fatality rate. Such misinformation may create unnecessary panic in the minds of policymakers, which may lead to wrong decisions. Several serology surveys have been carried out for the estimation of the prevalence of the SARS-CoV-2 virus (Sempos and Tian 2020), and the results suggest that the overestimation of the fatality rate in the population may be between fifty to eighty-five times the original fatality rate. The first serosurvey by ICMR in April–May 2020 in India estimated the true spread of the disease to be 6,468,388 (95% CI: 3,829,029–11,199,423). When compared with the fact that the laboratory-confirmed cumulative number of cases was only around 395,000 even on June 20th, 2020 (Murhekar et al. 2020), it may be said that the findings of the Indian serosurvey were largely in the lines of the findings of Sempos and Tian (2020).

The challenges to be faced by the ICMR in conducting a major survey like this are many. A good estimate requires a large sample size, lack of which may result in a very wide confidence interval, as reported above. Further testing for coronavirus is expensive. As per a recent report (Ghosh and Dasgupta 2020), the cost of the gold standard test for coronavirus, the reverse-transcription polymerase chain reaction test (RT-PCR test), costs the government around Rs. 1 Crore for 10,000 tests, which comes to around Rs. 1000 per test. The ICMR serosurveys used multiple tests to increase their efficacy: the samples were pre-processed using centrifugation, then subjected to commercial ELISA tests, and followed by Euroimmun SARS-CoV-2 ELISA tests for reconfirmation of positive cases (Murhekar et al. 2020). While the cost of the procedure was never reported, a study of the cost for these tests in commercial establishments in India suggests that the per-sample cost would have been close to that of the RT-PCR tests. Hence, controlling the cost, while increasing the sample size, will be vital for the subsequent surveys.

Secondly, as widely reported, although the RT-PCR tests themselves are highly accurate, the sample collection and preservation procedures for the coronavirus tests are not very simple, and introduce significant errors, both false positives, and false negatives, in the RT-PCR test results (Lakdawalla et al. 2020). The sensitivity and specificity of the ELISA assays in laboratory conditions were 92.4 and 97.9 percent, respectively. When we consider that serum was separation was done in local health facilities and transported to the laboratories in the designated ICMR institutes under cold chains, as reported by Murhekar et al. (2020), it is clear that the ELISA tests would also suffer the preservation issues. This also can pose significant challenges in estimating the correct prevalence rates.

It is interesting to observe that most experts recommend the use of pooled testing only for screening purpose. However, the data collected by pooled testing can also be effectively utilized for the estimation of the prevalence of the virus, which is a later development in pooled testing research (Roy and Banerjee 2019). This fact is not as widely known as the use of pooled testing for screening purposes in the general scientific community. We will revisit these results in Sect. 4, and show that pooled testing can be very effective in controlling cost, as well as controlling for estimation errors arising from false positives.

This article mainly serves two purposes. First, we revisit the statistical theory of the pooled testing procedure for screening, and also for the estimation of prevalence using basic probability theory (Feller 1968). Next, we discuss some practical issues arising out of the fact that these tests are imperfect. In other words, the tests may yield false positive and/or false negative results. Naturally, it raises an important practical question: what is the impact of these diagnostic errors on screening as well as on prevalence estimation? What can we do if we do not have an accurate idea about the rate of diagnostic errors? We study these effects using some hypothetical scenarios and discuss their implications in the context of the COVID-19 pandemic.

The rest of the article is divided as follows. In Sect. 2 we present Dorfman's methodology of pooled testing. In Sect. 3 we discuss the consequences when the test is imperfect. In Sect. 4, we discuss the estimation of prevalence using pooled data and compare it with the estimation procedure with individual data. Finally, we conclude with some remarks in Sect. 5.

2 Dorfman pooled testing technique

Pooled testing was originally introduced by Dorfman (1943) in infectious disease studies to reduce the cost and increase the speed of data collection. Typically, for an infectious disease, a sample of blood or urine is tested for the presence or absence of the disease. Rarer the disease, the more effective is the pooled testing technique. Dorfman's technique runs as follows. Suppose n samples of blood are to be tested for the presence or absence of a disease. A sample is either positive (disease present) or negative (disease absent). We attach a dichotomous variable Y to each sample: we assume $Y=1$ if the sample is positive, and $Y=0$ if it is negative. Instead of testing the samples individually, the samples are pooled into J groups of sizes n_1, \dots, n_j . Let Y_{ij} denote the Y -value for the i th sample in the j th group. The samples in each group are pooled and tested. Instead of observing Y_{ij} , in the pooled testing set-up, we observe $Y_j^* = \max_{i=1, \dots, n_j} Y_{ij}$, the maximum of $Y_{1j}, Y_{2j}, \dots, Y_{n_j j}$. Notice that this value is 0 if all Y_{ij} 's in the group equal to 0, and is 1 if even one of them is equal to 1. In other words, $Y_j^* = 0$ if and only if all samples in the j th group are negative, i.e., $Y_{ij} = 0, i = 1, \dots, n_j$ and $Y_j^* = 1$, if and only if at least one sample in the j th group is positive, i.e., $Y_{ij} = 1$ for at least one $i = 1, \dots, n_j$. If for a group, the test outcome is positive, then the samples in the group are individually tested to detect the positive samples, and hence the diseased individuals. If the result is negative for a group, then no further testing is required. Thus, for a positive test outcome, the number of tests required is one more than the group size, and for a negative test outcome, a single test is enough. If the chance of a positive test outcome for a group is small, in other words, the prevalence of the disease is low, then the average number of tests required would be much smaller in a group testing set-up than individual testing. Thus, it leads to considerable savings in the cost and time of testing. The process diagram is presented in Fig. 1 for a group of size k .

Now, suppose that the prevalence of a disease is p , and assume without loss of any generality $n_j = k$, i.e., all groups are of equal size, and consequently $n = Jk$. Notice that, $P(Y_j^* = 0) = (1 - p)^k$ and $P(Y_j^* = 1) = 1 - (1 - p)^k$. Further, suppose N_j denotes the number of tests for the j th group. Clearly, it is equal to $k + 1$ if $Y_j^* = 1$, and 1 if $Y_j^* = 0$. Thus, the expected number of tests for the j th group, which we henceforth write as $E(N_j)$, is given by

$$E(N_j) = (k + 1) - k(1 - p)^k. \quad (1)$$

If the value of p is small then $E(N_j)$ is close to 1 even for a sufficiently large group size k , and thus leading to a substantial saving in time and cost than individual testing.

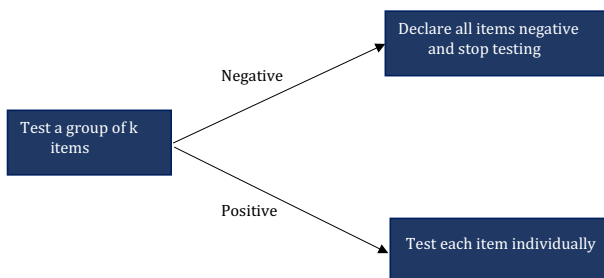


Fig. 1 Process diagram of Dorfman's algorithm for pooled testing

Notice that the technique is mainly useful when the prevalence is low. In particular, by rearranging (1), we can see that we require $p < 1 - \frac{1}{k^{1/k}}$ for a group size k to be beneficial. A little bit of trial and error shows that no benefit can be achieved if $p \geq 1 - \frac{1}{3^3} = 0.3066$, a fact first noted by Finucan (1964). Note, however, that this limit is valid only for perfect tests. Interesting situations emerge when the tests are less than perfect, which we discuss in the next section.

The simplicity of the idea, and its easy implementation, have led to its wide applicability in different fields, see Roy and Banerjee (2019) and the references therein. Its use in the screening of sexually transmitted diseases like chlamydia, gonorrhea (Lewis et al. 2012), and HIV (Litvak et al. 1994; Wein and Zenios 1996) is worth mentioning. American Red Cross and the European Blood Alliance routinely use it to screen donated blood for infectious diseases, see Aprahamian, Bish, and Bish (2019) and Bilder (2019). Given the huge shortage of test kits around the world, researchers worldwide are recommending the use of pooled testing for COVID-19 (Bilder et al. 2020; Abdalhamid et al. 2020; Lohse et al. 2020; Yelin et al. 2020).

3 Pooled testing for screening when the test is imperfect

Dorfman (1943) proposed the pooled testing technique assuming the test to be perfect, i.e., the chance of a false positive or a false negative test result to be zero. However, in reality, tests are far from perfect. In testing for the SARS-CoV-2 virus, two types of tests are mainly in use. As per Ferran (2020), “the first is a reverse-transcription polymerase chain reaction test or RT-PCR. This is the most common diagnostic test used to identify people currently infected with SARS-CoV-2. It works by detecting viral RNA in a person’s cells – most often collected from their nose. The second test being used is called a serological or antibody test. This test looks at a person’s blood to see if they have produced antibodies for the SARS-CoV-2 virus. If a test finds these antibodies, it means a person was infected and made antibodies in response”. How accurate are these tests? The accuracy of a test is measured by two numbers: sensitivity and specificity attached to it. If a test has a 5% chance of a false negative (positive) outcome, its sensitivity (specificity) is 95%. Although in a laboratory setting it is observed that the RT-PCR test has high sensitivity and specificity, in the real world testing condition the sensitivity and the specificity are usually much lower because of the testing conditions, the method of sample collection and the sample preservation technique are far from perfect. For example, in the real-world testing condition, the sensitivity of the RT-PCR test ranges between 66 and 80% (Ferran 2020). Thus, out of three infected people on average one may test negative. The sensitivity and specificity of the antibody tests are also found to be very high in a laboratory setting, but, then again, in real-world conditions, the accuracy is bound to suffer.

Using the same notations as above, we now find $E(N_j)$ assuming that the sensitivity and specificity of the test are S_e and S_p , respectively. Notice that due to the possible errors in testing, for the j th group, we now observe \hat{Y}_j^* instead of Y_j^* , a surrogate (or proxy) for Y_j^* , where S_e is the conditional probability of the test outcome coming positive for group j when it should, i.e., $P(\hat{Y}_j^* = 1 | Y_j^* = 1) = S_e$ and $P(\hat{Y}_j^* = 0 | Y_j^* = 0) = S_p$. A simple probability calculation then yields $P(\hat{Y}_j^* = 1) = P(Y_j^* = 1)S_e + P(Y_j^* = 0)(1 - S_p) = S_e + (1 - p)^k(1 - S_p - S_e)$, and hence,

$$E(N_j) = 1 + k[S_e + (1 - p)^k(1 - S_p - S_e)]. \tag{2}$$

The expected number of tests for n people would then be $J \times E(N_j)$, and hence compared to the individual testing, the reduction in the number of tests achieved by using group testing is $J \times (k - E(N_j))$. This reduction is a measure of the *efficiency* of the pooled testing procedure. Also, notice that for $S_p = S_e = 1$, i.e., when the test is perfect, Eq. (2) reduces to (1).

In the previous section, we saw that for a perfect test, i.e., when $S_p = S_e = 1$, we need $p < 0.3066$. However, when the test is less than perfect, this limit can vary drastically. We now provide an illustration. First, we note that, from (2), the maximum value of p for which group testing can be beneficial, satisfies

$$(1 - p)^k > \frac{\frac{1}{k} + S_e - 1}{S_p + S_e - 1}.$$

Now, this means that when $S_e < 1$, and $S_p + S_e > 1$, the left-hand side of the above inequality can be made negative by choosing a suitably large value of k , and hence *for any value of p* , there can be a benefit in performing a group test. For illustration, if $S_p = S_e = 0.9$, and $p = 0.9$, we require $0.9^k > \frac{1.25}{k} - 0.125$. The right-hand side of this inequality is clearly less or equal to zero for $k \geq 10$, and hence this inequality is achieved for all $k \geq 10$. (In fact, by trial and error, we can see that this inequality is valid for all positive integers $k \geq 2$.) Hence, for an imperfect test, it may be possible to see a benefit in group testing even when the disease is widely prevalent. It is, of course, true that more significant benefits are achieved when the values of p are relatively lower.

For a perfect test, there is no chance of misclassification of an infected as non-infected or vice versa, and hence the efficiency of the pooled testing procedure is measured only by $E(N_j)$. However, for an imperfect test, $E(N_j)$ is not enough to judge the efficiency of the pooled testing procedure. One also needs to measure the sensitivity and specificity of the pooled testing procedure, and most importantly, the probabilities of the diagnostic errors.

Let $T+$ and $T-$ respectively denote the events (situations) that an individual is tested positive and negative, respectively, by the pooled testing procedure. Further, let I (I^c) denote the events that an individual is infected (uninfected). Then one can easily check that the sensitivity, say, S_e^D , and specificity, say, S_p^D , of the Dorfman’s pooled testing procedure, are given by

$$S_e^D = P(T^+|I) = S_e^2, \tag{3}$$

and

$$S_p^D = P(T^-|I^c) = 1 - (1 - S_p)[S_e + (1 - S_p - S_e)(1 - p)^{k-1}] \tag{4}$$

respectively. The proofs of (3)–(4) are given in the Appendix.

Notice that S_e^D depends only on S_e while S_p^D depends on all three parameters S_e , S_p and p . As stated above, S_e^D and S_p^D provide useful information about the performance of the pooled testing procedure. However, the posterior or inverse probabilities $P(I^c|T^+)$ and $P(I|T^-)$, called the False Positive Predictive Value (FPPV) and the False Negative Predictive Value (FNPV) respectively, are often critical for assessing the performance of the test in a real-life situation. These are measures of diagnostic errors. The *FPPV* and the *FNPV* represent the proportion of misclassified individuals among those who are tested positive, and among those who are tested negative, respectively. In the context of testing for sera

samples of HIV virus, Litvak et al. (1994) proposed that the five key elements to be used to capture the overall performance of a pooled testing procedure should be: $E(N_j)$, S_e^D , S_p^D , $FPPV$ and $FNPV$.

Now, using Bayes' theorem (Feller 1968), one can easily show that

$$FPPV = \frac{(1 - p)(1 - S_p^D)}{(1 - p)(1 - S_p^D) + pS_e^D}; FNPV = \frac{p(1 - S_e^D)}{p(1 - S_e^D) + (1 - p)S_p^D}. \tag{5}$$

Notice that for individual testing $FPPV$ and $FNPV$, denoted henceforth $FPPVI$ and $FNPVI$ respectively, are obtained by substituting S_e^D and S_p^D by Se and S_p respectively in (5). In Tables 1 and 2, we furnish the values of efficiency, S_e^D , S_p^D , $FPPV$, $FNPV$, $FPPVI$, and $FNPVI$ for some appropriately chosen values of the sensitivity (Se), the specificity (S_p) and the prevalence (p). The ‘‘Eff’’ (efficiency) column gives the percentage reduction in the expected number of tests achieved by using the pooled testing than the individual testing. For a visual comparison of these factors, we also have plotted efficiency, S_e^D , S_p^D , $FPPV$, $FNPV$ for different values of p in Fig. 2.

Evidently, with an increase in p , and decrease in S_e and S_e , the efficiency reduces, as is evident from Fig. 2a. The value of $S_e^D (= S_e^2)$ is always less than, or equal to, S_e , as can also be observed in Fig. 2b. However, for a given set of values of p , S_e and S_p , the value of S_p^D is always more than S_p and for smaller values of S_p , this effect is substantial. Thus, the pooled testing procedure may lead to a significant improvement of the specificity, especially when the specificity of the individual test is low, see Fig. 2c. From Fig. 2d, we can observe that even a slight decrease in S_p leads to a substantial increase in the value of $FPPV$ for any given values of p and S_e . On the other hand, a change in the value of S_e has a negligible effect on the value of $FPPV$ for any given values of p and S_p . Also, with an increase in p , $FPPV$ decreases for any given values of S_e and S_p . Most importantly, compared to the individual testing, the pooled testing leads to a substantial reduction in the value of $FPPV$, which is extremely important from the point of view of its application in practice. Still, $FPPV$ can take very high values even for pooled testing for a low prevalence disease with low S_p . While this is undesirable, this compares favorably to individual testing. In contrast, for the range of values considered in Tables 1, 2 and Fig. 2 for p , S_e and S_p , the impact of changes in the values of these parameters has little effect on $FNPV$, as is obvious from Fig. 2e.

Notice that in case both the sensitivity and the specificity of the test are low, say, for example, $S_e = 0.8$ and $S_p = 0.7$, even with 10% prevalence, 68% (77%) of the pooled (individual) testing results would be falsely positive, which is extremely high. A recent meta-study by Kucirka et al. (2020), a group of medical professionals from the Johns Hopkins University, demonstrated that over different varieties of RT-PCR tests, which are the most commonly used tests for SARS-CoV-2, a best-case scenario is a $S_e \approx 0.8$. Another meta-study, by Cohen and Kessel (2020), has looked at thirty-seven different external quality assessment studies of different medical assays and found that in some cases, the specificity was as low as 0.83.

For testing for the SARS-CoV-2 virus, the above observations suggest that at places still at the initial phase of the pandemic, where the prevalence is low, (less than 10%,) pooled testing may lead to a substantially lower rate of false positive cases than individual testing, especially if a low specificity test is used. Admittedly, pooled testing leads to a slightly higher rate of false negative cases compared to individual testing, but as observed from Table 1, that effect is usually negligible.

Table 1 Optimal group size (k^*), efficiency in terms of reduction in test numbers as percentage of number of individuals tested (Eff), pooled testing sensitivity (S_e^D) and specificity (S_p^D) false positive predictive value (FPPV) and false negative predictive value (FNPV) corresponding to different prevalence rates (p), sensitivities (S_e) and specificities (S_p) for Dorfman's algorithm. False positive predictive value (FPPV) and false negative predictive value (FNPV) for individual tests are given within parentheses to facilitate comparison

Test quality	p	k^*	Eff (%)	S_e^D	S_p^D	FPPV (FPPV _i)	FNPV (FNPV _i)
$S_e = 1, S_p = 1$	0.01	11	80.44	1	1	0 (0)	0 (0)
$S_e = 1, S_p = 1$	0.02	8	72.58	1	1	0 (0)	0 (0)
$S_e = 1, S_p = 1$	0.05	5	57.38	1	1	0 (0)	0 (0)
$S_e = 1, S_p = 1$	0.10	4	40.61	1	1	0 (0)	0 (0)
$S_e = 0.99, S_p = 0.99$	0.01	11	79.65	0.98	0.999	0.09 (0.50)	0.0002 (0.0001)
$S_e = 0.99, S_p = 0.99$	0.02	8	71.87	0.98	0.999	0.07 (0.33)	0.0004 (0.0002)
$S_e = 0.99, S_p = 0.99$	0.05	5	56.83	0.98	0.998	0.04 (0.16)	0.0010 (0.0005)
$S_e = 0.99, S_p = 0.99$	0.10	4	40.30	0.98	0.997	0.02 (0.08)	0.0022 (0.0011)
$S_e = 0.99, S_p = 0.95$	0.01	11	76.07	0.98	0.993	0.41 (0.83)	0.0002 (0.0001)
$S_e = 0.99, S_p = 0.95$	0.02	8	68.47	0.98	0.991	0.30 (0.71)	0.0004 (0.0002)
$S_e = 0.99, S_p = 0.95$	0.05	5	53.74	0.98	0.989	0.18 (0.49)	0.0011 (0.0006)
$S_e = 0.99, S_p = 0.95$	0.10	4	37.67	0.98	0.985	0.12 (0.31)	0.0022 (0.0012)
$S_e = 0.99, S_p = 0.90$	0.01	11	71.59	0.98	0.981	0.65 (0.91)	0.0002 (0.0001)
$S_e = 0.99, S_p = 0.90$	0.02	8	64.22	0.98	0.978	0.52 (0.83)	0.0004 (0.0002)
$S_e = 0.99, S_p = 0.90$	0.05	5	49.87	0.98	0.973	0.34 (0.66)	0.0011 (0.0006)
$S_e = 0.99, S_p = 0.90$	0.10	4	34.39	0.98	0.966	0.24 (0.48)	0.0023 (0.0012)
$S_e = 0.99, S_p = 0.80$	0.01	12	62.69	0.98	0.943	0.85 (0.95)	0.0002 (0.0001)
$S_e = 0.99, S_p = 0.80$	0.02	9	55.75	0.98	0.936	0.76 (0.91)	0.0004 (0.0003)
$S_e = 0.99, S_p = 0.80$	0.05	6	42.41	0.98	0.924	0.59 (0.79)	0.0011 (0.0007)
$S_e = 0.99, S_p = 0.80$	0.10	4	27.83	0.98	0.917	0.43 (0.65)	0.0024 (0.0014)
$S_e = 0.99, S_p = 0.70$	0.01	13	53.86	0.98	0.886	0.92 (0.97)	0.0002 (0.0001)
$S_e = 0.99, S_p = 0.70$	0.02	9	47.42	0.98	0.879	0.86 (0.94)	0.0005 (0.0003)
$S_e = 0.99, S_p = 0.70$	0.05	6	35.05	0.98	0.863	0.73 (0.85)	0.0012 (0.0008)
$S_e = 0.99, S_p = 0.70$	0.10	5	21.74	0.98	0.839	0.60 (0.73)	0.0026 (0.0016)
$S_e = 0.95, S_p = 0.99$	0.01	11	80.07	0.90	0.999	0.10 (0.51)	0.0010 (0.0005)

Table 1 (continued)

Test quality	p	k^*	Eff (%)	S_e^D	S_p^D	$FPPV (FPPV_i)$	$FNPV (FNPV_i)$
$S_e=0.95,$ $S_p=0.99$	0.02	8	72.47	0.90	0.999	0.07 (0.34)	0.0020 (0.0010)
$S_e=0.95,$ $S_p=0.99$	0.05	5	57.74	0.90	0.998	0.04 (0.17)	0.0051 (0.0027)
$S_e=0.95,$ $S_p=0.99$	0.10	4	41.67	0.90	0.997	0.03 (0.09)	0.0107 (0.0056)
$S_e=0.95,$ $S_p=0.95$	0.01	11	76.49	0.90	0.993	0.43 (0.84)	0.0010 (0.0005)
$S_e=0.95,$ $S_p=0.95$	0.02	8	69.07	0.90	0.992	0.31 (0.72)	0.0020 (0.0011)
$S_e=0.95,$ $S_p=0.95$	0.05	5	54.64	0.90	0.989	0.19 (0.50)	0.0052 (0.0028)
$S_e=0.95,$ $S_p=0.95$	0.10	4	39.05	0.90	0.985	0.13 (0.32)	0.0109 (0.0058)
$S_e=0.95,$ $S_p=0.90$	0.01	11	72.01	0.90	0.982	0.67 (0.91)	0.0010 (0.0006)
$S_e=0.95,$ $S_p=0.90$	0.02	8	64.81	0.90	0.979	0.54 (0.84)	0.0020 (0.0011)
$S_e=0.95,$ $S_p=0.90$	0.05	6	50.82	0.90	0.971	0.38 (0.67)	0.0053 (0.0029)
$S_e=0.95,$ $S_p=0.90$	0.10	4	35.77	0.90	0.967	0.25 (0.49)	0.0111 (0.0061)
$S_e=0.95,$ $S_p=0.80$	0.01	12	63.15	0.90	0.944	0.86 (0.95)	0.0010 (0.0006)
$S_e=0.95,$ $S_p=0.80$	0.02	9	56.42	0.90	0.938	0.77 (0.91)	0.0021 (0.0013)
$S_e=0.95,$ $S_p=0.80$	0.05	6	43.47	0.90	0.926	0.61 (0.80)	0.0055 (0.0033)
$S_e=0.95,$ $S_p=0.80$	0.10	5	29.29	0.90	0.908	0.48 (0.65)	0.0118 (0.0069)
$S_e=0.95,$ $S_p=0.70$	0.01	13	54.35	0.90	0.888	0.92 (0.97)	0.0011 (0.0007)
$S_e=0.95,$ $S_p=0.70$	0.02	10	48.11	0.90	0.878	0.87 (0.94)	0.0023 (0.0015)
$S_e=0.95,$ $S_p=0.70$	0.05	6	36.11	0.90	0.866	0.74 (0.86)	0.0059 (0.0037)
$S_e=0.95,$ $S_p=0.70$	0.10	5	23.38	0.90	0.843	0.61 (0.74)	0.0127 (0.0079)
$S_e=0.90,$ $S_p=0.99$	0.01	11	80.59	0.81	0.999	0.10 (0.52)	0.0019 (0.0010)
$S_e=0.90,$ $S_p=0.99$	0.02	8	73.22	0.81	0.999	0.07 (0.35)	0.0039 (0.0021)
$S_e=0.90,$ $S_p=0.99$	0.05	5	58.87	0.81	0.998	0.04 (0.17)	0.0099 (0.0053)
$S_e=0.90,$ $S_p=0.99$	0.10	4	43.39	0.81	0.997	0.03 (0.09)	0.0207 (0.0111)
$S_e=0.90,$ $S_p=0.95$	0.01	11	77.01	0.81	0.993	0.45 (0.85)	0.0019 (0.0011)
$S_e=0.90,$ $S_p=0.95$	0.02	8	69.81	0.81	0.992	0.33 (0.73)	0.0039 (0.0021)

Table 1 (continued)

Test quality	p	k^*	Eff (%)	S_e^D	S_p^D	FPPV (FPPV _{<i>p</i>})	FNPV (FNPV _{<i>p</i>})
$S_e=0.90,$ $S_p=0.95$	0.05	6	55.82	0.81	0.988	0.22 (0.51)	0.0100 (0.0055)
$S_e=0.90,$ $S_p=0.95$	0.10	4	40.77	0.81	0.986	0.13 (0.33)	0.021 (0.0116)
$S_e=0.90,$ $S_p=0.90$	0.01	12	72.58	0.81	0.982	0.69 (0.92)	0.0020 (0.0011)
$S_e=0.90,$ $S_p=0.90$	0.02	9	65.59	0.81	0.978	0.57 (0.84)	0.0039 (0.0023)
$S_e=0.90,$ $S_p=0.90$	0.05	6	52.14	0.81	0.972	0.40 (0.68)	0.0102 (0.0058)
$S_e=0.90,$ $S_p=0.90$	0.10	4	37.49	0.81	0.968	0.26 (0.5)	0.0213 (0.0122)
$S_e=0.90,$ $S_p=0.80$	0.01	13	63.73	0.81	0.944	0.87 (0.96)	0.0020 (0.0013)
$S_e=0.90,$ $S_p=0.80$	0.02	9	57.25	0.81	0.939	0.79 (0.92)	0.0041 (0.0025)
$S_e=0.90,$ $S_p=0.80$	0.05	6	44.79	0.81	0.928	0.63 (0.81)	0.0107 (0.0065)
$S_e=0.90,$ $S_p=0.80$	0.10	5	31.33	0.81	0.912	0.49 (0.67)	0.0226 (0.0137)
$S_e=0.90,$ $S_p=0.70$	0.01	14	54.98	0.81	0.888	0.93 (0.97)	0.0022 (0.0014)
$S_e=0.90,$ $S_p=0.70$	0.02	10	49.02	0.81	0.880	0.88 (0.94)	0.0044 (0.0029)
$S_e=0.90,$ $S_p=0.70$	0.05	7	37.61	0.81	0.862	0.76 (0.86)	0.0115 (0.0075)
$S_e=0.90,$ $S_p=0.70$	0.10	5	25.43	0.81	0.848	0.63 (0.75)	0.0243 (0.0156)
$S_e=0.80,$ $S_p=0.99$	0.01	12	81.69	0.64	0.999	0.13 (0.55)	0.0036 (0.0020)
$S_e=0.80,$ $S_p=0.99$	0.02	9	74.75	0.64	0.999	0.09 (0.38)	0.0073 (0.0041)
$S_e=0.80,$ $S_p=0.99$	0.05	6	61.41	0.64	0.998	0.05 (0.19)	0.0186 (0.0105)
$S_e=0.80,$ $S_p=0.99$	0.10	4	46.83	0.64	0.998	0.03 (0.10)	0.0385 (0.0220)
$S_e=0.80,$ $S_p=0.95$	0.01	12	78.15	0.64	0.994	0.50 (0.86)	0.0036 (0.0021)
$S_e=0.80,$ $S_p=0.95$	0.02	9	71.42	0.64	0.992	0.38 (0.75)	0.0074 (0.0043)
$S_e=0.80,$ $S_p=0.95$	0.05	6	58.47	0.64	0.989	0.25 (0.54)	0.0188 (0.0110)
$S_e=0.80,$ $S_p=0.95$	0.10	5	44.29	0.64	0.985	0.18 (0.36)	0.0390 (0.0229)
$S_e=0.80,$ $S_p=0.90$	0.01	13	73.73	0.64	0.982	0.74 (0.93)	0.0037 (0.0022)
$S_e=0.80,$ $S_p=0.90$	0.02	9	67.25	0.64	0.980	0.61 (0.86)	0.0074 (0.0045)
$S_e=0.80,$ $S_p=0.90$	0.05	6	54.79	0.64	0.974	0.43 (0.70)	0.0191 (0.0116)

Table 1 (continued)

Test quality	p	k^*	Eff (%)	S_e^D	S_p^D	FPPV (FPPV _i)	FNPV (FNPV _i)
$S_e=0.80,$ $S_p=0.90$	0.10	5	41.33	0.64	0.966	0.32 (0.53)	0.0398 (0.0241)
$S_e=0.80,$ $S_p=0.80$	0.01	14	64.98	0.64	0.945	0.89 (0.96)	0.0038 (0.0025)
$S_e=0.80,$ $S_p=0.80$	0.02	10	59.02	0.64	0.940	0.82 (0.92)	0.0078 (0.0051)
$S_e=0.80,$ $S_p=0.80$	0.05	7	47.61	0.64	0.928	0.68 (0.83)	0.0200 (0.0130)
$S_e=0.80,$ $S_p=0.80$	0.10	5	35.43	0.64	0.919	0.53 (0.69)	0.0417 (0.0270)
$S_e=0.80,$ $S_p=0.70$	0.01	15	56.34	0.64	0.890	0.94 (0.97)	0.0041 (0.0029)
$S_e=0.80,$ $S_p=0.70$	0.02	11	50.95	0.64	0.883	0.90 (0.95)	0.0083 (0.0058)
$S_e=0.80,$ $S_p=0.70$	0.05	8	40.67	0.64	0.865	0.80 (0.88)	0.0214 (0.0148)
$S_e=0.80,$ $S_p=0.70$	0.10	6	29.91	0.64	0.849	0.68 (0.77)	0.0450 (0.0308)

Finally, it may be noted here that the impacts of the two diagnostic errors are asymmetric. In case of more false positive cases, more people are to be quarantined leading to a lot of economic, social, and emotional turmoil. On the other hand, in case of more false negative results, more infected people will be released into the population causing the disease to spread faster. Where detection and containment of the disease is the goal, the policy planner must make a trade-off. However, these considerations are not relevant when the goal is to the estimation of prevalence.

4 Estimation of prevalence using pooled testing data

In this section, first, we present the theory of estimation of the prevalence of a disease from the test data using basic probability theory (Feller 1968). Next, we study the impact of misspecification of sensitivity and specificity on the estimate of prevalence.

An accurate estimation of the prevalence of SARS-CoV-2-virus, being critical for assessing the lethality of the pandemic, is crucial for the policymakers to make strategic decisions. Recently, Ioannidis (2020), lamented in a highly critical opinion piece on COVID-19: “Three months after the outbreak emerged, most countries, including the U.S., lack the ability to test a large number of people and no countries have reliable data on the prevalence of the virus in a representative random sample of the general population.” Ioannidis argued that in the absence of such data, the estimate of fatality rate is bound to be a substantial overestimate of the true fatality rate if a significant percentage of COVID-19 cases are undetected. It is now well known that a significant percentage of COVID-19 cases are asymptomatic. Raman R Gangakhedkar, chief epidemiologist, ICMR, reported “Of 100 people with infection, 80 do not have symptoms,” (Thacker 2020). A lab in Iceland has suggested that 50% of the infected are asymptomatic, see John (2020). According

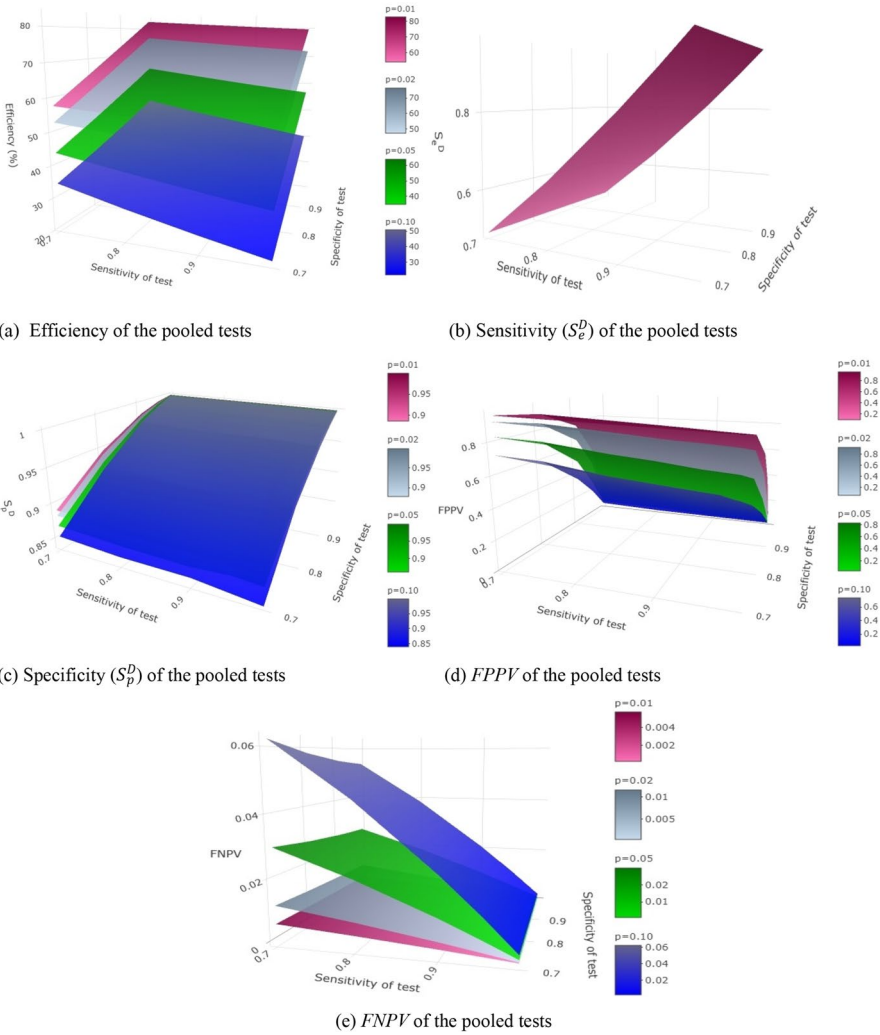


Fig. 2 Surface plots representing efficiency, S_e^D , S_p^D , $FPPV$ and $FNPV$ for different values of S_e and S_p for $p = 0.01, 0.02, 0.05$ and 0.10 . S_e^D is independent of p and hence only one common surface plot is produced

to a report by Arnold (2020), in a Boston homeless shelter, out of 400 guests staying there, 146 tested positive for COVID-19, but all were reported to be asymptomatic. If, for example, 80% of the cases are asymptomatic, then the estimate of fatality rate would naturally be five times the true fatality rate. While the reported percentage of asymptomatic cases vary from place to place, all agree that it is significant. Considering the possibility of substantial underestimation of infected cases in the absence of test data, Ioannidis has contended that the virus could be less deadly than people think, and destroying the economy in the effort to fight with the virus could be a “once-in-a-century-evidence-fiasco”.

A recent research paper published by Bendavid et al. (2021) lends support to Ioannidis’ contention. By estimating the prevalence of SARS-CoV-2 virus among the residents of Santa Clara County of California using antibody test data of a “properly selected” sample

Table 2 Expected values of the proportion of samples testing positive (π) corresponding to different prevalence rates, sensitivities (S_e) and specificities (S_p) for individual testing (π_I) Dorfman's algorithm (π_D). The π_D values are based on $k=4$, the most conservative pooling choice from Tables 1

Test Quality	p	π_I	π_D
$S_e = 1, S_p = 1$	0.01	0.01	0.01
$S_e = 1, S_p = 1$	0.02	0.02	0.02
$S_e = 1, S_p = 1$	0.05	0.05	0.05
$S_e = 1, S_p = 1$	0.10	0.10	0.10
$S_e = 1, S_p = 0.90$	0.01	0.11	0.02
$S_e = 1, S_p = 0.90$	0.02	0.12	0.03
$S_e = 1, S_p = 0.90$	0.05	0.14	0.07
$S_e = 1, S_p = 0.90$	0.10	0.19	0.13
$S_e = 1, S_p = 0.80$	0.01	0.21	0.05
$S_e = 1, S_p = 0.80$	0.02	0.22	0.07
$S_e = 1, S_p = 0.80$	0.05	0.24	0.11
$S_e = 1, S_p = 0.80$	0.10	0.28	0.18
$S_e = 1, S_p = 0.70$	0.01	0.31	0.11
$S_e = 1, S_p = 0.70$	0.02	0.31	0.12
$S_e = 1, S_p = 0.70$	0.05	0.34	0.16
$S_e = 1, S_p = 0.70$	0.10	0.37	0.23
$S_e = 0.90, S_p = 1$	0.01	0.01	0.01
$S_e = 0.90, S_p = 1$	0.02	0.02	0.02
$S_e = 0.90, S_p = 1$	0.05	0.05	0.04
$S_e = 0.90, S_p = 1$	0.10	0.09	0.08
$S_e = 0.90, S_p = 0.90$	0.01	0.11	0.02
$S_e = 0.90, S_p = 0.90$	0.02	0.12	0.03
$S_e = 0.90, S_p = 0.90$	0.05	0.14	0.06
$S_e = 0.90, S_p = 0.90$	0.10	0.18	0.11
$S_e = 0.90, S_p = 0.80$	0.01	0.21	0.05
$S_e = 0.90, S_p = 0.80$	0.02	0.21	0.06
$S_e = 0.90, S_p = 0.80$	0.05	0.23	0.10
$S_e = 0.90, S_p = 0.80$	0.10	0.27	0.15
$S_e = 0.90, S_p = 0.70$	0.01	0.31	0.10
$S_e = 0.90, S_p = 0.70$	0.02	0.31	0.11
$S_e = 0.90, S_p = 0.70$	0.05	0.33	0.15
$S_e = 0.90, S_p = 0.70$	0.10	0.36	0.21
$S_e = 0.80, S_p = 1$	0.01	0.01	0.01
$S_e = 0.80, S_p = 1$	0.02	0.02	0.01
$S_e = 0.80, S_p = 1$	0.05	0.04	0.03
$S_e = 0.80, S_p = 1$	0.10	0.08	0.06
$S_e = 0.80, S_p = 0.90$	0.01	0.11	0.02
$S_e = 0.80, S_p = 0.90$	0.02	0.11	0.03
$S_e = 0.80, S_p = 0.90$	0.05	0.13	0.05
$S_e = 0.80, S_p = 0.90$	0.10	0.17	0.09
$S_e = 0.80, S_p = 0.80$	0.01	0.21	0.05
$S_e = 0.80, S_p = 0.80$	0.02	0.21	0.06
$S_e = 0.80, S_p = 0.80$	0.05	0.23	0.13

Table 2 (continued)

Test Quality	p	π_I	π_D
$S_e = 0.80, S_p = 0.80$	0.10	0.26	0.10
$S_e = 0.80, S_p = 0.70$	0.01	0.31	0.10
$S_e = 0.80, S_p = 0.70$	0.02	0.31	0.11
$S_e = 0.80, S_p = 0.70$	0.05	0.33	0.14
$S_e = 0.80, S_p = 0.70$	0.10	0.35	0.18
$S_e = 0.70, S_p = 1$	0.01	0.01	0
$S_e = 0.70, S_p = 1$	0.02	0.01	0.01
$S_e = 0.70, S_p = 1$	0.05	0.03	0.02
$S_e = 0.70, S_p = 1$	0.10	0.07	0.05
$S_e = 0.70, S_p = 0.90$	0.01	0.11	0.02
$S_e = 0.70, S_p = 0.90$	0.02	0.11	0.02
$S_e = 0.70, S_p = 0.90$	0.05	0.13	0.04
$S_e = 0.70, S_p = 0.90$	0.10	0.16	0.07
$S_e = 0.70, S_p = 0.80$	0.01	0.20	0.05
$S_e = 0.70, S_p = 0.80$	0.02	0.21	0.05
$S_e = 0.70, S_p = 0.80$	0.05	0.22	0.08
$S_e = 0.70, S_p = 0.80$	0.10	0.25	0.11
$S_e = 0.70, S_p = 0.70$	0.01	0.30	0.10
$S_e = 0.70, S_p = 0.70$	0.02	0.31	0.10
$S_e = 0.70, S_p = 0.70$	0.05	0.32	0.13
$S_e = 0.70, S_p = 0.70$	0.10	0.34	0.16

of 3300 residents, it predicted that between 50 and 85 times more residents of the county were actually infected than what appeared in the official tallies, see Regaldo (2020). The authors claimed that “their data helps prove ... if undetected infections are as widespread as they think, then the death rate in the county may be less than 0.2%, about a fifth to a tenth other estimates.” The publication of this research paper immediately created waves in social media, in the press, and policy circles; if indeed, this number is not far from the truth, then the COVID-19 fatality rate is not very different from common flu. Realizing the findings’ serious policy implications, especially its lending support to the view of lifting the lockdown, some scientists immediately issued a caveat about the accuracy of the estimate. They expressed concerns about the flaws in the sample selection method (through Facebook advertisements), the statistical analysis carried out, and the reliability of the antibody tests (Bajak and Howe 2020; McCormick 2020).

We now discuss the methodology used by Bendavid et al. (2021) for prevalence estimation based on individual serological testing data. Let us denote the probability of a positive test result for an arbitrarily chosen individual by π . Clearly, for a sample of size n , the estimate of π is simply the proportion of people being identified as positive out of the sample. But we know from the previous section that

$$\pi = P(\text{Positive result}) = (1 - p)(1 - S_p) + pS_p \quad (6)$$

Plugging in the proportion of positive test results obtained from the sample, say $\hat{\pi}$, for π in (6), we obtain.

$$\hat{p} = \frac{\hat{\pi} + S_p - 1}{S_e + S_p - 1}. \tag{7}$$

However, for estimating prevalence using Dorfman’s algorithm, we simply need to replace S_e and S_p in (6) by S_e^D and S_p^D respectively (cf. Eqs. (3)–(4)). Thus, we have

$$\pi = (1 - S_p)(1 - S_p - S_e)(1 - p)^k + (1 - S_p)S_e(1 - p) + pS_e^2, \tag{8}$$

or, equivalently,

$$\pi = (1 - S_p - S_e)[(1 - S_p)(1 - p)^k - pS_e] + (1 - S_p)S_e. \tag{9}$$

Now, when Dorfman’s algorithm is used, and the estimate of π is given by $\hat{\pi}^D$, the sample proportion of positive results, we can show that $\hat{\pi}^D$ is unbiased and consistent for π , and moreover $\sqrt{n}(\hat{\pi}^D - \pi) \rightarrow N(0, C)$, where

$$C = \left(\pi - \pi^2 + \frac{k-1}{2}(1-p)^{k+2}(1-S_p)^3 + (1 - (1-p)^k)S_e(S_e p + (1 - S_p)(1-p))^2 \right).$$

The details are given in Appendix B.

Now, from (9) we observe that $\frac{d\pi}{dp} = (S_e + S_p - 1)[(1 - S_p)(1 - p)^{k-1} + S_e]$, whose sign is the same as that of $(S_e + S_p - 1)$; as long as $S_e + S_p \neq 1$. In practice, S_e and S_p are both high (> 0.5) and so $S_e + S_p$ is always positive. Therefore, π is a strictly monotonic function of p , and therefore, if the value of π is known, from (9) we can easily solve for p using a suitable numerical algorithm, e.g., a grid search algorithm. In the current paper, we achieve that using the *uniroot* function from the statistical software R, which computes the real roots of a polynomial within a given range by implementing an algorithm by Brent (1973) through successive linear interpolations. In Table 2, we report the values of π , the probability of testing positive, for the individual testing as well as for the Dorfman testing procedure, for different values of p , S_e and S_p . We denote by π_I (cf. Equation (6)) and π_D (cf. Equation (8)) the π values corresponding to the individual and Dorfman testing procedure, respectively. Figure 3 and Table 2 compare the values of π_I and π_D for different values of p . It is evident that individual testing leads to a much higher chance of testing positive compared to pooled testing when the test is not perfect.

Plugging in $\hat{\pi}^D$, the estimate of π obtained from the test data, in (8) or (9) and then solving it numerically for p , we can get an estimate of p , say \hat{p}^D . This estimate, although biased, is consistent for p , and $\sqrt{n}(\hat{p}^D - p) \rightarrow N(0, C')$, where $C' = \frac{C}{\left| \frac{d\pi}{dp} \right|}$, $\frac{d\pi}{dp} = (S_e + S_p - 1)((1 - S_p)k(1 - p)^{k-1} + S_e)$. This allows us to construct confidence intervals for p , for example, an approximate level- α confidence interval for p is given by $\hat{p}^D \pm z_{\alpha/2} \frac{\hat{C}'}{\sqrt{n}}$, where \hat{C}' is the plug-in estimator of C' , and $z_{\alpha/2}$ is the upper $\alpha/2$ th quantile of the standard normal distribution. The details of the derivations are given in Appendix C.

As noted in Sect. 3, both RT-PCR and antibody tests have high sensitivity and specificity in lab settings, but in a real-world situation, these values may be substantially lower than those obtained in lab settings. Suppose S_e^p and S_p^p are respectively the sensitivity and the specificity values as perceived and used by the scientist for estimating the prevalence from Eq. (7) or (8), whereas S_e^T and S_p^T are respectively the true sensitivity and specificity values in the field. In real-life situations, it is often the case that $S_e^T(S_p^T)$, is substantially less than $S_e^p(S_p^p)$, because the scientist’s perceived values are usually influenced by the values

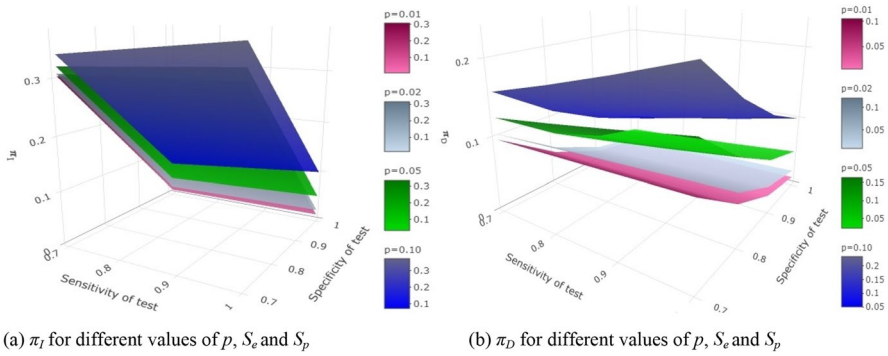


Fig. 3 Surface plot of π_I and π_D for $p=0.01, 0.02, 0.05$ and 0.10

under the lab setting. Thus, given the true value of π , the solution to Eq. (7) or (8) for p is equal to the true prevalence p_T if S_e and S_p are replaced S_e^T and S_p^T . On the other hand, given the true value of π , replacing S_e and S_p respectively by S_e^B and S_p^B in (7) and (8) would yield a solution p_p which is expected to deviate from the true prevalence p_T . We note that for \hat{p}_T and \hat{p}_p , obtained as above, $(\hat{p}_T - \hat{p}_p)$ is asymptotically consistent for the true bias $(p_T - p_p)$, provided the true sensitivity and specificity values are known. In fact, in such a scenario, we can also provide an approximate level- α confidence interval for the bias, which will be of the form $(\hat{p}_T - \hat{p}_p) \pm z_{\alpha/2} \frac{\hat{C}_E}{\sqrt{n}}$, where \hat{C}_E is the plug-in estimator of C'' . For detailed derivation and the expression for C_E , the readers can refer to Appendix D.

We study the effect of the misspecification of sensitivity and specificity values on the estimate of prevalence by evaluating the bias in the estimation of $(p_T - p_p)$. In Table 3, we report the values of the bias $(p_T - p_p)$ resulting from the individual testing as well as Dorfman’s pooled testing, denoted by $bias_I$ and $bias_D$, respectively, for different values of p, S_e^p, S_p^p, S_e^T and S_p^T . The same are also presented in Fig. 4.

The following patterns are visible from Table 3 and Fig. 4:

- (i) With the increase in the prevalence p , the bias is decreasing. Initially, it reduces from positive to zero and then becomes negative. We report the bias for four different values of p : 0.01, 0.02, 0.05 and 0.10.
- (ii) The effect of misspecification of S_e and S_p are asymmetric in nature. Misspecification of S_e has a little effect on the bias. However, the misspecification of S_p has a significantly large effect on the bias. Also, more is the deviation from the true values more is the effect on bias.
- (iii) The most interesting observation is, compared to the individual testing, pooled testing reduces the bias due to misspecification of S_e and S_p substantially. This is extremely important to know given the fact that misspecification of S_e and S_p are quite common.

In Table 2, we have reported the values of π , the probability of testing positive, for the individual testing as well as for the Dorfman testing procedure, for different values of p, S_e and S_p . Notice from Table 2 that, like the observation (ii) made above, the value of π is significantly influenced by S_p but not so by S_e . Also, it is important to note that, for a given value of S_p (say, 0.9) reduction in S_e (say, from 1 to 0.7) results in a reduction of π , though

Table 3 Bias $p_T - p_P$ when perceived sensitivity (S_e^p) and specificity (S_p^p) are higher than the true situations. Bias $bias_I$ corresponds to individual tests, while $bias_D$ corresponds to Dorfman tests. The Dorfman values are based on $k=4$, the most conservative pooling choice from Tables 1 and 2. A positive bias means the perceived prevalence is lower than the true prevalence, while a negative bias means the opposite

Perceived test quality	p	$S_e^p = S_e^t, S_p^p = S_p^t$ 10% less		S_e^t 10% less, $S_p^t = S_p^p$		Both S_e^t and S_p^t 10% less		Both S_e^t and S_p^t 20% less	
		bias _I	bias _D	bias _I	bias _D	bias _I	bias _D	bias _I	bias _D
$S_e = 1, S_p = 1$	0.01	0.10	0.01	0	0	0.10	0.01	0.20	0.04
$S_e = 1, S_p = 1$	0.02	0.10	0.01	0	0	0.10	0.01	0.19	0.04
$S_e = 1, S_p = 1$	0.05	0.10	0.02	-0.01	-0.01	0.09	0.01	0.18	0.04
$S_e = 1, S_p = 1$	0.10	0.09	0.03	-0.01	-0.02	0.08	0.01	0.16	0.03
$S_e = 0.99, S_p = 0.99$	0.01	0.10	0.01	0	0	0.10	0.01	0.20	0.04
$S_e = 0.99, S_p = 0.99$	0.02	0.10	0.02	0	0	0.10	0.01	0.20	0.04
$S_e = 0.99, S_p = 0.99$	0.05	0.10	0.02	-0.01	-0.01	0.09	0.01	0.18	0.04
$S_e = 0.99, S_p = 0.99$	0.10	0.09	0.03	-0.01	-0.02	0.08	0.01	0.16	0.03
$S_e = 0.99, S_p = 0.95$	0.01	0.11	0.02	0	0	0.10	0.02	0.21	0.05
$S_e = 0.99, S_p = 0.95$	0.02	0.10	0.02	0	0	0.10	0.02	0.20	0.05
$S_e = 0.99, S_p = 0.95$	0.05	0.10	0.03	-0.01	-0.01	0.10	0.02	0.19	0.05
$S_e = 0.99, S_p = 0.95$	0.10	0.10	0.04	-0.01	-0.02	0.09	0.01	0.17	0.04
$S_e = 0.99, S_p = 0.90$	0.01	0.11	0.03	0	0	0.11	0.02	0.22	0.06
$S_e = 0.99, S_p = 0.90$	0.02	0.11	0.03	0	0	0.11	0.02	0.22	0.06

Table 3 (continued)

Perceived test quality	p	$S_e^T = S_e^P, S_p^T, 10\%$ less		$S_e^T, 10\%$ less, $S_p^T = S_p^P$		Both S_e^T and S_p^T 10% less		Both S_e^T and S_p^T 20% less	
		bias _I	bias _D	bias _I	bias _D	bias _I	bias _D	bias _I	bias _D
$S_e = 0.99, S_p = 0.90$	0.05	0.11	0.03	-0.01	-0.01	0.10	0.02	0.20	0.06
$S_e = 0.99, S_p = 0.90$	0.10	0.10	0.04	-0.01	-0.02	0.09	0.02	0.18	0.05
$S_e = 0.99, S_p = 0.80$	0.01	0.13	0.04	0	0	0.12	0.04	0.25	0.09
$S_e = 0.99, S_p = 0.80$	0.02	0.12	0.04	0	0	0.12	0.03	0.24	0.08
$S_e = 0.99, S_p = 0.80$	0.05	0.12	0.04	-0.01	-0.01	0.11	0.03	0.23	0.08
$S_e = 0.99, S_p = 0.80$	0.10	0.11	0.05	-0.01	-0.02	0.10	0.03	0.20	0.06
$S_e = 0.99, S_p = 0.70$	0.01	0.14	0.05	0	0	0.14	0.05	0.28	0.11
$S_e = 0.99, S_p = 0.70$	0.02	0.14	0.05	0	0	0.14	0.05	0.28	0.11
$S_e = 0.99, S_p = 0.70$	0.05	0.14	0.05	-0.01	-0.01	0.13	0.04	0.26	0.10
$S_e = 0.99, S_p = 0.70$	0.10	0.13	0.06	-0.01	-0.02	0.12	0.03	0.23	0.08
$S_e = 0.95, S_p = 0.99$	0.01	0.11	0.02	0	0	0.10	0.01	0.21	0.05
$S_e = 0.95, S_p = 0.99$	0.02	0.10	0.02	0	0	0.10	0.01	0.20	0.05
$S_e = 0.95, S_p = 0.99$	0.05	0.10	0.02	-0.01	-0.01	0.10	0.01	0.19	0.04

Table 3 (continued)

Perceived test quality	p	$S_e^T = S_e^P, S_p^T, 10\%$ less		$S_e^T, 10\%$ less, $S_p^T = S_p^P$		Both S_e^T and S_p^T 10% less		Both S_e^T and S_p^T 20% less	
		bias _I	bias _D	bias _I	bias _D	bias _I	bias _D	bias _I	bias _D
$S_e = 0.95, S_p = 0.99$	0.10	0.10	0.03	-0.01	-0.02	0.09	0.01	0.17	0.03
$S_e = 0.95, S_p = 0.95$	0.01	0.11	0.02	0	0	0.11	0.02	0.22	0.06
$S_e = 0.95, S_p = 0.95$	0.02	0.11	0.02	0	0	0.11	0.02	0.21	0.06
$S_e = 0.95, S_p = 0.95$	0.05	0.11	0.03	-0.01	-0.01	0.10	0.02	0.20	0.05
$S_e = 0.95, S_p = 0.95$	0.10	0.10	0.04	-0.01	-0.02	0.09	0.01	0.18	0.04
$S_e = 0.95, S_p = 0.95$	0.01	0.12	0.03	0	0	0.12	0.03	0.23	0.07
$S_e = 0.95, S_p = 0.90$	0.02	0.12	0.03	0	0	0.11	0.03	0.23	0.07
$S_e = 0.95, S_p = 0.90$	0.05	0.11	0.03	-0.01	-0.01	0.11	0.02	0.21	0.06
$S_e = 0.95, S_p = 0.90$	0.10	0.11	0.04	-0.01	-0.02	0.09	0.02	0.19	0.05
$S_e = 0.95, S_p = 0.90$	0.01	0.13	0.04	0	0	0.13	0.04	0.26	0.09
$S_e = 0.95, S_p = 0.80$	0.02	0.13	0.04	0	0	0.13	0.04	0.26	0.09
$S_e = 0.95, S_p = 0.80$	0.05	0.13	0.04	-0.01	-0.01	0.12	0.03	0.24	0.08
$S_e = 0.95, S_p = 0.80$	0.10	0.12	0.05	-0.01	-0.02	0.11	0.03	0.21	0.06

Table 3 (continued)

Perceived test quality	p	$S_e^T = S_e^P, S_p^T, 10\%$ less		$S_e^T, 10\%$ less, $S_p^T = S_p^P$		Both S_e^T and S_p^T 10% less		Both S_e^T and S_p^T 20% less	
		bias _I	bias _D	bias _I	bias _D	bias _I	bias _D	bias _I	bias _D
$S_e = 0.95, S_p = 0.70$	0.01	0.15	0.05	0	0	0.15	0.05	0.30	0.12
$S_e = 0.95, S_p = 0.70$	0.02	0.15	0.05	0	0	0.15	0.05	0.30	0.12
$S_e = 0.95, S_p = 0.70$	0.05	0.15	0.06	-0.01	-0.01	0.14	0.04	0.28	0.11
$S_e = 0.95, S_p = 0.70$	0.10	0.14	0.06	-0.02	-0.02	0.12	0.04	0.25	0.08
$S_e = 0.90, S_p = 0.99$	0.01	0.11	0.02	0	0	0.11	0.01	0.22	0.05
$S_e = 0.90, S_p = 0.99$	0.02	0.11	0.02	0	0	0.11	0.01	0.22	0.05
$S_e = 0.90, S_p = 0.99$	0.05	0.11	0.03	-0.01	-0.01	0.10	0.01	0.20	0.05
$S_e = 0.90, S_p = 0.99$	0.10	0.10	0.04	-0.01	-0.02	0.09	0.01	0.18	0.04
$S_e = 0.90, S_p = 0.95$	0.01	0.12	0.02	0	0	0.12	0.02	0.23	0.06
$S_e = 0.90, S_p = 0.95$	0.02	0.12	0.03	0	0	0.11	0.02	0.23	0.06
$S_e = 0.90, S_p = 0.95$	0.05	0.11	0.03	-0.01	-0.01	0.11	0.02	0.21	0.06
$S_e = 0.90, S_p = 0.95$	0.10	0.11	0.04	-0.01	-0.02	0.09	0.02	0.19	0.04
$S_e = 0.90, S_p = 0.90$	0.01	0.12	0.03	0	0	0.12	0.03	0.25	0.08

Table 3 (continued)

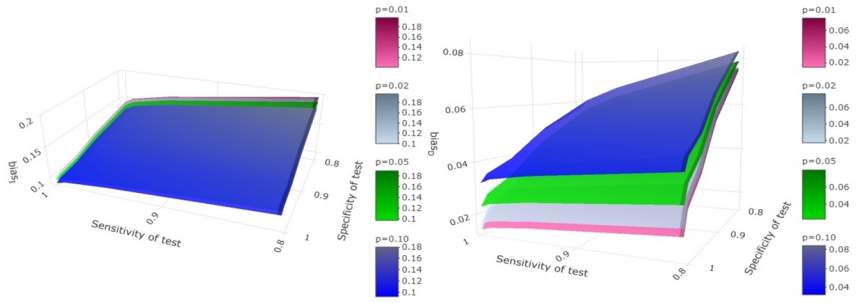
Perceived test quality	p	$S_e^T = S_e^P, S_p^T, 10\%$ less		$S_e^T, 10\%$ less, $S_p^T = S_p^P$		Both S_e^T and S_p^T 10% less		Both S_e^T and S_p^T 20% less	
		bias _I	bias _D	bias _I	bias _D	bias _I	bias _D	bias _I	bias _D
$S_e = 0.90, S_p = 0.90$	0.02	0.12	0.03	0	0	0.12	0.03	0.24	0.08
$S_e = 0.90, S_p = 0.90$	0.05	0.12	0.04	-0.01	-0.01	0.11	0.03	0.22	0.07
$S_e = 0.90, S_p = 0.90$	0.10	0.11	0.04	-0.01	-0.02	0.10	0.02	0.20	0.05
$S_e = 0.90, S_p = 0.90$	0.01	0.14	0.04	0	0	0.14	0.04	0.28	0.11
$S_e = 0.90, S_p = 0.80$	0.02	0.14	0.05	0	0	0.14	0.04	0.27	0.10
$S_e = 0.90, S_p = 0.80$	0.05	0.14	0.05	-0.01	-0.01	0.13	0.04	0.26	0.09
$S_e = 0.90, S_p = 0.80$	0.10	0.13	0.05	-0.01	-0.02	0.11	0.03	0.23	0.07
$S_e = 0.90, S_p = 0.70$	0.01	0.16	0.06	0	0	0.16	0.06	0.33	0.14
$S_e = 0.90, S_p = 0.70$	0.02	0.16	0.06	0	0	0.16	0.06	0.32	0.13
$S_e = 0.90, S_p = 0.70$	0.05	0.16	0.06	-0.01	-0.01	0.15	0.05	0.30	0.12
$S_e = 0.90, S_p = 0.70$	0.10	0.15	0.07	-0.02	-0.02	0.13	0.04	0.27	0.09
$S_e = 0.80, S_p = 0.99$	0.01	0.13	0.02	0	0	0.12	0.02	0.25	0.07
$S_e = 0.80, S_p = 0.99$	0.02	0.12	0.02	0	0	0.12	0.02	0.24	0.06

Table 3 (continued)

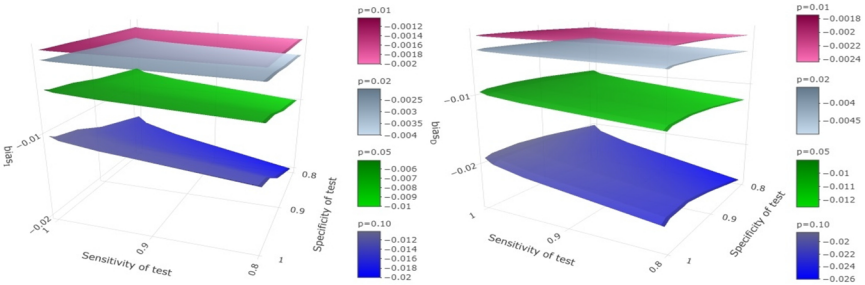
Perceived test quality	p	$S_e^T = S_e^P, S_e^T, S_e^P$ 10% less		S_e^T 10% less, $S_e^P = S_e^P$		Both S_e^T and S_e^P 10% less		Both S_e^T and S_e^P 20% less	
		bias _I	bias _D	bias _I	bias _D	bias _I	bias _D	bias _I	bias _D
$S_e = 0.80, S_p = 0.99$	0.05	0.12	0.03	-0.01	-0.01	0.11	0.02	0.23	0.06
$S_e = 0.80, S_p = 0.99$	0.10	0.11	0.04	-0.01	-0.02	0.10	0.01	0.20	0.05
$S_e = 0.80, S_p = 0.95$	0.01	0.13	0.03	0	0	0.13	0.03	0.26	0.08
$S_e = 0.80, S_p = 0.95$	0.02	0.13	0.03	0	0	0.13	0.03	0.26	0.08
$S_e = 0.80, S_p = 0.95$	0.05	0.13	0.04	-0.01	-0.01	0.12	0.02	0.24	0.07
$S_e = 0.80, S_p = 0.95$	0.10	0.12	0.05	-0.01	-0.02	0.11	0.02	0.21	0.05
$S_e = 0.80, S_p = 0.90$	0.01	0.14	0.04	0	0	0.14	0.04	0.28	0.10
$S_e = 0.80, S_p = 0.90$	0.02	0.14	0.04	0	0	0.14	0.03	0.27	0.09
$S_e = 0.80, S_p = 0.90$	0.05	0.14	0.05	-0.01	-0.01	0.13	0.03	0.26	0.08
$S_e = 0.80, S_p = 0.90$	0.10	0.13	0.05	-0.01	-0.02	0.11	0.03	0.23	0.07
$S_e = 0.80, S_p = 0.80$	0.01	0.16	0.06	0	0	0.16	0.05	0.33	0.13
$S_e = 0.80, S_p = 0.80$	0.02	0.16	0.06	0	0	0.16	0.05	0.32	0.13
$S_e = 0.80, S_p = 0.80$	0.05	0.16	0.06	-0.01	-0.01	0.15	0.05	0.30	0.11

Table 3 (continued)

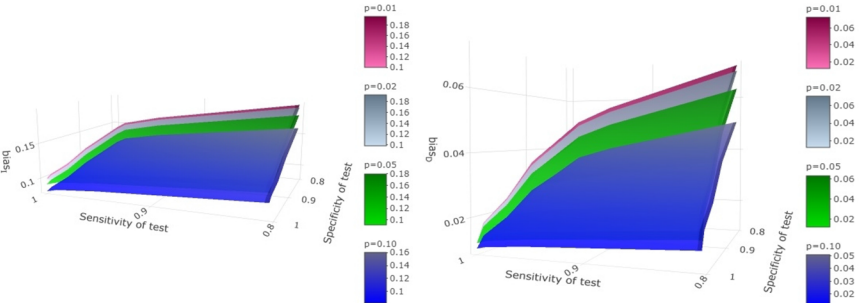
Perceived test quality	p	$S_e^T = S_e^P, S_p^T, 10\%$ less		$S_e^T, 10\%$ less, $S_p^T = S_p^P$		Both S_e^T and S_p^T 10% less		Both S_e^T and S_p^T 20% less	
		bias _I	bias _D	bias _I	bias _D	bias _I	bias _D	bias _I	bias _D
$S_e = 0.80,$ $S_p = 0.80$	0.10	0.15	0.07	-0.02	-0.02	0.13	0.04	0.27	0.09
$S_e = 0.80,$ $S_p = 0.70$	0.01	0.20	0.08	0	0	0.20	0.07	0.39	0.18
$S_e = 0.80,$ $S_p = 0.70$	0.02	0.20	0.08	0	0	0.19	0.07	0.38	0.18
$S_e = 0.80,$ $S_p = 0.70$	0.05	0.19	0.08	-0.01	-0.01	0.18	0.06	0.36	0.16
$S_e = 0.80,$ $S_p = 0.70$	0.10	0.18	0.08	-0.02	-0.03	0.16	0.05	0.32	0.12



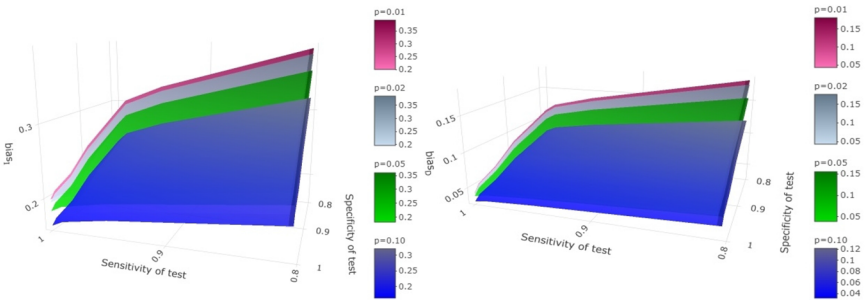
(a) $bias_I$ and $bias_D$ for $S_e^T = S_e^P, S_p^T$ 10% less than S_p^P .



(b) $bias_I$ and $bias_D$ for $S_p^T = S_p^P, S_e^T$ 10% less than S_e^P .



(c) $bias_I$ and $bias_D$ for S_p^T 10% less than S_p^P, S_e^T 10% less than S_e^P .



(d) $bias_I$ and $bias_D$ for S_p^T 20% less than S_p^P, S_e^T 20% less than S_e^P .

Fig. 4 Surface plots depicting $bias_I$ and $bias_D$ for different values of p, S_e^P, S_p^P, S_e^T and S_p^T

not a significant reduction. Thus, we may conclude that misspecification of S_p (S_e) has a large (negligible) effect on the value of π , and consequently, on the estimates of prevalence obtained by solving Eqs. (6) or (8), as the case may be. So, for accurate estimation of prevalence, it is very important to specify the value of S_p correctly, while misspecification of S_e has little effect.

Finally, we explain a subtle connection between the numbers reported in Table 2 and 3 with a specific example. Consider the case when prevalence p is equal to 0.01, and the true sensitivity (S_e^T) and specificity (S_p^T) are both equal to 0.9. From Table 2, we observe that the corresponding value of π (cf. Equation (6)) for individual testing is equal to 0.11. Suppose now the perceived sensitivity (S_e^P) and specificity (S_p^P) are equal to 1. To obtain the biased prevalence p corresponding to the perceived sensitivity and specificity, we plug in 0.11 for π on the left-hand side of Eq. (6), and $S_e = S_p = 1$ on the right-hand side. The solution can be obtained from Table 2 by observing the value of p corresponding to $S_e = S_p = 1$ and $\pi = 0.11$, which is clearly 0.11. Thus, the resulting bias due to misspecification is $0.11 - 0.01 = 0.10$, which is reported in Table 3. By using a similar argument, we can find the bias for pooled testing which is 0.01 using Table 2.

5 Discussions and concluding remarks

In this article, we have revisited the statistical theory behind Dorfman's pooled testing technique used for screening. We have explained the method of estimation of prevalence from individual testing and pooled testing data. Most importantly, we have provided insights, by looking into the practical issues that are being discussed in the scientific community, arising out of the fact that the tests for the SARS-CoV-2 virus are imperfect. We have illustrated that pooled testing is not only preferable for reducing the time and cost of screening and widening the net of testing, but it also helps in a significant reduction of misclassification among those who are tested positive. We have noted that for prevalence estimation, pooled testing is preferable to individual testing for a wide range of situations, especially when the specificity of the test is low. It helps in reducing the bias of the prevalence estimate significantly. A particularly interesting point that we have made is that for an imperfect test with a sensitivity less than 1, group testing can lead to cost savings even for a widely prevalent disease, which goes against the commonly held perception that group testing is useful only for low values of prevalence. It is worth mentioning, however, that the cost-saving benefits are the most pronounced for the lower prevalence rates.

We note that for estimating the prevalence of SARS-CoV-2 of Santa Clara County, California, Bendavid et al. (2021) collected data by conducting individual antibody tests on the 3300 subjects; a similar procedure was followed by ICMR for its serosurveys as well. Instead, had they used pooled testing technique, it would have been possible to collect the test data from a much larger sample for a similar cost and time. Also, as mentioned in the paper, given that the prevalence was quite low (estimated as 2.8%), adopting pooled testing technique would have made all the more sense in the context of their study.

As a final remark, we would like to emphasize that for countries with limited resources like India, pooled testing provides an avenue to achieve more without straining the resources too much and offers a way to follow the WHO guidance of testing enough members of the population (WHO 2020a). For the upcoming ICMR surveys in estimating COVID-19 spread in India, pooled testing can be effectively used, more so in regions

where the virus has not spread widely. In some “hot spots,” the prevalence of the disease may be seen to be more than 10% among the exposed population based on the present testing data where the savings would be less compared to areas where prevalence is less. However, when the goal is the estimation of the prevalence of the disease among the wider population, given that India’s number of cases per million population is still among the lowest in the world (Sarkar 2020; The Hindu 2020), the pooled testing methodology can help ICMR in achieving a significant reduction of cost and time, as well as in increasing accuracy and efficiency.

Appendix A: Computation of S_e^D and S_p^D

In the context of Dorfman’s testing procedure, S_e^D , the probability that a positive person can come positive is the probability that both the pooled test as well as the individual test comes correct, which is S_e^2 .

To compute the misclassification probability that a person is non-diseased and is found to be diseased, we first note that a non-diseased person, say Person j , can either be a part of a pool constituted entirely of non-diseased individuals (Case A) or a pool containing some diseased individuals (Case B). The probability of Case A is $(1 - p)^k$, and the probability of Case B is $(1 - p) - (1 - p)^k$.

Now, for Person j to be found to be diseased in Case A, the pooled test, and subsequently, the individual’s test, will both erroneously give a positive result, the probability of which is $(1 - S_p)^2$ for the original Dorfman algorithm. On the other hand, in Case B, the pooled test will be positive, but this time correctly, while the individual’s test will still be erroneously positive, the probability of which is $S_e(1 - S_p)$. Putting all together, the probability that a non-diseased person is found to be diseased is $[(1 - S_p)^2(1 - p)^k + S_e(1 - S_p)((1 - p) - (1 - p)^k)] / (1 - p) = (1 - S_p)S_e + (1 - S_p)(1 - S_p - S_e)(1 - p)^{k-1}$. Hence, S_p^D is $1 - (1 - S_p)S_e + (1 - S_p)(1 - S_p - S_e)(1 - p)^{k-1}$.

Appendix B: Derivation of the properties of $\hat{\pi}^D$

Let us define $\hat{Y}_{ij}^* = 1$ if the i th member of the j th pool is positive, and 0 otherwise, $i = 1, \dots, k$, $j = 1, \dots, J$. Therefore, we have $\hat{\pi}^D = \frac{1}{Jk} \sum_{j=1}^J \sum_{i=1}^k \hat{Y}_{ij}^*$, where $Jk = n$. As π is the probability of a positive result, $E(\hat{\pi}^D) = \pi$. This also means that $\hat{\pi}^D$ is consistent for π as $J \rightarrow \infty$, and k remains fixed. Now, noting that observations from different groups are independent,

$$\begin{aligned} V(\hat{\pi}^D) &= V\left(\frac{1}{n} \sum_{j=1}^J \sum_{i=1}^k \hat{Y}_{ij}^*\right) = \frac{1}{n^2} \sum_{j=1}^J V\left(\sum_{i=1}^k \hat{Y}_{ij}^*\right) = \frac{Jk}{n^2} V(\hat{Y}_{ij}^*) + \frac{Jk(k-1)/2}{n^2} (P(\hat{Y}_{i_1}^* = 1, \hat{Y}_{i_2}^* = 1) - \pi^2) \\ &= \frac{\pi - \pi^2}{n} + \frac{k-1}{2n} (P(\hat{Y}_{i_1}^* = 1, \hat{Y}_{i_2}^* = 1) - \pi^2), \end{aligned}$$

where

$$\begin{aligned} P(\hat{Y}_{i_1}^* = 1, \hat{Y}_{i_2}^* = 1) &= (1 - p)^k (1 - S_p)^3 (1 - p)^2 + (1 - (1 - p)^k) S_e (S_e p + (1 - S_p)(1 - p))^2 \\ &= (1 - p)^{k+2} (1 - S_p)^3 + (1 - (1 - p)^k) S_e (S_e p + (1 - S_p)(1 - p))^2. \end{aligned}$$

Therefore,

$$V(\hat{\pi}^D) = \frac{1}{n} \left(\pi - \pi^2 + \frac{k-1}{2} (1-p)^{k+2} (1-S_p)^3 + (1-(1-p)^k) S_e (S_e p + (1-S_p)(1-p))^2 \right).$$

Keeping k fixed, as $J \rightarrow \infty, n \rightarrow \infty$, and $\sqrt{n}(\hat{\pi}^D - \pi) \rightarrow N(0, C)$, where

$$C = \left(\pi - \pi^2 + \frac{k-1}{2} (1-p)^{k+2} (1-S_p)^3 + (1-(1-p)^k) S_e (S_e p + (1-S_p)(1-p))^2 \right).$$

Appendix C: Derivation of the properties of \hat{p}^D

From (9), we have $\frac{d\pi}{dp} = (S_e + S_p - 1) ((1 - S_p)k(1 - p)^{k-1} + S_e)$. Using the delta-method, we can write that.

$\hat{\pi}^D - \pi \approx (\hat{p}^D - p) \frac{d\pi}{dp}$. As long as $S_e + S_p \neq 1$, this shows that \hat{p}^D is a consistent estimator of p , and, moreover, $\sqrt{n}(\hat{p}^D - p) \rightarrow N(0, C')$, where $C' = C / \left| \frac{d\pi}{dp} \right|$.

Appendix D: Derivation of the properties of $\hat{p}_T - \hat{p}_P$

The estimates \hat{p}_T and \hat{p}_P will be obtained by plugging in $\hat{\pi}^D$ in both of the following two equations, and solving:

$$\pi = (1 - S_p^T - S_e^T) \left[(1 - S_p^T)(1 - p_T)^k - p_T S_e^T \right] + (1 - S_p^T) S_e^T \tag{C.1}$$

$$\pi = (1 - S_p^P - S_e^P) \left[(1 - S_p^P)(1 - p_P)^k - p_P S_e^P \right] + (1 - S_p^P) S_e^P \tag{C.2}$$

This tells us,

$$\hat{p}_T - p_T \approx (\hat{\pi}^D - \pi) / \frac{d\pi}{dp_T}, \text{ and } \hat{p}_P - p_P \approx (\hat{\pi}^D - \pi) / \frac{d\pi}{dp_P}.$$

In other words,

$$(\hat{p}_T - \hat{p}_P) - (p_T - p_P) \approx (\hat{\pi}^D - \pi) \left(\frac{1}{\frac{d\pi}{dp_T}} - \frac{1}{\frac{d\pi}{dp_P}} \right).$$

This tells us that $\hat{p}_T - \hat{p}_P$ will be consistent for $p_T - p_P$ for large sample sizes, and moreover,

$$\sqrt{n}((\hat{p}_T - \hat{p}_P) - (p_T - p_P)) \rightarrow N(0, C\varepsilon), \text{ where } C\varepsilon = C \left| \frac{1}{\frac{d\pi}{dp_T}} - \frac{1}{\frac{d\pi}{dp_P}} \right|.$$

Acknowledgements We thank two anonymous referees for their suggestions which have led to significant improvements of our paper.

Funding None.

Declaration

Conflict of interest The authors declare that they have no conflict of interest.

References

- Abdalhamid, B., Bilder, C.R., McCutchen, E.L., Hinrichs, S.H., Koepsell, S.A., Iwen, P.C.: Assessment of specimen pooling to conserve SARS CoV-2 testing resources. *Am. J. Clin. Pathol.* **153**, 715–718 (2020)
- Aguilar, J., Faust, J., Westafer, L., Gutierrez, J.: Investigating the impact of asymptomatic carriers on COVID-19 transmission. *medRxiv*. <https://doi.org/10.1101/2020.03.18.20037994> (2020)
- Alluri, A., Pathi, K.: India coronavirus: Should people pay for their own COVID-19 tests? BBC News (2020, April 20). <https://www.bbc.com/news/world-asia-india52322559>
- Aprahamian, H., Bish, D.R., Bish, E.K.: Optimal risk-based group testing. *Manag. Sci.* **65**, 4365–4384 (2019)
- Arnold, A.: How many people with the coronavirus are asymptomatic? The Cut (2020, April 21). <https://www.thecut.com/2020/04/how-many-people-with-the-coronavirus-are-asymptomatic.html>
- Bajak, A., Howe, J.: A study said covid wasn't that deadly. The right seized it. How coronavirus research is being weaponized. The New York Times (2020, May 14). <https://www.nytimes.com/2020/05/14/opinion/coronavirus-research-misinformation.html>
- Bendavid, E., Mulaney, B., Sood, N., Shah, S., Ling, E., Bromley-Dulfano, R., et al.: COVID-19 antibody seroprevalence in Santa Clara County, California. *Int. J. Epidemiol.* (2021). <https://doi.org/10.1093/ije/dyab010>
- Bilder, C.: Group testing for estimation. Wiley StatsRef: Statistics Reference Online (2019). <https://doi.org/10.1002/9781118445112.stat08231>
- Bilder, C., Iwen, P., Abdalhamid, B., Tebbs J., McMahan C.: Increasing testing capacity for SARS-CoV-2 by pooling specimens. *Significance* (2020, April 9). <https://www.significancemagazine.com/science/651-increasing-testing-capacity-for-SARS-CoV-2-by-pooling-specimens>
- Brent, R.: Algorithms for Minimization without Derivatives. Prentice-Hall, Englewood Cliffs, NJ (1973)
- Cohen, A.N., Kessel, B.: False positives in reverse transcription PCR testing for SARS-CoV-2. *medRxiv* (2020). <https://doi.org/10.1101/2020.04.26.20080911>
- Dorfman, R.: The detection of defective members of large populations. *Ann. Math. Stat.* **14**, 436–440 (1943)
- Feller, W.: An Introduction to Probability Theory and Its Applications, vol. 1. Wiley, London (1968)
- Ferran, M.: Coronavirus tests are pretty accurate, but far from perfect. The Conversation (2020, May 6). <https://theconversation.com/coronavirus-tests-are-pretty-accurate-but-far-from-perfect-136671>
- Finucan, H.M.: The blood testing problem. *J. r. Stat. Soc. Ser. C (appl. Stat.)* **13**(1), 43–50 (1964)
- Ghosh, A., Das Gupta, M.: Covid testing numbers on rise, but states are feeling the pinch of high RT-PCR costs. The Print (2020, July 5). <https://theprint.in/health/covid-testing-numbers-on-rise-but-states-are-feeling-the-pinch-of-high-rt-pcr-cost/454820/>
- Hanel, R., Thurner, S.: Boosting test-efficiency by pooled testing strategies for Sars-CoV-2. arXiv preprint [arXiv:2003.09944](https://arxiv.org/abs/2003.09944) (2020)
- ICMR: Advisory on feasibility of using pooled samples for molecular testing of COVID-19 (2020). <https://www.mohfw.gov.in/pdf/letterregguidanceonpoolingsamplesfortesting001.pdf>. Accessed 8 May 2020
- Ioannidis, J.P.: A fiasco in the making? As the coronavirus pandemic takes hold, we are making decisions without reliable data. *Stat.* **17** (2020, March 17). <https://www.statnews.com/2020/03/17/a-fiasco-in-the-making-as-the-coronavirus-pandemic-takes-hold-we-are-making-decisions-without-reliable-data/>
- John, T.: Iceland lab's testing suggests 50% of coronavirus cases have no symptoms. CNN (2020, April 3). <https://edition.cnn.com/2020/04/01/europe/iceland-testing-coronavirus-intl/index.html>
- Kucirka, L.M., Lauer, S.A., 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.* (2020). <https://doi.org/10.7326/M20-1495>
- Lakdawalla, D., Keeler, E., Goldman, D., Trish, E.: Getting Americans back to work (and School) with pooled testing. White paper, Leonard D. Schaeffer Center for Health Policy & Economics, University of Southern California (2020)
- Lewis, J.L., Lockary, V.M., Kobic, S.: Cost savings and increased efficiency using a stratified specimen pooling strategy for Chlamydia trachomatis and Neisseria gonorrhoeae. *Sex. Transm. Dis.* **39**, 46–48 (2012)

- Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., et al.: Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science* **368**(6490), 489–493 (2020)
- Litvak, E., Tu, X.M., Pagano, M.: Screening for the presence of a disease by pooling sera samples. *J. Am. Stat. Assoc.* **89**, 424–434 (1994)
- Lohse, S., Pfuhl, T., Berkó-Göttel, B., Rissland, J., Geißler, T., Gärtner, B., et al.: Pooling of samples for testing for Sars-CoV-2 in asymptomatic people. *Lancet Infect. Dis.* (2020). [https://doi.org/10.1016/S14733099\(20\)30362-5](https://doi.org/10.1016/S14733099(20)30362-5)
- McCormick: Why experts are questioning two hyped antibody studies in coronavirus hotspots. *The Guardian* (2020, April 23). <https://www.theguardian.com/world/2020/apr/23/coronavirus-antibody-studies-california-stanford>
- Medical Xpress, Johns Hopkins University: Asymptomatic spread of COVID-19 makes it harder to contain (2020 May 12). <https://medicalxpress.com/news/2020-05-asymptomatic-covid-harder.html>
- Mudur, G.S.: Third survey to look for Covid antibodies. *The Telegraph, India* (2020, December 12). <https://www.telegraphindia.com/india/third-survey-to-look-for-covid-antibodies/cid/1800285>
- Murhekar, M.V., et al.: Prevalence of SARS-CoV-2 infection in India: findings from the national serosurvey, May–June 2020. *Indian J. Med. Res.* **152**, 48–60 (2020)
- Pouwels, K.B., Roope, L., Barnett, A., Hunter, D.J., Nolan, T.M., Clarke, P.M.: Group testing for Sars-CoV-2: forward to the past? *PharmacoEcon Open* **4**(2), 207–210 (2020)
- Regalado, A.: Up to 4% of Silicon Valley is already infected with coronavirus. *MIT Technology Review* (2020, April 17). <https://www.technologyreview.com/2020/04/17/1000113/up-to-4-of-silicon-valley-already-infected-with-coronavirus/>
- Roy, S., Banerjee, T.: Estimation of log-odds ratio from group testing data using Firth correction. *Biom. J.* **61**, 714–728 (2019)
- Sarkar, S.: India's Covid-19 deaths per million among lowest globally: Health Ministry. *Hindustan Times* (2020, July 9). <https://www.hindustantimes.com/india-news/india-s-covid-19-cases-deaths-among-lowest-globally-per-mn-population-health-ministry/story-o86QTjmnQTzFTQLTYJjTN.html>
- Sempos, C., Tian, L.: Adjusting coronavirus prevalence estimates for laboratory test kit error. *medRxiv* (2020). <https://doi.org/10.1101/2020.05.11.20098202>
- SNS: ICMR approves IIT-Delhi's COVID-19 test, can significantly reduce cost. *The Statesman* (2020, April 25). <https://www.thestatesman.com/coronavirus/icmr-approvesiit-delhis-covid-19-test-can-significantly-reduce-cost-1502880185.html>
- Thacker, T.: No symptoms in 80% of Covid cases raise concerns. *The Economic Times* (2020, April 21). <https://economictimes.indiatimes.com/industry/healthcare/biotech/healthcare/no-symptoms-in-80-of-covid-cases-raise-concerns/articleshow/75260387.cms>
- The Hindu: India's COVID-19 cases per million among lowest in world (2020, December 13). <https://www.thehindu.com/news/national/coronavirus-indias-covid-19-cases-per-million-among-lowest-in-world/article33319889.ece>
- Wein, L.M., Zenios, S.A.: Pooled testing for HIV screening: capturing the dilution effect. *Oper. Res.* **44**, 543–569 (1996)
- World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19—11 March 2020. Geneva, Switzerland (2020a, March 11). <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19--3-march-2020>
- World Health Organization: Considerations for implementing and adjusting public health and social measures in the context of COVID-19, Interim guidance (2020b, November 4). <https://www.who.int/publications/i/item/considerations-in-adjusting-public-health-and-social-measures-in-the-context-of-covid-19-interim-guidance>
- Yelin, I., Aharony, N., Shaer-Tamar, E., Argoetti, A., Messer, E., Berenbaum, D., et al.: Evaluation of COVID-19 RT-qPCR test in multi-sample pools. *medRxiv* (2020)
- Yu, Y., Liu, Y., Luo, F., Tu, W., Zhan, D., Yu, G. et al.: COVID-19 asymptomatic infection estimation. *medRxiv* (2020). <https://doi.org/10.1101/2020.04.19.20068072>