medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

False Negative Mitigation in Group Testing for COVID-19 Screening

A. R. Alizad-Rahvar*, S. Vafadar†, M. Totonchi††, M. Sadeghi**
*School of Biological Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
†Laboratory of Biological Complex Systems and Bioinformatics (CBB), Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran
††Department of Genetics, Royan Institute for Reproductive Biomedicine, ACECR, Tehran, Iran
**National Institute for Genetic Engineering and Biotechnology, Tehran, Iran
Corresponding authors' email: *alizad@ipm.ir, **sadeghi@nigeb.ac.ir

Abstract

After lifting the COVID-19 lockdown restrictions and opening businesses, screening is essential to prevent the spread of the virus. Group testing could be a promising candidate for screening to save time and resources. However, due to the high false-negative rate (FNR) of the RT-PCR diagnostic test, we should be cautious about using group testing because a group's false-negative result identifies all the individuals in a group as uninfected. Repeating the test is the best solution to reduce the FNR, and repeats should be integrated with the group-testing method to increase the sensitivity of the test. The simplest way is to replicate the test twice for each group (the 2Rgt method). In this paper, we present a new method for group testing (the groupMix method), which integrates two repeats in the test. Then we introduce the adaptive version of both the groupMix and the 2Rgt methods. We compare these methods analytically regarding the sensitivity and the average number of tests. The tradeoff between the sensitivity and the average number of tests should be considered when choosing the best method for the screening strategy. We applied the non-adaptive groupMix method to screening 263 people and identified 2 infected individuals by performing 98 tests. This method achieved a 63% saving in the number of tests compared to individual testing. This method is currently applied to COVID-19 screening in the Clinical Genetic Laboratory at the Royan Institute, Tehran, Iran. Our experimental results show that in COVID-19 screening, the viral load can be low, and the group size should not be more than 6; otherwise, the FNR increases significantly. A web interface of the non-adaptive groupMix method is publicly available for laboratories to implement this method.

Index Terms

COVID-19 screening, false-negative mitigation, groupMix, group testing, pool testing, RT-PCR, SARS-CoV-2, sensitivity.

I. Introduction

In the post-COVID-19 era, most countries are trying to lift their lockdown restrictions. However, an infected person can remain completely asymptomatic and spread the virus. Hence, it is essential to increase screening tests to identify and quarantine the infected person and identify any other person who has been exposed to that individual. Unfortunately, the testing capacity of many countries is not sufficient, and they have to save their capacity for the second or even the third wave of the coronavirus outbreak.

In this situation, the group testing (or pool testing) technique can immediately and dramatically increase worldwide testing capacity by decreasing the required number of individual tests in a population. In this technique, separate samples are mixed together to create a sample pool. After a single diagnostic test has been performed on each group of samples, a negative result indicates that everyone in the group is uninfected; otherwise, at least one of the samples in the group is infected. Then, in conventional group testing [1], [2], an individual test on each sample in the infected group is performed to identify the infected samples. With this method, the number of tests is decreased by not testing the individual samples in an uninfected group.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Although different types of reverse-transcription polymerase chain reaction (RT-PCR) tests are the predominant diagnostic methods for detecting SARS-CoV-2, the accuracy of these methods for COVID-19 is still unknown [3]. False positives are rare for RT-PCR testing when primers and probes are designed appropriately. There are many reports of specificity of 100% for SARS-CoV-2 RT-PCR assays, based on the in vitro cross-reactivity assessment [4], [5], [6]. However, the clinical specificity, which is affected by the contamination of laboratory equipment and reagents or human error, could be less than 100%. Overall, the false positive rate (FPR) of SARS-CoV-2 RT-PCR diagnosis, without human error and contamination, can be considered to be zero.

The major problem in COVID-19 diagnosis with RT-PCR is false negatives. A false-negative rate (FNR) of up to 50% has been reported for RT-PCR-based SARS-CoV-2 diagnostic testing [4], but an FNR of 10-20% is more frequent in the reports [6], [7]. An unsuitable type of sample (e.g., a throat swab rather than a nasal swab [8]) and inadequate or inappropriate specimen collection, storage, and transport are responsible for a large portion of this high FNR. Moreover, the FNR of SARS-CoV-2 RT-PCR diagnosis is affected by the number of days that have elapsed since an individual first became infected. Indeed, the FNR is about 67% and 38% on day 4 and on the day of symptom onset (day 5), respectively. Three days after symptom onset (day 8), the FNR decreases to 20%, and then, it gradually increases again to 66% on day 21 [9].

A false-negative result puts the whole society at risk by falsely indicating that an infected person does not have an infection. Hence, this person, might move around the community and infect others. False negatives in group testing are much riskier than in individual testing. If a group of specimens is infected, each sample in the group can potentially be infected. Consequently, if the test result of this infected pool is a false negative, this result indicates that every person with a specimen in the pool is infection-free. Also, the mixing of specimens in group testing makes an infected specimen become diluted by the uninfected samples. Therefore, if the group size is not determined wisely, the infected specimen becomes undetectable, and the sensitivity of the test is reduced, resulting in false negatives (the dilution effect). Hence, the group testing methods must be made resistant to false-negative results. Unfortunately, the main focus of most of the studies in the field of group testing is only on reducing the average number of tests. Hence, more studies are needed to mitigate the effect of false negatives and increase the sensitivity of the test in group testing. This paper aims to address this need.

In this work, we propose a new method of group testing, called the groupMix method, for mitigating the false negatives. We propose non-adaptive (1 stage) and adaptive (2 stages) versions of this method. In adaptive (or sequential) group testing, we are dealing with a multi-stage procedure in which the tests of the next stage depend on the outcome of all previous stages. On the other hand, in non-adaptive group testing, all tests are known beforehand and are specified simultaneously. In general, adaptive group testing requires fewer tests, while non-adaptive group testing is more time-efficient [10]. In [11], the authors propose a non-adaptive method for false-negative mitigation of SARS-CoV-2 group testing. We propose a 2-stage adaptive version of this method and compare the sensitivity and the average number of tests of the adaptive and the non-adaptive versions of this method with those of the groupMix method. Then we present our test results for COVID-19 screening by using the non-adaptive groupMix method. Finally, we introduce our web interface, which will help laboratories to implement the groupMix method.

II. Group Testing Methods

The best practice to mitigate the effect of false negatives is to repeat the test [12]. If the false-positive results are ignored, a specimen can be diagnosed as positive if either test is positive. Provided that the FNR is 20%, and if we repeat the test on a sample twice independently, the chance of obtaining two false-negative results drops to 4%. Consequently, the sensitivity of the test increases from 80% (for a single test) to 96% (for two tests). This test replication should be integrated into the

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

group-testing procedure to mitigate the effect of false negatives. For instance, an adaptive screening procedure is introduced in [13] so that the negative groups of each stage are re-tested. If both tests of a group are negative, this group is considered as a negative group. Recently, a non-adaptive group testing method has been proposed for SARS-CoV-2 in [11] for false-negative mitigation. Here, we call this method 2-replicate group testing, denoted by 2Rgt. We introduce this method in the next section, and then, we propose an adaptive version of it. We will compare the results of our proposed method, groupMix, with the non-adaptive and adaptive 2Rgt's results.

A. Model and Notation

Assume that the prevalence (prior probability) of the disease in the population is p. We want to test N independent specimens by using group testing. The group size, i.e., the number of specimens in each pool, is n. The optimum group size, n_{opt}, minimizes the average number of tests, $\overline{T}$. The FNR of the diagnostic method for a single test is denoted by f_N, and the FPR is negligible. It is assumed that testing the pooled sample does not change f_N. In other words, the group size n is determined wisely to prevent the dilution effect and a drop in the sensitivity.

Denoting a group of samples by G, $G = 0$ means that all samples in G are uninfected. In contrast, $G = 1$ implies that G is a positive group; i.e., it has at least one infected specimen. By observing the results of group tests in different methods, we identify a group as infected or uninfected, denoted by $\hat{G} = 1$ and $\hat{G} = 0$, respectively. Here, we define two types of sensitivity. The *sample-level sensitivity*, S_s, refers to the probability of detecting an infected sample as positive. On the other hand, the *group-level sensitivity*, S_g, refers to the probability of positive detection of an infected group; i.e., $P(\hat{G} = 1 \mid G = 1)$. To be able to detect an infected sample, its corresponding pool should be diagnosed positive first, and then, its individual test should also be positive. Therefore, $S_s = S_g(1 - f_N)$. In comparing different methods of group testing, we compare S_g values to measure the accuracy of the methods.

B. Non-adaptive groupMix Method

The schematic diagram of the groupMix method is depicted in Fig. 1. This method has the following steps.

1) Pooling: We make groups of samples with the group size of n. Therefore, we have $m = \lceil N/n \rceil$ groups. We label each group as G_i where $1 \leqslant i \leqslant m$, and we call them *primary groups*.

2) Group mixing: Each primary group G_i $(2 \leqslant i \leqslant m-1)$ is mixed with G_{i-1} and G_{i+1} separately, giving the mixed groups $M_{i-1,i}$ and $M_{i,i+1}$, respectively, a group size of $2n$. In the case of the first (G_1) and last (G_m) groups, we mix them together and make $M_{m,1}$. With this form of mixing, each group exists in two mixed groups. Hence, each group will be tested twice by testing the mixed groups.

3) Group Testing: Each mixed group will be tested to identify the infected mixed groups. This form of group mixing results in m mixed groups, which equals the number of the primary groups G_i. Therefore, in this step, we need m tests.

4) Infected group detection: In this step, we detect the infected primary groups from the test results of the mixed groups. Different approaches can be applied for this purpose, based on the level of compromising the false negatives and the tradeoff between sensitivity and the number of tests. Assume that we want to detect the infected group in Fig. 1. Here, $M_{2,3}$ and $M_{3,4}$ are positive. We can have the following detection approaches.

- **Conservative group detection:** In a conservative approach, $M_{2,3} = 1$ indicates that both $\hat{G}_2$ and $\hat{G}_3$ should be considered as infected because both can cause a positive result for $M_{2,3}$. However, $\hat{G}_2 = 1$ should also make the test result for $M_{1,2}$ positive. Here, we assume pessimistically that the observed value of $M_{1,2} = 0$ is a false-negative result. Similarly, $M_{3,4} = 1$ results in $\hat{G}_3 = 1$ and $\hat{G}_4 = 1$. Hence, with this approach, the primary groups $\hat{G}_2$, $\hat{G}_3$, and $\hat{G}_4$ should be detected as positive in Fig. 1. Here, $\hat{G}_2 = 1$ and $\hat{G}_4 = 1$ are false positives due to our detection algorithm. Consequently, this method makes algorithmic false positives, increasing the number of tests. However, this approach

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

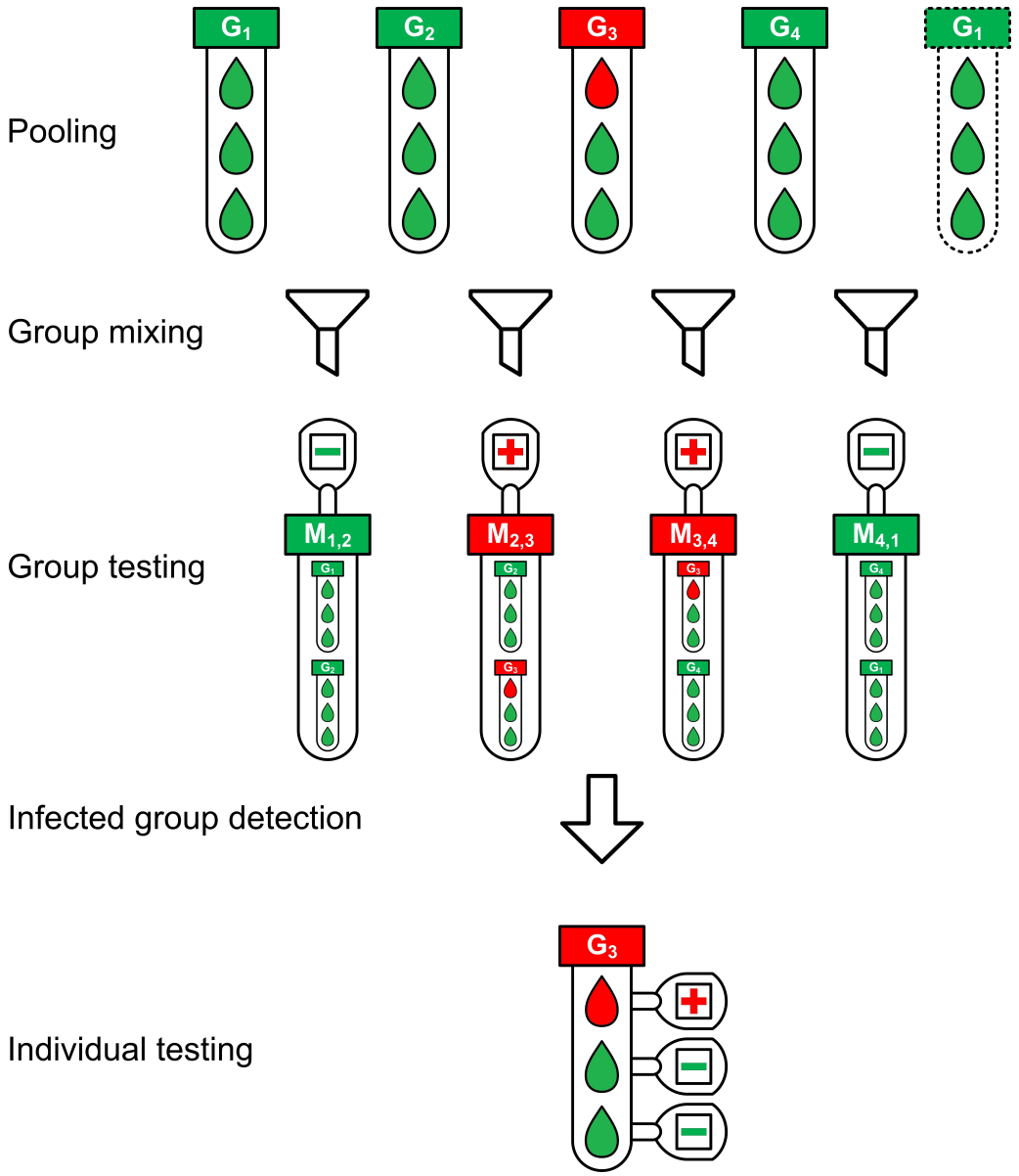

Fig. 1. The schematic diagram of the groupMix method for non-conservative infected group detection.

results in $S_g = 1 - f_N^2$. In summary, in conservative group detection, we consider $\hat{G}_i = 1$ and $\hat{G}_{i+1} = 1$ provided that $M_{i,i+1} = 1$.

• **Non-conservative group detection:** To reduce the number of algorithmic false positives, and consequently, to decrease the number of tests in conservative group detection, we should be tolerant of false negatives. To achieve this goal, we can detect only more probable infected groups, with the cost of less sensitivity. In this way, both $M_{2,3}$ and $M_{3,4}$ in Fig. 1 are positive, most likely because $G_3 = 1$. Although these positive results can also occur if $G_2 = 1$ and $G_4 = 1$, the lack of $M_{1,2} = 1$ and $M_{4,1} = 1$ makes it unlikely that G_2 and G_4 will both equal 1. In this approach, if, for example, G_2 is truly positive and $M_{1,2} = 0$ is a false-negative result, we will miss $\hat{G}_2 = 1$, resulting in less sensitivity. We will propose a 2-stage adaptive version of this method to resolve this problem and increase the sensitivity.

If there is not any false-negative result, we should always see a pair of positive mixed groups for an infected group G_i, i.e., $M_{i-1,i}$ and $M_{i,i+1}$. Therefore, provided that we observe only a single positive mixed group, i.e., $M_{i-1,i} = 0$, $M_{i,i+1} = 1$, and $M_{i+1,i+2} = 0$, then the other mixed group of the pair is not

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

diagnosed positive because of the false-negative result. Hence, a single positive mixed group $M_{i,i+1}$ indicates that both G_i and G_{i+1} can potentially be infected. Consequently, we consider both groups positive to test all of their specimens individually in the next step.

In summary, we have the following rules in non-conservative group detection.

1) Provided that the test result of both $M_{i-1,i}$ and $M_{i,i+1}$ are positive, $\hat{G}_i$ is considered as a positive group; otherwise, it is negative.
2) If we have a single positive mixed group, e.g., $M_{i,i+1}$, both $\hat{G}_i$ and $\hat{G}_{i+1}$ are considered as infected.

5) Individual testing: The specimens of all primary groups that are detected positive ($\hat{G}_i = 1$) are tested individually.

In Fig. 1, there are $N = 12$ samples with one infected sample. The pooling of samples is performed with a group size of $n = 3$. Hence, there are $m = 4$ primary groups in which G_3 is infected. Fig. 1 depicts the non-conservative group detection. We need 4 tests in the group-testing step and 3 individual tests in the last step, or a total of 7 tests. Using the conservative group detection, we need 4 tests in the group-testing step and 9 more individual tests, or a total of 13 tests. We can see the effect of the algorithmic false positives in increasing the number of tests.

C. Non-adaptive 2-replicate group test

In the non-adaptive 2-replicate group test (2Rgt) method, each group is tested twice. Provided that the result of at least one of the tests is positive, this group is diagnosed as a positive group. In this case, $S_g = 1 - f_N^2$. Note that the 2Rgt method has no algorithmic false positive.

In the case of Fig. 1, we need 11 tests if we use the 2Rgt method, while 7 and 13 tests are required if we use the non-conservative and the conservative groupMix method, respectively.

The analysis of the 2Rgt method in [11] has a mistake resulting in larger values of $\overline{T}$. We will discuss this mistake in Appendix A.

D. Adaptive groupMix Method

We propose a 2-stage adaptive version of the non-conservative groupMix method to increase its sensitivity. As explained in Sec. II-B, some cases are undetected by the non-conservative group detection. For instance, assume that $G_1 = 0$, $G_2 = 1$, $G_3 = 0$, and $G_4 = 1$. In this case, if we observe $M_{1,2} = 0$ (a false negative), $M_{2,3} = 1$, $M_{3,4} = 1$, and $M_{4,1} = 1$, the non-conservative method detects $\hat{G}_3 = 1$ and $\hat{G}_4 = 1$. Here, $\hat{G}_2 = 1$ is not detected because of the false-negative result. Since $\hat{G}_3 = 1$ is an algorithmic false positive, the individual test of its specimens will not show any infected sample. This observation indicates that $M_{2,3} = 1$ could be due to $G_2 = 1$, and implicitly shows that $M_{1,2} = 0$ is a false negative. Therefore, in the second stage, $\hat{G}_2$ is considered infected, and its specimens are tested individually. In this way, the sensitivity of the non-conservative groupMix method is increased by performing the second stage. In summary, the adaptive groupMix method is as follows.

1) Perform the non-adaptive groupMix method by using the non-conservative group detection.
2) Assume that the test of a mixed group is positive in the previous stage (e.g., $M_{i,i+1} = 1$), and that only one of its primary groups is detected as a positive group (for example, $\hat{G}_i = 0$ and $\hat{G}_{i+1} = 1$). Provided that the individual tests on the samples of the detected group (i.e., G_{i+1}) are all negative, perform the individual test on the specimens of the other undetected primary group (i.e., G_i).

This approach increases the time of the test, but time is not very critical in screening, and it is reasonable to spend more time to increase the sensitivity of the test.

E. Adaptive 2-replicate group test

Here, we present an adaptive version of the 2Rgt method to reduce the number of tests, while the sensitivity of the test, i.e., $S_g = 1 - f_N^2$, remains the same. The drawback of the non-adaptive 2Rgt

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

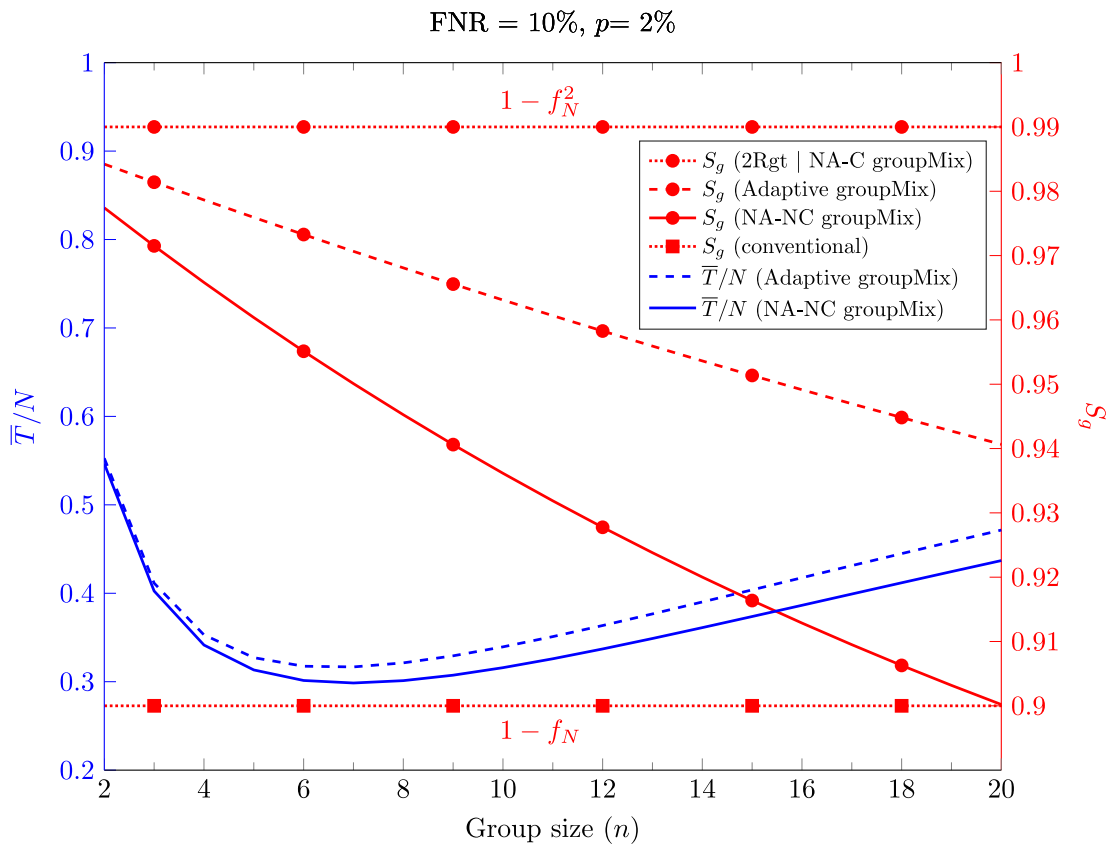

Fig. 2. The average number of tests per sample (left y-axis, blue curves) and group-level sensitivity S_g (right y-axis, red curves) of different methods against the group size.

method is that it uses two tests for each group and increases the required number of tests. Since observing only one positive result out of two tests is enough to consider a group as infected, we can separate these two tests on each group and avoid the second test for positive groups of the first stage. The summary of this method is as follows.

1) Test all groups once and identify the positive groups as infected.
2) Perform the second test only on the negative groups of the first stage and identify the positive groups.
3) Individually test the specimens of the positive groups of stages 1 and 2.

In this way, we reduce the number of tests by saving the second test of the infected groups identified in the first stage.

III. Analytical Results

In this section, we present the results of the analytical analysis of the different group-testing methods regarding S_g, and the average number of tests per sample, $\overline{T}/N$. The proof for the analyses is available in Appendix A. All analytical results are verified by simulation. In the following results, it is assumed that the FNR of the diagnostic test is 10%.

A. The effect of group size on $\overline{T}$ and S_g

First, we consider the effect of the group size on $\overline{T}/N$ and S_g in Fig. 2 for FNR = 10% and p = 2%. This figure has two different y-axes. The right axis shows S_g for the non-adaptive non-conservative (NA-NC), non-adaptive conservative (NA-C), and adaptive groupMix methods, the 2Rgt method (either the non-adaptive or adaptive version), and the conventional method. The left axis shows $\overline{T}/N$

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

for NA-NC and the adaptive groupMix methods. This figure indicates that for these specific values of FNR and prevalence, n_{opt}, which minimizes the average number of tests, is 7.

Moreover, Fig. 2 shows that S_g of both the NA-C groupMix and 2Rgt methods is independent of n, equal to $S_g = 1 - f_N^2$, which is 0.99 for $f_N = 0.1$. Similarly, the conventional group-test method has the fixed group-level sensitivity $S_g = 1 - f_N$ over n. However, S_g in either the NA-NC or adaptive groupMix method varies with n. In other words, with smaller value of n, a larger value of S_g is achieved. In this figure, it is obvious that there is a tradeoff between S_g and $\overline{T}/N$ for $n \leqslant 7$.

If the primary concern in group testing is only to reduce the average number of tests $\overline{T}$, then the group size should minimize $\overline{T}$. However, in practice, using the optimum value of n may not be possible in the presence of problems like the dilution effect. For example, the study in [14] shows that for detecting SARS-CoV-2 with the standard kits and protocols, a single infected sample can be detected in pools of up to 32 samples, with an estimated FNR of 10%. This finding could be valid for the specimen of an active symptomatic patient. However, in COVID-19 screening, we usually deal with asymptomatically-infected or pre-symptomatic individuals or people with very mild or atypical symptoms. These cases may not have a high viral load and detectable amounts of the virus in their specimen.

Consequently, in the screening scenario, the maximum group size could be much less than 32. Our experiments show that a group size of 6 is a good choice for group testing in COVID-19 screening; otherwise, the viral load in the pool becomes very low, and the FNR increases dramatically. A thorough investigation is required to determine the maximum group size for COVID-19 screening rather than for diagnostic testing for detecting active symptomatic patients.

B. Optimum group size

Figure 3(A) depicts n_{opt} against the prevalence percentage p for FNR = 10%. In the case of the groupMix methods, n_{opt} refers to the optimum group size of the primary group, i.e., G_i. Hence, the optimum size of the mixed group tested by the diagnostic method is $2n_{opt}$. Theoretically, the smaller values of p allow us to have larger groups of specimens. However, as we discussed above, the use of larger groups can reduce the sensitivity of the test.

C. Group-level sensitivity

Figure 3(B) shows S_g versus p for FNR= 10%. Overall, S_g of the NA-C groupMix method and the 2Rgt method (either non-adaptive or adaptive) does not vary with p and has the fixed value of $S_g = 1 - f_N^2$. Similarly, the conventional group-test method has the p-independent sensitivity of $S_g = 1 - f_N$. However, when we utilize the non-conservative infected group detection in the groupMix method, i.e., the NA-NC and the adaptive methods, S_g is a decreasing function of p. In non-conservative group detection, the sensitivity is compromised by accepting more false negatives to reduce the number of algorithmic false positives. The undetected cases in the non-conservative method occur more in higher values of p, resulting in less sensitivity. Even after a specific value of p, e.g., $p > 0.13$ for the NA-NC groupMix method and the FNR of 10%, S_g becomes less than that of the conventional method. Therefore, the usage of the NA-NC groupMix method is acceptable for small values of p, say, less than 5%. Moreover, Fig. 3(B) depicts that the adaptive groupMix method improves S_g significantly compared to the NA-NC method.

As we saw in Fig. 2, S_g of the NA-C groupMix, the 2Rgt, and the conventional methods does not vary with n. On the contrary, S_g of the NA-NC and the adaptive groupMix methods depends on n. For these two methods, we see S_g in Fig. 3(B) for two cases of $n = 3$ (i.e., a mixed group size of 6) and $n = n_{opt}$. For FNR = 10%, n_{opt} equals 3 for p greater than about 10%, resulting in the same S_g for two different cases of group size. For smaller values of prevalence, $n_{opt}>3$. Since S_g of the NA-NC and the adaptive groupMix methods is a decreasing function of n (Fig. 2), the group testing with $n = n_{opt}$ is less sensitive than $n = 3$ for $p < 10\%$.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

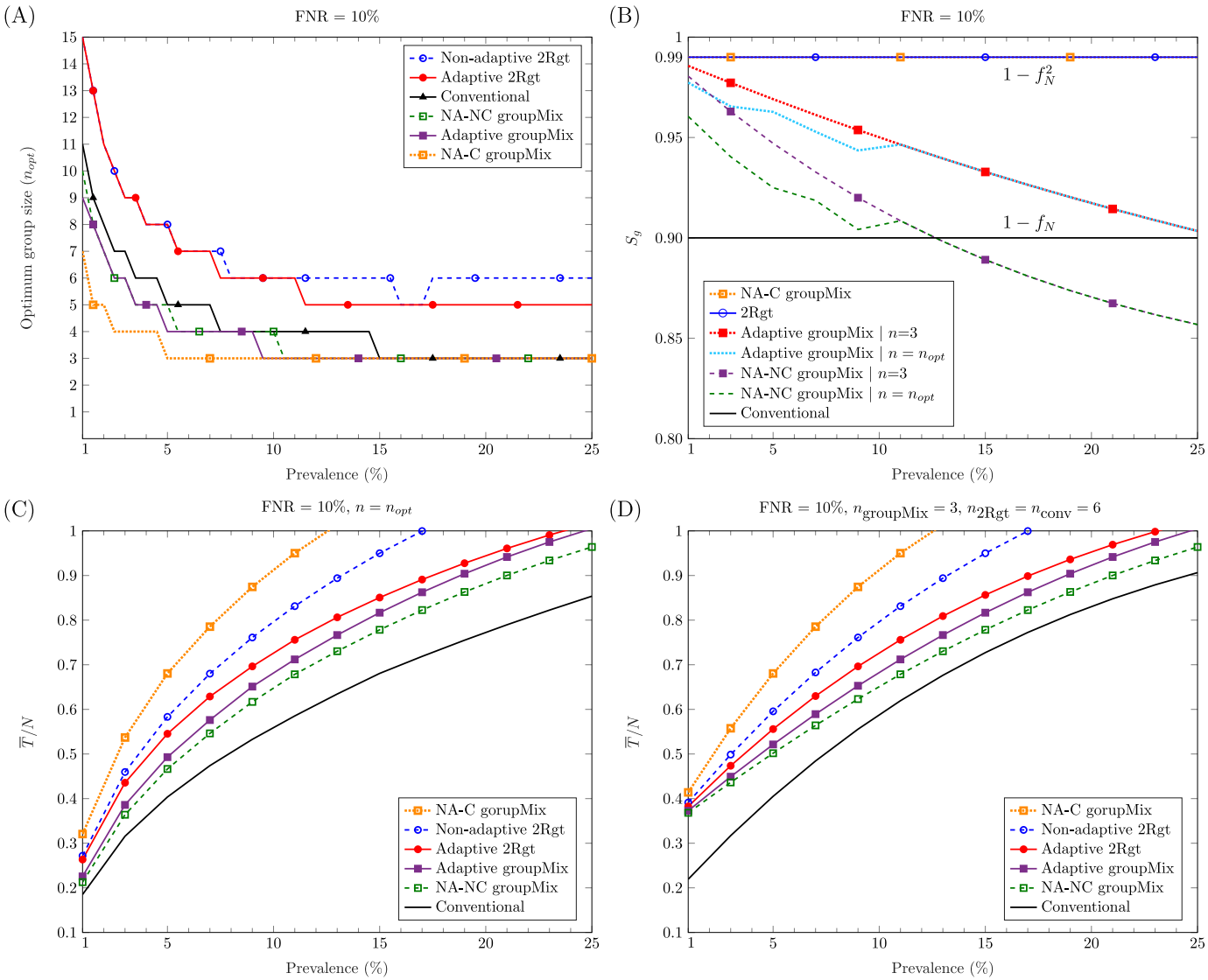

Fig. 3. Comparing different methods against the percentage of prevalence: (A) the optimum group size n_{opt} to minimize the average number of tests; (B) the group-level sensitivity S_g of the methods; (C) the average number of tests per sample for $n = n_{opt}$; (D) the average number of tests per sample for the group (or the mixed group in the groupMix method) size of 6.

D. Average number of tests

Figure 3(C) depicts the average number of tests per sample, $\overline{T}/N$, against the prevalence percentage p for $n = n_{opt}$. As discussed earlier, in a COVID-19 screening scenario, the group size should not be more than 6 to prevent the dilution effect. Hence, Fig. 3(D) shows $\overline{T}/N$ of different methods versus p for $n_{\text{groupMix}} = 3$ (the size of the mixed group is 6), $n_{\text{2Rgt}} = 6$, and $n_{\text{conventional}} = 6$. As these figures show, all methods proposed for false-negative mitigation require more tests compared to the conventional method.

Among the non-adaptive methods, the NA-NC groupMix method needs the least average number of tests. However, as mentioned before, its sensitivity reduces with p, and after a specific value of p, its sensitivity becomes even worse than that of the conventional group test method. Therefore, it is reasonable to use the NA-NC groupMix method for small values of p, say $p < 5\%$, to have both high sensitivity and less number of tests. In a comparison between the NA-C groupMix and the non-adaptive 2Rgt methods, which both have the same sensitivity of $S_g = 1 - f_N^2$, NA-C needs more tests due to algorithmic false positives. Therefore, if the goal is to reach $S_g = 1 - f_N^2$ by using a

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

non-adaptive method, the 2Rgt method is preferred.

In the case of the adaptive methods, the average number of tests in the adaptive 2Rgt method is less than that of the non-adaptive 2Rgt method. This reduction in the number of tests occurs because, in the adaptive 2Rgt method, we perform the second test only for the negative groups in the first stage, rather than using two tests for each group in the non-adaptive 2Rgt. In contrast, in groupMix methods using the non-conservative infected group detection, the adaptive version needs more tests compared to the non-adaptive one, i.e., the NA-NC method. This increase in the number of tests occurs because the adaptive version performs the full NA-NC test in the first stage, followed by the second stage.

In a comparison between the adaptive version of the groupMix and the 2Rgt methods, the former needs fewer tests than the latter. However, the sensitivity of the adaptive groupMix method is less than that of the adaptive 2Rgt, and reduces with an increase of p. Therefore, there is a tradeoff between S_g and $\overline{T}$ in decisions about choosing an adaptive group-test method for false-negative mitigation.

Considering the maximum group size (or the size of the mixed group in the groupMix method) of 6, compared to individual testing at 1% prevalence, we save 60% at the FNR of 10% for all proposed methods to mitigate the false negatives. For the optimum group size, this saving is between 70% to 80% at 1% prevalence, but it may result in less sensitive tests.

Overall, in choosing the best group-testing method to mitigate the false-negative results, we should determine the desired level of sensitivity and the amount of saving in the number of the tests. Moreover, the expected time of the test determines whether the adaptive method should be used.

IV. Experimental Methods

We performed COVID-19 screening in the Institute for Research in Fundamental Sciences (IPM), Tehran, Iran. A total of 263 individuals participated in the screening on five different days. These groups consisted of 78, 30, 31, 66, and 58 individuals, respectively. Swabs from the throat were collected and sent to the Clinical Genetic Laboratory at the Royan Institute, Tehran, Iran. Samples were collected between June 15-30, 2020.

The samples were pooled into the primary groups with a size of 3 prior to RNA extraction. Then, the primary groups were mixed according to the groupMix method's procedure, resulting in mixed groups of size 6.

A volume of 200 μL of the sample was mixed with 600 μL Lysis buffer, and RNA was extracted by using the Norgen Cell-Free RNA Purification Kit (Cat. No: 56300). 50 μL of Elution buffer was used in the extraction procedure.

The Novel Coronavirus (2019-nCoV) Real-Time Multiplex kit (Liferiver, Cat. No: ZJ0009) was utilized for real-time RT-PCR diagnosis. This kit detects the presence of SARS-CoV-2 RNA by detecting the three genes N, OFR1ab, and E. Reactions were heated to 45°C for 10 minutes (1 cycle) for reverse transcription and denatured in 95°C for 90 seconds (1 cycle). Then, 45 cycles of amplification were carried out in 95°C for 15 seconds and 58°C for 30 seconds. Fluorescence was measured at 58°C.

The non-adaptive non-conservative (NA-NC) groupMix method was performed to detect the infected specimens.

This study was approved by the ethical committee of the Royan institute with a waiver of informed consent due to de-identified nature of the data.

V. Experimental Results

On the second, fourth, and fifth days, all groups with 30, 66, and 58 individuals were negative. For these days, we performed 10, 22, and 20 tests, respectively. Therefore, we saved 67%, 67%, and 66% in the number of tests, compared to individual testing.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

In the first group, consisting of 78 individuals, we had 26 primary and mixed groups. With the NA-NC groupMix method, only the mixed group $M_{20,21}$ was positive. Since we should always have a pair of positive mixed groups, this result indicates that one mixed group is a false negative. Therefore, both primary groups G_{20} and G_{21} were candidates to be infected. Ultimately, by performing the individual test on these primary groups, one specimen in G_{20} was identified as infected. Consequently, we performed 32 tests to detect one infected sample out of 78 samples. Indeed, we had a 59% saving in the number of tests, compared to individual testing.

Regarding the third group consisting of 31 individuals, the mixed groups $M_{4,5}$ and $M_{5,6}$ were positive. Hence, G_5 was detected as the infected group. The individual test on G_5 revealed the presence of one infected sample in this group. Hence, we detected this infected sample by using 14 tests (i.e., a 55% saving).

In conclusion, 2 positive samples were successfully identified out of 263 by using 98 tests, i.e., a 63% saving in the number of tests compared to individual testing. Since the average prevalence in this population was about 0.8%, the average group-level sensitivity S_g of this screening was very close to $1 - f_N^2$.

VI. Discussion and Conclusion

In this paper, we investigated false-negative mitigation in group testing, focusing on massive testing for COVID-19 screening. Group testing is receiving attention as a strategy that can save time or resources for COVID-19 testing. Indeed, group testing is reasonable for screening since the percentage of infected people is very low, resulting in substantial savings. However, the high rate of false-negative results in RT-PCR-based COVID-19 diagnostic tests has been widely addressed recently as an important public health-related problem. This problem needs more attention in group testing because a false-negative result for a group of potentially infected individuals does not lead to isolating them, and they can quickly spread the virus in their community.

To mitigate the false-negative results and to increase the sensitivity of the group testing, we studied different strategies for repeating the test. By implementing r replicates in the test design, we can achieve the maximum group-level sensitivity of $S_g = 1 - f_N^r$, but there is a tradeoff between the sensitivity and the average number of tests; hence, we can reduce the average number of tests by compromising the false negatives. However, this compromising is negligible for the very low percentage of infected samples that usually occurs in screening.

Typically, in group-testing studies, researchers propose different methods for pooling and identifying the infected groups. Then, they find the optimum group size to minimize the average number of tests. However, we explained that in COVID-19 screening, the high viral load might not be available because we are dealing with asymptomatic, pre-symptomatic, or mild-symptomatic people. This fact is crucial in group testing because the dilution effect in the pooling of specimens can cause more false-negative cases. Therefore, the theoretical optimum group size of different methods may yield a severe dilution effect and high FNR in the screening scenario. Our experimental studies showed that the maximum group size in COVID-19 screening should be 6. However, systematic studies are still required to determine the maximum group size for COVID-19 screening.

VII. groupMix Web Interface

A web interface is developed to help laboratories implement the NA-NC groupMix method. This website is freely accessible at http://groupmix.ipm.ir. This web interface will be elaborated in Appendix B.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

References

[1] R. Dorfman, "The detection of defective members of large populations," *Ann. Math. Statist.*, vol. 14, no. 4, pp. 436–440, 12 1943.

[2] B. Cahoon-Young, A. Chandler, T. Livermore, J. Gaudino, and R. Benjamin, "Sensitivity and specificity of pooled versus individual sera in a human immunodeficiency virus antibody prevalence study," *J. Clin. Microbiol.*, vol. 27, no. 8, pp. 1893–1895, Aug 1989.

[3] V. C. Bachelet, "Do we know the diagnostic properties of the tests used in COVID-19? A rapid review of recently published literature," *Medwave*, vol. 20, no. 3, p. e7890, Apr 2020.

[4] A. N. Cohen and B. Kessel, "False positives in reverse transcription pcr testing for SARS-CoV-2," *medRxiv*, 2020. [Online]. Available: https://www.medrxiv.org/content/early/2020/05/20/2020.04.26.20080911

[5] V. M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D. K. Chu, T. Bleicker, S. Brunink, J. Schneider, M. L. Schmidt, D. G. Mulders, B. L. Haagmans, B. van der Veer, S. van den Brink, L. Wijsman, G. Goderski, J. L. Romette, J. Ellis, M. Zambon, M. Peiris, H. Goossens, C. Reusken, M. P. Koopmans, and C. Drosten, "Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR," *Euro Surveill.*, vol. 25, no. 3, 01 2020.

[6] J. P. Broughton, X. Deng, G. Yu, C. L. Fasching, J. Singh, J. Streithorst, A. Granados, A. Sotomayor-Gonzalez, K. Zorn, A. Gopez, E. Hsu, W. Gu, S. Miller, C.-Y. Pan, H. Guevara, D. Wadford, J. Chen, and C. Y. Chiu, "Rapid detection of 2019 novel coronavirus SARS-CoV-2 using a crispr-based detectr lateral flow assay," *medRxiv*, 2020. [Online]. Available: https://www.medrxiv.org/content/early/2020/03/27/2020.03.06.20032334

[7] D. Li, D. Wang, J. Dong, N. Wang, H. Huang, H. Xu, and C. Xia, "False-Negative Results of Real-Time Reverse-Transcriptase Polymerase Chain Reaction for Severe Acute Respiratory Syndrome Coronavirus 2: Role of Deep-Learning-Based CT Diagnosis and Insights from Two Cases," *Korean J Radiol*, vol. 21, no. 4, pp. 505–508, 04 2020.

[8] Y. Yang, M. Yang, C. Shen, F. Wang, J. Yuan, J. Li, M. Zhang, Z. Wang, L. Xing, J. Wei, L. Peng, G. Wong, H. Zheng, M. Liao, K. Feng, J. Li, Q. Yang, J. Zhao, Z. Zhang, L. Liu, and Y. Liu, "Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections," *medRxiv*, 2020.

[9] L. M. Kucirka, S. A. Lauer, O. Laeyendecker, D. Boon, and J. Lessler, "Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure," *Ann. Intern. Med.*, May 2020.

[10] D. Z. Du and F. K. Hwang, *Pooling designs and nonadaptive group testing : important tools for DNA sequencing*. New Jersey: World Scientific, 2006.

[11] R. Hanel and S. Thurner, "Boosting test-efficiency by pooled testing strategies for SARS-CoV-2," *arXiv*, 2020. [Online]. Available: https://arxiv.org/abs/2003.09944v1

[12] K. Ramdas, A. Darzi, and S. Jain, "'Test, re-test, re-test': using inaccurate tests to greatly increase the accuracy of COVID-19 testing," *Nat. Med.*, May 2020.

[13] E. Litvak, X. M. Tu, and M. Pagano, "Screening for the presence of a disease by pooling sera samples," *Journal of the American Statistical Association*, vol. 89, no. 426, pp. 424–434, 1994.

[14] I. Yelin, N. Aharony, E. Shaer Tamar, A. Argoetti, E. Messer, D. Berenbaum, E. Shafran, A. Kuzli, N. Gandali, O. Shkedi, T. Hashimshony, Y. Mandel-Gutfreund, M. Halberthal, Y. Geffen, M. Szwarcwort-Cohen, and R. Kishony, "Evaluation of COVID-19 RT-qPCR test in multi-sample pools," *Clin. Infect. Dis.*, May 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Appendix A
Analysis

In our analysis, we present S_g, and the average number of tests per sample, i.e., $\overline{T}/N$, for each method. Then, we explain how to obtain n_{opt}. First, we introduce some notations and basic equations to prove our analysis accordingly.

- $\#(x) \triangleq$ the number of the occurrence of event x
- $E[X] \triangleq$ the expected value of a random variable X
- $m \triangleq \#(\text{groups}) = N/n$ (m may not equal an integer, however, we do not round it up in our analysis, so that we can calculate the derivatives and simplify the equations. We will round up the all integer size values at the end of our analysis).
- TP_g: the detected positive group is a true positive
- FP_g: the detected positive group is a false positive
- The probability of having a positive group is
$$P_g \triangleq P(G=1) = 1 - P(G=0) = 1 - (1-p)^n.$$
- Group-level sensitivity (true-positive rate TPR_g):
$$S_g \triangleq P(\hat{G}=1 \mid G=1) = \frac{P(\hat{G}=1, G=1)}{P(G=1)} = \frac{P(\text{TP}_g)}{P_g}.$$
This equation yields $P(\text{TP}_g) = P_g S_g$.
- Group-level FPR:
$$\text{FPR}_g \triangleq P(\hat{G}=1 \mid G=0) = \frac{P(\hat{G}=1, G=0)}{P(G=0)} = \frac{P(\text{FP}_g)}{1-P_g}.$$
Hence, $P(\text{FP}_g) = (1-P_g)\ \text{FPR}_g$.
- $P_g^+ \triangleq P(\hat{G}=1) = P(\text{TP}_g) + P(\text{FP}_g) = S_g P_g + (1-P_g)\ \text{FPR}_g$.
- $n_g^+ \triangleq E[\#(\text{groups diagnosed positive})] = mP_g^+$.
- Calculation of $\overline{T}/N$ for non-adaptive group-test methods:
$$\begin{aligned}\overline{T} \triangleq E[T] &= \#(\text{group tests}) + E[\#(\text{individual tests})] \\ &= \#(\text{tests on each group}) \times m + E[\#(\text{groups diagnosed positive})] \times n \\ &= rm + n_g^+ n \\ \Rightarrow \frac{\overline{T}}{N} &= \frac{r}{n} + P_g^+, \end{aligned} \tag{1}$$
where $r = 1$ for the groupMix and the conventional methods, and $r = 2$ for the 2Rgt method.
- $\hat{M}_{i,i+1}$: the test result of the mixed group $M_{i,i+1}$

A. *Analysis of the conventional group-test method*

1) S_g:
$$\begin{aligned} S_g^{\text{conv}} &= P(\hat{G}=1 \mid G=1) \\ &= 1 - P(\hat{G}=0 \mid G=1) \\ &= 1 - f_N. \end{aligned}$$

2) $\overline{T}/N$: According to Eq. (1) we have
$$\begin{aligned} \frac{\overline{T}_{\text{conv}}}{N} &= \frac{1}{n} + P_g S_g^{\text{conv}} \\ &= \frac{1}{n} + (1-f_N) P_g. \end{aligned}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

B. *Analysis of the non-adaptive 2Rgt method*

Assume that $\hat{G}^{(1)}$ and $\hat{G}^{(2)}$ are the results of the first and the second test of G. $\hat{G}^{(1)}$ and $\hat{G}^{(2)}$ are dependent because they both depend on whether G is infected or not. However, they are conditionally independent given G; i.e.,

$$P(\hat{G}^{(1)}, \hat{G}^{(2)} \mid G) = P(\hat{G}^{(1)} \mid G) P(\hat{G}^{(2)} \mid G).$$

1) S_g:

$$\begin{aligned} S_g^{\text{NA-2Rgt}} &= P(\hat{G} = 1 \mid G = 1) \\ &= 1 - P(\hat{G} = 0 \mid G = 1) \\ &= 1 - P(\hat{G}^{(1)} = 0, \hat{G}^{(2)} = 0 \mid G = 1) \\ &= 1 - P(\hat{G}^{(1)} = 0 \mid G = 1) P(\hat{G}^{(2)} = 0 \mid G = 1) \\ &= 1 - f_N^2. \end{aligned}$$

2) $\overline{T}/N$: According to Eq. (1) we have

$$\begin{aligned} \frac{\overline{T}_{\text{NA-2Rgt}}}{N} &= \frac{2}{n} + P_g S_g^{\text{NA-2Rgt}} \\ &= \frac{2}{n} + (1 - f_N^2) P_g. \end{aligned}$$

As mentioned in Section II-C, there is a mistake in the analysis of the 2Rgt method in [11]. The researchers consider $\hat{G}^{(1)}$ and $\hat{G}^{(2)}$ independent. Therefore, there is no conditioning on G in their formulation, resulting in a larger value of $\overline{T}/N$ in their analysis.

C. *Analysis of the Adaptive 2Rgt method*

1) S_g: Assume that $\hat{G}^{(1)}$ and $\hat{G}^{(2)}$ are the test results of G in the first and the second stage of the adaptive 2Rgt method.

$$\begin{aligned} S_g^{\text{A-2Rgt}} &= P(\hat{G} = 1 \mid G = 1) \\ &= P(\hat{G}^{(1)} = 1 \mid G = 1) + P(\hat{G}^{(1)} = 0, \hat{G}^{(2)} = 1 \mid G = 1) \\ &= P(\hat{G}^{(1)} = 1 \mid G = 1) + P(\hat{G}^{(1)} = 0 \mid G = 1) P(\hat{G}^{(2)} = 1 \mid G = 1) \\ &= (1 - f_N) + f_N (1 - f_N) \\ &= 1 - f_N^2. \end{aligned}$$

S_g of the adaptive 2Rgt is the same as that of the non-adaptive 2Rgt method.

2) $\overline{T}/N$:

$$\begin{aligned} \overline{T}_{\text{A-2Rgt}} &= \#(\text{groups}) + \#(\text{groups diagnosed negative in 1st stage}) + E[\#(\text{individual tests})] \\ &= m + m[1 - P_g(1 - f_N)] + (n_g^+ \times n) \\ \Rightarrow \frac{\overline{T}_{\text{A-2Rgt}}}{N} &= [2 - P_g(1 - f_N)]\frac{1}{n} + (1 - f_N^2) P_g. \end{aligned}$$

The term $2 - P_g(1 - f_N)$ on the right-hand side is less than 2; hence, $\overline{T}/N$ in the adaptive 2Rgt method is less than that in the non-adptive 2Rgt method. Indeed, we save on average $NP_g(1 - f_N)/n$ tests with the adaptive approach.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

D. *Analysis of the NA-C groupMix method*

In the groupMix method, $\hat{M}_{i-1,i}$ and $\hat{M}_{i,i+1}$ are conditionally independent given G_i, i.e.,

$$P(\hat{M}_{i-1,i}, \hat{M}_{i,i+1} \mid G_i) = P(\hat{M}_{i-1,i} \mid G_i) P(\hat{M}_{i,i+1} \mid G_i).$$

Moreover, we use the following conditional probabilities in the proof of our analyses for the groupMix method.

1) $P(\hat{M}_{i-1,i} = 1 \mid G_i = 1) = 1 - f_N$
2) $P(\hat{M}_{i-1,i} = 0 \mid G_i = 1) = f_N$
3) $P(\hat{M}_{i-1,i} = 1 \mid G_i = 0) = P(G_{i-1} = 1) P(\hat{M}_{i-1,i} = 1 \mid G_{i-1} = 1, G_i = 0) = P_g(1 - f_N)$
4) $P(\hat{M}_{i-1,i} = 0 \mid G_i = 0) = 1 - P_g(1 - f_N)$

The same equations are valid for $\hat{M}_{i,i+1}$.

1) S_g:

$$\begin{aligned} S_g^{\text{NA-C}} &= P(\hat{G}_i = 1 \mid G_i = 1) \\ &= 1 - P(\hat{G}_i = 0 \mid G_i = 1) \\ &= 1 - P(\hat{M}_{i-1,i} = 0, \hat{M}_{i,i+1} = 0 \mid G_i = 1) \\ &= 1 - P(\hat{M}_{i-1,i} = 0 \mid G_i = 1) P(\hat{M}_{i,i+1} = 0 \mid G_i = 1) \\ &= 1 - f_N^2. \end{aligned}$$

2) $\overline{T}/N$: There is an algorithmic false positive in the groupMix method. Hence, we should first obtain the $\text{FPR}_g^{\text{NA-C}}$ to calculate P_g^+, and then calculate $\overline{T}/N$.

$$\begin{aligned} \text{FPR}_g^{\text{NA-C}} &= P(\hat{G}_i = 1 \mid G_i = 0) \\ &= 1 - P(\hat{G}_i = 0 \mid G_i = 0) \\ &= 1 - P(\hat{M}_{i-1,i} = 0, \hat{M}_{i,i+1} = 0 \mid G_i = 0) \\ &= 1 - P(\hat{M}_{i-1,i} = 0 \mid G_i = 0) P(\hat{M}_{i,i+1} = 0 \mid G_i = 0) \\ &= 1 - [1 - P_g(1 - f_N)]^2. \end{aligned}$$

Then, by using Eq. (1) we have

$$\begin{aligned} \frac{\overline{T}_{\text{NA-C}}}{N} &= \frac{1}{n} + P_g^+ \\ &= \frac{1}{n} + P_g S_g^{\text{NA-C}} + (1 - P_g)\ \text{FPR}_g^{\text{NA-C}} \\ &= \frac{1}{n} + (3 - 2f_N - f_N^2) P_g + (-3 + 4f_N - f_N^2) P_g^2 + (1 - f_N)^2 P_g^3. \end{aligned}$$

E. *Analysis of the NA-NC groupMix method*

1) S_g: In the NA-NC groupMix method, the primary group G_i is detected positive if the test results of both $M_{i-1,i}$ and $M_{i,i+1}$ are positive, or if we have a single positive mixed group, i.e., either $M_{i-1,i}$ or $M_{i,i+1}$. Therefore, we can calculate S_g as follows:

$$\begin{aligned} S_g^{\text{NA-NC}} &= P(\hat{G}_i = 1 \mid G_i = 1) \\ &= P(\hat{M}_{i-1,i} = 1, \hat{M}_{i,i+1} = 1 \mid G_i = 1) \\ &+ P(\hat{M}_{i-2,i-1} = 0, \hat{M}_{i-1,i} = 1, \hat{M}_{i,i+1} = 0 \mid G_i = 1) \\ &+ P(\hat{M}_{i-1,i} = 0, \hat{M}_{i,i+1} = 1, \hat{M}_{i+1,i+2} = 0, \mid G_i = 1). \end{aligned}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The second and the third probabilities in the above equation are equal. Hence, we have

$$\begin{aligned} S_g^{\text{NA-NC}} &= P(\hat{M}_{i-1,i} = 1 \mid G_i = 1)P(\hat{M}_{i,i+1} = 1 \mid G_i = 1) \\ &+ 2P(G_{i-1} = 0)P(\hat{M}_{i-2,i-1} = 0, \hat{M}_{i-1,i} = 1, \hat{M}_{i,i+1} = 0 \mid G_{i-1} = 0, G_i = 1) \\ &+ 2P(G_{i-1} = 1)P(\hat{M}_{i-2,i-1} = 0, \hat{M}_{i-1,i} = 1, \hat{M}_{i,i+1} = 0 \mid G_{i-1} = 1, G_i = 1). \end{aligned}$$

By applying the multiplying property of the conditional independence and using the equations explained at the beginning of Section A-D, we obtain

$$S_g^{\text{NA-NC}} = (1 - f_N^2) + 2(P_g^2 - 2P_g)f_N(1 - f_N)^2.$$

2) $\overline{T}/N$: For calculating $\text{FPR}_g^{\text{NA-NC}}$, the equations are like the equations for calculating S_g, except that the condition is $G_i = 0$. Ultimately, we obtain

$$\text{FPR}_g^{\text{NA-NC}} = P_g^2(1 - f_N)^2 + 2P_g[1 - P_g(1 - f_N)]f_N(1 - f_N).$$

Therefore, we have

$$\begin{aligned} \frac{\overline{T}_{\text{NA-NC}}}{N} &= \frac{1}{n} + P_g S_g^{\text{NA-NC}} + (1 - P_g)\ \text{FPR}_g^{\text{NA-NC}} \\ &= \frac{1}{n} + (1 + 2f_N - 3f_N^2)P_g + (1 - 10f_N + 15f_N^2 - 6f_N^3)P_g^2 + (-1 + 6f_N - 9f_N^2 + 4f_N^3)P_g^3. \end{aligned}$$

F. ***Analysis of the adaptive groupMix method***

In the adaptive groupMix method, we first perform the NA-NC method, and then, the second stage of the test is done based on the results of the first stage. Therefore, the S_g and $\overline{T}/N$ values of the adaptive groupMix method are calculated by using the sum of those values in the NA-NC method with the added values obtained by performing the second stage. In the following analyses, the superscript indexes (1) and (2) denote the first and the second stage of the adaptive method, respectively.

1) S_g:

$$\begin{aligned} S_g^{\text{A-groupMix}} &= \frac{P(\text{TP}_g)}{P_g} = \frac{P(\text{TP}_g^{(1)}) + P(\text{TP}_g^{(2)})}{P_g} \\ &= P(\hat{G}^{(1)} = 1 \mid G = 1) + P(\hat{G}^{(2)} = 1 \mid G = 1) \\ &= S_g^{\text{NA-NC}} + S_g^{(2)}, \end{aligned}$$

where $\text{TP}_g^{(1)}$ and $\text{TP}_g^{(2)}$ mean the true-positive detection of G that is detected in the first and second stages, respectively.

In the second stage, the following situations yield a true positive detection of G_i ($\hat{M}$ values are the test results of the mixed groups in the first stage):

$$\begin{aligned} S_g^{(2)} &= P(\hat{G}_i^{(2)} = 1 \mid G_i = 1) \\ &= P(\hat{M}_{i-2,i-1} = 1, \hat{M}_{i-1,i} = 1, \hat{M}_{i,i+1} = 0, G_{i-1} = 0 \mid G_i = 1) && \text{(I)} \\ &+ P(\hat{M}_{i-2,i-1} = 1, \hat{M}_{i-1,i} = 1, \hat{M}_{i,i+1} = 0, G_{i-1} = 1, \hat{G}_{i-1} = 0 \mid G_i = 1) && \text{(II)} \\ &+ P(\hat{M}_{i-1,i} = 0, \hat{M}_{i,i+1} = 1, \hat{M}_{i+1,i+2} = 1, G_{i+1} = 0 \mid G_i = 1) && \text{(III)} \\ &+ P(\hat{M}_{i-1,i} = 0, \hat{M}_{i,i+1} = 1, \hat{M}_{i+1,i+2} = 1, G_{i+1} = 1, \hat{G}_{i+1} = 0 \mid G_i = 1). && \text{(IV)} \end{aligned}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Because of the symmetry, the probabilities (I) and (III), and (II) and (IV) are equal. Therefore, we calculate only the probabilities (I) and (II) separately and then obtain the value of $S_g^{(2)}$.

$$\begin{aligned}
\text{(I)}: \quad & P(\hat{M}_{i-2,i-1}=1, \hat{M}_{i-1,i}=1, \hat{M}_{i,i+1}=0, G_{i-1}=0 \mid G_i=1) \\
&= P(G_{i-1}=0)P(\hat{M}_{i-2,i-1}=1, \hat{M}_{i-1,i}=1, \hat{M}_{i,i+1}=0 \mid G_{i-1}=0, G_i=1) \\
&= (1-P_g)P(\hat{M}_{i-2,i-1}=1 \mid G_{i-1}=0)P(\hat{M}_{i-1,i}=1 \mid G_{i-1}=0, G_i=1)P(\hat{M}_{i,i+1}=0 \mid G_i=1) \\
&= P_g(1-P_g)f_N(1-f_N)^2.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\text{(II)}: \quad & P(\hat{M}_{i-2,i-1}=1, \hat{M}_{i-1,i}=1, \hat{M}_{i,i+1}=0, G_{i-1}=1, \hat{G}_{i-1}=0 \mid G_i=1) \\
&= P(G_{i-1}=1, \hat{G}_{i-1}=0)P(\hat{M}_{i-2,i-1}=1, \hat{M}_{i-1,i}=1, \hat{M}_{i,i+1}=0 \mid G_{i-1}=1, G_i=1) \\
&= f_N(1-f_N)^2 \times P(G_{i-1}=1, \hat{G}_{i-1}=0).
\end{aligned}$$

In the above equation, $G_{i-1}=1$ and $\hat{G}_{i-1}=0$ means that the primary group G_{i-1} has at least one infected sample, but the individual test results of the infected samples in the first stage are all false negative. Therefore, $P(G_{i-1}=1, \hat{G}_{i-1}=0)$ denoted by $\sum(n,p,f_N)$ is as follows:

$$\sum(n,p,f_N) \triangleq P(G_{i-1}=1, \hat{G}_{i-1}=0) = \sum_{k=1}^{n} \binom{n}{k} (pf_N)^k (1-p)^{n-k},$$

where pf_N is the probability of a sample to be infected and diagnosed negative falsely.

Ultimately, we can obtain S_g of the adaptive groupMix method by using the above results.

$$\begin{aligned}
S_g^{\text{A-groupMix}} &= S_g^{\text{NA-NC}} + S_g^{(2)} \\
&= (1-f_N^2) + 2\left[(1-3P_g+P_g^2) + \sum(n,p,f_N)\right] f_N(1-f_N)^2.
\end{aligned}$$

2) $\overline{T}/N$*:* Similarly, we have

$$\text{FPR}_g^{\text{A-groupMix}} = \text{FPR}_g^{\text{NA-NC}} + \text{FPR}_g^{(2)}.$$

Equations (I) to (IV) used in calculation of $S_g^{(2)}$ can be used for calculating $\text{FPR}_g^{(2)}$ by replacing the condition by $G_i=0$. After performing the same calculations, we have

$$\text{FPR}_g^{\text{A-groupMix}} = P_g^2(1-f_N)^2 + 2[1-P_g(1-f_N)](1-f_N)\left[P_g f_N + (1-f_N)\sum(n,p,f_N)\right].$$

Finally,

$$\frac{\overline{T}_{\text{A-groupMix}}}{N} = \frac{1}{n} + P_g S_g^{\text{A-groupMix}} + (1-P_g)\ \text{FPR}_g^{\text{A-groupMix}}.$$

G. ***Finding the optimum group size***

To find the optimum group size n_{opt}, we should find the minimum value of $\overline{T}$ against n. Hence, we need to find the solution of $d\overline{T}/dn=0$. This solution is the same as the solution of $d(\overline{T}/N)/dn=0$.

By using Eq. (1) for non-adaptive group-testing methods, we have

$$\begin{aligned}
&\frac{d(\overline{T}/N)}{dn} = 0 \\
&\Rightarrow \frac{-r}{n^2} + \frac{dP_g^+}{dP_g} \times \frac{dP_g}{dn} = 0,
\end{aligned}$$

where

$$\frac{dP_g}{dn} = \frac{-1}{n}(1-P_g) \times \text{Ln}\,(1-P_g)\,.$$

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

n_{opt} can be founf by solving the above equation.

In the case of the adaptive groupMix method, P_g^+ is not only a function of P_g, and the derivative of $\sum(n, p, f_N)$ does not have a closed-form solution. Therefore, for this method, we find n_{opt} by using numerical computation. For the adaptive 2Rgt method, the calculation of $d(\overline{T}/N)/dn$ is straightforward when the above formula for dP_g/dn is used. Ultimately, the value of n_{opt} obtained by solving the equation $d(\overline{T}/N)/dn = 0$ should be rounded to the closest integer number.

Appendix B
groupMix Web Interface

In this appendix, we explain different parts of the groupMix web interface. In this web, the non-adaptive non-conservative (NA-NC) groupMix method is implemented. The following sections in this web interface can be navigated using the left side of the website.

A. Create Groups

In the "Create Groups" section (Fig. B.1), we should set three input values of the number of specimens, prevalence (%), and FNR (%). Then, the optimum size of the primary groups, i.e.,n_{opt}, the number of groups, and the average number of tests is calculated according to the provided input values. Under these values, the mixed groups and their corresponding specimen numbers are depicted. The group size can be changed to a different value.

For instance, in Fig. B.2, the input values, and the group size are set according to the case of group testing depicted in Fig. 1, where we have 12 samples and four mixed groups. After performing the tests on mixed groups, we can set the test results by clicking on the green icons in the "Test Results" column. The green icon indicates that the mixed group is uninfected, and it turns into a red color, indicating infected, by clicking on the green icon. After determining the infected mixed groups, the index of specimens that should be tested by individual testing, i.e., the samples of infected primary groups detected by the non-conservative group detection method, are listed at the bottom of the page. Finally, this test setup can be saved by clicking on the "Save Test Line" button.

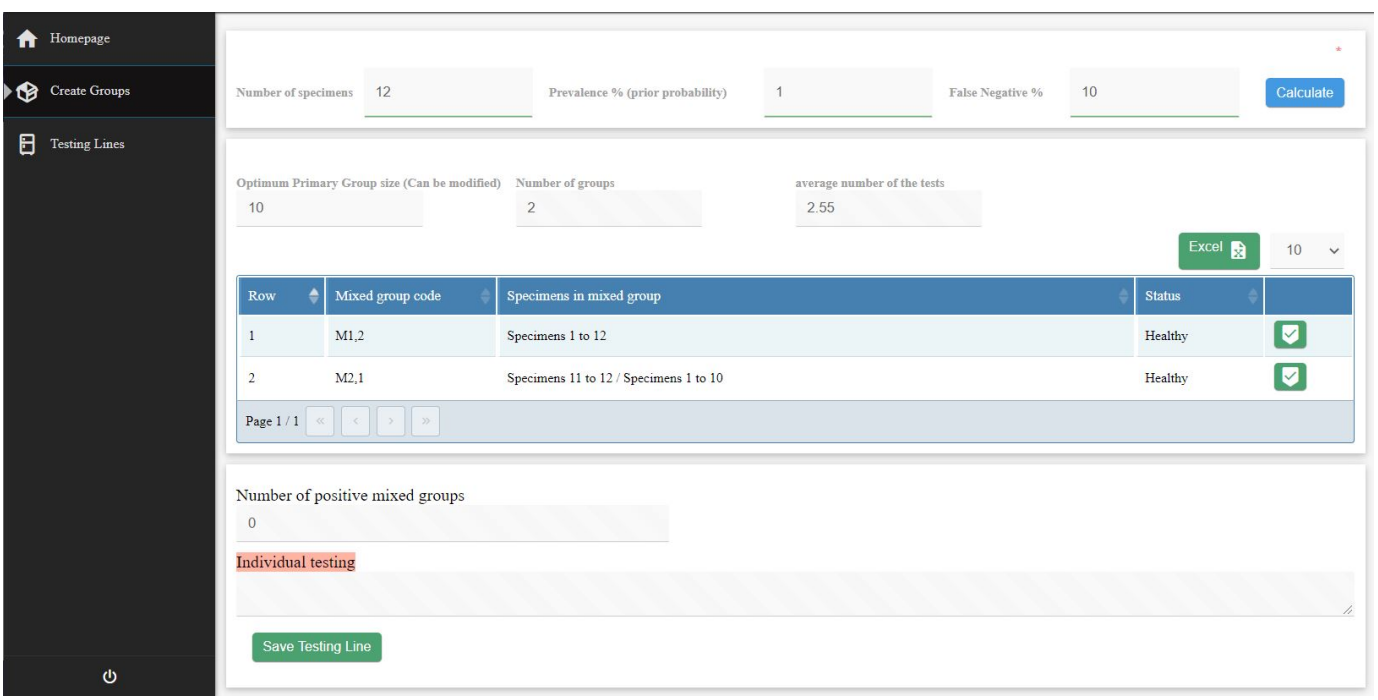

Fig. B.1. "Create Groups" section in the groupMix web interface.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20154070; this version posted August 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

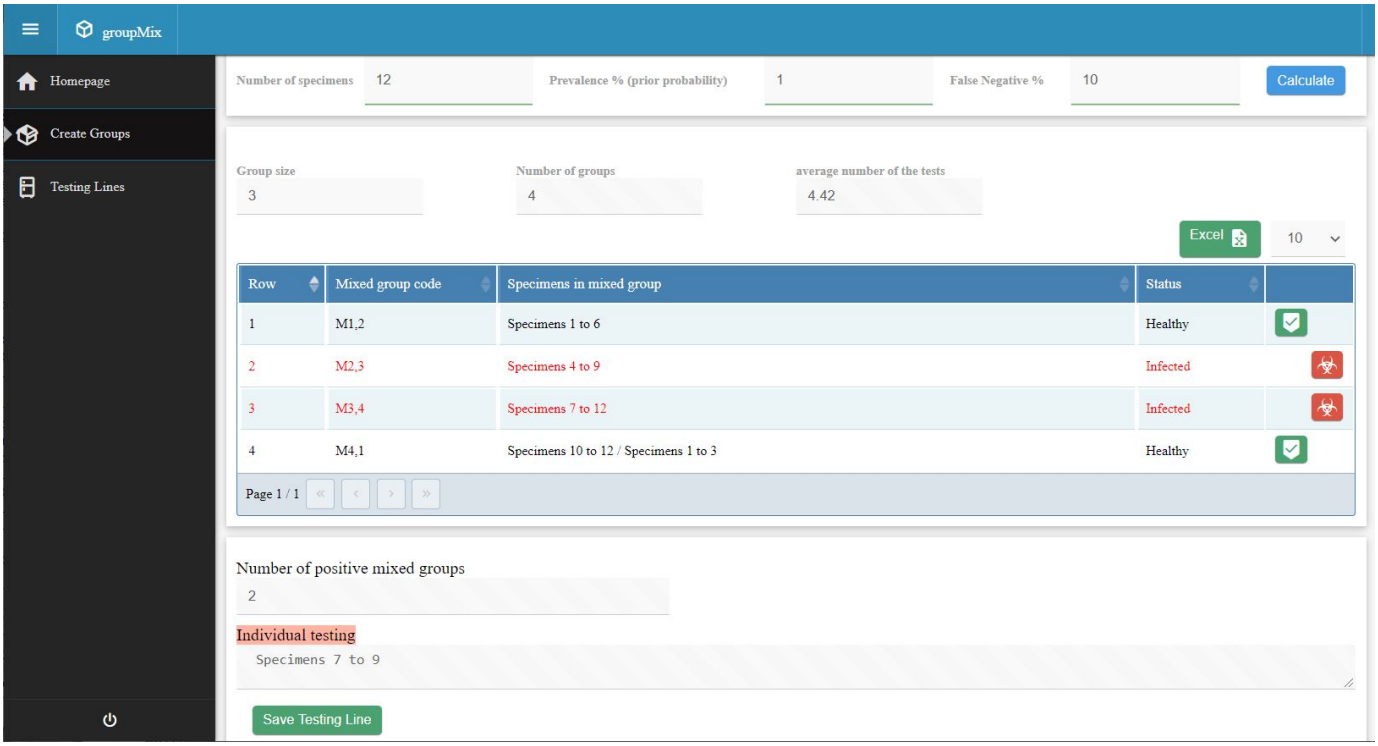

Fig. B.2. "Create Groups" section in the groupMix web interface for $N = 12$, $p = 1\%$, FNR= 10%, and $n = 3$.

Fig. B.3. "Test lines" section in the groupMix web interface.

B. Test Lines

In this section, we can see the list of saved test lines (Fig. B.3). Moreover, we can edit or delete each test line, and download the list of available test lines as an Excel spreadsheet.