

# Optimal Group Testing Strategy for the Mass Screening of SARS-CoV-2

Fengfeng Huang

*School of Management and Economics, University of Electronic Science and Technology of China, China, ffhuang@uestc.edu.cn*

Pengfei Guo

*College of Business, City University of Hong Kong, Kowloon, Hong Kong, penguo@cityu.edu.hk*

Yulan Wang\*

*Faculty of Business, The Hong Kong Polytechnic University, Kowloon, Hong Kong, yulan.wang@polyu.edu.hk*

---

**Abstract:** We analyze the group testing strategy that maximizes the efficiency of the SARS-CoV-2 screening test while ensuring its effectiveness, where the effectiveness of group testing guarantees that negative results from pooled samples can be considered presumptive negative. Two aspects of test efficiency are considered, one concerning the maximization of the welfare throughput and the other concerning the maximization of the identification rate (namely, identifying as many infected individuals as possible). We show that compared with individual testing, group testing leads to a higher probability of false negative results but a lower probability of false positive results. To ensure the test effectiveness, both the group size and the prevalence of SARS-CoV-2 must be below certain respective thresholds. To achieve test efficiency that concerns either the welfare throughput maximization or the identification rate maximization, the optimal group size is jointly determined by the test accuracy parameters, the infection prevalence rate, and the relative importance of identifying infected subjects. We also show that the optimal group size that maximizes the welfare throughput is weakly smaller than the one that maximizes the identification rate.

**Keywords:** Group testing, mass screening, test specificity, test sensitivity

---

---

\*Corresponding Author.

# 1 Introduction

Mass screening is an essential means to curb the spread of the coronavirus disease 2019 (COVID-19) due to the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. However, due to the limited test capacity, the extent of screening work usually falls short of our expectation. In order to reconcile the tight testing capacity and the urgent need for large-scale tests, “group testing” protocols have been put forward (see, e.g., Ackerman et al. (2020); Lohse et al. (2020); Shenta et al. (2020)) and into practice in several places, including Wuhan city of China (Walden and Zhao, 2020) and Germany (Healthcare in Europe, 2020). With group testing, samples from multiple subjects are pooled together and tested as one to check the presence of the virus. The pooled sample shall be tested negative if none of the sampled individuals is infected and positive otherwise. When a pooled sample is tested positive, each sample of the pool is then tested individually.

Group testing was developed by Dorfman (1943) as a cost efficient technique to eliminate syphilis men in army recruits and has been used afterwards in screening a variety of infections, including HIV, H1N1 influenza, malaria, Zika virus (see, e.g., Pilcher et al. (2002); Cowling et al. (2010); Hsiang et al. (2010); Kuehnert (2016), respectively). In these studies, the group testing strategy aims to minimize the expected number of tests. However, they usually overlook how group testing affects the overall test accuracy.

Virtually, no test is 100% accurate, because the distribution of the true positive result and that of the true negative result usually overlap and there does not exist a single decision criterion that could separate the samples cleanly (Metz, 1978; Wilson et al., 1968). When choosing the decision criteria for a test (which determine test specificity and sensitivity), several clinical factors shall be taken into consideration, including the disease prevalence and the relative importance of identifying infected subjects (Metz, 1978). Due to that pooling samples together alters the disease prevalence (among the pooled samples), the frequency of errors and the overall test accuracy of group testing are different from those of individual testing. In particular, for communicable diseases, the consequences associated with false positive results are more severe than those associated with false negative results, as false positives only lead to further testing whereas false negatives lead to further spread of the disease. Thus, it is needed to take into account the impacts of false positives and false negatives when developing a group testing strategy.

We develop a stylized model to investigate the optimal group testing strategy, i.e., the

number of individual samples to be mixed together as one, so that the efficiency of the test is maximized. Two aspects of test efficiency are considered, one concerning the maximization of the identification rate (namely, identifying as many infected individuals as possible) and the other concerning the maximization of the welfare throughput. The first aspect is in essence the same as minimizing the expected number of tests, which, as has been stated above, is the primary concern of the group testing strategy in the extant literature. The second aspect takes account of the consequences associated with false positive and false negative results, which, despite not having drawn enough attention in the group testing literature, have been widely recognized by the medical community; see, e.g., Wilson et al. (1968); Metz (1978).

When deciding the optimal group size, both aspects of test efficiency shall be subject to the group testing effectiveness requirement. For instance, the Food and Drug Administration of the United State suggests that group testing is effective only if each individual subject in a pooled sample can be considered presumptive negative whenever the pooled sample is tested negative (United State Food and Drug Administration, 2020).

We show that compared with individual testing, group testing leads to an increased probability of false negative results and a decreased probability of false positive results. The effectiveness of group testing requires both the infection prevalence rate and the group size of individual samples below certain thresholds. These thresholds are jointly determined by the test accuracy, which includes test specificity (the probability of correctly identifying an uninfected subject) and test sensitivity (the probability of correctly identifying an infected subject), and the relative importance of identifying infected subjects. We also show that the optimal group size hinges on the test accuracy, the infection prevalence rate, and the relative importance of identifying an infected subject. Moreover, the optimal group size under the welfare throughput maximization is weakly smaller than that under the identification rate maximization. When the agent aims to maximize the welfare throughput and test specificity is 100% (i.e., no false positive results), the optimal group size is the same as the one when the agent aims to maximize the identification rate, and it equals the minimum of two values with one corresponding to the group size that ensures the effectiveness of group testing and the other corresponding to the group size that minimizes the required testing times per sample.

We further conduct the extensive numerical studies by taking into account that (1) test outcomes are usually not dichotomous and (2) it is possible to alter the decision criteria to align with the purpose of the test in practice; see, e.g., Metz (1978); Wilson et al. (1968), where decreasing (increasing) test specificity is coupled with increasing (decreasing) test

sensitivity. Our numerical experiments reveal that a decision criterion associated with a slightly lower test specificity and a higher test sensitivity can improve the test efficiency of group-testing. Moreover, the relative advantage of group testing over individual testing in improving test efficiency diminishes when the infection prevalence rate is high.

The remainder of this paper is organized as follows. The related literature is reviewed in Section 2. Section 3 discusses the model setup. We analyze the optimal group sizes in Section 4. Concluding remarks are provided in Section 5.

## 2 Literature Review

Ever since group testing was developed by Dorfman in 1943 in eliminating syphilis man in army (Dorfman, 1943), the literature on group testing has being flourishing. Researchers have investigated the feasibility of group testing in a wide range of infections through experimental studies and clinic practice, including HIV, H1N1 influenza, malaria, and Zika virus; see Pilcher et al. (2002); Cowling et al. (2010); Hsiang et al. (2010); Kuehnert (2016). In response to the recent COVID-19 outbreak, group testing protocols for SARS-CoV-2 has surged (Ackerman et al., 2020; Lohse et al., 2020; Shenta et al., 2020). Group testing pools samples collected from multiple subjects into a single pool and tests the pooled sample for the presence of the virus; the pool shall be tested negative if none of the sampled individuals is infected and positive otherwise. Based on the group testing method, a protocol called the square array method was proposed by Phatarfod and Sudbury (1994) to further improve test efficiency. Under the square array method, samples are placed into a square array where each row and column forms a pool, and samples whose row and column pools both tested positive shall be deemed positive. The group testing strategy and its square-array derivative increase testing throughput, and have been analyzed by researchers from multiple disciplines by considering various perspectives, including search theories, learning theories, information theories, and control theories; see, e.g., Malyutov (2013); Mentus et al. (2020); Aldridge et al. (2019); Lin et al. (2020), respectively.

The aforementioned studies overlook the following two facts: One, the group testing strategy affects the frequencies of false positive and false negative results. Two, the consequences associated with false positive results are more severe than those associated with false negative results for communicable diseases (Wilson et al., 1968; Aprahamian et al., 2019). We note that Aprahamian et al. (2019) consider a certain type of consequence incurred by

the false positive and false negative results. Their group testing strategy aims to minimize a weighted sum of the expected number of false negative and false positive results, and the weights are assigned by the decision maker. Different from Aprahamian et al. (2019), the weights associated with false negative and false positive results in our welfare throughput maximization objective are jointly determined by the disease prevalence and the disease features.

Our work is also related to the studies on balancing the system performance (such as social/patient welfare and system congestion) and diagnosis accuracy, including Alizamir et al. (2013); Wang et al. (2010); Huang et al. (2022). Alizamir et al. (2013) consider a diagnostic process that consists of a sequence of imperfect tests. They find that the provider shall continue to test the patients until its belief that the patient is of a given type falls into a certain interval. Wang et al. (2010) consider a multi-server queueing system and demonstrate that with dual concerns over accuracy and congestion, increasing capacity might increase congestion. Huang et al. (2022) consider a similar diagnostic service but focus on examining how patients make their respective decisions based on their personal believes and how their decisions affect the social outcome.

Other related studies include Tu et al. (1995); Zenios and Wein (1998); Lin et al. (2020), which evaluate the effectiveness of group testing by using statistical or stylized models. Both Tu et al. (1995) and Zenios and Wein (1998) estimate the prevalence of HIV by adopting the group testing method. Tu et al. (1995) show that group testing not only reduces the probability of false positive results but also improves the accuracy of the estimator. Zenios and Wein (1998) take into account that HIV test outcomes are not dichotomous but continuous. They develop a parametric procedure and a hierarchical model to derive the decision criteria that yield accurate estimates. Lin et al. (2020) propose a group testing strategy to reduce the prevalence rate of COVID-19 in a closed community.

### 3 Model Setup

Consider a representative testing center (the agent) that provides SARS-CoV-2 (which causes COVID-19) mass screening tests to the public. The test capacity of the agent (namely, the number of tests the agent can perform at a time) is fixed, and his objective is to maximize the efficiency of group testing by choosing the appropriate group size. We consider two aspects of test efficiency, one concerning maximizing of the welfare throughput, and the other con-

cerning maximizing the identification rate, that is, identifying as many infected individuals as possible. Let  $p_0$  denote the infection prevalence rate; it measures the probability of an asymptomatic subject being infected. As of March 2022, the percentage of tested individuals being tested positive is no more than 20% in most countries (Ritchie et al., 2022), and such percentage shall be lower when mass screening is considered. Given such evidence, we consider that  $p_0 < 20\%$  in this study.

Denote a subject's true type by  $t$ , where  $t = 1$  indicates the subject (she) is infected and  $t = 0$  uninfected. The screening test categorizes subjects as positive ( $s = 1$ ) and negative ( $s = 0$ ). The screening test is not perfect. Depending on the subject's true type  $t$  ( $t \in \{0, 1\}$ ), the probability of  $s = t$  is as follows:

$$\alpha := P(s = 0|t = 0) \text{ and } \beta := P(s = 1|t = 1).$$

The parameters  $\alpha$  and  $\beta$  are known as test specificity (the probability of correctly identifying an uninfected subject) and test sensitivity (the probability of correctly identifying an infected subject), respectively. It is reported that the accuracy of SARS-CoV-2 screening testing usually has an average sensitivity of 84.3% and a specificity of between 96.6% and 99.7%; see, e.g., the systematic review of Bastos et al. (2020). Based on Bastos et al. (2020) and for the sake of analytical tractability, we restrict the parameter values to the case  $1.75 < \alpha + \beta < 2$ .

Given a subject's true type  $t = i$  ( $i \in \{0, 1\}$ ), she obtains a reward  $V_i$  if she is correctly identified ( $s = i$ ) and suffers a loss  $L_i$  if misidentified ( $s = 1 - i$ ).  $V_i$  is the potential benefit associated with either being assured of the health status when  $i = 0$  or receiving the proper treatment when  $i = 1$ , and  $L_i$  is the disutility/loss incurred from either not taking the proper treatment (or the potential pandemic containment failure) when  $i = 1$  or the unnecessary stress suffered by the concerned individual (or the health care system) when  $i = 0$ . Following Huang et al. (2022), we assume that  $\min\{\alpha V_0 - (1 - \alpha)L_0, \beta V_1 - (1 - \beta)L_1\} > 0$ , which indicates that even though the screening information is imperfect, it is valuable to all individuals. Define

$$\theta := \frac{V_1 + L_1}{V_0 + V_1 + L_0 + L_1},$$

which is an indicator of the relative importance of identifying an infected subject. As far as the communicable diseases are concerned, identifying positive cases are extremely important (Wilson et al., 1968), and hence, the value of  $\theta$  shall be large.

The group test is composed of two stages. In Stage 1, the agent pools samples into equal-sized groups and performs tests on the groups. If a group is tested negative, every sample in the group is considered to be uninfected; otherwise, it goes to Stage 2, in which every sample in the group is tested individually. Let  $x_K$  denote the prevalence of SARS-CoV-2 among the groups when the group size is  $K$  ( $K = 2, 3, \dots$ ). Then, we have

$$x_K = 1 - (1 - p_0)^K.$$

**Test Effectiveness.** When maximizing the efficiency of group testing, the agent is often subject to the effectiveness requirement. For instance, one important concern of the Food and Drug Administration of the United State with group testing is whether negative results from the pooled samples can be presumptive negative (United State Food and Drug Administration, 2020). Group testing can be regarded effective only when the expected reward of considering negative pooling test results as presumptive negative is greater than that as presumptive positive; that is,

$$(1 - g(x_K))V_0 - g(x_K)L_1 > g(x_K)V_1 - (1 - g(x_K))L_0, \quad (1)$$

where

$$g(x_K) := \frac{(1 - \beta)x_K}{(1 - \beta)x_K + \alpha(1 - x_K)},$$

and  $g(x_K)$  is the posterior probability that a group of size  $K$  contains at least one infected subject given that the test result is negative. When the condition stated in (1) is violated, the groups with negative results still need to proceed to Stage 2, and hence, group testing cannot reduce the required number of tests.

**Test Efficiency.** Given the group size  $K$ , let  $P_{ij}(K)$  denote the probability of identifying a type  $t = i$  ( $i \in \{0, 1\}$ ) subject as  $s = j$  ( $j \in \{0, 1\}$ ),  $E[N|K]$  the average number of tests required for the group, and  $\lambda(K) := E[N|K]/K$  the average number of tests conducted on each individual sample. Obviously, when samples are tested individually,  $P_{01}(1) = 1 - \alpha$ ,  $P_{10}(1) = 1 - \beta$ , and  $E[N|1] = \lambda(1) = 1$ .

As the agent's test capacity is fixed, maximizing the welfare throughput (the identification rate, respectively) shall be equivalent to maximizing the welfare throughput (the identification rate, respectively) per test. Denote the welfare throughput per test and the

identification rate per test as  $\mathcal{W}_t$  and  $\mathcal{W}_r$ , respectively. Then,

$$\mathcal{W}_t = \max_K \frac{(1 - p_0)[P_{00}(K)V_0 - P_{01}(K)L_0] + p_0[P_{11}(K)V_1 - P_{10}(K)L_1]}{\lambda(K)}, \quad (2)$$

$$\mathcal{W}_r = \max_K \frac{P_{11}(K)p_0K}{E[N|K]} = \max_K \frac{P_{11}(K)p_0}{\lambda(K)}, \quad (3)$$

both of which shall be subject to the effectiveness requirement indicated by (1). In the welfare throughput maximization objective function (2), the numerator is the expected reward of a representative subject, in which  $P_{00}(K)V_0 - P_{01}(K)L_0$  and  $P_{11}(K)V_1 - P_{10}(K)L_1$  are her expected rewards if she is uninfected ( $t = 0$ ) and infected ( $t = 1$ ), respectively. In the identification rate maximization objective function (3),  $P_{11}(K)p_0K$  is the expected number of infected individuals found in each group as  $p_0K$  is the expected number of infected individuals in a group.

## 4 Analysis

In this section, we first examine how group testing affects test outcomes and the average number of tests required per individual sample. Next, we derive the condition that ensures the test effectiveness. After that, we derive the optimal group sizes that ensure test efficiency in two aspects, welfare throughput maximization and identification rate maximization. Last, we conduct numerical experiments to strengthen our understanding of the implications of group testing.

First, we provide the following result.

**Lemma 1.** *Suppose that the group size is  $K(> 1)$ . Then,*

1. *the average number of tests required per individual sample is  $\lambda(K) = \frac{1}{K} + \beta - (\alpha + \beta - 1)(1 - p_0)^K$ ;*
2. *the false negative probability, i.e., the probability that an infected subject is tested negative, is  $P_{10}(K) = 1 - \beta^2$ ;*
3. *the false positive probability, i.e., the probability that an uninfected subject is tested positive, is  $P_{01}(K) = (1 - \alpha)[\beta - (\alpha + \beta - 1)(1 - p_0)^{K-1}]$ .*

Lemma 1 shows that given test specificity  $\alpha$  and test sensitivity  $\beta$ , group testing leads to a higher false negative probability (because  $1 - \beta^2 \geq 1 - \beta$ ) and a lower false positive



probability (because  $(1 - \alpha)[\beta - (\alpha + \beta - 1)(1 - p_0)^{K-1}] \leq 1 - \alpha$ ) compared to individual testing. Moreover, the false negative probability is independent of the group size whereas the false positive probability increases with the group size.

With regard to the effectiveness of the test, we obtain the following result.

**Lemma 2.** *To ensure the effectiveness of group testing, the infection prevalence rate shall satisfy*

$$0 < p_0 < 1 - \sqrt{\frac{(1 - \beta)\theta}{(1 - \beta)\theta + \alpha(1 - \theta)}} := \bar{p}_0.$$

*The group size  $K$  shall not exceed an upper bound  $\bar{K}$ , where*

$$\bar{K} = \left\lfloor \log_{1-p_0} \frac{(1 - \beta)\theta}{(1 - \beta)\theta + \alpha(1 - \theta)} \right\rfloor.$$

*Both  $\bar{p}_0$  and  $\bar{K}$  decrease with  $\theta$  (the relative importance of identifying an infected subject) and increase with test specificity  $\alpha$  and test sensitivity  $\beta$ .*

Lemma 2 shows that to ensure the effectiveness of group testing, the infection prevalence rate  $p_0$  shall not exceed an upper threshold  $\bar{p}_0$  that is jointly determined by the relative importance of identifying an infected subject  $\theta$  and the test accuracy parameters  $\alpha$  and  $\beta$ . It also suggests that government health agencies shall provide operational guidelines on the implementation of group testing such as the maximum pool size that depends on  $p_0$ ,  $\theta$ ,  $\alpha$ , and  $\beta$ . The results stated in Lemma 2 are also consistent with the observed practice. For example, regarding SARS-CoV-2 screening tests, the U.S. Food and Drug Administration authorized several public health laboratories to conduct group testing with 5 to 8 samples, depending on their respective test technology and accuracy (Schneider, 2020), and a group size of 10 to 25 is allowed in New York State (New York State, 2020).

We further investigate how the infection prevalence rate  $p_0$  affects the group size upper bound  $\bar{K}$  via a numerical experiment; see Figure 1. It shows that  $\bar{K}$  is decreasing with  $p_0$ . The relative importance of identifying infected subjects  $\theta$  also plays a critical role, where a higher  $\theta$  leads to a smaller  $\bar{K}$ . Figure 1 shows that given  $\alpha = 100\%$  and  $\beta = 82\%$ , if  $\theta = 99.0\%$ , group testing is effective only when the infection prevalence rate  $p_0$  is below 3.5%, while if  $\theta = 98.0\%$ , it is effective as long as  $p_0$  is no higher than 6.8%.

Based on Lemmas 1 and 2, we can obtain the following result regarding the optimal group size.

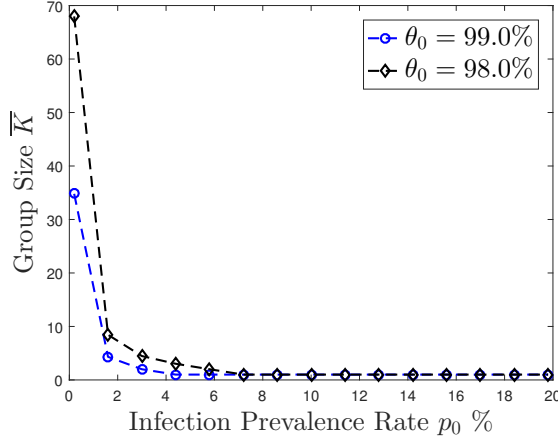


Figure 1: Group Size and Infection Prevalence Thresholds for Conducting Group Testing:  $\alpha = 100\%$  and  $\beta = 82\%$

**Proposition 1.** *To ensure the test effectiveness and efficiency, depending on whether the agent is concerned with the maximization of the welfare throughput or the identification rate, we have that:*

(i). *When the agent aims to maximize the welfare throughput, the optimal group size is*

$$K^* = \begin{cases} \min\{\tilde{K}, \bar{K}\} & \text{if } \alpha = 1, \\ \min\{\hat{K}, \bar{K}\} & \text{if } 1/2 < \alpha < 1; \end{cases}$$

(ii). *When the agent aims to maximize the identification rate, the optimal group size is*

$$K^* = \min\{\tilde{K}, \bar{K}\},$$

where  $\hat{K} < \tilde{K}$ ,  $\tilde{K}$  takes one of the following two values,  $\left\lfloor -\frac{\gamma}{\ln(1-p_0)} \right\rfloor$  or  $\left\lfloor -\frac{\gamma}{\ln(1-p_0)} \right\rfloor + 1$ , whichever yields a larger value of the objective function, and  $\gamma$  ( $0 < \gamma < 1$ ) satisfies

$$\ln(1-p_0) + (\alpha + \beta - 1)\gamma^2 e^{-\gamma} = 0.$$

Proposition 1 shows that under both test efficiency objectives—the welfare throughput maximization and the identification rate maximization—the magnitude of the optimal group size  $K^*$  highly depends on the prevalence of virus infection  $p_0$ , test accuracy parameters  $\alpha$  and  $\beta$ , and the relative importance of identifying an infected subject  $\theta$ . We can easily conclude from Proposition 1 that the optimal group size under the welfare throughput maximization is weakly smaller than that under the identification rate maximization. Interestingly, the optimal group size that maximizes the welfare throughput when  $\alpha = 100\%$  is the same as

that maximizes the identification rate, and it equals the minimum of  $\bar{K}$ , the largest allowable group size that ensures the effectiveness of group testing (see Lemma 2), and  $\tilde{K}$ , the group size that minimizes the required testing times per sample.

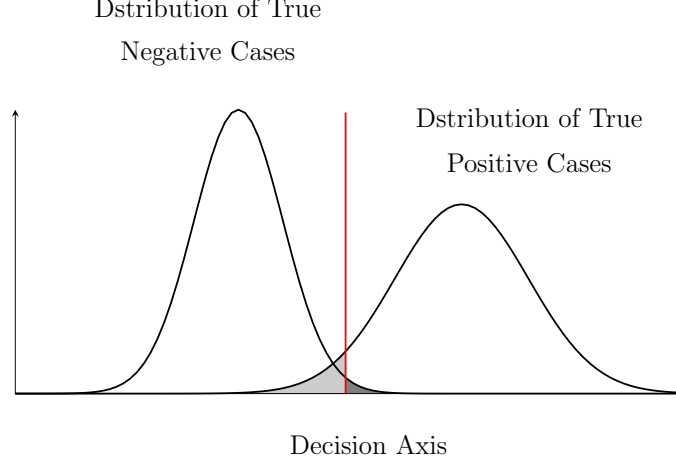


Figure 2: An Illustrative Example from Metz (1978)

**Numerical Study.** Before proceeding to the numerical experiments, we shall introduce the background knowledge about choosing the decision criteria that determine the magnitude of test specificity  $\alpha$  and test sensitivity  $\beta$  in the clinical environment. First, it is worth pointing out that the results of most existing tests are not dichotomy, because the distributions of the true positive test result and the true negative test result usually overlap and there does not exist any decision criteria that can separate the true positives and true negatives cleanly. Figure 2 excerpted from the theoretical guidelines of Metz (1978) illustrates this gist. According to Metz (1978), in Figure 2, any vertical line across the overlapped area of the two distributions can be a decision criterion that specifies a unique combination of test specificity and test sensitivity  $(\alpha, \beta)$ . Take the red line in Figure 2 As an example. The dark gray area formed by the red line and the distribution curve of true negative cases represents the false positive fraction and determines a unique value of  $1 - \alpha$ , while the light gray area formed by the red line and the distribution curve of true positive cases represents the false negative fraction and determines a unique value of  $1 - \beta$ . If we move the red line towards the left, test specificity  $\alpha$  decreases while test sensitivity  $\beta$  increases; *vice versa*. By varying the decision criteria via moving the red line, multiple combinations of test specificity and test sensitivity are obtained, based on which the medical community usually graphically plot a receiver operating characteristic (ROC) curve (Metz, 1978). Note that the ROC curve plots

the  $(1 - \alpha, \beta)$  pair rather than the  $(\alpha, \beta)$  pair.

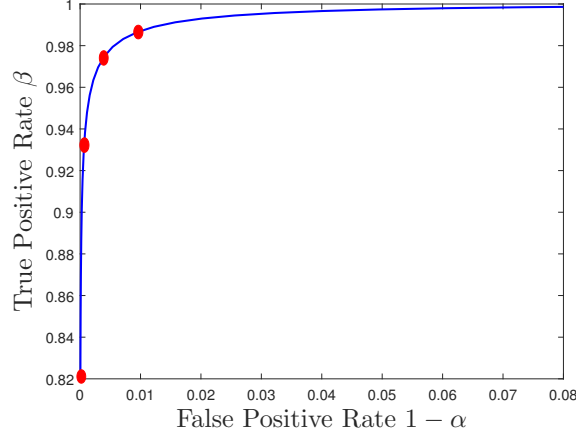


Figure 3: The Receiver Operating Characteristic Curve: True Negative  $\sim N(5, 1)$  and True Positive  $\sim N(10, 1.2)$

By adopting the theoretical guidelines of Metz (1978) along with the observed accuracy of SARS-CoV-2 tests in practice, we set the distributions of true negative and true positive results to be normally distributed with mean 5 and 10 and variance 1 and 1.2, respectively. We then explore a series of possible decision criteria and obtain the corresponding ROC curve as shown in Figure 3. We then select the following four test specificity and sensitivity combinations from the ROC curve— (100%, 82%), (99.9%, 93.8%), (99.5%, 97.7%), and (99.1%, 98.5%)— to conduct our numerical experiments, which are depicted as the red dots in Figure 3.

We now investigate how test specificity  $\alpha$  and test sensitivity  $\beta$  affect the optimal group size  $K^*$  and test efficiency. Our numerical experiments reveal that due to the high importance of identifying infected subjects for COVID-19 (i.e.,  $\theta$  is close to 1), the test effectiveness requirement characterized by the threshold group size  $\bar{K}$  (given in Lemma 2) is usually the binding factor, leading to that the two aspects of test efficiency—maximizing the welfare throughput and maximizing the identification rate—yield the same optimal group size. Hence, for illustration purpose, we only present the numerical outcomes under the welfare throughput maximization; see Figure 4. Figure 4(a) shows that the optimal group size  $K^*$  decreases with the infection prevalence rate  $p_0$ , and that  $K^*$  is weakly larger when test specificity  $\alpha$  decreases and test sensitivity  $\beta$  increases. Figure 4(b) depicts the relative efficiency improvement of group testing over individual testing. It shows that when  $0 < p_0 < 1\%$ , the test specificity and sensitivity pair  $(\alpha = 100\%, \beta = 82.0\%)$  leads to the highest test efficiency;

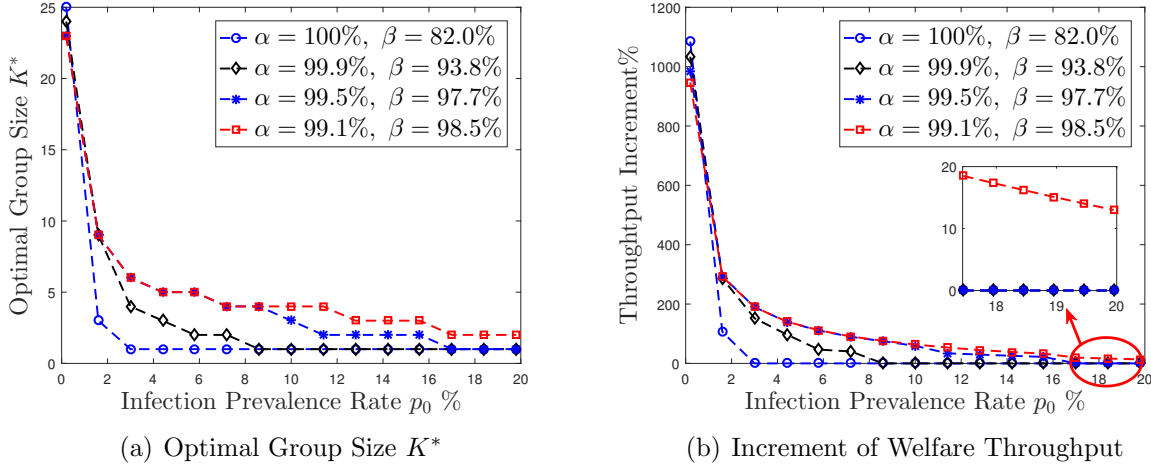


Figure 4: Optimal Group Testing Strategy That Maximizes the Welfare Throughput:  $\theta = 99\%$ ,  $V_0 = L_0$ ,  $V_1 = L_1$

when  $1\% \leq p_0 < 10\%$ , the test specificity and sensitivity pair  $(\alpha = 99.5\%, \beta = 97.7\%)$  leads to the highest test efficiency; and when  $10\% \leq p_0 < 20\%$ , the test specificity and sensitivity pair  $(\alpha = 99.1\%, \beta = 98.5\%)$  leads to the highest test efficiency. This indicates that as the infection prevalence rate  $p_0$  increases, to facilitate group testing, a decision criterion associated with a slightly lower test specificity  $\alpha$  and a higher test sensitivity  $\beta$  can improve test efficiency. Moreover, Figure 4(b) shows that for any combination of  $\alpha$  and  $\beta$ , the relative advantage of group testing over individual testing diminishes as  $p_0$  increases.

## 5 Conclusion

We consider an agent that provides SARS-CoV-2 screening tests to the public. The agent's objective is to maximize the efficiency of group testing. Two aspects of test efficiency are considered with one concerning the welfare throughput maximization and the other concerning the identification rate maximization. Apart from test efficiency, a group testing strategy must be "effective" in the sense that whenever a negative test result is obtained for a sample group, the group can be considered as presumptive negative. We show that compared with individual testing, group testing leads to a higher probability of false negative results but a lower probability of false positive results. To ensure the effectiveness of group testing, it requires that both the group size and the prevalence of SARS-CoV-2 are below certain thresholds. Moreover, the optimal group size that maximizes the welfare throughput is weakly smaller than the one that maximizes the identification rate, and the optimal group

sizes under both objectives—the welfare throughput maximization and the identification rate maximization—are jointly determined by test accuracy parameters, the infection prevalence rate, and the importance of identifying an infected subject. Our numerical study shows that the efficiency of group testing can be improved if we increase test sensitivity (the probability of correctly identifying infected subjects) by sacrificing test specificity (the probability of correctly identifying uninfected subjects). Moreover, the relative advantage of group testing over individual testing in improving test efficiency diminishes—or even vanishes—when the infection prevalence rate is high for all combinations of test specificity and test sensitivity.

## Acknowledgments

We are grateful to the department editor (Professor Joe Zhu), an anonymous associate editor, and two anonymous referees for their very helpful comments and suggestions. Fengfeng Huang’s work was supported by the National Natural Science Foundation of China (Grant number: 72101048). Pengfei Guo was supported in part by the Research Grants Council of Hong Kong (Grant number: 15502820). Yulan Wang is also affiliated with the Hong Kong Polytechnic University Shenzhen Research Institute and acknowledges the financial supports from the National Natural Science Foundation of China (Grant No. 71971184).

## References

- Ackerman, C.M., Myhrvold, C., Thakku, S.G., Freije, C.A., Metsky, H.C., Yang, D.K., Ye, S.H., Boehm, C.K., Kosoko-Thoroddsen, T.S.F., Kehe, J. and Nguyen, T.G., 2020. Massively multiplexed nucleic acid detection with Cas13. *Nature*, 582(7811), 277-282.
- Aldridge M., Johnson O., Scarlett J., 2019. Group testing: An information theory perspective. *Foundations and Trends in Communications and Information Theory*, 15(3-4), 196–392
- Alizamir, S., De Vericourt, F., Sun, P., 2013. Diagnosis accuracy under congestion. *Management Science*, 59(1), 157-171.
- Aprahamian, H, D.R. Bish, E.K. Bish., 2019. Optimal risk-based group testing. *Management Science*, 65(9), 4365-4384.
- Bastos, M.L., Tavaziva, G., Abidi, S.K., Campbell, J.R., Haraoui, L.P., Johnston, J.C., Lan, Z., Law, S., MacLean, E., Trajman, A. and Menzies, D., 2020. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. *bmj*, 370.
- Cowling, B.J., Lau, L.L., Wu, P., Wong, H.W., Fang, V.J., Riley, S. and Nishiura, H., 2010. Entry screening to delay local transmission of 2009 pandemic influenza A (H1N1). *BMC infectious diseases*, 10(1), 1-4.
- Dorfman, R., 1943. The detection of defective members of large populations. *Annals of Mathematical Statistics*, 14(4), 436-440.
- Healthcare in Europe. 2020, April 1. Corona ‘pool testing’ increases worldwide capacities many times over. <https://healthcare-in-europe.com/en/news/corona-pool-testing-increases-worldwide-capacities-many-times-over.html>.
- Huang, F., Guo, P. and Wang, Y., 2022. Modeling Patients’ Illness Perception and Equilibrium Analysis of Their Doctor Shopping Behavior. *Production and Operations Management*, 31(3), 1216-1234.
- Hsiang, M.S., Lin, M., Dokomajilar, C., Kemere, J., Pilcher, C.D., Dorsey, G. and Greenhouse, B., 2010. PCR-based pooling of dried blood spots for detection of malaria parasites:

- optimization and application to a cohort of Ugandan children. *Journal of clinical microbiology*, 48(10), 3539-3543.
- Kuehnert, M.J., 2016. Screening of blood donations for Zika virus infection—Puerto Rico, April 3–June 11, 2016. *MMWR. Morbidity and mortality weekly report*, 65.
- Lin, Y., Ren, Y., Wan, J. and Zhou, E., 2020. Group testing during the COVID-19 pandemic: optimal group size selection and prevalence control. *arXiv preprint arXiv:2008.06642*.
- Lohse, S., Pfuhl, T., Berko-Gottel, B., Rissland, J., Geibler, T., Gartner, B., Becker, S.L., Schneitler, S. and Smola, S., 2020. Pooling of samples for testing for SARS-CoV-2 in asymptomatic people. *The Lancet Infectious Diseases*.
- Malyutov M, 2013. Search for sparse active inputs: A review. In: Aydinian H, Cicalese F, Deppe C (eds) *Information Theory, Combinatorics, and Search Theory: In Memory of Rudolf Ahlswede*, Springer Berlin Heidelberg, Berlin, Heidelberg, 609–647.
- Mentus, C., Romeo, M. and DiPaola, C., 2020. Analysis and applications of adaptive group testing methods for COVID-19. *MedRxiv*.
- Metz, C.E., 1978. Basic principles of ROC analysis. In *Seminars in Nuclear Medicine*, 8(4), 283-298. WB Saunders.
- New York State, 2020, August 14. Governor Cuomo Announces New Testing Initiatives to Improve COVID-19 Detection & Control Across New York State. <https://www.governor.ny.gov/news/governor-cuomo-announces-new-testing-initiatives-improve-covid-19-detection-control-across-new>
- Ritchie, H., Mathieu, E., Rodés-Guirao L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D., Roser, M., 2022. Coronavirus Pandemic (COVID-19). Published online at OurWorldInData.org. Retrieved from: <https://ourworldindata.org/coronavirus-testing> on March 7, 2022.
- Phatarfod, R.M. and Sudbury, A., 1994. The use of a square array scheme in blood testing. *Statistics in medicine*, 13(22), 2337-2343.
- Pilcher, C.D., McPherson, J.T., Leone, P.A., Smurzynski, M., Owen-O’Dowd, J., Peace-Brewer, A.L., Harris, J., Hicks, C.B., Eron Jr, J.J. and Fiscus, S.A., 2002. Real-time,



- universal screening for acute HIV infection in a routine HIV counseling and testing population. *Jama*, 288(2), 216-221.
- Schneider, M., 2020, August 28. COVID-19 single-stage pooled testing method approved in Israel. <https://www.bioworld.com/articles/497142-covid-19-single-stage-pooled-testing-method-approved-in-israel>.
- Tu XM, Litvak E, Pagano M. 1995. On the Informativeness and accuracy of pooled testing in estimating prevalence of a rare disease: application to HIV screening. *Biometrika*, 82, 287-297.
- United State Food and Drug Administration, 2020, July 28. Does FDA have validation or other recommendations regarding SARS-CoV-2 diagnostic/screening tests for use with sample pooling? <https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/faqs-testing-sars-cov-2>
- Walden M., Zhao I., 2020, May 26. How China managed to test almost 1.5m people for coronavirus in a single day. <https://www.abc.net.au/news/2020-05-26/china-coronavirus-wuhan-testing-millions-in-10-days/12283202>.
- Wang, X., Debo, L. G., Scheller-Wolf, A., Smith, S. F., 2010. Design and analysis of diagnostic service centers. *Management Science*, 56(11), 1873-1890.
- Wilson, J.M.G., Jungner, G. and World Health Organization, 1968. Principles and practice of screening for disease.
- Shental, N., Levy, S., Wuvshet, V., Skorniakov, S., Shalem, B., Ottolenghi, A., Greenspan, Y., Steinberg, R., Edri, A., Gillis, R. and Goldhirsh, M., 2020. Efficient high-throughput SARS-CoV-2 testing to detect asymptomatic carriers. *Science Advances*, 6(37), p.eabc5961.
- Zenios SA, Wein LM, 1998. Pooled testing for hiv prevalence estimation: exploiting the dilution effect. *Statistics in Medicine*, 17(13), 1447-1467.

## Appendix: Proofs

**Proof of Lemma 1.** For a group test with size  $K > 1$ , if the test result is negative, the testing stops at Stage 1 with a total number of tests  $N = 1$ . Otherwise, the testing continues and proceeds to Stage 2, where additional  $K$  tests are needed. In this situation, the total number of tests  $N = 1 + K$ . The probability that the group is tested negative and that the group is tested positive can be respectively derived as

$$P^- = \alpha(1 - x_K) + (1 - \beta)x_K; \quad P^+ = (1 - \alpha)(1 - x_K) + \beta x_K.$$

Then, the average number of tests required for the group can be written as

$$E[N|K] = 1 + [(1 - \alpha)(1 - x_K) + \beta x_K]K.$$

Thus, the average number of tests needed for each individual sample in the group is

$$\lambda(K) = E[N|K]/K = 1/K + \beta - (\alpha + \beta - 1)(1 - p_0)^K.$$

At Stage 1, a pooled sample is tested false negative with probability  $(1 - \beta)x_K$  and tested true positive with probability  $\beta x_K$  when the group has at least one infected subject. If the group test result is true positive in Stage 1, each individual sample is re-tested in Stage 2. Recall that each individual sample is tested false negative with probability  $(1 - \beta)$ . Combining the above together, we can show that under the group testing, the expected number of false negative subjects can be derived as

$$FN = (1 - \beta)x_K \times \frac{Kp_0}{x_K} + \beta x_K \times (1 - \beta) \times \frac{Kp_0}{x_K} = (1 - \beta^2)p_0K,$$

where  $\frac{Kp_0}{x_K}$  is the expected number of infected subjects in a group that has at least one infected subject.

Similarly, at Stage 1, a pooled sample is tested false positive with probability  $(1 - \alpha)(1 - x_K)$  when the group has no infected subject. Then, at Stage 2, these  $K$  uninfected individual samples are re-tested and the result is false positive with probability  $(1 - \alpha)$ . When at Stage 1 a pooled sample is tested true positive (with probability  $\beta x_K$ ), the expected number of uninfected samples in the group (that has at least one infected subject) is  $K - K \frac{p_0}{x_K}$ . When the test proceeds to Stage 2, those unaffected samples are re-tested and the result is false positive with probability  $(1 - \alpha)$ . Combining the above together, we can show that under

the group testing, the expected number of false positive subjects can be derived as

$$\begin{aligned} FP &= (1 - \alpha)(1 - x_K) \times (1 - \alpha) \times K + \beta x_K \times (1 - \alpha) \times \left( K - K \frac{p_0}{x_K} \right) \\ &= (1 - p_0)(1 - \alpha)[\beta - (\alpha + \beta - 1)(1 - p_0)^{K-1}]K. \end{aligned}$$

The total numbers of true positive and true negative subjects in the group are  $p_0 K$  and  $(1 - p_0)K$ , respectively. Thus, the probability of false negative and false positive test results for the group are, respectively,

$$P_{10}(K) = 1 - \beta^2, \quad P_{01}(K) = (1 - \alpha)[\beta - (\alpha + \beta - 1)(1 - p_0)^{K-1}].$$

**Proof of Lemma 2.** Recall that the group testing strategy is effective only if the condition  $(1 - g(x_K))V_0 - g(x_K)L_1 > g(x_K)V_1 - (1 - g(x_K))L_0$  is satisfied, and that  $\theta = (V_1 + L_1)/(V_0 + V_1 + L_0 + L_1)$  and  $g(x_K) = \frac{(1-\beta)x_K}{(1-\beta)x_K + \alpha(1-x_K)}$ . We can show that this effectiveness condition is equivalent to

$$x_K < \frac{\alpha(1 - \theta)}{\alpha(1 - \theta) + (1 - \beta)\theta}.$$

Recall that  $x_K = 1 - (1 - p_0)^K$ . We can show that this requires

$$K < \log_{1-p_0} \frac{(1 - \beta)\theta}{(1 - \beta)\theta + \alpha(1 - \theta)}.$$

Since the group size is an integer, we can obtain that

$$\bar{K} = \left\lfloor \log_{1-p_0} \frac{(1 - \beta)\theta}{(1 - \beta)\theta + \alpha(1 - \theta)} \right\rfloor.$$

Considering that group testing requires  $K \geq 2$  (otherwise, it shall be tested individually), we further obtain that this requires

$$0 < p_0 < 1 - \sqrt{\frac{(1 - \beta)\theta}{(1 - \beta)\theta + \alpha(1 - \theta)}} := \bar{p}_0.$$

It can be easily shown that  $\frac{(1-\beta)\theta}{(1-\beta)\theta + \alpha(1-\theta)}$  increases with  $\theta$  and decreases with  $\alpha$  and  $\beta$ . It can be also easily shown that both  $\lfloor \log_{1-p_0} x \rfloor$  and  $1 - \sqrt{x}$  decrease with  $x$ . By using the chain rule, we can show that  $\bar{K}$  and  $\bar{p}_0$  decrease with  $\theta$  and increase with  $\alpha$  and  $\beta$ .

**Proof of Proposition 1.** It is easy to show that when  $0 < x < 1$ ,

$$x < -\ln(1 - x) < x/(1 - x). \quad (4)$$

By utilizing  $3/4 < \alpha + \beta - 1 < 1$  and  $p_0 < 20\%$ , we can show that

$$-\ln(1 - p_0) - (\alpha + \beta - 1)e^{-1} < \frac{p_0}{1 - p_0} - \frac{3}{4}e^{-1} < 0, \quad (5)$$

Define the following function over the range  $x \in [2, \infty)$ :

$$h(x) := 1/x - (\alpha + \beta - 1)(1 - p_0)^x + \beta. \quad (6)$$

Taking the first order derivative, we then have

$$h'(x) = -1/x^2 - (\alpha + \beta - 1)\ln(1 - p_0)(1 - p_0)^x.$$

It can be easily shown that  $-1/x^2$  and  $\ln(1 - p_0)(1 - p_0)^x$  are both increasing and concave in  $x$ . This implies that  $-1/x^2$  and  $\ln(1 - p_0)(1 - p_0)^x$  cross each other at most twice; in other words,  $h'(x)$  crosses 0 *at most* twice. We can show that

$$\begin{aligned} h'(2) &= -1/4 - (\alpha + \beta - 1)\ln(1 - p_0)(1 - p_0)^2 < 0, \\ h'(-1/\ln(1 - p_0)) &= -\ln(1 - p_0)[\ln(1 - p_0) + (\alpha + \beta - 1)e^{-1}] > 0, \end{aligned}$$

where the inequality  $<$  can be obtained from (4) and  $3/4 < \alpha + \beta - 1 < 1$ , and the inequality “ $>$ ” can be obtained based on (5). This implies that when  $2 \leq x \leq -1/\ln(1 - p_0)$ ,  $h'(x)$  crosses 0 once at a local minimum. Thus, for  $x > -1/\ln(1 - p_0)$ ,  $h'(x)$  crosses 0 *at most* once. Since  $h'(-1/\ln(1 - p_0)) > 0$ , the minimum on the interval shall be located at one of the two interval endpoints. We can also show that

$$h(-1/\ln(1 - p_0)) - \lim_{x \rightarrow +\infty} h(x) = -[\ln(1 - p_0) + (\alpha + \beta - 1)e^{-1}] < 0,$$

where the inequality follows from (5). Therefore, the global minimum of  $h(x)$  shall fall into the interval  $x \in [2, -1/\ln(1 - p_0)]$  and is determined by  $x = -\gamma/\ln(1 - p_0)$  with  $\gamma$  being the unique solution that satisfies

$$h'(-\gamma/\ln(1 - p_0)) = -\ln(1 - p_0)[\ln(1 - p_0)/\gamma^2 + (\alpha + \beta - 1)e^{-\gamma}] = 0. \quad (7)$$

**Welfare throughput maximization.** First, consider the case  $\alpha = 1$ . We can simplify the objective function (2) as follows:

$$\mathcal{W}_t = \max_K \frac{(1 - p_0)V_0 + p_0[\beta^2 V_1 - (1 - \beta^2)L_1]}{\lambda(K)}.$$

As the numerator is now a constant, the welfare throughput is maximized once the effective number of tests per individual sample  $\lambda(K)$  is minimized. We note that the function  $\lambda(K)$  is

equivalent to the function  $h(x)$  except that  $\lambda(K)$  is defined on positive integers. The group size that minimizes  $\lambda(K)$ ,  $\tilde{K}$ , shall be one of the two integers around  $x = -\gamma/\ln(1-p_0)$ , namely,

$$\tilde{K} \in \left\{ \left\lfloor -\frac{\gamma}{\ln(1-p_0)} \right\rfloor, \left\lfloor -\frac{\gamma}{\ln(1-p_0)} \right\rfloor + 1 \right\}, \quad (8)$$

whichever yields a larger value of the welfare throughput. Also, recall that the group size shall not exceed  $\bar{K}$ , and thus  $K^* = \min\{\bar{K}, \tilde{K}\}$ .

Next, consider  $1/2 < \alpha < 1$ . Define the following functions on the interval  $x \geq 2$ :

$$\begin{aligned} f(x) &= (1-p_0)[V_0 - (1-\alpha)[\beta - (\alpha + \beta - 1)(1-p_0)^{x-1}](V_0 + L_0)] \\ &\quad + p_0[V_1 - (1-\beta^2)(V_1 + L_1)], \\ F(x) &= f(x)/h(x). \end{aligned}$$

We can show that

$$\begin{aligned} f'(x) &= (1-\alpha)(\alpha + \beta - 1)(1-p_0)^x \ln(1-p_0)(V_0 + L_0) < 0, \\ F'(x) &= \frac{f'(x)h(x) - h'(x)f(x)}{h^2(x)}. \end{aligned}$$

Recall that  $h'(x) = 0$  when  $x = -\gamma/\ln(1-p_0)$ , in which  $\gamma$  is uniquely determined by (7). Then, we have

$$F'(-\gamma/\ln(1-p_0)) = f'(-\gamma/\ln(1-p_0)) / h(-\gamma/\ln(1-p_0)) < 0. \quad (9)$$

We have shown that  $h'(x)$  crosses 0 *at most* once when  $x > -\gamma/\ln(1-p_0)$ . We further have the following two scenarios:

1. If  $h'(x)$  does not cross 0, then  $h'(x) > 0$  for all  $x > -\gamma/\ln(1-p_0)$ , from which we can show that  $F'(x) < 0$ . Hence, the maximum value of  $F(x)$  on this interval shall be obtained at  $x = -\gamma/\ln(1-p_0)$ ;
2. If  $h'(x)$  indeed crosses 0 once, by using  $f'(x) < 0$ , we can show that  $F'(x)$  crosses 0 *at most* once. As  $\lim_{x \rightarrow \infty} F'(x) = 0$ , together with (9), we can show that  $F(x)$  decreases on this interval. Hence, the maximum value of  $F(x)$  shall be obtained at  $x = -\gamma/\ln(1-p_0)$ .

Thus, the maximum value of  $F(x)$  is obtained at  $x = -\gamma/\ln(1-p_0)$  for  $x > -\gamma/\ln(1-p_0)$ . This indicates that the global maximum of  $F(x)$  shall be found in the interval  $2 \leq x \leq$

$-\gamma/\ln(1-p_0)$ . Let  $x_0$  denote the global maximum of  $F(x)$ . By using (9), we can show that  $2 \leq x_0 < -\gamma/\ln(1-p_0)$ . Note that the objective function (2) is equivalent to the function  $F(x)$  except that (2) is defined on positive integers. Thus, the optimal group size, denoted as  $K_0$ , shall be one of the two integers around  $x_0$ , i.e,

$$\hat{K} \in \{\lfloor x_0 \rfloor, \lfloor x_0 \rfloor + 1\},$$

whichever yields a larger value of welfare throughput. As the group size shall not exceed  $\bar{K}$ ,  $K^* = \min\{\hat{K}, \bar{K}\} \leq \min\{\tilde{K}, \bar{K}\}$ , where the inequality is obtained because  $2 \leq x_0 < -\gamma/\ln(1-p_0)$ .

**Identification rate maximization.** From Lemma 1, we have  $P_{11}(K) = 1 - P_{10}(K) = \beta^2$ . By using (3), we can rewrite the agent's optimization problem as follows:

$$\begin{aligned} \mathcal{W}_r &= \max_K \frac{p_0 \beta^2}{\lambda(K)}, \\ s.t. \quad K &\leq \bar{K}. \end{aligned}$$

Then,  $\mathcal{W}_r$  is maximized whenever  $\lambda(K)$  is maximized. Obviously,  $\lambda(K) = h(K)$  for any integer  $K \geq 2$ . We have shown that  $h(x)$  is minimized at  $x = -\gamma/\ln(1-p_0)$ , where  $\gamma$  is determined by (7). By considering the effectiveness constraint  $K \leq \bar{K}$ , it is easy to show that  $\mathcal{W}_r$  is maximized at  $K^* = \min\{\bar{K}, \tilde{K}\}$ , where  $\tilde{K}$  is given by (8).