IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 16, Issue 3 Ser. III (May – June 2020), PP 58-61
www.iosrjournals.org

How to Select the Group Size for Group Testing to Reduce Cost and Increase Capacity for COVID-19 Test

Xueliang Pan, PhD
Department of Biomedical Informatics, the Ohio State University, Columbus, OH

Abstract: *The group test strategy, also known as pooling samples test /screening test / group sample test, has been used since WWII to test cases of HIV, chlamydia, malaria and influenza. It not only can increase test capacity, but also decrease test cost. Given the current shortage of COVID-19 tests, the group test strategy can be an appropriate and necessary tool for increasing testing and enabling safe returns to normal activity during the COVID-19 pandemic. Outside of China, the majority of COVID-19 tests are currently individual tests, which are less cost efficient than group testing. The selection of ideal group size is based on many factors including prevalence levels and operational challenges. The objective of this brief report is to recommend guide group size selection at various prevalence levels for a simple group test strategy, and discuss other pertinent issues governing group size selection.*

Date of Submission: 20-05-2020 Date of Acceptance: 05-06-2020

The shortage of the COVID-19 test kits and their high cost are emerging challenges to the timely detection of COVID-19 and the control of its spread. With limited number of test kits available and high cost burden in certain countries, there are calls for "group screening" test, or group test (Bilder, et al 2020). The group test strategy has been used since WWII to test cases of HIV, chlamydia, malaria and influenza. It can not only increase the test capacity, but also decrease the test cost per sample. The group test strategy is especially beneficial when implemented in large-scale testing, which is appropriate for the current COVID-19 pandemic. However, choosing an ideal group size for pooling samples is not easy, as it depends on several factors such as disease prevalence, operational challenges, and technical issues.

There are many different group screen testing algorithms. For example, one can separate 100 samples into 10 groups with group size of 10, or 14 groups with group size of 50 (one sample can be pooled into multiple groups). Some methods are mathematically beneficial but may be hard to implement due to increased operational mistakes, cost of time, or limited biological sample availability. The **group sample test strategy,** also known as Dorfman testing procedure (Dorfman 1943)is the first reported group test strategy in the literature. This test strategy is implemented as follows. Each sample is separated into 2-3 portions: one for group testing (portion A), one for individual testing if needed (portion B), and the rest as backup (optional). The first step is to combine a group of individual samples (portion A) as one pooled sample, and test the pooled sample. If the test result of the pooled sample is negative, then all samples in that group are declared negative. If the pooled sample is positive, then the second step is to test each individual sample (use portion B) in that group. This strategy can significantly reduce the number of tests needed, especially when the prevalence of the positive samples is low. The validity of this test strategy relies on the assumption that the test of the pooled sample will be positive only if one or more individual samples are positive, and that the accuracy of the test for the pooled sample is same for the individual sample.

The test proportion, defined as the ratio of the number of tests needed vs the total number of samples, is directly related to the total cost and the testing capacity. Assuming the test accuracy is perfect, the test proportion depends on the group size (n) and the prevalence of the positive samples in the population(P). To test 1,000,000 samples individually (group size n=1), we need 1,000,000 tests, i.e. the test proportion is100%. When the pooled sample has negative diagnosis, we only need 1 test to get the results of n samples, i.e. test proportion of 1/n. On the other hand, when the pooled sample is positive, we need n+1 tests to obtain the test results on n samples, i.e. test proportion is ((n+1)/n). Assuming each sample is independent to others, the probability that a pooled sample is negative (Pn) is the same as the probability that none of the samples in the group is positive. The number of positive tests (X) in n randomly selected samples follows the binominal distribution: X~Binom(n, P), where P is the prevalence of positive in the population. The probability that a pooled sample negative is Pn=Prob(X=0). The expected proportion of test needed can be calculated as

Test proportion= (Pn*1/n)+(1-Pn)*(1+1/n)=1+ 1/n - Pn.

The expected proportion of the tests needed is illustrated in figure 1 with various group size at different prevalence levels.

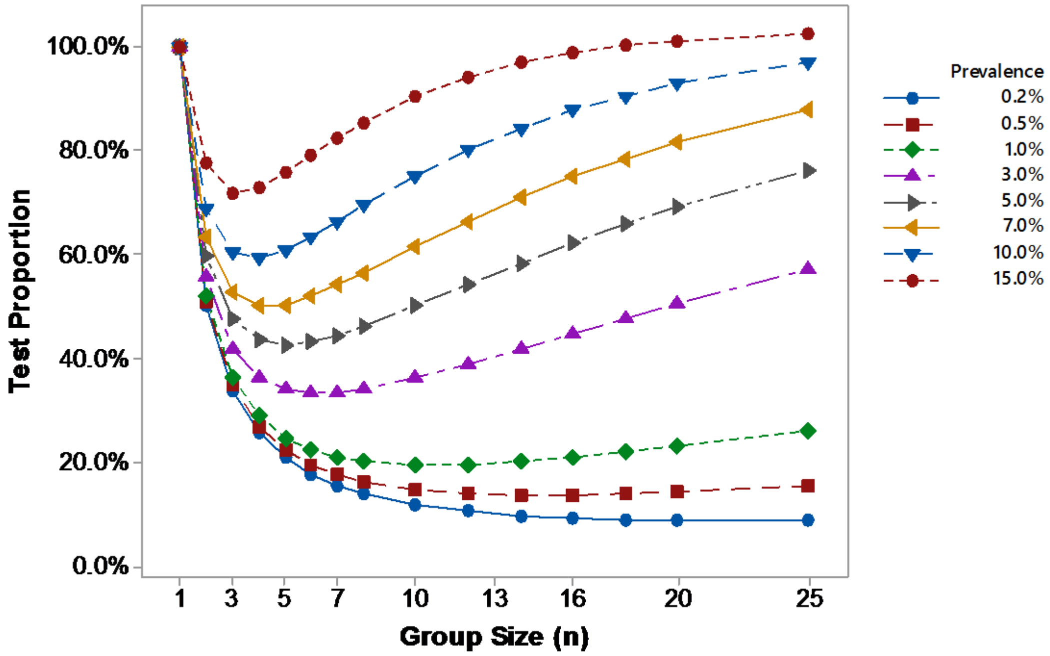

FIGURE 1: Test proportion using group sample test strategy at different prevalence levels

The optimal group size which yields the smallest test proportion can be obtained for each known prevalence levels. We noted that the group size(s) around the optimal group size also yield similar test proportions. This allows flexible group size selection in practice. Table 1summarizes the recommended group sizes at different expected prevalence levels. The test proportions for all these recommended group sizes are within 1% difference from the lowest test proportion.

Prevalence	Recommended Group Size	Estimated Test Proportion
0.20%	>14	8.9-9.9%
0.50%	10-20	13.9-14.9%
1%	8-14	19.6-20.3%
2%	6-10	27.4-28.3%
3%	5-8	33.4-34.1%
5%	4-6	42.6-43.5%
7%	4-5	50.2-50.4%
10%	3-4	59.4-60.4%
15%	3-4	71.9-72.8%

TABLE1: Recommended group sizes for population with different prevalence levels.

The test proportion is directly related to the cost and inversely related to the test capacity. For example, at prevalence level of 1%, with group size of 8-14, we expect the test proportion of ~20%, which mean only spend about 20% of the cost of the individual testing strategy. It also means the test capacity increased to about 1/20%=5 times of the test kits available.

At any selected group size, the smaller the prevalence level, the lower the test proportion, the lower the testing cost per sample, and the larger the testing capacity. Comparing to the optimal dose only recommendation, our multiple group size selection recommendation could provide increased flexibility for selecting ideal group sizes while prioritizing cost savings and increased testing capacity. If one believes the estimated prevalence level underestimates the true value, they can choose the least groups size from the recommended table. On the other hand, if one believes the given prevalence is higher than the true value, then they can choose the larger group size as suggested in the table. In practice, the group size selection must also consider the test reliability of the pooled samples, as well as the operational difficulties pertaining to time and cost of pooling samples. Normally, the operational and time cost for pooling samples likely increases as group size increases. In that case, we can choose the smaller groups size from the recommended group sizes. Another issue is that the knowledge of the estimated prevalence level may change overtime, requiring the selected group size be updated as data is collected overtime to maximize efficiency. If the updated group size includes the current group size, one does not need to change the group size to simplify the operation procedure. In addition, the prevalence levels sometime are associated with the patient's risk factors. In that case, the test proportion could be further decreased by stratifying the samples based on the patients' risk levels and selecting group size for each strata accordingly.

The simulation program is based on the open source R software platform. The R-code is attached for these who are interested to run their own simulation.

References:

[1]. Bilder CR, Iwen PC Abdalhamid A, Tebbs JM and McMahan CS (2020) Increasing testing capacity for SARS-CoV-2 by pooling specimens, *Significance*,https://www.significancemagazine.com/science/651-increasing-testing-capacity-for-sars-cov-2-by-pooling-specimens

[2]. Dorfman R. (1943) The detection of defective members of large populations. *The Annals of Mathematical Statistics* **14**, 436-440.

Attachment:

```
############################################################
# R Code
# n: group sample size;
# P: Prevalence, proportion of positive in population;
# X: number of positive sample in group size of nX~binom(n, P)
# Pn: probability of all negative samples in the group, or X=0;
#     Pn=dbinom(0, n, P) ;
# Pt: ration of tests needed and total sample size:
# if X>=1 then Pt=n+1/n and
# elseif X=0 then Pt=1/n
# expected test proportion=Pn*1/n + (1-Pn)*(1+1/n) = 1+1/n-Pn

# Samples are assumed to be independent

# written by Jeff Pan
# E-mail: Jeff.pan@osumc.edu
###########################################################

#######################################################
#step 1: create a matrix with various group size n and prevalence P
#####################################################
n=c(1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 16, 18, 20, 25)
p=c(0.001, 0.002, 0.005, 0.01, 0.02, 0.03, 0.05, 0.07, 0.10, 0.15)
a=length(n)
b=length(p)
row=a*b
col=4
out=matrix(rep(0, row*col), nrow=row, ncol=col)
colnames(out)=c("Group Size", "Prevalence", "Test Proportion", "Stdev")

for (i in (1:a)) {
        for (j in (1:b)) {
                out[((i-1)*b+j), 1]=n[i]
                out[((i-1)*b+j), 2]=p[j]}
}

#######################################################
#step 2: calculate the test proportion at different setting
#####################################################
```

```
for (i in (1:row)){
        Pn=dbinom(0, out[i,1], out[i,2])
        out[i,3]=1+1/out[i,1] - Pn
        out[i,4]=Pn*(1-Pn)
        if (out[i,1]<2 ) {
                out[i,3]=1
                out[i,4]=0}
}

######################################################
#step 3: view and save output results for further analysis
####################################################
out
write.csv(out, "P:/paper/out.csv")
```

Xueliang Pan, PhD. "How to Select the Size of Group Samples to Save Cost and IncreaseCapaci for COVID-19 Test." *IOSR Journal of Mathematics (IOSR-JM)*, 16(3), (2020): pp. 58-61.