Mathematics and Statistics 10(4): 808-815, 2022
DOI: 10.13189/ms.2022.100411 http://www.hrpub.org

Subset Intersection Group Testing Strategy

Maheswaran Srinivasan

Institute of Management Technology, Nagpur, Maharashtra, India

Received February 22, 2022; Revised May 17, 2022; Accepted June 22, 2022

Cite This Paper in the Following Citation Styles

(a): *[1] Maheswaran Srinivasan , "Subset Intersection Group Testing Strategy," Mathematics and Statistics, Vol. 10, No. 4, pp. 808 - 815, 2022. DOI: 10.13189/ms.2022.100411.*

(b): *Maheswaran Srinivasan (2022). Subset Intersection Group Testing Strategy. Mathematics and Statistics, 10(4), 808 - 815. DOI: 10.13189/ms.2022.100411.*

Abstract Many a time, items can be classified as defective or non-defective and the objective is to identify all the defective items, if any, in the population. The concept of group testing deals with identifying all such defective items using a minimum number of tests. This paper proposes probabilistic group testing through a subset intersection group testing strategy. The proposed algorithm 'Subset Intersection Group Testing Strategy' deals with dividing the whole population, if it is positive, into different rows and columns and individually testing all the defective rows and columns. Through this proposed strategy, the number of group tests is either always one when no defective is found or 1+r+c, where r and c denote the number of rows and columns, when at least one defective is found. The proposed algorithms are validated using simulation for different combinations of group size and the incidence probability of an item being defective (p) and implications are drawn. The results indicate that the average number of total tests required is smaller when p is small and considerably increases as p increases. Therefore, for the smaller values of p, this proposed strategy is more effective. Also, an attempt is made to estimate an upper bound for the number of tests through this strategy in various scenarios.

Keywords Probabilistic Group Testing, Expected Number of Tests, Relative Efficiency, Simulation

1. Introduction

While an objective is to identify and eliminate all the defective items in the population, testing only a sample of the population to decide about accepting or rejecting will not suffice. But at the same time, testing each and every unit of the population will not only be expensive but also a very cumbersome process. The concept of group testing deals with testing the group of items for the presence of at least one defective and testing each and every unit in the group to identify and eliminate the defectives ones if the group test is positive else with a single group test, a group is identified as non-defective. The idea of testing items in batches rather than individually testing items to minimize the number of tests to identify and eliminate defectives was discussed by Sobel et al. [1]. As the main objective is to identify and eliminate defective items in the group with the minimum number of tests, in group testing required the number of tests will be obviously a function of the number of defectives in the group.

The number of defectives in the population may be known or unknown, based on which the group testing method is categorized as combinatorial group testing (CGT) or probabilistic group testing (PGT). In CGT, the number of defectives is exactly known and the objective is to identify defective sets using the minimum number of tests whereas in PGT, the number of defectives is unknown and the objective is to minimize the expected number of tests using the suitable probability distribution of defectives to identify the defectives. As in many practical scenarios, the number of defectives in the population will be unknown. This study focuses on devising strategies for probabilistic group testing.

Group testing algorithms are classified as sequential and non-adaptive algorithms [2]. As the name suggests, in the sequential algorithm, the tests are carried out sequentially one by one and while determining the current test, the results of the previous tests are known whereas in the non-adaptive algorithm previous test result is unknown while determining the current test. An alternate approach

would be to use the combination of both sequential and non-adaptive algorithmic approaches namely a multistage algorithm. Since the goal of group testing is to minimize the number of tests to identify defectives, better to explore a sequential algorithm as it provides more information at the time of subsequent testing, this study also makes an attempt to devise a strategy using the sequential algorithm.

Unlike many other mathematical problems which can trace back to earlier centuries and divergent sources, the origin of group testing is pretty much pinned down to a fairly recent event- during World War II, and is usually credited to a single person- Robert Dorfman. The two economists Robert Dorfman and David Rosenblatt [3] were struck by the problem of the wastefulness of subjecting blood samples from millions of draftees to identical analyses in order to detect a few thousand cases of syphilis, while they were working for the price statistics branch of the research division in Washington during 1942 and the proposed solution to the same. It was shown in Robert Dorfman‘s article [3] that a different statistical approach can, under certain conditions, yield significant savings in effort and expense when a complete elimination of defective units is desired.

A few years later, Sobel et al. [4], the Bell Laboratories Scientists gave new meaning to the phrase ‘group testing’ by giving the subject a very thorough treatment and established many new problems for future studies in their 74-page paper. Though there were lots of researches taking place in the area of combinatorial group testing, not much works were carried out in the area of probabilistic group testing. F.k. Hwang’s [5] is a notable study that compares Li’s method, the binary search method and his own Hwang’s strategy.

The group testing design has been shown to be a compelling alternative to one-at-a-time testing in many areas where rare traits are of interest. [6,7] (discussed its use with vector-transfer experiments in plant pathology.

[8,9,10] used group testing to estimate HIV prevalence cost-efficiently. More recently, Xie et al. [11] demonstrate how pharmaceutical companies can use group testing to reduce costs in the early stages of drug discovery.

More recently, the concept of group testing was applied in the context of the development of DNA chips. In this application, the authors propose a microarray design methodology based on a group testing approach [12]. Ever since the concept of group testing was introduced much significant research had taken place with a good number of applications in different fields. According to Hughes et al. [13], pooled testing has been applied to problems in blood bank screening, drug discovery, epidemiology, veterinary medicine, and other areas. Besides this, the concept of group testing is also used for blood donation screening [14] and opportunistic testing of individuals for chlamydia [15], bovine viral diarrhea virus detection in cattle herds [16] and discovery of chemical compounds to use in new drugs [17]. Graham Hepworth et al. [18] used group testing in a study of virus infection levels in carnation plants life grown in glasshouses. According to Nguyen et al. [19], pooled testing is commonly used in public health settings, for both screening and surveillance of diseases and infections. Some of the applications include, but are not limited to, [20] used the concept of group testing human immunodeficiency virus testing programs and chlamydia and gonorrhea screening practices and Nicholas et al. [21] had used the concept of group testing to detect the West Nile Virus in mosquitos.

2. Proposed Algorithm

It is evident from the extant literature that less research work focused on developing strategies relevant to probabilistic group testing. Therefore, the current study focuses on probabilistic group testing.

The objective of the present study is to develop a strategy to identify and eliminate all the defective members of the large population using the minimum number of tests when the proportion number of defectives is relatively very less.

The proposed algorithm ‘Subset Intersection Group testing strategy’ deals with dividing the whole population, if it is positive, into different rows and columns and individually testing all the defective rows and columns. The proposed strategy is examined with the following assumptions. Firstly, the individual items being defective or not, is Bernoullian with the same ‘p’, that is each of the items in the group has the same probability of being defective p and the same probability of not being defective q=1-p irrespective of and independent of the quality of other elements in the group; secondly to test the entire group by one testing irrespective of its size to know whether the group has at least one defective item or is completely free from defective items is possible and finally, the cost of group and individual testing is assumed to be equal.

According to this proposed strategy, the whole group of items is arranged as a matrix of say p rows and q columns, defining blocks in two different ways (row blocks and column blocks) which are intersecting with one another. In this case, row blocks are tested and if blocks tested negative can be omitted, then column blocks are tested retaining only those column blocks that test positive. Thus if r row blocks and c column blocks are tested positive in a pxq design, one will obtain a small group of size rxc which contains all bad items in the group but at the cost of p+c group tests, this can be helpful when the incidence is quite small. This strategy is studied by simulation.

2.1. Test Strategy

The sequential stages of the proposed testing algorithm are as follows:

Stage 1: The group of n (n=pxq) items to be tested is arranged as a pxq matrix, each item being identified by its

position (i, j) in the matrix. The p rows are each tested as a block and those tested negative will be dropped from further testing. Let r1 be the number of rows tested positive. Each of these has at least one defective item. The q columns of this "residual' r1xq matrix are tested as blocks leaving say c1 defective columns. This residual set of r1xc1 (n1=r1xc1) elements will include all the defective elements in the group.

Stage 2 will now be tested in the same way by arranging these n1 elements (supplemented by some known good elements if necessary) into a matrix of size p2xq2. The process is repeated enough number of times. Thus at each stage, the group containing all the bad items is successively filtered out into the smaller and smaller group and when the size of this residual group falls below a threshold, all items are examined for defects. The total number of tests is thus

1+ (p1q1+p2q2+……. +pkqk) +pkqk

where k is the stage at which the residual group size falls below the threshold. Since the maximum of (ri, ci) is a lower bound to the number of defectives in the group this information may be used to advantage if interactive decision making is allowed. However, in the present study, only stage 1 is considered though as noted interactive decision making can effectively reduce the residual group size in the first stage.

If ci>ri, one can expect the number of defectives to be not more than 2 x ci and hence one can resize the matrix into 2 x ci x (ri/2) columns and rows which can lead to a considerable reduction in residual size.

2.2. Numerical Illustration

Let n=100, be the group size under consideration with the probability of an item being defective, p=0.02.

The following 100 items were generated randomly and arranged as a 10 x10 matrix with 1 indicating defective and 0 indicating non-defective items. We shall illustrate the procedure using this group data.

0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**
0	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

Step 1: Carry out a group test, gt=1, since it is positive the given group of 100 observations is arranged into a matrix of size 10x10 as explained above. Therefore, r=10, c=10.

Step 2: Now, carry out a group test for every row, and therefore the total number of group tests gt=gt+10=11; there are four rows (number of defective rows-ndr) ndr =4 which test positive, and the six rows which test negative will be dropped from further testing. Hence, the residual matrix will be

0	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**
0	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**

Step 3: Carry out a group test for each column, gt=gt+10=21, there are three columns (number of defective columns-ndc) ndc=3 columns testing positive, and the seven columns which test negative will be dropped from further testing. Therefore, the residual matrix will be of size 4x3 with each and every row as well as a column containing at least one defective.

0	**0**	**1**
0	**0**	**1**
1	**0**	**0**
0	**1**	**0**

Step 4: Carry out an individual test for all the items in the above matrix, it=4 x 3=12

Therefore, the total number of tests required for identifying all defectives in the group is T=1+r+c+ndr*ndc=1+10+10+4 x 3=33 i.e., gt=21 and it=12. Instead of carrying out 100 tests all defectives present in the group can be identified using 21 groups and 12 individual tests resulting in much saving time and expense in a number of tests.

3. Simulation

The effectiveness of the present strategy is examined through simulation. A data set X= {x_i ;(i =1, 2…. n)} is generated from uniform distribution U (0, 1). Then, this generated data set is converted to zeros and ones by applying the condition: if $x_i<p$ then x_i is equal to one and if $x_i \geq p$ then x_i is equal to zero. Further, zeros' in the data set represent to be non-defectives and 'ones' in the data set represent defective items. Simulation is carried out considering different combinations of group sizes and incidence probabilities. The combinations considered are group sizes n =100, 500 and 1000 with incidence probabilities p= 0.01, 0.05 and 0.10. Results indicating the number of individual and group tests with other relevant descriptive statistics are presented in Tables 1 to 3.

3.1. A Worst Case Scenario

A reasonable worst case analysis can be carried out as follows. Let n and p be the group size and incidence rates, then one can expect using the normal approximation for binomial the number of defectives to be not more than

$\mu+2\sigma<np+2\sqrt{npq}<np+2\sqrt{np}=d$, say with a chance of 97.5%. i.e., effectively the maximum number of defectives one can expect to have is effectively d. Hence, for r rows, c columns matrix formation, the number of tests $=1+r+c+d^2$ for a given n. r+c will be minimum when $r=c=\sqrt{n}$. Hence, in the worst case, the total number of group and item tests is to be not more than $1+2\sqrt{n}+(np+2\sqrt{np})^2$. Thus, for instance n=1000, p=0.01 we have $r=c\approx32$, np=10. Therefore, $d=10+2\sqrt{10}\approx17$ and hence the most of the time the number of tests will be not more than 1+64+289=354 for the group of 1000 items. If instead n=10000, r=c=100, np=100, $d=100+2\sqrt{100}=120$ and hence effectively upper bound to the total number of tests will be 1+2 x100 +14400 while for p=0.001, one gets $1+200+(10+2\sqrt{10})^2=490$ only. Thus, when the incidence rate is quite low, the subset intersection group test can be quite effective. Results of the upper bound of the number of defectives and the total number of tests using simulation for various combinations of n and p are presented in table 4.

Since d is not an effective upper bound (it grossly "over estimates") the number of defectives, d^2 is also not an effective upper bound for the residual set size. Hence, a second stage intersection testing of the residual set can reduce the number of tests to be much smaller value.

Let n and p be specified. Then the expected number of defectives=np and an effective number of defectives will be not more than $np+2\sqrt{npq}\approx np+2\sqrt{np}$ with a chance of it not less than 97.5%. When choosing r and c for reshaping the sequence of n items in the group, one has necessarily to have parallel test r+c times (with d>0) and hence, its minimum value, since r x c=n effectively $r=c=\sqrt{n}$. Hence, for group size n, the minimum number of tests required is $1+2\sqrt{n}$. However, when $np>2\sqrt{n}$, (as when p is relatively large enough) use of $r=c=\sqrt{n}$ often will not be of much use; the number of defectives will quite often be much larger than $\sqrt{n}$ and hence, each of the $\sqrt{n}$ rows and $\sqrt{n}$ columns can have at least one defective and hence, no rows or columns may be eliminated at all, making a complete enumeration a highly probable necessity. Hence, the logic of using $r=c=\sqrt{n}$ in the fair stage can be self-defeating.

An alternative strategy that is more likely to reduce the total number of tests is, to take r=(say) 1.5np, with c proportionately less than $\sqrt{n}$, making the number of group tests considerably lower than $\sqrt{n}$, but eliminating many rows as having no defectives, thus reducing the size of the residual set.

Simulation studies carried out taking this point into consideration show very promising results. Hence for various values of n and p the distribution of the number of tests appears to become nearly normal with $\sqrt{\beta_1}=0$ and $\beta_2=3$, where β_1 and β_2 denote skewness and kurtosis, see tables 1 to 3.

Table 1. Characteristics of subset intersection strategy for1000 runs with n=100

	P=0.01				P=0.05				P=0.10			
Descriptives→	Mean	S.D	Skewness	Kurtosis	Mean	S.D	Skewness	Kurtosis	Mean	S.D	Skewness	Kurtosis
number of defectives	0.961	0.978	1.0978	4.347	4.899	2.1103	0.4259	3.2195	9.913	3.027	0.3707	3.0867
number of group test	14.816	10.6388	-0.5289	1.278	23.885	1.6231	-14.0148	197.6092	36	0	*NaN	*NaN
number of individual test	5.734	5.1359	0.0841	1.4868	17.264	10.9886	1.1587	4.7535	31.501	11.3569	0.1699	2.6062
total number of test	20.55	15.281	-0.4424	1.2987	41.149	11.2853	0.8911	5.0915	67.501	11.3569	0.1699	2.6062
total cost	20.55	15.281	-0.4424	1.2987	41.149	11.2853	0.8911	5.0915	67.501	11.3569	0.1699	2.6062
relative saving %	79.45	15.281	0.4424	1.2987	58.851	11.2853	-0.8911	5.0915	32.499	11.3569	-0.1699	2.6062

*because p is large enough, there will be at least one defective in the generated 1000 trials and hence number. of group tests will be a constant, resulting in zero standard deviation.

Table 2. Characteristics of subset intersection strategy for1000 runs with n=500

	P=0.01				P=0.05				P=0.10			
Descriptives→	Mean	S.D	Skewness	Kurtosis	Mean	S.D	Skewness	Kurtosis	Mean	S.D	Skewness	Kurtosis
number of defectives	4.903	2.2514	0.4756	3.1544	25.025	4.8033	0.2107	3.0334	74.626	7.71	0.097	3.0746
number of group test	46.54	4.5792	-9.8346	97.8142	84	0	NaN	NaN	230	0	NaN	NaN
number of individual test	23.704	17.131	1.303	4.7799	148.394	29.3364	0.0282	2.7136	202.653	20.0098	-0.0504	2.9053
total number of test	70.244	18.3376	0.6099	5.7577	232.394	29.3364	0.0282	2.7136	432.653	20.0098	-0.0504	2.9053
total cost	14.0488	3.6675	0.6099	5.7577	46.4788	5.8673	0.0282	2.7136	86.5306	4.002	-0.0504	2.9053
relative saving %	85.9512	3.6675	-0.6099	5.7577	53.5212	5.8673	-0.0282	2.7136	13.4694	4.002	0.0504	2.9053

*because p is large enough, there will be at least one defective in the generated 1000 trials and hence number. of group tests will be a constant, resulting in zero standard deviation.

Table 3. Characteristics of subset intersection strategy for1000 runs with n=1000

	P=0.01				P=0.05				P=0.10			
Descriptives→	Mean	S.D	Skewness	Kurtosis	Mean	S.D	Skewness	Kurtosis	Mean	S.D	Skewness	Kurtosis
number of defectives	10.046	3.1451	0.38	3.4138	49.931	7.1037	0.1494	3.0609	149.87	11.4122	0.1129	3.0416
number of group test	66	0	NaN	NaN	159	0	NaN	NaN	455	0	NaN	NaN
number of individual test	80.151	43.1803	1.1569	6.0041	303.402	40.0083	0.0434	2.9131	406.899	29.6126	0.0782	2.9577
total number of test	146.151	43.1803	1.1569	6.0041	462.402	40.0083	0.0434	2.9131	861.899	29.6126	0.0782	2.9577
total cost	14.6151	4.318	1.1569	6.0041	46.2402	4.0008	0.0434	2.9131	86.1899	2.9613	0.0782	2.9577
relative saving %	85.3849	4.318	-1.1569	6.0041	53.7598	4.0008	-0.0434	2.9131	13.8101	2.9613	-0.0782	2.9577

*because p is large enough, there will be at least one defective in the generated 1000 trials and hence number. of group tests will be a constant, resulting in zero standard deviation.

Table 4. Upper bound for number of tests using subset intersection strategy for different values of n and p (worst case scenario)

n→ /p↓	100				500			
	Average no of defectives(d)	Upper bound for d	Upper bound for no. of tests	Observed maximum number of tests	Average no of defectives(d)	Upper bound for d	Upper bound for no. of tests	Observed maximum number of tests
0.001	0.1	0.7325	21.5365	34	0.5	1.9142	49.3856	69
0.002	0.2	1.0944	22.1978	34	1	3	54.7214	69
0.005	0.5	1.9142	24.6642	34	2.5	5.6623	77.7827	96
0.01	1	3	30	48	5	9.4721	135.4427	147
0.02	2	4.8284	44.3137	71	10	16.3246	312.2125	274
0.05	5	9.4721	110.7214	94	25	35	1270.7	315

n→ /p↓	1000				5000				10000			
	Average no of defectives(d)	Upper bound for d	Upper bound for no. of tests	Observed maximum number of tests	Average no of defectives(d)	Upper bound for d	Upper bound for no. of tests	Observed maximum number of tests	Average no of defectives(d)	Upper bound for d	Upper bound for no. of tests	Observed maximum number of tests
0.001	1	3	73.2456	98	5	9.4721	232.1427	243	10	16.3246	467.4911	563
0.002	2	4.8284	87.5593	98	10	16.3246	408.9125	416	20	28.9	1038.8	957
0.005	5	9.4721	153.9669	147	25	35	1367.4	1012	50	64.1	4315.2	2703
0.01	10	64.1	4256.6	426	50	64.1	4256.6	1866	100	120	14601	3702
0.02	20	28.9443	902.0164	511	100	120	14542	2036	200	228	52315	4053
0.05	50	64.1	4178.5	599	250	282	79454	2439	500	540	296920	4890

4. Conclusions and Scope for Further Study

As it is obvious through this strategy the number of group tests is either always one (if the group has no defective) or 1+r+c (when the group has at least one defective). It is observed from tables 1 to 3 that the average number of total tests required is smaller when p is small and considerably increases as p increases. Therefore, for the smaller values of p this strategy is more effective and results in more relative savings. Relative efficiency is calculated as the percentage of the ratio of a number of tests required using the proposed strategy to the number of tests required without using the proposed strategy. Another important feature of this strategy is that one can get lower and upper bounds to the number of defectives in the group: maximum (r, c) is a lower bound while rxc which is often large is not so effective upper bound, this feature is absent in other strategies proposed in probabilistic group testing by several researchers.

Other assumptions like the interdependence of items regarding their being defective (possibility of contagion) and the possibility of an upper bound for n, the group size for the test to unambiguously indicate non defectivity of the group though important on their own are kept for later study.

REFERENCES

[1] Sobel, M., Elashoff, R. M,. "Group Testing with a New Goal, Estimation", Biometrika, vol 62, no 1, pp.181–193, 1975. DOI: 10.2307/2334502

[2] Du, D., Hwang, F. K., Hwang, F., "Combinatorial group testing and its applications", World Scientific, vol. 12, 2000.

[3] Dorfman, R., "The Detection of Defective Members of Large Population", Annals of Mathematical Statistics, vol.14, pp.436-40, 1943.

[4] Sobel, M., & Groll, P.A., "Group testing to eliminate efficiently all defectives in a binomial sample", Bell System Technical Journal, vol 38 no 5, pp.1179-1252, 1959.

[5] Hwang, F.K., "A Note on Hyper Geometric Group Testing Procedures", SIAM Journal of Applied Mathematics, vol.34, no.2, pp.371-375, 1978.

[6] Tebbs, J. M., Swallow, W. H., "Estimating ordered binomial proportions with the use of group testing", Biometrika, vol 90, no 2, pp.471-477, 2003.

[7] Swallow, W., "Group testing for estimating infection rates and probabilities of disease transmission", Phytopathology, vol 75, pp. 882-889, 1985.

[8] Gastwirth, J. L., Hammick, P. A., "Estimation of the prevalence of a rare disease, preserving the anonymity of the subjects by group testing: Application to estimating the prevalence of AIDS antibodies in blood donors", Journal of statistical planning and inference, vol 22, no 1, pp.15-27, 1989.

[9] Gastwirth, J. L., Johnson, W. O., "Screening with cost-effective quality control: potential applications to HIV and drug testing", Journal of the American Statistical Association, vol 89 no 427, pp.972-981, 1994.

[10] Hammick, P. A., Gastwirth, J. L., "Group testing for sensitive characteristics: extension to higher prevalence levels", International Statistical Review/Revue Internationale de Statistique, pp. 319-331, 1994.

[11] Xie, M., Tatsuoka, K., Sacks, J., Young, S. S. "Group testing with blockers and synergis", Journal of the American Statistical Association, vol. 96 no. 453, pp.92-102, 2001.

[12] Schliep, A., Torney, D. C., Rahmann, S., "Group testing with DNA chips: generating designs and decoding experiments", In Computational Systems Bioinformatic, CSB2003, Proceedings of the 2003 IEEE Bioinformatics Conference. CSB2003, pp. 84-91 IEEE, 2003.

[13] Hughes-Oliver, J. M., "Pooling experiments for blood screening and drug discovery", In Screening, Springer, New York, pp. 48-68, 2006.

[14] Dodd, R. Y., Notari IV, E. P., Stramer, S. L., "Current prevalence and incidence of infectious disease markers and estimated window‐period risk in the American Red Cross blood donor population", Transfusion, vol 42, no 8, pp. 975-979, 2002.

[15] Mund, M., Sander, G., Potthoff, P., Schicht, H., Matthias, K., "Introduction of Chlamydia trachomatis screening for young women in Germany", JDDG: Journal der Deutschen Dermatologischen Gesellschaft, vol 6, no 12, pp. 1032-1037, 2008.

[16] Kennedy, J. A., Mortimer, R. G., Powers, B., "Reverse Transcription-Polymerase Chain Reaction on Pooled Samples to Detect Bovine Viral Diarrhea Virus by using Fresh Ear-Notch—Sample Supernatants", Journal of Veterinary Diagnostic Investigation, vol 18, no 1, pp.89-93, 2006.

[17] Remlinger, K. S., Hughes-Oliver, J. M., Young, S. S., Lam, R. L., "Statistical design of pools using optimal coverage and minimal collision", Technimetrics, vol 48, no 1, pp.133-143, 2006.

[18] Hepworth, G., Watson, R., "Debiased estimation of proportions in group testing", Journal of the Royal Statistical Society: Series C (Applied Statistics), vol 58, no 1, pp. 105-121, 2009.

[19] Nguyen, N. T., Aprahamian, H., Bish, E. K., Bish, D. R., "A methodology for deriving the sensitivity of pooled testing, based on viral load progression and pooling dilution", Journal of translational medicine, vol 17, no 1, pp. 1-10, 2019.

[20] Black, M. S., Bilder, C. R., Tebbs, J. M., "Optimal retesting configurations for hierarchical group testing", Journal of the Royal Statistical Society: Series C (Applied Statistics), vol 64, no 4, pp.693-710, 2015.

[21] Pritchard, N. A., Tebbs, J. M., "Estimating disease prevalence using inverse binomial pooled testing", Journal of agricultural, biological, and environmental statistics, vol 16 no 1, pp.70-87, 2011.