RT-PCR Pooling Analysis

Loay and Jayant

Group Testing for COVID-19: How to Stop Worrying and Test More

- Paper has considered very high infection reates as well, upto 33%. Other papers have only considered what happens when prevalence is below 5-10%, and have given no thought as to what happens for higher prevalences.
- Moreover, this paper has actually considered how false negatives and false positives come to be (Eg: Dilution, pooling, etc).
- Multiple group testing schemes for differing circumstances have been presented. For instance, different schemes have been recommended for various ratios of Number of infected people vs Total population

False Negatives and Positives

The false negative and positive rates have been well defined and simulated. We care more for false negatives than false positives. According to the paper, when the number of RNA copies (virus “particles”) are 5, the rate of a false positive is 0.07.

The paper also presents graphs regarding how the false negative rate changes with sensitivity and number of test replications.

Replication is necessary when the prior is more than 8% at a false negative and positive rates of 5% and 10%, respectively, for the testing technique.

For a very small rate of false negatives, we do not need to replicate a test more than twice.

Testing Schemes

Two paradigms of group testing discussed:

- Combinatorial Group Testing (CGT): These testing schemes require a fix or an upper bound on the number of infected samples. As long as the bound holds, the CGT schemes will always identify all the infected people. Eg: Adaptive Testing.
- Probabilistic group testing (PGT): In this scheme we require a prior probability which indicates probability of people infected in the Population that has already been tested.The probability of detecting all the infected patients is very close to 1 in this scheme.There exists a trade off between probability and average number of tests.

Practicality of GT

- GT is said to be efficient when the ratio of D/N (number of infected to total) is less than (3-root5)/2. In eneral, it needs to be below 33%.
- Effects of Dilution: If V is the amount of viral copies (RNA copies) present in a swab/sample, and L is the minimum number needed for an accurate test, then we wish to divide a pool into less than V/L samples. If W is the minimum load for a false negative, then we cannot do more than V/W replications of a single sample.
- Pooling Dilution:When a swabs of the infected population are mixed with the virus free swabs,the viral load of the original samples becomes diluted. Hence there is a maximum group size we can have for each value of D and pooling any group greater than that size leads to inaccurate results.

Algorithms

Three Algorithms have been presented:

- CGT Algorithms
  - GBS: Generalized Binary Splitting, similar to the binary search.
  - MST: Multistage Testing
- PGT Algorithms
  - NT: Nested Testing

GBS

GBS is simply recursion applied to the usual binary search algorithm.

In the worst cas scenario, the number of tests needed would be log(NCD)+D where D is the numer of infected samples.

GBS is more efficient for large values of N. For large values of D, GBS can fail (take more than N tests). This is an inherent drawback. Refer to graph in paper

Practical Considerations of GBS:

- Tests are performed in a sequential manner and hence this takes more time as we can’t conduct tests in a parallel fashion
- The GBS algorithm may not work properly if the value of D is underestimated. The number of infected patients must be <=D.To prevent this we pool the remaining N-D samples after this algorithm is over and test them to check if the value of D was not underestimated
- Inherent Repetition: Each positive sample in this Group is tested multiple times and hence the accuracy of this test is high. However in some cases certain samples will not be tested altogether.

Multi Stage Testing

This is a sort of mix of Adaptive and Non Adaptive testing.

Algorithm to be explained with reference to paper.

- PRACTICAL CONSIDERATIONS:
    1. All groups can be tester in parallel and hence the process can be sped up
    2. Each sample only undergoes a maximum of s tests and hence sample preparation becomes easier as we know the exact number
    3. Even if the value of D is underestimated, its value can be updated each time if we find out that more than D groups have tested positive.
    4. The maximum number of tests in MST is N

Nested Testing

Probability of detection is 1, regardless of the value of D

However, in some cases the total number of tests required may be momre than N. (Suboptimal)

NT relies heavily upon the availability of independent samples. (Geographically spaced)

Moreover, for a sample that tests negative, NT ensures it is tested twice. Low prob of false negative.

Computationally expensive.