A Systematic Study of Pooled Testing Procedures for SARS-CoV-2

David Zhou, Computer Science
Mentor: Dr. Huan Liu, PhD, Professor
School of Computing, Informatics, and Decision Systems Engineering

Motivation

Frequent testing is needed to monitor and limit the spread of COVID-19. However, its cost (US $30-50 per test [1]) and shortage of test kits hinder high-volume routine testing. Studies show that pooled testing is effective in reducing the number of tests needed to diagnose a population. This study compares several pooling algorithms, identifies their optimal parameters, and suggests the best pooling procedure for populations in Arizona.

Background

A literature search was conducted between 11/2020 to 03/2021. To generalize the methods, nineteen pooling procedures were placed among four categories:

- **Dorfman 2-stage:** a population is randomly partitioned into fixed-sized groups. Samples are combined in each group, then tested. A negative result implies the group is negative; a positive result requires individual retesting.
- **Multipool:** each individual is placed into multiple groups. The set of positive groups may be decoded to pinpoint positive samples without the need to retest.
- **Social network:** groups are formed based on relationships in the population (e.g., families, classrooms), since positive samples are likely to be positively correlated.
- **Binary search, poisoned wine:** groups are tested; positive groups are split, then retested. Continue until all groups are negative or the group size is equal to one.

The maximum recommended pool size is **64** [2].

Methodology

The following three metrics were used to compare the performance of each category:

- **Tests per individual**
- **Time to diagnose a population**
- **Prerequisites for constructing pools**

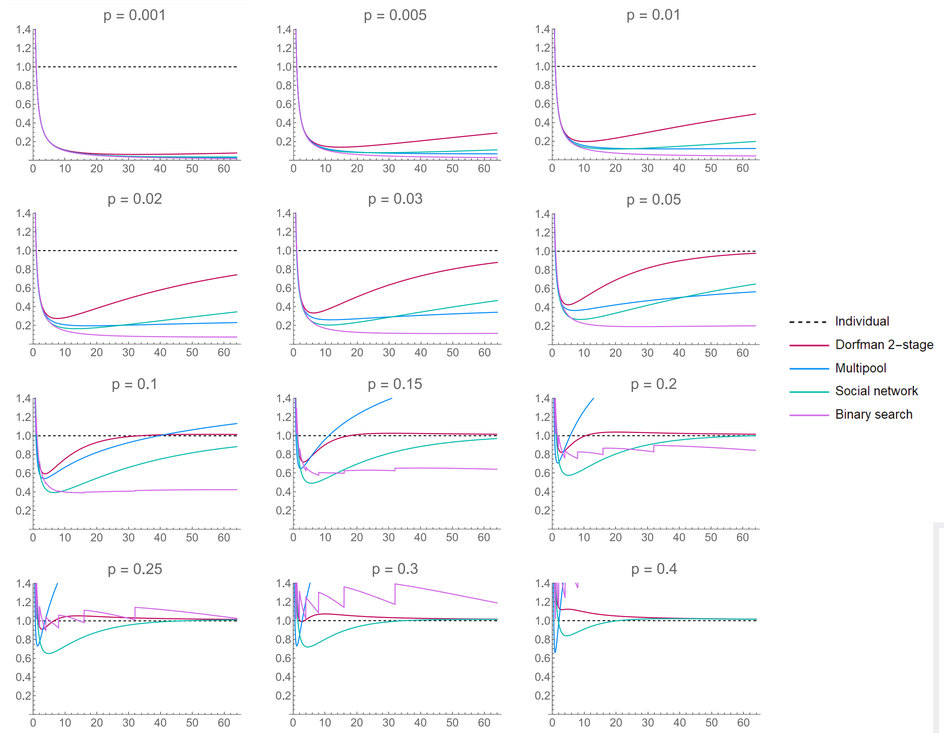
The minimum tests per individual was solved at each prevalence level *p* to identify the **optimal group size *n*** of each category.

Selected studies for numerical comparison of tests per individual:

- **Dorfman 2-stage:** Aragón-Caqueo, D. et al. [3]
- **Multipool:** Mutesa, L. et al. [4]
- **Social network:** Lin, Y. et al. [5]
- **Binary search, poisoned wine:** Srivastava, A. et al. [6]

Results

Tests per individual



Optimal group size

p	Dorfman 2-stage	Multipool	Social network	Binary search
	n			
0.001	32	64	58	64
0.005	15	64	26	64
0.01	11	35	19	64
0.02	8	17	14	64
0.03	6	12	11	46
0.05	5	7	9	28
0.1	4	3	7	16
0.15	3	2	6	8
0.2	3	2	5	8
0.25	3	1	5	4
0.3	3	1	5	1
0.4	3	1	5	1

Time to diagnose a population

procedure	# of stages
Dorfman 2-stage	2
Multipool	$1-3$
Social network	2
Binary Search	$\log_2 n$

Prerequisites for constructing pools

procedure	prerequisites
Dorfman 2-stage	none
Multipool	none
Social network	create social network graph
Binary Search	use results of previous stage

Suggested procedure for Arizona populations

Methodology

The procedure with the fewest tests per individual is selected based on current prevalence and group size.

Group size is limited by the daily number of conducted tests.

Prevalence and group size data were collected from azdhs.gov [7] on 11 April 2021

Location	p	n	procedure
Arizona	0.137	≤ 64	social net.
Apache	0.128	≤ 64	social net.
Cochise	0.118	≤ 64	social net.
Coconino	0.115	≤ 64	social net.
Gila	0.114	≤ 64	social net.
Graham	0.172	≤ 52	social net.
Greenlee	0.073	≤ 7	binary s.
La Paz	0.094	≤ 46	binary s.
Maricopa	0.143	≤ 64	social net.
Mojave	0.158	≤ 64	social net.
Navajo	0.152	≤ 64	social net.
Pima	0.117	≤ 64	social net.
Pinal	0.137	≤ 64	social net.
Santa Cruz	0.189	≤ 64	social net.
Yavapai	0.125	≤ 64	social net.
Yuma	0.194	≤ 64	social net.

Conclusions

- Pooled testing is much more efficient (~10x) than individual testing at low prevalence.
- Optimal group sizes decrease as prevalence increases.
- All procedures break even with individual testing around p = 0.3.
- Social network and binary search procedures are generally the most efficient in the number of tests per individual.
- Dorfman 2-stage and multipool procedures are less efficient, but no prior data is needed to construct pools. All their pools can be decided at once, saving time, whereas pools for social network and binary search depend on previous data.
- Dorfman 2-stage, multipool, and social network procedures take similar time, since they are derived from Dorfman's original pooling algorithm (1947).

Future Work

- Analysis on social-network-specific features (correlation constant, sparse vs. dense network) and multipool decoding algorithms.
- Consider test specificity and sensitivity.

References

[1] *Medicare Administrative Contractor COVID-19 test pricing*. (2021). U.S. Centers for Medicare & Medicaid Services. https://www.cms.gov/files/document/mac-covid-19-test-pricing.pdf

[2] Yelin, I., Aharony, N., Tamar, E. S., Argoetti, A., Messer, E., Berenbaum, D., Shafran, E., Kuzli, A., Gandali, N., Shkedi, O., Hashimshony, T., Mandel-Gutfreund, Y., Halberthal, M., Geffen, Y., Szwarcwort-Cohen, M., & Kishony, R. (2020). Evaluation of COVID-19 RT-qPCR Test in Multi sample Pools. *Clinical Infectious Diseases*, *71*(16), 2073–2078. https://doi.org/10.1093/cid/ciaa531

[3] Aragón-Caqueo, D., Fernández-Salinas, J., & Laroze, D. (2020). Optimization of group size in pool testing strategy for SARS-CoV-2: A simple mathematical model. *Journal of medical virology*, *92*(10), 1988–1994. https://doi.org/10.1002/jmv.25929

[4] Mutesa, L., Ndishimye, P., Butera, Y., Souopgui, J., Uwineza, A., Rutayisire, R., Musoni, E., Rujeni, N., Nyatanyi, T., Ntagwabira, E., Semakula, M., Musanabaganwa, C., Nyamwasa, D., Ndashimye, M., Ujeneza, E., Mwikarago, I. E., Muvunyi, C. M., Mazarati, J. B., Nsanzimana, S., … Ndifon, W. (2020). A strategy for finding people infected with SARS-CoV-2: Optimizing pooled testing at low prevalence. *MedRxiv*, 2020.05.02.20087924. https://doi.org/10.1101/2020.05.02.20087924

[5] Lin, Y.-J., Yu, C.-H., Liu, T.-H., Chang, C.-S., & Chen, W.-T. (2020). Positively Correlated Samples Save Pooled Testing Costs. *ArXiv:2011.09794 [Cs, Stat]*. http://arxiv.org/abs/2011.09794

[6] Srivastava, A., Mishra, A., Parekh, T. J., & Jena, S. (2020). Implementing Stepped Pooled Testing for Rapid COVID-19 Detection. *ArXiv:2007.09780 [q-Bio, Stat]*. http://arxiv.org/abs/2007.09780

[7] *Data Dashboard*. (2021, April 11). [Database]. Arizona Department of Health Services; Arizona Department of Health Services. https://www.azdhs.gov/preparedness/epidemiology-disease-control/infectious-disease-epidemiology/covid-19/dashboards/index.php

Acknowledgements

Many thanks to Dr. Huan Liu and Kaize Ding (Data Mining and Machine Learning [DMML] lab) for their guidance on the project.

FURI

Ira A. Fulton Schools of Engineering
Arizona State University