Welfare-Maximizing Pooled Testing

SIMON FINSTER
CREST-ENSAE
and
MICHELLE GONZÁLEZ AMADOR
UNU-MERIT & Maastricht University
and
EDWIN LOCK
University of Oxford
and
FRANCISCO MARMOLEJO-COSSÍO
Harvard University
and
EVI MICHA
University of Toronto
and
ARIEL D. PROCACCIA
Harvard University

This letter provides an overview of our recent work on COVID-19 testing mechanisms that appeared at EC'23. Large-scale testing is crucial in pandemics but resources are often prohibitively constrained. We study a scenario in which a population under lockdown utilizes a limited budget of tests to allow healthy individuals to resume in-person activities. Our work explores the optimal allocation of pooled tests in populations that are heterogeneous with respect to individual infection probabilities and utilities that materialize if included in a negative test (and being permitted to resume in-person activities). Non-overlapping allocations of tests, where no individual in the population is included in more than one pooled test, are both conceptually and logistically simpler to implement. We show that the welfare gain from overlapping testing over non-overlapping testing is bounded. Moreover, we design a heuristic mechanism for finding test allocations that is fast and empirically near-optimal. We also implement our mechanism in practice and provide experimental evidence on the benefits of utility-weighted pooled testing in a real-world setting. Our randomized trial at a higher education research institute in Mexico suggests that performance and mental health outcomes of participants under our testing mechanism are no worse than under the counterfactual of full access for individuals without testing.

Categories and Subject Descriptors: []:

General Terms: Algorithms; Theory; Economics; Experimentation; Performance
Additional Key Words and Phrases: pooled testing, welfare maximization, approximation guarantees, COVID-19, experiment, algorithms

Authors' addresses: simon.finster@ensae.fr, mgonzalez@merit.unu.edu, edwin.lock@cs.ox.ac.uk, fjmarmol@seas.harvard.edu, emicha@cs.toronto.edu, arielpro@seas.harvard.edu

1. INTRODUCTION

A challenging reality of pandemic response is that policymakers are often forced to make decisions with imperfect information. During the COVID-19 pandemic, governments across the globe imposed lockdowns to curb viral spread. These policies resulted from a lack of fine-grained information regarding infection prevalence among the population but came at a high economic and social cost [Deb et al. 2022; Camera and Gioffré 2021]. Population testing programs emerged as a viable alternative to blanket lockdowns, allowing individuals who are healthy to resume normal activities. In practice however, finances and resources for testing can be prohibitively constrained. As a result, comprehensive individual testing is often infeasible particularly in low- and middle-income countries [Kavanagh et al. 2020; Dhabaan et al. 2020; Abera et al. 2020].

In this letter, we highlight the main technical contributions and experimental results from our work [Finster et al. 2023]. We also describe insights that helped us bridge theory and practice. Our work focuses on optimally using a limited budget of tests to alleviate the costs of lockdown for a given population under the assumption that individuals who are verifiably healthy are allowed to return to in-person activities. To extend the reach of a limited testing budget, we make use of pooled testing. In a pooled test, samples of multiple individuals are pooled together and tested as one. If this test is positive, we know that at least one individual in the pool is infected; otherwise, all pooled individuals are healthy.

Pooled testing traces back to [Dorfman 1943], who devised the technique in 1943 to screen large numbers of soldiers for syphilis, and has since been the subject of a vast literature.[1] Our point of departure from prior work is the observation that individuals in the given population may have different utilities for resuming in-person activities, as well as different degrees of viral exposure resulting in different probabilities of being infected. The core challenge lies in determining the most effective allocation of tests to maximize the expected utility of individuals capable of resuming in-person activities.

Most importantly, our problem setting and computational techniques are heavily motivated by a collaboration with the Potosinian Institute of Scientific and Technological Research (IPICYT) during the COVID-19 pandemic. IPICYT is a public research institution in the state of San Luis Potosí, Mexico. During the latter stages of the pandemic, IPICYT was subjected to a full lockdown. Moreover, researchers and students had differing priorities for returning to on-site facilities for their work. IPICYT had limited resources for qPCR tests but their in-house testing facilities were able to process pooled tests (of up to 5 individuals in each pool). This motivated our research collaboration with IPICYT to develop a pooled testing mechanism that tests those who most urgently need to return to in-person activities.

When given a fixed budget of tests and a population to test, the number of ways to allocate pooled tests is vast. An important practical constraint from our partners at IPICYT was that no individual be included in more than one test in

[1] See [Aldridge et al. 2019] and [Du et al. 2000] for extensive surveys on computational techniques in pooled testing.

an allocation (we call this a *non-overlapping* testing allocation), as the logistical overhead of splitting samples into multiple tests was deemed to be prohibitive, and potentially prone to human error. For this reason, our paper starts by quantifying the loss incurred by confining allocations to be non-overlapping. Our theoretical analysis establishes that the worst-case ratio between an optimal overlapping and non-overlapping allocation of tests is bounded by a constant factor, supporting the prioritization of non-overlapping testing. Recognizing the computational complexity of determining an optimal testing allocation (that may include overlapping tests), we introduce a greedy polynomial-time algorithm for computing non-overlapping test allocations that is both conceptually simple and and computationally efficient. As we show, it also guarantees an approximate solution with at least one-fifth of the welfare achieved in the optimal non-overlapping testing regime. Moreover, it performs near-optimally in numerical experiments that use real-world data. The greedy algorithm is used in our randomized trial due to its effectiveness.

Our algorithmic testing framework was empirically validated in a small-scale randomized control trial (RCT) conducted in 2022 at IPICYT in Mexico with a population of 130 individuals. The goal of our RCT was to showcase the feasibility of our testing mechanism in practice, and to understand the impact on performance and mental health outcomes compared to full reopening without testing. The trial protocol, ensuring in-person access only for those with negative qPCR test results, curtailed contagion within the institution. Importantly, the results indicated no adverse effects on participants' productivity and mental health, comparing favorably to a counterfactual scenario of unrestricted access without testing and highlighting the effectiveness of the proposed testing approach in practice.

Related Work. During the COVID-19 pandemic, pooled testing emerged as resource-efficient testing strategy [Sanghani et al. 2021; Mutesa et al. 2021; Nalbantoglu 2020]. Our work relates to recent contributions such as [Lipnowski and Ravid 2021], which study optimal testing allocations with respect to welfare maximization in heterogeneous populations. Our work introduces a more general setup, acknowledging diverse health probabilities and utility variations among individuals. We address the complexities of overlapping and non-overlapping testing scenarios, considering practical algorithms for optimal testing allocations that account for heterogeneous utilities. Moreover, [Ely et al. 2021], [Brault et al. 2021], and [Gollier and Gossner 2020] explore diverse aspects of test allocation, from differential costs and sensitivities to early screening and infection prevalence estimation.

2. MAIN THEORETICAL RESULTS

We consider a heterogeneous population in which every individual is characterized by an independent probability of infection and a utility of returning to in-person activities. We also have a limited budget of available pooled tests. In each test, a limited number of samples of individuals are pooled.[2] The test result is negative if all the samples are negative, otherwise the test is positive. Due to independence,

[2]Pool sizes in pooled tests are limited due to biological constraints. Our partners in Mexico have replicated techniques from [Sanghani et al. 2021] to achieve a maximal pool size of 5 with saliva samples.

the probability of a negative result is equal to the product of the probabilities of the individuals included in the test being healthy. A *test allocation* indicates the samples of which individuals are assigned to each available test. The expected utility of an individual under a test allocation is equal to their probability of being included in at least one negative test multiplied by their utility to return to in-person activities. The goal is to maximize the *welfare*, i.e. the sum of expected utilities, earned under a test allocation.

Performance of Non-Overlapping Testing. We are particularly interested in *non-overlapping* test allocations, which include each individual in at most one test. In general, overlapping testing can achieve higher welfare than test allocations that are restricted to not overlap. However, non-overlapping test allocations are often strongly preferred for logistical reasons, as was the case with IPICYT. A natural question is to identify how much welfare may be lost by restricting to non-overlapping tests. If the difference in welfare achievable with overlapping and non-overlapping testing is not too large, even institutions with the logistical capacity to run overlapping tests may choose the latter to reduce costs.

Given a fixed population and testing budget, we define the *overlap welfare ratio* as the ratio of the welfare of an optimal test allocation over the welfare of an optimal non-overlapping test allocation. The *gain of overlapping* given a fixed budget is the maximum welfare ration across all possible populations and this budget. The main question that we aim to answer is:

How large can the gain of overlapping become?

Surprisingly, we find that this gain is a small constant for any value of the budget.

THEOREM 2.1. *For any budget, the gain of overlapping is at most* 4.

Algorithms for Computing Approximately Optimal Testing Allocations. Given Theorem 2.1, we focus on providing efficient algorithms for computing optimal non-overlapping testing allocations. When there is just one test available, we have two efficient algorithms for computing an approximately optimal testing allocation: a modification of the FPTAS of [Goldberg and Rudolf 2020] that accounts for pool size constraints, and a mixed integer conic optimization problem which we solve with commercial solvers in practice. In both cases, the algorithms efficiently compute allocations which achieve a $(1-\varepsilon)$ proportion of the optimal single test welfare for $\varepsilon > 0$.

When there are more than one test available, we design a greedy algorithm which repeatedly allocates a single test to untested individuals in the population, using our algorithms for one test. We prove that this greedy algorithm provides a constant factor approximation to the optimal non-overlapping test allocation welfare. In practice, we observed that greedy is near-optimal.

THEOREM 2.2. *For any population and testing budget, the greedy algorithm achieves at least* $\frac{1-\varepsilon}{5}$ *of the welfare of the optimal non-overlapping test allocation, for any* $\varepsilon > 0$.

3. RANDOMIZED CONTROLLED TRIAL

In order to provide causal evidence of the efficacy of the testing mechanism, we designed a two-group randomized controlled trial (RCT). The data collected during the RCT allowed us to a) validate our theoretical results using real-life parameter distributions (health probabilities and utilities), and b) evaluate population welfare (aggregate utilities of returning to onsite work as measured by productivity, performance, learning, stress, subjective well-being). Our results indicate that our algorithm and testing protocol are, indeed, a viable policy tool for pandemic response.

In September 2022, as campus facilities reopened, we conducted the RCT with a population of 130 individuals. The treatment group followed our algorithmic testing strategy, with access to campus granted only upon receiving a negative test result. The control group was granted permission to return to the institute without testing. (The treatment and control groups were instructed not to interact, in order to prevent contagion.) Testing services were provided by the National Laboratory of Agricultural, Medical, and Environmental Biotechnology (LANBAMA), a testing facility within IPICYT.

We worked with researchers at IPICYT to gather population data to estimate individuals' utilities for in-person access and their health probabilities. In collaboration with epidemiologists at IPICYT and in the local state of San Luis Potosí we were able to achieve reliable estimates of infection probabilities. We also allowed individuals to express onsite work preferences for two-day windows through the allocation of a virtual token budget. This helped us avoid scheduling individuals for testing on days they did not wish to access IPICYT facilities in the first place, and allocate more tests to particular days that proved more popular.

To coordinate the RCT, we developed a web application. The web app included intake and outtake surveys for treatment and control participants, from which we collected information with regards to their social and economic well-being, and their preferences and beliefs about their subjective well-being. These data were used to i) estimate static utilities and ii) evaluate the (static and dynamic) population impact of the testing mechanism. Through the web app we also managed their token budget for weekly onsite work preferences, communicated weekly testing schedules, and allowed the lab to anonymously communicate test results to treatment participants.

Evaluation and methods. In the RCT, we measured subjects' stress levels and subjective well-being (life satisfaction), as well as self-assessed performance, productivity, and learning. We obtain these measures through survey questions that subjects are invited to answer before (baseline) and after (endline) the trial period. The treatment effect on the treated was estimated with bivariate linear regressions, based on static scores as well as first differences[3], using the above-mentioned outcomes as dependent variables. The trial period was of low viral prevalence in San Luis Potosí. The general decrease in contagion rates was reflected in our study:

[3]We also performed equivalence tests, and multinomial logistic regression models for non-normally distributed outcomes. These robustness checks corroborate our results from our preferred model specifications. We also collected a number of covariates for further robustness checks and heterogeneity analyses of our estimations.

only one pool tested positive in the treatment group. No individuals self-reported having experienced symptoms in the control group. This low number means that we are unable to make any strong claims regarding the protocol's health benefits.

We find no statistical evidence that our protocol has a negative effect on participants' work/study performance, learning, or mental health, despite the increased effort in coordination it demands from them compared to a full reopening (the protocol followed by the control group). At the same time, our strategy ensures greater safety for all participating individuals compared to a full reopening without any safety mechanisms in place. We conjecture that accounting for welfare is the crucial ingredient in our mechanism, enabling in-person access for those who need and benefit from it the most.

4. BRIDGING RESEARCH AND PRACTICE

The social relevance of our research lies in its successful implementation and population impact, grounded in strong theoretical foundations. [Lock et al. 2021] highlight that a crucial component of creating successful research-to-practice pipelines is establishing strong partnerships with local regulatory agencies; these partnerships guarantee that applied research projects are in congruence with local realities, both in terms of administrative constraints and knowledge of the local flow of know-how.

In the case of Mexico, our main partners were the state research councils belonging to Mexico's National Network of State Councils of Science and Technology (REDNACECYT)[4]. The support of the local council of San Luis Potosí, COPOCYT, was instrumental in facilitating introductions and encouraging collaborations between local researchers and our team. It provided an important degree of trust when working with researchers at IPICYT and beyond, such as when we reached out to local epidemiologists at the National Autonomous University of San Luis Potosi (UASLP) to help inform local infection rates to be used in the randomized trial of our protocol.

Moreover, as the pandemic progressed, COPOCYT also played an important role in helping our project remain aligned with the changing nature of local administrative constraints. This was especially important with regards to government policy around testing, as our protocol made heavy use of novel pooled testing lab methodology, for which health and regulatory policies were still being developed; their support also extended to the alignment of our project to changing national and local policies regarding lockdown regulations, which continued to evolve alongside the pandemic. Ultimately, our communication with these local partners helped us maintain a positive feedback loop. This ensured that the implementation of our protocol was as effective as possible while avoiding unnecessary complexity.

5. FUTURE WORK

Our pooled testing framework belongs to a growing algorithmic literature on resource allocation that incorporates population heterogeneity. Our testing protocol

[4]REDNACECYT and its constituent councils operate under the purview of the Mexican National Council of Science and Technology (CONACYT), which in turn serves as Mexico's primary government entity in charge of the promotion, funding and regulation of scientific and technological activities.

offers institutions an economical and secure solution to safeguard their entire community.

Looking ahead, there are several immediate avenues for future exploration. On the theoretical front, there is a gap between our upper bound of 4 and lower bound of 7/6 on the overlap welfare ratio, as well as an upper bound of 5 on the approximation factor of the greedy algorithm. We anticipate that even tighter bounds are attainable when utilities are confined to a fixed number of values, such as in dichotomous or trichotomous populations.

The testing and re-integration policy we advocate is static, assuming a one-shot setting where a testing budget is fully utilized by a policymaker. Dynamic testing, on the other hand, could adjust allocations adaptively based on previous test results. Furthermore, policymakers may have access to different types of tests, each with varying costs and performance metrics (e.g., pool size and sensitivity). Determining optimal budget-constrained allocations in this heterogeneous test setting remains an open question.

Crucially, we hope that the insights gained from the performance and efficacy of our welfare-maximizing testing mechanism will aid in better protecting resource-constrained communities during infectious disease outbreaks. Our pooled testing protocol may extend to mass screening for HIV/AIDS and prove equally significant in pooled frameworks for organ donation.

REFERENCES

Abera, A., Belay, H., Zewude, A., Gidey, B., Nega, D., Dufera, B., Abebe, A., Endriyas, T., Getachew, B., Birhanu, H., et al. 2020. Establishment of COVID-19 testing laboratory in resource-limited settings: challenges and prospects reported from Ethiopia. *Global Health Action 13,* 1, 1–7.

Aldridge, M., Johnson, O., and Scarlett, J. 2019. Group Testing: An Information Theory Perspective. *Foundations and Trends® in Communications and Information Theory 15,* 3-4, 196–392.

Brault, V., Mallein, B., and Rupprecht, J.-F. 2021. Group testing as a strategy for COVID-19 epidemiological monitoring and community surveillance. *PLoS computational biology 17,* 3, 1–25.

Camera, G. and Gioffré, A. 2021. The economic impact of lockdowns: A theoretical assessment. *Journal of Mathematical Economics 97*, 1–14.

Deb, P., Furceri, D., Ostry, J. D., and Tawk, N. 2022. The economic effects of covid-19 containment measures. *Open Economies Review 33,* 1, 1–32.

Dhabaan, G. N., Al-Soneidar, W. A., and Al-Hebshi, N. N. 2020. Challenges to testing COVID-19 in conflict zones: Yemen as an example. *Journal of global health 10,* 1, 1–3.

Dorfman, R. 1943. The detection of defective members of large populations. *The Annals of Mathematical Statistics 14,* 4, 436–440.

Du, D., Hwang, F. K., and Hwang, F. 2000. *Combinatorial group testing and its applications.* Vol. 12. World Scientific, Singapore.

Ely, J., Galeotti, A., Jann, O., and Steiner, J. 2021. Optimal test allocation. *Journal of Economic Theory 193*, 1–17.

Finster, S., Amador, M. G., Lock, E., Cossio, F. M., Micha, E., and Procaccia, A. 2023. Welfare-maximizing pooled testing. In *ACM Conference on Economics and Computation 2023.* EC '23. Association for Computing Machinery, New York, NY, USA.

Goldberg, N. and Rudolf, G. 2020. On the complexity and approximation of the maximum expected value all-or-nothing subset. *Discrete Applied Mathematics 283*, 1–10.

Gollier, C. and Gossner, O. 2020. Group testing against Covid-19. Tech. rep., EconPol Policy Brief.

Kavanagh, M. M., Erondu, N. A., Tomori, O., Dzau, V. J., Okiro, E. A., Maleche, A., Aniebo, I. C., Rugege, U., Holmes, C. B., and Gostin, L. O. 2020. Access to lifesaving medical resources for African countries: COVID-19 testing and response, ethics, and politics. *The Lancet 395,* 10238, 1735–1738.

Lipnowski, E. and Ravid, D. 2021. Pooled testing for quarantine decisions. *Journal of Economic Theory 198*, 1–30.

Lock, E., Marmolejo-Cossío, F. J., Jonnerby, J., Rajgopal, N., Guzmán-Gutiérrez, H. A., Benavides-Vázquez, L. A., Tello-Ayala, J. R., and Lazos, P. 2021. Optimal testing and containment strategies for universities in mexico amid covid-19. In *Equity and Access in Algorithms, Mechanisms, and Optimization*. Proceedings of EAAMO'21, USA, 1–9.

Mutesa, L., Ndishimye, P., Butera, Y., Souopgui, J., Uwineza, A., Rutayisire, R., Ndoricimpaye, E. L., Musoni, E., Rujeni, N., Nyatanyi, T., et al. 2021. A pooled testing strategy for identifying SARS-CoV-2 at low prevalence. *Nature 589,* 7841, 276–280.

Nalbantoglu, O. U. 2020. Group testing performance evaluation for SARS-CoV-2 massive scale screening and testing. *BMC Medical Research Methodology 20,* 1, 1–11.

Sanghani, H. R., Nawrot, D. A., Marmolejo-Cossío, F., Taylor, J. M., Craft, J., Kalimeris, E., Andersson, M. I., and Vasudevan, S. R. 2021. Concentrating Pooled COVID-19 Patient Lysates to Improve Reverse Transcription Quantitative PCR Sensitivity and Efficiency. *Clinical Chemistry 67,* 5, 797–798.