Documentation and results accompanying the op-ed article "Omfattende og rettet testing for covid-19 er en billig, lite inngripende og effektiv intervensjonsstrategi" published in the Norwegian newspaper Aftenposten on June 16, 2020

Introduction

Previously, we have used our high-fidelity computational framework (HFM) for individual-based simulation of COVID-19 spread in Norway[1] to identify targeted pooled household testing (TPHT)[2] as a cost-effective and societally minimally disruptive intervention approach. Briefly, TPHT consists of selecting the largest households in a region with COVID-19 contagion and conducting tests weekly. Whenever a pooled sample from one of the selected households returns a positive result, the whole household is isolated. With the April 22, 2020 release of a preprint demonstrating that simple saliva samples are more accurate and sensitive than the standard nasopharyngeal swabs[3], and with the May 8 approval by the U.S. Food and drug administration (FDA) of the first device for home-collected saliva samples[4], it is feasible to implement TPHT large-scale testing.

Our individual-based computational framework is based on each individual following SEIR-type state transitions upon infection. The possibility of infection is induced through a multi-layer heterogeneous network based on a person's participation in groups, such as home or nursing home, work or different levels of school, and a random contact network that explicitly accounts for heterogeneity and for which a new instance is constructed every day[1]. The age distribution and number / size of groups in each layer follows detailed demographic data for each of Norway's municipalities[1]. We have assigned different probabilities of infection in each level, and these rates have been adjusted to make the model-predicted average COVID-19 hospitalization curve for Oslo follow empirical data[5]. Thus, it is straightforward to directly implement a variety of societal interventions, such as school closures and people working from home.

Additionally, this modelling framework directly accommodates heterogeneous introduction of social distancing effects through the (differential) reduction of spread probabilities in different layers (e.g. social distancing may be less effective in primary school than at work locations). The critical importance of heterogeneity in epidemiological modelling has been discussed elsewhere[6]. Note also, that with the HFM framework, it is not necessary to presume a reproduction number R, as it is not a fundamental parameter. Instead, R is a direct consequence of both the heterogeneous complex network and the choice of infection probabilities in each layer.

---

[1] https://s.ntnu.no/HFMdescription
[2] https://s.ntnu.no/covid19_tpht
[3] https://doi.org/10.1101/2020.04.16.20067835
[4] https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-first-diagnostic-test-using-home-collection-saliva

[5] See e.g. Results: May 7, 2020 section. https://s.ntnu.no/COVID19results
[6] See e.g. pages 7-14, https://www.nber.org/papers/w27007.pdf

Based on the successful model representation of the development in Oslo since medio March, we have used our HFM to assess the efficiency of the TPHT strategy for handling an unrealistically large insertion of COVID-19 cases. We investigate a set of scenarios that test the stability of the *status quo* in two major Norwegian cities in response to an abrupt insertion of a large number of infections. While the number of infections inserted (Trondheim: 1,000; Oslo: 5,000) into each city is significantly higher than what is reasonable to expect due to domestic or international travel, this computational experiment reveals the effectiveness and responsiveness of different testing approaches. Since the number of infected cases we introduce is large, we are of the opinion that this stress test is representative for extreme realizations of a large number and combination of scenarios. Interventions that are effective for such an exaggerated number of inserted infections will also be even more effective for a smaller (and more realistic) number of inserted infections.

Description and justification of major premises and approaches

Recently, there have been claims that up to 80 percent of people might not be susceptible to SARS-CoV-2 due to some unknown prior immunity factor[7]. Besides a recent paper in Cell[8] showing that T cells in blood drawn from people between 2015 and 2018 recognized and reacted to fragments of the Sars-CoV-2 virus in 40-60% of cases, there is scarce evidence for this claim. In terms of modelling, such cross-reactivity would in any case have to be included from the beginning of the epidemic and not something that is introduced to explain its fast decline. Due to the paucity of clear-cut evidence for the existence of a large percentage of insusceptible and non-infectious individuals, our modelling has been guided by the anticipation that the available data can be explained by the tremendous impact authority-imposed and self-imposed behavioral changes have on the infection dynamics. Epidemiological modelers have repeatedly the last 20 years underestimated the impact of self-imposed behavioral change and thus made very grim forecasts that did not materialize. Self-imposed behavioral change can easily explain the lack of exponential growth in the pre-lockdown phase in many countries, and we think that a larger part of the epidemiological modelling community should become more appreciative of this mechanism.

The parameters of our model are calibrated using hospital commitment data for Oslo through May 25th, using the progression of hospitalizations from early March to late May as the reference quantity. Details surrounding the specific parameter values chosen are presented on our web page (file: Parameters table – June 15, 2020). This specific parameter set is also used for simulations of Trondheim. Owing to the consistently low number of hospital commitments in Trondheim, thus dominated by stochastic effects, we did not use Trondheim data for model calibration. And it is comforting to observe that our projected spread pattern is consistent with existing data for Trondheim without any parameter fitting to these data.

It is important to pay attention to the system-level response after introduction of an infection spike: (1) effective stopping time, which captures how quickly the curves descend towards a low level, and (2) area under the curves, which corresponds to the total number of infected or deceased cases. These characteristics are key to evaluating the effectiveness of an intervention. Except for school closures, we have made the conservative choice of not including the effect of typical vacation travel in July and assume that people behave as normal. We believe this over-estimate the likelihood of disease spreading

[7] https://english.alarabiya.net/en/coronavirus/2020/06/06/Coronavirus-Up-to-80-percent-of-people-not-susceptible-to-COVID-19-says-data-expert
[8] https://doi.org/10.1016/j.cell.2020.05.015

during the summer months. Also, we have not included seasonal variation in COVID-19 infectiousness. While it is yet unknown if there exists such seasonal variation, it is not unreasonable owing to the various similarities between SARS-CoV-2 and various influenza viruses with known seasonal variation in their spread.

We define the effective stopping time for an intervention to be the number of days lapsed until the number of cases has gone through two halving cycles, i.e. 25% of the number of insertion cases. In Trondheim, for this particular set of simulations, the stopping time corresponds to the first date for which the number of active infections is reduced to 250 cases after the insertion.

We investigate the effect of several different scenarios: (1) no testing; (2) only testing symptomatic cases, and (3) a "Targeted pooled household testing" strategy (TPHT). Symptomatic testing means that absolutely all individuals that start displaying symptoms are isolated for the duration of their disease 2 days after symptoms occur. The 2-day lag is intended to be a conservative accounting for uncertainty in when an individual decides if they are symptomatic and a delay from test until a positive result is present. Since this scenario is identifying and testing absolutely all infected that display symptoms, it encompasses 100% compliance and 100% efficient contact tracing of symptomatic cases[9]. For TPHT we select a number of households (e.g. 5,000, 10,000 or 18,000 in Trondheim) among all households of size 4 and larger. Since a large majority of the Trondheim student population lives in shared housing situations, their presence is included in this approach. All results are based on 50 repeat simulations to assess the effect of stochasticity in the spreading computations.

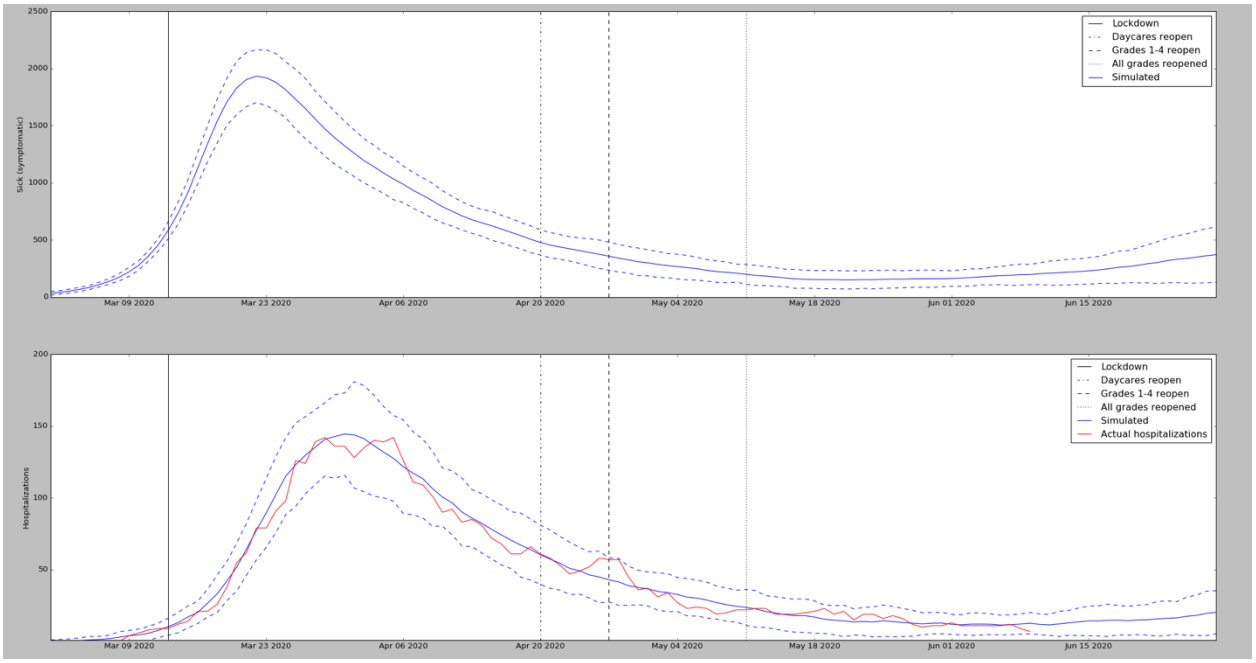
When comparing the different testing strategies, one should note that TPHT testing is scalable in the sense that the number of tests conducted can increase until the whole population is tested once a week. This is in stark contrast to voluntary symptomatic testing, where our scenario with perfect identification and subsequent isolation is the maximum achievable effect.

For Trondheim and Oslo, we present computer simulations of COVID-19 for three scenarios: one which follows a moderate natural progression in line with parameters determined from current evolution and where testing regimes are implemented early in response to a minor but significant increase in hospitalizations (50 hospitalizations in Oslo, 10 in Trondheim), and two scenarios conducted with an abrupt insertion of a large number of infected people in August. Here, the inserted cases are either spread evenly across the general population or concentrated amongst the student population (simply selected from ages 19-25).

Model calibration to actual hospitalization history in Oslo

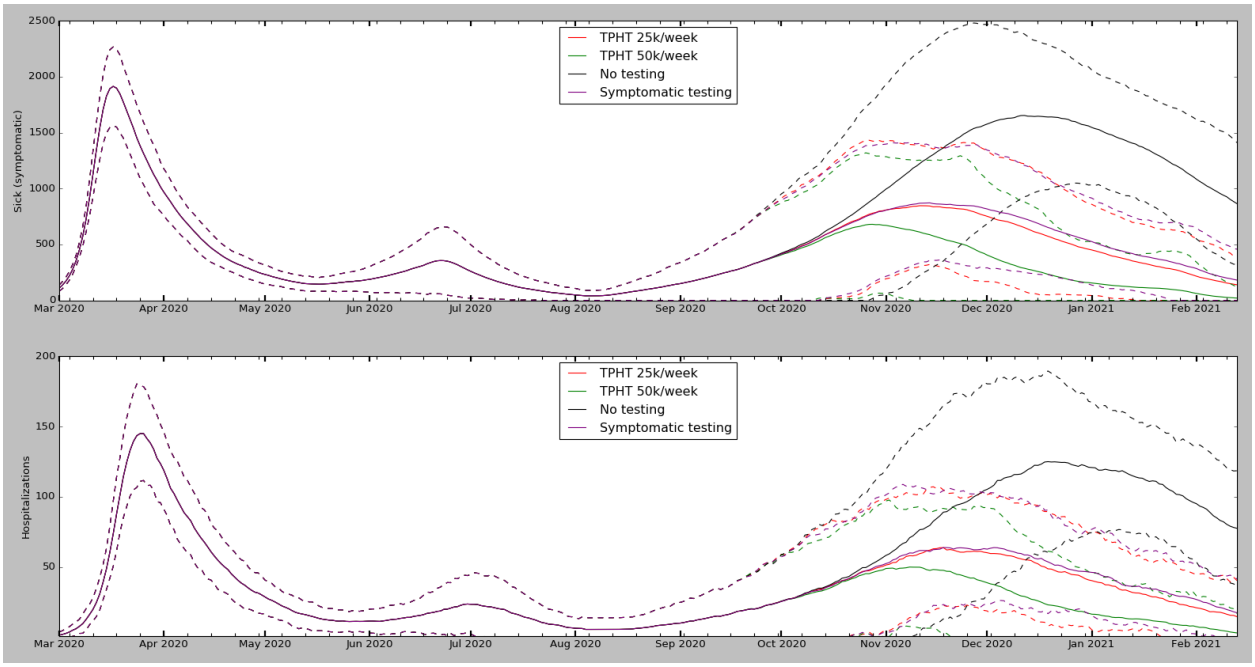
The following figure shows a comparison between recorded hospitalizations in Oslo as of June 10th, and the outcome of our simulations under current parameter choices (last updated June 1st, adapted to hospitalizations up to May 25th). Dates with closing or reopening of activity in the different layers are shown by the vertical bars. These interventions also take place in all following figures, but the corresponding lines are omitted as they would be highly compressed due to the extended time scale.

[9] We have not included testing of asymptomatic contacts in this first round due to lack of data concerning the efficiency of identifying asymptomatic cases through contact tracing. If such identification is highly successful, this will improve the intervention strength of the current policy somewhat. But the cost issue remains.



Model representation of the covid-19 development in Oslo since early March

The following figure shows computer simulation of COVID-19 spread in Oslo using our HFM framework, both without testing, with symptomatic testing and with TPHT testing implemented as a reactive measure initiated once a potential autumn wave reaches 50 hospitalized cases.



This figure is generated based on 50 separate simulations, and the solid lines represent the average behavior and the dotted lines the 2 standard deviation estimates. We see that in the current environment with the population implementing social distancing well, our simulations predict a markedly reduced R (approx. 1.2) compared to early March. Still, we find the possibility for a slight increase in contagion levels towards July. This is curbed by school closures. When they open again in August, the ensuing increase results in a broad peak stretching through the winter months. While this description is based on the average behavior, note that we find possible trajectories with no significant increase in cases throughout October (lower dotted black lines at or close to zero).

Observe that the change in population behavior in response to governmental social distancing guidelines is explicitly included in our model by reducing the contagion probabilities in several layers at the following dates: 13th of March: closing of all schools and daycares, 50% of workplaces, and an 83% reduction of infectivity in the random layer; 20th of April: re-opening of daycares, 27th of April: re-opening of schools up to fourth grade, 11th of May, re-opening of all schools. This change is sufficient to match the empirical trajectories of hospitalizations. Thus, following Occam's razor, it is quite unnecessary to appeal to esoteric effects such as "immunological dark matter"[10] or pre-existing immunity[11].

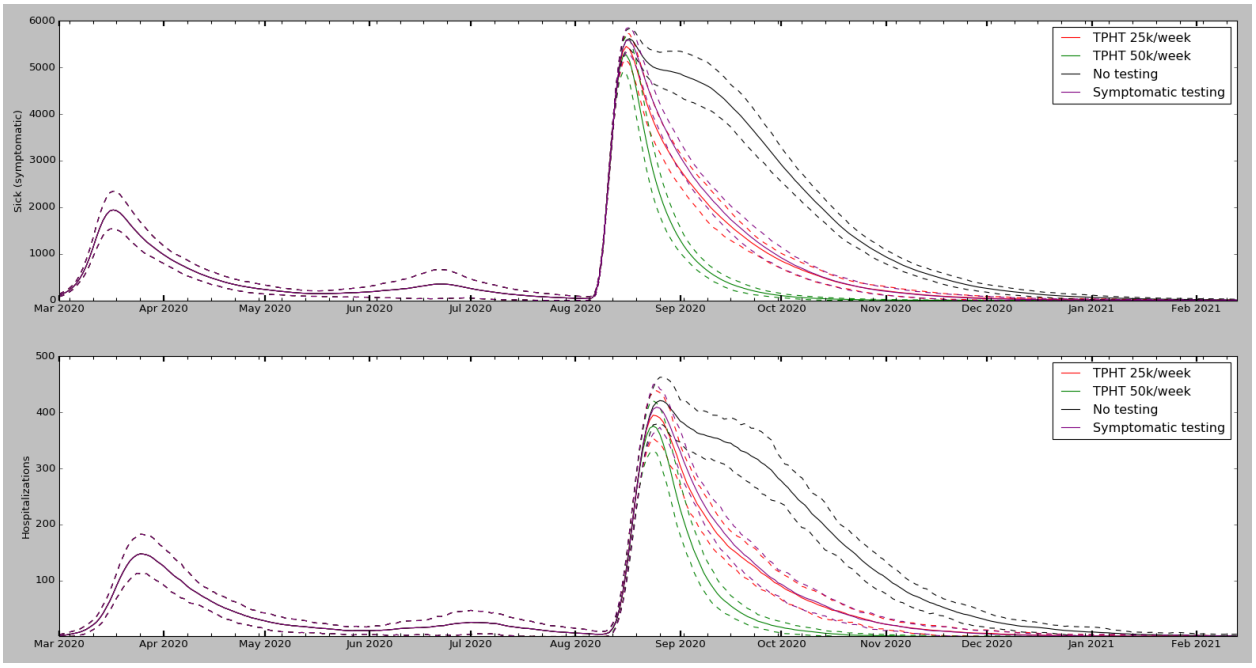
In this figure, we also show the effect of 3 different testing regimes: symptomatic testing, and responsive TPHT of 25,000 households per week and 50,000 per week. Note that the symptomatic testing, when conducted with perfect accuracy shows a similar effect to TPHT for 25k tests. However, symptomatic testing cannot do any better. Instead, this figure demonstrates that doubling the test capacity for TPHT to 50k per week is able to markedly drop the number of infected. Note that the lower

[10] https://www.theguardian.com/world/2020/may/31/covid-19-expert-karl-friston-germany-may-have-more-immunological-dark-matter
[11] https://doi.org/10.1101/2020.03.24.20042291

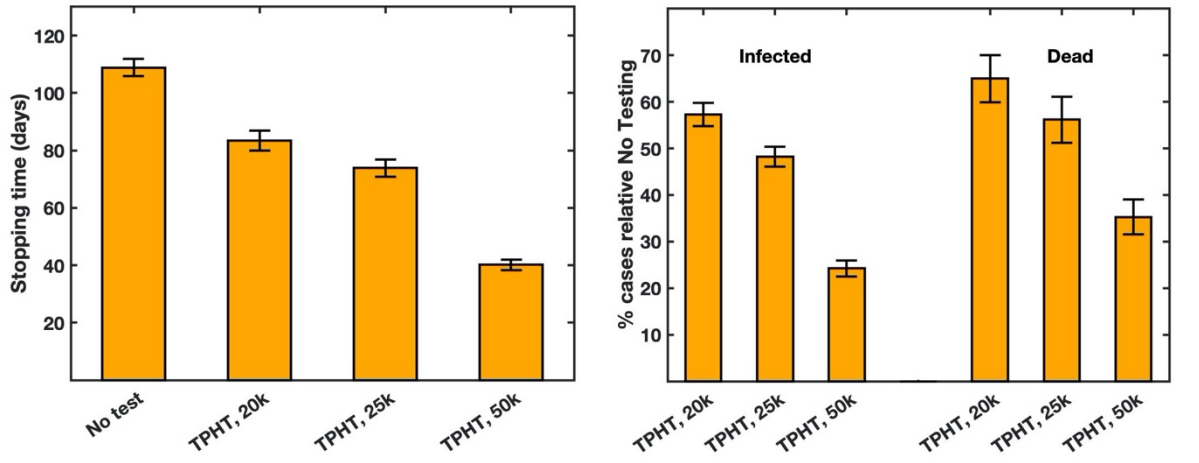
dotted line for 50k TPHT is at zero for almost the whole time-course. This means that within the possible outcomes are realizations where 50k TPHT almost completely eradicates COVID-19 spread in Oslo.

Insertion of 5,000 infected people in Oslo early August

In the following figure, we display results from the abrupt insertion of 5,000 infected people in early August. This spike is followed by either no testing, testing of all symptomatic cases, and the two levels of TPHT testing. Here, it is evident that a perfect implementation of symptomatic testing of absolutely all people displaying symptoms (and contact tracing) is about as effective as 25k TPHT. However, increasing the number of TPHT tests two-fold produces a significant decrease in the number of cases that is not possible to achieve with only symptomatic testing.



To further explore the effect of TPHT on the spiking of 5,000 infected inhabitants in August, we calculate the stopping time and total number of infected and dead for several levels of TPHT testing. Here, we define the stopping time as the number of days lapsed until the number of infected caused by the spike reaches 25% of the spike value, i.e. two cycles of halving.

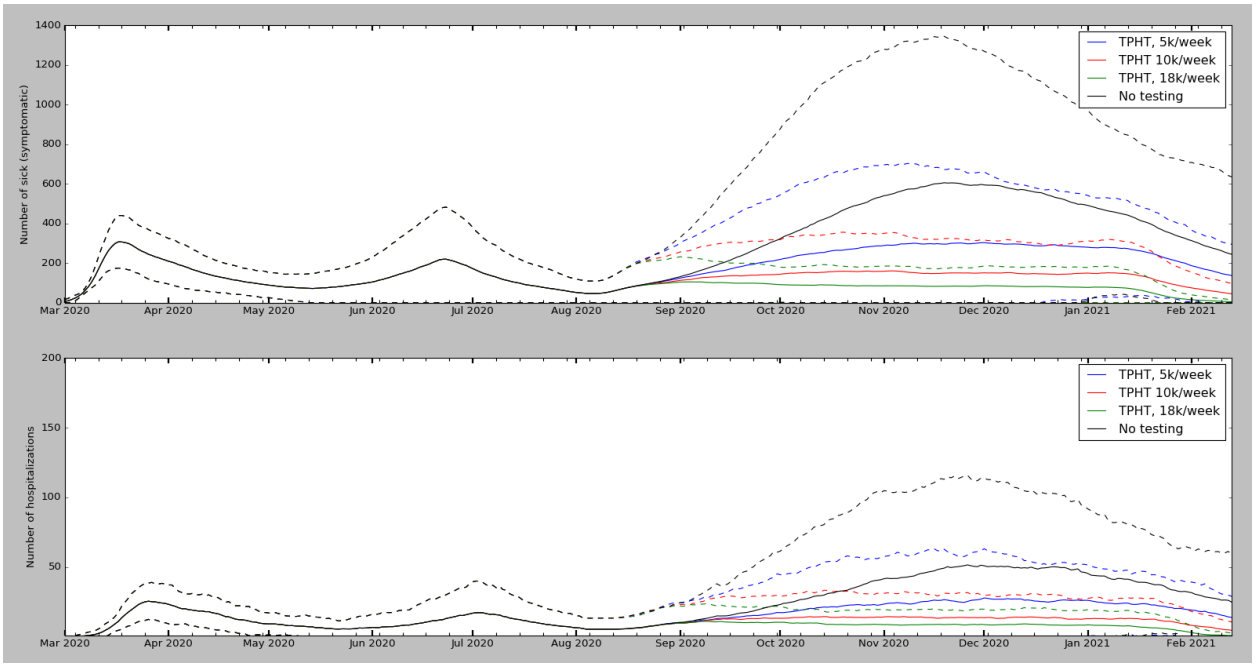


Effect of testing strategies in Oslo. Left panel: calculated stopping time after stress test of inserting 5,000 infected people into Oslo mid-August. Right panel: total number of infected or dead at simulation stop for different numbers of TPHT tests. Results are relative to total number of cases found in the No Test simulation (Average Infected: 95,133; Average Dead: 492)

Note that the TPHT testing of the 50k largest households in Oslo gives rise to a stopping time of 40 days. Since the typical infection trajectory from exposure to recovery for the individuals involved in the initial spike is about 2 weeks, and thus is an absolute shortest stopping time, this intervention is very effective.

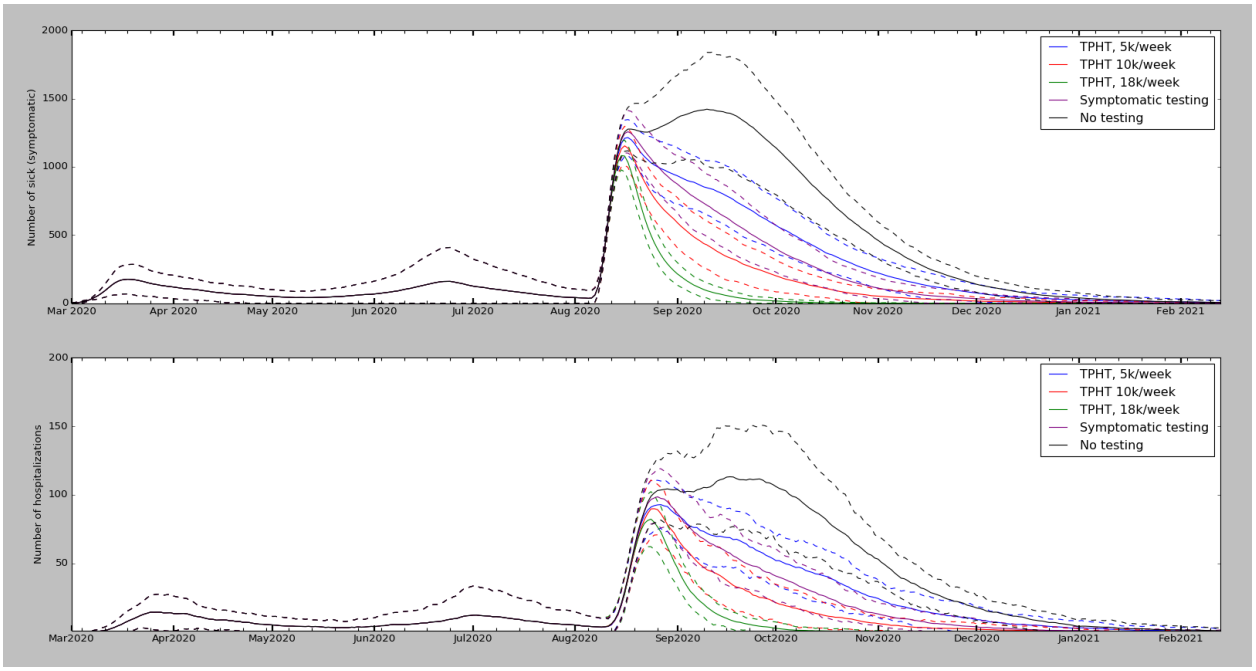
Insertion of 1,000 infected people in Trondheim early August

The following figure shows computer simulation of COVID-19 spread in Trondheim, when we follow a strategy of implementing testing as a reactive measure once a potential autumn wave reaches 10 hospitalized cases. Dashed lines correspond to upper and lower confidence intervals (2 standard deviations over 50 simulations):

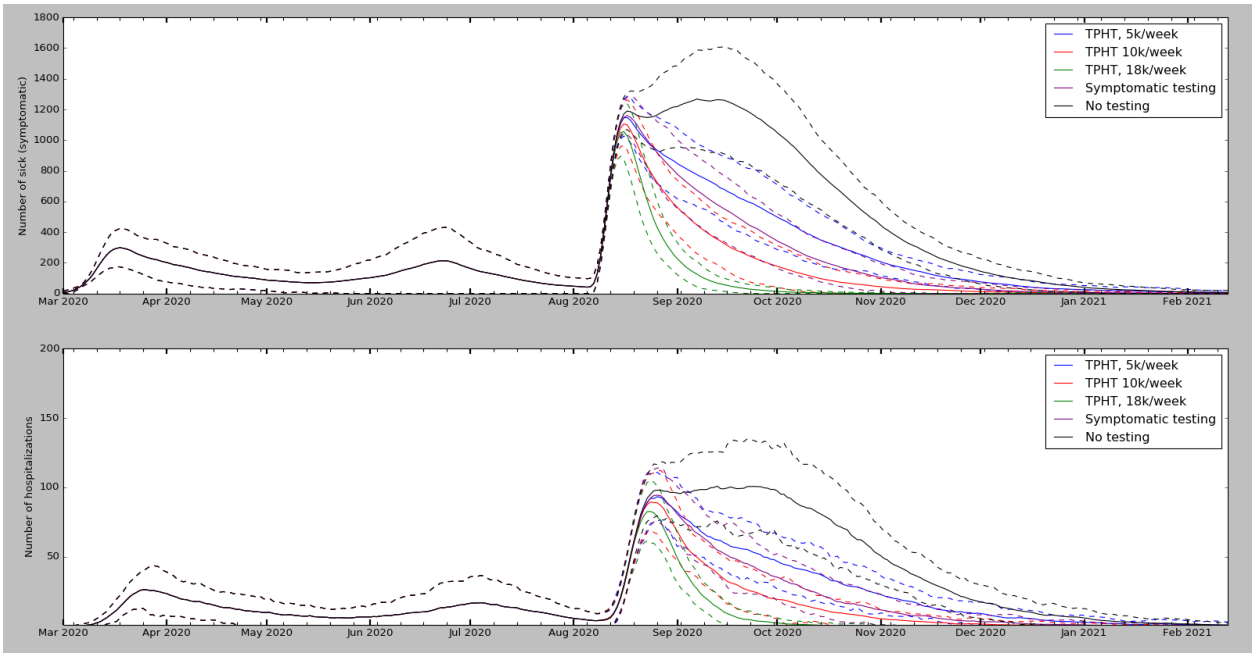


We note that our projected spread pattern is consistent with existing data for Trondheim without any parameter fitting to these data.

Similarly to the simulation of an infection spike in Oslo, the following figure shows an abrupt introduction of 1,000 infected individuals in Trondheim at the beginning of August. The infected people are randomly chosen from the population of Trondheim. Dashed lines correspond to upper and lower confidence intervals (2 standard deviations over 50 simulations):

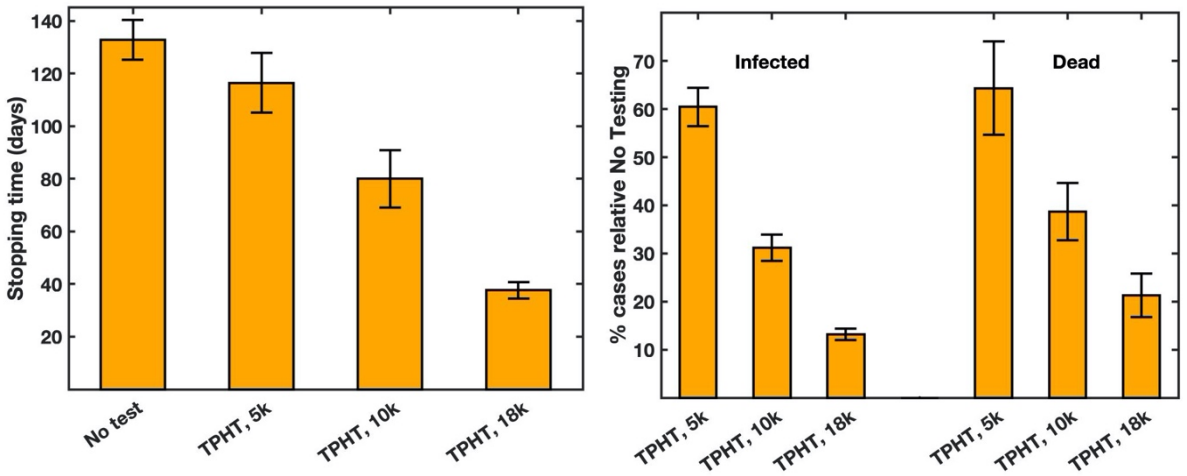


In the following figure, we show computer simulations of time development of COVID-19 spread in Trondheim with abrupt insertion of 1,000 infected students (randomly chosen individuals from the age bracket 19-25) in August. Dashed lines correspond to upper and lower confidence intervals (2 standard deviations over 50 simulations):



We notice that the concentrated spiking of the typical student population age groups results in very similar spread patterns as was found for the general Trondheim population. Additionally, we note that perfect testing of symptomatic cases gives a response a little more effective than that of 5k TPHT. However, testing only 5% of the Trondheim population (10k) using TPHT clearly gives a stronger response in suppressing the spread.

Effect of testing strategies in response to an abrupt insertion of 1,000 infected inhabitants in August for 5,000, 10,000 or 18,000 weekly TPHT tests.



Effect of testing strategies in Trondheim. Left panel: calculated stopping time after stress test of inserting 1,000 infected people into Trondheim mid-August. Right panel: total number of infected or dead at simulation stop for different numbers of TPHT tests. Results are relative to total number of cases found in the No Test simulation (Average Infected: 32,408; Average Dead: 154)

Discussion

Across our simulations, we note that an effective testing regime has a substantial effect on the spread of COVID-19 under currently determined infectivity parameters. The current policy of the Norwegian government is testing of symptomatic cases. While potentially effective, it also has some drawbacks in that COVID-19 seems to have a considerable potential for asymptomatic or pre-symptomatic contagion, and/or that infected individuals may not feel their symptoms to be strong enough to warrant seeking medical attention. In order to circumvent these issues, we propose a strategy where we perform systematic regular testing of larger households (in this document, 4 members and above, but the threshold can be adjusted to fit testing capacity) which we call Targeted Pooled Household Testing (TPHT).

The TPHT strategy is motivated by lessons from network epidemiology, in which it is found that preventing spread from highly-connected individuals ("hubs" in network parlance, or potential "super spreaders" in terms more familiar to the general public) is a highly efficient way of reducing R. Traditionally, this prevention of spread would be done by means of vaccination, which is currently unavailable. However, effective testing and quarantine can serve as an alternate prevention mechanism. While there are certainly other traits which may make an individual susceptible to act as a super spreader (and a targeted testing regime could also be applied to these individuals), large households are tempting targets for multiple reasons. First, household interactions, due to their frequency and closeness, are particularly likely to cause COVID-19 infection. Second, typical non-pharmaceutical interventions (quarantine/lockdown measures, face masks mandates, and similar) are harder to implement in the household sphere, and so individuals in large households are more likely to remain vulnerable even if such measures are imposed. Third, identifying the number of individuals in a

household is substantially easier than identifying individuals with a large amount of interactions in other social layers.

TPHT (or any other form of systematic testing of at-risk individuals) does not conflict with the current policy and can be used to supplement rather than replace it. We find that for Trondheim, a TPHT regime with 5,000 tests a week is comparable in efficiency to a regime in which every symptomatic case is identified and quarantined.

Compared to symptomatic testing, the main weakness of TPHT is the substantially larger number of tests required while the prevalence of the disease is low. However, TPHT has benefits in that it can be scaled according to testing capacity, and in that it does not rely on contact tracing in order to be effective. As contact tracing currently costs more than the testing itself (a cost which would rise even further if COVID-19 makes a resurgence in Norwegian society in the fall), and as advances in testing of COVID-19 through saliva is likely to lead to a substantial reduction in the cost of testing, it is quite possible for cost/benefit comparisons of the two strategies to shift in favor of TPHT in the coming months.