



Smart Containment with Active Learning (SCALE): Smart Testing Strategy

Version 2: May 19th, 2020

This SCALE [operational plan](#) lays out an evidence-responsive COVID-19 response strategy that seeks to minimize disease spread and limit adverse socio-economic impact by targeting response to the current & projected level of prevalence in the smallest feasible geographical units. The immediate need is large-scale testing to establish the level of prevalence in each area. This includes syndromic and sentinel surveillance, contact tracing and testing, and frontline worker testing. This document lays out the SCALE smart testing strategy. This is a living document and will be updated based on learnings.

Smart Containment with Active Learning (SCALE) is a multidisciplinary policy response to COVID-19 that draws on the expertise of researchers and practitioners in public health, infectious diseases, epidemiology, economics, policy and public management, technology and data science as well as business & non-profit leaders. We have assembled a broad coalition of experts from leading institutions. As part of SCALE, we are producing a series of ancillary documents that provide more detail on specific topics.

The lead contributors for this document include Dr Ali Cheema (LUMS), Dr Jishnu Das (Georgetown / Centre for Policy Research), Dr Adnan Q. Khan (LSE), Dr Asim I. Khwaja (Harvard), Dr Farhan Majid (University of Georgia), Dr Aryn A. Malik (Yale), Dr Tyler McCormick (University of Washington), Dr Saad B. Omer (Yale), Omer Qasim (CERP), Maroof Syed (CERP).

For comments, please reach out to us at: covidrapidresponse@cerp.org.pk.



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#).

Step 1: Define the lowest sampling grids “S” for City/Districts and conduct data diagnosis

- ‘S’ should be as small as minimally feasible (200+ households)
- Shape files (geo-referenced boundaries) will be required for ‘S’
- Ideal ‘S’ are census block/mauzas as that will allow for overlaying census information (population density, distribution of age etc.)
- Alternative options: If census blocks/mauzas not available, then the following can also be used:
 - ◆ UCs
 - ◆ Geo-referenced grids
 - ◆ Creating geo-referenced polygons using cell phone towers
- Conduct a data diagnosis to assess the usability of the data¹ that is being collected and understand its workflow

Inputs	Outputs
<ol style="list-style-type: none"> 1. Shape files of ‘S’ grids 2. Census data on ‘S’ grids 3. Geo-locations of isolation centers, quarantine facilities hospitals and labs conducting PCR Tests 4. Data Diagnosis 	N/A

Step 2a: Contact Tracing Testing

- Identifying contacts of known positive cases - there are two ways to obtain this:
 - ◆ First preference: Through Call Detail Records (CDR) and/or Cell Phone Tower Data. The steps involved would be as follows:
 - Provide telecom operators with cell phone numbers of infected patients
 - Identify exposed people - identify individuals that were in the same physical location (same tower) during the span of 15 minutes and/or 30 minutes overlapping window in the past 14 days (ideally 3 weeks). If this number is not too large, use it. If it is too large, narrow it down further to exposed *contacts*
 - Define contact list - ask telecom operators to identify all numbers called by the infected person’s number for at least 14 days (ideally 3 weeks) prior to the person’s infection start date

¹ Health, location, and contact tracing data is personal data. Information collectors should ensure that the data collected is accurate and secure. The integrity of data can be improved by cross-referencing it with reputable databases and by providing access for the consumer to verify it. Information collectors can keep data secure by protecting against both internal and external security threats. They can limit access within their organization to only necessary employees to protect against internal threats, and they can use encryption and other computer-based security systems to stop outside threats.

- Narrow down list of exposed *people* to exposed *contacts* by excluding everyone from exposed people list who are not in contact list
 - ◆ Second preference: Actual contact tracing done through in-person surveys/phone calls/robocalls (will capture contact details and work/home location of contacts)
- Test & administer basic in-field survey (survey captures basic demographics useful in refining disease model) to contacts and geotag where they live/work and assign to appropriate 'S' grid

Inputs	Outputs
1. Individual level data on contact tracing (CDR or captured through surveys/robocalls) 2. Home and work geo-locations of traced contacts	1. Overlaying the data from inputs on appropriate 'S' grids

Step 2b: Exposed Frontline Worker Testing

- Identify all exposed frontline workers. These are workers (medical staff, caretakers, police and government officials, retailers, etc.) who are/have been in direct contact with infected people (with 10 minutes or more of exposure)
- Test & administer in-field for contacts and geotag where they (i) work and (ii) live and assign to appropriate 'S' grids

Inputs	Outputs
1. Individual level data on exposed front line workers 2. Home and work geo-location of frontline workers	1. Overlaying the data from inputs on appropriate 'S' grids

Step 3: Testing in 'S' grids

- Prioritize PCR testing in 'S' grids of the following types - start testing in Category 1 first and then move to Categories 2 and 3:
- ◆ Category 1
 - 'S' grids that have no infection detected yet but are next to a grid which has an infection
 - 'S' grids that have a high number of frontline workers working or living in them
 - 'S' grids with people who have high mobility/travel/connectivity
 - ◆ Category 2:

- 'S' grids that have high population density
- 'S' grids with large fraction of high health risk people (elderly etc)
- ◆ Category 3:
 - 3a: all remaining uninfected grids
 - 3b: all grids that already have been detected with high infection (PCR Testing in such areas is not as informative but these grids are top priority for seroprevalence testing)

→ Conduct the following types of prevalence testing:

◆ **Screening and Testing:**

- Ideal - but may be in limited supply; Instead of randomly testing people in 'S' grid, test those who are High Infection Susceptibility (i.e. people who have higher likely of infection) since this gives a more efficient way of testing (i.e. will need to use less tests to detect infection). These can be defined as people with (i) high number of physical contacts/interactions and/or (ii) high mobility
 - For PCR Testing, pool 10 or 20 tests each depending on what you have been instructed to do. Follow Pooled Testing protocol. Make sure everyone is also administered an in-field basic survey
- **Preferable method is using CDR data:**
 - High physical contacts: The following complementary approaches can be used
 - ◆ run algorithm for each phone number by seeing in one week (can use the most recent full week - make sure weekend and weekdays are included) what is the total number of OTHER phone numbers that were in the same physical location as they were in an overlapping 30 minute time window - list (phone no & location) of the top 5% percentile in this measure
 - ◆ run algorithms on CDR data to identify high contact nodes who are in contact with more people/links (restrict these to those within the same city and also see which of their edges are more active for further contact tracing)
 - High mobility folk: These are people who move around a lot (so could be more likely to get exposed) - run algorithm to count how many unique towers (i.e. different locations) the person has been at - can refine this over time to maybe weigh locations that are hot spots - provide list (phone & location) of the top 5% percentile in this measure
 - Build profiles of high contact/mobility individuals by combining CDR data with administrative records
- **If CDR data not available, then define High physical contacts/mobility by:**
 - Randomly arriving at the GPS pin location provided in 'S' Grid - Test & Survey member of household who has highest contact/travel in past 14 days (may not be owner but could be

domestic help who does regular shopping etc.) AND ask the household head who is a person with high contacts nearby OR who is their friend nearby (people who are identified as friends will likely have higher connectivity) and then go and Test & Survey them

◆ **Sentinel Surveillance PCR Testing**

- Identify main 2-3 sewage collection points in ‘S’ Grid and collect sample by using protocols similar to those for polio environmental surveillance through sewage sampling
 - This is an experimental approach and will require a pilot before being rolled out on a broader scale. Areas where grid maps of sewerage are available would be required
- Take samples and note the GPS location where the sample was taken

◆ **Phone-based Syndromic Surveys**

- Call randomly selected numbers (if possible prioritize numbers that show high mobility or physical contact in past 2-3 weeks) and administer survey of symptoms on the phone (this can also be done through robo-calls)

Inputs	Outputs
<ol style="list-style-type: none"> 1. Categorize ‘S’ Grids <ol style="list-style-type: none"> a. Identify high mobility and high contact individuals to administer appropriate tests 2. Contact details and biographic data on high mobility/high contact individuals 	<ol style="list-style-type: none"> 1. COVID19 detection PCR tests and field survey on people with high mobility and high no of contacts 2. Sentinel Surveillance by testing water samples and recording GPS location where the sample was collected from 3. Syndromic Surveys through Robocalls and Phone CATI 4. Positive cases overlaid on respective ‘S’ grids 5. Action items for stakeholders to take in respective ‘S’ grids <ol style="list-style-type: none"> a. SOPs for dealing with Positive cases (option for isolation at home)

Step 4: Follow-up socio-economic and broader health surveys

- All individuals tested are administered short phone surveys that will capture their socio-economic as well as broader health situation. This will be used to see what other costs they are incurring and what help may be useful to provide to them

Inputs	Outputs
1. Contact details and relevant biographic data of all individuals tested	<ol style="list-style-type: none">1. CATI phone surveys to capture socio-economic and broader health concerns2. Overlay the information on appropriate 'S' grids3. Devising action items for the stakeholders in 'S' grids