Project Baseline
by _verily

Baseline Site Selection Tool Presentation

AKA: "BSST" when appropriate

Baseline Site Selection Tool: Putting COVID models to work planning vaccine trials

Disclaimer

THIS CONTENT IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS CONTENT, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

The Challenge

COVID vaccine trials are designed around a simple idea:
- Recruit a diverse group of volunteers, ideally reflecting the people most at risk,
  centered around multiple medical centers which can run a trial
- Split the volunteers into placebo and vaccinated groups
- Wait to see how many placebo cases get COVID vs the vaccinated group
- ⇒ If many (say 150 of 15k) placebo cases get COVID, and very few of the vaccinated,
  you might have an effective vaccine (safety, how long immunity lasts are separate)

Trial planners need significant lead time to pick and prepare sites, but even the best models cannot make reliable localized predictions beyond 3-4 weeks.
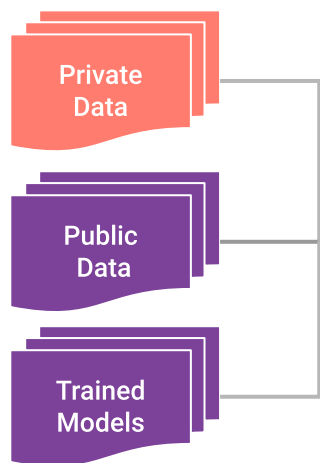
So how do we pick trial sites? Trust our favorite model anyway? Pick at random?
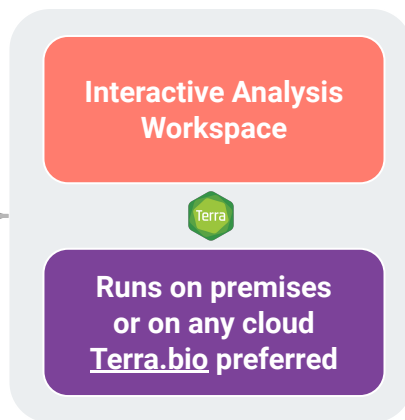
Vaccine trial site selection

Working with vaccine makers, we've developed a suite of tools for vaccine trial planning:

- Take as input any localized ensemble of models and supporting demographic data

- Encode trial rules, prospective sites, and expected recruitment data to dynamically calculate expected numbers and demographic breakdown of patients in the placebo group who may get COVID. (The number of placebo vs vaccinated cases is a critical determinant of the efficacy of the vaccine)

- Interactive interface to allow trial planners to explore scenarios and react to evolving trial conditions e.g. to predict consequences of turning up a new site, or mitigation strategies for low recruitment at particular sites

- Optimization algorithms to calculate recommended sites to reach a faster trial outcome or a more representative participant pool
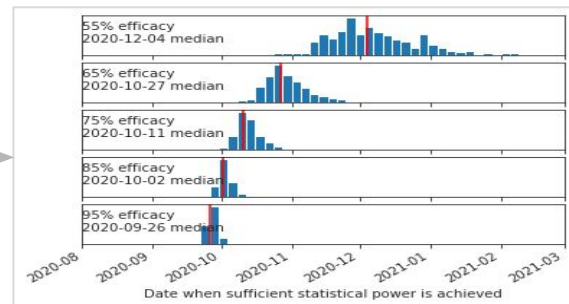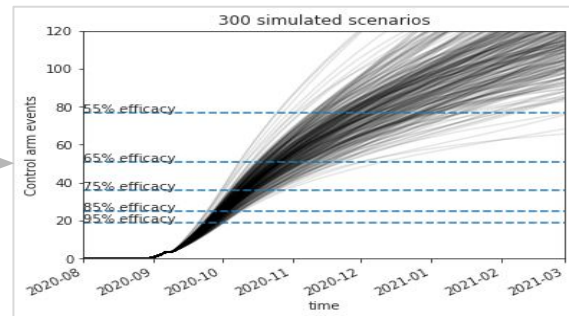
Baseline Site Selection Tool

Private Data

Public Data

Trained Models

Interactive Analysis Workspace

Terra

Runs on premises or on any cloud Terra.bio preferred

Add private data to your own curated public data and models.

Explore different scenarios and assumptions.

View projections of enrollment, events, and statistical significance.

How we predict/quantify trial resolution date

1. Encode detailed trial rules

2. Project recruitment rate at each trial site

3. Choose models of future incidence, *locally* at each site

4. Combine recruitment + incidence to forecast control arm cases

5. Predict distribution of end-times

6. Predict demographic distribution of the enrolled cohort

Trial end-dates: example "back of the envelope"

- At **70% vaccine efficacy**, need to reach **50 events** in the control arm

- Annualized Incidence varies from **0.1% -> 10%**

- Assume **70%** of covid infections reach clinical endpoint for vaccine trial

- Assume **15k** participants in control arm

- Assume recruitment delay of **4 weeks**, and **observation delay of 4 weeks**

Average Incidence	Control Arm Cases/yr	Time to reach 50 cases	Delays	Total Trial Time
0.1 %	10.5	5 years	8 weeks	260 weeks
10 %	1050	2.5 weeks	8 weeks	10.5 weeks

Illustrative example: Recruitment

Recruited participants over time
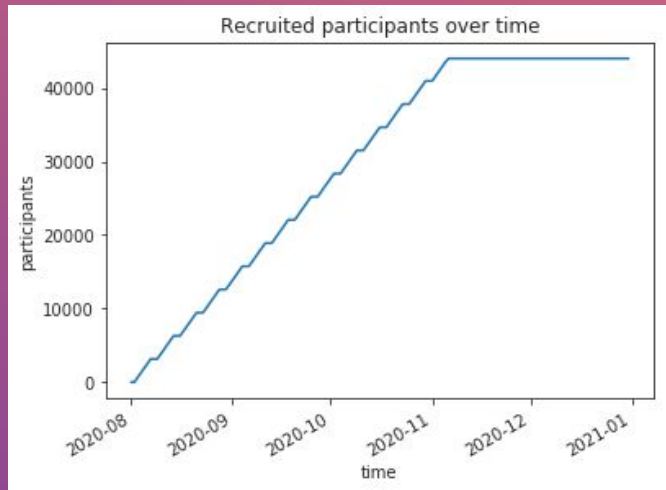
To demonstrate how this works, we show an illustrative example.

64 sites obtained from clinicaltrials.gov.
One dropped because missing forecast.

We assume trial starts Aug 1, with sites recruiting at 50 participants/week/site, recruiting Mon-Fri each week. Participants are assumed uniformly sampled from the county.

44k participant cap.

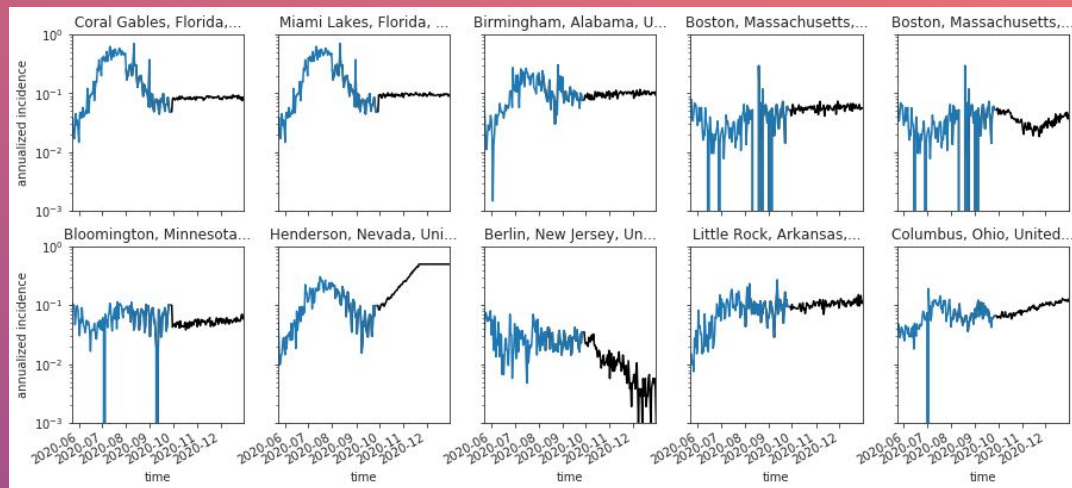For recruitment, tool is set up to handle additional complexities, including:
- Different time varying site capacities
- Enrichment for minority, elderly, comorbidities.

Illustrative example:
Control arm events from forecasts

Blue = groundtruth; black = forecast

A simulated scenario

observation delay (4 weeks)
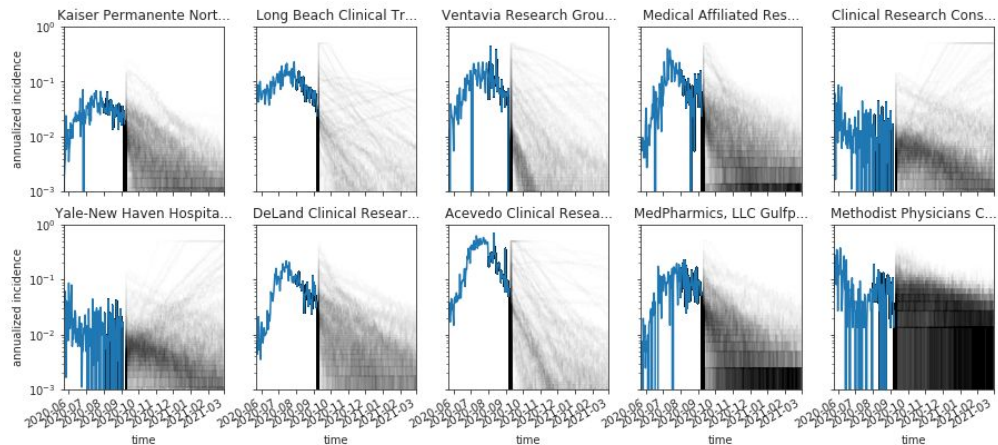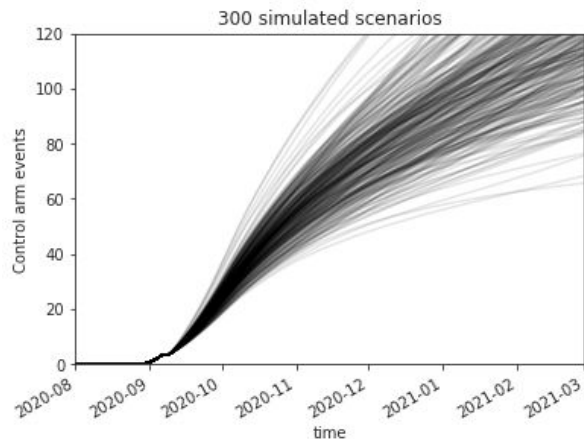
Slow ramp-up because of rolling recruitment
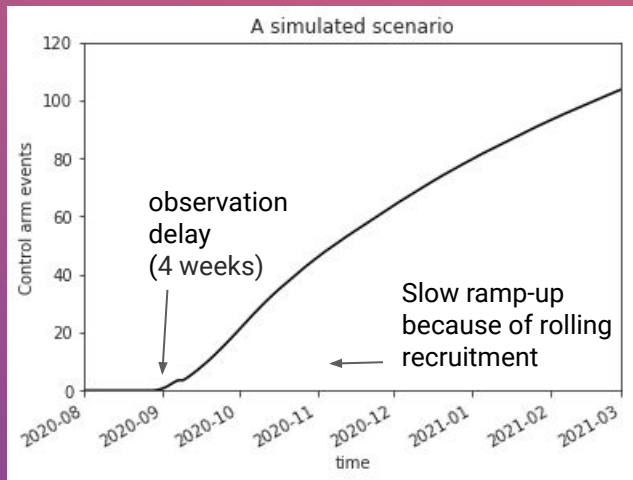
Each forecast of control arm events integrates incidence forecasts from 63 different sites.

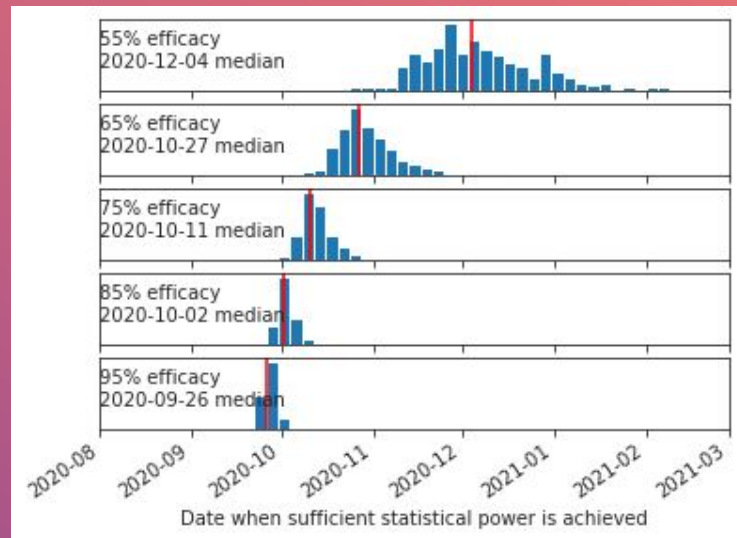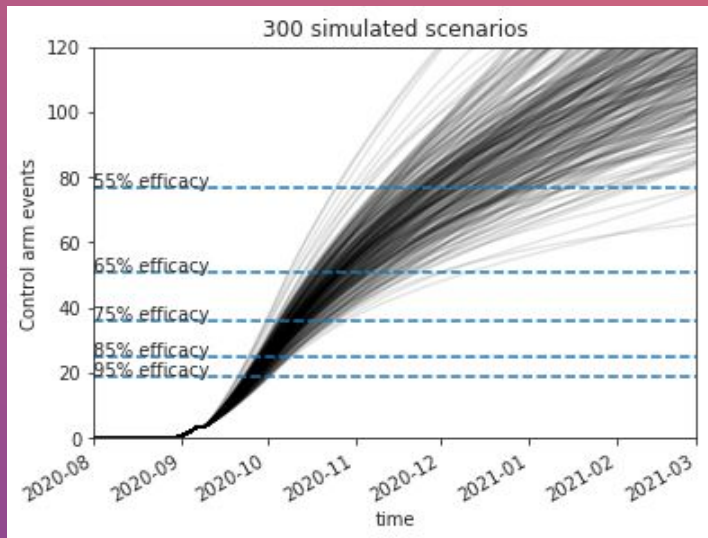Underlying incidence forecast (10 of 63 sites shown here)

Illustrative example:
Control arm events from forecasts

Underlying incidence forecast (10 of 144 sites shown here)

Uncertainty in forecast leads to many possible forecasts that are consistent with the model.

Forecasting scenarios predict end-date distribution

End date depends on vaccine efficacy, as efficacy determines the needed number of events in the control arm.

View the impact of a single site on the whole trial
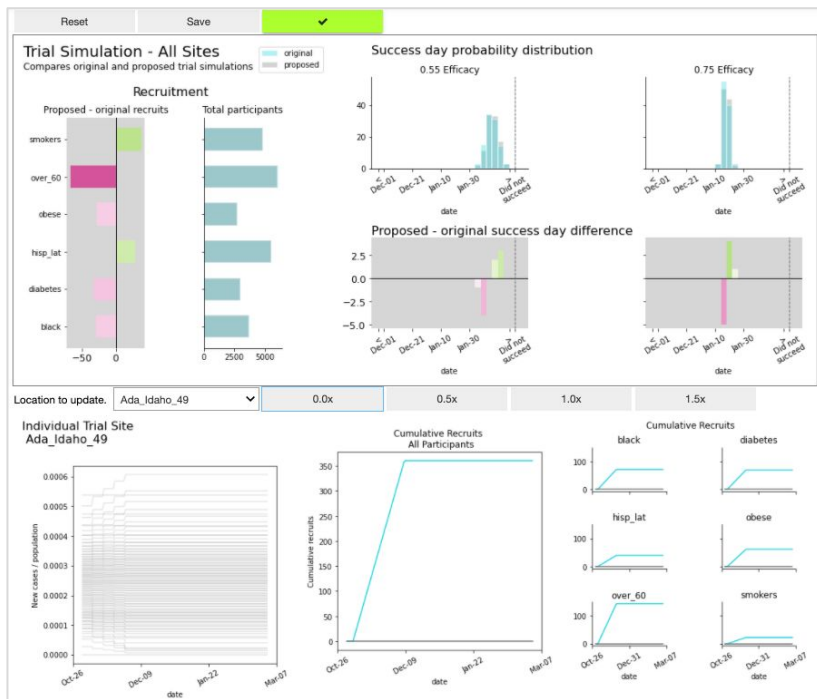
Individual sites can be adjusted, trial simulation recomputes in real time.

Baseline Site Selection Tool allows trial planners to view the impact of a single site on the whole trial

Sortable table allows trial planners to identify key sites.

'population_fraction' Sort by: population ∨												
location	SR1	SR2	frac_cap	population	black	diabetes	hisp_lat	obese	over_60	smokers	proposed_events	original_events
Los_Angeles_California_19	California	Los Angeles County	1.0x	10,103,711	24%	12%	11%	14%	21%	13%	0.30	0.30
Los_Angeles_California_15	California	Los Angeles County	1.0x	10,103,711	14%	18%	35%	4%	20%	25%	0.19	0.19
Los_Angeles_California_18	California	Los Angeles County	1.0x	10,103,711	5%	19%	35%	15%	38%	28%	0.01	0.01
Los_Angeles_California_17	California	Los Angeles County	1.0x	10,103,711	3%	5%	23%	3%	23%	12%	0.07	0.07

Project Baseline by verily
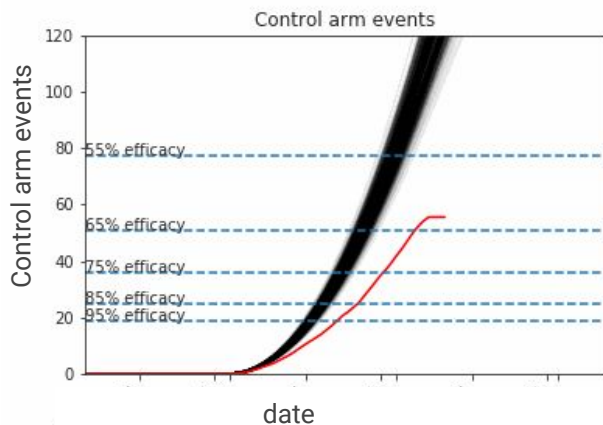
Optimization:
Changing many sites

Site planning can be turned into a mathematical optimization problem:

Given the trial planner specifications (site lists, recruitment schedules) the tool solves an optimization problem to achieve a desired potential outcome (minimize time to success, diversity targets, etc.)
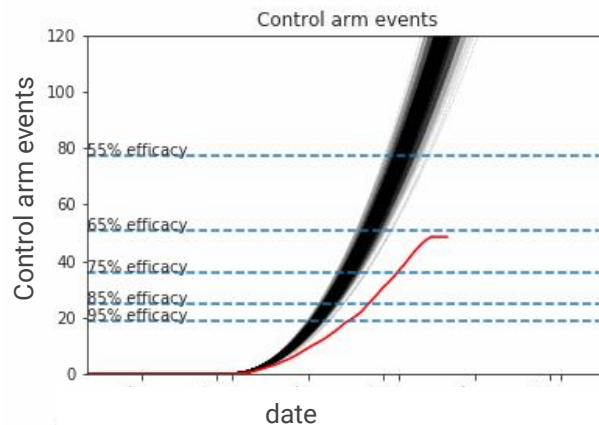
The resulting site plan reflects the logical consequence of the trial planners beliefs about what is likely to happen in the future (eg recruitment and incidence models)

Project Baseline by <u>verily</u>

Site optimization can be carried out "on the fly"

All sites activated

Activation optimized every week

Important links

Public repo of localized epi data from counties/cities around the world,

with complete tracking of provenance and open source code for processing.

Free to use, and may be updated from time to time:

https://github.com/GoogleCloudPlatform/covid-19-open-data

Metis code open sourced and available for download:

https://github.com/verilylifesciences/metis

Ready to run on Terra:

https://app.terra.bio/#workspaces/verily-metis/Metis-toolkit-for-vaccine-trial-planning

Project Baseline by verily