COVID-19 Vaccine Predictions: User Guide

We have developed an interactive web tool that enables users to generate predictions of when a COVID-19 vaccine will be available and accessible. Incorporating data from numerous health experts, this tool takes into account both the process of research and development, or R&D (the approval process whereby a body defined by the World Health Organization as a "stringent regulator" approves the vaccine), and manufacturing. The tool provides global estimates for policymakers, researchers, journalists, and others—and therefore does not provide estimates for specific countries. The goal of our tool is to help users understand the COVID-19 vaccine portfolio, to gauge whether the portfolio for COVID-19 vaccines is sufficiently diversified, and to generate predictions of how long it will take to approve and manufacture vaccines. We released preliminary results in a recent paper, and will be updating the data for this tool at regular intervals.

What you can learn from the tool

Using the tool, you can learn about whether the vaccine portfolio is sufficiently diversified, and can generate predictions for:

- When a COVID-19 vaccine will be approved by a stringent regulator.
- When everyone in the world could have access to a COVID-19 vaccine.
- When there will be enough vaccines available to vaccinate all health workers.
- The likelihood that a vaccine purchased by my country will be approved, and when.

How to use the tool

Our tool has two main settings, "quick set parameters" and "full parameters." Below we outline the options available with these two settings. Users can generate information of how likely a vaccine is to pass through different phases of development, the type of vaccines that are most likely to succeed, and the importance of funding and platform type. This is combined with overall portfolio data to estimate the likelihood of a vaccine being approved by a certain date.

Users can adjust information on timelines for scaling up production of vaccines, as well as global manufacturing capacity, and generate predictions for when there will be enough vaccines to vaccinate people globally.

The "Quick set" setting

Users can:

- Easily adjust timelines and probabilities of success from the custom data on the status of vaccine candidates around the world that we supply. We will update this data at regular intervals, users can then select "optimistic" or "pessimistic" settings. The optimistic or pessimistic inputs alter all probability of success and or timeline by the multiple of the number the user inputs. See below to the data section for more information.

The "Full parameters" setting

Users can customize the parameters, to have more control over the model. These include:

- Changing the average trial times and the probability of success for reaching the next phase of the trial.
- Indicating the importance of funding and platform type.
- Moderating according to the total number of clinical trials that can be run.
- Incorporating the odds that an unforeseen event will slow down all clinical trials.
- Including scale up time for manufacturing.
- Taking into account global manufacturing capacity

For many of the inputs, we collect a most likely, best case and worst-case input, in each simulation the model randomly choose a number between the best and worst case scenario, using a triangle distribution. Any number between the best- and worst-case scenario can be chosen, but those closer to the most likely input will be selected more often.

What our tool cannot do

- Give anything other than *predictions* for when a COVID-19 vaccine might be ready.
- Outline how effective a vaccine will be.
- Say exactly which vaccine will be approved or manufactured first.
- Provide country-specific information regarding either the R&D process, or the manufacturing process.
- Inform when your government/health provider will procure enough vaccines for their constituents.
- Outline how a specific country will roll out vaccine coverage.

Method for our tool

Our model uses a Monte Carlo simulation, which applies probabilistic rules repeatedly, to simulate many future outcomes and understand how likely something is to occur. If, for example you want to know how likely it is that you roll two dice and both landed on six, you could create a model that randomly chooses a number between one and six, twice. Each simulation would see these two random numbers selected and would map a plausible outcome for what could happen in the real world. While no single simulation would tell you how likely getting two sixes is, if you were to run this model 10,000 times and aggregate the results the proportion of times that the model randomly chooses two sixes will be very close to the real life chance of rolling a double six (2.78 percent).

There are over 200 vaccine candidates currently competing to fight COVID-19, and we have collected data on all of these to feed into our tool. When we run our model, each simulation should similarly create plausible timelines for different vaccines. As a user, it's important to

remember that the best results come only when the timelines are aggregated. This is when it is able to provide the clearest overview.. Running our model several hundred times should allow the user to understand timelines and the diversity within the COVID-19 portfolio. Because of the inherent randomness of the simulation, each time the model is run the outputs are likely to differ very slightly even when the inputs are the same. As a user, this might mean that you plug in the same information, but the probability of getting a vaccine in the next year varies between 99% in one run and 99.2% in the next. You can find the full methodology here.

Results

Our model then gives results based on how long it will take to approve a vaccine, the number of vaccines that are likely to be approved, and how long it will take until we have enough vaccines to reach different target groups.

What data has been plugged into our tool?

All our default inputs come from a mix of expert interviews and our internal expertise. The easiest way to tweak the model is to adjust the timeline and or probability of success for all inputs, using our quick set parameters. Because each adjustment changes multiple inputs, the impact is several-fold larger than the adjustment.

The tool runs on two types of data: objective and subjective inputs. These collect factual information on all of the vaccines in the portfolio such as their stage of development, funding category, and platform. We intend to update this information once a month, and users can download and edited data in this table to add or remove vaccines, or change vaccines stage of development, funding type or platform. You can then run the analysis on this customised portfolio (though for this to work the format of the original table needs to be strictly adhered to).

The second set of data is subjective inputs. These give the user the ability to change information on how likely a given vaccine is to get through different stages of development, how important factors such as funding are, and how long manufacturing scale up will take. Users of the online tool can change any of these parameters in the model to get different results. We are exploring ways of updating this information too, by redoing interviews every few months.

Finally, all of our analysis takes place using a master data sheet, where we have attempted to track all vaccine candidates, for this we use a large range of data sources, including the London School of Hygiene and Tropical Medicines vaccine tracker. Users can download the CSV file that this information is stored in, remove, add in or edit the information on vaccines in the portfolio and then upload it back to run the model on their own portfolio. It is critical that the data uploaded follows our format exactly so that the information can be read by the model.

Using/sharing our work

All of our work is open source, you can find and use our code here. Please be sure to cite our work when using it elsewhere and get in touch if you've any questions.