Vaccine Effectiveness (VE) Model

Poster Title: An Open-Source Probabilistic Modeling Framework for Personal Pandemic Risk Assessment

By Quantum Risk Analytics, Inc.: https://pandemonium.dev/

Motivation

Vaccine Effectiveness (VE) varies with time since vaccination, disease outcome that is protected against, vaccine type(s), and number of vaccinations. It is useful to model VE since data on VE is incomplete and is heavily marred by uncertainties. Furthermore, people have a need to make vaccination decisions in real time, before meta-analyses are available for new vaccines or variants.

Probabilistic Programming for Modeling Vaccine Effectiveness

VE and its temporal variation are considerably uncertain due to numerous factors, including the coarse temporal resolution of VE data, the dearth of VE data in the first two weeks after the most recent vaccination, the receipt of literature-undocumented combinations of vaccines, and "study effects" stemming from differences in study design and/or in the studied populations. To best incorporate the available VE data, including its uncertainties, we apply probabilistic programming in the creation of our vaccine effectiveness model. Instead of treating model parameters as fixed quantities, probabilistic programming languages treat these parameters as random variables whose initial-guess or "prior" probability distributions are revised in light of the observed data into "posterior" or final distributions.

Multiple stages are involved in the probabilistic program we developed.

- **Stage I:** As stated before, VE (ϵ) varies with time (t) and with other features. We model $\epsilon(t)$ for the different scenarios using a family of parabolic-exponential functions[1], which are determined by 5 parameters:
  - ϵ_{max}: $the\ maximum\ of\ VE$
  - $t_{\epsilon max}$: $the\ time\ after\ last\ dose\ at\ which\ \epsilon_{max}\ is\ reached$
  - r: $the\ decay\ rate\ of\ VE$
  - ϵ_{final}: $the\ final\ value\ of\ VE\ as\ t\ approaches\ \infty$
  - ϵ_{init}: $the\ initial\ value\ of\ VE$
  - ➢ Here, we assume that VE starts from ϵ_{init} at the time of the last dose, then reaches ϵ_{max} at $t_{\epsilon max}$, and finally decrease to ϵ_{final} with decay rate r.

[1] This function is typically an upside-down parabola for small t, and then becomes an exponential decay curve (with possibly non-zero asymptote) for large t. In rare exceptions, the curve behaves differently: a vaccine may be counterproductive, and thus the parabola curves in the opposite direction and potentially has VE < 0 where the extremum is actually a minimum. Occasionally, it can have an inflection point where the derivative vanishes as the 2 halve parabolae can curve in opposite directions. (Since rare, the "max" subscript is used for convenience, recognizing it is occasionally a misnomer.)

- ➢ Also, in this stage, we specify how our modeled VE curves relate to the reported observations by making reasonable assumptions about the structure of the distributions from which the observations come.

- ❖ **Stage II:** Here, we further specify that the stage I parameters derive from combinations of more fundamental stage II parameters. These represent global or "baseline" effects, single effects, and paired interaction effects. The single and paired interaction effects are those imparted from the different features that may be present (shown in the table below) and from their paired interactions. There is also a separate age effect.
- ❖ **Stage III:** The prior distributions for parameters in stage II are parameterized using stage III parameters. These latter parameters are themselves random variables, whose prior distributions we specify here using fixed stage III hyperparameters.

After training, our model outputs the parameters that generate the VE curves that holistically fit the observed VE data the best. Also, it can predict continuous VE curves for unobserved vaccine combinations and times since vaccination.

The multi-stage probabilistic model above is coded using Python and Pyro, a powerful probabilistic programming framework built on top of PyTorch that allows for flexible probabilistic modeling and inference.

Data

The observed VE data are collected from 16 different studies including the systematic review of Higdon et al. (2022)[2] and three studies within (Stowe et al. 2022, Ranzani et al. 2022, Florentino et al. 2022) for VE data from the first two weeks after vaccinations. There are 31 features (see below). After preprocessing, we end up with a total of 341 observed VE values and 197 observed combinations of features. This data is randomly split into a testing and a training set with a 1:4 ratio. The 31 features, including the combination of vaccines, variant types, dosage number, etc., are listed in the table below. Categorical variables are represented using "one-hot encoding", such that only one intra-category member at a time is allowed.

[2] Higdon et al., (2022): https://doi.org/10.1016/S1473-3099(22)00409-1

Column ID	True Meaning
0	Vaccine: 2-Dose Vaccine: Pfizer
1	Vaccine: 2-Dose Vaccine: Moderna
2	Vaccine: 2-Dose Vaccine: Sinovac
3	Vaccine: 2-Dose Vaccine: AstraZeneca
4	Vaccine: 1-Dose Vaccine: Janssen
5	Vaccine: 1-Dose Vaccine: None
6	Vac Type: mRNA
7	Vac Type: WV
8	Vac Type: ChAdOx1
9	Vac Type: Ad26
10	Vac MFGR: Pfizer
11	Vac MFGR: Moderna
12	Vac MFGR: Sinovac
13	Vac MFGR: AstraZeneca
14	Vac MFGR: Janssen
15	Vac MFGR: None
16	Vac Stage: Partial
17	Vac Stage: Booster
18	Vac Stage: Primary
19	Eff Type: Preventing Disease
20	Eff Type: Preventing Infection
21	Eff Type: Preventing Severe Disease
22	Variant: Omicron
23	Other Demo: None
24	Prior Vaccine: None
25	Prior Vaccine: 2-Dose Vaccine: Moderna
26	Prior Vaccine: 2-Dose Vaccine: Sinovac
27	Prior Vaccine: 2-Dose Vaccine: AstraZeneca
28	Prior Vaccine: 2-Dose Vaccine: Pfizer
29	Heterogeneous Vax: True
30	Bivalent: True
31	Severe: True