Comments on "Draft Master Protocol Assessment of Risk of Safety Outcomes Following COVID-19 Vaccination"

Observational Health Data Sciences and Informatics (OHDSI)

Contact:
George Hripcsak, MD, MS
Columbia University Irving Medical Center
OHDSI Coordinating Center
622 W 168th St PH20
New York, NY 10032
gh13@cumc.columbia.edu

Primary commenters: Patrick Ryan, Martijn Schuemie, Daniel Prieto-Alhambra, George Hripcsak

We applaud FDA CBER for releasing its vaccine safety risk protocol for public comment. The analyses are carefully designed and account well for issues related to vaccine safety.

Negative controls to detect and control residual bias

The protocol specifies multiple alternative study designs that may be applied for COVID-19 vaccine safety surveillance, including self-controlled case series, self-controlled risk interval, and comparative cohort designs. Each of these designs aims to produce an effect estimate that presents the strength of a casual association. Each analysis will produce a series of statistics, such as the point estimate, confidence interval, and p-value. As Table 3 in the protocol highlights, each design carries with it a set of limitations that can potentially bias that effect estimate and corresponding statistics. Therefore, regardless of the design selected for any given vaccine-outcome question, quantifying the magnitude of systematic error should be required, and re-calibrating the statistics to account for this error should be strongly recommended to avoid misinterpretation of results. OHDSI's EUMAEUS study (https://ohdsi-studies.github.io/Eumaeus/Protocol.html) outlines how using a large sample of **negative control outcomes** (and generating synthetic positive controls based on those negative controls) can be used to estimate systematic error for a given analysis. Our preliminary results, based on historical performance of methods from other vaccines (H1N1, seasonal influenza, zoster, and HPV), suggest that all of these designs demonstrate error that substantially deviates from nominal expectations of unbiased estimators (e.g., the coverage probability of 95% confidence intervals is much less than 95%, the Type 1 error rate using a threshold of p=0.05 is much higher than 5%). These preliminary results are consistent with prior published methodological research [Schuemie et al., Harvard Data Science Review 2020 https://pubmed.ncbi.nlm.nih.gov/33367288/] which found uncalibrated estimates for the same designs subject to substantial systematic error when applied to known drug-outcome pairs. Empirical calibration [Schuemie et al. PNAS 2018 https://pubmed.ncbi.nlm.nih.gov/29531023/] can be used to incorporate the observed systematic error into effect estimates, allowing for empirical verification that nominal operating characteristics (e.g., 95% coverage probability, 5% Type I error rate) are maintained. The OHDSI open-source tools can be used to support the FDA team in the identification and evaluation of candidate negative control outcomes upon request; the list of negative controls used in EUMAEUS may serve as a useful starting point.

Exposures

We appreciate the recognition that COVID-19 vaccination codes may change over time and need to be reviewed. To support this effort, the Standardized Vocabularies maintained by the OHDSI community are regularly updated based on source vocabulary releases. At present time, in addition to containing the necessary CPT-4 and NDC codes for the Pfizer, Moderna, Janssen, and AstraZeneca COVID-19 vaccines, the OHDSI vocabularies also contains corresponding CVX codes and mappings to RxNorm concepts. The OHDSI Standardized Vocabularies (available at: athena.ohdsi.org) may become a useful resource for FDA, particularly if linkage to state immunization registries or use of electronic health record data are considered for future studies.

Outcomes

While the master protocol provides a list of pre-specified Adverse Events of Special Interest (AESI) that will be considered as study outcomes, we believe the guidance included in this protocol has broader applicability to other potential safety outcomes that may emerge as we learn more about the effects of COVID-19 vaccines. We note that each outcome requires phenotype development and evaluation, and are subject to measurement error with imperfect sensitivity and specificity and index date misspecification. We applaud FDA for sharing their suggested ICD-10-CM diagnosis code lists to be used for the pre-specified outcomes. Given that choices in phenotype algorithms, such as codesets, care setting, and clean window length, can potentially have a substantial implication on the magnitude of measurement error associated with the outcome, and the recognition that medical record review may not be feasible, we encourage sensitivity analyses to be performed using alternative phenotype definitions to evaluate the robustness of findings. OHDSI has developed alternative phenotype definitions for the AESIs, and have characterized the incidence rates using these definitions as part of its ongoing Covid19AESIIncidenceCharacterization study (https://github.com/ohdsi-studies/Covid19VaccineAesiIncidenceCharacterization). The first publication from this work is currently under review and available as pre-print (Li et al, https://www.medrxiv.org/content/10.1101/2021.03.25.21254315v2). We recommend considering to use OHDSI's CohortDiagnostics tool to perform phenotype evaluation as an initial step of study diagnostics prior to unblinding results.

Self-controlled Designs

We support the recommendation to consider self-controlled case series designs. Prior empirical research has suggested that self-controlled case series has greater discrimination than comparator cohort or case-control designs [Ryan et al, Drug Safety 2013, https://pubmed.ncbi.nlm.nih.gov/24166231/; Schuemie et al. Harvard Data Science Review 2020 https://pubmed.ncbi.nlm.nih.gov/33367288/]. As generically defined, a person can contribute to an estimate within a SCCS design if there is at least one outcome event, at least one day of 'at-risk' time, and at least one day of 'unexposed' time. In this regard, a person serves as her own control by enabling a comparison of the incidence rate during the at-risk time with the incidence rate during the unexposed time. Statistical power is determined both by the number of qualifying individuals (cases), as well as the duration of 'at-risk' and 'unexposed' time within each individual.

We note that the protocol suggests limiting the SCCS analyses to control 'unexposed' time that follows the post-vaccination 'at-risk' period, and not considering **pre-exposure time** as part of the 'unexposed' intervals. This decision has two objective consequences: 1) it will limit the statistical power of the SCCS analysis, by reducing the amount of 'unexposed' time that is used to estimate the control incidence rate within each individual, and 2) it will reduce the timeliness in detection as patients will need to be observed for some duration longer than the 'at-risk' period before they could potentially qualify for the SCCS analysis (whereas, use of a pre-exposure control window would allow patients to qualify the analysis one day after vaccination). We recognize this design choice is intended to mitigate the risk of violating the SCCS model assumption that occurrence of an event does not influence subsequent exposure. It is unclear in this context what is the appropriate bias-variance tradeoff to make. Given that statistical power and time-to-detect are likely important considerations to FDA, we recommend further evaluation on this design choice. The EUMAEUS study will provide some empirical guidance on implications of alternative choices within the SCCS design. We should caution that the extent to which those findings will be generalizable to the current COVID-19 vaccine context is uncertain, since the negative control outcomes used in the methods evaluation are unlikely to have similar 'contra-indication' effect on vaccine exposure as is likely hypothesized for some of the AESIs for COVID-19 immunization. The impact of this choice likely will vary by outcome. For example, it seems conceivable that a patient experiencing a recent event of anaphylaxis could influence timing of vaccination (so the SCCS model assumption could be violated), but less plausible that appendicitis in the prior year is meaningfully impacting immunization.

One concern related to self-controlled designs is that they do not account for **anchoring** surrounding the vaccination. We define here anchoring as the tendency to see more outcomes near medical encounters, even for encounters not related to the outcome of interest. Our preliminary work in characterizing incidence of AESIs shows that for the shortest time at risk, 0-1 days, the effect can be huge, up to 100 times for anaphylaxis. For longer times at risk of 28 to 42 days, the effect can be up to 3 times for many outcomes. We do not yet know how patterns of healthcare utilization around COVID-19 vaccination will be observed. Influenza vaccines are often given during other medical visits, but COVID-19 vaccines may be given predominantly as part of non-medical events. We therefore believe it will be important to assess the degree to which anchoring is observed, potentially studying negative control outcomes that are not believed to be caused by vaccination and see if their rates rise after vaccination. Our preliminary work suggests anchoring on influenza vaccine administration or well visits has a smaller effect than anchoring on any type of visit.

We note that while we detect **seasonal effects**, they are much smaller than other effects observed from our AESI incidence characterization study. Only anaphylaxis has a significant effect, with higher rates (1.3 times) during summer. Others are within 10%.

Section 5.2.3 (Page 12, bullet (iii)) suggests that patients will be excluded if they have a AESI event during the outcome-specific **clean window** prior to exposure. It is likely a minor practical consideration, but we note that the SCCS design relies on time intervals, so it is possible to exclude the 'clean window' as 'immortal' since no event can be observed during this time; after that exclusion, as long as an individual still has at least one event, at least one day of 'at risk' time and at least one day of 'unexposed' time, then they remain qualified for the analysis and can contribute to the SCCS estimation.

Cohort Design

For controlling confounding, we suggest **large-scale propensity score** adjustment rather than the fixed covariates specified in the protocol. Prior research has shown that using all available pre-exposure covariates results in better performance than empirically selected covariates [Tian et al. International Journal of Epidemiology 2018 https://pubmed.ncbi.nlm.nih.gov/29939268/], and that manually selected covariates can be susceptible to residual confounding from the variables not included in the model [Weinstein et al. Drug Safety 2020 https://pubmed.ncbi.nlm.nih.gov/32500272/]. It ensures that measured confounders are included, does not induce increased bias from instruments and colliders, and may account for unmeasured confounders [Schuemie et al. JAMIA 2020 https://pubmed.ncbi.nlm.nih.gov/32827027/; Hripcsak et al. JAMA Internal Medicine 2020 https://pubmed.ncbi.nlm.nih.gov/32065600/]..

Results from our empirical evaluations suggest that **stratification and matching** are superior ways to balance samples than inverse probability treatment weighting (IPTW). The probable reason for this is that even when using stabilized weights, extreme weights can introduce bias, and trimming extreme weights can introduce further bias [Schuemie et al. Harvard Data Science Review 2020 https://pubmed.ncbi.nlm.nih.gov/33367288/].

We agree with the mention of **negative controls,** and we suggest that large numbers of controls should be employed to detect residual bias and provide the opportunity for calibration. These can be taken from our EUMAEUS vaccine safety experiment.

We note that the cohort design may similarly be subject to **anchoring**. We again suggest testing this also using negative controls.