Title:

ISPE-endorsed guidance in using electronic health records for comparative effectiveness research in COVID-19: opportunities and trade-offs

Authors: Grammati Sarri[1], Dimitri Bennett[2], Thomas Debray[3], Anouk Deruaz-Luyet[4], Montse Soriano Gabarró[5], Joan A. Largent[6], Xiaojuan Li[7], Wei Liu[8], Jennifer L. Lund[9], Daniela C. Moga[10], Mugdha Gokhale[9,11], Christopher T. Rentsch[12], Xuerong Wen[13], Chen Yanover[14], Yizhou Ye[15], Huifeng Yun[16], Andrew R. Zullo[17] and Kueiyu Joshua Lin[18]

Affiliations:
[1]RWE Strategy Lead, Visible Analytics Ltd, Oxford, UK

[2]Takeda Global Evidence and Outcomes, Takeda Pharmaceuticals USA, Inc., Cambridge, MA, USA and Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA

[3]Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, The Netherlands amd Smart Data Analysis and Statistics, Utrecht, The Netherlands

[4]Global Epidemiology and Real-World Evidence CoE, Corporate Medical Affairs, Boehringer Ingelheim International GmbH, Ingelheim-am-Rhein, Germany

[5]Bayer Partnerships and Integrated Evidence Generation Office, Integrated Evidence Generation & Business Innovation, Medical Affairs & Pharmacovigilance, Bayer AG, Berlin, Germany

[6]Real-World Solutions, IQVIA, Mission Viejo, CA, USA

[7]Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA

[8]Division of Epidemiology, Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

[9] Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[10]Department of Pharmacy Practice and Science, College of Pharmacy, University of Kentucky, Lexington, KY, USA

[11]Department of Epidemiology, Merck, West Point, PA, USA

[12]Faculty of Epidemiology and Population Health, Department of Non-communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK; Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA

[13]Health Outcomes, Pharmacy Practice, College of Pharmacy, University of Rhode Island, Kinston, RI, USA

[14]KI Research Institute, Kfar Malal, Israel

[15]Global Epidemiology, Pharmacovigilance and Patient Safety, AbbVie Inc., North Chicago, IL, USA

[16]Department of Epidemiology, University of Alabama at Birmingham, Birmingham, AL, USA

[17]Department of Health Services, Policy, and Practice, Brown University School of Public Health, Providence, RI, USA; Department of Epidemiology, Brown University School of Public Health, Providence, RI, USA; Center of Innovation in Long-Term Services and Supports, Providence Veterans Affairs Medical Center, Providence, RI, USA; and Department of Pharmacy, Lifespan-Rhode Island Hospital, Providence, RI, USA

[18]Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

Word count: 4894

Target Journal: BMJ Research Methods and Reporting or BMJ Analysis (https://www.bmj.com/about-bmj/resources-authors/article-types)

Abstract

As the scientific research community around the world fights tirelessly against the COVID-19 pandemic, the need for comparative effectiveness research (CER) on preventive and therapeutic interventions for COVID-19 is immense. Randomized controlled trials markedly underrepresent the frail and complex patients seen in routine care, and they do not typically have data on long-term treatment effects. The increasing availability of electronic health records (EHRs) for clinical research offers the opportunity to generate timely real-world evidence reflective of routine care for optimal management of COVID-19. However, there are many potential threats to the validity of CER based on EHR data that are not originally generated for research purposes. To ensure unbiased and robust results, we need high-quality healthcare databases, rigorous study designs, and proper implementation of appropriate statistical methods. We aimed to describe opportunities and challenges in EHR-based CER for COVID-19-related questions and to introduce best practices in pharmacoepidemiology to minimize potential biases. We structured our discussion into the following topics: 1) Study population identification based on exposure status; 2) Ascertainment of outcomes; 3) Common biases and potential solutions; and 4) Data operational challenges specific to COVID-19 CER using EHR. We provide structured guidance for the proper conduct and appraisal of drug and vaccine effectiveness and safety research using EHR data for the pandemic. This manuscript is endorsed by the International Society for Pharmacoepidemiology (ISPE)

Introduction

Coronavirus disease 2019 (COVID-19), caused by the Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2), emerged in 2019 as a major and urgent public health emergency worldwide (1). With the number of known cases and deaths rising exponentially (as of August 2021, there were over 213 million confirmed cases and over 4.4 million deaths) (2), public health control measures have focused on improving preventive strategies, including the introduction of non-pharmaceutical interventions (NPI), improving testing facilities, and restrictive social measures. Declared as a pandemic on March 11, 2020, COVID-19 unavoidably continues to place a huge strain on our activities of daily living while posing significant health, social, economic, and environmental challenges with major implications for the entire global community (3).

The scientific research community has tirelessly worked on the fight against the virus, the disease, and its complications. COVID-19 vaccines have been introduced (4) (5) (6), and therapies are being developed and studied (7). While large-scale randomized controlled trials (RCTs) remain pivotal to determine the drug efficacy and safety for regulatory approval purposes (8), many medications and vaccines have received emergency use authorizations (EUA) or fast-tracked approval, which resulted in their extensive use in wider populations (beyond the targeted populations recruited in RCTs). In addition, some primary concerns were raised about the use of RCTs to generate generalizable evidence in COVID-19. The selected participants in the RCTs may be unlikely to represent the frail and complex patients seen in routine care, and the short duration of RCTs did not allow for the generation of findings on long-term outcomes. Therefore, we need to also rely on non-randomized studies to generate real-world evidence (RWE) on the effectiveness and safety of preventive and therapeutic interventions. RWE can be generated through a range of applications like individualized prescribing, post-marketing surveillance, and can support policy or reimbursement decisions. Databases that reflect routine care delivery include electronic health records (EHRs), administrative health insurance claims, disease and product registries, and other non-research-specific data sources (for example, social media) (9). The advantages of RWE providing timely, generalizable evidence from a large, diverse group of patients are well recognized, particularly for a public health crisis like the COVID-19 pandemic (10) (11). To capitalize on the RWE offerings to inform the management of the COVID-19, we need access to high-quality healthcare data, rigorous study designs, and proper implementation of appropriate statistical methods to ensure unbiased and robust results (12).

EHRs have been widely used throughout the pandemic to generate evidence for risk stratification of patients, prognostic and risk factor identification, natural disease history investigation, and outcomes of interest that may be used in comparative effectiveness research (CER) (12), many of which are not routinely available in claim databases. In addition to its timely availability for research purposes, EHRs capture information on key factors for patient phenotyping and confounding adjustment, including inpatient medication use, disease, and patient characteristics (such as vital signs, laboratory test, and imaging results, smoking status, body mass index (BMI), and code status). EHRs also provide clinical notes and reports that are often important for validation studies (13). Furthermore, the growing availability and utilization of EHRs from different populations across healthcare systems with federated data networks and multi-database infrastructure involving multiple countries are also contributing to increasing opportunities for urgent and critical CER evidence on COVID-19 treatments (14). On the contrary, the availability of insurance claims and registry data is typically lagged, which poses a barrier to timely availability of such evidence.

However, controversies around the validity of findings from previous studies using real-world data (RWD) assessing the effectiveness of COVID-19 treatments have sparked skepticism around the use of

this evidence to inform clinical decision-making in the management of the pandemic (15). In addition to concerns with the veracity and appropriateness of specific data sources, some of these controversies arose from the inherent limitations of RWE that are not unique in COVID-19 CER, including data quality, missing data, and confounding bias (16). Some challenges are EHR-specific, such as misclassification of key information due to EHR data-discontinuity (17), converting unstructured free-text data into structured data (18), and harmonization of data across EHR systems in a multi-center study (19). Given the growing requirement for optimizing the potential of RWD for assessing the real-time effectiveness and safety of COVID-19 treatments, there is a compelling need for setting up clear guidance to ensure the results of these observational studies are reliable and valid for decision making. The purpose of this paper is to discuss the opportunities, unique challenges, and potential solutions when using EHR data for CER to inform the delivery of care in response to public health crises such as the COVID-19 pandemic. We hope this will equip the readers with a non-exhaustive list of tools to implement and interpret quality RWE using EHRs in a pandemic setting.
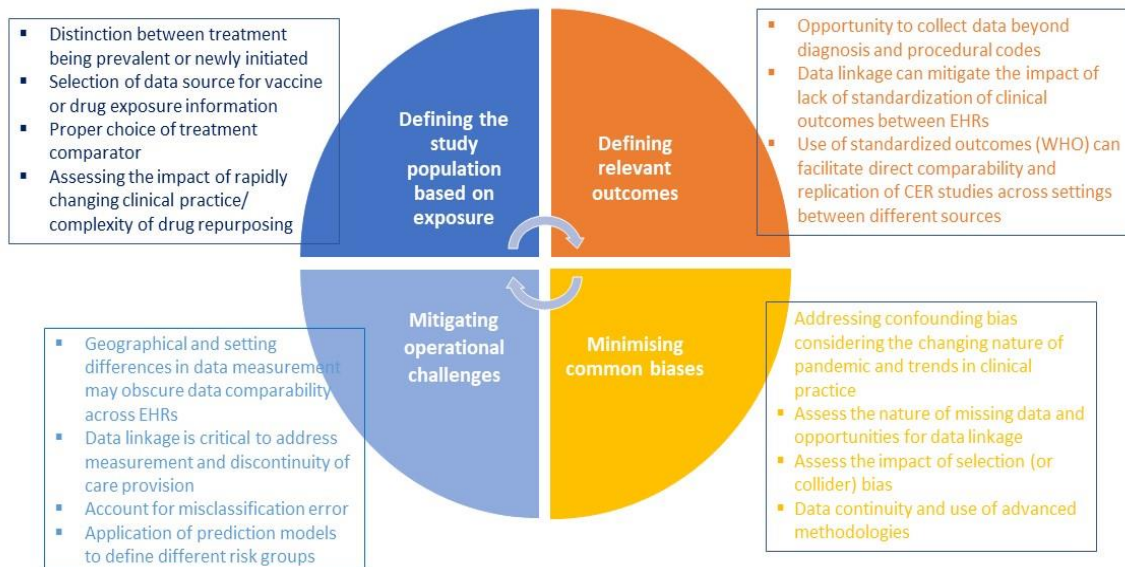
Methodological approach

This discussion paper has been formulated based on the information retrieved from a targeted literature review conducted in March 2021 (for full details, please see Appendix A) and the discussions held among the participants of the International Society for Pharmacoepidemiology (ISPE) Comparative Effectiveness Research (CER) Special Interest Group (SIG) working group.

The purpose of these discussions was to identify opportunities, challenges, and good research practices around the use of EHRs in COVID-19 CER to inform the content of this paper which was structured in two areas: a) **methodological issues** including how to define "exposure" and "outcomes" in COVID-19 CER, how to minimize confounding and information bias, and other related methodological issues and b) data **operational challenges** specific to COVID-19 CER. A summary of these considerations is presented in Figure 1. Some of these considerations may be more relevant or applicable for one type of EHRs over another (inpatient, outpatient EHRs, post-acute care, long-term care settings or linked EHRs). It is important to note that our discussion focused on CER issues arising after the selection of EHR data sources has been tested for validity and reliability. The urgency of the COVID-19 pandemic may impose an urgency to explore new data sources; therefore, its fitness-for-use for research purposes should be thoroughly investigated. We also encourage our readers to consider our paper alongside previous relevant guidance related to the design of non-randomized studies, data collection, source validation, results reproducibility, and how to reliably synthesize results from RCTs and non-randomized studies and on general topics regarding the use of this evidence in CER (supplementary references).

Figure 1. Summary of main considerations in EHRs for COVID-19 CER

Considerations in EHRs for COVID-19 CER

Defining the study population based on exposure
- Distinction between treatment being prevalent or newly initiated
- Selection of data source for vaccine or drug exposure information
- Proper choice of treatment comparator
- Assessing the impact of rapidly changing clinical practice/ complexity of drug repurposing

Defining relevant outcomes
- Opportunity to collect data beyond diagnosis and procedural codes
- Data linkage can mitigate the impact of lack of standardization of clinical outcomes between EHRs
- Use of standardized outcomes (WHO) can facilitate direct comparability and replication of CER studies across settings between different sources

Mitigating operational challenges
- Geographical and setting differences in data measurement may obscure data comparability across EHRs
- Data linkage is critical to address measurement and discontinuity of care provision
- Account for misclassification error
- Application of prediction models to define different risk groups

Minimising common biases
- Addressing confounding bias considering the changing nature of pandemic and trends in clinical practice
- Assess the nature of missing data and opportunities for data linkage
- Assess the impact of selection (or collider) bias
- Data continuity and use of advanced methodologies

Methodological issues using EHRs in COVID-19 CER

Defining the study population-based on preventive or treatment exposure

In CER, the study population is typically defined by use or non-use of specific preventive and therapeutic interventions. For the purposes of this paper, exposure refers to pharmacological and non-pharmacological interventions used to prevent or treat COVID-19, including vaccines and therapeutics. EHR data typically allow researchers to define COVID-19 by positivity of laboratory results. We recommend using a case definition based on International Classification of Diseases (ICD) 10 diagnosis codes (e.g., U07.1) or positive results of a Nucleic Acid Amplification Test (NAAT) or Reverse transcription polymerase chain reaction (RT-PCR) since rapid antigen tests have variable performance when validated against the PCR test results, especially for asymptomatic patients or those with symptom onset more than 1 week ago (20). Antibody has not been routinely used for COVID-19 diagnosis and should not be used to define COVID-19 in CER (21). The key considerations related to defining treatment exposure data using EHR data are as follows:

➢ ***Selecting the appropriate source(s) for vaccine or drug exposure information:*** EHR systems may have various sources of medication information, each with various details and validity regarding preventive or therapeutic interventions. Prescribing (i.e., order) data are typically available in EHRs, and in many settings, electronic medication administration record (eMAR) data as well. Although drug administration is not observed for drugs dispensed in pharmacies, the likelihood of misclassification is lower in the dispensing data than in prescribing data because dispensing is one step closer to actual ingestion of medication. EMAR data have reliable inpatient, emergency room, or outpatient on-site drug use information. However, researchers should not only be concerned about ascertainment bias in relation to identification of cases but also assess data completeness for the specific drugs of interest and work with clinical experts within the system to identify these gaps

6

as data completeness is contextual and is determined through an understanding of specific data needs (22).

In addition, it is important to note that vaccine information may be incompletely captured or missing in the EHRs as most COVID-19 vaccinations programs may occur in mass vaccination sites, pharmacies, and other settings where often no health insurance claims are submitted. Information on whether an individual has been vaccinated or not relies, at times, on the individual reporting such information to their healthcare professional. The propensity of misclassifying or missing information on the vaccination status is highly dependent on local vaccination policy settings. For example, in settings where vaccines are mandatory, there is an opportunity for data linkage to vaccination records or to the collection of patient-generated information, and such information should be reported in EHR-based vaccine studies. Moreover, in settings where significant under-recording of vaccination status may be an issue, a correction factor in outcomes analysis can be applied using a standard methodology to correct exposure misclassification.

EHR data also often include structured and unstructured data, and consideration of the source and accuracy of drug exposure information is important. Unstructured pharmacy data usually take the form of free-text fields in which providers record information about prescriptions, and therefore capture information with varying degrees of completeness. Although unstructured clinical notes and images may contain medications not available in prescribing or dispensing data, such as over-the-counter drug or supplement information, natural language processing (NLP) of the free-text notes is needed. NLP may also help extract medical indications, such as thrombotic events or bleeding risks that are contained in the unstructured EHR. However, developing a valid NLP module often requires manual chart review to establish annotated dataset (the "gold-standard"), which is resource- and time-consuming.

➢ *Appropriate choice of treatment comparator:* Operationalization of an exposure-comparator definition must be mapped to the specified research question (23). However, given the dynamic nature of COVID-19 (i.e., rapidly evolving knowledge of the disease's natural history) and the lack of standard treatment guidelines, particularly at the early phase of the pandemic, it can be challenging to find appropriate therapeutic comparators. For example, each of the following comparisons may be faced with different methodological challenges: (a) comparing initiation of different treatment options (e.g., a drug or vaccine) (24), (b) comparing initiation of different doses of the same therapeutic agent, or (c) comparing any use versus no use of a treatment or comparison of different drug sequential drug therapy strategies (25). The use of an active-comparator, new-user study design is generally more desirable [comparisons (a) and (b)] as it aims to mitigate biases by firstly restricting the study to individuals with an indication for treatment and without contraindications, while also aligning individuals at the same point in time to start follow-up (i.e., treatment initiation) and ensuring the correct temporality between covariate and exposure assessment. (26) Researchers should pay close attention to the evolving guidelines as to indication of different treatment options in terms of disease severity (e.g., monoclonal antibodies are recommended to initiate early in the disease course (27) whereas systemic steroids are indicated for patients with moderate-to-severe disease (28)) while considering the use of active comparators of the same administration route.

Comparison between treatments given at different disease stages can lead to refractory confounding bias. Therefore, it is not surprising that researchers may experience difficulties in identifying a comparable alternative treatment in a newly emerging disease like COVID-19, which led to many non-user comparisons. Non-user comparisons in the non-randomized settings are

subject to two types of bias: 1) immortal time bias: researchers often need an exposure assessment period to determine non-use status (e.g., no use in the first 48 hours after admission) as immortal time bias can occur if the start of the follow-up began before the end of this exposure assessment period (e.g., the follow-up starts on the admission date) because the non-users "cannot die" until the end of assessment period before inclusion to the study and 2) confounding bias: because prescribing is highly informed by prognostic factors, non-users are either much healthier individuals for whom no treatment is needed or are individuals a grave prognosis for whom many aggressive treatments may be withheld. Such confounding may not be addressable if some prognostic factors are unmeasured in the study database. In addition, because COVID-19 care is rapidly changing as research findings emerge, careful consideration of the time trend of clinical practice is critical when choosing an appropriate comparator (29).

➢ *The distinction between prevalent vs. newly initiated users:* A new-user design is recommended in CER because the hazards of medical treatment may be different for a new user compared to a chronic user who had tolerated it before cohort entry. However, it may be challenging to distinguish new initiators from prevalent users of a drug in EHR data because some patients may have inadequate or no baseline data to determine prior drug exposure, especially for inpatient treatment studies. In other words, misclassification of prior use can occur if such use is recorded in other EHR systems for patients cared for by providers using different EHR systems. Some EHR systems have medication reconciliation data that are routinely recorded at specific medical encounters, including office visits, on hospital admission, and at discharge from hospitalization, where the providers record the medications that patients take at home, including the ones not prescribed or dispensed from the EHR system. This additional medication information, when accurate, may enhance the identification of prior drug exposure and reduce misclassification of prevalent versus new users.

➢ *The complexity of drug repurposing for COVID-19:* Several drugs assessed for utility in the prevention or treatment of COVID-19 were originally indicated for other conditions but repurposed for COVID-19 (i.e., "off-label use") (30). Drug repurposing brings additional challenges to balancing the confounders at baseline since the same drug may be used to treat COVID-19 or the original indications (e.g., some may use angiotensin-converting enzyme inhibitors to prevent COVID-19 infection while others use it to treat hypertension). Another challenge that may affect the specificity of exposure definitions is the *frequent switching and discontinuation of therapies* during the earlier phase in the pandemic when evidence was sparse and guidelines are rapidly evolving. One strategy is to use time-varying exposure definitions such as dynamic treatment strategies with proper adjustment for time-varying confounding by marginal structural models or other g-methods (31). It is also important to account for the reasons that give rise to the switch or discontinuation in these models (32) (33).

Defining outcomes relevant for COVID-19

➢ *The opportunity of collecting data on specific clinical outcomes:* While COVID-19 primarily affects the lungs, causing interstitial pneumonitis and severe acute respiratory distress syndrome (ARDS), it also affects multiple organs. A growing literature has identified some of the short- and long-term effects of COVID-19 on key markers of dysfunction in several organ systems (respiratory, cardiovascular, immune, musculoskeletal, hepatic, renal, and neurological). The rich and comprehensive clinical data contained in the EHRs often grant opportunities to ascertain these outcomes not only based on diagnosis and procedure codes but also abnormal vital signs, laboratory test results, or imaging findings (34). In addition, for newly used diagnosis and procedure codes, such as ICD diagnosis code of COVID-19 and its complications, it is also possible to use EHR data for validation of the outcome definitions by chart review.

➢ *The challenge of standardizing clinical endpoints across EHR systems:* Many COVID-19 CER studies investigated hospitalization, intubation, intensive care unit (ICU) admission, and death as the

outcomes of interest. These events are typically well-captured in EHR data sources, except for out-of-hospital death data, for which linkage to death records is often recommended. Attention should be paid to the interpretation of these outcomes as proxies of COVID-19 disease progression or recovery in the real-world setting, in the rapidly progressing pandemic, as some of these outcomes such as ICU admission and oxygen use might not be routinely collected in all data sources and influenced by the institute care protocol, hospital capacity or supply. For example, ICU admission can be misclassified due to some units being repurposed as ICU in response to the patient surge and potentially driven by the incidence rates in each period, in which case specific interventions that indicate critical illness may be a more reliable outcome, such as mechanical ventilation, extracorporeal membrane oxygenation (ECMO) or use of vasopressors.

➢ ***The lack of harmonization in the collection and reporting of outcomes:*** In EHRs, the lack of direct comparability of results from studies across different health systems and geographic areas due to lack of harmonization in the data collection and reporting is a well-documented challenge and this is not unique in COVID-19 research. However, in situations like the COVID-19 pandemic, comparability in outcome collection and definition is a critical issue that must be addressed to halt the unprecedented havoc on public health and economies. In response to this issue, organizations such as the World Health Organization (WHO) have produced guidance on the minimum set of common outcome measures for studies of COVID-19 with the aim of enabling direct comparability and replication of CER studies across different settings (35). Several consortiums have also been established to propose approaches for the aggregation of EHRs and related data to answer COVID-19 research questions by introducing specific diagnosis and procedure codes for identifying the COVID-19 related outcomes and establishing robust cohort definitions to ensure reproducibility and harmonization of concepts across different care settings (36). For instance, the core outcomes developed by the WHO research group include viral burden (quantitative PCR or cycle threshold), patient survival (mortality at hospital discharge or at 60 days), and patient progression (hospital stay length, need for mechanical ventilation) and focus on the acute phase of COVID-19 disease, whereas routinely collected safety data in EHRs such as QT-prolongation and diagnosis of arrhythmias may provide additional information regarding the CER of treatments in diverse patient populations (35). Currently, there is a lack of consensus (standard framework) on how to best evaluate long-term effects of COVID-19 disease, including ambiguity in defining the "long COVID", related and concurrent disorders and whether these are related to disease long-term itself or as a result of therapeutic/vaccines safety outcomes (37). Ongoing research funding has been made available to support research into "long COVID," including developing an EHR-based registry detailing symptoms linked to patients samples that may further characterize this disease (38).

Minimizing common biases

➢ ***Confounding bias:*** As previously noted, EHRs contain rich clinical data typically not available in insurance claims data; many of which are potential confounders in COVID-19 CER, including vital signs (e.g., oxygen saturation, blood pressures, body temperature), lifestyle factors (e.g., BMI, smoking status, alcohol consumption (39)), laboratory tests (e.g., C-reactive protein and lactate dehydrogenase, D-dimer, and other biomarkers for inflammation or disease severity), and imaging findings (e.g., chest X-ray or Computed tomography evidence of pulmonary infiltration or embolism) (40). In a rapidly evolving pandemic like COVID-19, it is crucial to adjust for potential confounding by calendar time trends in clinical practices. It is also important to consider changes in data availability and quality over time (41). Besides, much of the essential confounder information, such as patient-reported symptoms, severity, stage, the prognosis of disease, and functional status (42), is recorded in free-text notes or reports in EHRs, although this may not be consistent across hospital and EHR systems. While substantially underutilized for confounding adjustment, adding unstructured information can potentially enhance researchers' ability to reduce confounding after using NLP to

convert the free-text data into an analyzable format (18) (33). Differences in interventions between health settings could also be explored using advanced techniques, such as the use of high-dimensional propensity scores with machine learning algorithms or instrumental variable analysis, can be considered for adjustment for proxies of unmeasured confounding (43). However, for instrumental variable analyses, it is challenging to identify a valid instrument and often requires strong assumptions. For instance, prescriber preference has been used as a potential instrumental variable, but if the preference of different prescribers is linked to their quality of care that is associated with the outcome of interest (which is often the case), the assumption of the instrumental variable being only linked to the outcome through the treatment is violated (44).

➢ **Missing data:** To properly handle missing data, investigators need to understand the mechanism of missingness. Missing data may occur 1) "missing completely at random" (MCAR, i.e., missingness is independent of all factors; e.g., missing a batch of laboratory results due to fire or a natural disaster); 2) "missing at random" (MAR, i.e., missingness is only dependent on observed data; e.g., missing laboratory results in the rehabilitation facilities but no other facilities when the type of facilities is observed); 3) "missing not at random" (MNAR, i.e., missingness is dependent on unobserved data; e.g., missing a specific laboratory test and or imaging results due to differential ordering pattern of the physicians but the reasons underlying the decisions are not measured). Under MCAR, performing analysis using only those with complete data will not result in bias but may reduce statistical power. Under MAR, investigators need to collect and adjust for these factors underlying missingness using proper methods (e.g., multiple imputation, maximum likelihood-based methods, or inverse probability weighting)(45). However, these methods are less appropriate when the prevalence of missing data is very high. Under MNAR, bias is generally expected, and investigators should attempt to assess the magnitude of such impact on the study estimates (46).

In studies using RWD, investigators typically assume the absence of recording of a disease state (e.g., having a negative test for COVID-19 diagnosis) as the absence of the condition, thus EHR-based CER often turns missing data into misclassification of the study variables. Unlike claims data in which the enrollment of the insurance coverage has well-documented start- and end-date, there is no "enrollment" or "membership" defined in an EHR. Therefore, it is possible that some medical care of the study participants was provided in another EHR system and not captured by the study EHR. EHR discontinuity (e.g., receiving care outside of a particular EHR system) has been shown to be associated with a large amount of information bias in essential variables in CER (17). Applying a prediction model to identify patients with high EHR continuity and restrict the analysis among these patients can substantially reduce such biases (47). It has been demonstrated that the patients with high EHR data continuity have similar co-morbidity profiles compared to those with low EHR data continuity in a given study EHR if we compare their insurance claims data that are not affected by data discontinuity (13). Data linkage of EHRs with other data sources (e.g., claims data with a shorter time lag, such as local insurance plan data or state-reported Medicaid data or leveraging novel electronic data collection methods such as software application on smartphones to capture data in the real-world setting) is important to address information bias due to EHR data discontinuity, although this process is often complicated by privacy concerns (e.g., the requirement of patient identifiers for data linkage), different clinical terminologies, technical specifications, and functional capabilities of different data sources (48) (49).

➢ **Selection bias (or collider bias):** It can occur if restricting an analysis to those people who with a cohort-qualifying event such as hospitalization with COVID-19, been tested for active infection or who have volunteered their participation in a prospective study (i.e., conditioning on a collider variable) (50). It can also happen with researchers who included only patients without missing data when the missingness did not occur completely at random (i.e., "no missing data" effectively

becomes the cohort inclusion criterion). The spurious association is expected if the collider variable is simultaneously associated with the treatment and outcome of interest. Such bias can be addressed by inverse probability weighting with the weights being the reciprocal of the probability of being selected into the cohort, conditioning on the predictors of the cohort-qualifying event (32).

Operational challenges using EHRs in COVID-19 CER

Operational challenges due to the COVID-19 pandemic and the health care systems response and differences in data measurement across hospitals may hinder the comparability of outcomes across different settings and obstruct the possibility of combined data from different databases. Geographic variation in patient management strategies and the inability to capture exposure information from EHR data sources can also be an important challenge. Different countries or regions within countries may adopt different strategies regarding the initiation of treatment in inpatient or outpatient settings. In addition, the availability of EHRs that are more readily available for research differs from country to country. A previous study has shown the importance of differentiating between categories of patients admitted to hospitals and triaged to home; these care choices may not reflect similar patient physiology but instead reflect local care provision (51). Data measurement (e.g., safety outcomes) or detailed record-keeping on patients' regular monitoring may also significantly be impacted by the availability of medical staff and the emergency caused by the unpredicted number of patients admitted with COVID-19 across different settings. To overcome such operational challenges, data linkage is particularly important in studies using EHR data for COVID-19 CER (34). Filling the data gaps by generating linkable identifiers is critical to address mismeasurement and discontinuity of care provision due to the lack of a centralized health care system between health care providers (hospitals, nursing homes, general practitioners). However, caution should be paid around the potential risk of selection bias caused by incomplete linkage.

Lastly, it is also important to assess treatment effect heterogeneity by patient characteristics, care setting, and time considering changes in clinical practice over time and with variants of concerns. There is abundant literature on EHR-based prognostic prediction models, which can be informative to define different risk groups based on variables available in EHR (52). Given the substantial differences in public health policy, care delivery systems, and EHR data structures, it is often recommended to stratify CER analyses by geography, healthcare systems, and study databases. The expected large underlying effect of heterogeneity can also be addressed by adopting random effect models. Recently, several extensions to these models have been proposed to allow for heterogeneity across methods of imputation and adjustment for measurement errors in COVID-19 research (53).

Summary and Conclusions

The COVID-19 pandemic has presented an unprecedented need for timely and reliable assessment of safety and effectiveness of therapeutic and preventive interventions and the wide availability of RWD can play a significant role in the generation of new knowledge.

➤ EHRs represent one important source of RWD that may be critical in developing RWE to inform healthcare decision-making for COVID-19 without the need for primary data collection, something that would further negatively impact an already overburdened healthcare system. However, the urgent need of the public health emergency of the pandemic to generate "fast" conclusions should not come at the expense of methodological rigor and trust in pharmacoepidemiology science.

➤ Researchers should assess the common principles of understanding how data from EHRs are generated, data quality elements and reporting practices in the context of unique challenges

that COVID-19 disease presents. Engagement with systems generating the data will provide an important insight of the data origins and data gaps.

➢ Our recommendations start with careful construction of the research questions which requires setting up clear definitions of study population based on treatment exposure and proper choice of the comparator groups considering the changing knowledge about the COVID-19 disease and understanding the rapidly changing clinical practice patterns over time. EHR data contain rich clinical data for assessing relevant clinical endpoints, and the identification of potential key confounders and effect modifiers relevant for answering the specific CER questions regarding therapeutic and vaccine interventions for COVID-19. Furthermore, challenges related to use of this data for reliable analysis, such as transforming free-text unstructured EHR data to analyzable data set, ensuring EHR continuity and addressing the impact of missing information, will require detailed data analysis protocol and knowledge of advanced methodologies. Data harmonization and outcome reporting standardization are also pivotal for a valid pooled CER analysis as often integrating data from multiple EHR systems is needed to accrue sufficient power and to demonstrate the generalizability of the study findings across settings.

In conclusion, EHRs provide an opportunity to perform rapid COVID-19 CER due to availability of both structured and unstructured data such as laboratory and imaging data as well as relevant confounders, independent risk factors, and potential for data linkages to create a holistic view of patient management and outcomes. However, unique features of COVID-19, including varying disease presentation, disease measurement and constantly changing clinical management during an emerging pandemic, require special consideration to ensure a robust methodological approach has been followed that produces reliable evidence to support healthcare decision-making. Producing reliable comparative evidence and ensuring its rapid translation into trustworthy clinical and policy decision-making requires organized coordination and collaboration between clinicians, researchers, and health care policymakers. The pandemic of COVID-19 highlights not only the challenges in using EHRs to produce robust analyses of CER of therapeutic interventions and vaccines in a constantly changing environment, but also its unique potential to generate generalizable information from real-world, heterogeneous populations. This paper provided structured guidance for the proper conduct and appraisal of drug and vaccines effectiveness and safety research using EHR for the ongoing and future pandemics.

Collaborators
Guidance from the Working Group of the International Society for Pharmacoepidemiology Comparative Effectiveness Research Special Interest Group

Contributors
All authors conceived and developed the content of this manuscript. GS and JL drafted the manuscript and responded to other authors' comments. All authors reviewed, contributed to revisions and approved the final version of the manuscript.

Disclaimer

This article reflects the views and opinions of the authors and inputs that members of the working group shared in their personal capacity and do not represent the position of their respective bodies. The views expressed are those of the authors and do not necessarily represent the views of the Food and Drug Administration or the U.S. government.

Competing interests

DB is an employee of Takeda.
ARZ has received grant funding from Sanofi Pasteur direct to Brown University for research on the epidemiology of infections and vaccinations among nursing home residents and infants.
ADL is an employee of Boehringer Ingelheim International.
MSG is an employee of Bayer AG and owns stock in the company.
YY is an employee of AbbVie and may own AbbVie stock or options.
JL's spouse is an employee of GlaxoSmithKline and owns stock in the company.
JAL is an employee of IQVIA and owns stock in the company
MG is an employee of Merck and owns stock in the company
HY has research grant from Pfizer for unrelated work

References

1. WHO announces COVID-19 outbreak a pandemic [Internet]. [cited 2021 Aug 26]. Available from: https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic

2. WHO Coronavirus (COVID-19) Dashboard [Internet]. [cited 2021 Oct 10]. Available from: https://covid19.who.int

3. Fegert JM, Vitiello B, Plener PL, Clemens V. Challenges and burden of the Coronavirus 2019 (COVID-19) pandemic for child and adolescent mental health: a narrative review to highlight clinical and research needs in the acute phase and the long return to normality. Child and Adolescent Psychiatry and Mental Health. 2020 May 12;14(1):20.

4. Commissioner O of the. FDA Approves First COVID-19 Vaccine [Internet]. FDA. FDA; 2021 [cited 2021 Oct 10]. Available from: https://www.fda.gov/news-events/press-announcements/fda-approves-first-covid-19-vaccine

5. Regulatory approval of Pfizer/BioNTech vaccine for COVID-19 [Internet]. GOV.UK. [cited 2021 Oct 10]. Available from:

https://www.gov.uk/government/publications/regulatory-approval-of-pfizer-biontech-vaccine-for-covid-19

6. PINHO AC. EMA recommends first COVID-19 vaccine for authorisation in the EU [Internet]. European Medicines Agency. 2020 [cited 2021 Oct 10]. Available from: https://www.ema.europa.eu/en/news/ema-recommends-first-covid-19-vaccine-authorisation-eu

7. Chakraborty I, Maity P. COVID-19 outbreak: Migration, effects on society, global environment and prevention. Sci Total Environ. 2020 Aug 1;728:138882.

8. Dean NE, Gsell P-S, Brookmeyer R, Crawford FW, Donnelly CA, Ellenberg SS, et al. Creating a Framework for Conducting Randomized Clinical Trials during Disease Outbreaks. N Engl J Med. 2020 Apr 2;382(14):1366–9.

9. Radawski CA, Hammad TA, Colilla S, Coplan P, Hornbuckle K, Freeman E, et al. The utility of real-world evidence for benefit-risk assessment, communication, and evaluation of pharmaceuticals: Case studies. Pharmacoepidemiology and Drug Safety. 2020;29(12):1532–9.

10. Kruse CS, Stein A, Thomas H, Kaur H. The use of Electronic Health Records to Support Population Health: A Systematic Review of the Literature. J Med Syst. 2018 Sep 29;42(11):214.

11. Franklin JM, Lin KJ, Gatto NM, Rassen JA, Glynn RJ, Schneeweiss S. Real-World Evidence for Assessing Pharmaceutical Treatments in the Context of COVID-19. Clin Pharmacol Ther. 2021 Apr;109(4):816–28.

12. International electronic health record-derived COVID-19 clinical course profiles: the 4CE consortium | npj Digital Medicine [Internet]. [cited 2021 Aug 25]. Available from: https://www.nature.com/articles/s41746-020-00308-0

13. Lin KJ, Rosenthal GE, Murphy SN, Mandl KD, Jin Y, Glynn RJ, et al. External Validation of an Algorithm to Identify Patients with High Data-Completeness in Electronic Health Records for Comparative Effectiveness Research. Clin Epidemiol. 2020;12:133–41.

14. Madhavan S, Bastarache L, Brown JS, Butte AJ, Dorr DA, Embi PJ, et al. Use of electronic health records to support a public health response to the COVID-19 pandemic in the United States: a perspective from 15 academic medical centers. J Am Med Inform Assoc. 2021 Feb 15;28(2):393–401.

15. Two elite medical journals retract coronavirus papers over data integrity questions | Science | AAAS [Internet]. [cited 2021 Aug 26]. Available from: https://www.sciencemag.org/news/2020/06/two-elite-medical-journals-retract-coronavirus-papers-over-data-integrity-questions

16. Wells BJ, Chagin KM, Nowacki AS, Kattan MW. Strategies for handling missing data in electronic health record derived data. EGEMS (Wash DC). 2013;1(3):1035.

17. Lin KJ, Glynn RJ, Singer DE, Murphy SN, Lii J, Schneeweiss S. Out-of-system Care and Recording of Patient Characteristics Critical for Comparative Effectiveness Research. Epidemiology. 2018 May;29(3):356–63.

18. Luo Y, Thompson WK, Herr TM, Zeng Z, Berendsen MA, Jonnalagadda SR, et al. Natural Language Processing for EHR-Based Pharmacovigilance: A Structured Review. Drug Saf. 2017 Nov;40(11):1075–89.

19. Burn E, You SC, Sena A, Kostka K, Abedtash H, Abrahao MTF, et al. Deep phenotyping of 34,128 patients hospitalised with COVID-19 and a comparison with 81,596 influenza patients in America, Europe and Asia: an international network study. medRxiv. 2020 Jun 28;2020.04.22.20074336.

20. Dinnes J, Deeks JJ, Berhane S, Taylor M, Adriano A, Davenport C, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. Cochrane Database Syst Rev. 2021 Mar 24;3:CD013705.

21. Scourfield DO, Reed SG, Quastel M, Alderson J, Bart VMT, Teijeira Crespo A, et al. The role and uses of antibodies in COVID-19 infections: a living review. Oxf Open Immunol. 2021;2(1):iqab003.

22. Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. J Biomed Inform. 2013 Oct;46(5):10.1016/j.jbi.2013.06.010.

23. D'Arcy M, Stürmer T, Lund JL. The importance and implications of comparator selection in pharmacoepidemiologic research. Curr Epidemiol Rep. 2018 Sep;5(3):272–83.

24. Effect of Tocilizumab in Hospitalized Patients with Severe COVID-19 Pneumonia: A Case-Control Cohort Study [Internet]. [cited 2021 Aug 25]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7603074/

25. Association of mortality and aspirin prescription for COVID-19 patients at the Veterans Health Administration [Internet]. [cited 2021 Aug 25]. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0246825

26. Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. Curr Epidemiol Rep. 2015 Dec;2(4):221–8.

27. Anti-SARS-CoV-2 Monoclonal Antibodies | COVID-19 Treatment Guidelines [Internet]. [cited 2021 Aug 26]. Available from: https://www.covid19treatmentguidelines.nih.gov/therapies/anti-sars-cov-2-antibody-products/anti-sars-cov-2-monoclonal-antibodies/

28. Corticosteroids [Internet]. COVID-19 Treatment Guidelines. [cited 2021 Aug 26]. Available from:

https://www.covid19treatmentguidelines.nih.gov/therapies/immunomodulators/corticosteroids/

29. Methodological quality of COVID-19 clinical research | Nature Communications [Internet]. [cited 2021 Aug 26]. Available from: https://www.nature.com/articles/s41467-021-21220-5

30. Rentsch CT, Beckman JA, Tomlinson L, Gellad WF, Alcorn C, Kidwai-Khan F, et al. Early initiation of prophylactic anticoagulation for prevention of COVID-19 mortality: a nationwide cohort study of hospitalized patients in the United States. medRxiv. 2020 Dec 11;2020.12.09.20246579.

31. Li X, Young JG, Toh S. Estimating Effects of Dynamic Treatment Strategies in Pharmacoepidemiologic Studies with Time-varying Confounding: A Primer. Curr Epidemiol Rep. 2017 Dec;4(4):288–97.

32. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000 Sep;11(5):550–60.

33. Introduction to g methods | International Journal of Epidemiology | Oxford Academic [Internet]. [cited 2021 Aug 26]. Available from: https://academic.oup.com/ije/article/46/2/756/2760169?login=true

34. Dagliati A, Malovini A, Tibollo V, Bellazzi R. Health informatics and EHR to support clinical research in the COVID-19 pandemic: an overview. Briefings in Bioinformatics. 2021 Mar 1;22(2):812–22.

35. A minimal common outcome measure set for COVID-19 clinical research - PubMed [Internet]. [cited 2021 Aug 25]. Available from: https://pubmed.ncbi.nlm.nih.gov/32539990/

36. The COVID-19 Core Outcome Set Project | Cochrane Methods [Internet]. [cited 2021 Aug 26]. Available from: https://methods.cochrane.org/news/covid-19-core-outcome-set-project-invitation-take-part

37. Rando HM, Bennett TD, Byrd JB, Bramante C, Callahan TJ, Chute CG, et al. Challenges in defining Long COVID: Striking differences across literature, Electronic Health Records, and patient-reported information. medRxiv. 2021 Mar 26;2021.03.20.21253896.

38. NIH launches new initiative to study "Long COVID" [Internet]. National Institutes of Health (NIH). 2021 [cited 2021 Sep 6]. Available from: https://www.nih.gov/about-nih/who-we-are/nih-director/statements/nih-launches-new-initiative-study-long-covid

39. Umnuaypornlert A, Kanchanasurakit S, Lucero-Prisno DEI, Saokaew S. Smoking and risk of negative outcomes among COVID-19 patients: A systematic review and meta-analysis. Tob Induc Dis. 2021;19:09.

40. A multi-center study of COVID-19 patient prognosis using deep learning-based CT image analysis and electronic health records — COVID-19 Research Collaborations [Internet].

[cited 2021 Aug 26]. Available from: https://covid19.elsevierpure.com/en/publications/a-multi-center-study-of-covid-19-patient-prognosis-using-deep-lea

41. Bayley KB, Belnap T, Savitz L, Masica AL, Shah N, Fleming NS. Challenges in using electronic health record data for CER: experience of 4 learning organizations and solutions applied. Med Care. 2013 Aug;51(8 Suppl 3):S80-86.

42. S K, Aj Z, C G, J G, A S, L M-H, et al. Can we predict the severe course of COVID-19 - a systematic review and meta-analysis of indicators of clinical outcome? medRxiv : the preprint server for health sciences [Internet]. 2020 Dec 11 [cited 2021 Aug 26]; Available from: https://pubmed.ncbi.nlm.nih.gov/33200148/

43. Addressing unmeasured confounding in comparative observational research - Zhang - 2018 - Pharmacoepidemiology and Drug Safety - Wiley Online Library [Internet]. [cited 2021 Aug 26]. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/pds.4394

44. The complex role of the prescribing physician for confounder adjustment in comparative effectiveness research methods [Internet]. [cited 2021 Aug 26]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4347927/

45. Carpenter JR, Smuk M. Missing data: A statistical framework for practice. Biom J. 2021 Jun;63(5):915–47.

46. Missing clinical and behavioral health data in a large electronic health record (EHR) system - PubMed [Internet]. [cited 2021 Aug 26]. Available from: https://pubmed.ncbi.nlm.nih.gov/27079506/

47. Lin KJ, Singer DE, Glynn RJ, Murphy SN, Lii J, Schneeweiss S. Identifying Patients With High Data Completeness to Improve Validity of Comparative Effectiveness Research in Electronic Health Records Data. Clin Pharmacol Ther. 2018 May;103(5):899–905.

48. Lo W-C, Wang F-C, Lin L-Y, Jyan H-W, Wu H-C, Huang Y-L, et al. Enhancing Data Linkage to Break the Chain of COVID-19 Spread: The Taiwan Experience. Journal of Medical Internet Research. 2021 May 7;23(5):e24294.

49. Bhattacharya A, Collin SM, Stimson J, Thelwall S, Nsonwu O, Gerver S, et al. Healthcare-associated COVID-19 in England: a national data linkage study [Internet]. 2021 Feb [cited 2021 Aug 26] p. 2021.02.16.21251625. Available from: https://www.medrxiv.org/content/10.1101/2021.02.16.21251625v1

50. Griffith GJ, Morris TT, Tudball MJ, Herbert A, Mancano G, Pike L, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. Nat Commun. 2020 Nov 12;11(1):5749.

51. Turcato G, Zaboli A, Pfeifer N. The COVID-19 epidemic and reorganisation of triage, an observational study. Intern Emerg Med. 2020 Aug 9;1–8.

52. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal | The BMJ [Internet]. [cited 2021 Aug 26]. Available from: https://www.bmj.com/content/369/bmj.m1328

53. Evaluation of COVID-19 vaccine effectiveness [Internet]. [cited 2021 Sep 14]. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-vaccine_effectiveness-measurement-2021.1