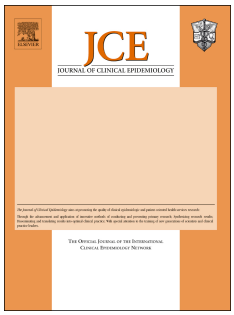

NON-INTERVENTIONAL STUDIES IN THE COVID-19 ERA: METHODOLOGICAL CONSIDERATIONS FOR STUDY DESIGN AND ANALYSIS

Anne M. Butler, PhD, Mehmet Burcu, PhD, MS, Jennifer B. Christian, PharmD, MPH, PhD, FISPE, Fang Tian, PhD, MPH, MHS, Kathleen M. Andersen, PhD, William A. Blumentals, PhD, Karen E. Joynt Maddox, MD, MPH, G. Caleb Alexander, MD, MS

PII: S0895-4356(22)00291-8

DOI: https://doi.org/10.1016/j.jclinepi.2022.11.011

Reference: JCE 10957

To appear in: *Journal of Clinical Epidemiology*

Received Date: 13 August 2022

Revised Date: 27 October 2022

Accepted Date: 9 November 2022

Please cite this article as: Butler AM, Burcu M, Christian JB, Tian F, Andersen KM, Blumentals WA, Joynt Maddox KE, Alexander GC, NON-INTERVENTIONAL STUDIES IN THE COVID-19 ERA: METHODOLOGICAL CONSIDERATIONS FOR STUDY DESIGN AND ANALYSIS *Journal of Clinical Epidemiology* (2022), doi: https://doi.org/10.1016/j.jclinepi.2022.11.011.

NON-INTERVENTIONAL STUDIES IN THE COVID-19 ERA: METHODOLOGICAL CONSIDERATIONS FOR STUDY DESIGN AND ANALYSIS

Anne M. Butler, PhD[1,2]; Mehmet Burcu, PhD, MS[3]; Jennifer B. Christian, PharmD, MPH, PhD, FISPE[4,5]; Fang Tian, PhD, MPH, MHS[6]; Kathleen M. Andersen, PhD[7,8]; William A. Blumentals, PhD[9]; Karen E. Joynt Maddox, MD, MPH[10,11]; G. Caleb Alexander, MD, MS[7,8,12,13]

1. Department of Medicine, Division of Infectious Diseases, Washington University School of Medicine, St. Louis, MO
2. Department of Surgery, Division of Public Health Sciences, Washington University School of Medicine, St. Louis, MO
3. Department of Epidemiology, Merck & Co., Inc., Rahway, NJ, United States
4. Center for Advanced Evidence Generation, Real World Solutions, IQVIA
5. Clinical Epidemiology and Health Services Research, Weill Cornell Medicine
6. Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD
7. Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
8. Center for Drug Safety and Effectiveness, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
9. Department of Epidemiology and Benefit-Risk, Sanofi, Cambridge, MA 02142
10. Department of Medicine, Division of Cardiology, Washington University School of Medicine, St. Louis, MO
11. Center for Health Economics and Policy, Institute for Public Health at Washington University, St. Louis, MO
12. Division of General Internal Medicine, Johns Hopkins Medicine, Baltimore, MD
13. Division of Endocrinology, Johns Hopkins Medicine, Baltimore, MD

CORRESPONDENCE
Anne M. Butler, PhD
Division of Infectious Diseases
John T. Milliken Department of Medicine
Washington University School of Medicine
4523 Clayton Ave., MSC 8051-0043-15
St. Louis, MO 63110
Phone: (314) 273-1975; Fax: (314) 454-5392
Email: anne.butler@wustl.edu

ABSTRACT

The global COVID-19 pandemic has generated enormous morbidity and mortality, as well as large health system disruptions including changes in use of prescription medications, outpatient encounters, emergency department admissions, and hospitalizations. These pandemic-related disruptions are reflected in real-world data derived from electronic medical records, administrative claims, disease/medication registries, and mobile devices. We discuss how pandemic-related disruptions in healthcare utilization may impact the conduct of non-interventional studies designed to characterize the utilization and estimate the effects of medical interventions on health-related outcomes. Using hypothetical studies, we highlight consequences that the pandemic may have on study design elements including participant selection and ascertainment of exposures, outcomes, and covariates. We discuss the implications of these pandemic-related disruptions on possible threats to external validity (participant selection) and internal validity (e.g., confounding, selection bias, missing data bias). These concerns may be amplified in populations disproportionately impacted by COVID-19, such as racial/ethnic minorities, rural residents, or people experiencing poverty. We propose a general framework for researchers to carefully consider during the design and analysis of non-interventional studies that use real-world data from the COVID-19 era.

KEYWORDS

COVID-19; real-world data; real-world evidence; methodology; study design; data analysis.

WHAT IS NEW?

- The authors discuss how pandemic-related disruptions in healthcare utilization may impact the conduct of non-interventional studies designed to characterize the utilization and estimate the effects of medical interventions on health-related outcomes.
- These concerns may be amplified in studies of populations that have been disproportionately impacted by COVID-19, such as racial/ethnic minorities, rural residents, or people experiencing poverty.

- Using hypothetical studies, we highlight consequences that the pandemic may have on study design elements including participant selection and ascertainment of exposures, outcomes, and covariates.
- We propose a general framework for researchers to carefully consider during the design and analysis of non-interventional studies that use real-world data from the COVID-19 era.

DISCLOSURES: After the completion of this work, Dr. Andersen became a full-time employee of Pfizer Inc. Dr. Alexander is prior Chair and a current member of the FDA Peripheral and Central Nervous System Advisory Committee; is a consultant and holds equity in Monument Analytics, a health care consultancy whose clients include the life sciences industry as well as plaintiffs in opioid litigation; and is a prior member of OptumRx's National P&T Committee. This arrangement has been reviewed and approved by Johns Hopkins University in accordance with its conflict of interest policies. Dr. Butler receives investigator-initiated funding from Merck & Co., Inc., Rahway, NJ, United States. Dr. Burcu is an employee of and owns stock in Merck & Co., Inc., Rahway, NJ, United States. Dr. Christian is an employee of and owns stock at IQVIA. Dr. Blumentals is an employee of and owns stock in Sanofi stock. Dr. Joynt Maddox serves on the Health Policy Advisory Committee for Centene Corp and has received research funding from Humana. Dr. Tian is an employee of the U.S. Food and Drug Administration. The article reflects the views and opinions of the authors and should not be construed to represent the views and opinions of the U.S. Food and Drug Administration. This manuscript was endorsed by the International Society for Pharmacoepidemiology (ISPE).

FINANCIAL SUPPORT: No specific funding for this work was provided. Dr. Butler was supported by a grant from the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (KL2 TR002346). Dr. Andersen received doctoral training support from the National Heart, Lung and Blood Institute Pharmacoepidemiology T32 Training Program (T32HL139426-03).

INTRODUCTION

The global COVID-19 pandemic (1) has caused staggering morbidity and mortality as well as substantial disruption to the healthcare sector including loss of health insurance (2, 3), decreases in preventive care (4), deferred elective and specialty care (5, 6), delays in disease diagnoses (7), changes in patient care-seeking behavior (8), and reductions in medical utilization and expenditures (9). For example, early in the course of the pandemic, there were increases in telemedicine programs (10) but decreases in prescription refills (11, 12), in-person outpatient visits (13), emergency department visits (14, 15), and hospitalizations relative to prior years (4, 16).

These pandemic-related disruptions will have important implications for the design and conduct of non-interventional studies that use real-world data, including those derived from electronic medical records (EMRs), administrative claims, disease/medication registries, and mobile devices. Researchers seeking to perform non-interventional studies with real-world data and policymakers seeking to understand them will need to consider how the pandemic may create new challenges and exacerbate perennial challenges inherent to observational research. Given changes in healthcare-seeking behavior during the pandemic, several issues are important to consider in the design and conduct of studies using real-world data from the pandemic era. These concerns may be amplified in studies of populations that have been disproportionately impacted by COVID-19, such as racial/ethnic minorities, rural residents, or people experiencing poverty (17, 18).

We focus on the importance of considering pandemic-related healthcare disruptions during the planning stage of non-interventional studies that use real-world data. We use hypothetical study examples to demonstrate the implications of pandemic-related changes on study design elements, and consequently on study validity. We focus on the following study design elements: participant selection, exposure assessment, outcome assessment, covariate assessment, and accounting for competing risks. We discuss the implications of these pandemic-related disruptions on possible threats to external validity (participant selection) and internal validity (e.g., confounding, selection bias, missing data bias), and propose

considerations and approaches to characterize or mitigate these potential biases (**Table**). We refer broadly to pandemic-related disruptions, but recognize that the timing and magnitude of these disruptions varies by geography and setting of health care delivery – some disruptions subsided early in the pandemic whereas others persisted (19). The proposed framework extends beyond recent work focused exclusively on using real-world data to study COVID-19-related treatments (20, 21), and builds upon previous work addressing the use of real-world data to study non-COVID-19-related topics (22, 23). While not exhaustive of all potential study design considerations, our work proposes a general framework reflective of an evolving conceptualization of the conduct of non-interventional studies using real-world data from the COVID-19 era. This manuscript was endorsed by the International Society for Pharmacoepidemiology (ISPE).

PARTICIPANT SELECTION: HOW HAS THE ABILITY TO SAMPLE A TARGET POPULATION CHANGED OVER TIME?

External validity refers to the extent to which an internally valid effect measured in a study sample is an unbiased estimator of the treatment effect in the target population (24). An important component of external validity is generalizability, which concerns making inference from a possibly non-random sample back to the target population (25-27). Representative sampling is not required for studies of causal inference (28), but is key for interpretation of results with respect to generalizability. During the pandemic, identification of a random sample of the target population using real-world data may be challenging due to substantial changes in patterns of health-seeking behavior, healthcare utilization, and insurance enrollment. For example, systematic differences between the target population and those who sought care through a healthcare encounter, from which the study population is sampled, (5, 29) may threaten external validity.

Consider a cohort study using a database with comprehensive pharmacy dispensing data that sought to evaluate the risk of antibiotic treatment failure among women coded for urinary tract infection (UTI) in the outpatient setting and dispensed a same-day antibiotic prescription. Over-

the-phone prescriptions likely increased during COVID, and consequently, fewer women would meet study eligibility criteria in some countries due to absence of a UTI code typically assigned during an in-person encounter. If patients with worse symptoms are more likely to seek in-person care, cohort restriction to in-person encounters may result in a sicker study population, who may also be at higher risk of the outcome. This change may result in a different sample average treatment effect than the average treatment effect in the target population.

Another important consequence of pandemic-related disruptions is their disproportionate effect on specific populations. For example, Medicaid discontinuity decreased during the pandemic due to increased Medicaid enrollment (via unemployment and loss of private insurance) as well as prohibitions on Medicaid disenrollment (30); consequently, studies of Medicaid-insured individuals may include study populations with different characteristics before versus after the pandemic. In another example, widespread and urgent implementation of telemedicine programs nationwide in response to the pandemic may have disproportionately served younger, well-resourced patients with stable internet access and comfort with technology (4, 31); consequently, studies of telemedicine may include a low proportion of older adults, racial/ethnic minorities, rural residents, or people experiencing poverty.

Approaches to Characterize or Mitigate Bias Related to Participant Selection

Given concern that the pandemic may result in a lower proportion of in-person healthcare encounters but among a sicker or frailer sample of the population, it is important to assess temporal changes in key patient- and provider-level characteristics. For example, researchers can investigate changes in severity of illness via proxies, such as receipt of kidney imaging to rule out ascension of the infection from the bladder among patients with UTI. To address changing study population characteristics, analyses may account for calendar time via stratification or other adjustment methods.

Under certain conditions, direct standardization (via g-formula (32) or adjustment formula (33, 34)), or inverse probability of sampling weights, can be used to estimate the population average

treatment effect when the study sample is not a random sample of the target population. These estimators use data from the study sample on the exposure–outcome relationship and data from the target population on either (a) the distribution of pretreatment covariates for the g-formula estimator or (b) the sampling probabilities conditional on pretreatment covariates for the inverse probability of sampling estimator (25). Details and illustrative examples are available in Lesko et al. (25).

When possible, researchers should seek and demand data sources (directly or via data linkages) that enable study of historically marginalized populations. For example, study inclusion of low-income populations or individuals living in residential facilities (e.g., jails/prisons) can enable results that generalize to important sub-populations, and ultimately contribute to overcoming a long history of underrepresentation in epidemiology and health services research (35). Robust feasibility assessments of data options are needed to identify and select fit-for-purpose data sources (36). Additionally, collaborations and initiatives across sectors such as government, clinical practice, epidemiology, and informatics can improve the presence or accuracy of variables (e.g., race/ethnicity, socio-economic status) critical for studying historically marginalized populations.

EXPOSURE ASSESSMENT: CAN EXPOSURES BE ACCURATELY IDENTIFIED?

Typically, studies of medications define the start date of the treatment exposure window based on the dispensing date of the first or second prescription of interest, after requiring a washout period without documentation of drug dispensing. Similarly, discontinuation dates are often defined as the dispensing date of the last prescription date plus the sum of the days' supply and a grace period (37). However, studies that use real-world data to define treatment exposure are susceptible to bias from unobservable factors including drug initiation, timing of initiation, and adherence. Furthermore, studies reflecting the pandemic era may be susceptible to additional biases related to exposure misclassification given disruptions to prescription drug utilization such as drug stockpiling (11, 12, 22), increased flexibility in take-home scheduling (e.g., methadone (38)), and decreased adherence to chronic medications (39).

For example, dates of drug initiation and discontinuation may be less reliable given drug stockpiling. Prescription fills for lisinopril (angiotensin-converting enzyme [ACE] inhibitor) and losartan (angiotensin receptor blocker [ARB]) peaked the week after the March 13, 2020 declaration of a national emergency by the U.S. (11). In a new-user study of ACE inhibitor users versus ARB users—restricted to patients who did not use either drug during the washout period—it is plausible that continuous users who stockpiled ACE inhibitors or ARBs may be misclassified as new users. In this scenario, inclusion of prevalent users may introduce prevalent user biases, particularly if the proportion of person-time is substantial among prevalent users. This "depletion of susceptibles" phenomenon occurs because prevalent users only represent the "survivor" subset of all initiators; it excludes individuals who became nonadherent due to experiencing events during the early period of pharmacotherapy, which may lead to substantial bias.

Approaches to Characterize or Mitigate Bias Related to Exposure Assessment

Assessing temporal trends in the prevalence of exposure by calendar time relative to COVID-19 may be useful for identifying the impact of pandemic-related changes in healthcare utilization, such as declines in laboratory tests to inform medication prescribing (6, 8). Approaches are also available to address misclassification or partial missingness of exposure. In some settings, restricting the study period to the pre-pandemic era may reduce misclassification. Also, stratifying analyses by calendar time relative to COVID-19 will provide effect estimates by time period, providing insight into the robustness of results. These analyses can be performed as primary, secondary, or sensitivity analyses, as applicable. To assess the possible impact of drug stockpiling on exposure misclassification, investigators can vary the washout period duration (e.g., 6, 12, or 18 months) to identify new-users or vary the grace period that defines date of discontinuation. To evaluate the extent of pandemic-related nonadherence to medications (i.e., persistence and implementation), investigators can assess individual-level drug dispensing data by applying common methods such as the refill gap method, the anniversary model, the proportion of days covered, and the medication possession ratio (40). In the setting of

235 comparative effectiveness research, inverse probability weighting, g-estimation, and
236 instrumental variable estimation can reduce bias introduced by nonadherence (41).

237 There are also additional approaches, which are outside of the main scope of this paper. For
238 example, in scenarios when exposure information is partly missing, two-stage g-computation
239 designs offer principled approaches to handle the missingness (42). Two-stage g-computation
240 estimators leverage partially observed information on the full study sample and complete
241 exposure information on a subset to estimate causal effects. Additional methods available to
242 address exposure misclassification include simple bias analysis (43-45), probabilistic bias
243 analysis (43, 44, 46-48), Bayesian bias analysis (43, 49-51), modified maximum likelihood (52-
244 55), multiple imputation for measurement error (56, 57), and regression calibration (58-60).
245 Funk and Landi summarize these methods and provide examples (61).

246

247 **OUTCOME ASSESSMENT: CAN OUTCOMES BE ACCURATELY IDENTIFIED?**

248 In non-interventional studies using real-world data, outcomes are commonly defined based on
249 diagnoses, procedures, prescriptions, or healthcare encounters recorded by clinicians in routine
250 care or by health insurance companies for billing purposes. Pandemic-induced changes to
251 patient care-seeking behaviors, healthcare utilization, and (in some countries) insurance
252 coverage may impact the accuracy, completeness, and timeliness of outcome assessment.
253 Thus, outcome misclassification may occur more commonly in studies using real-world data
254 generated in the pandemic. Consider a cohort study comparing risk of venous
255 thromboembolism (VTE) between initiators of different oral contraceptives, where VTE is
256 defined by the presence of diagnosis codes and radiology reports. Compared to the pre-
257 pandemic era, VTE diagnosis among patients with mild/moderate symptoms may occur less
258 frequently due to declines in in-person physical exams and diagnostic testing (62). Typically,
259 nondifferential outcome misclassification results in bias toward the null and underestimates of
260 VTE risk, whereas differential outcome misclassification (i.e., differential VTE underdiagnosis by
261 formulation) may result in bias that either exaggerates or underestimates VTE risk (61).

Selection bias due to differential loss-to-follow-up, also known as informative censoring (63), represents another common threat to internal validity in cohort studies. The current standard in epidemiological research is to treat loss-to-follow-up, such as health plan disenrollment, as a censoring event that occurs independently of the outcome, and thus as a type of noninformative censoring. However, recent evidence of differential risk of disenrollment by patient- and health-plan level characteristics raises the possibility that treating health plan disenrollment as an independent censoring event may bias descriptive statistics as well as estimates of causal effect (64, 65). Consequently, selection bias may be more pronounced in studies using real-world data collected during the pandemic, since pandemic-related increases in unemployment and subsequent health plan disenrollment resulted in more loss-to-follow-up in specific subpopulations.

In addition, due to changes in healthcare utilization patterns during the pandemic, the length of time between disease diagnosis and treatment initiation may increase due to delays in receiving medical care and treatments (5, 66). Consider a cohort study designed to quantify progression-free and overall survival among patients with resectable pancreatic cancer receiving neoadjuvant chemotherapy followed by surgery. But, during the pandemic, cancer patients may experience longer delays in initiation of treatment, impacting patient outcomes regardless of treatment type (due to delayed treatment initiation). Therefore, time from diagnosis to treatment initiation may differ by time relative to the pandemic (e.g., pre-pandemic versus pandemic era).

Approaches to Characterize or Mitigate Bias Related to Outcome Assessment

Little is known about the performance of algorithms using data from the COVID-19 era, thus, validation studies are critically needed. In addition, other standard approaches are available to address biases related to outcome assessment. Restricting or stratifying analyses by calendar time relative to COVID-19 or leveraging additional data may improve the validity and completeness of outcomes. For example, the accuracy of an algorithm to identify VTE may be enhanced by leveraging pharmacy claims or unstructured EHR notes. Furthermore, EHR data can be linked with additional data sources, such as the master death file, hospital data and/or

the national death index. Rivera et al. and Pratt et al. provide guidance on the appropriateness, feasibility, evaluation, and reporting of data linkages (67, 68).

To mitigate outcome misclassification, researchers ought to select outcomes robust to pandemic-related changes, such as outcomes that avoid reliance on outpatient and emergency department encounters that declined during the pandemic. For example, in the aforementioned UTI study, selecting a primary outcome that almost exclusively requires hospitalization, such as pyelonephritis, may be advantageous. Also, sensitivity analyses are needed to quantify the impact of outcome misclassification on estimates and the uncertainty around these estimates. Available methods are described above (43-46, 48, 49, 51, 52, 55-57, 61).

To mitigate selection bias, advanced statistical methods are available to handle outcome missingness or loss-to-follow-up. First, it is important to understand which individuals are at greatest risk of missing outcome data or loss-to-follow-up during COVID-19, and whether or not data are missing at random, missing completely at random, or missing not at random. Multiple imputation is an appropriate framework for dealing with data when missing at random or completely at random, and offers protection in some missing not at random contexts (69-71). Other analytic approaches, such as Bayesian imputation or doubly robust estimators may be explored (72-74), and sensitivity analysis may be conducted to evaluate the robustness of the different approaches. In scenarios with high potential for selection bias, appropriate methods to address selection bias should be applied including standard regression adjustment, joint modeling, and inverse probability of censoring-weighted estimation (63, 64, 74-83).

To understand differential delays between disease diagnosis and treatment between comparator groups if selected from different time periods, descriptive analyses may be conducted. For example, the time between diagnosis and treatment initiation can be assessed to detect irregularities during the pandemic; this parameter can be used as a descriptive statistic or a proxy measure to obtain confounding control. Sensitivity analyses are needed to address calendar-time (via restriction, stratification, modeling, matching, or weighting).

COVARIATE ASSESSMENT: IDENTIFYING POTENTIAL CONFOUNDERS AND EFFECT MODIFIERS

Confounding bias arises in the presence of nonexchangeability across exposure groups—due to the imbalance of (typically) causes of the outcome across levels of the exposure (84)—and is an inherent limitation of non-interventional studies due to lack of randomization. Missing data on covariates that affect the exposure and independently affect the outcome may bias effect estimates due to residual confounding. The ability to ascertain information on potential confounders and account for confounding in studies using real-world data may be particularly challenging during COVID-19. Pandemic-related stay-home orders resulted in a decline of in-person healthcare encounters and an increase in telemedicine visits, resulting in absence of routinely collected clinical measurements (e.g., body mass index, blood pressure) and laboratory results (e.g., hemoglobin A1c). This is problematic in real-world data studies because the absence of a documented diagnosis or procedure is typically interpreted as the absence of the condition. And, the extent of missingness may be differential by specialty or disease/indication due to varying impacts of the pandemic on healthcare delivery. Thus, estimates of treatment effects may be subject to more residual confounding during COVID-19 due to missing information on a variety of potential confounders.

A new consideration for studies using pandemic-era data is the possibility for COVID-19 infection history or COVID-19 vaccination status to be confounders of the exposure-outcome relationship under study. The ability for COVID-19 infection history or vaccination status to bias treatment effect estimates will be wide-ranging given the substantial COVID-19 burden of acute clinical manifestations and post-acute sequelae (85). For example, in a study aiming to estimate the comparative risk of cardiovascular events for several antidiabetic medications, COVID-19 infection history may be a potential confounder if recipients of a particular antidiabetic medication have a differential history of COVID-19 infection and differential risk of the outcome. COVID-19 infection history and COVID-19 vaccination status may also be effect measure modifiers, resulting in treatment effect heterogeneity.

Approaches to Characterize or Mitigate Bias Related to Covariate Assessment

There are several rigorous epidemiologic study design decisions that can reduce or detect the potential for confounding by measured and unmeasured confounders. First, in studies of the comparative effects of medical interventions, use of an active comparator new-user study design is a powerful tool to balance the treatment groups with respect to patient characteristics and to reduce measured and unmeasured confounding by indication and frailty bias (86, 87). The active comparator aspect of the design requires restriction of the study population to patients who receive the medical intervention of interest or an alternative intervention commonly administered to patients with the same indication and without contraindications. Second, assessment of baseline confounders during an all-available covariate assessment period is generally preferable to the fixed-duration approach (e.g., 6 months pre-index) because it performs similarly or reduces overall confounder misclassification by increasing the sensitivity of confounder measurement (88-90). Third, since missingness of potential confounders likely varies over calendar time, study designs that account for time (e.g., calendar-time specific propensity scores, matching on index date) may mitigate differences between exposure groups (91, 92). Fourth, negative control exposures and outcomes can detect, quantify, and correct for uncontrolled confounding (93-95). Lastly, propensity score calibration or data linkages can reduce missingness of potential confounders.

In addition, standard statistical approaches, such as multivariable regression models or propensity score methods, can be employed to account for confounding when potential confounders are adequately measured. Analytic approaches to handle missing data on potential confounders include complete-case analysis, last observation carried forward, the missingness pattern approach, multiple imputation, and inverse-probability-of-missingness weighting (96). Instrumental variable methods can address uncontrolled confounding when potential confounders are not adequately measured, as long as a suitable instrument exists (97, 98). However, the use of calendar time as an instrument—a common choice in comparative effectiveness studies—will likely violate the instrumental variable assumptions during COVID-19, since calendar time may affect an outcome, such as mortality, in ways other than through the exposure (99).

Sensitivity analyses are needed to quantify the impact of covariate misclassification on estimates and the uncertainty around these estimates. Available methods include regression calibration (100, 101) as well as those described above (43-49, 51, 52, 56-58, 60, 61).

To address the possibility of confounding by COVID-19 infection history, investigators may treat documented COVID-19 infection as: a) an exclusion criterion if diagnosed on or before the index date; b) a censoring event if diagnosed during the follow-up period; or c) a confounder in the analysis. Treatment effect heterogeneity by COVID-19 infection history should also be evaluated, particularly in populations with high COVID-19 prevalence. However, these approaches may be limited by variability in testing capacity for SARS-CoV-2 infection and validity of COVID-19 diagnostic codes across settings and calendar time (102-104).

ACCOUNTING FOR COMPETING RISKS

In time-to-event analyses, individuals are observed from the start of the follow-up period until the occurrence of the event of interest, a competing event, or a censoring event. Competing events preclude the outcome of interest from occurring. For example, patients who die before experiencing a stroke will never be observed to have the event of interest.

Given that COVID-19 caused substantial excess mortality, it is crucial to treat mortality as a competing event in studies with populations who experienced high COVID-19-related mortality, such as older adults or institutionalized individuals. Even though competing events are ubiquitous in epidemiological data, researchers frequently simplify analyses by treating competing events as censoring events, thus generating estimates of conditional risk defined as the risk that would be observed if all competing events were prevented without altering the hazard of the event of interest (43, 105, 106). However, the assumptions necessary for interpreting conditional risks limits the utility of these estimates for measuring public health impact. Since the degree of inflation of the estimate is proportional to the incidence of the competing event, censoring the competing events when competing events are common can dramatically distort estimates of risk (107).

Approaches to Characterize or Mitigate Bias Related to Competing Risks

The competing-risk approach provides an estimate of the total amount of the event of interest that will occur in the population, which may provide estimates with greater accuracy and precision for health care policy and planning, compared to censoring the competing events. Researchers should assess their data sources with regards to the identifiability and frequency of competing risks in their study sample (e.g., mortality), and consider using analytic methods to account for competing risks. The Aalen-Johansen estimator is straightforward to implement and can be used to generate interpretable, policy-relevant estimates of risk in the presence of competing events (108). Edwards et al. provide detail on competing risks methodology and provide applied examples from non-interventional studies using real-world data (107).

DISCUSSION

Non-interventional studies are important for augmenting RCT evidence and generating evidence from real-world populations, but designing high-quality studies in the COVID-19 era requires careful consideration. We outline several challenges inherent to non-interventional research, and discuss how disruptions to healthcare utilization and outcomes during the COVID-19-pandemic pose challenges to the validity of non-interventional studies. Our proposed framework addresses several important methodological considerations in the design of non-interventional studies using pandemic-era real-world data. Researchers may find the guidance particularly useful for studies of populations that have been disproportionately impacted by COVID-19, such as racial/ethnic minorities, rural residents, or people experiencing poverty. Detailed attention to study design and analytic decisions have broad implications for the quality of future studies using real-world data from the pandemic-era. Our recommendations will foster improvements in the design and conduct of future non-interventional studies using real-world data and enhance the ability of future studies to provide rigorous evidence that is critical to patients, caregivers, clinicians, payers, policymakers, and other stakeholders.

References

1. World Health Organization;Pageshttps://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline?gclid=Cj0KCQiA7qP9BRCLARIsABDaZzgrZXiLYiVNWbEXCaC9NJ4fx0vFF_c1hbkseUVQgrZVygxH7TrRrlAaAmcMEALw_wcB on February 2, 2021.
2. Agarwal SD, Sommers BD**.** Insurance Coverage after Job Loss - The Importance of the ACA during the Covid-Associated Recession. N Engl J Med. 2020;383(17):1603-6.
3. Blumenthal D, Fowler EJ, Abrams M, Collins SR**.** Covid-19 - Implications for the Health Care System. N Engl J Med. 2020;383(15):1483-8.
4. Whaley CM, Pera MF, Cantor J, Chang J, Velasco J, Hagg HK, et al. Changes in Health Services Use Among Commercially Insured US Populations During the COVID-19 Pandemic. JAMA Netw Open. 2020;3(11):e2024984.
5. Jazieh AR, Akbulut H, Curigliano G, Rogado A, Alsharm AA, Razis ED, et al. Impact of the COVID-19 Pandemic on Cancer Care: A Global Collaborative Study. JCO Glob Oncol. 2020;6:1428-38.
6. Meredith JW, High KP, Freischlag JA**.** Preserving Elective Surgeries in the COVID-19 Pandemic and the Future. Jama. 2020;324(17):1725-6.
7. Kaufman HW, Chen Z, Niles JK, Fesko YA**.** Changes in Newly Identified Cancer Among US Patients From Before COVID-19 Through the First Full Year of the Pandemic. JAMA Netw Open. 2021;4(8):e2125681.
8. Alexander GC, Tajanlangit M, Heyward J, Mansour O, Qato DM, Stafford RS**.** Use and Content of Primary Care Office-Based vs Telemedicine Care Visits During the COVID-19 Pandemic in the US. JAMA Netw Open. 2020;3(10):e2021476.
9. McWilliams JM, Russo A, Mehrotra A**.** Implications of Early Health Care Spending Reductions for Expected Spending as the COVID-19 Pandemic Evolves. JAMA Intern Med. 2021;181(1):118-20.
10. Patel SY, Mehrotra A, Huskamp HA, Uscher-Pines L, Ganguli I, Barnett ML**.** Trends in Outpatient Care Delivery and Telemedicine During the COVID-19 Pandemic in the US. JAMA Intern Med. 2021;181(3):388-91.
11. Vaduganathan M, van Meijgaard J, Mehra MR, Joseph J, O'Donnell CJ, Warraich HJ**.** Prescription Fill Patterns for Commonly Used Drugs During the COVID-19 Pandemic in the United States. Jama. 2020;323(24):2524-6.
12. Selke Krulichová I, Selke GW, Bennie M, Hajiebrahimi M, Nyberg F, Fürst J, et al. Comparison of drug prescribing before and during the COVID-19 pandemic: A cross-national European study. Pharmacoepidemiol Drug Saf. 2022;31(10):1046-55.
13. Baum A, Kaboli PJ, Schwartz MD**.** Reduced In-Person and Increased Telehealth Outpatient Visits During the COVID-19 Pandemic. Ann Intern Med. 2021;174(1):129-31.
14. Sokolski M, Gajewski P, Zymliński R, Biegus J, Berg JMT, Bor W, et al. Impact of Coronavirus Disease 2019 (COVID-19) Outbreak on Acute Admissions at the Emergency and Cardiology Departments Across Europe. Am J Med. 2021;134(4):482-9.
15. Nijman RG, Honeyford K, Farrugia R, Rose K, Bognar Z, Buonsenso D, et al. Presentations of children to emergency departments across Europe and the COVID-19 pandemic: A multinational observational study. PLoS Med. 2022;19(8):e1003974.
16. Jacobs EA, Ogedegbe O, Fihn SD**.** Elective Care and Health Services Research in the COVID-19 Era. JAMA Netw Open. 2020;3(11):e2025731.
17. Butler SM**.** Four COVID-19 Lessons for Achieving Health Equity. JAMA. 2020;324(22):2245-6.

18. Metzl JM, Maybank A, De Maio F**.** Responding to the COVID-19 Pandemic: The Need for a Structurally Competent Health Care System. Jama. 2020;324(3):231-2.
19. Coughlin SS, Yiğiter A, Xu H, Berman AE, Chen J**.** Early detection of change patterns in COVID-19 incidence and the implementation of public health policies: A multi-national study. Public Health Pract (Oxf). 2021;2:100064.
20. Pottegård A, Kurz X, Moore N, Christiansen CF, Klungel O**.** Considerations for pharmacoepidemiological analyses in the SARS-CoV-2 pandemic. Pharmacoepidemiol Drug Saf. 2020;29(8):825-31.
21. 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;109(4):816-28.
22. Reynolds MW, Petruski-Ivleva N, Samant M, Trivedi V, Dreyer NA, Graff J, et al. The Impact of COVID-19 on Real-World Health Data and Research: Understanding the implications of the pandemic on real-world health research: key considerations and solutions [White paper]. In: IQVIA, ed; 2021.
23. Webster-Clark M**.** Ways COVID-19 may impact unrelated pharmacoepidemiologic research using routinely collected data. Pharmacoepidemiol Drug Saf. 2021;30(3):400-1.
24. Cole SR, Stuart EA**.** Generalizing evidence from randomized clinical trials to target populations: The ACTG 320 trial. Am J Epidemiol. 2010;172(1):107-15.
25. Lesko CR, Buchanan AL, Westreich D, Edwards JK, Hudgens MG, Cole SR**.** Generalizing Study Results: A Potential Outcomes Perspective. Epidemiology. 2017;28(4):553-61.
26. Stuart EA, Cole SR, Bradshaw CP, Leaf PJ**.** The use of propensity scores to assess the generalizability of results from randomized trials. J R Stat Soc Ser A Stat Soc. 2001;174(2):369-86.
27. Pearl J**.** Generalizing experimental findings. J Causal Inference. 2015;3:259–66.
28. Rothman KJ, Gallacher JE, Hatch EE**.** Why representativeness should be avoided. Int J Epidemiol. 2013;42(4):1012-4.
29. Cutler D**.** How Will COVID-19 Affect the Health Care Economy? Jama. 2020;323(22):2237-8.
30. Corallo BM, S. 2022;Pages. Accessed at Kaiser Family Foundation at https://www.kff.org/coronavirus-covid-19/issue-brief/analysis-of-recent-national-trends-in-medicaid-and-chip-enrollment/ on October 20, 2022.
31. Gmunder KN, Ruiz JW, Franceschi D, Suarez MM**.** Demographics associated with US healthcare disparities are exacerbated by the telemedicine surge during the COVID-19 pandemic. J Telemed Telecare. 2021:1357633x211025939.
32. Robins J**.** A new approach to causal inference in mortality studies with sustained exposure periods - application to control of the healthy worker survivor effect. Math Model. 1986;7:1393-512.
33. Pearl J**.** Causality. New York, NY: Cambridge University Press; 2009.
34. Maldonado G, Greenland S**.** Estimating causal effects. Int J Epidemiol. 2002;31:422-9.
35. Kaviani P, Landi SN, McKethan A, Brookhart MA, McGrath LJ**.** Who are we missing? Underrepresentation of data sources used for pharmacoepidemiology research in the United States. Pharmacoepidemiol Drug Saf. 2020;29(11):1494-8.
36. Gatto NM, Campbell UB, Rubinstein E, Jaksa A, Mattox P, Mo J, et al. The Structured Process to Identify Fit-For-Purpose Data: A Data Feasibility Assessment Framework. Clin Pharmacol Ther. 2022;111(1):122-34.
37. Stürmer T, Wang T, Golightly YM, Keil A, Lund JL, Jonsson Funk M**.** Methodological considerations when analysing and interpreting real-world data. Rheumatology (Oxford). 2020;59(1):14-25.

38. Substance Abuse and Mental Health Services Administration 2022;Pageshttps://www.samhsa.gov/medication-assisted-treatment/statutes-regulations-guidelines/methadone-guidance on October 26, 2022.
39. Clement J, Jacobi M, Greenwood BN**.** Patient access to chronic medications during the Covid-19 pandemic: Evidence from a comprehensive dataset of US insurance claims. PLoS One. 2021;16(4):e0249453.
40. Rasmussen L, Wettermark B, Steinke D, Pottegård A**.** Core concepts in pharmacoepidemiology: Measures of drug utilization based on individual-level drug dispensing data. Pharmacoepidemiol Drug Saf. 2022;31(10):1015-26.
41. Hernán MA, Hernández-Díaz S**.** Beyond the intention-to-treat in comparative effectiveness research. Clin Trials. 2012;9(1):48-55.
42. Breger TL, Edwards JK, Cole SR, Westreich D, Pence BW, Adimora AA**.** Two-stage g-computation: Evaluating Treatment and Intervention Impacts in Observational Cohorts When Exposure Information Is Partly Missing. Epidemiology. 2020;31(5):695-703.
43. Lash TL VT, Haneuse S, Rothman KJ**.** Modern Epidemiology. Fourth ed. Philadelphia, PA: Wolters Kluwer; 2021.
44. Lash TL FM, MacLehose R**.** Applying Quantitative Bias Analysis to Epidemiologic Data. Second ed. Switzerland: Springer International Publishing; 2021.
45. Schneeweiss S**.** Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. Pharmacoepidemiol Drug Saf. 2006;15(5):291-303.
46. Lash TL, Fink AK**.** Semi-automated sensitivity analysis to assess systematic errors in observational data. Epidemiology. 2003;14(4):451-8.
47. Fox MP, Lash TL, Greenland S**.** A method to automate probabilistic sensitivity analyses of misclassified binary variables. Int J Epidemiol. 2005;34(6):1370-6.
48. Greenland S**.** Basic methods for sensitivity analysis of biases. Int J Epidemiol. 1996;25(6):1107-16.
49. MacLehose RF, Olshan AF, Herring AH, Honein MA, Shaw GM, Romitti PA**.** Bayesian methods for correcting misclassification: an example from birth defects epidemiology. Epidemiology. 2009;20(1):27-35.
50. Chu H, Wang Z, Cole SR, Greenland S**.** Sensitivity analysis of misclassification: a graphical and a Bayesian approach. Ann Epidemiol. 2006;16(11):834-41.
51. McCandless LC, Gustafson P, Levy A**.** Bayesian sensitivity analysis for unmeasured confounding in observational studies. Stat Med. 2007;26(11):2331-47.
52. Rosner B, Willett WC, Spiegelman D**.** Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. Stat Med. 1989;8(9):1051-69; discussion 71-3.
53. Neuhaus J**.** Bias and efficiency loss due to misclassified responses in binary regression. Biometrika. 1999;86(4):43–55.
54. Lyles RH, Tang L, Superak HM, King CC, Celentano DD, Lo Y, et al. Validation data-based adjustments for outcome misclassification in logistic regression: an illustration. Epidemiology. 2011;22(4):589-97.
55. Magder LS, Hughes JP**.** Logistic regression when the outcome is measured with uncertainty. Am J Epidemiol. 1997;146(2):195-203.
56. Bang H, Chiu YL, Kaufman JS, Patel MD, Heiss G, Rose KM**.** Bias Correction Methods for Misclassified Covariates in the Cox Model: comparison offive correction methods by simulation and data analysis. J Stat Theory Pract. 2013;7(2):381-400.

57. Cole SR, Chu H, Greenland S**.** Multiple-imputation for measurement-error correction. Int J Epidemiol. 2006;35(4):1074-81.
58. Lunt M, Glynn RJ, Rothman KJ, Avorn J, Stürmer T**.** Propensity score calibration in the absence of surrogacy. Am J Epidemiol. 2012;175(12):1294-302.
59. Sturmer T, Glynn RJ, Rothman KJ, Avorn J, Schneeweiss S**.** Adjustments for unmeasured confounders in pharmacoepidemiologic database studies using external information. Med Care. 2007;45(10 Supl 2):S158-65.
60. Sturmer T, Schneeweiss S, Avorn J, Glynn RJ**.** Adjusting effect estimates for unmeasured confounding with validation data using propensity score calibration. Am J Epidemiol. 2005;162(3):279-89.
61. Funk MJ, Landi SN**.** Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. Curr Epidemiol Rep. 2014;1(4):175-85.
62. Ishigami J, Kou M, Ding N, Matsushita K**.** Cardiovascular Disease and Coronavirus Disease 2019: Epidemiology, Management, and Prevention. Curr Epidemiol Rep. 2021:1-8.
63. Hernan MA, Hernandez-Diaz S, Robins JM**.** A structural approach to selection bias. Epidemiology. 2004;15(5):615-25.
64. Howe CJ, Cole SR, Lau B, Napravnik S, Eron JJ, Jr. Selection Bias Due to Loss to Follow Up in Cohort Studies. Epidemiology. 2016;27(1):91-7.
65. Butler AM, Todd JV, Sahrmann JM, Lesko CR, Brookhart MA**.** Informative censoring by health plan disenrollment among commercially insured adults. Pharmacoepidemiol Drug Saf. 2019;28(5):640-8.
66. Carrieri D, Peccatori FA**.** Seeing beyond COVID-19: understanding the impact of the pandemic on oncology, and the importance of preparedness. Hist Philos Life Sci. 2020;42(4):57.
67. Pratt NL, Mack CD, Meyer AM, Davis KJ, Hammill BG, Hampp C, et al. Data linkage in pharmacoepidemiology: A call for rigorous evaluation and reporting. Pharmacoepidemiol Drug Saf. 2020;29(1):9-17.
68. Rivera DR, Gokhale MN, Reynolds MW, Andrews EB, Chun D, Haynes K, et al. Linking electronic health data in pharmacoepidemiology: Appropriateness and feasibility. Pharmacoepidemiol Drug Saf. 2020;29(1):18-29.
69. Harel O, Mitchell EM, Perkins NJ, Cole SR, Tchetgen Tchetgen EJ, Sun B, et al. Multiple Imputation for Incomplete Data in Epidemiologic Studies. Am J Epidemiol. 2018;187(3):576-84.
70. Schafer JL, Graham JW**.** Missing data: our view of the state of the art. Psychol Methods. 2002;7(2):147-77.
71. Rubin DB**.** Multiple Imputation After 18+ Years. Journal of the American Statistical Association. 1996;91(434):473-89.
72. Erler NS, Rizopoulos D, Jaddoe VW, Franco OH, Lesaffre EM**.** Bayesian imputation of time-varying covariates in linear mixed models. Stat Methods Med Res. 2019;28(2):555-68.
73. Erler NS, Rizopoulos D, Rosmalen J, Jaddoe VW, Franco OH, Lesaffre EM**.** Dealing with missing covariates in epidemiologic studies: a comparison between multiple imputation and a full Bayesian approach. Stat Med. 2016;35(17):2955-74.
74. Perkins NJ, Cole SR, Harel O, Tchetgen Tchetgen EJ, Sun B, Mitchell EM, et al. Principled Approaches to Missing Data in Epidemiologic Studies. Am J Epidemiol. 2018;187(3):568-75.
75. Sun B, Perkins NJ, Cole SR, Harel O, Mitchell EM, Schisterman EF, et al. Inverse-Probability-Weighted Estimation for Monotone and Nonmonotone Missing Data. Am J Epidemiol. 2018;187(3):585-91.
76. Robins JM, Finkelstein DM**.** Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics. 2000;56(3):779-88.

77. Hernan MA, McAdams M, McGrath N, Lanoy E, Costagliola D. Observation plans in longitudinal studies with time-varying treatments. Stat Methods Med Res. 2009;18(1):27-52.
78. Daniel RM, Kenward MG, Cousens SN, De Stavola BL. Using causal diagrams to guide analysis in missing data problems. Stat Methods Med Res. 2012;21(3):243-56.
79. Westreich D. Berkson's bias, selection bias, and missing data. Epidemiology. 2012;23(1):159-64.
80. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol. 2008;168(6):656-64.
81. Howe CJ, Cole SR, Chmiel JS, Munoz A. Limitation of inverse probability-of-censoring weights in estimating survival in the presence of strong selection bias. Am J Epidemiol. 2011;173(5):569-77.
82. Greenland S, Pearl J. Adjustments and their Consequences—Collapsibility Analysis using Graphical Models. International Statistical Review / Revue Internationale de Statistique. 2011;79(3):401-26.
83. Hernan MA RJ. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC; 2020.
84. Westreich D. Epidemiology by Design: A Causal Approach to the Health Sciences; 2019.
85. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. Nature. 2021;594(7862):259-64.
86. Lund JL, Richardson DB, Sturmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. Curr Epidemiol Rep. 2015;2(4):221-8.
87. D'Arcy M, Stürmer T, Lund JL. The importance and implications of comparator selection in pharmacoepidemiologic research. Curr Epidemiol Rep. 2018;5(3):272-83.
88. Connolly JG, Schneeweiss S, Glynn RJ, Gagne JJ. Quantifying bias reduction with fixed-duration versus all-available covariate assessment periods. Pharmacoepidemiol Drug Saf. 2019;28(5):665-70.
89. Nakasian SS, Rassen JA, Franklin JM. Effects of expanding the look-back period to all available data in the assessment of covariates. Pharmacoepidemiol Drug Saf. 2017;26(8):890-9.
90. Conover MM, Stürmer T, Poole C, Glynn RJ, Simpson RJ, Jr., Pate V, et al. Classifying medical histories in US Medicare beneficiaries using fixed vs all-available look-back approaches. Pharmacoepidemiol Drug Saf. 2018;27(7):771-80.
91. Greifer N, Stuart EA. Matching Methods for Confounder Adjustment: An Addition to the Epidemiologist's Toolbox. Epidemiol Rev. 2022;43(1):118-29.
92. Mack CD, Glynn RJ, Brookhart MA, Carpenter WR, Meyer AM, Sandler RS, et al. Calendar time-specific propensity scores and comparative effectiveness research for stage III colon cancer chemotherapy. Pharmacoepidemiol Drug Saf. 2013;22(8):810-8.
93. Lipsitch M, Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology (Cambridge, Mass.). 2010;21(3):383-8.
94. Tchetgen Tchetgen E. The control outcome calibration approach for causal inference with unobserved confounding. Am J Epidemiol. 2014;179(5):633-40.
95. Li KQ, Shi X, Miao W, Tchetgen ET. Double Negative Control Inference in Test-Negative Design Studies of Vaccine Effectiveness. ArXiv. 2022.
96. Leyrat C, Carpenter JR, Bailly S, Williamson EJ. Common Methods for Handling Missing Data in Marginal Structural Models: What Works and Why. Am J Epidemiol. 2021;190(4):663-72.
97. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. Pharmacoepidemiol Drug Saf. 2010;19(6):537-54.
98. Brookhart MA, Schneeweiss S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. Int J Biostat. 2007;3(1):Article 14.

99. Mack CD, Brookhart MA, Glynn RJ, Meyer AM, Carpenter WR, Sandler RS, et al. Comparative Effectiveness of Oxaliplatin Versus 5-flourouricil in Older Adults: An Instrumental Variable Analysis. Epidemiology. 2015;26(5):690-9.
100. Logan R, Spiegelman D 2012. [cited 2014 July 25];Pageshttp://www.hsph.harvard.edu/donna-spiegelman/software/blinplus-macro/.
101. Spiegelman D, McDermott A, Rosner B**.** Regression calibration method for correcting measurement-error bias in nutritional epidemiology. Am J Clin Nutr. 1997;65(4 Suppl):1179s-86s.
102. Bodilsen J, Leth S, Nielsen SL, Holler JG, Benfield T, Omland LH**.** Positive Predictive Value of ICD-10 Diagnosis Codes for COVID-19. Clin Epidemiol. 2021;13:367-72.
103. Wu G, D'Souza AG, Quan H, Southern DA, Youngson E, Williamson T, et al. Validity of ICD-10 codes for COVID-19 patients with hospital admissions or ED visits in Canada: a retrospective cohort study. BMJ Open. 2022;12(1):e057838.
104. Lynch KE, Viernes B, Gatsby E, DuVall SL, Jones BE, Box TL, et al. Positive Predictive Value of COVID-19 ICD-10 Diagnosis Codes Across Calendar Time and Clinical Setting. Clin Epidemiol. 2021;13:1011-8.
105. Prentice RL, Kalbfleisch JD, Peterson AV, Jr., Flournoy N, Farewell VT, Breslow NE**.** The analysis of failure times in the presence of competing risks. Biometrics. 1978;34(4):541-54.
106. S G**.** Summary measures of population health: Concepts, ethics, and applications; 2002:291-302.
107. Edwards JK, Hester LL, Gokhale M, Lesko CR**.** Methodologic Issues When Estimating Risks in Pharmacoepidemiology. Curr Epidemiol Rep. 2016;3(4):285-96.
108. Aalen OO JS**.** An Empirical Transition Matrix for Non-Homogeneous Markov Chains Based on Censored Observations. 1978;5(3):141-50.

Table. Considerations and Approaches To Address COVID-19 Pandemic-Induced Threats To Validity

COVID-19 Pandemic-Induced Considerations	*Possible Threats to Validity*	*Approaches to Characterize or Mitigate Bias*
PARTICIPANT SELECTION: HOW HAS THE ABILITY TO SAMPLE A TARGET POPULATION CHANGED OVER TIME?		
• Stay-home orders led to missed or delayed healthcare encounters altogether, resulting in under-diagnoses and under-treatment • Patient populations who sought treatment via healthcare encounters during pandemic may comprise non-random subset of target population (e.g., more sick/frail) • Some populations may be disproportionately impacted by COVID-19 (e.g., racial/ethnic minorities, rural residents, low-income) • Changes over time in patient populations (e.g., telemedicine, Medicaid-insured)	• Reduced generalizability (external validity) due to inability to identify representative sample of target population	• Assess changes over time of baseline characteristics by exposure group, such as pre-COVID-19 versus COVID-19 eras • Account for calendar time via restriction or stratification • Use direct standardization (g-formula, adjustment formula) or inverse probability of sampling weights to estimate the population average treatment effect when study sample is not random sample of target population • Conduct robust feasibility assessment of database options to select fit-for-purpose data sources, including careful consideration of ability to study historically marginalized populations (overall or subgroup analyses) Leverage additional data sources*
EXPOSURE ASSESSMENT: CAN EXPOSURES BE ACCURATELY IDENTIFIED?		
• Stay-home orders resulted in stockpiling of drugs, resulting in limited ability to accurately identify start and stop dates of prescription drugs • Increased flexibility in take-home scheduling for scheduled medications • Possible decreases in medication adherence due to pandemic-related barriers to health care access	• Exposure misclassification via inability to identify new-users of drugs or drug discontinuation • Exposure misclassification due to non-adherence • Missing data bias	• Assess temporal trends in prevalence of exposure • Account for calendar time via restriction or stratification • Apply two-stage g-computation designs for handling missing exposure information • Conduct sensitivity analyses to vary exposure definitions and quantify potential impact of exposure misclassification • Assess adherence using refill gap method, anniversary model, proportion of days covered, medication possession ratio • Use inverse probability weighting, instrumental variable estimation, or g-estimation to reduce bias from nonadherence • Leverage additional data sources*
OUTCOME ASSESSMENT: CAN OUTCOMES BE ACCURATELY IDENTIFIED?		
• Incomplete ascertainment of outcomes with lower severity of illness (that did not require healthcare encounter) • Prolonged time between visits delay initiation of treatmncompent progression • Loss of employment/insurance coverage may result in loss-to-follow-up/missing outcomes	• Outcome misclassification • Selection bias due to differential loss-to-follow-up • Missing data bias	• Consider outcome definitions that are robust to pandemic-related changes in healthcare utilization • Assess and describe temporal trends in prevalence of outcome • Account for calendar time via restriction, stratification, matching, weighting, or multivariable adjustment • Sensitivity analyses varying algorithms to define outcomes and quantify potential impact of outcome misclassification

• Some populations may be disproportionately impacted by COVID-19 (e.g., racial/ethnic minorities, rural residents, low-income) • Delay in treatment initiation impacts outcomes regardless of treatment		• Leverage additional data sources*
COVARIATE ASSESSMENT: IDENTIFYING POTENTIAL CONFOUNDERS AND EFFECT MEASURE MODIFIERS		
• Stay-home orders resulted in missed or delayed healthcare encounters altogether, resulting in the absence of recorded data • Increased telemedicine leads to missing data on clinical measurements (e.g., height, weight, blood pressure) and routine laboratory results (e.g., hemoglobin A1c)	• Confounding due to unmeasured or poorly measured covariates • Confounding and/or effect measure modification by COVID-19 status • Missing data bias	• Assess temporal trends in prevalence of potential confounders • Define potential confounders using all-available covariate assessment period • Consider strategies for missing confounders such as complete-case analysis, last observation carried forward, the missingness pattern approach, multiple imputation, and inverse-probability-of-missingness weighting • Use negative controls to quantify uncontrolled confounding • Restrict/stratify study population by COVID-19 infection history • Censor patient follow-up on COVID-19 infection diagnosis date • Account for COVID-19 status as a confounder via restriction, stratification, matching, weighting, or multivariable adjustment • Account for calendar time via restriction, stratification, matching, weighting, or multivariable adjustment • Use restriction, stratification, matching, weighting, or multivariable adjustment to account for factors associated with differential impact of COVID-19 (e.g., race/ethnicity, income) • Sensitivity analyses varying algorithms to define covariates (e.g., duration of look-back period) and quantify potential impact of confounding • Replicate analyses among different data sources • Leverage additional data sources*
ACCOUNTING FOR COMPETING RISKS		
• COVID-19 resulted in high mortality, particularly in certain subpopulations • Mortality is a competing event which precludes occurrence of many outcomes of interest	• Inflation of risk estimates when competing events are treated as censoring events	• Assess frequency of competing events • Use methods to estimate risk accounting for competing events • Compare results to those obtained using analytic simplifications commonly used to handle competing events, such as treating competing events as censoring events

***Leveraging additional data sources can enrich existing data in several ways (e.g., capture structured and unstructured data fields; include various types of healthcare encounters and communications such as office visits, telehealth visits, phone/email communications; inclusion of a different population).

WHAT IS NEW?

- The authors discuss how pandemic-related disruptions in healthcare utilization may impact the conduct of non-interventional studies designed to estimate the utilization and effects of medical interventions on health-related outcomes.
- These concerns may be amplified in studies of populations that have been disproportionately impacted by COVID-19, such as racial/ethnic minorities, rural residents, or people experiencing poverty.
- Using hypothetical studies, we highlight consequences that the pandemic may have on study design elements including participant selection and ascertainment of exposures, outcomes, and covariates.
- We propose a general framework for researchers to carefully consider during the design and analysis of non-interventional studies that use real-world data from the COVID-19 era.

DISCLOSURES: After the completion of this work, Dr. Andersen became a full-time employee of Pfizer Inc. Dr. Alexander is prior Chair and a current member of the FDA Peripheral and Central Nervous System Advisory Committee; is a consultant and holds equity in Monument Analytics, a health care consultancy whose clients include the life sciences industry as well as plaintiffs in opioid litigation; and is a prior member of OptumRx's National P&T Committee. This arrangement has been reviewed and approved by Johns Hopkins University in accordance with its conflict of interest policies. Dr. Butler receives investigator initiated funding from Merck & Co., Inc. Dr. Burcu is an employee of and owns stock in Merck & Co., Inc. Dr. Christian is an employee of and owns stock at IQVIA. Dr. Blumentals is an employee of and owns stock in Sanofi stock. Dr. Joynt Maddox serves on the Health Policy Advisory Committee for Centene Corp and has received research funding from Humana. Dr. Tian is an employee of the U.S. Food and Drug Administration. The article reflects the views and opinions of the authors and should not be construed to represent the views and opinions of the U.S. Food and Drug Administration.

FINANCIAL SUPPORT: No specific funding for this work was provided. Dr. Butler was supported by a grant from the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (KL2 TR002346). Dr. Andersen received doctoral training support from the National Heart, Lung and Blood Institute Pharmacoepidemiology T32 Training Program (T32HL139426-03).

Author Statement

Butler, Burcu: Conceptualization **Butler, Burcu, Christian, Tian, Andersen, Blumentals**: Writing- Original draft preparation. **Butler, Burcu, Alexander**: Supervision, Project Administration**: Butler, Burcu, Christian, Tian, Andersen, Blumentals, Joynt Maddox, Alexander:** Writing- Review and Editing