Biases in evaluating the safety and effectiveness of drugs for covid-19: designing real-world evidence studies

Christel Renoux, Laurent Azoulay, and Samy Suissa

Correspondence to Dr. Samy Suissa, Centre for Clinical Epidemiology, Jewish General Hospital 3755 Cote Ste-Catherine, H4.61, Montreal, Québec, Canada H3T 1E2 (e-mail: samy.suissa@mcgill.ca) Tel: 514-340-7593 Fax: 514-340-7564

Main point of contact during editing:
Dr. Christel Renoux, Centre for Clinical Epidemiology, Jewish General Hospital 3755 Cote Ste-Catherine, H-416, Montreal, Québec, Canada H3T 1E2 (e-mail: christel.renoux@mcgill.ca) Tel: 514-340-7563 Fax: 514-340-7564

Author affiliations: Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada (Christel Renoux, Laurent Azoulay, and Samy Suissa); Centre for Clinical Epidemiology, Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Canada (Christel Renoux, Laurent Azoulay, and Samy Suissa); Department of Neurology and Neurosurgery, McGill University, Montreal, Canada (Christel Renoux); and Gerald Bronfman Department of Oncology, McGill University, Montreal, Canada (Laurent Azoulay).

This study was supported by infrastructure funding from the Canadian Institutes of Health Research and the Canadian Foundation for Innovation. The funding source for the work had no role in the design of the study; the collection, analysis, and interpretation of the data; and the decision to approve publication of the finished manuscript.

Conflict of interest: CR reports no conflict of interest. LA consulted for Janssen and Pfizer for work unrelated to this manuscript. SS participated in advisory board meetings or as speaker for Atara, Boehringer-Ingelheim, Bristol-Myers-Squibb, Merck, Pfizer, Seqirus.
All authors had access to the data and a role in writing the manuscript.

Running head: Biases in evaluating drugs for covid-19

Abstract

The coronavirus disease 2019 (COVID-19) pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has led to an unprecedented effort to generate real-world evidence on the safety and effectiveness of various treatments. A growing number of observational studies evaluating the effects of certain drugs have been conducted, including several assessing whether hydroxychloroquine improves outcomes in infected individuals and whether renin-angiotensin-aldosterone system inhibitors have detrimental effects. We review and illustrate how immortal time bias and selection bias were present in several of these studies. Understanding these biases and how they can be avoided may prove important for future observational studies assessing the effectiveness and safety of potentially promising drugs during the COVID-19 pandemic.

Keywords: bias; cohort studies; covid-19; epidemiology

Abbreviations:

COVID-19: coronavirus disease 2019

Severe acute respiratory syndrome coronavirus 2: SARS-CoV-2

Angiotensin-converting enzyme inhibitors: ACEIs

Angiotensin II receptor blockers: ARBs

Randomized controlled trials: RCTs

The coronavirus disease 2019 (COVID-19) pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has led to an unprecedented effort to generate real-world evidence on the safety and effectiveness of various treatments. While the randomized controlled trial (RCT) is widely accepted as the design providing the most definitive results, the generalizability of its findings to the real-world setting can be challenging. Indeed, RCTs often use strict selection criteria and treatment adherence that may differ with the real-world setting.[1] Moreover, compared with observational studies, RCTs may take longer to implement, and therefore their findings may take longer to reach the scientific community, which is a particular concern in the context of a rapidly evolving pandemic. Thus, by leveraging the rapidly accumulating data on patients hospitalized with COVID-19, a growing number of observational studies evaluating the effects of certain drugs have been conducted and published at an impressive pace. This includes assessing the effectiveness of hydroxychloroquine and evaluating whether angiotensin-converting enzyme inhibitors (ACEIs) and angiotensin II receptor blockers (ARBs) have detrimental effects. While the urge to generate rapid information to guide clinical practice is understandable, only well-conducted observational studies will provide results that can inform clinical decision-making, health policies, and future research.

While observational studies have the potential to complement the results of RCTs, they can be methodologically complex. Aside from adequate control of confounding, studies of drug effectiveness and safety present unique challenges that stem from the time-varying nature of drug exposure. Indeed, in many studies, cohort entry is not defined by the date of treatment initiation but may correspond, for instance, to the date of disease diagnosis or the hospital admission date for a COVID-19 infection. Thus, in these real-world scenarios, patients may start

treatment at different times following the cohort entry date. While these considerations are irrelevant in RCTs where the date of randomization always defines cohort entry and patients considered exposed to the drug for the entire follow-up, the proper definition of cohort entry and exposure over time is inherently more complex in cohort studies.

These critical aspects of drug effect cohort studies are often underappreciated. Indeed, whereas confounding issues are subjected to intense and often exclusive scrutiny, properly defining and analyzing treatment status over follow-up time is equally important to avoid introducing immortal time bias.[2,3] Immortal time corresponds to a period of follow-up (person-time) during which, by design, the outcome of interest (death or another outcome) cannot occur. In cohort studies assessing drug effectiveness and safety, immortal time is typically introduced when treatment status is defined based on a prescription issued or received at some point during follow-up. The time period between disease diagnosis or cohort entry and the first treatment prescription is necessarily *immortal* since the patient had to survive or be outcome-free (owing to the censoring of events in the analysis) to be classified as exposed. Immortal time bias is then introduced when this immortal period between diagnosis or cohort entry and the first treatment prescription is misclassified as exposed rather than correctly accounted for as unexposed. Immortal time bias is also introduced when this immortal period is excluded, with cohort entry defined as the date of treatment initiation for exposed patients and defined as the date of disease diagnosis or hospitalization for non exposed patients. Immortal time bias is common in cohort studies of drug effects and systematically biases the results downward in favor of the treatment under study. Consequently, this bias can make harmful treatments appear neutral, and neutral treatments appear protective.

Most recent cohort studies assessing the effectiveness of hydroxychloroquine on mortality in patients hospitalized with COVID-19 determined exposure, in their primary analysis, based on treatment received at *any time* during follow-up or typically within 48 hours of hospitalization. However, patients were considered exposed as of the date of hospital admission, thereby introducing immortal time bias (Figure 1A).[4-9] Indeed, the time period between cohort entry (date of hospitalization) and treatment initiation is *immortal* since the patients had to survive or be outcome-free to be classified as treated. This immortal period between the date of hospitalization and treatment initiation was misclassified as exposed rather than correctly accounted for as unexposed. Immortal time bias was also introduced when patients receiving hydroxychloroquine during hospitalization were followed from the date of starting treatment while patients not exposed were followed from the date of hospitalization. In this instance, the immortal period between the date of hospitalization and treatment initiation was excluded from the analysis.[10] The magnitude of the bias and its overall impact on the results depends on the duration of the immortal period, the duration of follow-up, the number of exposed in the cohort, and the event rate. Although we focused on studies assessing the effectiveness of hydroxychloroquine on mortality, immortal time bias was also introduced in studies evaluating other drugs. For instance, immortal time bias was present in a recent cohort study evaluating the association between ACEIs/ARBs and all-cause mortality in patients with hypertension hospitalized with COVID-19.[11] Indeed, patients who received ACEs/ARBs at *any time* during the hospitalization were considered exposed from the date of admission until the end of follow-up, regardless of the timing of treatment initiation.

Another critical design aspect is the selection of patients to be included in, and particularly excluded from, the cohort. As such, the exclusion of patients based on an event or

treatment occurring at some point during follow-up can lead to selection bias (Figure 1B). For example, in some cohort studies, exposure was based on receiving hydroxychloroquine within 48 hours of hospitalization; patients initiating treatment more than 48 hours after hospital admission were excluded in the primary analysis.[8,10] This exclusion may introduce selection bias in addition to immortal time bias. Both immortal time and selection bias were also at play in a study where the cohort was restricted to patients with at least six days of follow-up and a minimum of three days of treatment with hydroxychloroquine, but cohort entry was defined as the date of hospital admission.[12] Also, this study did not have any comparator group. Similarly, excluding patients who did not experience the outcome or were not yet discharged by the end of the study period is incorrect.[4] A flowchart describing cohort selection with numbers of patients excluded and reasons for exclusion should therefore be provided to assess the potential for such bias. It should be noted that these methodological issues are introduced by the investigators at the design or analysis stage, and thus, are not inherent 'flaws' of cohort studies. Moreover, as these biases are information and selection biases, methods used to deal with confounding, such as propensity scores, would not correct these biases.

Several options can be used at the design or analytical stage to prevent these biases. One approach is to define exposure at the date of hospitalization (cohort entry) and thus consider as exposed only those patients initiating the drug of interest at the date of hospitalization while all other patients not exposed at cohort entry, including patients initiating the drug later during follow-up, are considered unexposed. Similarly, exposure could be defined based on treatment initiation in the first 48 hours of hospitalization with cohort entry accordingly moved to 48 hours after the date of hospitalization. Although this option is easy to implement, it only assesses the effectiveness of the drug when initiated at or soon after the date of hospitalization and does not

optimally use the information from all patients exposed to the drug of interest over time. Moreover, the study population does not include patients who die early after hospitalization. Finally, one caveat of this approach is that it may introduce some exposure misclassification. A second approach is to use a time-varying exposure definition at the analytic stage. The study cohort comprises all consecutive patients hospitalized with COVID-19 during a specific time period with cohort entry defined as the date of hospitalization. For each patient, each day of follow-up is classified as either exposed or not exposed to the drug of interest, allowing patients to move from a period of non-exposure to a period of exposure. Once treatment is initiated, patients can be considered exposed for the remainder of the follow-up regardless of treatment discontinuation (analogous to an intention to treat approach), censored when treatment is stopped, or their person-days of follow-up after treatment cessation classified as unexposed. A grace period can be added after treatment discontinuation where patients are still considered exposed to account for the residual biological effect of the drug under study. Time-dependent Cox proportional hazards models are then used to estimate hazard ratios for the association between current use of the drug under study and the outcome of interest.

Finally, a design approach aimed at emulating a target trial in this setting can be the prevalent new-user cohort design.[13] Briefly, the base cohort consists of all consecutive patients hospitalized with COVID-19 during a specific time period. This base cohort includes patients initiating treatment (for example, hydroxychloroquine) at various time points during hospitalization and patients not treated during the entire follow-up. Among this base cohort, each patient initiating hydroxychloroquine is matched 1:1 (or 1:n) to a patient not exposed to hydroxychloroquine up to the same point in time. Thus starting chronologically to emulate the randomized trial process, each patient initiating hydroxychloroquine is matched without

replacement to a patient not exposed to hydroxychloroquine up to this point in time. The point of hydroxychloroquine initiation is used to define the time-based exposure set which include all potential comparator patients with the same duration of follow-up since entry into the base cohort (date of hospitalization) and determine the point at which one patient starting hydroxychloroquine is matched to one comparator patient. This approach allows to take into account the time since hospitalization admission and provides a similar time point during hospitalization to measure characteristics for the treated patient and the matched comparator. Time conditional propensity scores can be used to identify the comparator patient most similar to the patient who initiated hydroxychloroquine. They are time conditional because they depend on the time-varying patient characteristics measured at the point of the time-based exposure sets, and the positivity assumption is verified conditionally within each exposure set. To compute the propensity of initiating hydroxychloroquine versus no treatment as a function of the time-varying patient characteristics measured at the point of the exposure set, conditional logistic regression is used to conserve the matching induced by the exposure set. For the matched pairs formed, cohort entry is defined as the date of hydroxychloroquine initiation and the corresponding date for the matched comparator patient. A strenght of this design is its flexibility with the possibility to compare multiple treatment regimens. As such, patients switching to a different drug during hospitalization may be compared to patients continuing on the first drug initiated during hospitalization while matching on duration of previous treatment. Similar approaches have been proposed by others, such as creating sequential cohorts at predetermined time intervals.[14,15]

Aside from these design issues, the potential for confounding by indication is always a concern in drug effectiveness studies, particularly in the current pandemic where no treatments are available. Indeed, off-label treatments are typically given preferentially to moderate to severe

patients at the time of hospital admission or to those with a worsening condition during follow-up; in extreme situations, the treatment is given for compassionate use to highly severe patients. Thus, depending on the clinical context, the confounding may be intractable so that available statistical methods will not be able to control for this bias. The rapidly changing treatment recommendations may create an additional challenge in adequately balancing the exposure groups. Finally, traceable and transparent data are prerequisites to the above considerations, as recently reminded to the scientific community.

In summary, real-world data are useful to complement evidence from RCTs and can even predict their results in some settings.[16] However, recently published cohort studies assessing the effectiveness and safety of drugs in patients hospitalized with COVID-19 illustrate the importance of carefully designing and analyzing such studies. While much attention is paid to confounding, fundamental methodological principles must also be applied to derive meaningful conclusions. Methods exist, such as the prevalent new-user design, that allow to avoid these biases and permit a proper control for confounding.[13] Otherwise, ill-designed observational studies may have detrimental consequences on clinical decision-making, informing future clinical trials, and ultimately the credibility of observational studies. These methodologic principles may prove important for the many future observational studies on potentially promising drugs in the midst of the COVID-19 pandemic.

Acknowledgements: Dr. Christel Renoux (Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada; Centre for Clinical Epidemiology, Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Canada; and Department of Neurology and Neurosurgery, McGill University, Montreal, Canada) reports no conflict of interest. Dr. Christel Renoux is a recipient of a chercheur-boursier salary award from the Fonds de recherche du Québec – Santé. Dr. Laurent Azoulay (Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada; Centre for Clinical Epidemiology, Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Canada; and Gerald Bronfman Department of Oncology, McGill University, Montreal, Canada) consulted for Janssen and Pfizer for work unrelated to this manuscript. Dr. Laurent Azoulay is a recipient of a chercheur-boursier salary award from the Fonds de recherche du Québec – Santé. Dr. Samy Suissa (Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada; and Centre for Clinical Epidemiology, Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Canada) participated in advisory board meetings or as speaker for Atara, Boehringer-Ingelheim, Bristol-Myers-Squibb, Merck, Pfizer, Seqirus. Dr. Samy Suissa is the recipient Distinguished James McGill Chair. All authors had access to the data and a role in writing the manuscript.

Data Availability Statement: no data available

References

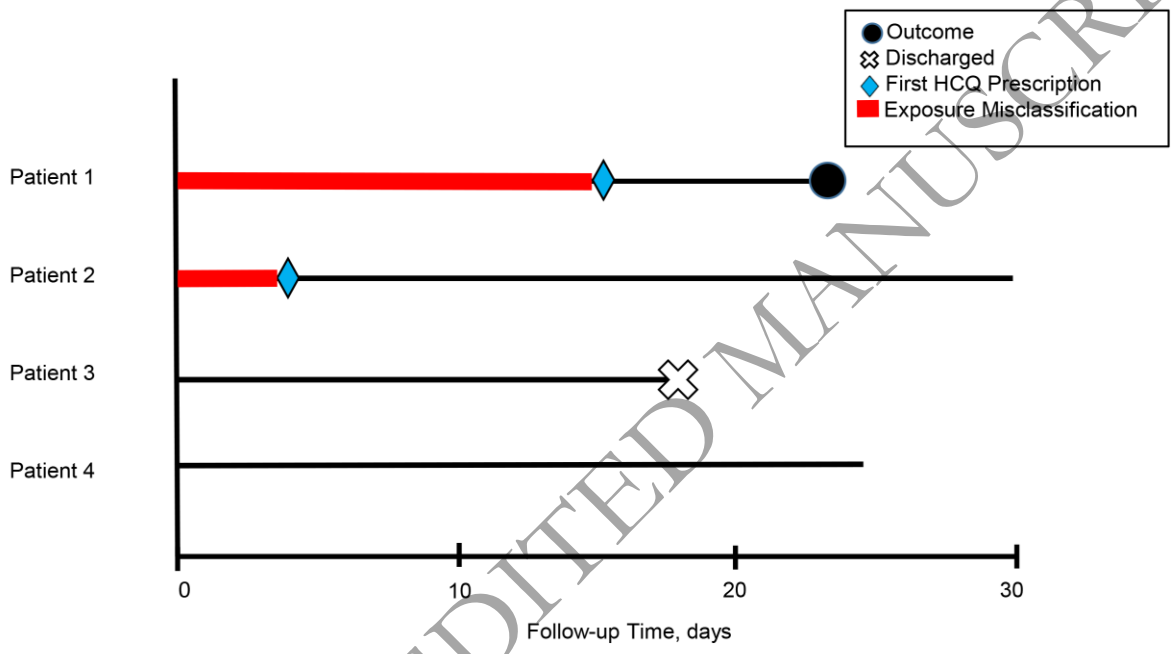
1. Westreich D, Edwards JK. Invited commentary: every good randomization deserves observation. *Am J Epidemiol.* 2015;182(10):857-860.
2. Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol.* 2008;167(4):492-499.
3. Levesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ.* 2010;340:b5087.
4. Arshad S, Kilgore P, Chaudhry ZS, et al. Treatment with hydroxychloroquine, azithromycin, and combination in patients hospitalized with COVID-19. *International Journal of Infectious Diseases.* 2020;97:396-403.
5. Catteau L, Dauby N, Montourcy M, et al. Low-dose hydroxychloroquine therapy and mortality in hospitalised patients with COVID-19: a nationwide observational study of 8075 participants. *International journal of antimicrobial agents.* 2020;56(4):106144.
6. Geleris J, Sun Y, Platt J, et al. Observational Study of Hydroxychloroquine in Hospitalized Patients with Covid-19. *N Engl J Med.* 2020 Jun 18;382(25):2411-2418.
7. Ip A, Berry DA, Hansen E, et al. Hydroxychloroquine and tocilizumab therapy in COVID-19 patients-An observational study. *PLoS One.* 2020;15(8):e0237693.
8. Mahevas M, Tran VT, Roumier M, et al. Clinical efficacy of hydroxychloroquine in patients with covid-19 pneumonia who require oxygen: observational comparative study using routine care data. *BMJ.* 2020;369:m1844.
9. Rosenberg ES, Dufort EM, Udo T, et al. Association of Treatment With Hydroxychloroquine or Azithromycin With In-Hospital Mortality in Patients With COVID-19 in New York State. *JAMA.* 2020 Jun 23;323(24):2493-2502.
10. Paccoud O, Tubach F, Baptiste A, et al. Compassionate use of hydroxychloroquine in clinical practice for patients with mild to severe Covid-19 in a French university hospital. *Clin Infect Dis.* 2020 Jun 18;ciaa791. [Online ahead of print] DOI: 10.1093/cid/ciaa791.
11. Zhang P, Zhu L, Cai J, et al. Association of Inpatient Use of Angiotensin Converting Enzyme Inhibitors and Angiotensin II Receptor Blockers with Mortality Among Patients With Hypertension Hospitalized With COVID-19. *Circ Res.* 2020 Jun 5;126(12):1671-1681.
12. Gautret P, Lagier JC, Parola P, et al. Clinical and microbiological effect of a combination of hydroxychloroquine and azithromycin in 80 COVID-19 patients with at least a six-day follow up: A pilot observational study. *Travel medicine and infectious disease.* 2020;34:101663.
13. Suissa S, Moodie EE, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. *Pharmacoepidemiol Drug Saf.* 2017;26(4):459-468.
14. Hernan MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology.* 2008;19(6):766-779.
15. Ross ME, Kreider AR, Huang YS, Matone M, Rubin DM, Localio AR. Propensity Score Methods for Analyzing Observational Data Like Randomized Experiments: Challenges and Solutions for Rare Outcomes and Exposures. *Am J Epidemiol.* 2015;181(12):989-995.
16. Patorno E, Schneeweiss S, Gopalakrishnan C, Martin D, Franklin JM. Using Real-World Data to Predict Findings of an Ongoing Phase IV Cardiovascular Outcome Trial: Cardiovascular Safety of Linagliptin Versus Glimepiride. *Diabetes Care.* 2019;42(12):2204-2210.

Figure 1. Illustration of immortal time and selection biases in a hypothetical cohort of patients hospitalized with COVID-19

Panel A. Patients 1 and 2 receive a first prescription of hydroxychloroquine at some point during hospitalization, but are considered exposed from the date of hospital admission, thereby introducing immortal time bias caused by exposure misclassification (red line).

Panel B. At the end of the study period, patients 2 and 4 have not experienced the outcome but are not yet discharged. These two patients, still alive and unexposed for a period of time, are incorrectly excluded (red dashed lines) therefore introducing selection bias. Abbreviation: HCQ, hydroxychloroquine.

A)



B)