HHS Public Access

Author manuscript

Clin Pharmacol Ther. Author manuscript; available in PMC 2023 January 01.

Published in final edited form as:

Clin Pharmacol Ther. 2022 January ; 111(1): 243–251. doi:10.1002/cpt.2400.

Longitudinal data discontinuity in electronic health records and consequences for medication effectiveness studies

Kueiyu Joshua Lin[1], **Yinzhu Jin**[1], **Joshua Gagne**[1], **Robert J. Glynn, ScD, PhD**[1], **Shawn N. Murphy**[2], **Angela Tong**[1], **Sebastian Schneeweiss**[1]

[1]Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School

[2]Mass General Brigham Research Information Science and Computing, Massachusetts General Hospital Department of Neurology, and Harvard Medical School

Abstract

Electronic health records (EHR) discontinuity, i.e., receiving care outside of the study EHR system, can lead to information bias in EHR-based real-world evidence (RWE) studies. An algorithm has been previously developed to identify patients with high EHR-continuity. We sought to assess whether applying this algorithm to patient selection for inclusion can reduce bias caused by data-discontinuity in 4 RWE examples. Among Medicare beneficiaries aged >=65 years from 2007 to 2014, we established four cohorts assessing drug effects on short-term or long-term outcomes, respectively. We linked claims data with two US EHR systems and calculated %bias of the multivariable-adjusted effect estimates based on only EHR vs. linked EHR-claims data since the linked data capture medical information recorded outside of the study EHR. Our study cohort included 77,288 patients in system 1 and 60,309 in system 2. We found the sub-cohort in the lowest quartile of EHR-continuity captured 72–81% of the short-term and only 21–31% of the long-term outcome events, leading to %bias of 6–99% for the short-term and 62–112% for the long-term outcome examples. This trend appeared to be more pronounced in the example using a non-user comparison rather than an active comparison. We did not find significant treatment effect heterogeneity by EHR-continuity for most subgroups across empirical examples. In EHR-based RWE studies, investigators may consider excluding patients with low algorithm-predicted EHR-continuity as the EHR data capture relatively few of their actual outcomes, and treatment effect estimates in these patients may be unreliable.

Keywords

data leakage; care continuum; patient connectedness; loyalty cohort; data completeness

Corresponding author: Kueiyu Joshua Lin, Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, 1620 Tremont St. Suite 3030, Boston, MA 02120, Phone: (617) 278-0930; Fax: (617) 232-8602. jklin@bwh.harvard.edu.

Author contributions:

K.J.L. and S.S. wrote the manuscript. K.J.L. and S.S. designed the research. All authors performed the research. Y.J. analyzed the data.

Supplementary File:

1. Supplemental Material.docx

Introduction:

Large comparative effectiveness research (CER) studies are often needed in a timely fashion as new medications are marketed with limited information about their effectiveness in routine care. Much effectiveness research and pragmatic randomized trials work with secondary healthcare data.[1–3] In the US, there has been a remarkable growth in electronic health record (EHR) databases availability for clinical research purposes in the last decade.[4,5] EHR data contain rich clinical information essential for patient phenotyping and confounding adjustment that is not available in other administrative databases, which has substantially expanded researchers" capacity in CER.[6,7] and clinical decision support tool development.[8] However, except for those based in integrated healthcare delivery systems, most US EHR systems do not comprehensively capture medical encounters across all care settings (e.g., ambulatory office, emergency room, hospitals, etc.). We define EHR-discontinuity as "receiving care outside the reach of a given EHR system." Our prior work showed that EHR-discontinuity could cause a substantial amount of misclassification of the study variables because medical information recorded at a facility outside of a given EHR system is "invisible" to the investigators and therefore often assumed to be absent in the study.[9] In contrast, insurance claims data have defined enrollment (start and end) dates and recording of all covered healthcare encounters across care settings and locations, although the level of clinical detail is less than in an EHR system.[10] Linking EHR with claims data could potentially address bias due to EHR-discontinuity, but such linkage is often not feasible for governance reasons and privacy and compliance concerns (e.g., sensitive identifiers required for reliable linkage may not be accessible). Insufficient overlap between databases is another common reason that limits the usability of data linkage.

To reduce information bias (i.e., misclassification of the study variables) for comparative effectiveness research based on EHR data alone, we previously developed and validated a prediction algorithm to identify patients with high EHR-continuity.[11] We found that patients in the top quintile of predicted EHR-continuity had 3.5–5.8 fold less misclassification of 40 clinical factors commonly used as drug exposure, confounders, and outcome variables in comparative effectiveness research studies compared to those in the lower quintiles of predicted EHR-continuity.[11] However, the influence of such information bias on comparative safety and effectiveness analyses is likely context-specific. For example, the influence of EHR-continuity on an outcome that requires longer follow-up after a chronic medication (e.g., heart failure or malignancy after taking an antidiabetic) may differ from that of an acute outcome after a short-term medication exposure (e.g., hyperkalemia after antibiotic use). Therefore, we aimed to assess the validity of comparative effectiveness and safety study results based on EHR data alone in patients with high vs. low predicted EHR-continuity when compared to the gold-standard estimates based on EHR linked with insurance claims where medical information outside of study EHR is also available. The empirical examples included different medication comparisons in relation to both short-term and long-term outcomes.

Methods:

Data sets:

We linked longitudinal claims data from the US Medicare system to EHR data from two medical care delivery networks. The first network (EHR system1) consists of 1 tertiary hospital, 2 community hospitals, and 19 primary care centers. The second network (EHR system 2) includes 1 tertiary hospital, 1 community hospital, and 18 primary care centers. The EHR database contains information on patient demographics, medical diagnoses, procedures, medications, and various clinical data. The Medicare claims data contain information on demographics, enrollment start- and end-dates, dispensed medications and performed procedures, and medical diagnoses.[10] In the prior study, the EHR system 1 was used for training and system 2 for validating the EHR-continuity prediction model.[11]

Study population:

The study cohort consists of the Medicare fee-for-service beneficiaries aged 65 years and older with at least 365 days of continuous enrollment in Medicare (including inpatient, outpatient, and prescription coverage) from 2007/1/1 to 2014/12/31 and with at least one EHR encounter in EHR system 1 or 2 during their active Medicare enrollment period. Among these patients, we established 4 comparative cohorts: 1) Comparing the effect of two antibiotics on a short-term outcome (A-STO): hyperkalemia within 30 days after new use of trimethoprim/sulfamethoxazole vs. cephalexin[12,13]; 2) Comparing the effect of two antibiotics on a long-term outcome (A-LTO): Clostridium difficile infection (CDI) in the year following new use of trimethoprim/sulfamethoxazole vs. cephalexin[14]; 3) Comparing the effect of a gastroprotective agent vs. non-use on a long-term outcome (GN-LTO): pneumonia in the year following new use of a proton-pump inhibitor (PPI) vs. non-use[15]; 4) Comparing the effect of two gastroprotective agents on a long-term outcome (GG-LTO): pneumonia in the year following new use of a PPI vs. histamine type-2 receptor antagonists (H2RA).[16] We chose to contrast empirical examples involving antibiotics vs. gastroprotective agents because the former tends to be used for a shorter duration, which may be relevant when assessing the impact of EHR continuity. New use was defined as having a medication record for the drug of interest without any use in the preceding 365 days in the EHR. Non-user of PPI cohort was established by risk-set sampling[17] a non-user for each PPI user, matched on the calendar date. In each example, the cohort entry date (CED) was the day of the medication start or the risk-set sampling date (Figure 1).

Exposure and outcome definition:

Medication use was determined based on the prescribing and dispensing and medication reconciliation data available in the EHR (Table S1) to ensure the comparisons between study cohorts identified based on EHR alone vs. linked EHR-claims data were performed in the same populations. The outcome definitions were based on the International Classification of Diseases (ICD) diagnosis codes recorded in the EHR using outcome definitions validated in the literature when available.[18–19] The hyperkalemia outcome was defined by the presence of an ICD diagnosis code or laboratory results because a prior study has suggested that a definition relying on coded diagnosis alone underestimates the clinically evident hyperkalemia (Table S2).

Algorithm-predicted EHR-continuity:

We used a previously validated algorithm to predict EHR-continuity on a yearly basis, which has been shown to be highly correlated with measured EHR-continuity and the degree of misclassification of study variables in the external set [11,20] (Table S3). The model predictors of the EHR-continuity are mainly indicators related to primary care follow-up in the study EHR, including (a) codes for a routine-care office visit; (b) preventive interventions or screening tests; (c) recording of diagnoses or medications in the EHR; (d) presence and numbers of certain types of encounters in the EHR; and (e) seeing the same provider repeatedly in the system (Table S4).

Covariates:

The pre-exposure covariate assessment period was 365 days before (and including) the cohort entry date. We assessed the following covariates: 1) demographic variables: age, sex, race, and ethnicity; 2) co-morbidities: coronary artery disease, venous thromboembolism, hypertension, diabetes, hyperlipidemia, atherosclerosis, heart failure, stroke, stroke, myocardial infarction, gastrointestinal and other bleeds, peripheral vascular disease, liver/kidney diseases, dementia); 3) prior medication use: aspirin, other antiplatelet agents, nonsteroidal anti-inflammatory drugs, anticoagulants, antihypertensive agents, antiarrhythmics, statins, antidiabetics, acid suppressants; 4) healthcare use variables: number of medications, hospitalizations, hospital days, and office visits (see detailed definition of each covariate in Table S5).

Statistical analysis:

To select covariates for adjustment for each empirical example, we entered a total of 70 variables described above in the least absolute shrinkage and selection operator (LASSO) for each example in relation to the study-specific outcome.[21] We then built multivariate-adjusted Cox proportional hazards models that included the LASSO-selected covariates[22] to estimate the hazard ratio (HR) and the 95% confidence intervals (CI). We included 12 covariates in the A-STO, 8 in the A-LTO, 20 in the GN-LTO, and 23 in the GG-LTO examples (see list of covariates in each example in Table S6). We used an as-started follow-up model akin to intention-to-treat (ITT) analysis and patients were followed until the earliest of the following: 1) loss of Medicare coverage; 2) death; 3) 2014/12/31, the end of the study period. All the statistical analyses were conducted with SAS 9.4 (SAS Institute Inc., Cary, NC).

Performance evaluation:

The objective of this study is not to assess the causal effect of drug effect on the outcomes but to quantify the discrepancy between the estimates based on EHR vs. linked EHR-claims data. Therefore, the bias of interest is defined as "*the deviation of the estimates based on EHR alone from that based on linked EHR-claims data (to assess outcome and covariates),*" since claims data capture medical information recorded outside of study EHR. Comparing HR based on only EHR (HR_{EHR}) with the linked EHR-claims data ($HR_{EHR+claims}$), we calculated the proportion of the outcome events captured by the study EHR and %bias on the logarithmic scale: %bias = Exp {absolute value [LN (HR_{EHR}) – LN ($HR_{EHR+claims}$)]}*100%

- 100%. We *a priori* specified %bias < 10% as acceptable performance. Within each EHR system, we calculated these metrics by quartiles of predicted EHR-continuity. To assess the presence of treatment effect heterogeneity by EHR-continuity (i.e., different effect estimates in the subgroups defined by EHR-continuity), we calculated the ratio of adjusted $HR_{EHR+claims,}$ comparing each quartile to the top quartile of EHR-continuity (e.g., $HR_{EHR+claims}$ in patients with lower 25% of EHR-continuity divided by $HR_{EHR+claims}$ in patients with top 25% of EHR-continuity). We tested the presence of interaction by a product term between the EHR-continuity subgroups and the treatment variable. The study was reviewed and approved by the Institutional Review Board (IRB) of the Brigham and Women's Hospital (IRB protocol number: 2017P002659).

Results:

Study population:

Our study cohort included a total of 77,288 patients in system 1 and 60,309 patients in system 2. In system 1, we identified 6,404 trimethoprim/sulfamethoxazole (mean age 76.5±7.6, 63.7% female), 5,339 cephalexin (mean age 76.6±7.7, 58.3% female), 28,657 PPI (mean age 76.3±75.9, 59.2% female), 28,801 PPI non-users (mean age 75.8±7.3, 59.3% female), and 8,069 H2RA new users (mean age 75.9±7.5, 62.2% female). In system 2, we identified 4,436 trimethoprim/sulfamethoxazole (mean age 75.2±7.2, 64.3% female), 4,524 cephalexin (mean age 75.8±7.3, 59.2% female), 22,442 PPI (mean age 75.5±7.2, 62.4% female), 22,529 PPI non-users (mean age 74.9±6.9, 64.7% female), and 6,378 H2RA (mean age 75.2±7.1, 63.7% female) new users (Table 1.; cohort formation in Table S7).

Comparing outcome events captured by predicted EHR-continuity:

We observed a decreasing trend in the proportion of total outcome events captured by the study EHR from the highest to lowest predicted EHR quartiles across empirical examples in both systems 1 and 2 (Table 1*, $p<0.001$ for all examples*). The trend was more pronounced for the long-term versus short-term outcomes. For example, comparing the top vs. lowest quartile, the proportion of outcome events captured by EHR went from 94% to 81% for A-STO and from 70% to 21% for A-LTO. A similar trend was observed in system 2 (Table 1*, $p<0.001$ for all examples*).

Comparing incidence rates (IR) by predicted EHR-continuity:

Compared to using EHR+claims data, we found that EHR data alone consistently underestimated IRs. The underestimation of the IRs was more severe for patients with lower than higher predicted EHR-continuity and more pronounced for the long-term than short-term outcomes. For example, for A-STO in system 1, the IR per 100 person-year (PY) in the exposed group was 27.12 based on EHR only vs. 29.04 based on EHR+claims data in the top quartile of EHR-continuity. The corresponding IRs were 20.24 based on EHR only vs. 23.65 based on EHR+claims data in the lowest quartile of EHR-continuity. In contrast, for A-LTO in system 1, the corresponding IRs in the exposed group were 1.57 based on EHR only vs. 2.37 based on EHR+claims data in the top quartile of EHR-continuity, and 0.44 vs. 1.89 in the lowest quartile of EHR-continuity (Table 1). A similar trend was observed

in system 2 (Table 1, *p<0.001 for decreasing IR based on EHR alone by EHR-continuity quartiles for all examples in both systems*).

Comparing %bias by predicted EHR-continuity:

Comparing HRs based on EHR only to those based on EHR+claims data, we found that %bias was consistently smaller in patients with higher EHR-continuity in both crude (Table 1) and adjusted analysis (Figure 2&3, *p<0.001 for increasing bias% in the lower EHR-continuity quartiles for all examples in both systems*). This trend appeared to be more pronounced for examples with long-term outcomes. For example, comparing the adjusted HR in the top vs. lowest quartile of EHR-continuity in system 1, the %bias was 2% vs. 6% for A-STO and 10% vs. 62% for A-LTO (Figure 2). The %bias appeared to be more evident for non-user comparison (GN-LTO) than for active comparison (GG-LTO). For example, comparing the adjusted HR in the top vs. lowest quartile of EHR-continuity in system 2, the %bias was 4% vs 35% for GN-LTO and 5% vs. 13% for GG-LTO (Figure 3).

Representativeness of estimates in patients with high vs. low EHR-continuity:

Based on ratios of adjusted $HR_{EHR+claims}$ comparing each quartile to the top quartile of EHR-continuity, we did not find significant treatment effect heterogeneity by EHR continuity for most of the EHR-continuity subgroups across empirical examples. Among the 24 interaction comparisons, only in two comparisons for GG-LTO did we observe borderline significant associations (Ratio=0.81–0.83, Table 2).

Discussion:

Based on two academic EHR systems in the metropolitan Boston area, we evaluated a previously developed algorithm to identify patients with high EHR-continuity in four real-world evidence studies comparing the effects of antibiotics and gastroprotective agents in relation to short-term and long-term outcomes. We found that analyses in patients in the lower 25–50% of predicted EHR-continuity substantially under-captured outcome events and under-estimated their incidence. Our findings suggest that patients with low predicted EHR-continuity contributes relatively few outcome events to the study (as compared to the total number of events that these patient experience based on the claims data) and the information that they do add may be unreliable. We did not find statistically significant treatment effect heterogeneity by EHR-continuity for most subgroups across empirical examples.

Our finding needs to be interpreted in context. Our results were based on only four examples that considered the intended duration of medication use and the immediacy of the outcome occurrence. Further testing in a wide variety of studies may be warranted before generalizing our findings to other research questions. Also, we used two urban academic EHR systems. While the EHR-continuity algorithm was also validated in another EHR system,[20] there could be other EHR systems with different data availability or structure that can affect EHR-based RWE studies. Also, our study population included only patients aged 65 years or older. As medical-seeking behavior may differ by age group, the findings may not be generalizable to a younger population. Besides, in patients with the lower 2

quartiles of EHR-continuity, the estimates based on EHR alone are imprecise with wide confidence intervals due to under-capturing of the outcome events in the EHR, partially accounting for the observed increased discrepancies between estimates based on EHR alone vs. EHR plus claims. Because EHR-based estimates in both subgroups are highly imprecise, random variability could explain why %bias is greatest in the 3rd rather than 4th quartile of EHR-continuity in some examples. Taken together, we recommend viewing our results as descriptive rather than prescriptive and that investigators and decision-makers use caution when generalizing to a different EHR setting or research question. It is also important to note that the objective of this study is not to assess the medications' causal effects on the outcomes but to quantify the discrepancy between the estimates based on EHR vs. linked EHR-claims data. Therefore, the results on the medications' effects on the clinical outcomes should not be overinterpreted.

We observed a pattern that the information bias due to EHR-discontinuity appears more pronounced for long-term (e.g., assessed over a year) rather than short-term outcomes (e.g., evaluated in the first 30 days). In patients in the lowest quartile of predicted EHR-continuity, the proportion of outcomes captured by EHR data was 72–81% for the short-term outcome and 21–49% for the long-term outcome. When designing an EHR-based CER study, it is important to consider the "observability" of the outcome in the study database. For example, when assessing the effect of an inpatient medication on short-term outcomes observable within the index admission (e.g., inpatient mortality, transfer to an intensive care unit, or respiratory failure requiring mechanical ventilation)[24], the EHR will be less susceptible to information bias due to EHR-discontinuity. However, continuity should be considered for longer-term outcomes.

Our findings also suggest that the information bias due to EHR-discontinuity is more pronounced for the non-use comparison than an active comparator design.[25] In comparative effectiveness research, an active comparator design is often recommended to improve confounding adjustment,[26] although finding a clinically meaningful comparator is not always feasible. The criterion of having a medication initiation at the index date in both comparison arms by design requires each study participant to have an EHR medication record at cohort entry, making it more likely that follow-up visits will be observable in the same system (since the physicians prescribing the medications may more likely have a subsequent encounter to follow up the treatment effects). Such an EHR-continuity enhancement is not expected in the non-user group of a non-user comparison unless the investigator explicitly requires that patients in the non-user group also have a medical encounter on the cohort entry date. Therefore, researchers need to pay close attention to potential bias due to EHR-discontinuity when comparing a treatment with non-use based on only EHR.

It is important to consider the generalizability of the findings when restricting the study cohort to those with high EHR-continuity. The algorithm to predict EHR-continuity mainly includes indicators related to primary care follow-up in the study EHR. [11,20] It is possible that the algorithm could identify a sub-cohort that overrepresents patients with higher medical complexity. However, we found no statistical evidence of effect modification by EHR-continuity quartiles. It could indicate that that the estimates obtained in those with high

EHR-continuity can be representative of that of the general population with all available data (EHR plus claims data). Moreover, some EHR systems may capture larger or smaller proportions of patients' overall care. For example, an integrated delivery system may have less overall discontinuity as compared to a single academic medical center such that patients in the lower quartiles of predicted EHR-continuity may still have relatively complete capture in the EHR data. While we focused on quartiles and two academic EHR systems in the Boston areas, other thresholds for discontinuity may be relevant in other settings. Another limitation of this study is that the study EHR research database does not contain reliable information on medication days or quantity supply, so we cannot perform "as-treated" analyses based on empirical duration of treatment.

In conclusion, in EHR-based RWE studies, analyses among patients with low EHR-continuity tend to substantially underestimate the incidence of the outcomes. Investigators may consider excluding patients with lower algorithm-predicted EHR-continuity as the EHR data capture a relatively small proportion of outcome events for these patients, and what little statistical information these patients contribute may be unreliable. Such exclusion does not substantially affect the generalizability of the results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding:

This project was supported by NIH Grant R01LM012594

Conflict of interest disclosure: Dr. Schneeweiss participates in investigator-initiated grants to the Brigham and Women's Hospital from Bayer, Vertex, and Boehringer Ingelheim, unrelated to the topic of this study. He is a consultant to Aetion Inc., a software manufacturer of which he owns equity. His interests were declared, reviewed, and approved by the Brigham and Women's Hospital and Mass General Brigham System in accordance with their institutional compliance policies. Dr. Gagne is currently an employee of Johnson & Johnson, Inc. All other authors have no conflict of interest to disclose.

References:

1. Avorn J Powerful medicines : the benefits, risks, and costs of prescription drugs, (Vintage Books, New York, 2005).
2. Strom BL, Kimmel SE & Hennessy S Textbook of pharmacoepidemiology, (Wiley Blackwell, Chichester, West Sussex England; Hoboken, NJ, 2013).
3. Schneeweiss S & Avorn J A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 58, 323–337 (2005). [PubMed: 15862718]
4. Randhawa GS Building electronic data infrastructure for comparative effectiveness research: accomplishments, lessons learned and future steps. J Comp Eff Res 3, 567–572 (2014). [PubMed: 25494561]
5. Corley DA, Feigelson HS, Lieu TA & McGlynn EA Building Data Infrastructure to Evaluate and Improve Quality: PCORnet. J Oncol Pract 11, 204–206 (2015). [PubMed: 25980016]
6. Desai RJ, et al. Development and Preliminary Validation of a Medicare Claims-Based Model to Predict Left Ventricular Ejection Fraction Class in Patients With Heart Failure. Circ Cardiovasc Qual Outcomes 11, e004700 (2018). [PubMed: 30562067]
7. Desai RJ, et al. Effectiveness of angiotensin-neprilysin inhibitor treatment versus renin-angiotensin system blockade in older adults with heart failure in clinical care. Heart (2021).

8. Lin KJ, et al. Prediction Score for Anticoagulation Control Quality Among Older Adults. J Am Heart Assoc 6(2017).
9. Lin KJ, et al. Out-of-system Care and Recording of Patient Characteristics Critical for Comparative Effectiveness Research. Epidemiology 29, 356–363 (2018). [PubMed: 29283893]
10. Hennessy S Use of health care databases in pharmacoepidemiology. Basic & clinical pharmacology & toxicology 98, 311–313 (2006). [PubMed: 16611207]
11. Lin KJ, et al. Identifying Patients With High Data Completeness to Improve Validity of Comparative Effectiveness Research in Electronic Health Records Data. Clin Pharmacol Ther 103, 899–905 (2018). [PubMed: 28865143]
12. Antoniou T, et al. Trimethoprim-sulfamethoxazole induced hyperkalaemia in elderly patients receiving spironolactone: nested case-control study. Bmj 343, d5228 (2011). [PubMed: 21911446]
13. Antoniou T, et al. Trimethoprim-sulfamethoxazole-induced hyperkalemia in patients receiving inhibitors of the renin-angiotensin system: a population-based study. Archives of internal medicine 170, 1045–1049 (2010). [PubMed: 20585070]
14. Lo Vecchio A & Zacur GM Clostridium difficile infection: an update on epidemiology, risk factors, and therapeutic options. Curr Opin Gastroenterol 28, 1–9 (2012). [PubMed: 22134217]
15. Laheij RJ, et al. Risk of community-acquired pneumonia and use of gastric acid-suppressive drugs. Jama 292, 1955–1960 (2004). [PubMed: 15507580]
16. Wang Y, et al. Efficacy and safety of gastrointestinal bleeding prophylaxis in critically ill patients: an updated systematic review and network meta-analysis of randomized trials. Intensive Care Med 46, 1987–2000 (2020). [PubMed: 32833040]
17. Goldstein L & Langholz B Risk set sampling in epidemiologic cohort studies. Statistical Science 11, 35–53, 19 (1996).
18. Raebel MA, et al. The positive predictive value of a hyperkalemia diagnosis in automated health care data. Pharmacoepidemiol Drug Saf 19, 1204–1208 (2010). [PubMed: 20878650]
19. Drahos J, Vanwormer JJ, Greenlee RT, Landgren O & Koshiol J Accuracy of ICD-9-CM codes in identifying infections of pneumonia and herpes simplex virus in administrative data. Annals of Epidemiology 23, 291–293 (2013). [PubMed: 23522903]
20. Lin KJ, et al. External Validation of an Algorithm to Identify Patients with High Data-Completeness in Electronic Health Records for Comparative Effectiveness Research. Clin Epidemiol 12, 133–141 (2020). [PubMed: 32099479]
21. Tibshirani R The lasso method for variable selection in the Cox model. Stat Med 16, 385–395 (1997). [PubMed: 9044528]
22. Cox DR Regression models and life tables (with discussion). Journal of the Royal Statistical Society [B] 34, 187–220 (1972).
23. Armitage P Tests for Linear Trends in Proportions and Frequencies. Biometrics 11, 375–386 (1955).
24. Lin KJ, et al. Pharmacotherapy for Hospitalized Patients with COVID-19: Treatment Patterns by Disease Severity. Drugs 80, 1961–1972 (2020). [PubMed: 33151482]
25. Rassen JA, Murk W & Schneeweiss S Real-world evidence of bariatric surgery and cardiovascular benefits using electronic health records data: A lesson in bias. Diabetes Obes Metab (2021).
26. Johnson ES, et al. The incident user design in comparative effectiveness research. Pharmacoepidemiology and drug safety 22, 1–6 (2013).

Study highlights

What is the current knowledge on the topic?

Electronic health records (EHR) discontinuity (i.e., receiving care outside of an EHR) can lead to a substantial amount of information bias in EHR-based comparative effectiveness research (CER)

What question did this study address?

What is the impact of algorithm-predicted EHR-continuity on estimates in 4 CER examples?

What does this study add to our knowledge?

We found that analyses in patients in the lower predicted EHR-continuity substantially under-captured outcome events and under-estimated their incidence. Our findings suggest that patients with low predicted EHR-continuity contribute relatively few outcome events to the study, and the information that they do add may be unreliable. We did not find significant treatment effect heterogeneity by EHR-continuity across empirical examples.

How might this change clinical pharmacology or translational science?

Investigators may consider excluding patients with low algorithm-predicted EHR-continuity as the EHR data capture relatively few of their actual outcomes, and treatment effect estimates in these patients may be unreliable.

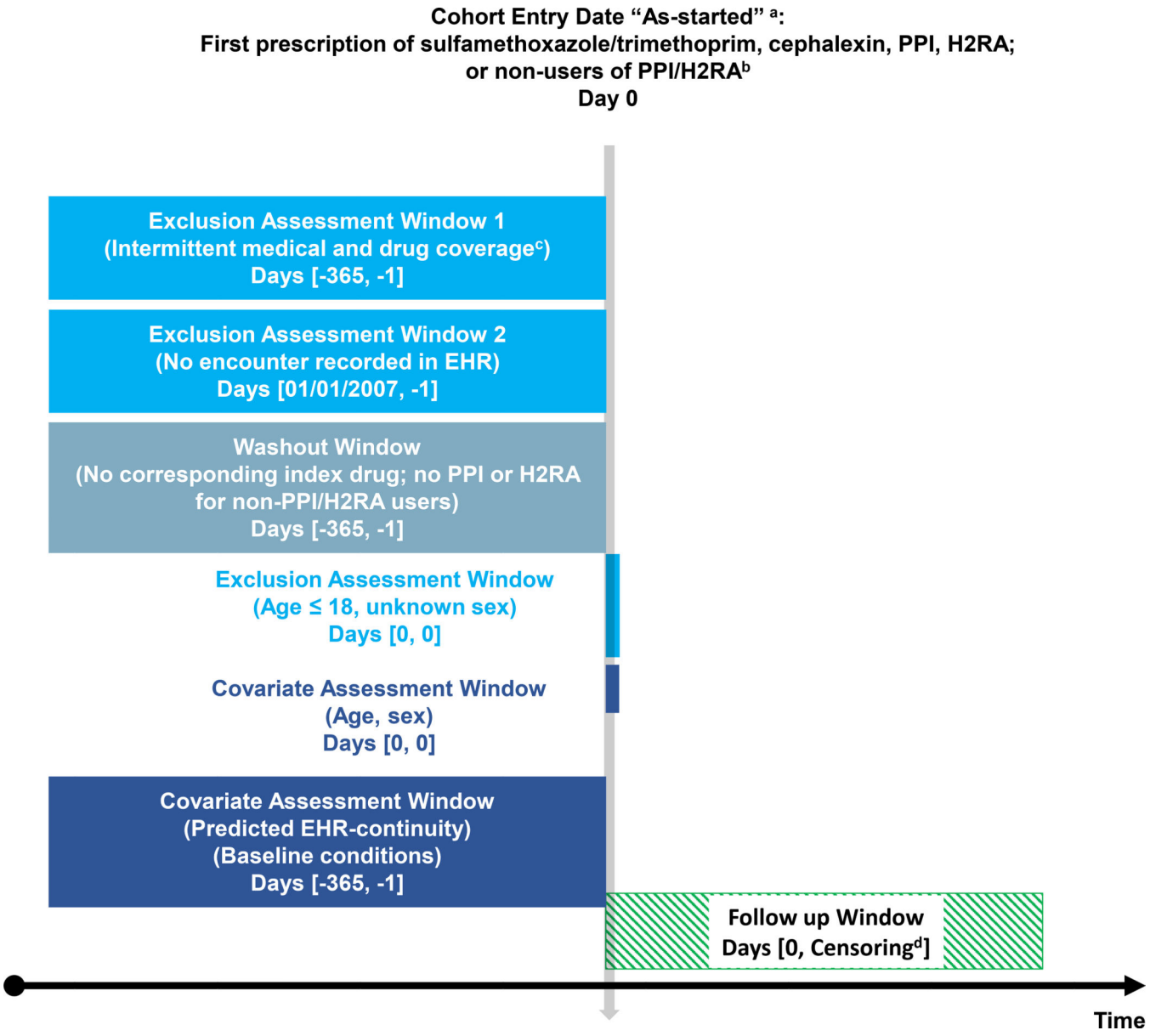

Figure 1. Study design diagram

a. As-started: treatment classification as the first drug exposure for the pre-specified duration

b. Non-PPI/H2RA users were 1:1 matched to PPI users by random sampling from the patient pool who met other inclusion/exclusion criteria.

c. Up to 31-day gaps in medical or pharmacy enrollment allowed

d. Earliest of outcome of interest, death, disenrollment, 365 days of follow-up, end of the study period

EHR = electronic health records

PPI = Proton pump inhibitor

H2RA = Histamine type-2 receptor antagonists

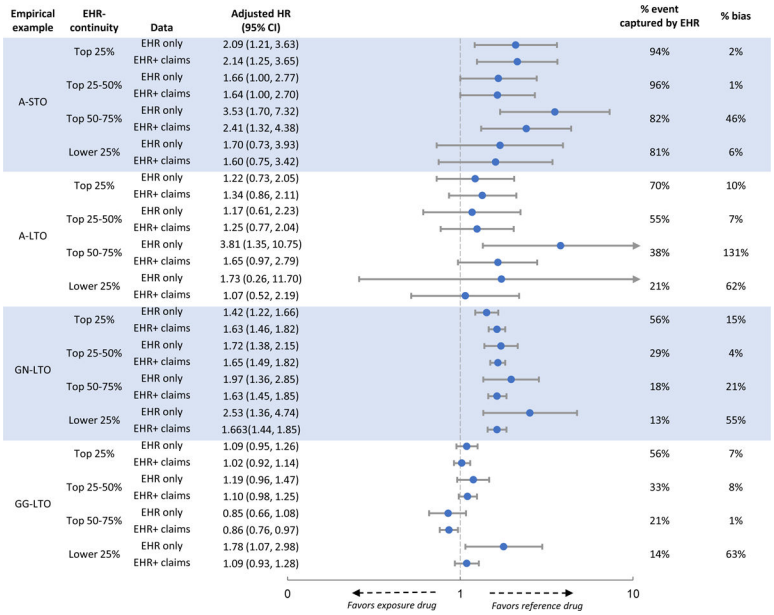

Figure 2. Comparison between patients with high vs. low EHR-continuity in system 1

EHR= electronic health records, HR= adjust hazard ratio, CI= confidence interval, Ref.= referent group

A-STO: comparing the effect of two Antibiotics on a short-term outcome

A-LTO: comparing the effect of two Antibiotics effect on a long-term outcome

GN-LTO: Comparing the effect of a Gastroprotective agent vs. non-use on a long-term outcome

GG-LTO: Comparing the effect of two Gastroprotective agents on a long-term outcome

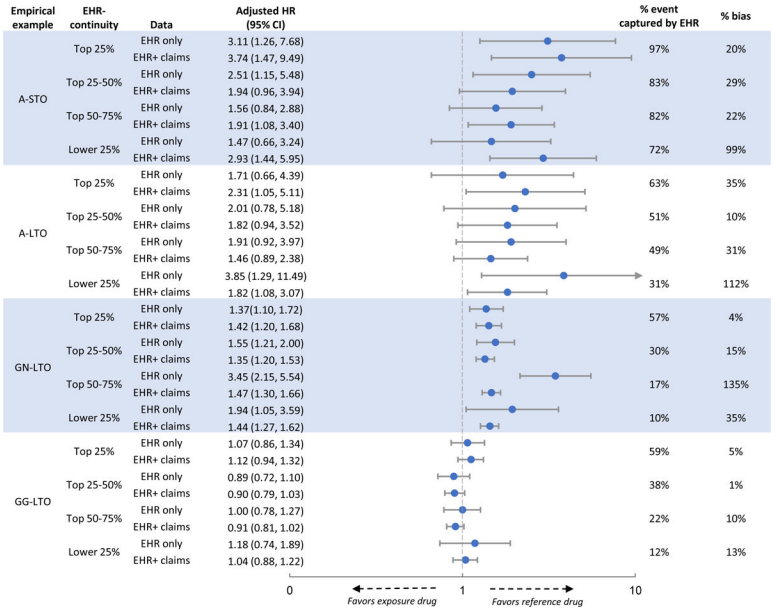

Figure 3. Comparison between patients with high vs. low EHR-continuity in system 2

EHR= electronic health records, HR= adjust hazard ratio, CI= confidence interval, Ref.= referent group

A-STO: comparing the effect of two Antibiotics on a short-term outcome

A-LTO: comparing the effect of two Antibiotics effect on a long-term outcome

GN-LTO: Comparing the effect of a Gastroprotective agent vs. non-use on a long-term outcome

GG-LTO: Comparing the effect of two Gastroprotective agents on a long-term outcome

Clin Pharmacol Ther. Author manuscript; available in PMC 2023 January 01.

Table 1

Comparing the study cohort with high vs. low EHR-continuity (unadjusted analyses)

Hospital	Empirical example	EHR-continuity	Data	Exposure, event #/IR per 100 PY	Reference, event #/IR per 100 PY	Crude HR (95% CI)	% outcome event captured by EHR	% bias
System 1	A-STO	Top 25%	EHR only	43/27.12	16/12.27	2.20 (1.27, 3.84)	94%	0%
			EHR+ claims	46/29.04	17/13.04	2.22 (1.30, 3.81)		
		Top 25–50%	EHR only	47/32.16	21/17.52	1.83 (1.14, 2.94)	96%	1%
			EHR+ claims	49/33.52	22/18.36	1.82 (1.14, 2.90)		
		Top 50–75%	EHR only	46/39.87	<11/8.53	4.63 (2.22, 9.68)	82%	41%
			EHR+ claims	53/46.08	13/13.89	3.29 (1.81, 5.98)		
		Lower 25%	EHR only	18/20.24	<11/9.23	2.19 (0.95, 5.02)	81%	18%
			EHR+ claims	21/23.65	11/12.72	1.85 (0.89, 3.84)		
	A-LTO	Top 25%	EHR only	26/1.57	18/1.31	1.20 (0.72, 2.01)	70%	13%
			EHR+ claims	39/2.37	24/1.75	1.35 (0.86, 2.12)		
		Top 25–50%	EHR only	22/1.50	13/1.06	1.40 (0.74, 2.66)	55%	6%
			EHR+ claims	41/2.80	23/1.88	1.48 (0.90, 2.42)		
		Top 50–75%	EHR only	21/1.79	<11/0.41	4.31 (1.61, 11.51)	38%	145%
			EHR+ claims	45/3.89	21/2.19	1.76 (1.08, 2.89)		
		Lower 25%	EHR only	<11/0.44	<11/0.23	1.95 (0.36, 10.68)	21%	41%
			EHR+ claims	17/1.89	12/1.36	1.39 (0.66, 2.90)		
	GN-LTO	Top 25%	EHR only	730/9.97	299/4.22	2.34 (2.05, 2.68)	56%	11%
			EHR+ claims	1329/18.98	503/7.19	2.59 (2.33, 2.87)		
		Top 25–50%	EHR only	427/6.93	168/2.41	2.83 (2.37, 3.38)	29%	6%
			EHR+ claims	1410/24.86	612/9.04	2.68 (2.43, 2.94)		
		Top 50–75%	EHR only	173/4.19	78/1.25	3.29 (2.53, 4.28)	18%	20%
			EHR+ claims	880/23.26	497/8.19	2.75 (2.46, 3.07)		
		Lower 25%	EHR only	146/2.83	32/0.68	4.04 (2.73, 5.98)	13%	66%
			EHR+ claims	1015/21.29	381/8.40	2.44 (2.16, 2.76)		
	GG-LTO	Top 25%	EHR only	699/10.52	244/10.58	1.00 (0.87, 1.14)	56%	1%
			EHR+ claims	1245/19.59	438/19.85	0.99 (0.90, 1.10)		

Hospital	Empirical example	EHR-continuity	Data	Exposure, event #/IR per 100 PY	Reference, event #/IR per 100 PY	Crude HR (95% CI)	% outcome event captured by EHR	% bias
		Top 25–50%	EHR only	390/6.81	106/6.23	1.10 (0.89, 1.35)	33%	2%
			EHR+ claims	1189/22.26	330/20.67	1.08 (0.96, 1.21)		
		Top 50–75%	EHR only	238/4.63	76/6.08	0.77 (0.60, 0.98)	21%	8%
			EHR+ claims	1167/25.01	340/30.50	0.83 (0.74, 0.93)		
		Lower 25%	EHR only	149/2.83	16/1.49	1.90 (1.15, 3.14)	14%	61%
			EHR+ claims	1033/21.26	180/17.95	1.18 (1.01, 1.37)		
System 2	A-STO	Top 25%	EHR only	23/31.35	<11/10.33	3.03 (1.28, 7.13)	97%	4%
			EHR+ claims	24/32.74	<11/10.34	3.16 (1.35, 7.41)		
		Top 25–50%	EHR only	21/23.24	<11/9.01	2.56 (1.17, 5.62)	83%	20%
			EHR+ claims	24/26.61	11/12.41	2.13 (1.07, 4.26)		
		Top 50–75%	EHR only	36/41.35	17/18.03	2.27 (1.29, 4.00)	82%	24%
			EHR+ claims	47/54.27	18/19.10	2.81 (1.65, 4.80)		
		Lower 25%	EHR only	24/24.24	<11/8.74	2.74 (1.31, 5.72)	72%	37%
			EHR+ claims	36/36.58	11/9.62	3.75 (1.91, 7.38)		
	A-LTO	Top 25%	EHR only	11/1.39	<11/0.83	1.67 (0.66, 4.25)	63%	29%
			EHR+ claims	19/2.40	<11/1.11	2.16 (0.98, 4.79)		
		Top 25–50%	EHR only	13/1.41	<11/0.64	2.17 (0.86, 5.43)	51%	17%
			EHR+ claims	24/2.61	13/1.39	1.85 (0.96, 3.58)		
		Top 50–75%	EHR only	22/2.60	11/1.10	2.28 (1.11, 4.71)	49%	30%
			EHR+ claims	41/4.87	27/2.73	1.75 (1.08, 2.82)		
		Lower 25%	EHR only	15/1.53	<11/0.33	4.60 (1.53, 13.79)	31%	128%
			EHR+ claims	38/3.93	23/1.88	2.02 (1.21, 3.39)		
	GN-LTO	Top 25%	EHR only	283/7.62	120/3.32	2.29 (1.85, 2.85)	57%	3%
			EHR+ claims	494/13.68	218/6.11	2.23 (1.90, 2.62)		
		Top 25–50%	EHR only	307/6.12	108/2.16	2.79 (2.24, 3.48)	30%	30%
			EHR+ claims	929/19.71	434/8.96	2.15 (1.92, 2.41)		
		Top 50–75%	EHR only	184/4.79	58/1.10	4.25 (3.15, 5.75)	17%	79%
			EHR+ claims	881/25.05	518/10.17	2.37 (2.12, 2.64)		
		Lower 25%	EHR only	117/2.27	28/0.50	4.41 (2.88, 6.76)	10%	100%

Hospital	Empirical example	EHR-continuity	Data	Exposure, event #/IR per 100 PY	Reference, event #/IR per 100 PY	Crude HR (95% CI)	% outcome event captured by EHR	% bias
			EHR+ claims	970/20.37	479/8.94	2.21 (1.97, 2.46)		
	GG-LTO	Top 25%	EHR only	260/8.00	117/9.10	0.88 (0.72, 1.09)	59%	0%
			EHR+ claims	439/13.90	198/15.88	0.88 (0.76, 1.03)		
		Top 25–50%	EHR only	262/6.08	117/8.75	0.70 (0.58, 0.86)	38%	11%
			EHR+ claims	715/17.47	285/22.70	0.78 (0.69, 0.89)		
		Top 50–75%	EHR only	242/4.94	83/6.33	0.79 (0.63, 0.99)	22%	5%
			EHR+ claims	1111/24.74	354/30.00	0.83 (0.75, 0.93)		
		Lower 25%	EHR only	127/2.41	19/1.82	1.32 (0.82, 2.12)	12%	11%
			EHR+ claims	1009/20.75	169/17.30	1.19 (1.01, 1.39)		

EHR= electronic health records, IR=incidence rate, PY=person-years, HR = hazard ratio, CI= confidence interval, Ref= referent group

A-STO: comparing the effect of two Antibiotics on a short-term outcome

A-LTO: comparing the effect of two Antibiotics effect on a long-term outcome

GN-LTO: Comparing the effect of a Gastroprotective agent vs. non-use on a long-term outcome

GG-LTO: Comparing the effect of two Gastroprotective agents on a long-term outcome

Table 2

Treatment effect heterogeneity by EHR-continuity

Hospital	Empirical example	EHR-continuity	$HR_{EHR+claims}$ (95% CI)	Ratio of $HR_{EHR+claims}$ (95% CI) *	p for interaction**
System 1	A-STO	Top 25%	2.14 (1.25,3.65)	Ref	ref
		Top 25–50%	1.64 (1.00,2.70)	0.76 (0.37,1.56)	0.4612
		Top 50–75%	2.41 (1.32,4.38)	1.23 (0.55,2.74)	0.6213
		Lower 25%	1.60 (0.75,3.42)	0.68 (0.28,1.69)	0.4094
	A-LTO	Top 25%	1.34 (0.86,2.11)	Ref	ref
		Top 25–50%	1.25 (0.77,2.04)	0.97 (0.50,1.88)	0.9319
		Top 50–75%	1.65 (0.97,2.79)	1.07 (0.55,2.10)	0.8418
		Lower 25%	1.07 (0.52,2.19)	0.88 (0.37,2.10)	0.7785
	GN-LTO	Top 25%	1.63 (1.46,1.82)	Ref	ref
		Top 25–50%	1.65 (1.49,1.82)	0.96 (0.83,1.11)	0.5911
		Top 50–75%	1.63 (1.45,1.85)	0.99 (0.85,1.15)	0.8778
		Lower 25%	1.63 (1.44,1.85)	0.97 (0.83,1.14)	0.7162
	GG-LTO	Top 25%	1.02 (0.92,1.14)	Ref	ref
		Top 25–50%	1.10 (0.98,1.25)	1.10 (0.94,1.29)	0.2577
		Top 50–75%	0.86 (0.76,0.97)	0.83 (0.71,0.98)	0.027
		Lower 25%	1.09 (0.93,1.28)	1.09 (0.90,1.32)	0.3901
System 2	A-STO	Top 25%	3.74 (1.47,9.49)	Ref	ref
		Top 25–50%	1.94 (0.96,3.94)	0.60 (0.20,1.81)	0.3685
		Top 50–75%	1.91 (1.08,3.40)	0.66 (0.24,1.81)	0.4231
		Lower 25%	2.93 (1.44,5.95)	0.91 (0.31,2.70)	0.8687
	A-LTO	Top 25%	2.31 (1.05,5.11)	Ref	ref
		Top 25–50%	1.82 (0.94,3.52)	0.80 (0.28,2.24)	0.6663
		Top 50–75%	1.46 (0.89,2.38)	0.61 (0.24,1.55)	0.302
		Lower 25%	1.82 (1.07,3.07)	0.73 (0.28,1.87)	0.5102
	GN-LTO	Top 25%	1.42 (1.20,1.68)	Ref	ref
		Top 25–50%	1.35 (1.20,1.53)	0.93 (0.76,1.13)	0.4671
		Top 50–75%	1.47 (1.30,1.66)	0.94 (0.77,1.14)	0.5187
		Lower 25%	1.44 (1.27,1.62)	0.93 (0.76,1.14)	0.4888
	GG-LTO	Top 25%	1.12 (0.94,1.32)	Ref	ref
		Top 25–50%	0.90 (0.79,1.03)	0.81 (0.66,0.99)	0.0393
		Top 50–75%	0.91 (0.81,1.02)	0.83 (0.68,1.01)	0.0631
		Lower 25%	1.04 (0.88,1.22)	0.94 (0.75,1.18)	0.5792

EHR= electronic health records, $HR_{EHR+claims}$ = adjusted hazard ratio based on the linked EHR-claims data, CI= confidence interval, Ref= referent group

*Ratios of adjusted $HR_{EHR+claims}$ compared between top 25% vs. top 25–50%, 50–75%, and lower 25%of predicted EHR-continuity

** P testing for interaction between adjusted $HR_{EHR+claims}$ in top 25% vs. top 25–50%, 50–75%, and lower 25%of predicted EHR-continuity

A-STO: comparing the effect of two Antibiotics on a short-term outcome

A-LTO: comparing the effect of two Antibiotics effect on a long-term outcome

GN-LTO: Comparing the effect of a Gastroprotective agent vs. non-use on a long-term outcome

GG-LTO: Comparing the effect of two Gastroprotective agents on a long-term outcome