

RESEARCH

Open Access



# A strategy to identify event specific hospitalizations in large health claims databases

Joshua Lambert<sup>1\*</sup>, Harpal Sandhu<sup>2</sup>, Emily Kean<sup>1</sup>, Teenu Xavier<sup>1</sup>, Aviv Brokman<sup>3</sup>, Zachary Steckler<sup>3</sup>, Lee Park<sup>3</sup> and Arnold Stromberg<sup>3</sup>

## Abstract

**Background:** Health insurance claims data offer a unique opportunity to study disease distribution on a large scale. Challenges arise in the process of accurately analyzing these raw data. One important challenge to overcome is the accurate classification of study outcomes. For example, using claims data, there is no clear way of classifying hospitalizations due to a specific event. This is because of the inherent disjointedness and lack of context that typically come with raw claims data.

**Methods:** In this paper, we propose a framework for classifying hospitalizations due to a specific event. We then tested this framework in a private health insurance claims database (Symphony) with approximately 4 million US adults who tested positive with COVID-19 between March and December 2020. Our claims specific COVID-19 related hospitalizations proportion is then compared to nationally reported rates from the Centers for Disease Control by age.

**Results:** Across all ages (18+) the total percentage of Symphony patients who met our definition of hospitalized due to COVID-19 was 7.3% which was similar to the CDC's estimate of 7.5%. By age group, defined by the CDC, our estimates vs. the CDC's estimates were 18–49: 2.7% vs. 3%, 50–64: 8.2% vs. 9.2%, and 65+: 14.6% vs. 28.1%.

**Conclusions:** The proposed methodology is a rigorous way to define event specific hospitalizations in claims data. This methodology can be extended to many different types of events and used on a variety of different types of claims databases.

**Keywords:** Claims data, Classification, COVID-19, Methodology, COVID-19 research database

## Background

Prescription and health insurance claims providers can deliver unique patient-level retail pharmacy, diagnosis, and procedure data. These data can range in size and complexity depending on the provider [1–3]. Successfully using these data in medical research is not an easy task and requires some key considerations [1]. Some of this difficulty comes from the lack of structure and context to how certain International Classification

of Diseases (ICD)-10, Current Procedural Terminology (CPT), or drug codes are grouped with one another around an event of interest. For example, a hospitalization CPT code does not link to the diagnosis code that caused it, nor does the drug code that was prescribed because of the event. This disjointness is a major hurdle for researchers hoping to harness these large claims data for their research question of interest.

## Motivation

These issues arose in our own research, where we sought to use a large healthcare claims database called Symphony. Like many others, Symphony Health Solutions

\*Correspondence: lambejw@ucmail.uc.edu

<sup>1</sup> College of Nursing, University of Cincinnati, Procter Hall, Cincinnati, OH 45219, USA

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

includes data on retail pharmacy claims, medical claims, and readmittance claims.

Symphony Health is a leading provider of high-value data for biopharmaceutical manufacturers, healthcare providers, and payers. The company helps clients understand disease incidence, prevalence, progression, treatment, and influences along the patient and prescriber journeys by connecting and integrating a broad set of primary and secondary data. Symphony Health derived data improves health management decisions, and helps clients drive revenue growth while providing critical insights on how to effectively adapt to the changing healthcare ecosystem.

For each diagnosis, procedure, or prescription event, Symphony provided us with a patient ID, relevant code (e.g., ICD-10 code), and date of the event. However, it was not known whether a diagnosis for a patient is a primary, secondary, or subsequent diagnosis. No range or specific information about the date of the event was provided (example: no range of when hospitalization, or procedure occurred. Rather just a single date for each event). No enrollment criteria (example: at least 6 months continuous coverage) were used when Symphony constructed the database. With these obstacles in mind, we sought to use these data to reconstruct whether a patient who tested positive for COVID-19 was admitted to the hospital because of that diagnosis. We found that researchers like us (using Symphony Health Data), overcame this hurdle in a variety of ways.

### Literature review

To uncover how researchers deal with this lack of structure in claims data, a comprehensive literature search was conducted by a health sciences librarian (E. K.). EBSCOhost Academic Search Complete, Business Source Complete, CINAHL Plus with Full Text, MEDLINE with Full Text, and OmniFile Full Text Mega (H.W. Wilson) were searched from the dates of inception through August 2021. Additionally, the search consisted of a combination of keywords and equivalent subject headings representing “Symphony Health Solutions” as a company or a reference to the use of Symphony data. The results from the Symphony search were combined with a broad variety of terms representing the concepts of classification or categorization. An English language limit was applied to results, and after deduplication, 52 articles were retrieved.

Of the 52 articles retrieved, 32 articles were deemed relevant for this study. Of the 32 remaining articles, a clear pattern was identified as to how the authors chose to reconstruct the events of interest. Eighteen studies analyzed Symphony data using one timeline [4–21] to

reconstruct the event. For example, Hampp et al. [7] used the Symphony Health Solutions PHAST Prescription Monthly database to investigate the antidiabetic drug use in the US population during a predefined single timeline. Six studies looked at data with one timeline but multiple follow-ups within a time range [22–27]. Multiple timelines were used by eight of the retrieved papers [2, 28–34]. For example, Brixner et al. [2] conducted a longitudinal study using patient-level Symphony Health Solutions administrative claims data to assess the effectiveness of the HUMIRA Complete PSP in patients receiving adalimumab (ADA) treatment for broad range of diagnoses. They required the patients to have  $\geq 2$  claims which were at least 30 days apart to be included in the study.

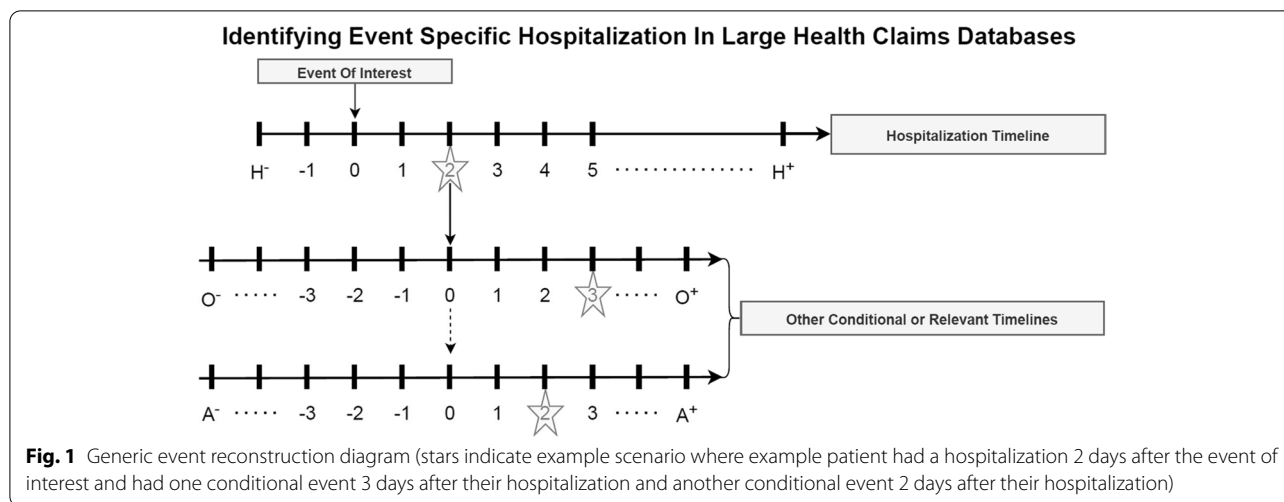
Using this past work, as well as our own personal experience with claims data, we developed a generic methodology to reconstruct event specific hospitalizations. This proposed methodology is meant to act as a guide for how researchers can utilize health claims data in a more rigorous way.

## Methods

### Event reconstruction strategy

Our event reconstruction strategy centers on the overlap of various event horizons (timelines) of interest. Specifically, when it comes to identifying event specific hospitalizations, the hospitalization event horizon is an important one to define. In Fig. 1, the hospitalization

horizon is defined as  $H^-$  to  $H^+$  where each endpoint is an integer ( $Z$ ). Event specific hospitalizations, by definition, must have the event of interest occur within the hospitalization horizon ( $H^-$ ,  $H^+$ ). Other relevant/conditional event horizons may be defined ( $O^-$  to  $O^+$  or  $A^-$  to  $A^+$ ) to sensitize the definition around the time of the hospitalization. The other event horizon(s) may act as validation event(s) which may come from a contextual understanding of the problem or a literature search. If a specific patient has the event of interest within the hospitalization horizon and, if necessary, the other relevant/conditional event horizons occur within the designated horizon around the hospitalization then the patient is said to have had the event specific hospitalization. All other patients can be thought of as not having the event. In Fig. 1 (as represented by stars) an example patient had a hospitalization 2 days after the event of interest and had one conditional event 3 days after their hospitalization and another conditional event 2 days after their hospitalization. Because this patient had events within the designated horizons ( $H^-$  to  $H^+$ ,  $O^-$  to  $O^+$  and  $A^-$  to  $A^+$ ) they are said to have the event specific hospitalization.



**Results**

The COVID-19 research database enables public health and policy researchers to use real-world data to better understand and combat the COVID-19 pandemic. In June 2021, via the COVID-19 research database, we gained access to the Symphony Health Data.

**Symphony data**

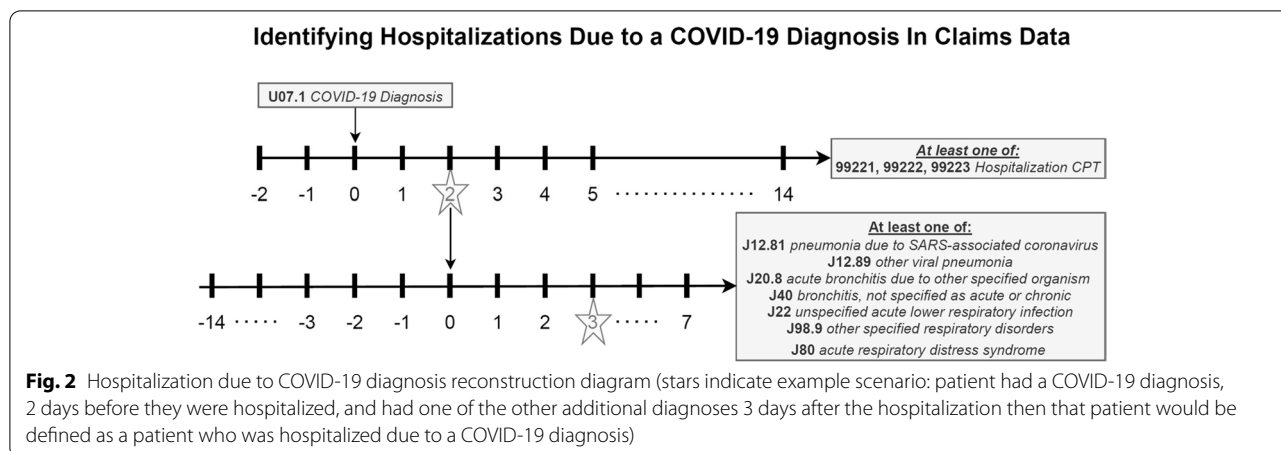
Our symphony data had approximately 4 million patients who tested positive for COVID-19 between 03/01/2020 and 12/31/2020. While we had data after 12/31/2020, our study focused on 2020 due to access to the Food and Drug Administration (FDA) Emergency Use Authorization (EUA) COVID-19 vaccines in early 2021 and beyond which we felt would deduct from our focused study of interest. Patient records are de-identified and minimal demographic information (Age, Sex, first two digits of the patient’s residential zip code) is known about the unique patients within the dataset. Patient level CPT, diagnosis, and prescription codes were available from late 2018 to mid-2021. Using these data, which are contained in different tables accessible via the snowflake SQL platform, we utilized our event reconstruction strategy where we intended to reconstruct which, of the 4 million patients who tested positive for COVID-19, were hospitalized due to the COVID-19 diagnosis.

**Hospitalization due to COVID-19 diagnosis reconstruction**

Using our generic definition, defined above, and outlined in Fig. 1, our clinical and research team decided the necessary event horizons endpoints. First, a hospitalization CPT code of at least one of 99,221, 99,222, or 99,223 needed to occur in the -2 to 14-day timeline from COVID diagnosis (U07.1). Within our claims

data diagnoses were not ranked as they are in some claims data, so the COVID diagnosis could have been any ranked diagnosis for the specific patient (principal, secondary, ...). The time lag and lead were determined as claims do not always mimic the actual timeline that the patient experienced. A sensitivity analysis of these endpoints.

showed that most of the diagnoses and hospitalizations that met our -2 to +14 criteria actually occurred very close (-1 to +1) to one another. If multiple hospitalization CPT codes or multiple COVID-19 diagnoses codes occurred for a specific patient, then the minimum distance between all possible combinations of diagnoses and hospitalizations were considered. As a tie breaker the earliest minimum combination which met our criteria was considered as the COVID-19 hospitalization for that specific patient. If one of the combinations met the criteria of -2 to +14 then the patient was said to have met the first part of the criteria for being classified as hospitalized due to a COVID-19 diagnosis. As a validation, patients needed to have at least one of a set of additional diagnoses which occur around the time (-14 to +7) of a hospitalization. This set was again, determined by our clinical and research team. These were: pneumonia due to SARS-associated coronavirus (J12.81), other viral pneumonia (J12.89), acute bronchitis due to other specified organism (J20.8), bronchitis not specified as acute or chronic (J40), unspecified acute lower respiratory infection (J22), other specified respiratory disorders (J98.9), or acute respiratory distress syndrome (J80). For example (See Fig. 2), if a patient had a COVID-19 diagnosis, 2 days before they were hospitalized, and had one of the other additional diagnoses 3 days after the hospitalization then that patient would be defined as a patient who was hospitalized due to a COVID-19 diagnosis.



**Comparison of proposed methodology and centre for disease control and prevention (CDC) estimates**

The CDC provides estimates [35] of the number of symptomatic COVID-19 illness and the aggregate number of hospitalizations from February 2020 through May 2021. Unfortunately, we were unable to attain estimates for the same time period under investigation in our study (March 1<sup>st</sup>, 2020, through December 31<sup>st</sup>, 2020). Our methodology was developed independently of the CDC data, and the CDC data is meant to provide a type of external validation to our methodology.

Table 1 provides a comparison of how our methodology on the Symphony data compares to the CDC’s population estimates. Using the proposed methodology on the Symphony data, 2.7% of the patients in the age group of 18–49 years old were hospitalized due to a COVID-19 diagnosis. The CDC’s estimated that 3% of 18–49 years old were hospitalized due to symptomatic COVID-19. In the age group of 50–64 years old, 8.2% of the Symphony patients were hospitalized as compared to the 9.2% estimated by the CDC. In the age group of 65+ years old, 14.6% of the Symphony patients were hospitalized as compared to the 28.1% estimated by the CDC. Across all age groups the total percentage

of Symphony patients hospitalized due to COVID-19 was 7.3% which is similar to the estimate of 7.5% by the CDC.

The difference observed in the 65+ population is worrisome and warrants serious attention. Considering Medicaid claims data are not included in our Symphony data, this discrepancy is less surprising. In the United States, those who are 65+ years old qualify for Medicare and may not be privately insured. Also, those patients who are only insured by Medicaid or not insured at all would not be captured by the Symphony data. One explanation as to why so few 65+ in the Symphony data are not hospitalized due to a COVID-19 diagnosis (as compared to the CDC’s estimates) is that they may be wealthier (can afford private insurance) and are therefore healthier on average.

While our sample is older on average than the general population, our methodology’s overall estimates for the percent of COVID-19 diagnoses who were hospitalized due to a COVID-19 diagnosis is close to the CDC’s overall estimates (7.3% vs. 7.5%). The main place to compare our methodology to the CDC estimates is in the portion of the US adult population who is likely privately insured (18–64 years). Within this group, our

**Table 1** Comparison of symphony data and CDC estimates

Age Groups	CDC Estimated COVID-19 Cases	Symphony Estimated COVID-19 Cases	CDC Estimated Hospitalizations	Symphony Estimated COVID-19 Hospitalizations	CDC Estimated Percentage of Hospitalizations <sup>(a)</sup>	Symphony Estimated Percentage of Hospitalizations <sup>(b)</sup>
18–49 years	51,581,445	1,867,749	1,533,679	49,638	3.0%	2.7%
50–64 years	17,377,602	1,045,133	1,604,612	85,271	9.2%	8.2%
65+ years	10,005,696	1,061,390	2,808,089	155,439	28.1%	14.6%
All ages	78,964,743	3,974,272	5,946,380	290,348	7.5%	7.3%

<sup>a</sup> CDC Estimated Percentage of Cases = CDC Estimated COVID-19 Hospitalizations/CDC Estimated COVID-19 Cases

<sup>b</sup> Symphony Estimated Percentage of Hospitalizations = Symphony Estimated Hospitalizations/Symphony Estimated COVID-19 Cases

methodology seems to capture a percent of the sample similar to that of the CDC estimates.

## Discussion

An extensive literature search identified 32 articles which sought to define events using the Symphony Health database. From these articles a clear pattern emerged. Researchers used these claims data to restructure events using one or more time horizons. This review led us to define a generic methodology which can be used by future researchers hoping to define event specific hospitalizations within their own data. In addition, this methodology can easily be adapted to be used for other diseases and medical events. Our methodology is not specific to one database and can be extended to other claims databases as well. While our attempt to validate this method did yield similar findings to the CDC's estimates, no method is without its fair share of limitations. Currently, the Symphony Health Database are not open source or easily available without considerable cost to the researcher, which may stifle an attempt to validate this work. The inherent lack of structure and context in claims data thwart any efforts to know for certain how accurate the methodology truly is. Defining other conditional and relevant timelines is a very important step, yet there is no clear way of defining what these should be for a given event. Clinical guidance should determine the time horizon values used in defining overlapping events. Even still, these time horizon values are subject to a misunderstanding of claims data as well as systematic bias in the way claims data are recorded.

## Recommendations for replicating this work

The framework outlined in this manuscript is purposely generic so to be as useful across, what is, a diverse landscape of available claims data. We have several recommendations for replicating this framework in the readers own claims data. The researcher should first start by understanding the basic structure of the claims data available to them. For example, some claims datasets contain information batched on the claim level (what diagnoses, procedures, prescriptions happened during a timeline) or can be presented individually and not batched together at the claim level. The Symphony data available to us, for this paper, was separate non-batched data. Sometimes claims data contain what seem to be duplicate records which will need special consideration or rules on how to address them (example: remove duplicates or consider as separate events). Considering the characteristics and definition of the event of interest, the researcher may first define relevant conditional event horizons in the framework meant to validate that the event really did happen. For example, if mechanical

ventilation due to COVID-19 is an event of interest, a conditional event horizon between hospitalization and mechanical ventilation may be added to Fig. 2. Any other available data (example: inpatient, admission data) that the researcher has for the patient could be used for refining the event of interest or validating the event (example: prescribed medication, or surgical procedures). Some claims data have diagnosis codes ranked (principal, secondary, ...). If available, this rank can be used to strengthen the event reconstruction strategy. After identifying patients with the defined event, we recommend that the researcher compare the result with a constructed external benchmark, if available. If discrepancies exist between the framework estimates and the external benchmark this could be due to the errors in the reconstruction strategy or a bias in the available claims data. Researchers should be aware that claims data typically contain temporal aberrations which should be built into event reconstruction timelines. For example, it is possible that in the claims timeline it shows a hospital admittance on November 1st, and a COVID-19 diagnosis the following day on November 2nd. Whether these events are close enough to be considered tied together is up to the researcher and their personal event reconstruction strategy. These choices and decisions on how to construct the target event timeline, conditional/validation event horizons all impact the downstream construction of the events used for later analysis. Researchers should be transparent about this ambiguity within the limitation's section of their corresponding academic works.

## Conclusions

In this manuscript we have defined a generic methodology to rigorously define event specific hospitalizations from claims data. Our attempt to validate this methodology vs. the CDC's estimates showed similar estimates within those likely to be privately insured (18–64 years old). We believe this methodology will be useful for other researchers hoping to leverage large claims databases.

## Abbreviations

ICD: International Classification of Diseases; CPT: Current Procedural Terminology; ADA: Adalimumab; FDA: Food and Drug Administration; EUA: Emergency Use Authorization; CDC: Centre for Disease Control and Prevention.

## Acknowledgements

Not applicable.

## Authors' contributions

JL is the Co-PI of the funded project, corresponding author, wrote and coordinated the main ideas and findings in this manuscript. HS is the other Co-PI for the funded project and assisted in defining the methodology. EK is Science Librarian and completed the literature search. TX assisted with manuscript preparation and submission, editing, as well as general feedback regarding the methodology. JL, AB, ZS, LP, AS helped create the methodology within this

paper. ZS, LP, AS both assisted with editing and providing feedback for this manuscript. All authors read and approved the final manuscript.

### Funding

R21 National Library of Medicine 7R21LM013683-02.

### Availability of data and materials

The data that support the findings of this study are available from the COVID-19 Research Database, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the COVID-19 Research Database.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>College of Nursing, University of Cincinnati, Procter Hall, Cincinnati, OH 45219, USA. <sup>2</sup>Department of Bioengineering, University of Louisville Speed School of Engineering, Eastern Pkwy, Louisville, USA. <sup>3</sup>Dr. Bing Zhang Department of Statistics, University of Kentucky, Lexington, USA.

Received: 28 January 2022 Accepted: 18 May 2022

Published online: 26 May 2022

### References

- Konrad R, Zhang W, Bjarnadóttir M, Proaño R. Key considerations when using health insurance claims data in advanced data analyses: an experience report. *Health Syst (Basingstoke)*. 2019;9(4):317–25.
- Brixner D, Rubin DT, Mease P, Mittal M, Liu H, Davis M, et al. Patient support program increased medication adherence with lower total health care costs despite increased drug spending. *J Manag Care Spec Pharm*. 2019;25(7):770–9.
- Ash A, Porell F, Gruenberg L, Sawitz E, Beiser A. Adjusting Medicare capitation payments using prior hospitalization data. *Health Care Financ Rev*. 1989;10(4):17–29.
- Biskupiak J, Oderda G, Brixner D, Tang D, Zacker C, Dalal AA. Quantification of economic impact of drug wastage in oral oncology medications: comparison of 3 methods using palbociclib and ribociclib in advanced metastatic breast cancer. *J Manag Care Spec Pharm*. 2019;25(8):859–66.
- Chan SS, Chappel AR, Maddox KEJ, Hoover KW, Huang Y-LA, Zhu W, et al. Pre-exposure prophylaxis for preventing acquisition of HIV: A cross-sectional study of patients, prescribers, uptake, and spending in the United States, 2015–2016. *PLoS Med*. 2020;17(4):e1003072.
- Dellay B, Sexter A, Wang JH, Hess GP, Ray Kim W, Israni AK. Impact of Sofosbuvir-Based Therapy on Liver Transplant Candidates with Hepatitis C Virus Infection. *Pharmacotherapy*. 2019;39(4):424–32.
- Hampp C, Borders-Hemphill V, Moeny DG, Wysowski DK. Use of antidiabetic drugs in the U.S., 2003–2012. *Diabetes Care*. 2014;37(5):1367–74.
- Hess GP, Natarajan P, Faridi KF, Fievitz A, Valsdottir L, Yeh RW. Proprotein convertase subtilisin/kexin type 9 inhibitor therapy: payer approvals and rejections, and patient characteristics for successful prescribing. *Circulation*. 2017;136(23):2210–9.
- Laliberté F, Cloutier M, Nelson WW, Coleman CI, Pilon D, Olson WH, et al. Real-world comparative effectiveness and safety of rivaroxaban and warfarin in nonvalvular atrial fibrillation patients. *Curr Med Res Opin*. 2014;30(7):1317–25.
- Lam NN, Schnitzler MA, Segev DL, Hess GP, Kasiske BL, Randall HB, et al. Diabetes mellitus in living pancreas donors: use of integrated national registry and pharmacy claims data to characterize donation-related health outcomes. *Transplantation*. 2017;101(6):1276–81.
- Maron MS, Hellawell JL, Lucove JC, Farzaneh-Far R, Olivetto I. Occurrence of clinically diagnosed hypertrophic cardiomyopathy in the united states. *Am J Cardiol*. 2016;117(10):1651–4.
- Miller C, Apple S, Paige JS, Grabowsky T, Shukla O, Agnese W, et al. Current and future projections of amyotrophic lateral sclerosis in the united states using administrative claims data. *Neuroepidemiology*. 2021;55(4):275–85.
- Myung G, Nelson WW, McMahon MA. Effects of repository corticotropin injection on medication use in patients with rheumatologic conditions: A claims data study. *Journal of Pharmacy Technology*. 2017;33(4):151–5.
- Null K, Kumar V, Lissos T, Luo M. Infusion administration billing for vedolizumab and infliximab in inflammatory bowel disease. *J Med Econ*. 2018;21(11):1102–9.
- Parks MM, Secemsky EA, Yeh RW, Shen C, Choi E, Kazi DS, et al. Longitudinal management and outcomes of acute coronary syndrome in persons living with HIV infection. *Eur Heart J Qual Care Clin Outcomes*. 2021;7(3):273–9.
- Randall HB, Alhamad T, Schnitzler MA, Zhang Z, Ford-Glanton S, Axelrod DA, et al. Survival implications of opioid use before and after liver transplantation. *Liver Transpl*. 2017;23(3):305–14.
- Rice JB, Panaccio MP, White A, Simes M, Billmyer E, Downes N, et al. Consequences of insurance denials among U.S. patients prescribed repository corticotropin injection (Acthar Gel) for nephrotic syndrome. *Curr Med Res Opin*. 2021;37(3):431–41.
- Rohr MK, Mikuls TR, Cohen SB, Thorne JC, O'Dell JR. Underuse of methotrexate in the treatment of rheumatoid arthritis: A national analysis of prescribing practices in the US. *Arthritis Care Res (Hoboken)*. 2017;69(6):794–800.
- Xu J, Davis CS, Cruz M, Lurie P. State naloxone access laws are associated with an increase in the number of naloxone prescriptions dispensed in retail pharmacies. *Drug Alcohol Depend*. 2018;1(189):37–41.
- Tang Y, Huang X, Liu J, Shankar RR, Ganz ML, Rajpathak S. The effects of a sitagliptin formulary restriction program on diabetes medication use. *Am Health Drug Benefits*. 2017;10(9):456–62.
- Zhang KH, Singh H, Twardowski R, Lu C, Wan Y, Null K. Pgi4 retrospective study of dose escalation with adalimumab, infliximab, and vedolizumab in the treatment of inflammatory bowel disease. *Value Health*. 2020;23:S143.
- Butler J, Yang M, Manzi MA, Hess GP, Patel MJ, Rhodes T, et al. Clinical course of patients with worsening heart failure with reduced ejection fraction. *J Am Coll Cardiol*. 2019;73(8):935–44.
- Joshi K, Lafeuille M-H, Brown B, Wynant W, Emond B, Lefebvre P, et al. Baseline characteristics and treatment patterns of patients with schizophrenia initiated on once-every-three-months paliperidone palmitate in a real-world setting. *Curr Med Res Opin*. 2017;33(10):1763–72.
- Lee J, Pelkey R, Gubitosa J, Henrick MF, Ganz ML. Comparing Healthcare Costs Associated with Oral and Subcutaneous Methotrexate or Biologic Therapy for Rheumatoid Arthritis in the United States. *Am Health Drug Benefits*. 2017;10(1):42–9.
- Patel AK, Duh MS, Barghout V, Yenikomshian MA, Xiao Y, Wynant W, et al. Real-world Treatment Patterns Among Patients With Colorectal Cancer Treated With Trifluridine/Tipiracil and Regorafenib. *Clin Colorectal Cancer*. 2018;17(3):e531–9.
- Pergolizzi JV, Ben-Joseph R, Chang C-L, Hess G. Predicting medication persistence to buprenorphine transdermal system. *Pain Pract*. 2015;15(2):140–9.
- Rymer JA, Fonseca E, Bhandary DD, Kumar D, Khan ND, Wang TY. Difference in Medication Adherence Between Patients Prescribed a 30-Day Versus 90-Day Supply After Acute Myocardial Infarction. *J Am Heart Assoc*. 2021;10(1): e016215.
- Emond B, Rossi C, Côté-Sergent A, Dunn K, Lefebvre P, Lafeuille M-H, et al. Weight Change and Predictors of Weight Change Among Patients Initiated on Darunavir/Cobicistat/Emtricitabine/Tenofovir Alafenamide or Bictegravir/Emtricitabine/Tenofovir Alafenamide: A Real-World Retrospective Study. *J Health Econ Outcomes Res*. 2021;8(1):88–98.
- Lentine KL, Lam NN, Schnitzler MA, Hess GP, Kasiske BL, Xiao H, et al. Predonation prescription opioid use: A novel risk factor for readmission after living kidney donation. *Am J Transplant*. 2017;17(3):744–53.

30. Matias G, Haguinet F, Lustig RL, Edelman L, Chowell G, Taylor RJ. Model estimates of the burden of outpatient visits attributable to influenza in the United States. *BMC Infect Dis.* 2016;16(1):641.
31. Mehta D, Davis M, Epstein AJ, Wensel B, Grinnell T, Williams GR. Impact of early initiation of eslicarbazepine acetate on economic outcomes among patients with focal seizure: results from retrospective database analyses. *Neurol Ther.* 2020;9(2):585–98.
32. Potnis A, VanMeter S, Stange J. Prevalence of Hepatic Encephalopathy from a Commercial Medical Claims Database in the United States. *Int J Hepatol.* 2021;8(2021):8542179.
33. Ryan KJ, Nero D, Feinberg BA, Lee CH, Pimentel R, Gajra A, et al. Real-world incidence and cost of pneumonitis post-chemoradiotherapy for Stage III non-small-cell lung cancer. *Future Oncol.* 2020;16(1):4303–13.
34. Song Y, Betts KA, Lu Y, Singh R, Clewell J, Griffith J. Economic Burden of Switching to Different Biologic Therapies Among Tumor Necrosis Factor Inhibitor-Experienced Patients with Psoriatic Arthritis. *Rheumatol Ther.* 2019;6(2):285–97.
35. Estimated COVID-19 Burden' | CDC [Internet]. [cited 2021 Oct 25]. Available from: <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/burden.html>

### Publisher's note

Springer nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

