Please provide feedback

Causal Discovery on the Effect of Antipsychotic Drugs on Delirium Patients in the ICU using Large EHR Dataset

Riddhiman Adib   RIDDHIMAN.ADIB@MARQUETTE.EDU
Department of Computer Science
Marquette University
Milwaukee, Wisconsin, USA

Md Osman Gani   MOGANI@UMBC.EDU
Department of Information Systems
University of Maryland, Baltimore County
Baltimore, Maryland, USA

Sheikh Iqbal Ahamed   SHEIKH.AHAMED@MARQUETTE.EDU
Department of Computer Science
Marquette University
Milwaukee, Wisconsin, USA

Mohammad Adibuzzaman   ADIBUZZA@OHSU.EDU
Oregon Clinical and Translational Research Institute
Oregon Health & Science University
Portland, Oregon, USA

Abstract

Delirium occurs in about 80% cases in the Intensive Care Unit (ICU) and is associated with a longer hospital stay, increased mortality and other related issues. Delirium does not have any biomarker-based diagnosis and is commonly treated with antipsychotic drugs (APD). However, multiple studies have shown controversy over the efficacy or safety of APD in treating delirium. Since randomized controlled trials (RCT) are costly and time-expensive, we aim to approach the research question of the efficacy of APD in the treatment of delirium using retrospective cohort analysis. We plan to use the Causal inference framework to look for the underlying causal structure model, leveraging the availability of large observational data on ICU patients. To explore safety outcomes associated with APD, we aim to build a causal model for delirium in the ICU using large observational data sets connecting various covariates correlated with delirium. We utilized the MIMIC III database, an extensive electronic health records (EHR) dataset with 53,423 distinct hospital admissions. Our null hypothesis is: there is no significant difference in outcomes for delirium patients under different drug-group in the ICU. Through our exploratory, machine learning based and causal analysis, we had findings such as: mean length-of-stay and max length-of-stay is higher for patients in Haloperidol drug group, and haloperidol group has a higher rate of death in a year compared to other two-groups. Our generated causal model explicitly shows the functional relationships between different covariates. For future work, we plan to do time-varying analysis on the dataset.

1. Background and Problem Statement

With a focus on the theoretical development of causal inference methodologies in the previous three chapters, this chapter aims to discuss a practical, real-world application of the causal inference framework to untangle unknown healthcare information. For this purpose, we look into Delirium patients in the ICU.

Delirium (or acute brain failure) Girard et al. (2008) is a disorder or disruption of consciousness, presenting with a reduced capacity to focus, sustain, or shift concentration. Delirium occurs in about 80% cases in the Intensive Care Unit (ICU) and is associated with a more extended hospital stay, increased mortality for each additional day with Delirium in the ICU Pisani et al. (2009) and other clinical complications such as self-extubation and removal of catheters. Two of the significant issues in diagnosing and treating delirium patients are:

- Currently, no biomarker exists to diagnose Delirium; rather, Delirium is diagnosed with subjective assessment tools such as the confusion assessment method (CAM) Inouye et al. (1990); Wassenaar et al. (2015). This diagnosis requires the presence of a physician active in the medical center and makes the diagnosis and detection of Delirium patients in the real-world challenging.

- Delirium is commonly treated with antipsychotic drugs (APD) Girard et al. (2008) such as Haloperidol, Ziprasidone, Olanzapine, etc. However, multiple randomized controlled trials (RCTs) have shown either conflicting or inconclusive results about the efficacy of APD in the treatment of delirium Neufeld et al. (2016); Hatta et al. (2014). This has created a controversy over the efficacy or safety of APD in treating Delirium.

RCTs have been considered the gold standard since the 1960s Greene and Podolsky (2012). The goal was to identify the causation of diseases and understand the causal effect of drugs by the regulatory bodies such as the FDA and clinical communities. The key ideas behind RCT are:

- By random assignment of treatment or interventions, the confounding bias, i.e., the bias due to the assignment of treatment or presence of other variables, can be removed from the estimand, including the unobserved confounders.

- By comparing similar population groups of treatment and control arm, an estimation can be made about treatment efficacy in the target population group.

However, RCTs have their own set of challenges as well. RCTs have become increasingly time-consuming, costly, and are often infeasible for safety and efficacy reasons Frieden (2017). Thus there is a need to find alternatives to RCTs, possibly to detect causal effects from other sources of information and aid in removing controversies of treatments in the field.

Recent advances in technology and the adoption of computerized systems in routine healthcare have enabled the collection and curation of large volumes of data during routine healthcare, albeit with confounding biases. At the same time, recent advances in the theory
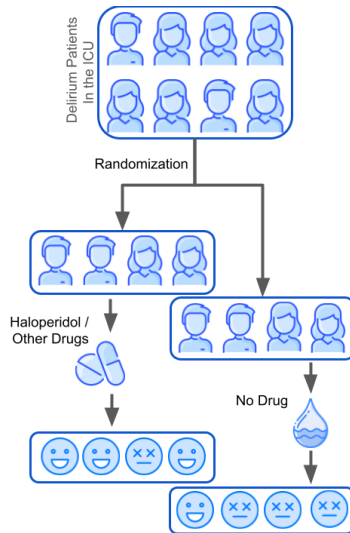
Figure 1: RCT for Antipsychotics-based treatment for Delirium

of causal model, more specifically structure causal models (SCM), provides the framework for adjusting for these confounding biases in many cases. This removal of biases can be done (sometimes even if the confounders are unobserved) from observational data using adjustment formulas such as backdoor/front door criterion Bareinboim and Pearl (2016b); Pearl et al. (2016); Spirtes et al. (2000). However, this approach requires developing a graphical representation of the problem domain with meticulous scrutiny of the variables' relationship, structure learning algorithms, clinical experience, and existing literature.

With the availability (collection, storage, maintenance) of large-scale data in different branches of science (healthcare, finance, sociology) and technology (connected health, smart home), opportunities exist to extract necessary information from it. Big data has aided in the revolution of neural networks, advanced reinforcement learning, and improved statistical machine learning methodologies. Most research advancements integrating big data revolve around curve fitting and correlation. However, without causal relationships, scientists lack the power of intervention or to even explore hypothetical scenarios (counterfactuals).

Big data is responsible for many breakthroughs and advancements in healthcare, contributing to improved treatment policy solutions and collaborated information from multiple sources. Although most breakthroughs are based on predictive models, causal relationships are more crucial for healthcare. This has led to countless experimental trials (randomized controlled trials, case-control studies) on finding the efficacy or impact of an intervention on target outcomes. Causal inference leverages big data and contributes to finding causal information, sometimes even without experiments. One of the strengths of causal inference methodologies is to draw conclusions on causal effects from observational data. Causal inference and its potential with big data are not limited to healthcare only; it has shown great potential in other fields (finance, sociology, law) as well Hofer et al. (2008); Nishioku et al. (2009); Semmler et al. (2005). Artificial intelligence is iteratively improved with research work and is getting better at decision making and predictive modeling.

Since RCTs cost a lot in terms of money and time, emulation of RCTs from the observational dataset can help reduce them. It also would aid in using datasets from all over the world to find causation in other diseases and health complexities. While RCTs are the gold standard for identifying causal effects of interventions, it is time-consuming and costly. On the other hand, the data collected during routine care, such as electronic health records (EHR), might also be valuable to generate insight, identify the disease pathway and estimate the effect of interventions using recent advances in methods for causal inference.

We aim to study the efficacy of APD in the treatment of Delirium using retrospective cohort analysis. We plan to use the Causal inference framework to look for the underlying causal structure model, leveraging the availability of large observational data on ICU patients. It will help us to untangle the causal relationship between variables and look into the counterfactual world (what-if). To explore safety outcomes associated with APD, our research work targets to develop a causal model for Delirium in the ICU using large observational data sets. We aim to utilize the MIMIC III database, an extensive electronic health records (EHR) dataset with 53,423 distinct hospital admissions Adibuzzaman et al. (2016). Our null hypothesis is: that there is no significant difference in outcomes for delirium patients under different drug-group in the ICU. If successful, our research work should help clear the common controversy over prescribing APDs as well as shed light on the underlying causal mechanism triggering Delirium in ICU patients. In other words, we propose the following specific aims.

1. Create and curate three cohorts for patients with Delirium in the ICU from MIMIC EHR data.

2. Develop structural causal models (SCM) with the domain expertise to integrate clinical knowledge and probabilistic information from the data to estimate the causal effect of interventions.

3. Validate the models with statistical methods and independent data sets.
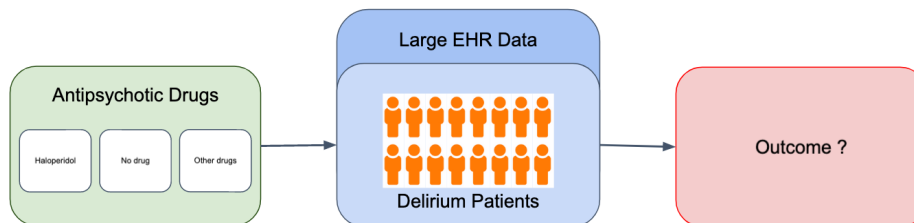

Figure 2: Target observational study from large EHR data

Epidemiologists have continuously involved causal inference tools, such as causal structure learning algorithms, in identifying underlying causal structures. The process is impactful since it generates a causal model based on the information available (data, literature, expertise), leading to a better understanding of the disease ecosystem; and estimates causal effects based on that. This creates a potential to explore Delirium-treatment-related controversies through observational datasets. Different studies have taken different paths; few

studies Vitolo et al. (2018); Adegunsoye et al. (2019) have used specific SLA algorithms to detect a causal DAG applicable for a targeted research question, whereas others Schlüter et al. (2019); Arendt et al. (2016) have assumed the causal structure from literature, and validated them using datasets available. We plan to create a similar computational pipeline for Delirium patients in the ICU inspired by these.

2. Method

To create a data cohort on Delirium patients in the ICU, along with relevant covariates, we seek help from MIMIC-III Johnson et al. (2016), a publicly available large electronic healthcare dataset. MIMIC-III is curated for twelve (12) years (2001-2012) and holds information on around 53k distinct hospital admissions with around 40k distinct patient histories. The database is well-maintained, de-identified, and open for researchers (with necessary and relevant access protocol) to explore and investigate.

The general process starts with appropriate data mining and data preparation process. We plan to extract information regarding Delirium patients (based on relevant ICD-9 code) and related covariates (decided upon exploring literature). We then move forward with the data analysis protocol, which consists of three (3) types of analysis:

- Exploratory analysis: to explore data distribution and dataset properties

- Machine learning-driven analysis: to infer primary point of interest (i.e., primary outcome) based on all available covariates, as a standard approach to prediction

- Causal analysis:

  - Causal structure generation: to regenerate underlying causal model through various structure learning algorithms

  - Causal effect estimation: to evaluate the 'true' causal effect of treatment on our defined points of interests

3. Results

This section describes our data curation protocol in detail, along with data exploration and analysis. We present our general findings based on those steps taken.

3.1. Covariate Selection

We start the process by defining the research questions *(Is Haloperidol better at treating Delirium patients in the ICU, compared to no antipsychotics or other antipsychotics, such as Ziprasidone, Olanzapine, etc.?)*. We formulate this question based on controversies present in existing literature *(described in the background section)*. Our null hypothesis is that there is no significant difference in target outcomes for Delirium patients under different antipsychotics treatment groups in the ICU. We define the treatment as the antipsychotics prescribed after being diagnosed with Delirium in the ICU, with three different arms (Haloperidol, no antipsychotics, and other antipsychotics). Our primary outcomes are (1) patient death in hospital and (2) patient death timeline (death in 30 days / 90 days /

sex	age	race	icd9 codes
sofa	apsiii	surgery	pneumonia
sepsis	dementia	alzheimers	depression
anxiety	met. acidosis	airway obs.	copd
liver disease	heart disease	mechvent.	mechvent. count
time to mechvent.	**drug group**	drug categories count	drug timelength
death in hospital	**death timeline**	**length of stay**	**time in mechvent.**

Table 1: Features in MIMIC-Delirium

a year / survived more than a year). Our secondary outcomes are (1) length of stay in the ICU and (2) time put in mechanical ventilation. A total of fifty (50) relevant covariates are explored and marked, which are closely correlated with our points of interests (primary and secondary outcomes) for Delirium patients in the ICU. However, due to the lack of availability of all covariates in the observational dataset, we opt for the most significant twenty-eight (28) covariates, as listed in Table 1. Here, the drug group (Haloperidol, no drug, other drugs) is the treatment provided. Primary outcomes are death in hospital & death timeline, and secondary outcomes are the length of stay & time in mechvent.

These all together defined a target trial that we plan to emulate. The target trial is inspired by existing RCTs done on delirium patients to find the effects of antipsychotics and is designed to minimize the effect of confounding variables and (selection) bias. Based on these, we start our data curation process from the MIMIC-III dataset.

3.2. Data Curation Process

To determine eligible Delirium patients, we look into patients with ICD-9 code 293.0 (Delirium due to conditions classified elsewhere) ICD9Data.com. We extract relevant information about the patients from admissions, icu_stays, and diagnoses_icd table to form the base dataset. We then infuse it with information from cptevents, d_icd_diagnoses and prescriptions tables, and other views presented in the public repository of the database (sofa, apsiii, ventdurations) MIT-LCP. We merge all information together to create our target dataset of 1398 patients. We name this curated dataset as *MIMIC-Delirium* for future references.

3.3. Data Overview & Exploratory Insights

After our data curation to create MIMIC-Delirium dataset, we successfully extract 1671 ICU stays with 1445 hospital admission counts on 1398 unique patients and their relevant 28 covariate information. In terms of treatment provided in the ICU, we found 681 *(40.75%)* were given Haloperidol, 528 *(31.60%)* were given other antipsychotics and 462 *(27.65%)* were given no antipsychotics. In terms of outcome, 311 *(18.61%)* had death in 30 days, 108 *(6.46%)* had death in 90 days, 175 *(10.47%)* had death in a year, and 253 *(15.14%)* survived at least a year (information on 821 *(49.13%)* were unknown). Among the common associated diseases in the ICU, 375 *(22.44%)* had Sepsis, 484 *(28.96%)* had Pneumonia, 1035 *(61.94%)* had (a variation of) heart diseases, and 97 *(5.80%)* had (a variation of) liver diseases. Figure 4 shows the general data distribution on age in years *skewed to right since*
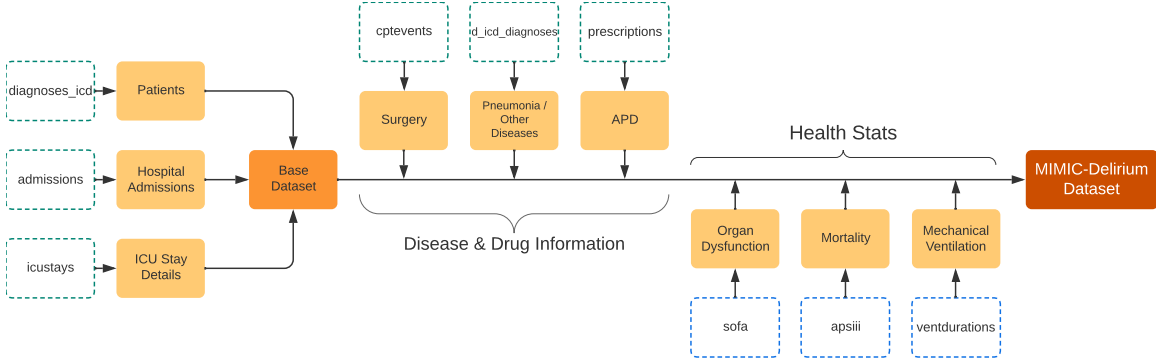
Figure 3: Data mining protocol (simplified)

Delirium is frequent in elderly population and length-of-stay in days *(skewed to left since higher number of ICU stay is severe and rare).*
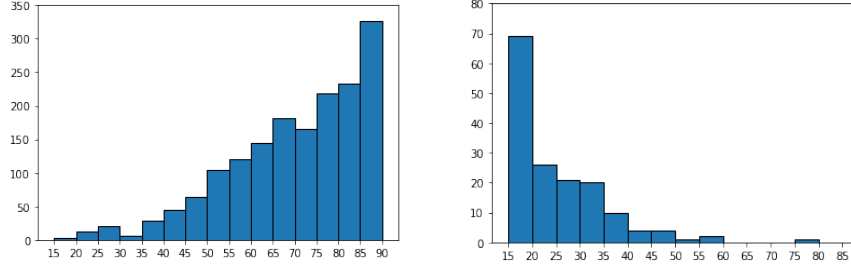

Figure 4: Data distribution on age in years *(left)* and length-of-stay in days *(right)*

Additionally, we had findings such as mean length-of-stay and max length-of-stay is higher for patients in the Haloperidol drug group, most patients, who were given multiple APD, were given Haloperidol, The Haloperidol group has a higher death rate in a year than the other two groups, etc.

For the statistical analyses, we conducted a one-way between-subjects ANOVA to compare the effect of the drug group on length of stay in Haloperidol, no drug, and other drugs group. With $p < 0.05$, we found a significant effect of the drug group on the length-of-stay. Post hoc comparisons by the Tukey HSD test indicate that the mean score for the Haloperidol group (mean: 7.47, deviation: 8.55) was significantly higher compared to no drug group (mean: 4.12, deviation: 5.66) and other drugs group (mean: 5.44, deviation: 6.14).

3.4. Predictive Analysis on MIMIC-Delirium dataset

Before our deep dive into causal exploration, we briefly explored the MIMIC-Delirium dataset for predictive analysis. We employed standard supervised classification algorithms on the complete dataset, with all 24 covariates (discarding the output features) as features and death in hospital as the label. We deployed 10-fold cross-validation with Logistic

Regression, Support Vector Machine, and XGBoost algorithm. Mean accuracy with Logistic Regression is 89.71%, mean accuracy with SVM is 89.11%, and test-mlogloss-mean for XGBoost (with 50 rounds of boosts) is 0.2724. For XGBoost, we also find that length-of-stay and age have the highest impact in predicting outcome death in this case, which is self-explanatory. Figure 5 shows the general correlation between features as a heatmap.
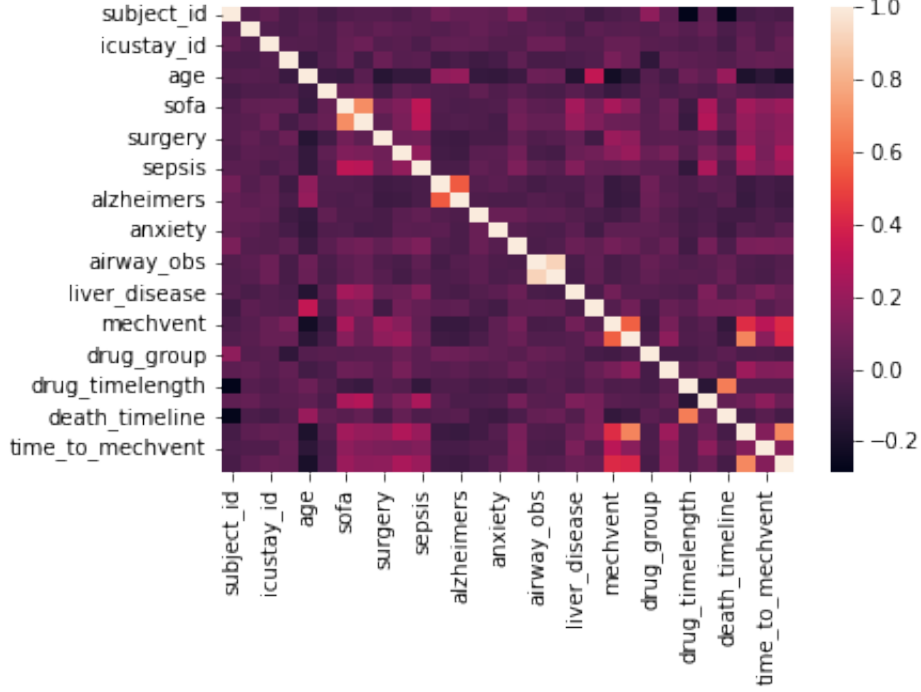

Figure 5: Correlation heatmap of MIMIC-Delirium

3.5. Causal Analysis on MIMIC-Delirium dataset

Our causal analysis is built upon two steps: (1) causal structure generation and (2) causal effect estimation (based on causal structure generated).

3.5.1. CAUSAL STRUCTURE GENERATION

To generate the most feasible underlying causal structure from the MIMIC-Delirium dataset, we rely on causal structure learning algorithms (SLA), with assumptions of causal sufficiency and faithfulness. Specifically, we apply eight (8) causal structure learning algorithms: (1) PC, (2) FCI, (3) GES, (4) GIES, (5) GDS, (6) LINGAM, (7) MMHC, and (8) MMTABU, with help from existing R libraries: (1) pcalg Kalisch et al. (2012); Hauser and Bühlmann (2012) and (2) pchc Tsagris (2021). With the application of these SLAs, we have eight (8) individual causal graphs. However, we apply majority voting to each edge to merge all this information together. This merging defines an edge as being present in the final graph if it is present in more than 50% cases (more than four graphs). Although this

	Causal effect of drug group on:			
	death in hospital	death timeline	length of stay in days	time in mech. vent.
Hal. vs. No Drug	0.0310	-0.1291	1.8372	12.3007
Other Drug vs. No Drug	0.0216	0.0373	1.6102	8.1912
Hal. vs. Other Drug	0.0113	-0.1386	-0.0533	4.4827

Table 2: Outcomes estimation in Average Treatment Effects (ATE)

is a straightforward and naive solution to merge multiple causal graphs, we employ this ensembling method since no standard has been established in the literature yet. Figure 6 shows the final merged causal graph generated.

3.5.2. CAUSAL EFFECT ESTIMATION

With the causal structure generated, we now focus on causal effect estimation. For this purpose, we employ the pipeline proposed by Microsoft **Do-Why** library Sharma et al. (2019):

- Modeling

- Identification

- Estimation

- Refutation

With modeling completed as part of the causal structure generation step, we now focus on causal effect identification and estimation. Based on the causal structure generated, we identify the conditional probability equation for the four target outcomes. Specifically, we express the do-calculus operations Bareinboim and Pearl (2016a) in order to 'virtually' manipulate the outcomes. The do-calculus equations are presented below:

- $P(death\_in\_hosp|do(drug\_group)) = \sum_{age} P(death\_in\_hosp|drug\_group, age)P(age)$
- $P(death\_timeline|do(drug\_group)) = \sum_{age} P(death\_timeline|drug\_group, age)P(age)$
- $P(los\_days|do(drug\_group)) = \sum_{heart\_disease, mechvent} P(los\_days|drug\_group, heart\_disease, mechvent)P(heart\_disease, m$
- $P(time\_in\_mechvent|do(drug\_group)) = \sum_{age, mechvent} P(death\_in\_hosp|drug\_group, age, mechvent)P(age, mechvent)$

We now find the causal effect estimates based on these causal expressions identified. In Table 2, we present the causal effect estimations, as Average Treatment Effects (ATE), for treatment, aka, drug group on the four target outcomes. As shown in the table, the causal effect of treatment on death in Delirium and death timeline is very close. However, any drug, Haloperidol (1.8372) and other drugs (1.6102), does much better in reducing hospital length of stay compared to the no drug patient group (-0.0533). In addition to that, any drug performs better (8.1912) in reducing time in mechanical ventilation compared to no drug (4.4827), and Haloperidol does better (12.3007) than any other drugs (8.1912).
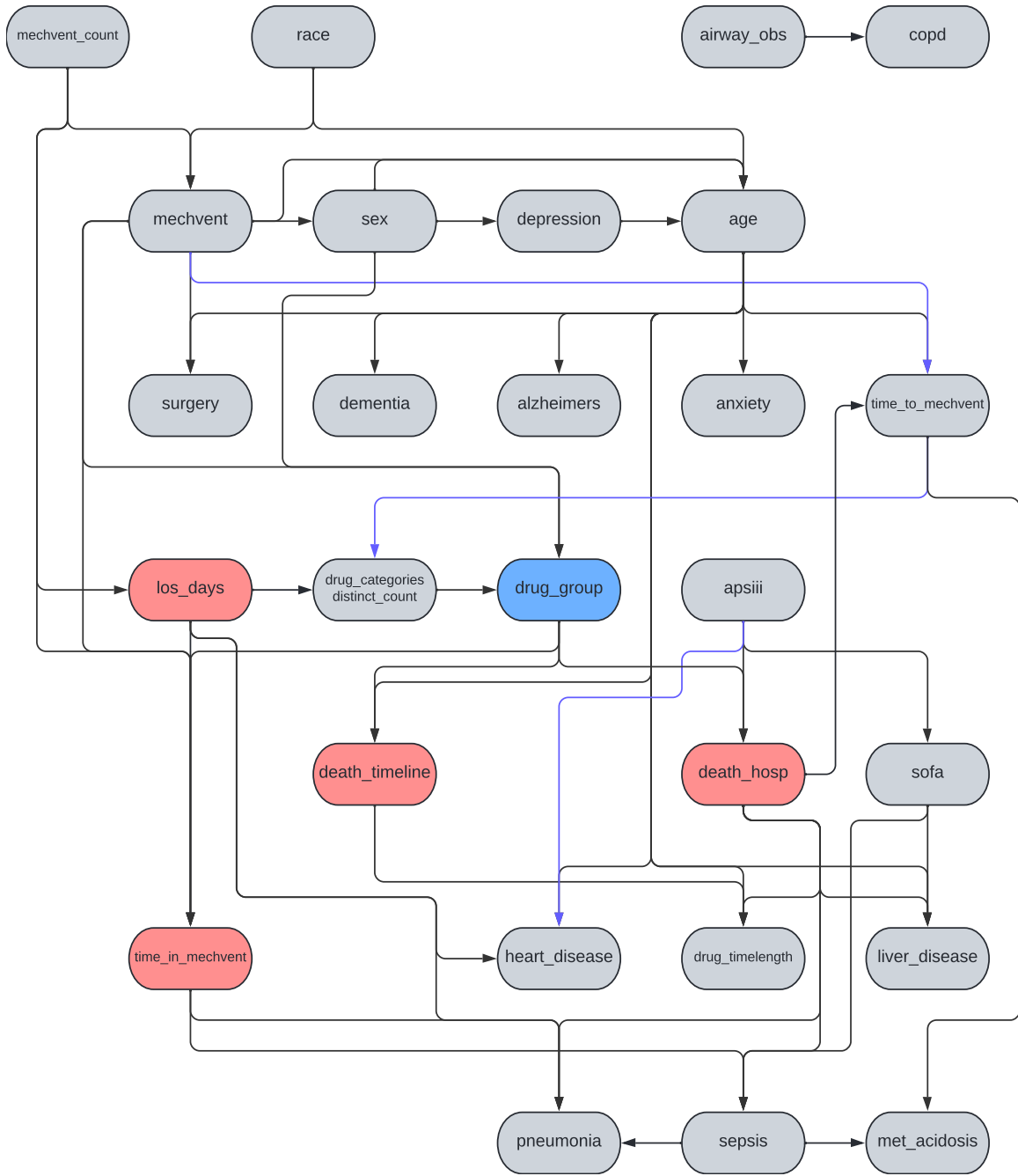
Figure 6: Combined Causal Graph for Delirium in the ICU *(blue: treatment, red: primary and secondary outcomes)*

We now move to the final stage of causal effect estimation, which is the refutation of the estimated effect. We do so in four different steps: by adding a random common cause to the causal model, adding an unobserved common cause to the causal model, using a Placebo treatment, and using a subset of data. The expectation for these four is that:

- **Adding a random common cause**: should not change the estimated outcome from before since this should be adjusted by use of do-calculus expressions

- **Adding an unobserved common cause**: should change the estimated outcome from before since the unobserved confounder induces non-removable biases in the system

- **Using a placebo treatment**: should be close to zero since placebo treatment should not have any impact on the outcome

- **Using a subset of data**: should not change the estimated outcome from before since underlying data distribution did not change

Application of these four steps results in the following values, which also align with our expectations for a stable causal model and estimated effect:

- Estimated effect: 0.0309

- Add a random common cause: 0.0310

- Add an Unobserved Common Cause: 0.0262

- Use a Placebo Treatment: 0.0003

- Use a subset of data: 0.0319

4. Discussion

We have explored a potential observational study on Delirium patients in the ICU in this study. Our curated dataset is analyzed through two lenses: regular observational analysis and 'simulated' randomized controlled trial through the structural theory of causation. We have multiple novel contributions to this research work:

- Our observational study creates a prospective data cohort *(MIMIC-Delirium)* for Delirium patients

- Data properties for *MIMIC-Delirium* provides insight into the general patient demography in the ICU

- Machine learning-driven analysis on *MIMIC-Delirium* presents usage of prediction-based computational modeling and these algorithm's general strengths

- Causal analysis on *MIMIC-Delirium* found:

    – No significant impact (X) of Antipsychotics choice in one of the primary outcomes, death in hospital

- No significant impact (X) of Antipsychotics choice in length of stay in the ICU; however, usage of any drug shows better outcome (X) compared to that with no drugs

- Haloperidol performs better (X) in affecting time in mechanical ventilation, compared to the similar impact of usage of other drugs or no drugs,

Our study relies on a few underlying assumptions. We assume that the Delirium patients in the ICU represent general Delirium demography since it occurs more frequently (80% cases in ICU) in the ICU compared to other traditional medical settings. Additionally, in generating the causal structure, we did not incorporate any background knowledge from peer-reviewed literature because of the existing controversies over the usage and benefits of Antipsychotics on the Delirium population *(discussed in the background section)*. One of the critical limitations of our study is the lack of involvement of experienced physicians actively working in the ICU. Their involvement can aid in disputing general confusion in different parts of the study; however, bias from their understanding needs to be handled by involving multiple physicians. This limitation can be mitigated in future work.

In summary, our proposed analysis and pipeline create pathways for similar studies, especially in the healthcare research domain. The abundance of curated large electronic healthcare data presents a potential to find unexplored insights in a specific population group. Causal inference, especially the structural theory of causation, holds the potential to handle such research questions, look for causal insights, and report them appropriately.

References

Ayodeji Adegunsoye, Justin M Oldham, Shashi K Bellam, Steven Montner, Matthew M Churpek, Imre Noth, Rekha Vij, Mary E Strek, and Jonathan H Chung. Computed tomography honeycombing identifies a progressive fibrotic phenotype with increased mortality across diverse interstitial lung diseases. *Annals of the American Thoracic Society*, 16(5):580–588, 2019.

Mohammad Adibuzzaman, Ken Musselman, Alistair Johnson, Paul Brown, Zachary Pitluk, and Ananth Grama. Closing the data loop: An integrated open access analysis platform for the mimic database. In *2016 Computing in Cardiology Conference (CinC)*, pages 137–140. IEEE, 2016.

Linn Håkonsen Arendt, Cecilia Høst Ramlau-Hansen, Allen J Wilcox, Tine Brink Henriksen, Jørn Olsen, and Morten Søndergaard Lindhard. Placental weight and male genital anomalies: a nationwide danish cohort study. *American journal of epidemiology*, 183(12): 1122–1128, 2016.

Elias Bareinboim and Judea Pearl. Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences*, 113(27):7345–7352, 2016a.

Elias Bareinboim and Judea Pearl. Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences*, 113(27):7345–7352, 2016b.

Thomas R Frieden. Evidence for health decision making—beyond randomized, controlled trials. *New England Journal of Medicine*, 377(5):465–475, 2017.

Timothy D Girard, Pratik P Pandharipande, and E Wesley Ely. Delirium in the intensive care unit. *Critical care*, 12(S3):S3, 2008.

Jeremy A Greene and Scott H Podolsky. Reform, regulation, and pharmaceuticals—the kefauver–harris amendments at 50. *New England Journal of Medicine*, 367(16):1481–1483, 2012.

Kotaro Hatta, Yasuhiro Kishi, Ken Wada, Toshinari Odawara, Takashi Takeuchi, Takafumi Shiganami, Kazuo Tsuchida, Yoshio Oshima, Naohisa Uchimura, Rie Akaho, et al. Antipsychotics for delirium in the general hospital setting in consecutive 2453 inpatients: a prospective observational study. *International journal of geriatric psychiatry*, 29(3): 253–262, 2014.

Alain Hauser and Peter Bühlmann. Characterization and greedy learning of interventional markov equivalence classes of directed acyclic graphs. *The Journal of Machine Learning Research*, 13(1):2409–2464, 2012.

Stefan Hofer, Christian Bopp, Christian Hoerner, Konstanze Plaschke, Rita M Faden, Eike Martin, Hubert J Bardenheuer, and Markus A Weigand. Injury of the blood brain barrier and up-regulation of icam-1 in polymicrobial sepsis. *Journal of Surgical Research*, 146 (2):276–281, 2008.

ICD9Data.com. Icd-9-cm diagnosis code 293.0 : Delirium due to conditions classified elsewhere. URL http://www.icd9data.com/2015/Volume1/290-319/290-294/293/293.0.htm.

Sharon K Inouye, Christopher H van Dyck, Cathy A Alessi, Sharyl Balkin, Alan P Siegal, and Ralph I Horwitz. Clarifying confusion: the confusion assessment method: a new method for detection of delirium. *Annals of internal medicine*, 113(12):941–948, 1990.

Alistair EW Johnson, Tom J Pollard, Lu Shen, H Lehman Li-wei, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G Mark. Mimic-iii, a freely accessible critical care database. *Scientific data*, 3:160035, 2016.

Markus Kalisch, Martin Mächler, Diego Colombo, Marloes H Maathuis, and Peter Bühlmann. Causal inference using graphical models with the r package pcalg. *Journal of statistical software*, 47:1–26, 2012.

MIT-LCP. mimic-code/sofa.sql at main · mit-lcp/mimic-code. URL https://github.com/MIT-LCP/mimic-code/blob/main/mimic-iii/concepts/severityscores/sofa.sql.

Karin J Neufeld, Jirong Yue, Thomas N Robinson, Sharon K Inouye, and Dale M Needham. Antipsychotic medication for prevention and treatment of delirium in hospitalized adults: a systematic review and meta-analysis. *Journal of the American Geriatrics Society*, 64 (4):705–714, 2016.

Tsuyoshi Nishioku, Shinya Dohgu, Fuyuko Takata, Tomoaki Eto, Naoko Ishikawa, Kota B Kodama, Shinsuke Nakagawa, Atsushi Yamauchi, and Yasufumi Kataoka. Detachment of brain pericytes from the basal lamina is involved in disruption of the blood–brain barrier

caused by lipopolysaccharide-induced sepsis in mice. *Cellular and molecular neurobiology*, 29(3):309–316, 2009.

Judea Pearl, Madelyn Glymour, and Nicholas P Jewell. *Causal inference in statistics: A primer*. John Wiley & Sons, 2016.

Margaret A Pisani, So Yeon Joyce Kong, Stanislav V Kasl, Terrence E Murphy, Katy LB Araujo, and Peter H Van Ness. Days of delirium are associated with 1-year mortality in an older intensive care unit population. *American journal of respiratory and critical care medicine*, 180(11):1092–1097, 2009.

Daniela K Schlüter, Rowena Griffiths, Abdulfatah Adam, Ashley Akbari, Martin L Heaven, Shantini Paranjothy, Anne-Marie Nybo Andersen, Siobhán B Carr, Tania Pressler, Peter J Diggle, et al. Impact of cystic fibrosis on birthweight: a population based study of children in denmark and wales. *Thorax*, 74(5):447–454, 2019.

Alexander Semmler, Torsten Okulla, Magdalena Sastre, Lucia Dumitrescu-Ozimek, and Michael T Heneka. Systemic inflammation induces apoptosis with variable vulnerability of different brain regions. *Journal of chemical neuroanatomy*, 30(2-3):144–157, 2005.

Amit Sharma, Emre Kiciman, et al. DoWhy: A Python package for causal inference. https://github.com/microsoft/dowhy, 2019.

Peter Spirtes, Clark N Glymour, Richard Scheines, and David Heckerman. *Causation, prediction, and search*. MIT press, 2000.

Michail Tsagris. *pchc: Bayesian Network Learning with the PCHC and Related Algorithms*, 2021. URL https://CRAN.R-project.org/package=pchc. R package version 0.6.

Claudia Vitolo, Marco Scutari, Mohamed Ghalaieny, Allan Tucker, and Andrew Russell. Modeling air pollution, climate, and health data using bayesian networks: A case study of the english regions. *Earth and Space Science*, 5(4):76–88, 2018.

A Wassenaar, MHWA van den Boogaard, Theo van Achterberg, AJC Slooter, MA Kuiper, ME Hoogendoorn, KS Simons, E Maseda, N Pinto, C Jones, et al. Multinational development and validation of an early prediction model for delirium in icu patients. *Intensive care medicine*, 41(6):1048–1056, 2015.