

# *Penemuan Hubungan Kausal dari Lama Rawat Inap di Unit Perawatan Intensif: Pendekatan Algoritma Peter-Clark pada Analisis Data Diagnosis dan Laboratorium*

Ismail Syababun Halim<sup>1</sup>, Anastasia Mia Martalia<sup>2</sup>, Muhammad Helmi Hibatullah<sup>3</sup>  
Nugraha Priya Utama<sup>4</sup>, Ayu Purwarianti<sup>5</sup>

<sup>1,2,3,4,5</sup>School of Electrical Engineering and Informatics - Bandung Institute of Technology

<sup>4,5</sup>Center of Excellence for AI, Computer Vision, NLP, and Big Data Analytics - Bandung Institute of Technology  
Jl. Ganesa, Lb. Siliwangi, Kec. Coblong, Kota Bandung, Indonesia  
23524072@std.stei.itb.ac.id

---

## Abstrak

Dalam sistem pelayanan kesehatan, Unit Perawatan Intensif (ICU) merupakan komponen penting untuk menangani pasien dalam kondisi kritis yang membutuhkan pemantauan intensif. Namun, durasi rawat inap atau *Length of Stay* (LoS) seorang pasien di ICU biasanya sangat bervariasi dan perpanjangan LoS berdampak signifikan pada beban biaya, penggunaan sumber daya, dan efisiensi pelayanan di rumah sakit. Penelitian ini bertujuan untuk mengidentifikasi faktor kausal yang memengaruhi LoS ICU menggunakan Algoritma Peter-Clark (PC) untuk penemuan kausal atau *Causal Discovery*. Data yang digunakan berasal dari MIMIC-IV, sebuah basis data klinis komprehensif dari Beth Israel Deaconess Medical Center tahun 2008–2019, yang mencakup demografi pasien, kode diagnosis ICD, dan hasil pemeriksaan laboratorium. Metode yang digunakan meliputi penerapan Algoritma PC, yang dipilih karena kemampuannya pada data berdimensi tinggi dengan *Fisher's Z-test* untuk pengujian independensi, yang diimplementasikan pada berbagai tingkat signifikansi ( $\alpha = 0.01, 0.05, 0.1$ ). Validasi dilakukan melalui 500 iterasi bootstrap untuk mengetahui stabilitas dari struktur graf kausal. Hasil analisis menunjukkan enam variabel yang secara konsisten menjadi penyebab langsung LoS ICU diantaranya *admission type*, *APR-DRG severity*, *high mortality risk*, *category*, *flag*, dan *anchor age*. Struktur kausal yang dihasilkan memberikan gambaran hubungan sebab-akibat yang stabil dan signifikan antar variabel klinis, yang dapat digunakan untuk mendukung pengambilan keputusan berbasis data dalam manajemen pasien dan alokasi sumber daya ICU. Studi ini juga menegaskan potensi pendekatan *Causal Discovery* dalam analitik layanan kesehatan, khususnya dalam memahami faktor determinan LoS ICU secara mendalam.

**Kata Kunci:** Unit Perawatan Intensif, Lama Rawat Inap, Penemuan Kausal, Algoritma PC, Fisher's Z-test, MIMIC-IV

## Abstract

*In healthcare systems, the Intensive Care Unit (ICU) is a crucial component for managing critically ill patients who require intensive monitoring. However, the length of stay (LoS) in the ICU often varies significantly across patients, and prolonged LoS imposes substantial burdens on hospital costs, resource utilization, and overall service efficiency. This study aims to identify the causal factors influencing ICU LoS using the Peter-Clark (PC) Algorithm for causal discovery. The dataset was obtained from MIMIC-IV, a comprehensive clinical database from Beth Israel Deaconess Medical Center covering the years 2008–2019, which includes patient demographics, ICD diagnosis codes, and laboratory test results. The method involves applying the PC Algorithm, selected for its robustness in high-dimensional data, combined with Fisher's Z-test for conditional independence testing, implemented at multiple significance levels ( $\alpha = 0.01, 0.05, \text{ and } 0.1$ ). Validation was performed using 500 bootstrap iterations to assess the stability of the resulting causal graph. The analysis revealed six variables that consistently act as direct causal drivers of ICU LoS: admission type, APR-DRG severity, high mortality risk, category, flag, and anchor age. The resulting causal structure shows stable and significant cause-effect relationships among clinical variables, providing a data-driven foundation for decision-making in patient management and ICU resource allocation. This study highlights the potential of causal discovery approaches in healthcare analytics, particularly in understanding the key determinants of ICU length of stay.*

---

**Keywords:** Intensive Care Unit, Length of Stay, Causal Discovery, PC Algorithm, Fisher's Z-test, MIMIC-IV

---

## I. PENDAHULUAN

Dalam dunia kesehatan (*healthcare*), Unit Perawatan Intensif (ICU) merupakan salah satu bagian penting dalam sistem pelayanan kesehatan karena ICU merupakan tempat untuk menangani pasien dengan kondisi kritis yang memerlukan pemantauan khusus serta perawatan medis yang intensif. Namun, lamanya setiap pasien menjalani perawatan di ICU biasanya bervariasi (*length of stay*). Menurut penelitian yang dilakukan di Saudi Arabia, ditemukan bahwa beberapa faktor lamanya pasien di rawat di ICU di antaranya faktor usia, pasien berjenis kelamin laki-laki, pasien yang akan atau sudah melaksanakan operasi, pasien yang memiliki riwayat darah tinggi, obesitas, diabetes, penyakit ginjal kronis, dan pasien yang memiliki masalah disfungsi sistem imun [1]. Durasi proses perawatan pasien di ICU yang berkepanjangan tidak hanya meningkatkan risiko komplikasi, tetapi juga membebani sistem pelayanan kesehatan yang mana menyebabkan peningkatan biaya operasional serta menjadi terbatasnya sumber daya ICU, selain itu lamanya pasien dirawat di ICU menyebabkan meningkatnya kecemasan pada keluarga pasien [2].

Banyak penelitian sebelumnya di Indonesia telah mengeksplorasi terkait faktor-faktor yang berkontribusi terhadap durasi lamanya seorang pasien menjalani perawatan di ICU dengan menggunakan pendekatan statistik maupun model prediktif berbasis *machine learning* atau *deep learning*. Namun, pendekatan ini biasanya hanya menjelaskan terkait korelasi antara faktor tersebut tanpa memperhitungkan hubungan sebab-akibat (*causal inference*) antara faktor-faktor atau variabel yang digunakan [3]. Menurut data Laporan Kementerian Kesehatan (Kemenkes) Republik Indonesia per 2023 [4] rasio ketersediaan ruangan rawat inap per 1000 penduduk berada di 1.38 dimana statusnya cukup sesuai dengan ketentuan ketercukupan oleh World Health Organization (WHO) sedangkan pada salah satu studi kasus di RSUP Dr. Sardjito okupansi kebutuhan ICU berasal dari IGD mencapai 52,68% dengan hampir 92% ditolak dengan 68% ditolak karena keterbatasan tempat pada tahun 2021 [5]. Ketersediaan ICU dan bagaimana Rumah Sakit mengatur manajemen ICU dapat meningkatkan persentase okupansi ICU rujukan berasal dari IGD. Oleh karena itu, untuk dapat memahami faktor-faktor yang paling berpengaruh dan menjadi penyebab lama perawatan inap (*length of stay*) seorang pasien di ICU, dibutuhkan pendekatan *causal inference* untuk dapat mengidentifikasi hubungan kausal dari data yang diteliti. Kausalitas, yang melibatkan perbedaan

antara sebab dan akibat, sangat penting untuk memahami hubungan kompleks dalam data [6][7].

*Causal inference* atau yang lebih dikenal dengan hubungan sebab-akibat antara faktor risiko dengan hasil merupakan salah satu hal yang sangat penting dalam dunia kesehatan. Pada *causal inference*, dicari faktor yang terjadi atas hasil [8][9]. Hubungan sebab-akibat antara faktor risiko seperti gaya hidup, riwayat kesehatan, faktor genetik, serta lingkungan dari pasien dengan hasil kesehatan pasien seperti lama rawat inap pasien, jenis penyakit yang diderita, serta respon pasien terhadap alat bantu medis maupun obat inilah yang dipelajari dalam dunia kesehatan karena hal tersebut sangat dibutuhkan untuk pengambilan keputusan klinis dan kebijakan kesehatan yang tepat.

Penelitian ini merupakan studi pertama di Indonesia yang menerapkan algoritma PC (Peter-Clark) pada dataset MIMIC-IV untuk analisis kausal LoS ICU, dengan pendekatan studi *causal inference* LoS ICU. Studi ini berbeda dari penelitian sebelumnya karena menekankan analisis kausal serta memvalidasi hasil dengan *bootstrapping* dan perbandingan dengan algoritma Structural Equation Modeling (SEM) dan Greedy Equivalence Search (GES).

Dataset MIMIC-IV sendiri merupakan dataset rekam medis dari pasien ICU yang dikumpulkan oleh Beth Israel Deaconess Medical Center [10]. Dataset ini menyediakan informasi yang sangat kaya tentang pasien, termasuk data klinis, hasil pemeriksaan laboratorium, serta kode diagnosis ICD yang dapat digunakan untuk analisis kausal.

Algoritma PC merupakan salah satu metode yang dapat digunakan dalam pendekatan *causal discovery* yang mampu mengidentifikasi hubungan sebab-akibat antara variabel dengan asumsi *sparsity* dan independensi kondisi tertentu [11]. Algoritma ini nantinya akan digunakan untuk membangun struktur keterkaitan (*causal*) dari faktor-faktor yang memengaruhi lama rawat inap di ICU berdasarkan data hasil ICD Diagnoses dan Laboratory Test. Algoritma PC digunakan pada penelitian ini karena algoritma PC menggunakan *Causal Markov Condition* dan *Faithfulness*, tanpa asumsi distribusi variabel tertentu, sehingga cocok untuk data laboratorium dan diagnosis campuran seperti data MIMIC-IV [12]. Dengan menerapkan pendekatan *causal inference* menggunakan algoritma PC pada dataset MIMIC-IV, diharapkan dapat diperoleh pemahaman yang lebih mendalam tentang faktor-faktor yang benar-benar berpengaruh terhadap lama rawat inap (*length of stay*) seorang pasien di ICU.

Untuk menilai stabilitas empiris fitur grafik yang dihasilkan oleh algoritma PC, *bootstrapping*

nonparametrik sering digunakan dalam literatur. Teknik ini menghasilkan sejumlah besar sampel *bootstrap*, masing-masing dibangun dengan mengambil sampel  $N$  observasi dengan penggantian dari kumpulan data asli, dan menjalankan kembali algoritma PC pada setiap kumpulan sampel. Frekuensi di mana setiap *edge* atau struktur- $v$  muncul kembali di seluruh replikasi bootstrap berfungsi sebagai skor stabilitas empiris: fitur yang muncul dalam proporsi sampel yang tinggi (umumnya  $\geq 70\%$ ) diartikan sebagai *robust* [13], sedangkan fitur dengan frekuensi lebih rendah mungkin mencerminkan variabilitas sampling atau sensitivitas terhadap fluktuasi sampel terbatas [14].

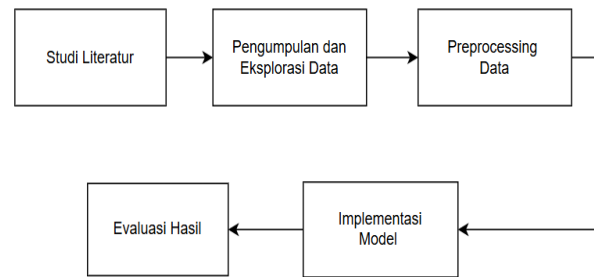
Setelah penipisan berbasis kendala dan orientasi grafik tak berarah oleh algoritma PC, dua kerangka kerja komplementer sering digunakan dalam penemuan kausal: Structural Equation Modeling (SEM) dan Greedy Equivalence Search (GES) [15]. SEM menganggap setiap variabel  $X_i$  sebagai fungsi (mungkin non-linier) dari penyebab langsungnya  $Pa(X_i)$  ditambah suku gangguan independen  $\epsilon_i$ , sehingga membingkai inferensi kausal sebagai estimasi sistem persamaan, seperti pada persamaan (1); di bawah kondisi *identifiability* yang sesuai (misalnya, *non-Gaussianity* atau *additive noise*), SEM dapat memulihkan DAG yang unik daripada hanya kelas ekuivalensinya.

$$X_i = f_{i_i}(Pa(X_i), \epsilon_i) \quad (1)$$

Sebaliknya, GES adalah pencarian berbasis skor yang menavigasi ruang kelas ekuivalensi Markov [16]. Dimulai dari grafik kosong, ia secara *greedy* menambahkan *directed edge* yang inklusinya paling meningkatkan skor *goodness-of-fit* yang dipilih (seperti BIC), kemudian memetakan hasilnya ke kelas ekuivalensi yang sesuai dan akhirnya secara *greedy* menghapus *edge* untuk lebih mengoptimalkan skor. Dalam batas sampel besar, GES konvergen ke kelas ekuivalensi yang sebenarnya di bawah asumsi Markov dan *Faithfulness* yang sama tetapi tanpa memerlukan uji kemandirian bersyarat [17].

## II. METODE PENELITIAN

Penelitian ini bertujuan untuk menemukan hubungan kausal antara hasil diagnosis (kode ICD), data laboratorium, dan lama perawatan pasien di ICU (*ICU Length of Stay* atau ICU LOS) menggunakan pendekatan *causal discovery* dengan algoritma PC (Peter-Clark). Solusi dirancang untuk menganalisis data observasional dan memodelkan hubungan antarvariabel dengan graf kausal terarah (*Directed Acyclic Graphs*). Langkah-langkah solusi yang dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

a) **Studi Literatur:** Pada tahap studi literatur ini difokuskan pada tiga poin utama yaitu teknik *causal discovery* berbasis *constraint-based*, khususnya algoritma PC, penerapan algoritma kausal dalam domain perawatan intensif (ICU), dan karakteristik serta penggunaan database MIMIC-IV dalam penelitian klinis.

b) **Pengumpulan dan Eksplorasi Data:** Data dikumpulkan dari basis data Medical Information Mart for Intensive Care (MIMIC-IV), yang berisi rekam medis elektronik teranonimkan dari pasien perawatan intensif di Beth Israel Deaconess Medical Center (2008–2019) [10]. MIMIC-IV merupakan set data rekam medis elektronik yang dapat diakses secara bebas, memuat data klinis dari pasien di unit gawat darurat dan unit perawatan intensif pada sebuah pusat medis akademik besar. MIMIC-IV telah menerapkan langkah-langkah kerahasiaan data yang ketat melalui proses de-identifikasi komprehensif. Ini mencakup penghapusan identifikasi pribadi pasien dan penggunaan berbagai teknik *obfuscation* data. Sebagai contoh, penyedia set data menerapkan metode sandi acak untuk mengaburkan informasi sensitif. Selain itu, detail temporal spesifik—seperti tahun-tahun pasti—digeneralisasi menjadi rentang tahun untuk lebih melindungi anonimitas pasien. Sebelum *preprocessing* yang intensif, *exploratory data analysis* (EDA) dilakukan untuk memahami struktur, konten, dan kualitas awal set data. Tinjauan skema memeriksa hubungan antara tabel dan karakteristik data, didukung oleh keluaran `df.info()`. Dokumentasi resmi MIMIC-IV dikonsultasikan untuk memahami terminologi medis dan makna kolom data. Profiling data awal mengungkapkan banyak nilai yang hilang pada kolom-kolom kunci, yaitu *deathtime* atau waktu kematian pasien, *discharge\_location* atau lokasi kepulangan pasien, *dod* atau tanggal kematian pasien, *drg\_severity* atau tingkat keparahan penyakit, dan *drg\_mortality* atau tingkat kematian penyakit. Selain itu, teridentifikasi juga volume data yang besar pada tabel *labevents* yang berisi hasil pemeriksaan laboratorium pasien selama perawatan yang menjadi tantangan pemrosesan utama. Eksplorasi ini juga

mengidentifikasi pola pada nilai *drg\_type* atau tipe *diagnosis related group* (DRG) serta implikasinya terhadap data keparahan dan mortalitas.

c) *Preprocessing Data*: Tahapan ini mencakup pembersihan data dan seleksi fitur. Sebelum eksperimen, data melalui serangkaian tahapan pengolahan yang meliputi pembersihan, transformasi, ekstraksi fitur, dan integrasi untuk memastikan kualitas dan relevansi data. Tahap pembersihan data berfokus pada penghapusan nilai kosong, duplikat, dan outlier di berbagai tabel seperti *admissions*, *patients*, *d\_icd\_diagnoses*, *diagnoses\_icd*, dan *drgcodes*. Ini termasuk penghapusan kolom yang tidak relevan, pengisian nilai *null* yang sesuai, standardisasi format kode, dan penanganan duplikasi kode ICD yang disengaja. Tujuannya adalah untuk menghasilkan dataset yang akurat, konsisten, dan siap untuk dianalisis.

Setelah data dibersihkan, dilakukan transformasi untuk mengubahnya ke format yang lebih terstruktur. Pada tahap ini, diagnosis diagregasi ke tingkat kunjungan pasien, dan penanda kunjungan berulang (*is\_first\_stay*, *is\_subsequent\_stay*) ditambahkan. Kode DRG dipisahkan menjadi kategori APR-DRG dan MS/HCFA-DRG, serta ditambahkan indikator risiko keparahan dan mortalitas tinggi. Diagnosis juga dikelompokkan ke dalam domain klinis berdasarkan *chapter* ICD, dengan penambahan fitur temporal seperti *prev\_stay\_primary\_icd\_code* dan *prev\_stay\_clinical\_domain* untuk analisis kunjungan berulang.

Selanjutnya, proses dilanjutkan dengan ekstraksi fitur untuk memperkaya dataset dengan informasi baru yang relevan, seperti diagnosis primer, jumlah diagnosis per kunjungan, urutan kunjungan, dan frekuensi kunjungan. Informasi kode APR-DRG dan MS-DRG diekstrak bersama indikator risiko. Domain klinis dari kode ICD juga diekstrak untuk analisis agregat berdasarkan jenis penyakit. Seluruh data dari berbagai tabel seperti *admissions*, *patients*, *icu\_stay*, *labevents*, *d\_labitems*, *d\_icd\_diagnoses*, *diagnoses\_icd*, dan *drgcodes* kemudian diintegrasikan menjadi satu dataset komprehensif. Penggabungan ini dilakukan secara bertahap menggunakan kunci penghubung seperti *subject\_id* dan *hadm\_id* untuk memastikan semua informasi pasien terkait rawat inap, ICU, tes laboratorium, dan diagnosis tersedia dalam satu sumber terpadu, mempermudah analisis kausal dan predictive modeling.

Terakhir, sebelas variabel dari dataset komprehensif sebelumnya dipilih berdasarkan relevansi klinisnya yang telah terbukti terhadap lama rawat inap atau *length of stay* (LoS) di ICU.

1. *valuenum* atau nilai numerik laboratorium dan *flag* atau penanda abnormalitas biner. Kedua variabel ini dipilih karena gangguan laboratorium awal sering kali menjadi pertanda ketidakstabilan fisiologis dan disfungsi organ yang dapat memperpanjang LoS [18].
2. *anchor\_age* atau usia pasien saat admisi ICU. Variabel ini menangkap resiliensi fisiologis pasien karena individu yang lebih tua umumnya memiliki beban komorbiditas yang lebih tinggi dan cenderung pulih lebih lambat [14].
3. *gender* atau jenis kelamin biologis. Jenis kelamin berdampak pada respons imun, presentasi penyakit, serta keputusan perawatan yang terkait dengan variasi LoS [14].
4. *admission\_type* atau jenis admisi seperti darurat, elektif, atau transfer. Variabel ini mencerminkan karakteristik tingkat keparahan dan kompleksitas yang lebih tinggi pada admisi non-elektif, yang berkorelasi dengan durasi rawat inap yang lebih panjang [18].
5. *apr\_drg\_severity* atau klasifikasi keparahan dan *high\_mortality\_risk* atau risiko mortalitas APR-DRG disertakan untuk mengukur tingkat keparahan penyakit dan memprediksi pemanfaatan sumber daya [19].
6. *icd\_chapter* atau kode *International Classification of Diseases* (ICD) tingkat tinggi dipertahankan untuk merepresentasikan kelompok diagnosis yang berbeda beserta jalur klinisnya yang unik [17].
7. *category* atau kategori yang lebih luas (misalnya, lab, vital, dan administratif) dan *clinical\_domain* (misalnya neurologi dan kardiologi) diturunkan dari variabel-variabel sebelumnya. Hal ini dilakukan guna meningkatkan interpretasi dan mengungkap pola LoS yang spesifik per domain [18].
8. Terakhir, *los* (*length of stay*) atau lama waktu rawat inap ICU itu sendiri sebagai keluaran target disiapkan untuk pemodelan, agar dapat mendukung analisis efisiensi ICU, implikasi biaya, dan luaran pasien [19].

Fitur-fitur ini kemudian menjalani pengkodean label untuk variabel kategoris dan standardisasi menggunakan *StandardScaler* yang menghasilkan kumpulan data pemodelan akhir. Pengkodean variabel yang digunakan dirinci dalam Tabel 1.

**Tabel 1. Variable Encoding**

Variable	Encoding/Scaling Detail
valuenum	Distandardisasi: $(x - 92.37) / 3965.7$
anchor_age	Distandardisasi: $(x - 62.87) / 16.05$
flag_encoded	0 = abnormal, 1 = normal
category_encoded	0 = Chemistry, 1 = Hematology
admission type encoded	0-7 untuk tipe admission
clinical domain encoded	0-20 sesuai chapter ICD
icd chapter encoded	0-20 sesuai huruf ICD
gender encoded	0 = F, 1 = M
high_mortality_risk_encoded	0 = false, 1 = true
apr_drg_severity	1 (Mild) - 4 (Extreme)

d) Implementasi Model: Penelitian ini mengimplementasikan algoritma PC (Peter-Clark) sebagai metode utama untuk menemukan hubungan kausal yang memengaruhi lama rawat inap di ICU. Algoritma PC merupakan algoritma berbasis batasan yang mendukung penemuan sebab-akibat dan pembelajaran struktur kausal. Algoritma ini secara khusus cocok untuk tugas-tugas yang bertujuan mengidentifikasi fitur-fitur prediktif terhadap suatu *outcome* di berbagai lingkungan ( $\epsilon$ ). Mengingat penelitian ini menggunakan data yang dapat dikonseptualisasikan berasal dari konteks yang berbeda (misalnya, pengukuran lab rutin versus titik data spesifik terkait admisi), kemampuan algoritma PC untuk menangani data multi-lingkungan adalah sebuah keuntungan [20]. Objektif fundamental dari algoritma ini adalah untuk membangun graf kausal berdasarkan data yang diobservasi, di bawah asumsi kunci *Causal Markov Condition* dan *Faithfulness*.

Proses penemuan kausal dilakukan menggunakan algoritma PC dari *library causal-learn* pada set data yang telah melalui *preprocessing*. Pertama, algoritma secara iteratif menguji independensi kondisional antara semua pasangan variabel untuk menyimpulkan *Directed Acyclic Graph* (DAG) yang mendasari dan merepresentasikan potensi hubungan kausal. Untuk setiap uji independensi kondisional, digunakan uji Z Fisher (diimplementasikan sebagai *fisherz* dalam *causal\_learn.utils.cit*) yang sangat sesuai untuk data laboratorium dan diagnostik yang kontinu serta terstandarisasi. Untuk menguji bagaimana pemilihan ambang batas signifikansi memengaruhi struktur kausal yang disimpulkan, dijalankan algoritma PC di bawah tiga tingkat alfa yang berbeda ( $\alpha = 0.01, 0.05, \text{ dan } 0.1$ ). Hal ini dilakukan untuk membandingkan stabilitas dan *sparsity* dari graf yang dihasilkan [13].

e) Evaluasi Hasil: Pada fase evaluasi, fokus utamanya adalah memahami dan memvalidasi

keluaran model. *Directed Acyclic Graphs* (DAG) kausal divisualisasikan menggunakan *networkx* dan *matplotlib.pyplot*, dengan arah hubungan diinterpretasikan dalam konteks klinis. Stabilitas model dinilai melalui *bootstrapping* [21], menghasilkan 500 sampel yang masing-masing terdiri dari 100.000 titik data acak. *Edge* (hubungan) yang muncul pada lebih dari 50% iterasi dianggap stabil. Hasilnya divalidasi lebih lanjut melalui analisis komparatif dengan *Structural Equation Modeling* (SEM) [22] dan *Greedy Equivalence Search* (GES) yang memberikan beragam perspektif mengenai hubungan kausal. Selain itu, dilakukan juga kuantifikasi kesesuaian antara *subgraph* yang disimpulkan dan *subgraph* dari *surrogate ground truth* menggunakan presisi, *recall*, *F1-score*, dan *Structural Hamming Distance* (SHD).

### III. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil penemuan kausal (*causal discovery*) menggunakan algoritma PC (Peter-Clark) dan uji stabilitas dengan *bootstrapping*, perbandingan dengan menggunakan pendekatan *Structural Equation Modeling* (SEM) dan *Greedy Equivalence Search* (GES), serta evaluasi kinerja terhadap *surrogate ground truth*.

#### 1) Hasil Penemuan Kausal (Causal Discovery) dengan Algoritma PC

Algoritma PC diterapkan pada *dataset* yang telah disiapkan untuk mengidentifikasi potensi hubungan kausal yang memengaruhi *Length of Stay* (LoS) di ICU. Eksperimen dilakukan menggunakan nilai alfa ( $\alpha$ ) sebesar 0.01, 0.05, dan 0.1 untuk menguji sensitivitas struktur graf yang ditemukan terhadap ambang batas signifikansi statistik. Seperti yang dirangkum dalam Tabel 2, nilai *alpha* yang lebih kecil memberlakukan kriteria yang lebih ketat untuk mempertahankan *edge* (hubungan).

Pada tingkat *alpha* ( $\alpha$ ) 0.01, 0.05, dan 0.1, variabel-variabel yang teridentifikasi memiliki hubungan kausal langsung dengan LoS disajikan sebagai berikut: Graf kausal (Gambar 2, 3, dan 4) dan penyebab langsung LoS yang teridentifikasi (Tabel 3) ditampilkan.

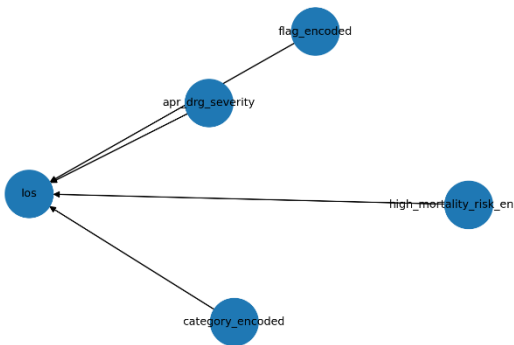
**Tabel 2. Pengaruh Nilai Alpha ( $\alpha$ ) terhadap Struktur DAG pada Algoritma PC**

Nilai Alpha	Kriteria Penghapusan Edge	Struktur Hasil DAG	Risiko
Kecil (misalnya: 0,01)	Sangat ketat, edge mudah dihapus jika tidak	Graf menjadi lebih <i>spars</i>	Risiko <i>false negative</i> (hubungan nyata tidak

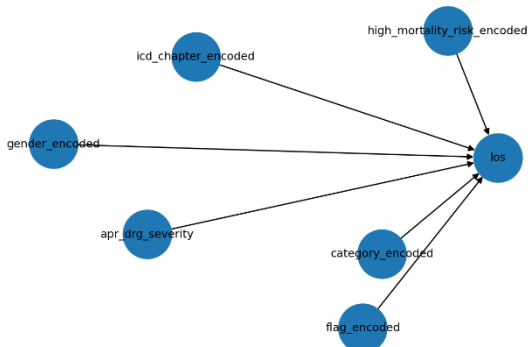
	signifikan	(jarang)	terdeteksi)
Sedang (misalnya: 0,05)	Moderat, seimbang antara deteksi dan penghapusan <i>edge</i>	Struktur moderat, cocok sebagai baseline	Risiko <i>false positive</i> dan <i>false negative</i> relatif seimbang
Besar (misalnya: 0,1)	Longgar, <i>edge</i> sulit dihapus	Graf menjadi lebih <i>dense</i> (padat)	Risiko <i>false positive</i> (mendeteksi hubungan yang tidak nyata)

Tabel 3. Algoritma PC: Variabel yang Secara Langsung Menyebabkan LoS ( $\alpha = 0,01; 0,05; 0,1$ )

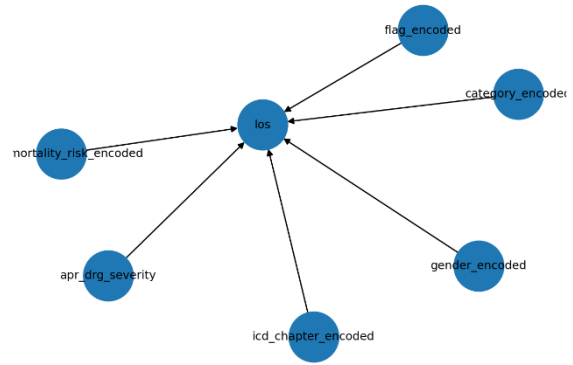
Variabel yang Terhubung ke LoS	$\alpha$		
	0,01	0,05	0,10
flag_encoded	✓	✓	✓
category_encoded	✓	✓	✓
icd_chapter_encoded	-	✓	✓
apr_drg_severity	✓	✓	✓
high_mortality_risk_encoded	✓	✓	✓
gender_encoded	-	✓	✓



Gambar 2. Graf Kausal dari Algoritma PC ( $\alpha = 0,01$ )



Gambar 3. Graf Kausal dari Algoritma PC ( $\alpha = 0,05$ )



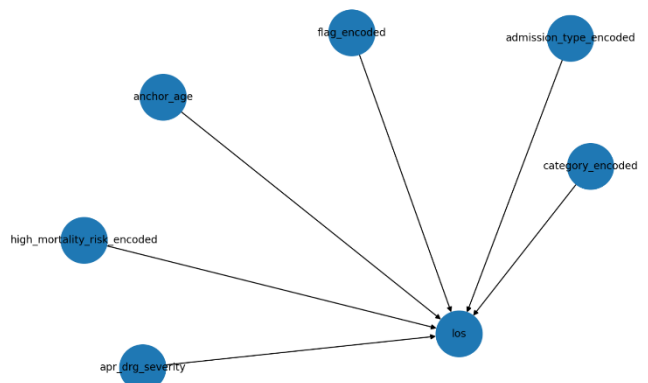
Gambar 4. Graf Kausal dari Algoritma PC ( $\alpha = 0,1$ )

2) Uji Stabilitas Algoritma PC dengan Bootstrapping

Untuk mengevaluasi *robustness* temuan algoritma PC (dengan menggunakan  $\alpha = 0.05$  untuk setiap *run* dalam *bootstrap*), metode *bootstrapping* digunakan secara iterasi sebanyak 500 kali. Setiap iterasi menggunakan sampel acak sebanyak 100.000 titik data. Graf kausal stabil yang dihasilkan dari proses ini disajikan pada Gambar 5. Enam variabel utama secara konsisten muncul sebagai penyebab langsung LoS (Tabel 4): *flag encoded*, *category encoded*, *apr drg severity*, *high mortality risk encoded*, *admission type encoded*, dan *anchor age*.

Tabel 4. Algoritma PC dengan Bootstrapping: Variabel yang Secara Langsung Menyebabkan LoS

No	Simpul
1	flag_encoded
2	category_encoded
3	apr_drg_severity
4	high_mortality_risk_encoded
5	admission_type_encoded
6	anchor_age



**Gambar 5. Graf Kausal Stabil dari Algoritma PC dengan *Bootstrapping* (500 iterasi,  $\alpha = 0,05$ )**

**3) Hasil *Structural Equation Modeling* (SEM)**

Pada tahap validasi komparatif, *Structural Equation Modeling* (SEM) diimplementasikan untuk melihat variabel yang terhubung secara langsung ke LoS. Analisis SEM mengidentifikasi sepuluh variabel dengan koefisien ( $\beta$ ) yang signifikan secara statistik memengaruhi LoS (nilai- $p \leq 0.001$ ), yang rinciannya disajikan dalam Tabel 5. Diagram jalur SEM yang bersesuaian ditampilkan pada Gambar 5. Tabel 5 menampilkan hasil *Structural Equation Modeling* (SEM) untuk setiap variabel yang mengarah ke LoS dengan nilai- $p \leq 0.001$ .

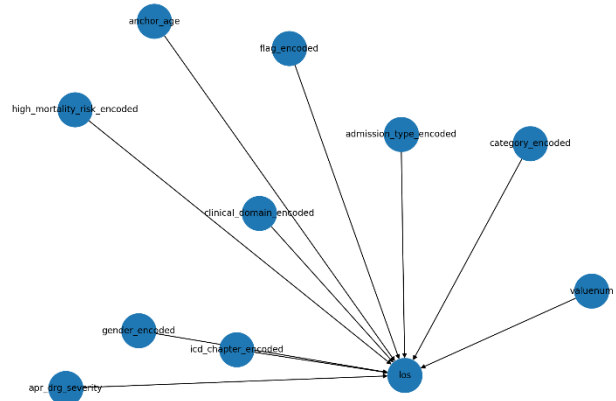
Keterangan dari SEM menunjukkan dampak positif yang kuat dari *high\_mortality\_risk\_encoded* dan *apr\_drg\_severity* terhadap LoS. Sebaliknya, 'normal' *flag\_encoded* dan 'Hematology' *category\_encoded* terkait dengan LoS yang lebih singkat.

**4) Hasil *Greedy Equivalence Search* (GES)**

Algoritma GES mengidentifikasi model yang lebih parsimonious, dengan hubungan kausal langsung ke LoS hanya dari *flag\_encoded* dan *category\_encoded* (Tabel 6). Graf kausal GES ditunjukkan pada Gambar 7.

**Tabel 5. Hasil *Structural Equation Modeling* (SEM) dengan  $p$ -value  $\leq 0.001$**

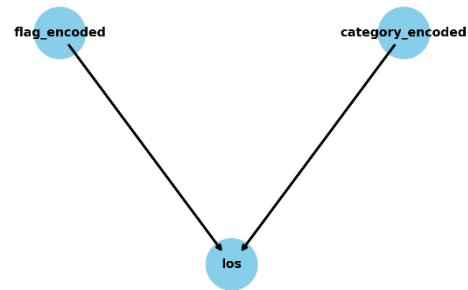
No	Variabel yang Terhubung ke LoS	Estimate ( $\beta$ )
1	<i>flag_encoded</i>	-0.450
2	<i>category_encoded</i>	-0.402
3	<i>apr_drg_severity</i>	0.757
4	<i>high_mortality_risk_encoded</i>	2.939
5	<i>admission_type_encoded</i>	0.307
6	<i>anchor_age</i>	-0.049
7	<i>valuenum</i>	-0.011
8	<i>gender_encoded</i>	0.211
9	<i>clinical_domain_encoded</i>	-0.030
10	<i>icd_chapter_encoded</i>	0.045



**Gambar 6. Graf Kausal dari *Structural Equation Model* (SEM)**

**Tabel 6. Algoritma GES: Variabel yang Secara Langsung Menyebabkan LoS**

No	Simpul
1	<i>flag_encoded</i>
2	<i>category_encoded</i>



**Gambar 7. Graf Kausal dari Algoritma *Greedy Equivalence Search* (GES)**

GES mengindikasikan bahwa *flag\_encoded* 'normal' secara langsung mengurangi LoS. Interpretasi untuk *category\_encoded* (Hematologi vs. Kimia) yang meningkatkan LoS, sebagaimana diturunkan dari teks sumber untuk GES, berlawanan dengan temuan SEM dan memerlukan peninjauan cermat untuk pelaporan yang konsisten.

**Tabel 7. Perbandingan Hasil dan Cara Kerja Algoritma PC, SEM, dan GES**

Algoritma	Hasil	Cara Kerja
PC	6 variabel menuju LoS	Melakukan serangkaian tes untuk melihat apakah dua variabel benar-benar saling terkait setelah mempertimbangkan variabel lain. Jika hubungan tidak kuat, maka <i>edge</i> dipotong, sehingga graf menjadi lebih jarang ( <i>sparse graph</i> ) [13].

SEM	10 variabel menuju LoS	Mencari model yang paling cocok secara keseluruhan (misalnya, dengan AIC/BIC) di semua hubungan yang ditentukan, sambil menambahkan variabel tersembunyi guna menjelaskan varians yang belum tertangkap. Ini menghasilkan jaringan yang lebih padat dan dapat mengalami <i>overfitting</i> jika kompleksitas model tidak dibatasi dengan cermat [23].
GES	2 variabel menuju LoS	Secara bertahap menambah atau menghapus <i>edge</i> di graf untuk mendapatkan skor terbaik (BIC), dengan memberi penalti jika graf terlalu rumit. GES kurang efektif dalam mengidentifikasi hubungan kausal yang sebenarnya ada, pada dataset dengan jumlah variabel sedang, karena penalti globalnya untuk <i>edge</i> yang lemah [24].

**Tabel 8. Perbandingan Hasil dan Implikasi Algoritma PC, SEM, dan GES**

Algoritma	Hasil	Cara Kerja
PC	6 variabel menuju LoS	Lebih robust terhadap asosiasi palsu ( <i>spurious associations</i> ), namun bisa gagal dalam mendeteksi hubungan yang lemah, terutama jika <i>threshold</i> dibuat lebih ketat.
SEM	10 variabel menuju LoS	Dapat melihat hubungan langsung dan tidak langsung, tetapi model bisa menjadi terlalu rumit dan membuat hubungan yang sebenarnya palsu, jika data yang digunakan terlalu sedikit.
GES	2 variabel menuju LoS	Menghasilkan graf yang sederhana dan mudah untuk diinterpretasikan, tetapi dapat melewatkan beberapa hubungan yang sebenarnya ada.

**5) Evaluasi Kinerja dan Analisis Surrogate Ground Truth**

Penggunaan *surrogate ground truth* sebagai evaluator yang *established* untuk mengevaluasi algoritma PC menggunakan pendekatan uji stabilitas berbasis *bootstrapping* (sebagaimana dirinci dalam metodologi, yang selanjutnya disempurnakan untuk mengidentifikasi *edge* paling kuat yang muncul di berbagai ambang stabilitas dari 500 *run bootstrap*). *Surrogate ground truth* konsensus ini (Tabel 9) terdiri dari tujuh *edge* langsung ke LoS.

Kinerja algoritma PC dibandingkan dengan *surrogate* untuk *subgraph* dari *edge* yang diarahkan

ke LoS ditunjukkan pada Tabel 10. Pada  $\alpha = 0.01$ , Algoritma PC mencapai presisi sempurna (1.0) untuk *edge* yang terdeteksi ke LoS, mengidentifikasi 4 *true positives* tanpa *false positives* sama sekali, menghasilkan *F1-Score* terbaik (0.800) dan SHD terendah (2) untuk *subgraph* ini. Namun, dua hubungan yang hadir dalam kebenaran dasar *surrogate* (*anchor\_age* → LoS dan *admission\_type\_encoded* → LoS) secara konsisten terlewatkan di semua nilai alfa (FN=2), mengindikasikan keterbatasan dalam mendeteksi sinyal spesifik yang lebih lemah ini. *Recall* tetap konstan pada 0.667. Ini menunjukkan bahwa  $\alpha = 0.01$  menawarkan struktur yang paling *robust* untuk subset hubungan kausal ke LoS yang teridentifikasi.

**Tabel 9. Surrogate Ground Truth: Variabel yang Secara Langsung Menyebabkan LoS**

No	Simpul
1	flag_encoded
2	category_encoded
3	apr_drg_severity
4	high_mortality_risk_encoded
5	admission_type_encoded
6	anchor_age

**Tabel 10. Metrik Kinerja Algoritma PC terhadap Surrogate Ground Truth (Directed Subgraph → LoS)**

$\alpha$	TP	FP	FN	Precision	Recall	F1-Score	SHD
0.01	4	0	2	1.000	0.667	0.800	2
0.05	4	2	2	0.667	0.667	0.667	4
0.1	4	2	2	0.667	0.667	0.667	4

**IV. KESIMPULAN**

Penelitian ini bertujuan untuk menemukan faktor-faktor penyebab langsung dari lama rawat inap atau *Length of Stay* (LoS) di ICU dengan menerapkan algoritma Peter-Clark (PC) pada data klinis MIMIC-IV. Melalui pengujian stabilitas menggunakan 500 iterasi *bootstrapping*, penelitian ini berhasil mengidentifikasi enam faktor kausal yang secara konsisten dan *robust* memengaruhi LoS secara langsung, yaitu: *admission type* (tipe admisi), *APR-DRG severity* (tingkat keparahan APR-DRG), *high mortality risk* (risiko mortalitas tinggi), *category* (kategori tes klinis), *flag* (penanda abnormalitas laboratorium), dan *anchor age* (usia pasien). Keenam faktor ini menjadi hipotesis kausal utama yang divalidasi lebih lanjut melalui perbandingan dengan metode lain. Hasilnya menunjukkan bahwa model *Structural Equation Modeling* (SEM)

mengidentifikasi set variabel yang lebih luas (sepuluh faktor), sementara *Greedy Equivalence Search* (GES) menghasilkan model yang jauh lebih sederhana dengan hanya dua faktor. Hal ini menempatkan hasil dari algoritma PC sebagai temuan yang paling seimbang, menghindari potensi kompleksitas berlebih dari SEM dan simplifikasi berlebih dari GES.

Untuk mengevaluasi kinerja, struktur kausal yang ditemukan oleh algoritma PC diuji terhadap *surrogate ground truth* yang dibangun dari hasil bootstrapping yang paling stabil. Hasil evaluasi menunjukkan bahwa pengaturan alpha yang lebih ketat ( $\alpha=0.01$ ) memberikan performa terbaik. Pada tingkat signifikansi ini, algoritma PC mencapai presisi sempurna (1.0) dengan tidak ada satupun hubungan palsu (*false positive*) yang terdeteksi, serta menghasilkan F1-Score tertinggi (0.800) dan *Structural Hamming Distance* (SHD) terendah (2). Namun, presisi ini datang dengan konsekuensi dua hubungan kausal yang lebih lemah, yaitu dari *anchor age* dan *admission type* ke LoS, tidak terdeteksi (*false negative*), sehingga menghasilkan nilai *recall* sebesar 0.667. Temuan ini menegaskan bahwa pemilihan ambang batas statistik yang ketat sangat penting untuk menghasilkan struktur kausal yang paling andal dan dapat dipercaya dari data klinis yang kompleks, meskipun beberapa sinyal kausal yang lebih lemah mungkin terlewatkan.

## UCAPAN TERIMA KASIH

Ucapan terima kasih kami ucapkan kepada Bapak Ibu yang telah membimbing kami dalam penulisan penelitian jurnal ini.

## REFERENSI

- [1] K. K. Alharbi, T. J. Arbaein, A. A. Alzhrani, A. M. Alzahrani, S. S. Monshi, A. F. M. Alotaibi, A. I. Aljasser, K. T. Alruhaimi, S. D. K. Alotaibi, A. K. Alsultan, M. S. Arafat, A. Aldhabib, dan E. E. AbdEllatif, "Factors affecting the length of stay in the intensive care unit among adults in Saudi Arabia: A cross-sectional study," *National Library of Medicine*, 2023. [Online]. Tersedia: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10649797>.
- [2] I. Rosidawati dan S. Hodijah, "Hubungan antara lama rawat dengan tingkat kecemasan keluarga pasien di ruangan intensive care unit RSUD Dr. Seokardjo Kota Tasikmalaya," *Jurnal Kesehatan Masyarakat Unesa*, vol. 7, no. 1, 2019. [Online]. Tersedia: <https://doi.org/10.36085/jkmu.v7i1.308>.
- [3] R. A. Fauzi, A. Y. Jufan, dan Y. Widyastuti, "Identifikasi Faktor Risiko Prediktif Prolonged Length of Stay (PLOS) di ICU RSUP Dr. Sardjito Yogyakarta," *Jurnal Komplikasi Anestesi*, vol. 9, no. 1, pp. 50-57, 2023. [Online]. Tersedia: <https://doi.org/10.22146/jka.v9i1.8517>
- [4] Kementerian Kesehatan Republik Indonesia, (2024) Profil Kesehatan Indonesia 2023. [Online]. Tersedia: <https://kemkes.go.id/id/profil-kesehatan-indonesia-2023>.
- [5] N. K. A. Astiti, H. I. Kusumawati, dan Sutono, "Analisis data sistem rujukan terintegrasi (SISRUTE) di instalasi gawat darurat (IGD) infeksi RSUP Dr. Sardjito," *Jurnal Manajemen Pelayanan Kesehatan*, vol. 26, no. 1, pp. 1–7, 2023.
- [6] A. Lamsaf, R. Carrilho, J. C. Neves, dan H. Proença, "Causality, machine learning, and feature selection: A survey," *Sensors*, vol. 25, no. 8, p. 2373, 2025. [Online]. Tersedia: <https://doi.org/10.3390/s25082373>.
- [7] M. Z. Naser, "Discovering causal models for structural, construction and defense-related engineering phenomena," *Defence Technology*, 2024. [Online]. Tersedia: <https://doi.org/10.1016/j.dt.2024.04.007>.
- [8] Y. S. R. Nur, A. Sa'adah, D. Aldo, dan Faizah, "Implementation of PC algorithm to the incidence factor of stunting disease," dalam *2024 International Conference on Information Technology Research and Innovation (ICITRI)*, 2024. [Online]. Tersedia: <https://doi.org/10.1109/ICITRI62858.2024.10699213>.
- [9] Y. S. R. Nur, A. Sa'adah, D. Aldo, dan B. Masulah, "Causal modeling of factors in stunting using the Peter-Clark and Greedy Equivalence Search algorithms," *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 10, no. 3, 2025. [Online]. Tersedia: <https://doi.org/10.33480/jitk.v10i3.6184>.
- [10] A. Johnson, L. Bulgarelli, T. Pollard, S. Horng, L. A. Celi, dan R. Mark, (2022) MIMIC-IV (Version 2.0). PhysioNet. [Online]. Tersedia: <https://doi.org/10.13026/7vcr-e114>.
- [11] H. Gong, "Evaluation and comparison of causal discovery algorithms," Master degree project, KITopen, 2024. [Online]. Tersedia: <https://doi.org/10.5445/IR/1000164209>.
- [12] A. Ruiz de Villa, *Causal Inference for Data Science*. Shelter Island, NY: Manning, 2024.
- [13] P. Spirtes, "Introduction to causal inference," *Journal of Machine Learning Research*, vol. 11, pp. 1643–1662, 2010.
- [14] A. B. Böhmer, K. S. Just, R. Lefering, et al., "Factors influencing lengths of stay in the intensive care unit for surviving trauma patients: a retrospective analysis of 30,157 cases," *Crit Care*, vol. 18, no. R143, 2014. [Online]. Tersedia: <https://doi.org/10.1186/cc13976>.
- [15] P. Spirtes, C. Glymour, dan K. Zhang, "Review of causal discovery methods based on graphical models," *Frontiers in Genetics*, vol. 10, p. 524, 2019. [Online]. Tersedia: <https://doi.org/10.3389/fgene.2019.00524>.
- [16] R. Rahmadi, P. Groot, M. Heins, H. Knoop, dan T. Heskes, "Causality on cross-sectional data: Stable specification search in constrained structural equation modeling," *Applied Soft Computing*, vol.

- 52, pp. 687–698, 2017. [Online]. Tersedia: <https://doi.org/10.1016/j.asoc.2016.10.003>.
- [17] B. Miazgowski, C. Pakulski, dan T. Miazgowski, “Length of stay in emergency department by ICD-10 specific and non-specific diagnoses: A single-centre retrospective study,” *Journal of Clinical Medicine*, vol. 12, no. 14, p. 4679, 2023. [Online]. Tersedia: <https://doi.org/10.3390/jcm12144679>
- [18] L. Hempel, S. Sadeghi, and T. Kirsten, “Prediction of Intensive Care Unit Length of Stay in the MIMIC-IV Dataset,” *Applied Sciences*, vol. 13, no. 12, p. 6930, 2023. [Online]. Tersedia: <https://doi.org/10.3390/app13126930>.
- [19] J. Liu, E. Larson, A. Hessels, B. Cohen, P. Zachariah, D. Caplan, and J. Shang, “Comparison of Measures to Predict Mortality and Length of Stay in Hospitalized Patients,” *Nursing Research*, vol. 68, no. 3, pp. 200-209, May/Jun. 2019. [Online]. Tersedia: <https://doi.org/10.1097/NNR.0000000000000350>.
- [20] J. Kaddour, A. Lynch, M. J. Kusner, dan R. Silva, “Causal machine learning: A survey and open problems,” 2022. [Online]. Tersedia: <https://arxiv.org/abs/2206.15475>.
- [21] K. Debeire, A. Gerhardus, J. Runge, dan V. Eyring, “Bootstrap aggregation and confidence measures to improve time series causal discovery,” dalam *Proceedings of Machine Learning Research*, vol. 236, 2024, pp. 1–28. [Online]. Tersedia: <https://arxiv.org/abs/2306.08946>.
- [22] Y. S. R. Nur, R. Rahmadi, dan C. Effendy, “Causal modeling between factors on quality of life in cancer patients using S3C-latent algorithm,” *Journal of Information Systems Engineering and Business Intelligence*, vol. 7, no. 1, pp. 74–83, 2021. [Online]. Tersedia: <https://doi.org/10.20473/jisebi.7.1.74-83>.
- [23] J. C. Hurley, “How to apply structural equation modelling to infectious diseases concepts,” *Clinical Microbiology and Infection*, vol. 28, no. 12, pp. 1567–1571, 2022. [Online]. Tersedia: <https://doi.org/10.1016/j.cmi.2022.05.028>.
- [24] Y. Wen, J. Huang, S. Guo, Y. Elyahu, A. Monsonego, H. Zhang, Y. Ding, dan H. Zhu, “Applying causal discovery to single-cell analyses using CausalCell,” *eLife*, vol. 12, p. e81464, 2023. [Online]. Tersedia: <https://doi.org/10.7554/eLife.81464>.