



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Causal Analysis for Multivariate Integrated Clinical and Environmental Exposures Data

Sharmin Akter

RADI 6336641



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Background

- Understanding causal relationships between clinical outcomes and environmental exposures is crucial.
- Such insights can enhance public health interventions and personalized medicine.
- Traditional models often focus on associations rather than causation.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Causal Relationship vs Correlation

Causal Relationship:

- **Exposure:** Smoking cigarettes
- **Outcome:** Lung cancer

Here, the link is **causal** because:

- Smoking introduces carcinogens into the lungs.
- These carcinogens damage DNA and cells.
- Long-term exposure **directly leads to** higher incidence of lung cancer.



Causal Relationship vs Correlation

Correlation:

- Data shows that **ice cream sales** and **drowning incidents** both **increase in the summer**.
- A **positive correlation** exists: as one goes up, so does the other.

But... Does Ice Cream Cause Drowning?

- **No!** This is a **spurious correlation** — they're related **because of a third factor: hot weather**.
- Hot weather increases:
 - Ice cream sales (people want to cool off).
 - Swimming activity (more people go to pools/lakes).
 - Thus, risk of drowning also increases.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Study Objective

- Apply a causal inference algorithm to an EHR dataset.
- Identify causal relationships between various factors and asthma attacks.
- Utilize simulated interventions to assess causal effects.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Data Source

- Data obtained from the Integrated Clinical and Environmental Exposures Service (ICEES).
- Focused on patients with asthma or related respiratory conditions.
- Sample size: 14,937 patients.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Features Analyzed

- Demographic factors: sex, race.
- Clinical measures: obesity, prednisone use.
- Environmental exposures: airborne particulate matter, proximity to major roadways, residential density.
- Outcome measure: annual number of ED or inpatient visits for respiratory issues.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Methodology Overview

1. Extracted a multivariate feature table from ICEES.
2. Applied a Principal Component algorithm for causal inference.
3. Conducted simulated interventions to observe effects on asthma attack frequency.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

1. Feature Selection

Feature variable	Variable definition and enumeration
Sex	Male (0), Female (1)
Race	White, Black African American, Asian, Native Hawaiian/Pacific Islander, American/Alaskan Native, Other, Unknown
Prednisone	Common medication for asthma-like conditions (1 = Yes, 0 = No)
Obesity	Diagnostic code for obesity anytime over 'study' period (1 = Yes, 0 = No)
Airborne particulate exposure	Abbreviated herein as "PM2.5Exposure". US Environmental Protection Agency estimated maximum daily exposure to particulate matter ≤ 2.5 -microns in diameter over 'study' period, binned using pandas.cut
Roadway exposure	Abbreviated herein as "RoadwayExposure". US Department of Transportation estimated distance in meters from residential household to nearest major roadway or highway (1 = 0–49, 2 = 50–99, 3 = 100–149, 4 = 150–199, 5 = 200–249, 6 = ≥ 250 m)
Residential density	Abbreviated herein as "EstResidentialDensity". US Census Bureau American Community Survey 2007–2011 estimated total population [block group], binned according to US Census Bureau definitions
Emergency Department or inpatient visits	Abbreviated herein as "TotalEDInpatientVisits". Total number of emergency department or inpatient visits for respiratory issue(s) over the 'study' period (0, 1, 2, 3, ...)

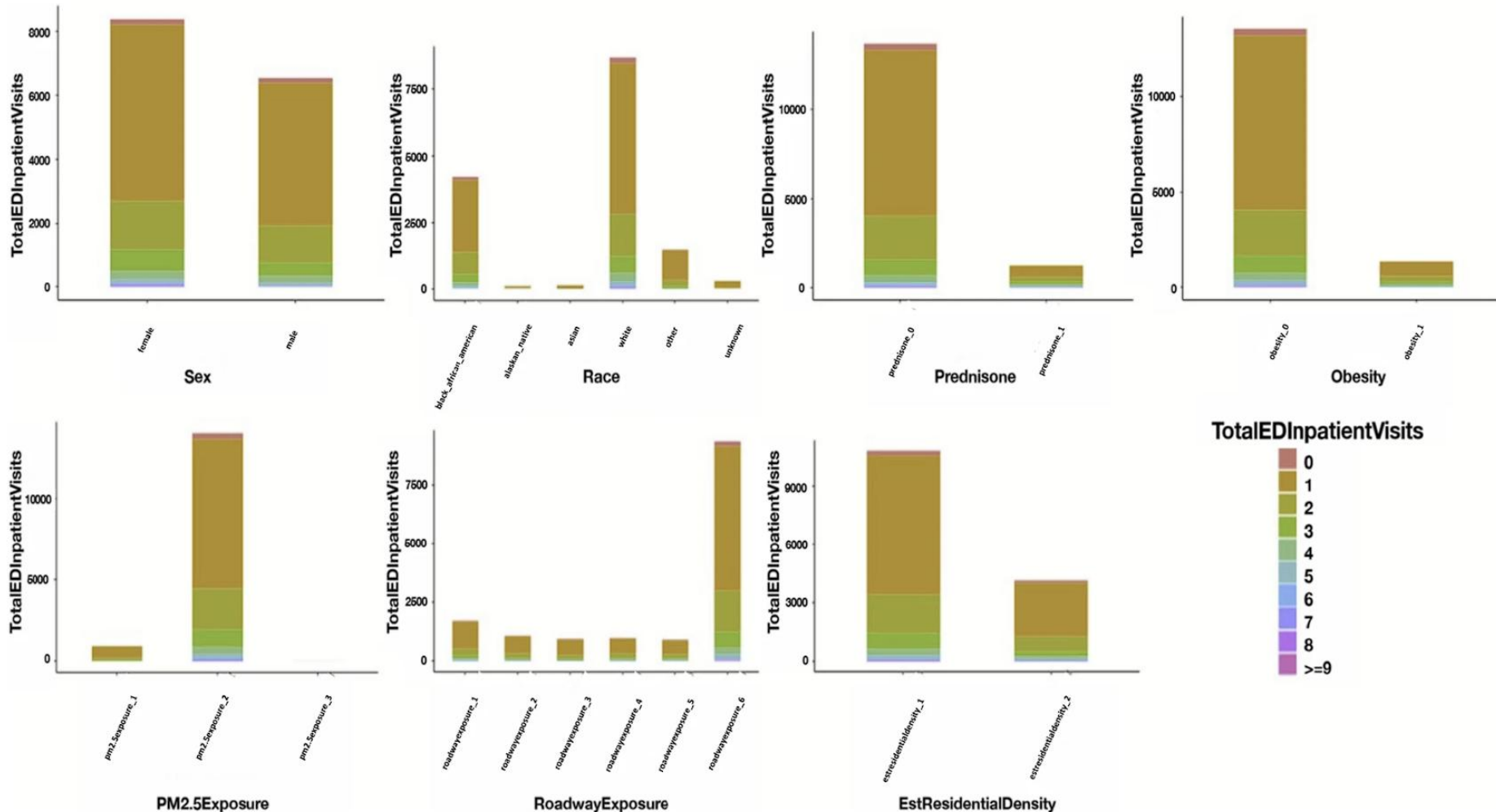


Mahidol University

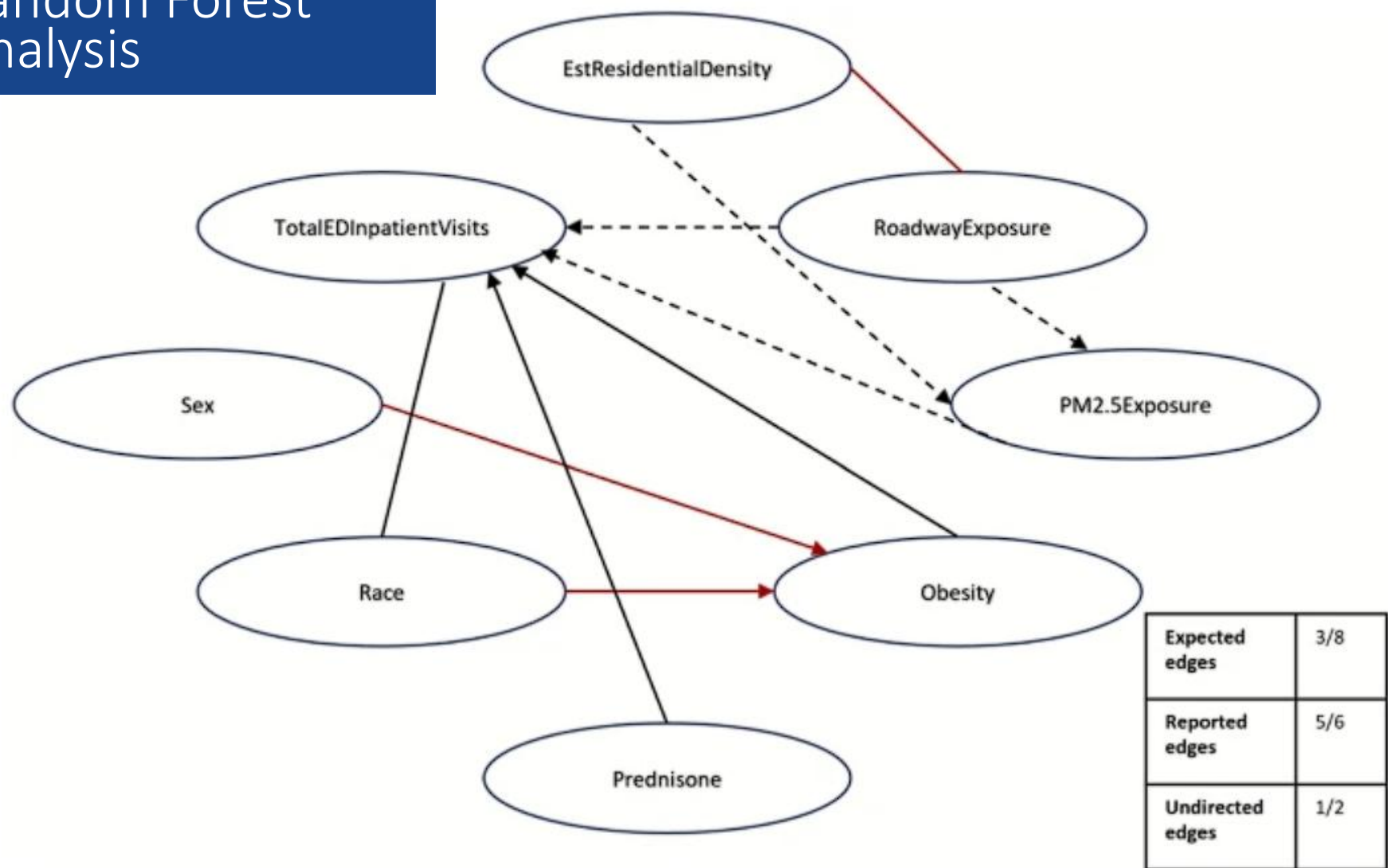
Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

1. Feature Selection



Random Forest Analysis



Inferred causal graph. Solid black lines represent inferred expected edges based on subject matter expertise combined with published literature (true positives), dashed lines represent missed expected edges (false negatives), and red lines represent unexpected edges, meaning not expected based on subject matter expertise or the published literature (false positives)



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Key Findings – Direct Effects

- Obesity and prednisone use directly increase the frequency of asthma attacks.
- These factors have a measurable causal impact on respiratory-related hospital visits.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Key Findings – Indirect Effects

- Sex and race influence asthma attacks indirectly through their relationship with obesity.
- Highlights the interconnectedness of demographic and clinical factors.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

2. Causal Inference Approach

- Principal Component algorithm identifies significant causal relationships.
- Differentiates between direct and indirect effects.
- Enhances understanding beyond mere associations.



Causal Graph Reconstruction Algorithm

Two Main Approaches to Causal Network Inference:

1. Score-Based Approach:

1. Uses a **Bayesian scoring function** to evaluate how well a graph fits the data.
2. Involves **search heuristics** to explore possible graph structures.
3. Strength: Can handle interventional data.
4. Limitation: **Not scalable** for large datasets or complex networks.

2. Constraint-Based Approach:

1. Uses **conditional independence tests** to infer causal relationships.
2. Starts with a **fully connected undirected graph**, then removes edges based on statistical tests.
3. Follows **d-separation rules** to determine independencies.



3. Simulated Interventions

- To test Claim
- (a), created a mutilated network by fixing the state of Obesity to 1, which means Obesity is forced to be present.
- (b), fixed the state of Prednisone to be 1, meaning that prednisone use was forced to be present.
- (c), fixed the state of Sex to be Male.

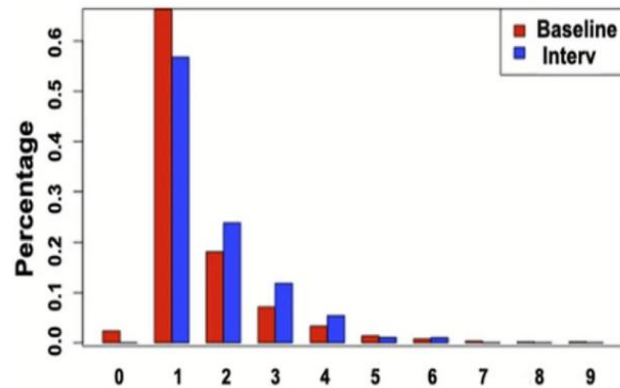


Mahidol University

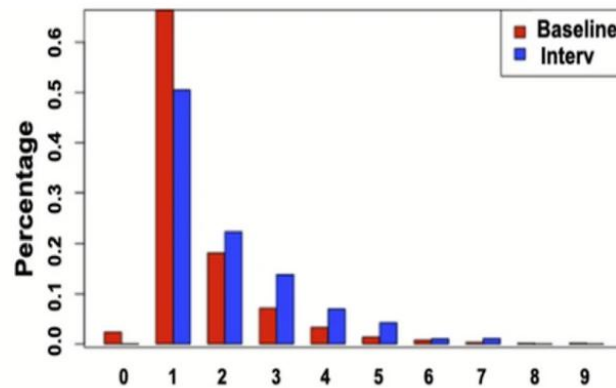
Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

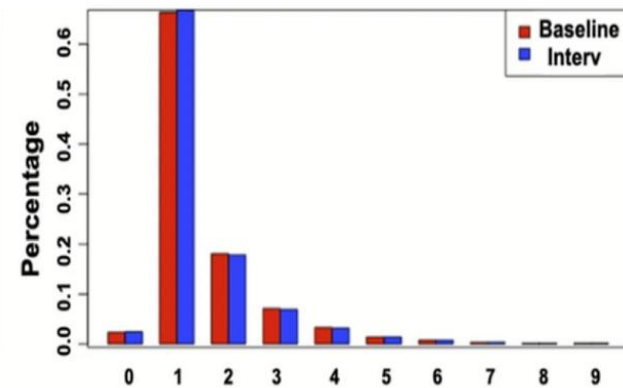
From: [Causal analysis for multivariate integrated clinical and environmental exposures data](#)



(a) TotalEDInpatientVisits



(b) TotalEDInpatientVisits



(c) TotalEDInpatientVisits

The change in the mean number (% increase) of TotalEDInpatientVisits after each intervention: **a** 0.5681 to 0.6642 mean number of visits (9.62% increase) for Obesity; **b** 0.5681 to 0.7271 mean number of visits (15.90% increase) for Prednisone; and **c** 0.5681 to 0.5722 mean number of visits (0.42% increase) for Sex. Interv = intervention



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Results from Simulated Interventions

- Simulated scenarios where all patients are obese or using prednisone.
- Observed a rightward shift in the probability distribution of asthma attacks.
- Confirms the direct causal effects identified.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Sex-Based Intervention Findings

- Simulating all patients as female did not significantly alter asthma attack frequency.
- Suggests sex alone is not a direct causal factor in this context.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Implications for Public Health

- Targeting obesity and prednisone use could reduce asthma-related hospital visits.
- Interventions should consider both direct and indirect causal pathways.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Advantages of Causal Modeling

- Provides deeper insights than traditional association models.
- Helps in designing effective, targeted interventions.
- Enhances the precision of public health strategies.



Conclusion

- Causal analysis reveals critical insights into factors influencing asthma attacks.
- Obesity and prednisone use are key direct causal factors.
- Demographic factors exert indirect effects through clinical variables.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Limitations

- Study focused on a specific patient population.
- Potential unmeasured confounders could influence results.
- Further research needed to generalize findings.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Future Research Directions

- Expand causal analysis to other health conditions.
- Incorporate additional environmental and genetic factors.
- Develop more sophisticated models for complex interactions.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Clinical Recommendations

- Monitor and manage obesity in patients with respiratory conditions.
- Evaluate the necessity and dosage of prednisone use.
- Consider demographic factors when assessing patient risk profiles.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Policy Implications

- Public health policies should address obesity as a modifiable risk factor.
- Guidelines for prednisone use may need revision based on causal impacts.
- Resource allocation can be optimized by understanding causal relationships.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Integration into Healthcare Systems

- Incorporate causal analysis findings into electronic health records.
- Use insights to inform clinical decision support systems.
- Train healthcare professionals on the importance of causal factors.



Mahidol University

Faculty of Medicine Ramathibodi Hospital

Department of Clinical Epidemiology and Biostatistics

Reference:

- <https://bmcmmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-025-02849-4#Tab1>
- <https://medium.com/walmartglobaltech/fundamentals-of-causal-discovery-and-causal-impact-analysis-41be185259b6>
- <https://www.indeed.com/career-advice/career-development/how-to-perform-causal-analysis>