# Aubrey Barnard

Machine Learning Computer Scientist

Biostatistics and Medical Informatics • University of Wisconsin–Madison

barnard@cs.wisc.edu • 608-443-9487

pages.cs.wisc.edu/~barnard • github.com/afbarnard • www.linkedin.com/in/afbarnard

June 23, 2025

## Expertise

Probabilistic graphical models, causality in observational data, medical data, temporal models, multi-relational rule learning, databases.

## Education

**PhD**, Computer Sciences                                                                        2019

> University of Wisconsin–Madison
> Dissertation: Causal Discovery of Adverse Drug Events in Observational Data
> Advisor: David Page
> Minor: Statistics

**MS**, Computer Sciences                                                                        2010

> University of Wisconsin–Madison

Special student (no degree)                                                              Spring 2005

> University of Wisconsin–Madison

**BA** *magna cum laude*, Music Theory & Composition, Computer Science                      2004

> Saint Olaf College, Northfield, MN

## Research Positions

Research Scientist, Obstetrics and Gynecology, UW–Madison                          2022–present

> Constructed matched case–control studies from electronic health records databases from multiple health systems. Analyzed them with machine learning to identify early signs of ovarian cancer. Python, SQL, Scikit-Learn, GCP, Azure.

> Mentored graduate students.

> Built tool to automatically clean laboratory test results and infer normal ranges based on convolutions and Gaussian mixture models. Python, Scikit-Learn.

> Wrote and submitted grants proposing (1) detecting early-stage ovarian cancer with machine learning analysis of observational studies, and (2) scalable causal model structure learning using pairwise log-linear Markov networks.

Proposed ovarian cancer case–control study to institutional review board. Shepherded it through approvals.

Postdoctoral Research Fellow, Biostatistics and Medical Informatics, UW–Madison          2020–2022

Invented algorithm for Bayesian network structure learning via convex optimization. Julia, Python.

Applied Bayesian network structure learning to build a causal model of dermatitis related to environmental exposures in farm environments. Python.

Wrote and was awarded grants on causal structure learning. Contributed grant sections proposing causal modeling for COVID-19.

Mentored research intern in comparing Bayesian network structure learning algorithms.

Research Assistant, Biostatistics and Medical Informatics, UW–Madison          2008–2019

Invented causal discovery machine learning and temporal inverse probability weighting methods for discovering differences between brand and generic versions of medications by analyzing controlled before–after studies. Python, Scikit-Learn.

Applied survival analysis to estimate the effects of common medications on the longevity of patients using electronic health records data. Python, R.

Developed algorithm for learning the structure of causal dynamic Bayesian networks by fitting temporal Markov networks to medical event sequences. Python, Julia.

Researched scores for identifying causal relationships among proposed cause–effect pairs. Reduced confounding by adjusting scores with a probabilistic model of patient event sequences. Python, Go, Fortran, C/C++, SQL, AWS.

Phenotyped adverse effects of drugs by learning relational rules with inductive logic programming. Prolog, Python, SQL.

Implemented statistical relational model that probabilistically combined relational rules using a tree-augmented naïve Bayesian network. Java.

Cleaned, transformed, and analyzed relational electronic health records data containing millions of patients and 50k types of events. Unix core utilities, SQL, Python.

Analyzed experimental results with statistical analyses corrected for multiple testing. Visualized data and results with charts and plots. Python, R, Matplotlib, Gnuplot, Scikit-Learn, Java.

Managed 1000s of long-running, parallel computing jobs. Linux, Bash, Make.

Organized and supervised research group meetings.

Applied Scientist Intern, Comprehend Medical, Amazon          2017

Developed recurrent neural network model of medical event sequences for summarizing medical histories of patients. Visualized clusters of patients with t-SNE. Python, PyTorch, AWS.

Research Assistant, Computer Science, Saint Olaf College          2003

Refactored and extended software for drawing fractals. Designed and built web tool for degree contracts. Java, GUI, HTML, XML, DTD, XSL.

## Current Research Projects

Identifying ovarian cancer earlier by inspecting electronic health records. Python, SQL, Scikit-Learn.

Pairwise interactions are sufficient for independence testing; generalized Hammersley–Clifford theorem. Theory, Julia.

Bayesian network structure learning via convex optimization. Theory, algorithms, Julia.

Efficiently enumerating relevant cycles. Theory, algorithms, Python.

Scalable matching. Algorithms, Python.

Speeding up cross validation with experimental design. Algorithms, Python.

Principled, statistical comparison of graphs for evaluating structure learning. Theory, Python.

Any-time inference for log-linear Markov networks via decreasing likelihood enumeration. Algorithms.

Faster and more optimal inductive logic programming via frequent itemset mining. Algorithms, Python.

Replacing noisy-OR. Theory.

## Publications

### Manuscripts

Pairwise Interactions are Sufficient for Independence Testing. **Aubrey Barnard**, Scott Alfeld. In preparation.

### Dissertation

Causal Discovery of Adverse Drug Events in Observational Data. **Aubrey Barnard**. PhD Dissertation, Computer Sciences, UW–Madison. 2019

### Conference Papers

Temporal Inverse Probability Weighting for Causal Discovery in Controlled Before–After Studies: Discovering ADEs in Generics. **Aubrey Barnard**, Peggy Peissig, David Page. Causal Learning and Reasoning 4. 2025

Causal Structure Learning via Temporal Markov Networks. **Aubrey Barnard**, David Page. Probabilistic Graphical Models 9. 2018

Identifying Adverse Drug Events by Relational Learning. David Page, Vítor Santos Costa, Sriraam Natarajan, **Aubrey Barnard**, Peggy Peissig, Michael Caldwell. AAAI 26. 2012

### Workshop Papers

An Authentication System for Student and Faculty Projects. **Aubrey Barnard**, Richard Brown, Theodore Johnson. Midwest Instruction and Computing Symposium. 2004

Extreme Programming in the Liberal Arts Classroom: A Progress Report. Richard A. Brown, **Aubrey F. Barnard**, Matthew T. Bills, Michael W. Bongard, Aaron F. Etshokin,Theodore M. Johnson, Michael R. Zahniser. Midwest Instruction and Computing Symposium. 2004

Google Scholar Profile

## Fellowships and Awards

Postdoctoral Traineeship, Computation and Informatics in Biology and Medicine, National Library of Medicine. 2020–2022

American Family Funding Initiative Data Science Grant, American Family Insurance Data Science Institute. 2020–2021

Traineeship, Computation and Informatics in Biology and Medicine, National Library of Medicine. 2013–2015

Best focus talk, National Library of Medicine Informatics Training Conference. 2015

## Talks

Causal Model Structure Learning with Log-Linear Markov Networks. Computation and Informatics in Biology and Medicine Seminars. 2022

Discovering Adverse Effects of Generic Drugs with Temporal Inverse Probability Weighting. Computation and Informatics in Biology and Medicine Seminars. 2020, 2021

Causal Discovery of Adverse Drug Events in Observational Data. UW–Madison Computer Sciences PhD defense. 2019

Identifying Adverse Drug Events using Markov Networks and Temporal Dependence. National Library of Medicine Informatics Training Conference. **Best focus talk award**. 2015

Finding Adverse Drug Events in Observational Medical Data using Markov Networks. National Library of Medicine Informatics Training Conference. 2014

Identifying Adverse Drug Events in Observational Medical Data. Computation and Informatics in Biology and Medicine Seminars. 2014

## Posters

Temporal Inverse Probability Weighting for Causal Discovery in Controlled Before–After Studies: Discovering ADEs in Generics. Causal Learning and Reasoning 4. 2025

Causal Structure Learning via Temporal Markov Networks. Computation and Informatics in Biology and Medicine Retreat. 2018

Causal Structure Learning via Temporal Markov Networks. Probabilistic Graphical Models 9. 2018

Identifying Adverse Drug Events in Observational Medical Data using Markov Networks. Computation and Informatics in Biology and Medicine Retreat. 2014

Identifying Adverse Drug Events with Relational Learning. Computation and Informatics in Biology and Medicine Retreat. 2013

Identifying Adverse Drug Events by Relational Learning. AAAI 26. 2012

Identifying Adverse Drug Events with Relational Learning. Observational Medical Outcomes Partnership Symposium. 2012

## Teaching

CS Department Tutor, Computer Sciences Learning Center, UW–Madison 2018

> Helped drop-in students with introductory and intermediate programming assignments using teaching techniques learned in the course Theory and Practice of CS Education.

Private Tutor 2012

> Instructed student in political science. Weekly meetings for a semester.

Private Tutor 2010

> Instructed student in math, statistics. Weekly meetings for a semester.

Academic Match Tutor, Greater University Tutoring Service, UW–Madison 2008

> Instructed student in Matlab for engineering assignments. Weekly meetings for ½ semester.

Teaching Assistant, Computer Science, Saint Olaf College 2001–2004

> Developed course materials for teaching extreme programming, graded assignments, led lectures, answered questions and taught during computer lab hours.

## Other Employment

Programmer and Technician, Electronic Data Interchange, Epic Systems Corporation 2006–2007

> Configured and customized network interfaces between Epic's ambulatory electronic medical records software and external systems for laboratory, pharmacy, etc. Supported hospitals in the installation, operation, and maintenance of such interfaces. HL7, Caché / MUMPS, VB.

Project Assistant, Center for Limnology, UW–Madison 2005

> Processed, analyzed, and visualized gigabytes of data from a water flow simulator in support of the hydrologic and biogeochemical fluxes in land–water mosaics project. Java, Excel, VBA.

Consultant Programmer, Dunn County Health Department 2004

> Designed and implemented interactive applet for educating the public on indoor air quality and healthy homes. Java, GUI, XML.

## Leadership

Coordinator, AI Reading Group 2017–2019

> Solicited and organized presentations on artificial intelligence and machine learning for weekly meetings during the spring and fall semesters. Moderated discussions. Created and maintained web page and meeting archive. Presented when needed.

Coordinator, Time Series Analysis Reading Group                                     Summer 2015
       Organized and moderated weekly discussions of textbook material on time series analysis.

## **Reviewing**

| | |
|---|---|
| NeurIPS | 2019, 2021 |
| ICML | 2020 |
| AAAI | 2015, 2016, 2017 |
| UAI | 2017 |
| KDD | 2015 |
| ECML-PKDD | 2013 |

## **Lectures / Informal Presentations**

| | |
|---|---|
| Inference via low-dimensional couplings.  AI Reading Group. | 2018 |
| Mastering the game of Go without human knowledge.  AI Reading Group. | 2017 |
| Finding optimal Bayesian networks.  AI Reading Group. | 2017 |
| Statistics Done Wrong.  AI Reading Group. | 2016 |
| AlphaGo.  AI Reading Group. | 2016 |
| Lasso.  AI Reading Group. | 2015 |
| Causal inference from observational data.  AI Reading Group. | 2015 |
| Markov chain Monte Carlo.  AI Reading Group. | 2014 |
| Stochastic processes.  AI Reading Group. | 2013 |
| Relational dependency networks.  AI Reading Group. | 2012 |
| Scaling Markov logic networks with Tuffy.  AI Reading Group. | 2012 |

## **Open Source Software**

esal: Event sequence analysis library.  Python.

Roc: Evaluating classification results with ROC and PR curves.  Java.

go-lbfgsb: Go interface for Fortran L-BFGS-B optimizer.  Go, C, Fortran.

libDAI: Extended Python interface for C++ probabilistic inference library.  Python, C++, Swig.

## **Computer Languages**

Recent heavy use: Python, Bash, SQL, Julia, Make, LaTeX.

Recent light use: R, HTML, CSS, C++.

Previous heavy use: Java, Prolog, Scheme, Go, C, C++, HTML, CSS, XML, DTD.

Previous light use: Matlab / Octave, Basic, OCaml, Fortran, MUMPS / Caché, Maple / Maxima, XSLT, Perl.

Learning: Rust.

## Human Languages

German: CEFR B1 / Limited Working Proficiency (self-assessed)

French: CEFR A2 / Elementary Proficiency (self-assessed)

## References

Irene Ong
Associate Professor
Obstetrics and Gynecology, Biostatistics and Medical Informatics
University of Wisconsin–Madison
irene.ong@wisc.edu
https://www.obgyn.wisc.edu/directory/profile/60039538

David Page
Professor
Biostatistics and Bioinformatics
Duke University
david.page@duke.edu
https://scholars.duke.edu/person/david.page

Mark Craven
Professor
Biostatistics and Medical Informatics
University of Wisconsin–Madison
craven@biostat.wisc.edu
https://www.biostat.wisc.edu/~craven/

Scott Alfeld
Associate Professor
Computer Science
Amherst College
salfeld@amherst.edu
https://www.amherst.edu/people/facstaff/salfeld