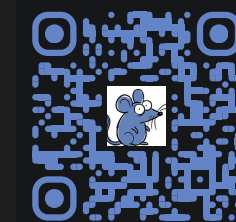


Can and Should Large Language Models be Used as Tools for "Thinking Epidemiologically"?



Ashley I Naimi, PhD | Dept of Epidemiology
Professor | Emory University

✉ ashley.naimi@emory.edu
🌐 <https://ainaimi.github.io/>



EMORY

ROLLINS
SCHOOL OF
PUBLIC
HEALTH

AI Implicates Many Things at Once

- **Ethical** — plagiarism, attribution, fraud, deception
- **Legal** — copyright, liability, regulation, disclosure
- **Environmental** — energy, water, carbon, compute
- **Social** — labor, inequities, trust, attention
- **Scientific** — discovery, evidence, reproducibility, verification
- **Educational** — learning, assessment, de-skilling, mentorship
- **More ...**

These pull in many directions: genuinely useful; overhyped and opaque distractions; quietly corrosive to scientific and societal norms.

This talk: Caveat Emptor

This talk is based largely on my experiences

There is a need in the field for a more systematic evaluation of these tools

(Also, students, beware ...)

Not all AI is created equal

There is a (usually) large difference between:

- free and paid frontier models
- terminal-based (e.g., Claude-Code), desktop-based (e.g., Claude Cowork), and web chat interface
- open (e.g., models available via Ollama) and closed (e.g., frontier models)

Not all problems that AI can help with are created equal

Problems fall on a spectrum:

Verifiable

Unverifiable



Citations
Plot creation
Administrative Tasks

Original Research
Solutions to Certain Math Problems
Certain Analysis

Where AI is useful in epidemiology

1. Literature Discovery and Synthesis
2. Knowledge Management & Idea Interrogation
3. Coding and Code Review
4. Reproducibility Infrastructure
5. Administrative Support

Some General Principles

- **Provide Context and Constraints**
- **Record and Document Output**
- **Verify and Unit Test**

An example: Data Science Pipelines



THE PREPRINT SERVER FOR HEALTH SCIENCES

HOME | SUBMIT | FAQ | BLOG | ALERTS / RSS | RESOURCES
| ABOUT

Search



Advanced Search

Follow this preprint Previous

Next

Environmental Chemicals as Modifiers of the Association between Age and Ovarian Reserve

Posted December 11, 2025.

Ashley I Naimi, Edward H. Kennedy, Ya-Hui Yu, Russ Hauser, Lidia Minguéz-Alarcon, Audrey Gaskins

doi: <https://doi.org/10.64898/2025.12.09.25341902>

This article is a preprint and has not been peer-reviewed [what does this mean?]. It reports new medical research that has yet to be evaluated and so should *not* be used to guide clinical practice.

Download PDF

Email

Print/Save Options

Share

Author Declarations

Citation Tools

Data/Code

Get QR code

Subject Area

Sexual and Reproductive Health

Abstract

Full Text

Info/History

Metrics

Preview PDF

ABSTRACT

Reviews and Context

An example: Data Science Pipelines

The screenshot shows a GitHub repository page for 'EARTH_Analysis' (Public). At the top, there are navigation buttons: Pin, Unwatch (1), Fork (0), and Star (0). Below this, the repository structure is shown with a search bar and buttons for 'Add file' and 'Code'. A commit history table lists recent updates by user 'ainaimi'. The main content area displays the README file, titled 'EARTH Treatment Heterogeneity Analysis', which describes the analysis pipeline for environmental endocrine disrupting chemicals. The right sidebar contains sections for 'About', 'Releases', 'Packages', and 'Contributors'.

EARTH_Analysis Public

main 1 Branch 0 Tags

Go to file Add file Code

Commit	Message	Time
ainaimi	updated makefile after removing sensitivity analysis	3356a7b · 6 months ago 35 Commits
code	removed eversmok since it is a function of smokstat.	6 months ago
.gitignore	updating all program files just before branch to explore a n...	7 months ago
EARTH_Analysis.Rproj	committed modified Rproj file.	7 months ago
Makefile	updated makefile after removing sensitivity analysis	6 months ago
README.md	updated readme file to reflect removal of sensitivity analysis	6 months ago

README

EARTH Treatment Heterogeneity Analysis

This repository contains the complete analysis pipeline for examining whether and how environmental endocrine disrupting chemicals modify the well established association between age and antral follicle count (a marker of fertility) in 775 women aged 21 to 46 years.

The analysis uses data from the EARTH (Environment and Reproductive Health) Study and employs the doubly robust (DR) learner, a method to estimate conditional average treatment effects. We use this learner to estimate

About

No description, website, or topics provided.

- Readme
- Activity
- 0 stars
- 1 watching
- 0 forks

Releases

No releases published
[Create a new release](#)









Packages

No packages published
[Publish your first package](#)

Contributors 1

- ainaimi Ashley I Naimi

An example: Data Science Pipelines

Name
 .gitignore
 0_data_gen.R
 1_data_man.R
 2_chem_analysis.R
 3_IF_scores_gen.R
 4_IF_scores_analysis.R
 EARTH_Analysis_Report.qmd
 Rplots.pdf

Reproducibility Makefile Pipeline

```
# Makefile for EARTH Analysis Pipeline
# Run entire pipeline with: make all
# Clean outputs with: make clean

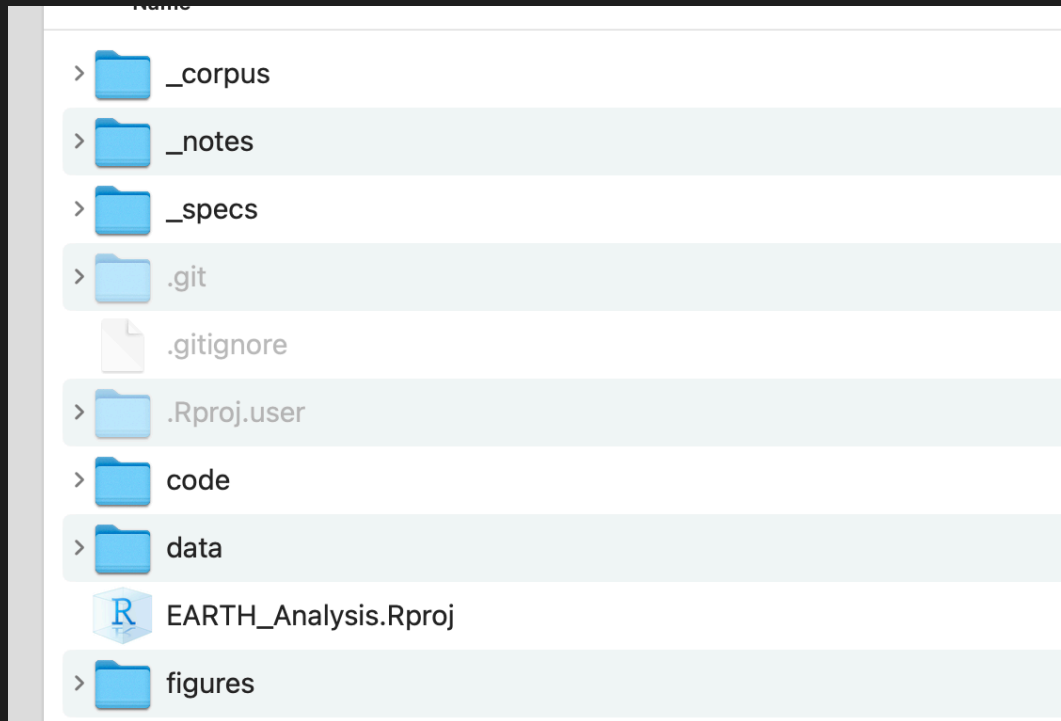
# =====
# VARIABLES SECTION
# =====

# Define the command to run R scripts
R := Rscript

# Define the command to render the Quarto report, placing output in output directory
QUARTO := quarto render code/EARTH_Analysis_Report.qmd --output-dir ../output

# =====
# DIRECTORY PATHS
# =====
```

General Principles



- The **_corpus** folder: context material directly related to reproducibility and Makefiles
- The **_notes** folder: a space to record and document output and interactions
- The **_specs** folder: context and constraints that shape AI "behavior"

Our charge

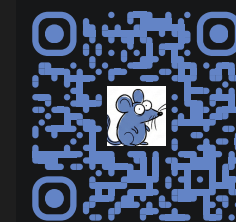
AI is a Tool

Our work is science, which is more than just an assortment of tools and methods.

AI changes what is "cheap". It does not change what we owe society at large.

There are principled and effective ways for using AI to do good science.

Can and Should Large Language Models be Used as Tools for "Thinking Epidemiologically"?



Ashley I Naimi, PhD | Dept of Epidemiology
Professor | Emory University

✉ ashley.naimi@emory.edu
🌐 <https://ainaimi.github.io/>



EMORY

ROLLINS
SCHOOL OF
PUBLIC
HEALTH