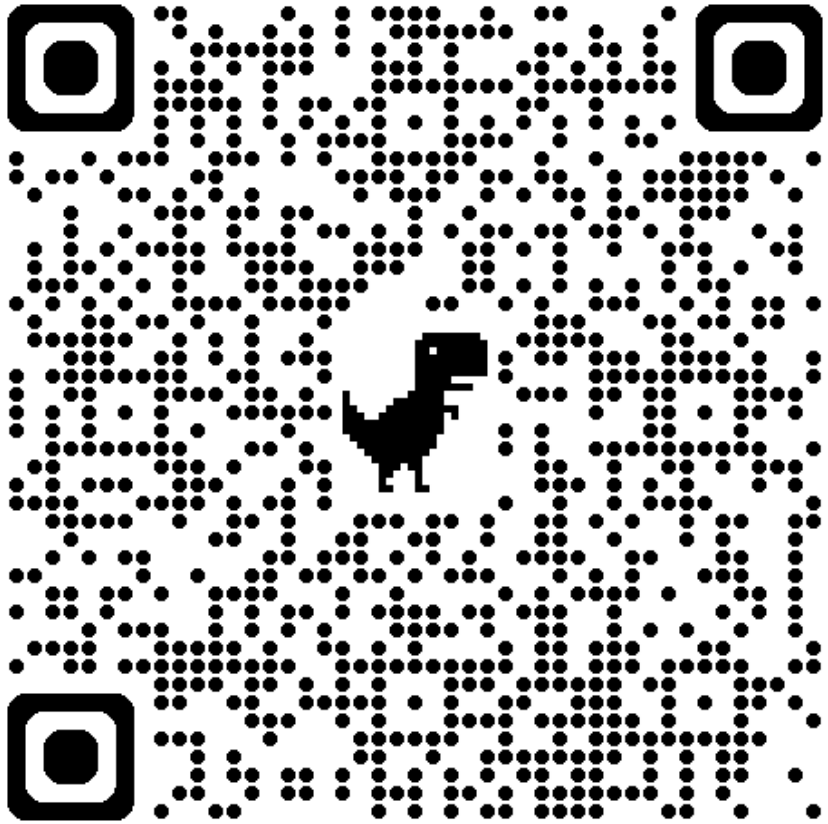
Fred Hutch Cancer Center

Responsible AI with the WHI Data Resource

Stephen Salerno, PhD
Fred Hutchinson Cancer Center

May 8, 2026

A QR code for these slides



Scope and Disclaimer

The recommendations that follow are intended as ***high-level, general guidance*** based on my experience working at the intersection of AI, statistics, and biomedical research.

Individual Perspective
Reflects my perspective as a researcher in this space.

Not Official WHI Guidance
Is not official guidance from the Women's Health Initiative.

Not Fred Hutch Policy
Does not represent policy from the Fred Hutchinson Cancer Center.

Not CDO Recommendations
Are not recommendations from the Office of the Chief Data Officer.

Appropriate use of AI in WHI-related research should always be evaluated ***case by case***, in consultation with relevant data governance and institutional guidance.

*Any concepts or examples are meant to be **illustrative** rather than exhaustive and may not apply to every setting.*

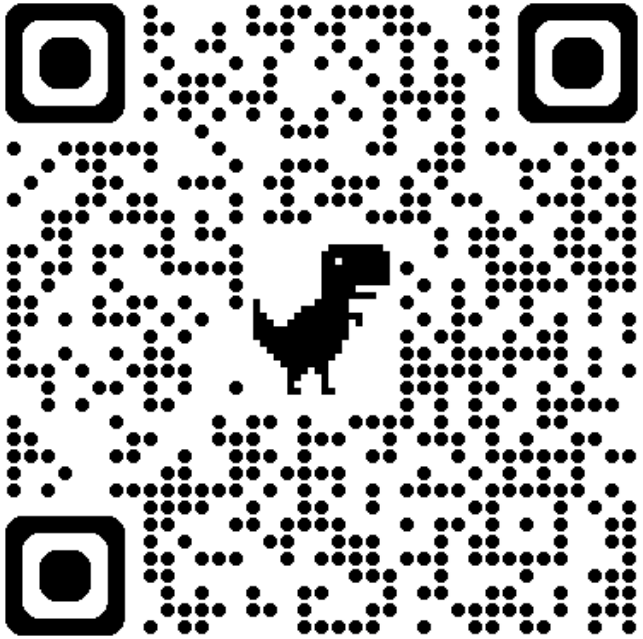
WHI Data Use Agreement

Statement on the Use of AI Tools (www.whi.org/md/DUA):

"The use of AI tools (like ChatGPT, Google Bard, natural language processing (NLP) servers, and others) is increasing rapidly in the research sphere. One concern for WHI is how these tools protect the security and privacy of the information they collect. Many AI vendors state that they will use content entered into their services to train AI models, meaning data entered into the AI product could be accessed by subsequent users of the product or could be vulnerable to a data breach at the AI vendor. Additionally, many of these AI vendors are external service providers and therefore operate outside of the governance of institutional data policies and agreements.

When an investigator signs a data use agreement (DUA) to gain access to WHI data, they agree not to share data with commercial entities or anyone else without authorization from WHI. **Be aware that entering any individual-level WHI data (or unpublished summarized data) into an external AI tool could constitute a disclosure of data that is not compliant with your DUA and not consistent with participant consent.** The few exceptions to this would require a signed legal agreement between your institution and the AI vendor and would need WHI review.

Questions about usage of AI and machine-learning tools related to WHI data should be directed to the WHI Help Desk (helpdesk@whi.org)."



Why does this conversation matter now?

AI is already part of the ***research pipeline***

Literature Search Assistants

Coding Copilots

Statistical Workflow Automation

Manuscript Editing/Drafting

Multimodal Modeling

Clinical Prediction Tools

Question: How do we use these responsibly with WHI data?

What makes WHI different?

Thinking about the WHI data resources

Longitudinal
Data spanning many collection periods.

Clinically Rich
Deep medical and health insights.

High-Impact for Policy
Informing critical public health decisions.

Multi-Investigator
Collaborative and complex access needs.

Sensitive
Requires strict privacy safeguards.

Access-Restricted
Controlled and authorized usage only.

These unique characteristics drive the need for ***responsible AI use!***

This means AI use must preserve ***trust***, ***privacy***, and ***interpretability***.

Responsible AI use depends on *where* it enters the pipeline

AI for research can potentially be used to act on:

Protected Data

Sensitive Participant Information

Methods & Logic

Code, Analysis Decisions, Interpretation

Outputs

Manuscript Text, Figures, Results

Each stage presents ***different risk levels.***

We'll consider a simple, three-tier risk framework:

01 Lowest Risk
AI assists writing or coding

- Summarizing relevant literature
- Drafting methods/analysis plans
- Suggesting general coding syntax

02 Moderate Risk
AI analyzes summaries/results

- Exploratory data summaries
- Defining new variables
- Interpreting model results/figures

03 Highest Risk
AI interacts directly with data

- Uploading datasets into LLM tools
- Using cloud-based predictive models
- Generating text based on unpublished work

These require the strictest governance.

Categories of AI Tools Researchers Use

Four major types:

Closed LLMs

ChatGPT, Claude, Gemini

Proprietary models hosted by external providers

Open LLMs

LLaMA, Mistral, Falcon

Models with public weights that can be hosted locally

Coding Agents

Copilot, Cursor, Codex

AI tools for code generation and workflow acceleration

Domain ML Models

Vision Models, Survival Nets

Specialized tasks like computer vision or survival prediction

Each behaves differently with respect to ***privacy***.

Closed LLMs

Proprietary models hosted by external providers

Examples

- ChatGPT
- Claude
- Gemini
- Codex
- Copilot

Strengths

- Strong reasoning capabilities
- Writing/editing assistance
- Advanced coding support
- Statistical/logical explanations

Risks

- External inference servers/web exposure
- Unclear data use and retention policies
- Possible training reuse of data or data leakage

Recommendation: Never paste protected WHI data into closed LLMs.

Open LLMs

Models with public weights that can be hosted locally

Examples

- LLaMA (Meta)
- Mistral / Mixtral
- Gemma
- Falcon
- Private Deployments

Strengths

- Potential secure hosting
- Reproducible pipelines
- HIPAA-style compatibility
- Auditable/version control
- Can customize cohort metadata/performance

Risks

- Local infrastructure need
- Task-specific performance gaps
- Version tracking requirements
- Data exposure if misconfigured

Recommendation: Potentially viable for protected WHI data. Needs rigorous AI protection and data governance plans, version control, and approved environments.

Coding Agents

AI tools for code generation and workflow acceleration

Examples

- GitHub Copilot
- Cursor
- Codex

Strengths

- Programming workflows
- Generate example code
- Assist with debugging
- Language translation
- Improve documentation
- Simulation study setup

Risks

- Incorrect statistical code
- Hallucinates functions
- Non-standard shortcuts
- External prompt transmission
- Data exposure via prompts

Recommendation: Potentially lower risk for WHI projects when used for code generation strictly *without* data exposure. But always review code and output for accuracy.

Domain ML Models

Specialized tasks like computer vision or survival prediction

Examples

- Vision transformers
- CNNs for radiology
- Survival deep learning
- Genomic foundation
- Models for extracting text

Strengths

- Enable analysis of complex modalities
- Scalable prediction from imaging and biosignals
- Integrate datasets
- Improve phenotyping
- Supports discovery

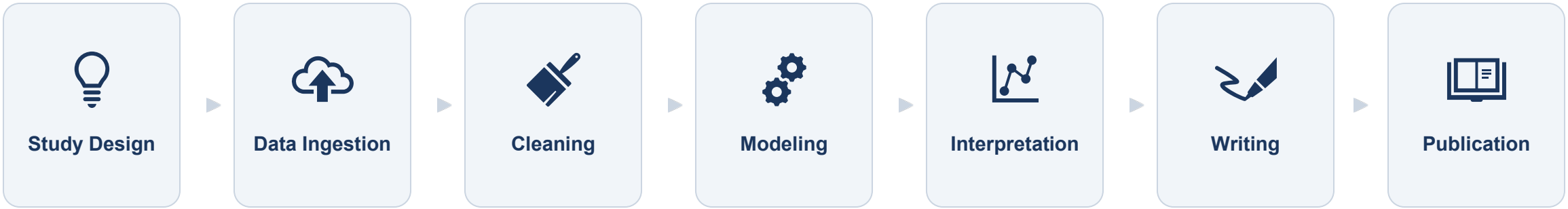
Risks

- Encode unknown biases
- External checkpoints may violate governance rules
- May require cloud GPUs
- Limited interpretability or reproducibility
- Re-identification risks

Recommendation: Use only when approved and within approved secure compute environments. Document model source, checkpoints, preprocessing steps, and configurations.

AI Can Enter the Pipeline at Many Stages

A typical research workflow:



Each stage has different ***governance*** needs.

Study Design

Ideation, power calculations, literature review

Examples

- Brainstorming research questions
- Drafting power calculations
- Summarizing public literature
- Generating dummy data structures

Governance Context

- Typically involves no participant data, making it a potentially low risk entry point for AI
- Focus on using AI to speed up the methodological and structural planning

Recommendation: Can utilize general-purpose AI tools for administrative and structural tasks while ensuring no sensitive institutional or pre-publication proprietary data are shared.

Data Ingestion

Extracting, summarizing, and modeling raw or derived data

Examples

- Summarizing participant-level tables
- Extracting signals from imaging
- Embedding clinical text
- Multimodal integration

Concerns

- PHI Exposure
- Re-identification risk
- Cloud storage leakage
- Vendor retraining risk

Recommendation: Only use with a pre-approved plan, within approved secure compute environments, and employ local inference models and institutionally managed containers.

Writing Code

Syntax generation, templates, and example frameworks

Safe Uses

- Generating regression syntax
- Producing visualization templates
- Developing simulation frameworks
- Documenting human-written code

Example Prompts

“Write an R script that demonstrates fitting a Cox model with interaction terms.”

This is generally safe, as no participant-level data is required for AI to generate this code.

“Write an R script that analyzes the data in ./phi_data.csv using a Cox model with interaction terms.”

This is not safe, as protected data are given as context for AI to generate this code.

Recommendation: Do not give data as context. AI for code generation is typically appropriate when working with public libraries, syntax structures, and non-PHI simulation parameters.

Statistical Modeling Decisions

Variable selection, hyperparameters, and analysis planning

Examples of AI Use

- Variable selection suggestions
- Hyperparameter tuning advice
- Model comparisons
- Simulation planning

Potential Risks

- Hidden assumptions in model logic
- Overconfidence in specific methods
- Hallucinated or incorrect references

Recommendation: Treat AI as another collaborator, not an authority (and seek humans who are authorities). Always verify outputs through independent statistical validation.

Interpreting Results

Summarizing outputs, interpreting the meaning of statistical measures

Examples of AI Use

- Summarizing regression outputs
- Translating concepts like hazard ratios
- Generating figure captions

Potential Risks

- Incorrect inference claims
- Causal misinterpretation
- Exaggeration of certainty

Recommendation: Do not put unpublished results from analysis of protected data into LLMs. General AI interpretations must be investigator-reviewed to ensure accuracy and validity.

Writing Assistance (*Lower Risk if Used Properly*)

High-level guidelines for safe and appropriate usage

Safe Uses

- Editing grammar
- Formatting methods
- Restructuring paragraphs
- Summarizing literature

Not Appropriate

- Generating new/first drafts of text
- Uploading data to write conclusions
- Fabricating citations
- Drafting undisclosed analyses

Recommendation: Do not use to generate text or to write conclusions for unpublished results. Always check journal/institutional AI policies for writing assistance and disclose the use of AI when allowed.

A Special Case: Training AI Models

When AI model development is itself the research objective

Common Examples

Increasingly, new research is being proposed that seeks to:

- Train prediction models on tabular data
- Integrate biomarkers into existing ML workflows
- Use free text narratives with structured data

Workflow Requirements

Specialized infrastructure is often necessary for model training:

- GPU-enabled compute
- Pretrained checkpoints or open-weight models
- Containerized workflows
- Reproducible pipelines and version tracking

Governance Expectations

Projects that propose to train AI models typically require:

- Formal analysis plan and model documentation
- Specification of compute and storage locations
- Confirmation that training occurs within secure institutional infrastructure

Recommendation: Work within approved secure environments using locally hosted, open-weight models and institutionally approved containers. Have detailed AI protection, infrastructure, model training, and data governance plans before proposing these types of projects.

Reproducibility Concerns with LLMs

LLMs introduce additional challenges:

- **Non-Deterministic Outputs:** Identical prompts can yield different results with no way of ‘seeding’ a response.
- **Version Drift:** Model updates are irregular and can change behavior and quality over time.
- **Undocumented Reasoning:** Internal logic for arriving at a result is often opaque and cannot be replicated.
- **Hidden Training Data:** Exact data sources are rarely disclosed and can leak information/lead to potential plagiarism.

Best Practices:

- Record Model Name
- Record Version
- Record Prompts
- Record Date

Treat prompts as ***analytic parameters*** in research documentation.

Some Final Recommendations for WHI Projects

To maintain data integrity and research transparency, ensure the appropriate:

- ❑ **Data Privacy:** Never upload participant-level data to external LLM APIs.
- ❑ **Infrastructure:** Prefer local inference models for sensitive workflows.
- ❑ **Documentation:** Log AI-assisted steps in analysis documentation.
- ❑ **Reproducibility:** Treat prompts as part of reproducible research.
- ❑ **Disclosure:** Disclose AI assistance in manuscripts.

Prioritize ***data security*** and ***methodological rigor*** in all AI implementations.

Documentation Expectations for AI-Assisted Research

Some minimum reporting standards

Model Identity
Specify the exact model name and version number used.

Intended Purpose
Define the specific role of AI within your research workflow.

Data Privacy
Note if any protected or sensitive participant data were involved.

Human Oversight
Confirm that all AI outputs underwent human verification.

Analytical Impact
Describe how AI influenced the final research conclusions.

Reproducibility Log
Provide a log of all prompts used to ensure analysis can be replicated.

Think of this as your ***"AI Methods Section"***

Practical Safe-Use Checklist for WHI Investigators

Some critical questions to ask before using AI:

- ❑ Does this tool see protected data?
- ❑ Does this tool store prompts externally?
- ❑ Can this step be reproduced later?
- ❑ Would I disclose this in a methods section?

If unsure, consult WHI and the appropriate data governance officials at your institution.

Opportunities for WHI in Responsible AI

WHI is uniquely positioned to:

Governance Standards
Set cohort-level AI governance standards

Reproducible Workflows
Develop reproducible AI workflows

Reporting Norms
Define reporting norms for AI-driven research

Multimodal Modeling
Enable privacy-preserving multimodal modeling

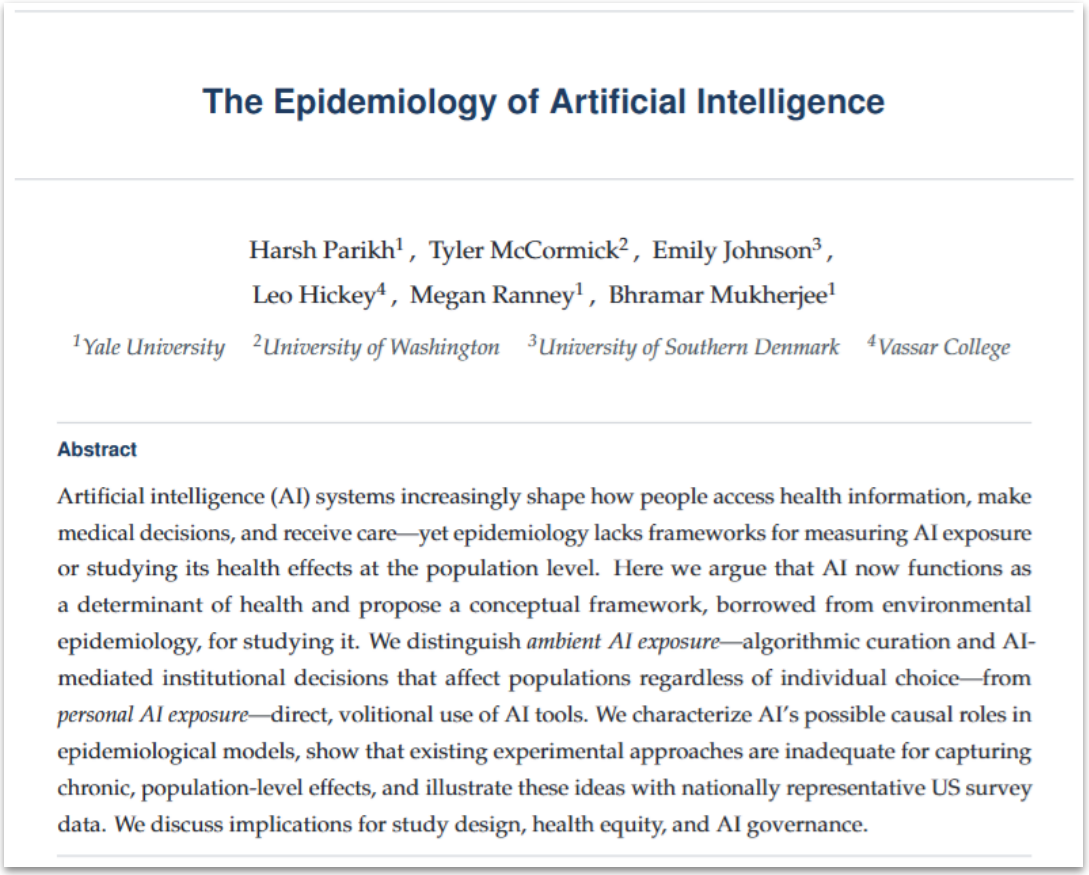
Exciting time to be thinking nationally about ***ethical*** cohort-scale AI use

AI as a Scientific Exposure

Recognizing its influence

- It is spreading rapidly across science and society
- It spans individuals, institutions, and populations
- It influences public health research and practice
- It has the capacity to do good or harm

Source: https://arxiv.org/abs/2604.14086



The Epidemiology of Artificial Intelligence

Harsh Parikh[1], Tyler McCormick[2], Emily Johnson[3],
Leo Hickey[4], Megan Ranney[1], Bhramar Mukherjee[1]

[1]*Yale University* [2]*University of Washington* [3]*University of Southern Denmark* [4]*Vassar College*

Abstract

Artificial intelligence (AI) systems increasingly shape how people access health information, make medical decisions, and receive care—yet epidemiology lacks frameworks for measuring AI exposure or studying its health effects at the population level. Here we argue that AI now functions as a determinant of health and propose a conceptual framework, borrowed from environmental epidemiology, for studying it. We distinguish *ambient AI exposure*—algorithmic curation and AI-mediated institutional decisions that affect populations regardless of individual choice—from *personal AI exposure*—direct, volitional use of AI tools. We characterize AI's possible causal roles in epidemiological models, show that existing experimental approaches are inadequate for capturing chronic, population-level effects, and illustrate these ideas with nationally representative US survey data. We discuss implications for study design, health equity, and AI governance.

Summary

AI has the potential to safely accelerate research if we are always thinking critically about

Protecting Our Data
Protected data stay protected

Reproducibility
Methods stay reproducible

Investigator-Led Interpretation
Interpretation stays investigator-led

Transparency
Documentation stays transparent

Responsible AI use is not about ***restriction***, it is about ***scientific integrity***.

Additional Helpful Resources

Courtesy of Dr. Carrie Wright

AI Attribution Toolkit

https://aiattribution.github.io/

OTTR AI Use Cheatsheet

https://www.ottrproject.org/cheatsheets/ai_everyday_uses.html

Presentation on AI Ethics

https://docs.google.com/presentation/d/15RmlGNAontco7VF4EdtQFeaOl_1QKc1TVMCb79vKKz0/edit?slide=id.g39de1049851_0_188#slide=id.g39de1049851_0_188

Fred Hutch Cancer Center

Thank You!

E: ssalerno@fredhutch.org