Systematic Review Workflow Can Already Incorporate NAMs in Evidence Syntheses

Andrew Rooney, PhD

Acting Chief, Integrative Health Assessments Branch
Division of Translational Toxicology
National Institute of Environmental Health Sciences

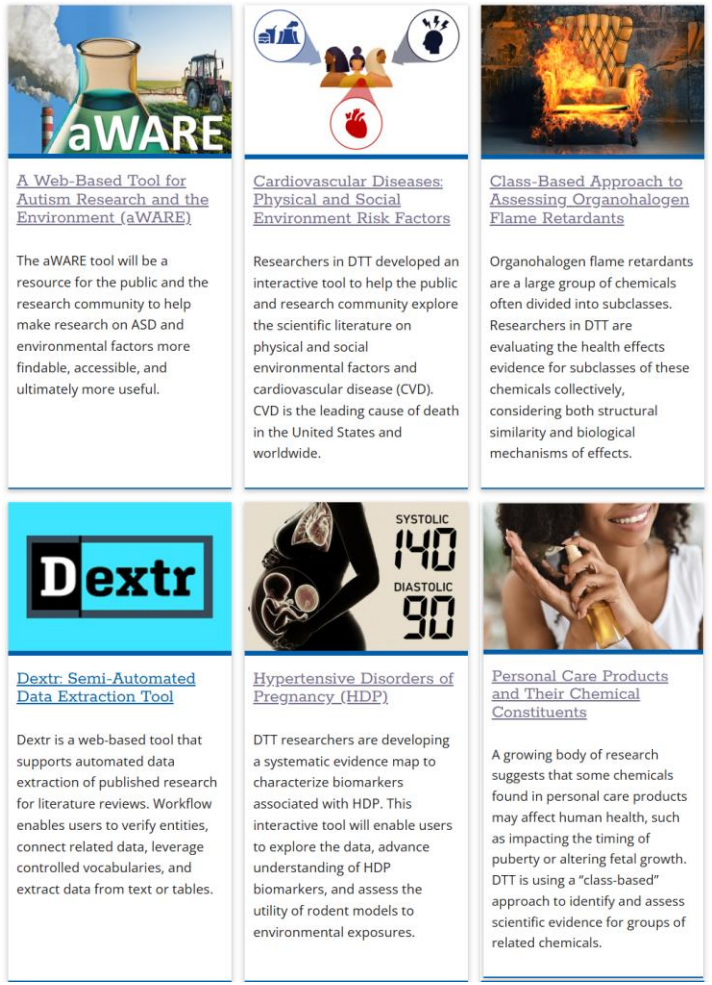
Background - IHAB

DTT Integrative Health Assessments Branch

Links previous Office of Health Assessment and Translation (OHAT) and Report on Carcinogens (RoC)

- **Critically assess the scientific evidence** that environmental substances cause adverse health effects
- **Conclusions to inform** public, government, scientific, and medical communities
- **Promote understanding and advance methodologies** using the highest standards in rigor and reproducibility
- **Develop innovative informatic approaches** anchored in systematic review requirements to support decision making in health effects research and policy

IHAB Projects and Tools



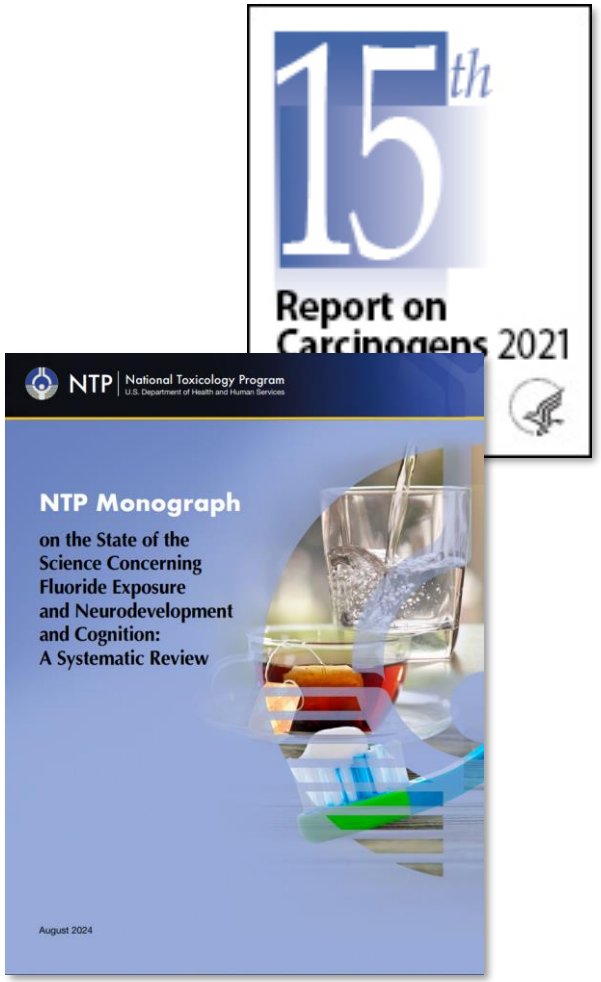
https://www.niehs.nih.gov/research/atniehs/labs/iha/projects

Rigorous methods to support decision making

Rigor, reproducibility, transparency

- Systematic reviews are essential for evidence-based assessments of potential human health effects from chemical or environmental exposures

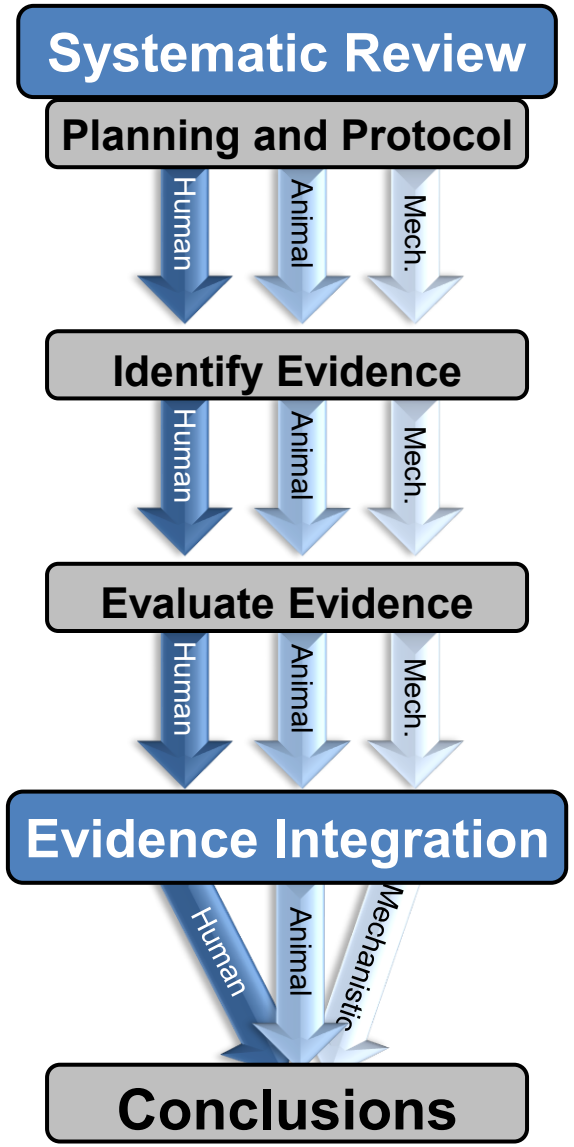
- These methods minimize bias and ensure transparency in identifying, capturing, critically assessing, and synthesizing research to answer specific questions



Across all evidence

Systematic reviews can assess all data types

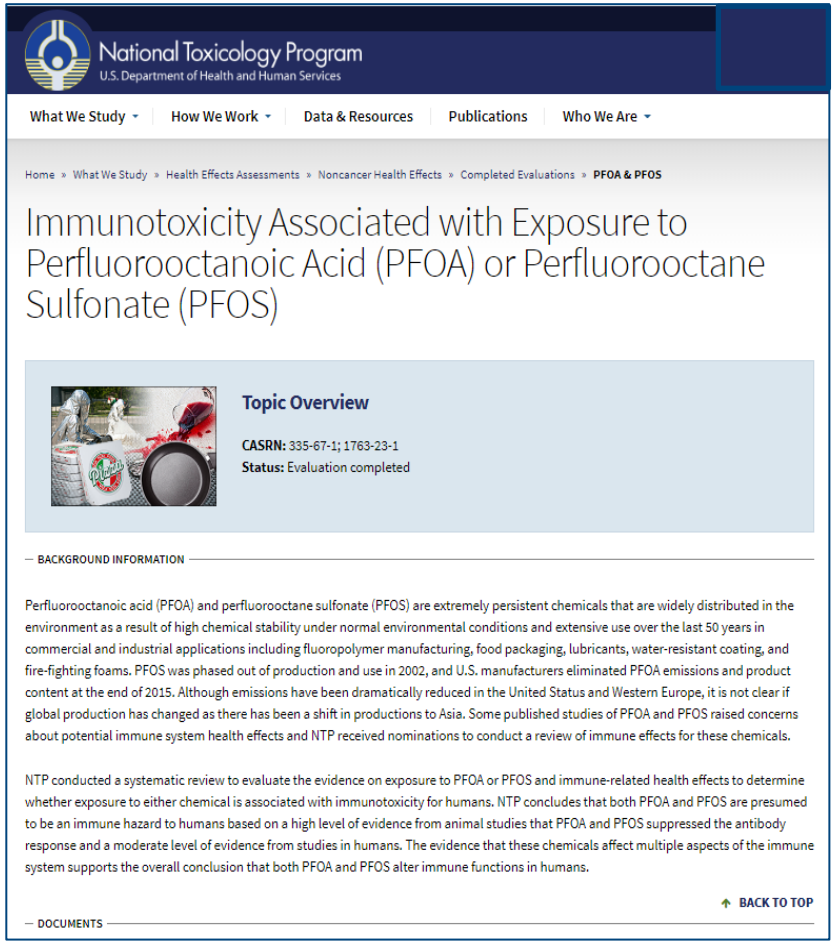
- **Established workflows**
  - Evaluating evidence across studies of the same or similar endpoints (apples to apples)
  - Growing experience with mechanistic data and NAMs
- **Evidence Integration**
  - Consideration of factors that could increase or decrease confidence in a body of data
  - Includes assessment of consistency both within and across evidence streams
  - Could be used to transparently evaluate available data to anchor NAMs data



In vitro data and NAMS

Example consideration of in vitro data

- **Objective:** to develop hazard conclusions on the association between exposure to PFOA or PFOS and immunotoxicity
- **Health effects**:
  - Strongest and sensitive data were on immunosuppression
  - Suppression of the antibody response
- Data from all three evidence streams
  - Human studies 
  - Animal studies 
  - In vitro studies 



https://ntp.niehs.nih.gov/go/749926

Apples to apples and mechanisms

We are not as good at assessing study quality or risk of bias for NAMs studies

Antibody Response and mechanisms

Human Data

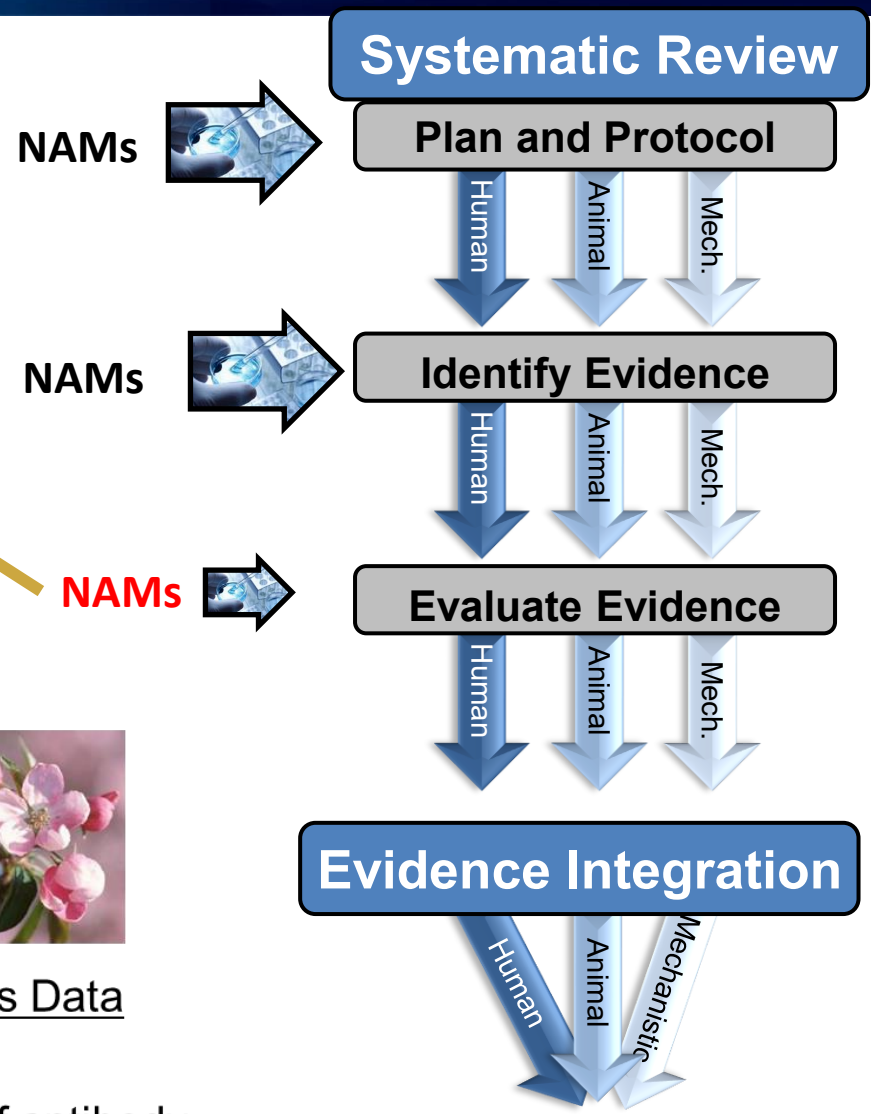
- antibodies to vaccines
- anti-rubella IgM
- anti-rubella IgG

Experimental Animal Data

- antibodies to T-cell antigens
- anti-SRBC IgM
- anti-SRBC IgG

In vitro and NAMs Data

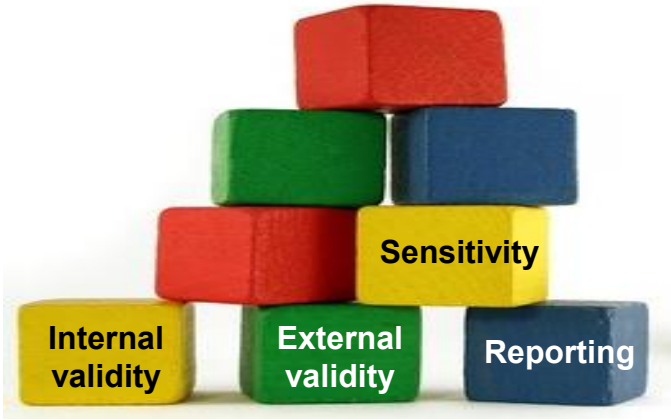
- *in vitro* IgM
- mechanisms of antibody production/response



Considering quality

Risk of bias in context

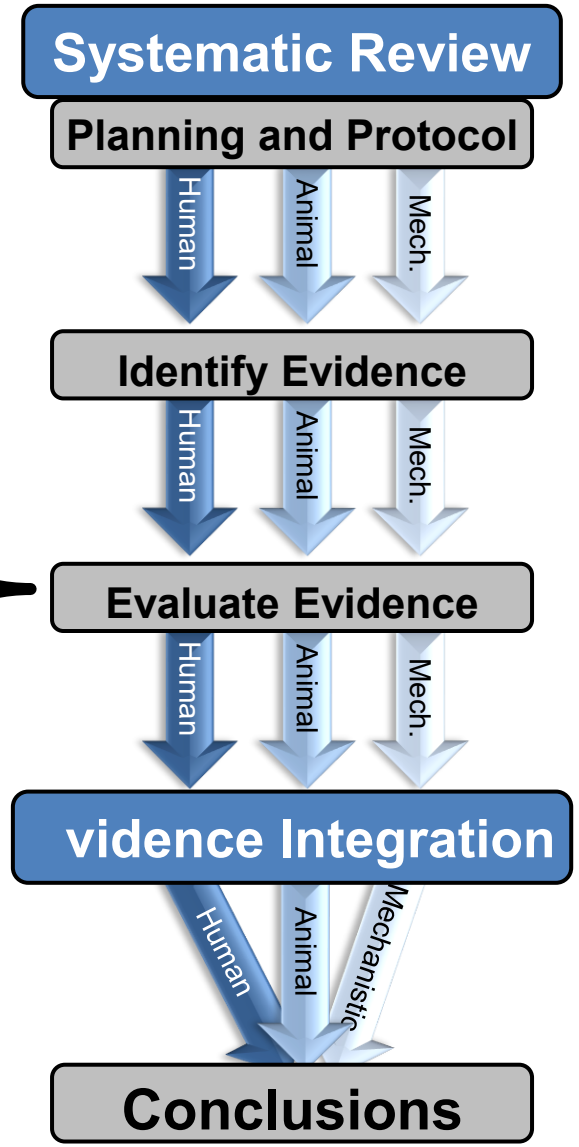
- **Reporting quality**
  - How well was the study reported?
- **Risk of bias (also called internal validity)**
  - How credible are the findings based on study design and conduct?
  - Methodological characteristics of a study that may lead to systematic errors in magnitude or direction of the results
- **Directness and Applicability (also called external validity)**
  - How well does the study address the topic under review?
    - Population of interest? Upstream indicator of primary health outcome
    - Route, timing/duration of exposure and health outcome assessment
    - Relevance of animal model for human health



Why assess risk-of-bias (RoB) in systematic reviews?

- Critical, transparent, and consistent evaluation of the body of evidence is required for a systematic review
  - Absence of RoB severely limits utility
  - Is it a systematic review without RoB?
- Characterization of the RoB of individual studies informs confidence in the body of evidence
  - Conclusions, confidence, and any discussion of uncertainty should consider entire body of evidence, and
  - With particular attention to the RoB in the best evidence

Risk-of-bias of individual studies

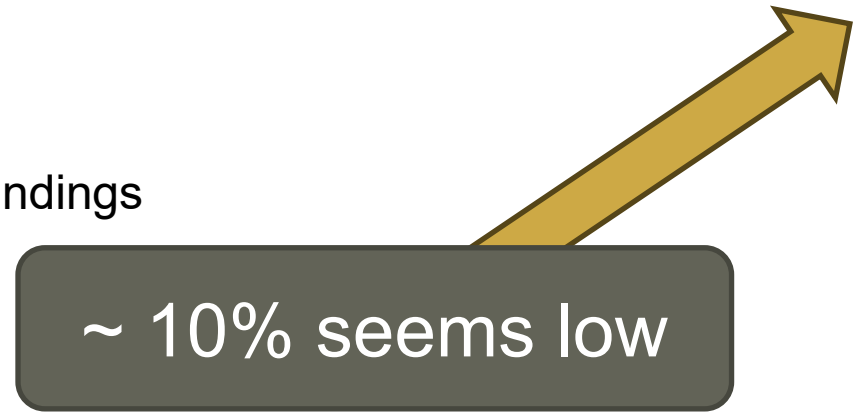


Risk of bias
is a critical aspect for
systematic reviews in
toxicology and public health,
but assessing it takes too long…

Time and labor-intensive process

Estimated percent of time for each step

Step	Percent Time
Problem Formulation and Protocol Development	8 %
Identify the Evidence (*primarily split* ***screening*** *and* ***extracting***)	39 %
Evaluate the Evidence	9 %
Summarize and Synthesize Data	5 %
Integrate Evidence and Report Findings	20 %
Project Management	19 %

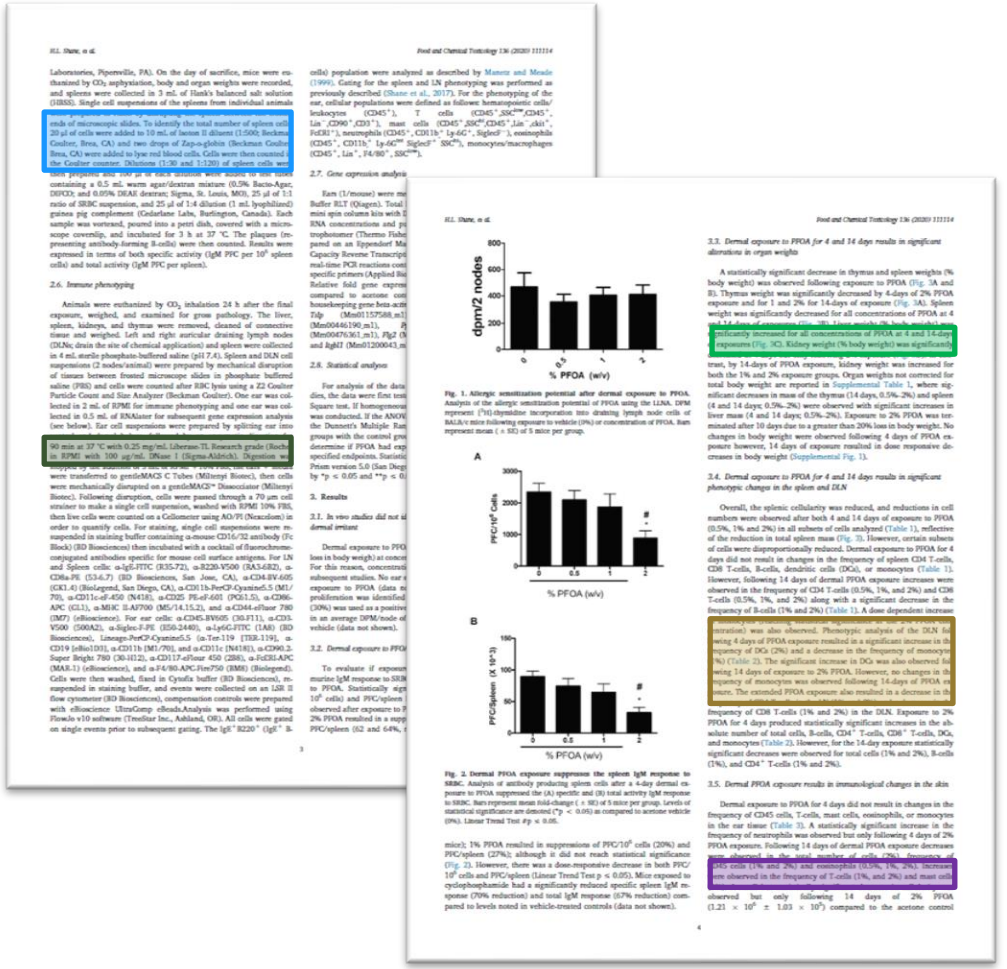


Adapted from Clark J et. al.. 2020. A full systematic review was completed in 2 weeks using automation tools: A case study. J Clin Epidemiol. 121:81-90. https://dx.doi.org/10.1016/j.jclinepi.2020.01.008

Opportunities for ML, LLM, and AI - what about Risk of Bias?

Methods for using AI in literature search and screening are increasingly common … including data extraction tools

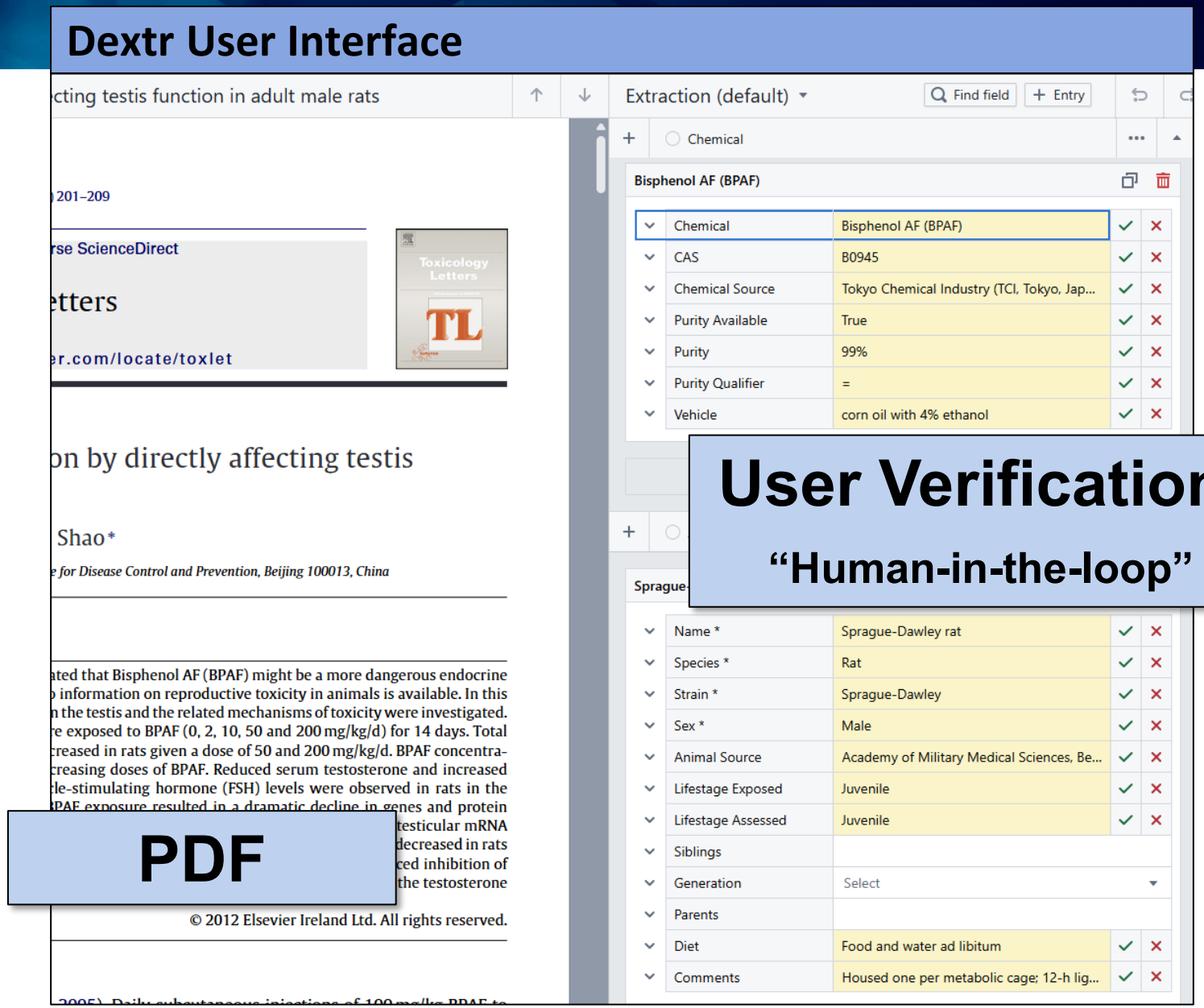
- Machine learning (ML) approaches have been used for years
- The emergence of large language models (LLMs) provides opportunities to move forward with less effort on training data sets
- These approaches can identify specific text or sections of text in published research



AI assisted extraction

Dextr

- Web-based user interface that processes full text PDFs using structure data extraction forms and AI-assisted models.
- Researcher verifies (or edits) model predictions to confirm extraction element.
- QC Module enables dual extractors or AI-assisted verification.



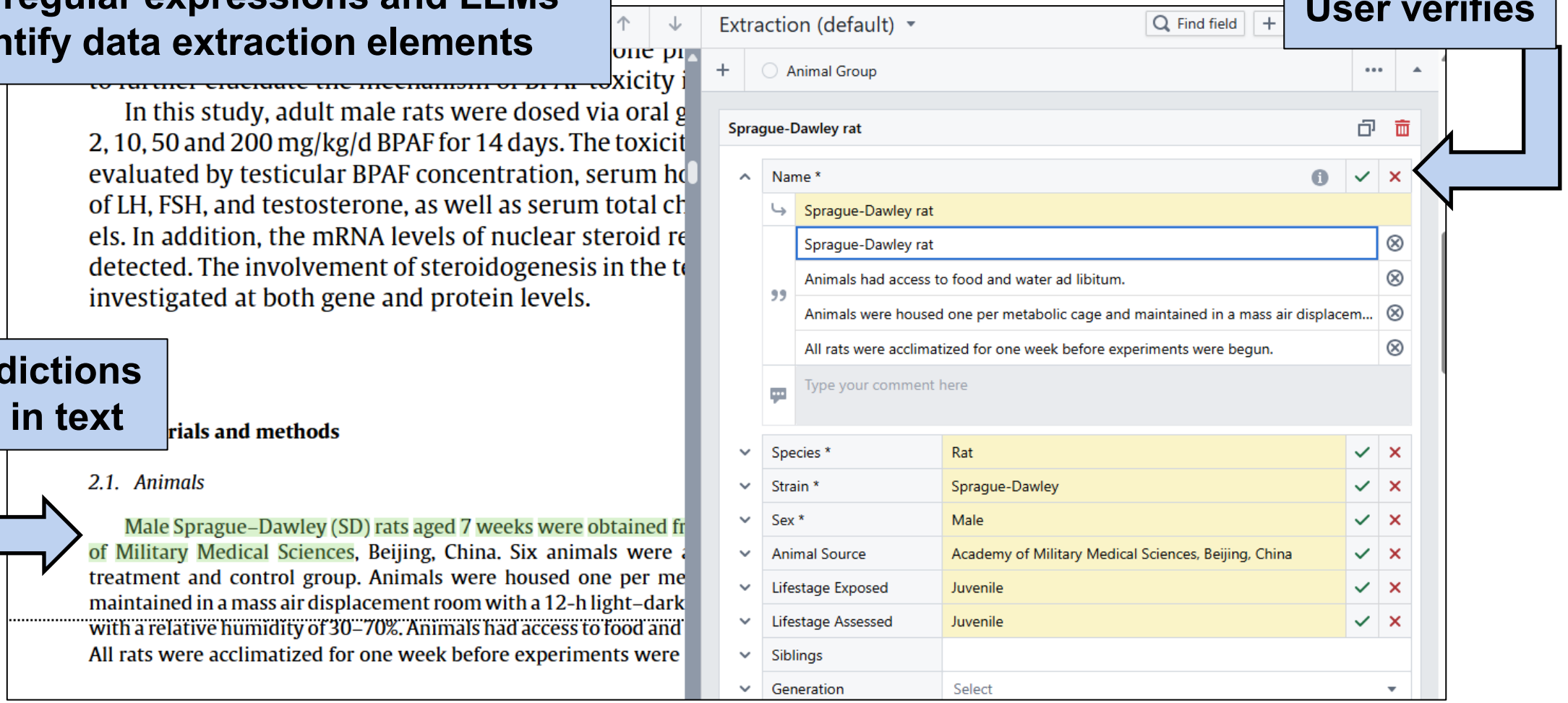
AI assisted extraction - Dextr

Dextr Integrates Multiple Models

NLP, regular expressions and LLMs identify data extraction elements

User verifies

model predictions identified in text



Risk of bias includes simple and complex concepts

Randomization

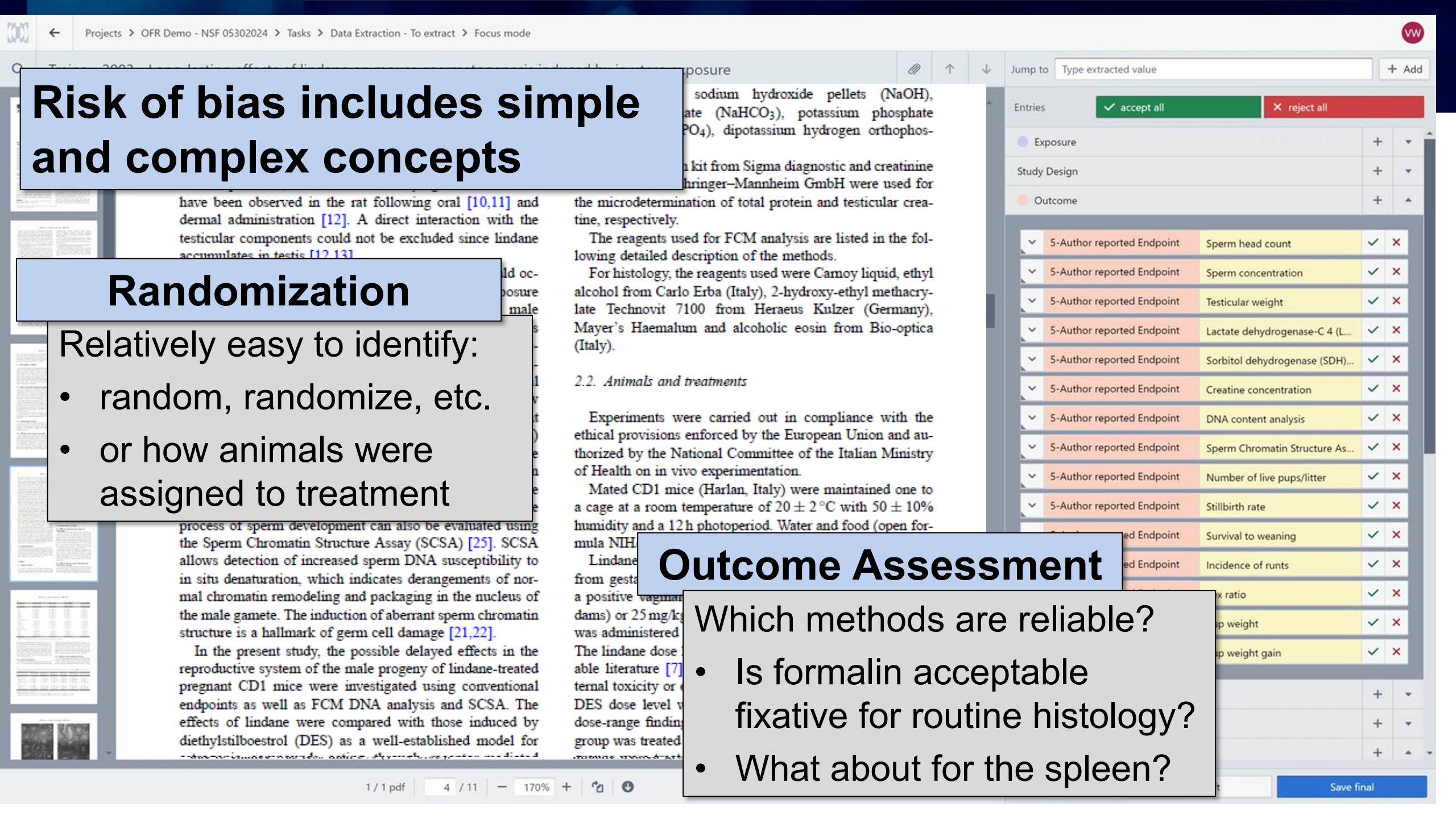
Relatively easy to identify:

- random, randomize, etc.
- or how animals were assigned to treatment

Outcome Assessment

Which methods are reliable?

- Is formalin acceptable fixative for routine histology?
- What about for the spleen?



Accuracy, verification, and acceptability

Risk of bias is already highly debated

- **Current state:**
  - Subject matter experts are critical to manual work
  - Caution and perhaps fear of AI
  - Need for transparency, standards for AI use
- **Question**: How can AI help right now?
  - We can already do the easy stuff
  - Highlight key areas of text to speed expert review
  - “human-in-the-loop” approaches
- **Question**: How do we move forward?
  - Develop, test, validate the new approaches
  - Recognize we are facing a skeptical community

Thank you