Lessons learned in protocol development and implementation of multi-endpoint systematic reviews

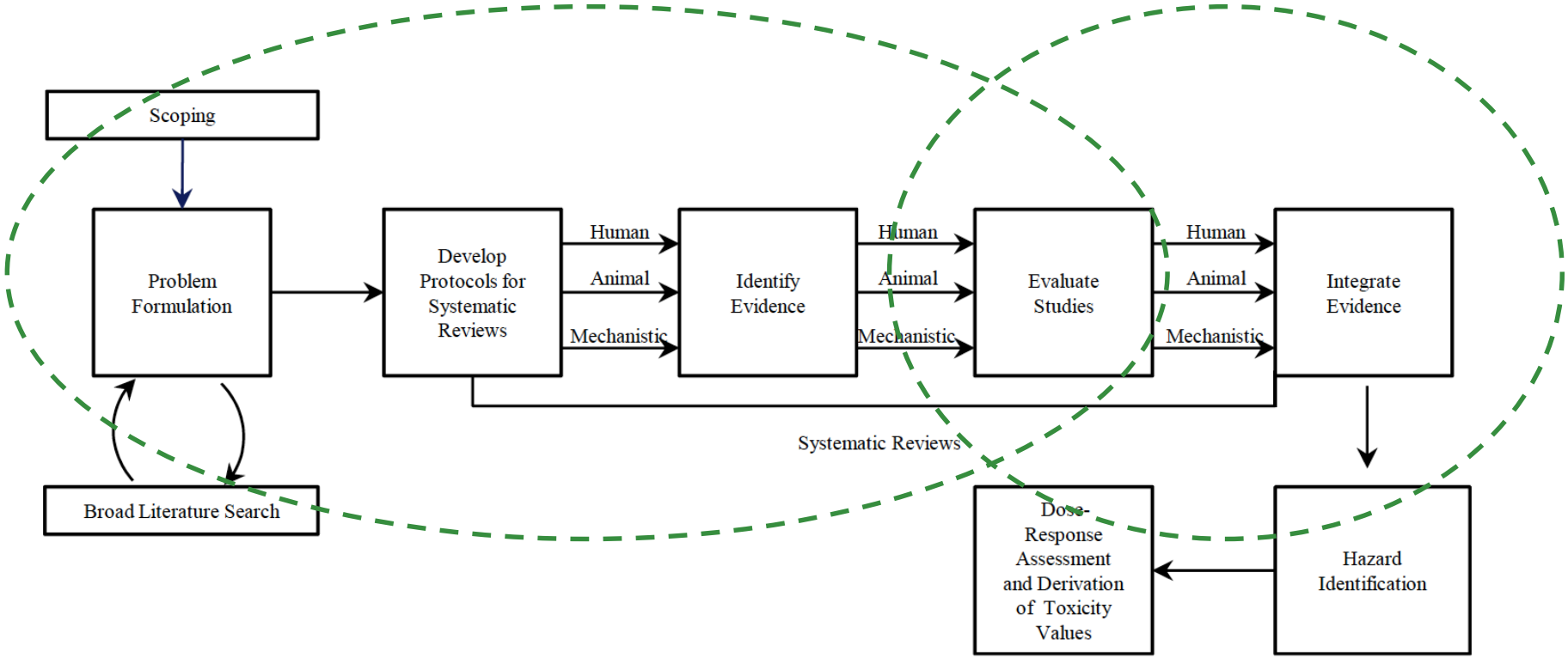
Daniele Wikoff, PhD
Health Sciences Practice Leader
ToxStrategies, Inc

ToxStrategies

ToxStrategies

- Toxicologist, ~10 years in consulting
- Consulting firm that provides toxicology and risk assessment services to private and public organizations
  - ~30 scientists (toxicologists, engineers, statisticians)
- Responsible for integrating systematic review into our health sciences practice – generally a “user” of what has already been established in EBT

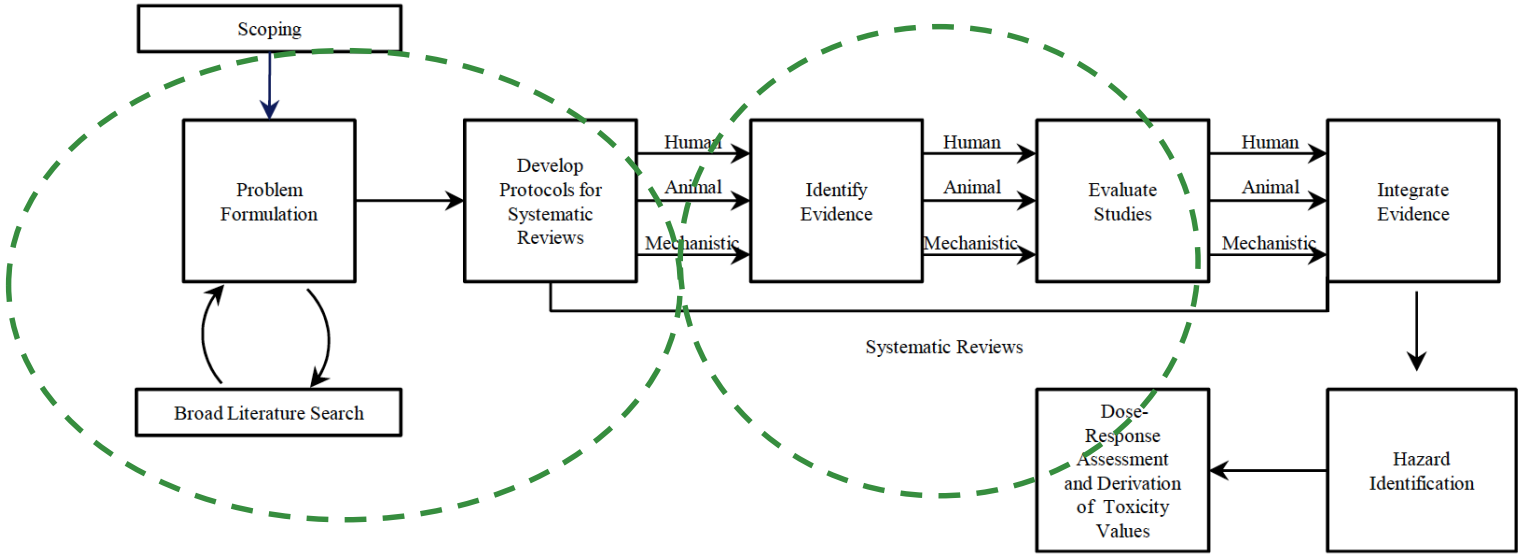
Lessons learned from ToxStrategies



NRC 2014. FIGURE S-1 Systematic review in the context of the IRIS process. The committee views public input and peer review as integral parts of the IRIS process, although they are not specifically noted in the figure.

ToxStrategies

Lessons learned from a (slightly) different perspective



Protocol development

- Developing SR team (and roles)
- PECO – multiple outcomes (and multiple endpoints)
- Publication

Implementation

- Literature searching and screening
- Individual study assessment

Practical lessons

- General conduct of SRs
- Resource allocation and time

ToxStrategies

My disclosure

The term “systematic review” gets used very loosely in the field of toxicology.

There are many “correct” ways to do systematic review, as long as the key elements are maintained (particularly emphasis on problem formulation, protocol development, transparency, and documentation).

Systematic review is an excellent tool but it does not eliminate scientific judgments (and subjectivity), rather it provides a platform to integrate and evaluate scientific judgment and data, and present findings in a systematic way.

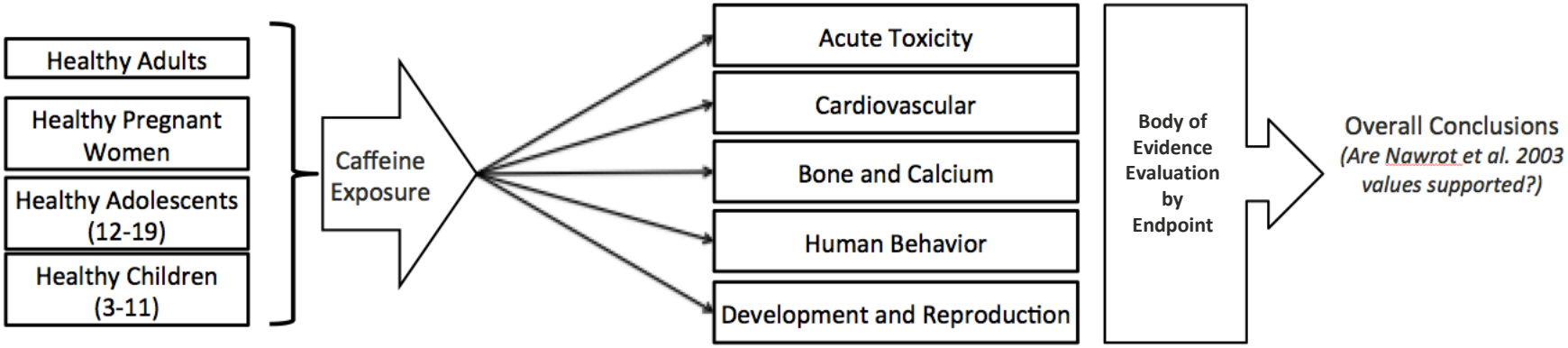
Background on projects used as case examples

Lessons come from a variety of projects

Project	Streams	Hazard	Risk	Client Type
Assessment of adverse effects of caffeine	Human (including TK)	X		Non-profit scientific foundation
Derivation of inhalation toxicity value for environmental compound	Human, animal, mechanistic	X	X	Private
Assessment of compound in common consumer product	Human, animal, mechanistic (including TK)	X		Government

And others…

Analytic framework for the adverse effects of caffeine

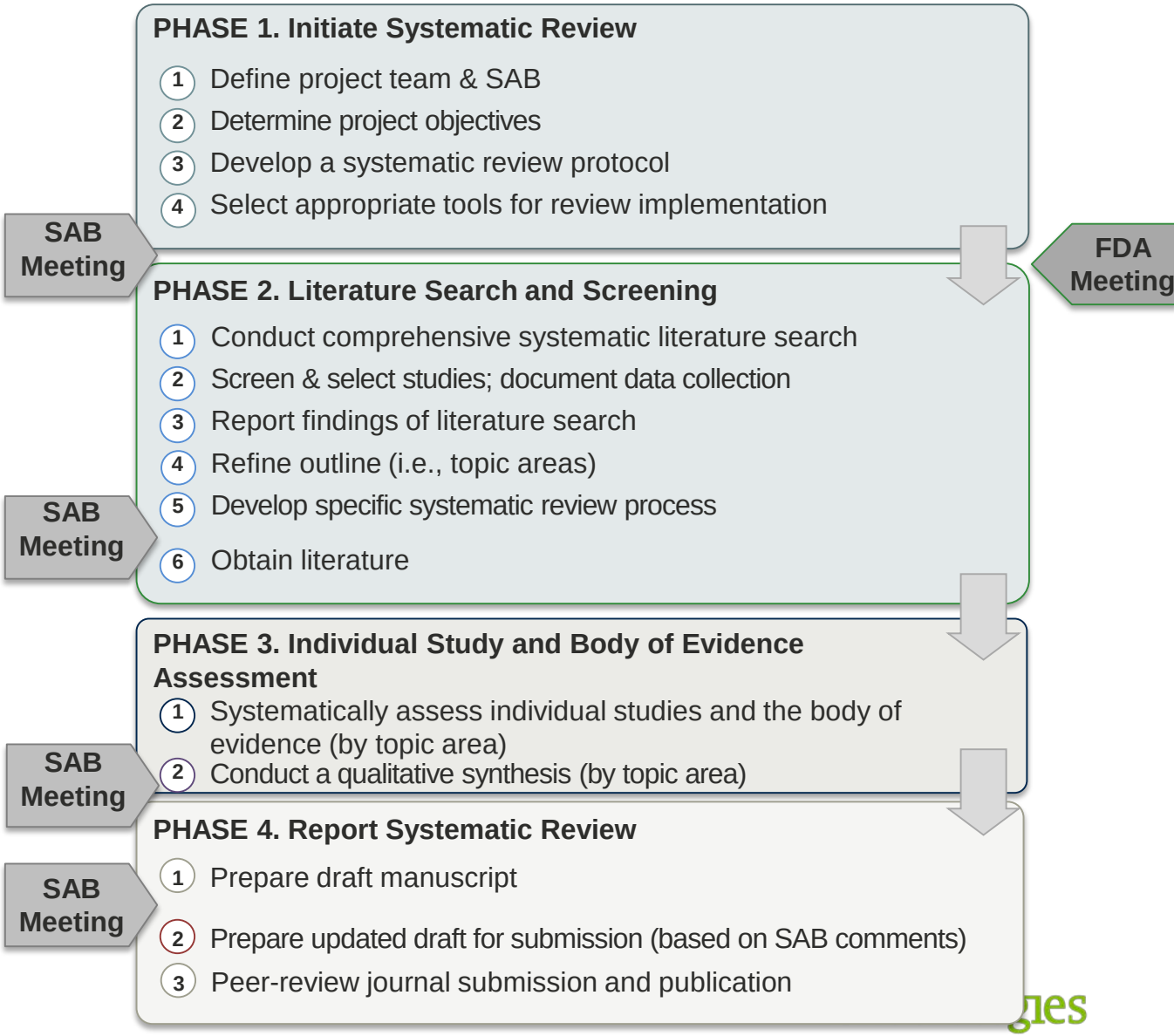


- **PECO – based on updating Nawrot et al 2003:**
  - **For [population], is caffeine intake above [dose], compared to intakes [dose] or less, associated with adverse effects on [endpoint]?**

ToxStrategies

Project overview

- 5 SRs conducted using IOM framework
  - OHAT Handbook and RoB Tool
- Team: SAB + ToxStrategies
- Sponsor (ILSI-North America Caffeine Working Group)
- FDA awareness
- ~2 years in duration (scoping to publication)



Lessons learned in problem formulation and protocol development

Develop SR team and roles

Establishing the Review Team (per IOM):

"The review team is composed of individuals who will manage and conduct the review. The objective of organizing the review team is to pull together a group of researchers as well as key users and stakeholders who have the necessary skills…."

- Multidisciplinary (SR methodology, information specialists, technical experts) and balanced

Developing roles

- Type (e.g., team leader, evidence analysts, technical experts, SAB, etc.)
- Application/tasks (e.g., protocol development, screening, data extraction, individual study assessment)

Sponsor Input and Role

Provided at each phase of the project, per IOM:

- "While an SR should respond to the sponsor's questions, the sponsor should not overly influence the SR process. The relationship between the sponsor and the SR review team needs to be carefully managed to balance the competing goals of maintaining the scientific independence of the SR team and the need for oversight to ensure the quality and timeliness of their work."
- "Sponsors should not be allowed to delay or prevent publication of an SR in a peer-reviewed journal and should not interfere with the journal's peer review process"

ILSI North America Caffeine Working Group is the Sponsor

This work was supported by the North America Branch of the International Life Sciences Institute (ILSI N.A) and through unrestricted grants to ILSI N.A. by the American Beverage Association (ABA) and the National Coffee Association (NCA) ILSI NA is a public, non-profit foundation that provides a forum to advance understanding of scientific issues related to the nutritional quality and safety of the food supply by sponsoring research programs, educational seminars, and workshops, and publications. ILSI NA receives support primarily from its industry membership. ABA is the national trade association that represents the U.S. non-alcoholic beverage industry. NCA is the national trade association that represents the U.S. coffee industry. The opinions expressed herein are those of the authors and do not necessarily represent the views of the funding organization.

ILSI NA mission: scientific integrity

ToxStrategies

Case Study: Caffeine - Roles

Entity	Description	Roles
ToxStrategies	Scientists with a range of expertise (caffeine, toxicology, epidemiology, systematic review, literature searching, etc.)	A, D, F, I, P *Develop and perform the SR (consistency in application of SR process, independent assessment, documentation)*
SAB	Multidisciplinary experts (systematic review, behavior, cardiovascular, bone & calcium, acute, pharmacokinetics – PhD's and MDs from academic, private, and clinical practices)	I, A *Provide input, review, and approval*
Sponsor	Members of the ILSI-North America Caffeine working group	I, A* **Budgetary*

A = approve; D = develop; F = facilitate; I = input; P = perform; X = not involved

Lessons Learned

- *Diverse team strengthens the SR (but also requires flexibility and influences timeline)*
- *Refined and clear roles = smooth(er) implementation*
- *Establishes independence of the scientific assessment*

ToxStrategies

PECO – Lessons learned

Word choice can be difficult – every word counts

- PECO "guides" all decisions – when in doubt, revisit the PECO

Important to have contextual key questions

Determination of issues and rationale is best informed by a multidisciplinary team

PECO structure varies depending on application and objectives – hazard, risk, candidate datasets/values, etc.

Case study: PECO for multiple endpoints

- **For [population], is caffeine intake above [exposure], compared to intakes [comparator] or less, associated with adverse effects on [outcome]?**

Population	Healthy Adults	Healthy Pregnant Women	Healthy Adolescents	Healthy Children
Exposure	> # mg/day	> # mg/day	> # mg/kg-day	> # mg/kg-day
Comparator	≤ # mg/day	≤ # mg/day	≤ # mg/kg-day	≤ # mg/kg-day
Outcome	Adverse effect (five outcomes)			

ToxStrategies

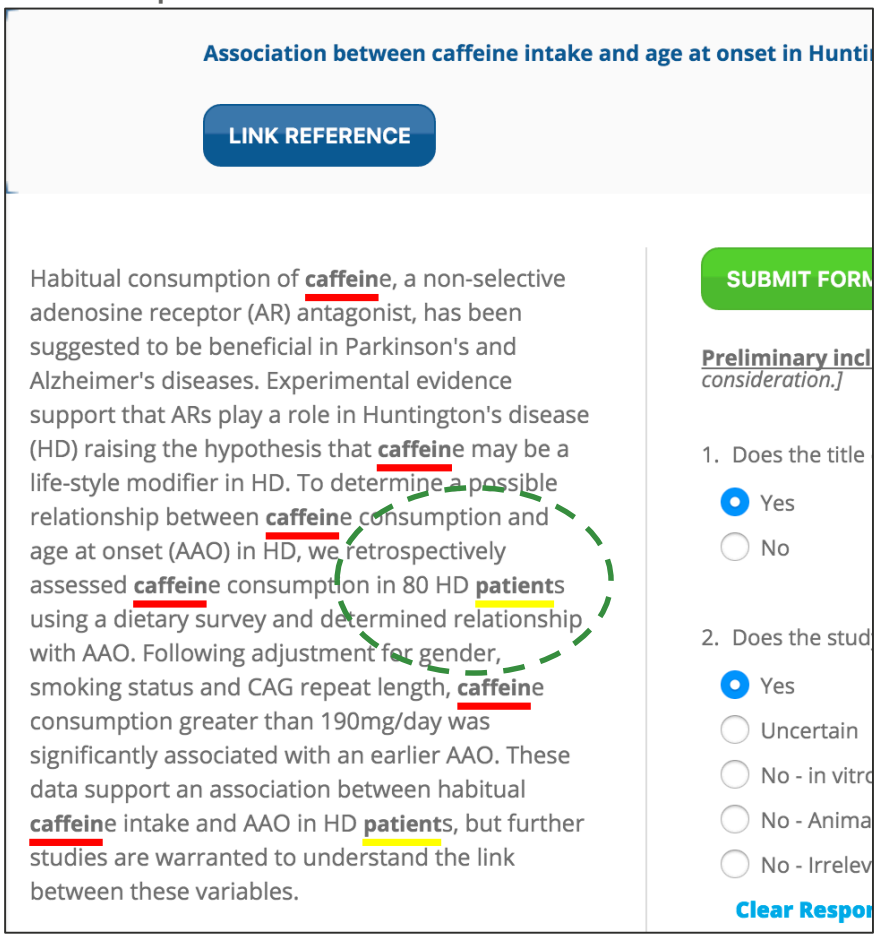
Population: “healthy”

Issue: Determining inclusion/exclusion based on healthy populations

- Compromised physical function; caffeine as a therapeutic
- Physicians and SR expert played a key role in helping with this determination

1. Define healthy
   - Subjects who were not specifically described as hospitalized, diagnosed with disease, and/or receiving medical treatment for a disease at the time of the study
2. Impact on SR
   - Many abstracts/papers excluded based on definition (retained if included healthy control arm or similar)
     - E.g., diabetes, Parkinson’s

Example:

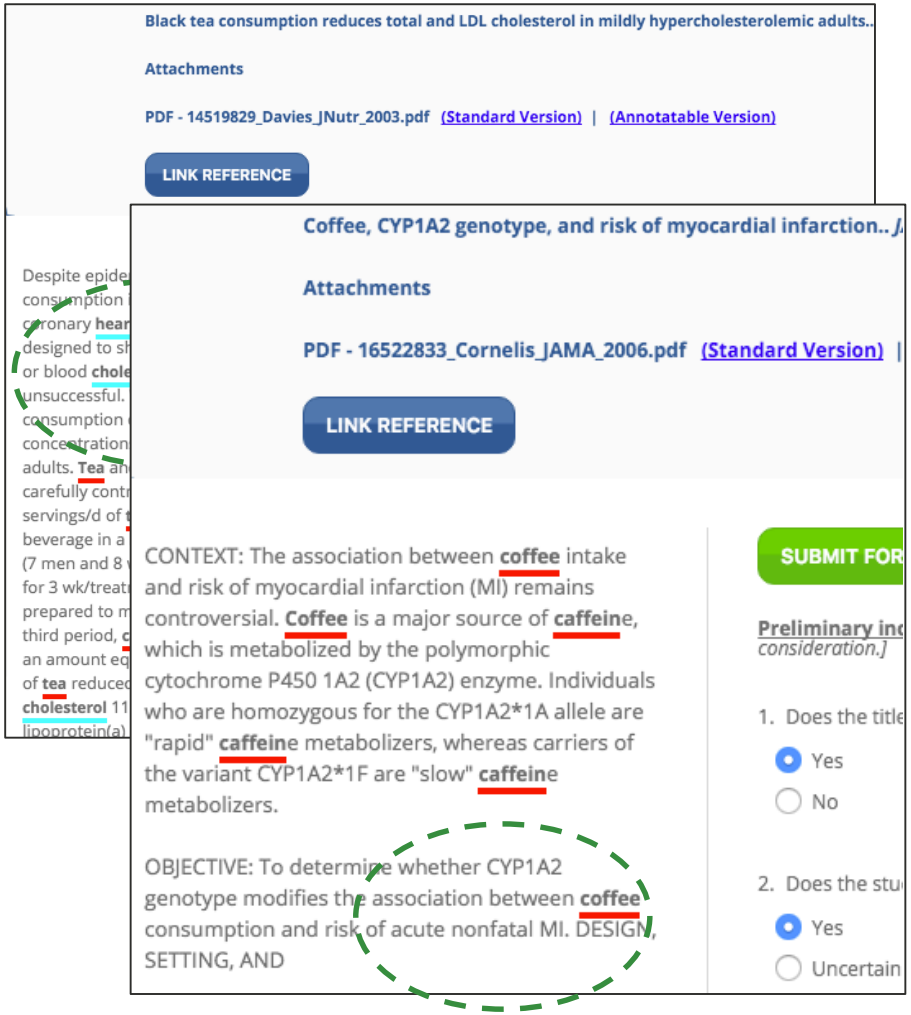


ToxStrategies

Outcome: “adverse”

Issue: Nawrot et al. 2003 update to adverse effects

1. Define adverse
   - Use effects in Nawrot?
   - Subjective, difficult to draw a line
   - Decision: be conservative and comprehensive
2. Impact on SRs
   - Pilot searches utilized “adverse” term (and related) – too restrictive
   - Extracted information based on author conclusion as well as analyst conclusion
   - Characterization of effects in subgroups (e.g., physiological/clinical, order in progression of effect, etc.)



ToxStrategies

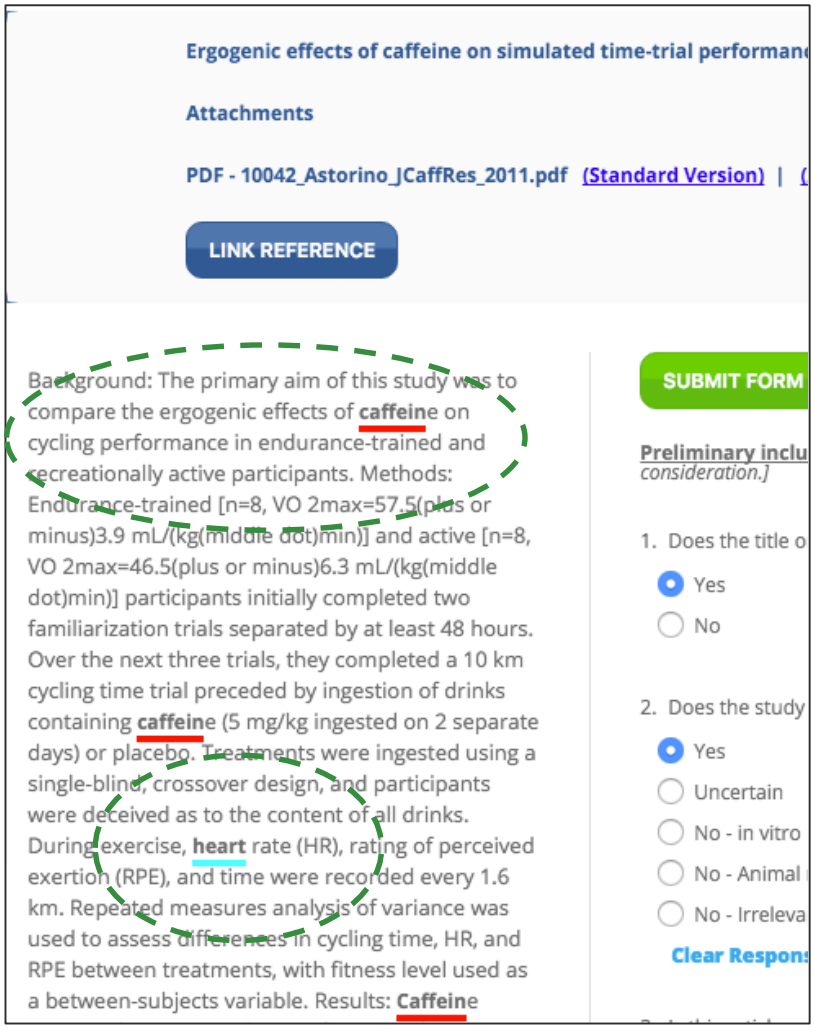
Outcome: "adverse" – an added layer of complexity

Issue: Caffeine exposure has been associated with benefits and adverse effects

- Adverse effects in studies that are evaluating benefits? (e.g., RCTs)
  - Controlled exposure, etc.
- OR lack of adverse effects in studies evaluating benefits?
  - Null findings on potentially adverse outcomes (e.g., heart rate)

Decision

- Include: studies reporting data associated with adverse effects within a benefit/therapy study.
- Exclude: Studies assessing only beneficial or therapeutic endpoints or outcomes following exposure to caffeine.



ToxStrategies

Outcome: multiple endpoints, multiple outcomes

Issue: Outcome of each review associated with many endpoints

Decision: Comprehensive and inclusive (consistent w/ characterization of hazards)

Impact: Frameworks and implementation activities need to be accommodating

Cardiovascular: Endpoint Keywords Identified From Abstracts (281)		
acute stiffness (pulse wave velocity)	chest pain and palpitation	stroke volume and cardiac contractility
arrhythmia	cholesterol	supra ventricular dysrhythmias
atrial fibrillation	endothelial function/performance	ventricular function
blood pressure	heart rate	
cardiovascular disease	hypertension	
cerebral blood flow	myocardial blood flow	

Protocol publication

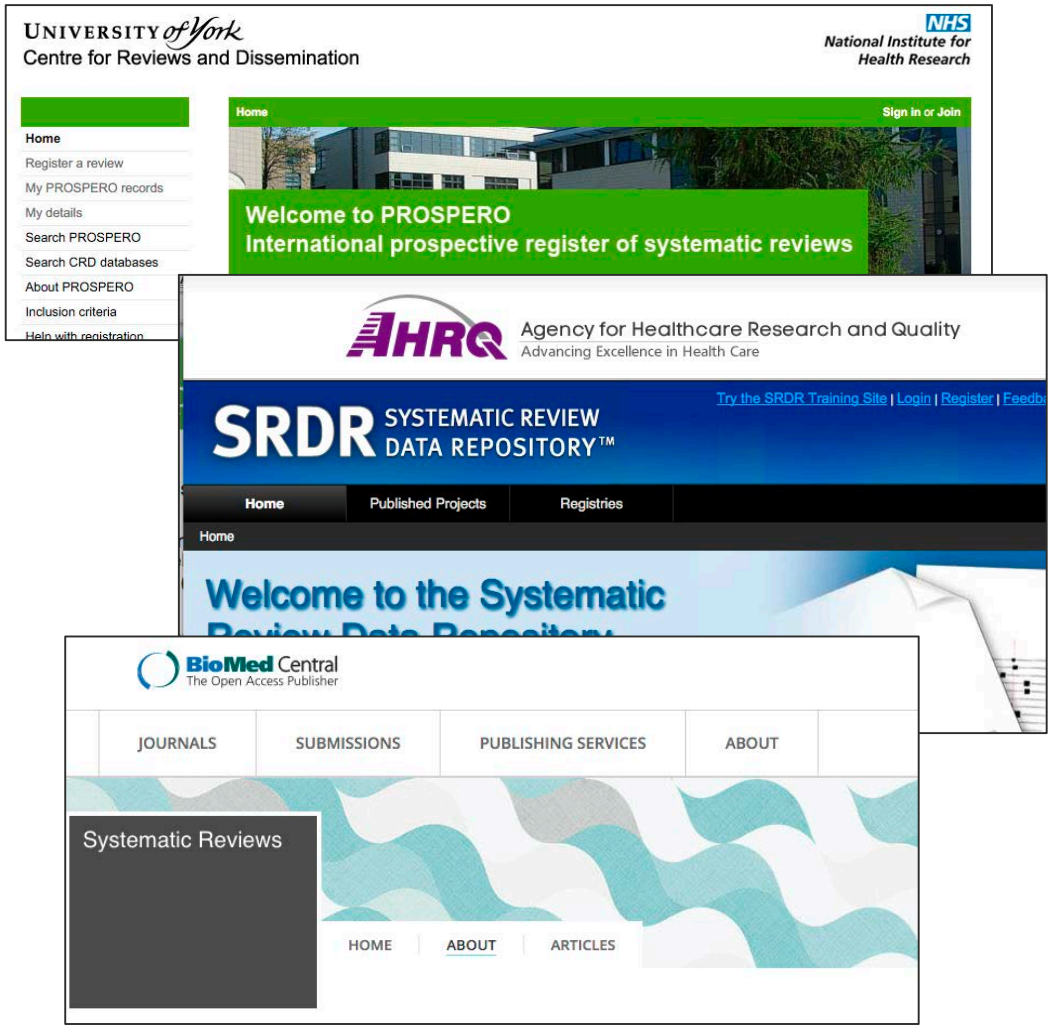
Issue: publication is the “gold standard”

- Considerations
  - Level of transparency (proprietary investigations)
  - Platform for publication

ToxStrategies minimum standard: internal comprehensive documentation (including SOPs)

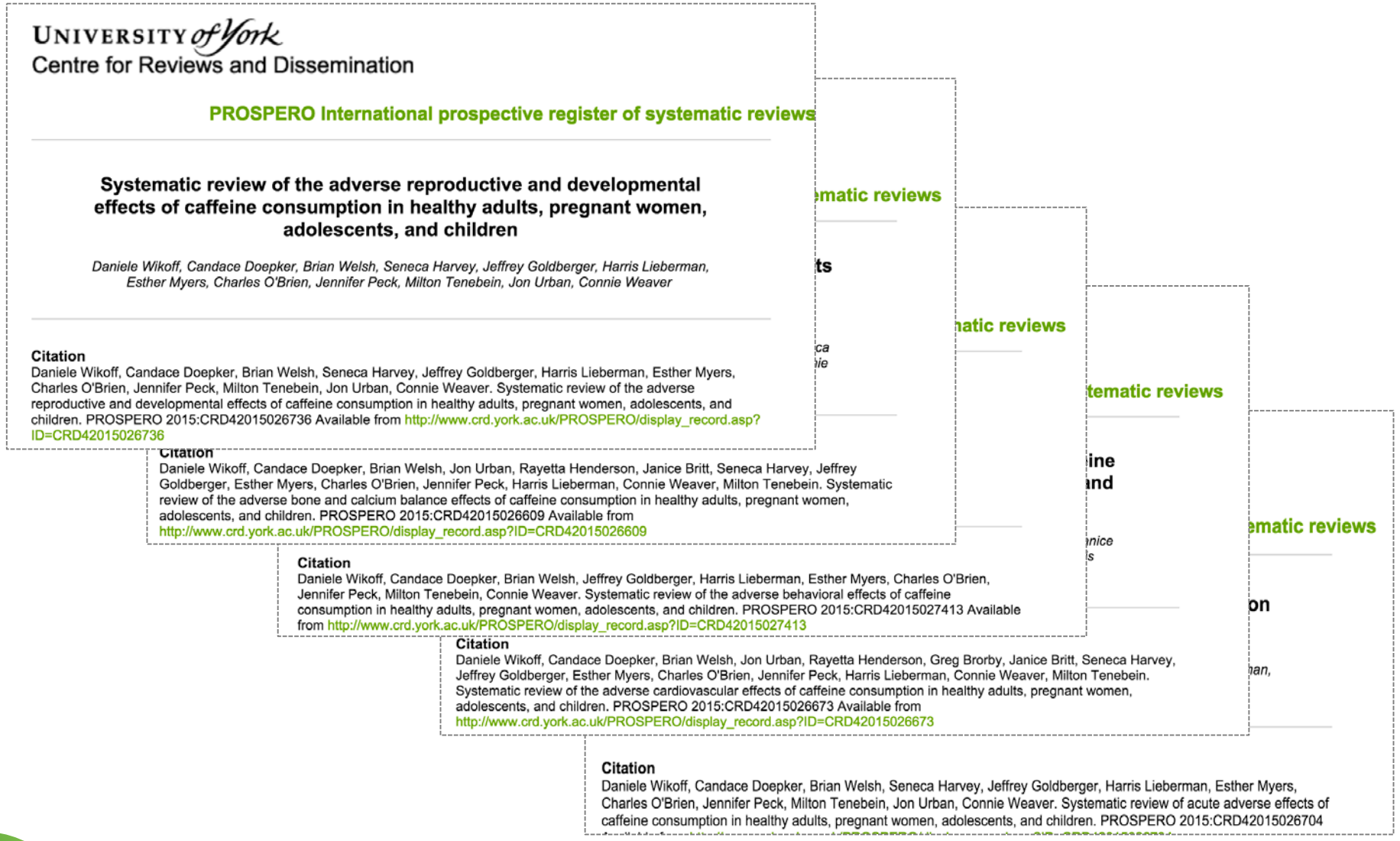
Decisions implemented:

- Publication on PROSPERO
- Methods/supplementary to publication
- Appendix to reports



ToxStrategies

Case study: Protocol publication in PROSPERO



http://www.crd.york.ac.uk/PROSPERO/

ToxStrategies

Lessons learned in protocol implementation

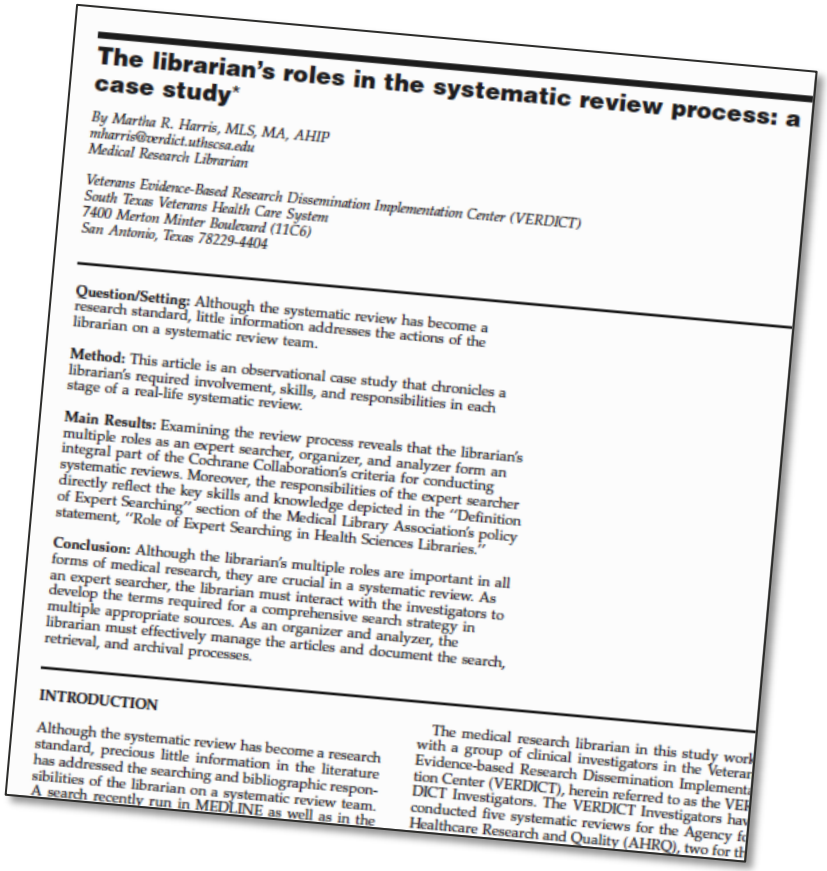
Role of librarian/information specialist

Per IOM, review team includes a librarian/information specialists trained in searching bibliographic databases

In our experience, we have utilized librarian/information specialists assists with:

- Conducting preliminary searches
- Development of overall search strategy – including selection of databases
- Development of search strings (syntax specific to individual databases)
- Assist with refinements and rationale to protocol (particularly as part of pilot)

Practical application note: not all of us have "easy" access to such specialists



The librarian's roles in the systematic review process: a case study*

By Martha R. Harris, MLS, MA, AHIP
mharris@verdict.uthscsa.edu
Medical Research Librarian

Veterans Evidence-Based Research Dissemination Implementation Center (VERDICT)
South Texas Veterans Health Care System
7400 Merton Minter Boulevard (11C6)
San Antonio, Texas 78229-4404

Question/Setting: Although the systematic review has become a research standard, little information addresses the actions of the librarian on a systematic review team.

Method: This article is an observational case study that chronicles a librarian's required involvement, skills, and responsibilities in each stage of a real-life systematic review.

Main Results: Examining the review process reveals that the librarian's multiple roles as an expert searcher, organizer, and analyzer form an integral part of the Cochrane Collaboration's criteria for conducting systematic reviews. Moreover, the responsibilities of the expert searcher directly reflect the key skills and knowledge depicted in the "Definition of Expert Searching" section of the Medical Library Association's policy statement, "Role of Expert Searching in Health Sciences Libraries."

Conclusion: Although the librarian's multiple roles are important in all forms of medical research, they are crucial in a systematic review. As an expert searcher, the librarian must interact with the investigators to develop the terms required for a comprehensive search strategy in multiple appropriate sources. As an organizer and analyzer, the librarian must effectively manage the articles and document the search, retrieval, and archival processes.

INTRODUCTION

Although the systematic review has become a research standard, precious little information in the literature has addressed the searching and bibliographic responsibilities of the librarian on a systematic review team. A search recently run in MEDLINE as well as in the

The medical research librarian in this study work with a group of clinical investigators in the Veteran Evidence-based Research Dissemination Implement tion Center (VERDICT), herein referred to as the VE DICT Investigators. The VERDICT Investigators hav conducted five systematic reviews for the Agency fo Healthcare Research and Quality (AHRQ), two for th

J Med Libr Assoc 93(1) January 2005

ToxStrategies

Databases - everyone uses PubMed, right?

Starting point for most health-related searches (though terms/syntax are rarely reported)

Is one database sufficient?

How choose (other) databases? Depends on many things…

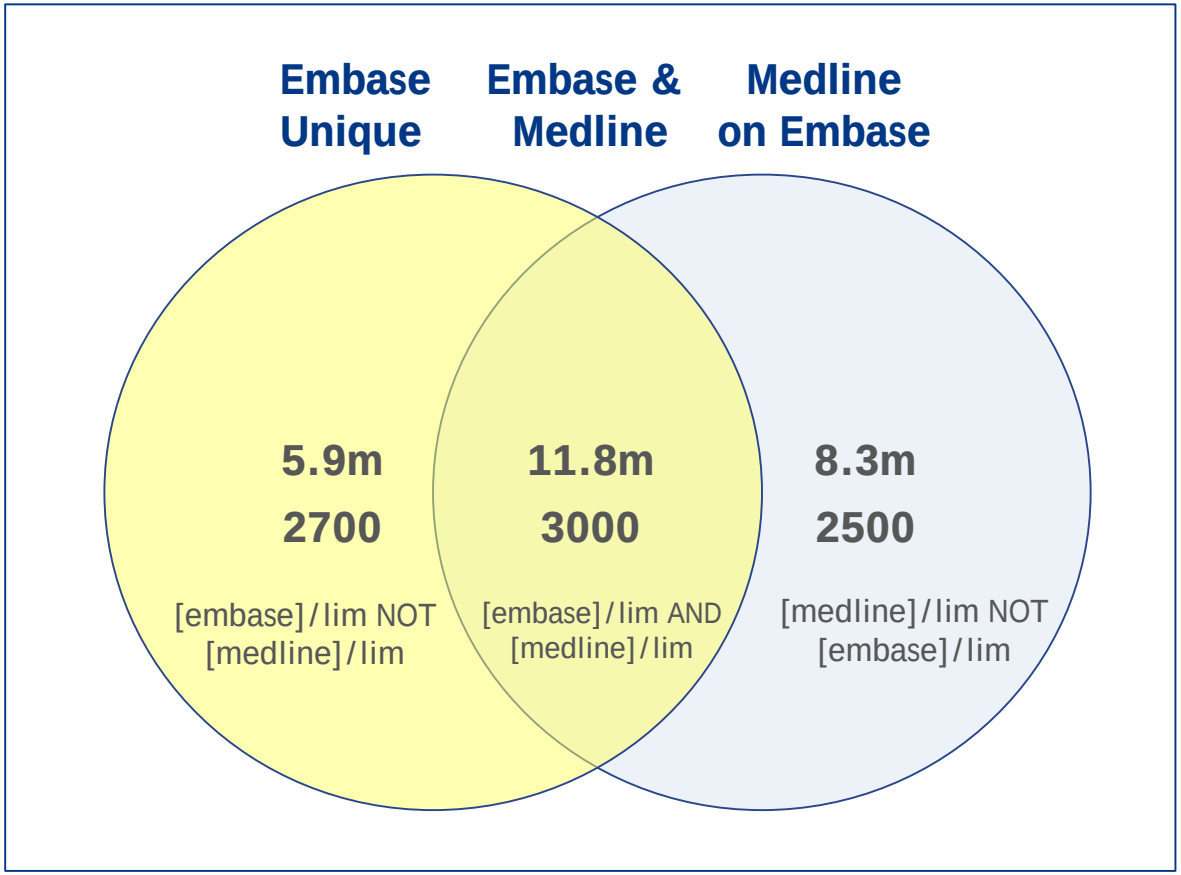
- Objectives (e.g, turn over every stone, only publicly available, etc.)
- PECO (e.g., some databases specific/unique outcomes [DART, ReproEXPERT], exposures/chemicals [SciFinder], document types [ExPub, CEBS], etc)
- Practical considerations (e.g., facilitation with software, resources/access, overlap in aggregation)

Per the IOM:

The appropriate sources of information for an SR depend on the research question, analytic framework, outcomes of interest, study population, etc.

Embase vs. Medline (content)

- >30 million records (>8,500 journals, grey literature)
  - Over 2700 journals not indexed on MEDLINE
- Emtree Life Science thesaurus, which has over twice as many terms as the MEDLINE thesaurus (MeSH)
- Lesson: not an “either/or”
  - Includes most of Medline
  - Lag for indexing Medline



ToxStrategies

Developing database-specific syntax (some examples)

Operators	Syntax example	Impact
AND Both words or phrases in the record (e.g. title, abstract, keywords, etc)	'caffeine' AND 'addiction'	Tends to focus or narrow a search.
OR At least one word or phrase in the record (e.g. title, abstract, keywords, etc)	'caffeine' OR 'coffee'	Tends to broaden or make a search more inclusive. Useful when terms are related.
NOT Word or phrase after NOT must be excluded	'caffeine' NOT 'adverse'	Tends to focus a search. May be useful for exclusion terms.
Wildcards	'caffeine' AND cardio*	Broad search on the root, though can be too broad (example gave *only* 43 results and an error message)
MeSH	'caffeine' AND cardiovascular[MeSH]	Established indexed groupings, focused results (example gave 3080 results)

Wildcard search for 'cardio*' used only the first 600 variations. Lengthen the root word to search for all endings.

Are these really that important?

Of 105 Medline search strategies examined, 63 were assessed; 31 were excluded because they were inadequately reported…..

Most (90.5%) of the assessed search strategies contained > or =1 errors...

The most common search errors were missed MeSH terms (44.4%), unwarranted explosion of MeSH terms, and irrelevant MeSH or free text terms (28.6%)

Missing spelling variants, combining MeSH and free text terms in the same line, and failure to tailor the search strategy for other databases occurred with equal frequency (20.6%)

Logical operator error occurred in 19.0% of searches

J Clin Epidemiol. 2006 Oct;59(10

Errors in search strat

Sampson M[1], McGowan J.

Author information

Abstract

OBJECTIVE: Errors in th
studied the frequency

STUDY DESIGN A
systematic revie
more section
assessed
the eligi ctronic search

RESULTS: Of 105 MEDLIN
were duplicates of assessed search strategies. Most (90.5%) of the assessed search strategies contained > or =1 errors (median 2, interquartile range [IQR] 1.0-3.0). Errors that could potentially lower recall of relevant studies were found in 82.5% (median 1, IQR 1.0-2.0) and inconsequential errors (to the evidence base) were found in 60.3% (median 1, IQR 0.0-1.0) of the search strategies. The most common search errors were missed MeSH terms (44.4%), unwarranted explosion of MeSH terms (38.1%), and irrelevant MeSH or free text terms (28.6%). Missed spelling variants, combining MeSH and free text terms in the same line, and failure to tailor the search strategy for other databases occurred with equal frequency (20.6%). Logical operator error occurred in 19.0% of searches.

CONCLUSION: When the MEDLINE search strategy used in a systematic review is reported in enough detail to allow assessment, errors are commonly revealed. Additional peer review steps are needed to ensure search quality and freedom from errors.

ToxStrategies

Tools to facilitate the SR

http://systematicreviewtools.com/index.php

Many tools available

- Selection dependent on many variables
  - objectives, team members, and platforms for communication and reporting, etc.
- Vary in capabilities and transparency
  - Protocols, literature screening and selection, library/source management, data extraction, text mining, meta-analysis, etc.
- Vary substantially in price

DRAGON
An online tool for systematic review

SRDR SYSTEMATIC REVIEW DATA REPOSITORY™

DistillerSR

MENDELEY

ENDNOTE WEB®

covidence

ToxStrategies

Case example: caffeine – librarian role

1. Selected three databases to be included in the search
   - Considerations for the voluminous body of literature, endpoints of interest, etc.
2. Crafted search syntax for each database
   - Involved *many* iterations
3. Result: outcome-specific search strings

Lessons Learned

- *Librarians/information specialists significantly enhance the literature search strategy – but it is critical that the (technical) project team be highly engaged in decisions regarding the search strategy*

ToxStrategies

Case example: caffeine – searching multiple databases for multiple outcomes

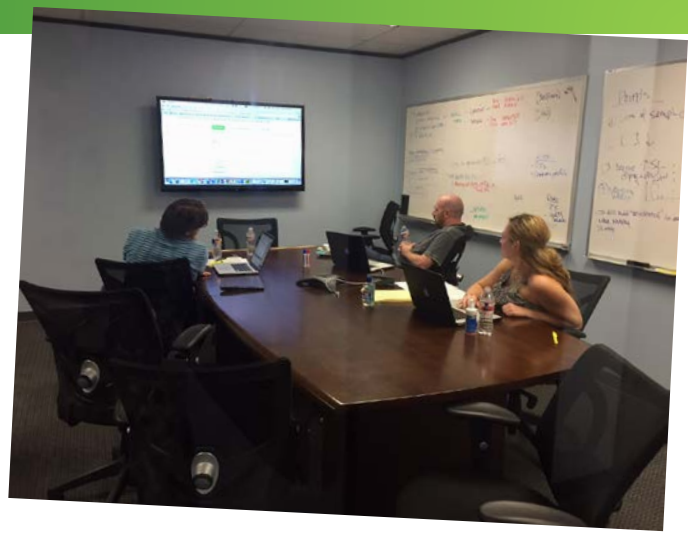
Issue: Independent searches resulted in large number of hits across databases

Solution: Run concatenated outcome search, upload all findings in DistillerSR, and categorize during screening (more efficient)

- "Automatic" de-duplication
- Identify articles that address multiple outcomes

Combination	# Hits
Individual	
Acute Toxicity	2337
Behavior	3942
Bone	622
Cardio	1058
PK	1254
Reproductive	875
Sum of individual searches	**10088**
Concatenated search string (all outcomes)	**4942**

Case example: caffeine-screening pilot



Extensive multi-day pilot conducted with all screeners present

- Consistent with IOM framework

Initiated prior to final search; tested using preliminary and final search strings

Developed, tested, and modified DistillerSR screening forms

Developing and refined internal “SOP”

- Minor additions/clarifications to protocol

Result: homogeneity in responses across reviewers

- Builds confidence

Lessons Learned

1. *Pilot is critical for consistency; plan for refinements to the protocol and consider “SOPs”*
2. *Documentation can be cumbersome – but it’s critical*
3. *SR software (with audit log) ensures transparency and completeness*

Challenges in individual study assessment

Design and execution of data extraction

- Generic templates (Distiller SR forms) designed to cover all outcomes
- Include author information and conclusions as well as topic/project specific fields that involve interpretation in order to apply to the PECO (and to easily apply in the body of evidence synthesis)

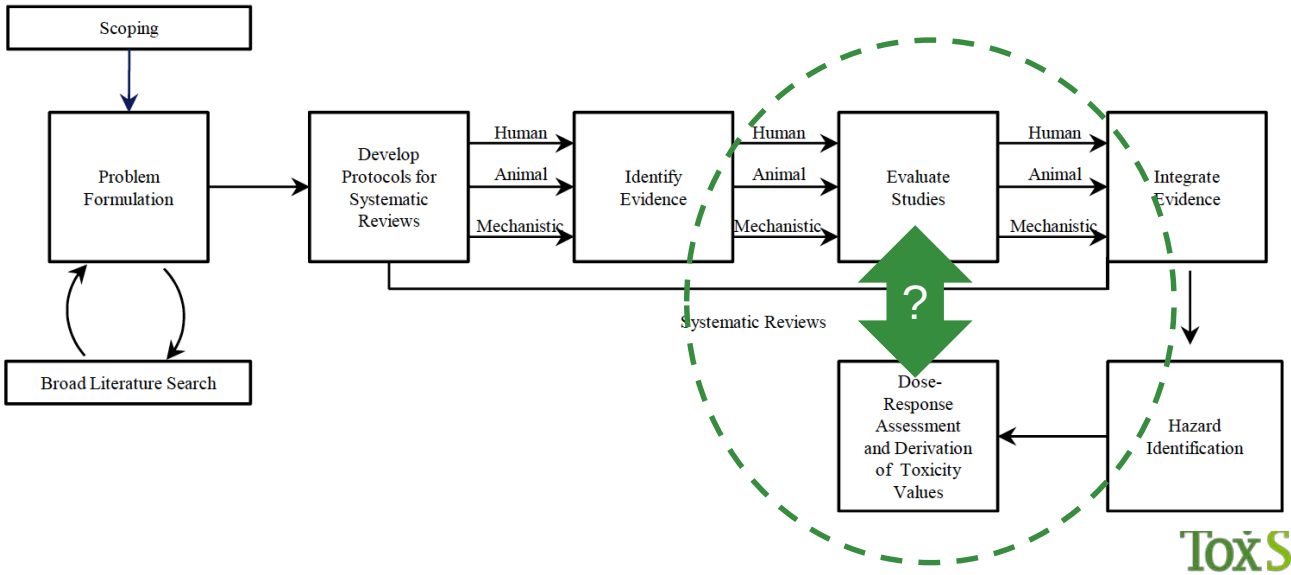
Selection of “grading” criteria

- Applicable to all outcomes and streams (requires flexibility) and/or aspects of PECO
  - Going beyond RoB?
- Most appropriate to assess overall project objectives (e.g., hazard, risk, causation) as well as data types (including contextual data)
  - May require multiple systems/framework or tiered approach

Challenges in individual study assessment (cont.)

Application of “grading” criteria

- Interpretation of criteria (not always directly applicable, or well described etc.)
  - Tailor guidance/interpretation to PECO (less so in “broad” applications”)
  - Requires extensive pilot and generation of internal SOPs
- Use of results in subsequent analyses
  - Inclusion/exclusion, subgroup analyses, etc.
  - Identification of candidate datasets for development of toxicity values
- Available frameworks still seem to require expert judgment when applying to a body of evidence/conclusions (as well as identifying candidate datasets)



Practical Lessons Learned

Overall “practical” lessons

1. Every SR is different; each SR has many, many moving parts
2. Problem formulation and protocol development are not only fundamental, but are vital to the SR (and can be difficult)
3. SR approaches vary within EBT – particularly for hazard, risk, causation, etc.
   - Broad PECO (and multi-endpoint assessments) (may) require a (slightly) different application of SR methodologies
4. Selection of appropriate software tools greatly enhances efficiency (and compliance with SR methodologies)
   - Documentation can be cumbersome, but it’s critical
5. The field of EBT is evolving
   - We are gaining resources, but we are still learning how to consistently apply such
6. Cannot ignore the role of resources and time

Overall “practical” lessons (cont.) - Time

Phase	Time	Notes
Problem formulation and protocol development	Variable *(not measured in minutes or hours)*	Includes preliminary searching, collaborations, publication etc.
Literature searching and database generation	Variable	Number of databases, software compliance, de-duplication, hand-searching
Pilot screening	1-5+ minutes/hit	Discussion, documentation, protocol refinements, form development/refinement; # of outcomes
Title and abstract Review*	~1-2 minutes + conflict/group review	Specificity of SOP + inclusion/exclusion; Level of documentation
Pilot study assessment	>2 hours/study	Discussion, documentation, protocol refinements, form development/refinement
Individual study assessment*	1.5-3 hours (with outliers)	(obtain/manage papers), data extraction, grading; # of endpoints, complexity

**Discussed in OHAT 2015*

Questions and Discussion

Thank you!