Making the most of *existing evidence* in the OHDSI *evidence generation* environment

An update from the Knowledge Base (Laertes) workgroup

Richard D Boyce, Martijn Schuemie, Anthony Sena

April 12th, 2016

OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

University of Pittsburgh DBMI Department of Biomedical Informatics

Outline for today's update

1. Motivation for the work and the current status of the LAERTES evidence base
2. Use of the evidence base to generate "negative controls"
3. Demo of the negative control workflow within Atlas
4. Discussion

Some of you might remember…

OMOP's Vision of Risk Identification - 2013

http://omop.org/node/649

The 2013 OMOP symposium presented a vision…

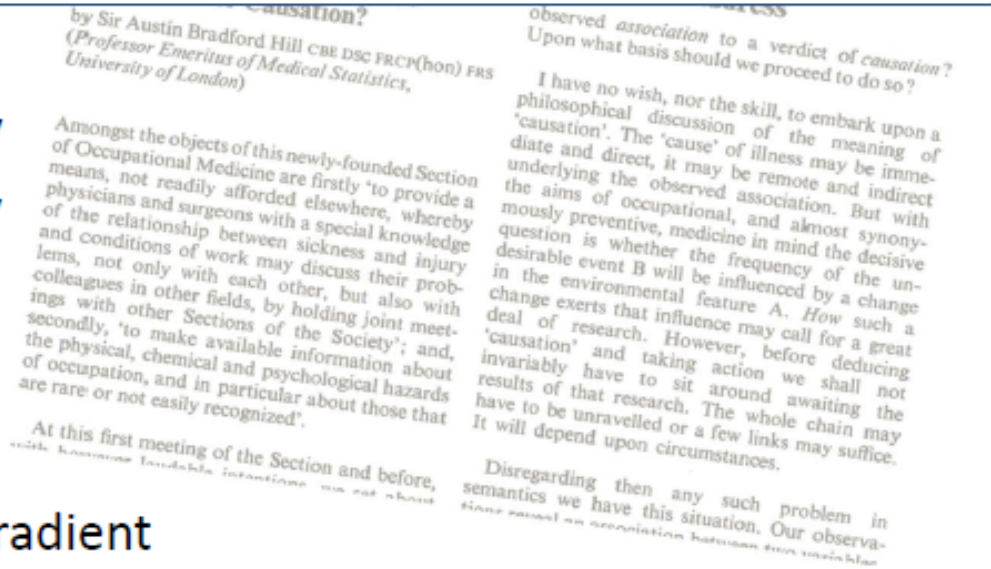
- ***Large-scale evidence generation***
  - Large patient-level datasets offer unprecedented opportunities for evidence generation
  - We no longer need to constrain ourselves to custom- selecting one piece of evidence at a time
  - What we do need is a standardized, systematic approach to interrogate, evaluate, and synthesize **all the diverse evidence** that is at your disposal to guide your decision-making

OBSERVATIONAL
MEDICAL
OUTCOMES
PARTNERSHIP

To go forward, we must go back

"What aspects of that association should we especially consider before deciding that the most likely interpretation of it is causation?"

- Strength
- Consistency
- Temporality
- Plausibility
- Experiment
- Coherence
- Biological gradient
- Specificity
- Analogy



by Sir Austin Bradford Hill CBE DSC FRCP(hon) FRS
(Professor Emeritus of Medical Statistics,
University of London)

Amongst the objects of this newly-founded Section of Occupational Medicine are firstly 'to provide a means, not readily afforded elsewhere, whereby physicians and surgeons with a special knowledge of the relationship between sickness and injury and conditions of work may discuss their problems, not only with each other, but also with colleagues in other fields, by holding joint meetings with other Sections of the Society'; and, secondly, 'to make available information about the physical, chemical and psychological hazards of occupation, and in particular about those that are rare or not easily recognized'.

At this first meeting of the Section and before,

observed *association* to a verdict of *causation*?
Upon what basis should we proceed to do so?

I have no wish, nor the skill, to embark upon a philosophical discussion of the meaning of 'causation'. The 'cause' of illness may be immediate and direct, it may be remote and indirect underlying the observed association. But with the aims of occupational, and almost synonymously preventive, medicine in mind the decisive question is whether the frequency of the undesirable event B will be influenced by a change in the environmental feature A. *How* such a change exerts that influence may call for a great deal of research. However, before deducing 'causation' and taking action we shall not invariably have to sit around awaiting the results of that research. The whole chain may have to be unravelled or a few links may suffice. It will depend upon circumstances.

Disregarding then any such problem in semantics we have this situation. Our observa-

Austin Bradford Hill, "The Environment and Disease: Association or Causation?," *Proceedings of the Royal Society of Medicine*. 58 (1965). 295-300.

OBSERVATIONAL
MEDICAL
OUTCOMES
PARTNERSHIP

Coherence

Sir Bradford-Hill 1965:
"The cause-and-effect interpretation of our data should not seriously conflict with the generally known facts of the natural history and biology of the disease."

OMOP 2013:
Observational data is only one piece of the puzzle. We need a interactive framework to explore observational analysis results alongside other evidence from published literature, product labeling, spontaneous adverse event reporting, and biomedical ontologies.

"Like a photo mosaic, a clear and understandable image of a potential drug safety issue can emerge when the relevant sources of evidence are brought together"

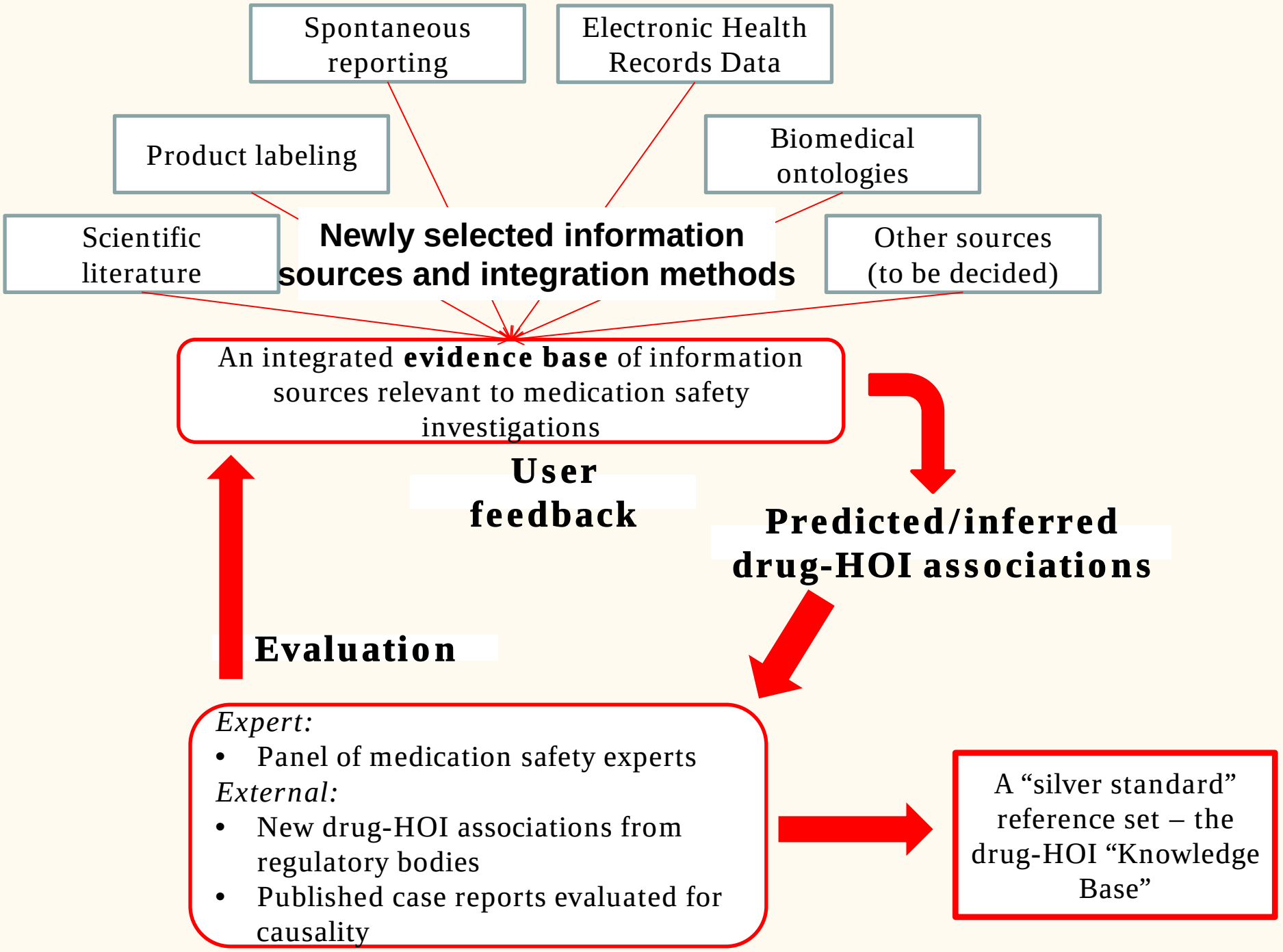
Boyce RD, Ryan PB, Norén GN, Schuemie MJ, Reich C, Duke J, Tatonetti NP, Trifirò G, Harpaz R, Overhage JM, Hartzema AG, Khayter M, Voss EA, Lambert CG, Huser V, Dumontier M. Bridging islands of information to establish an integrated knowledge base of drugs and health outcomes of interest. Drug Saf. 2014 Aug;37(8):557-67. doi: 10.1007/s40264-014-0189-0. PubMed PMID: 24985530

Knowledgebase (LAERTES) Workgroup

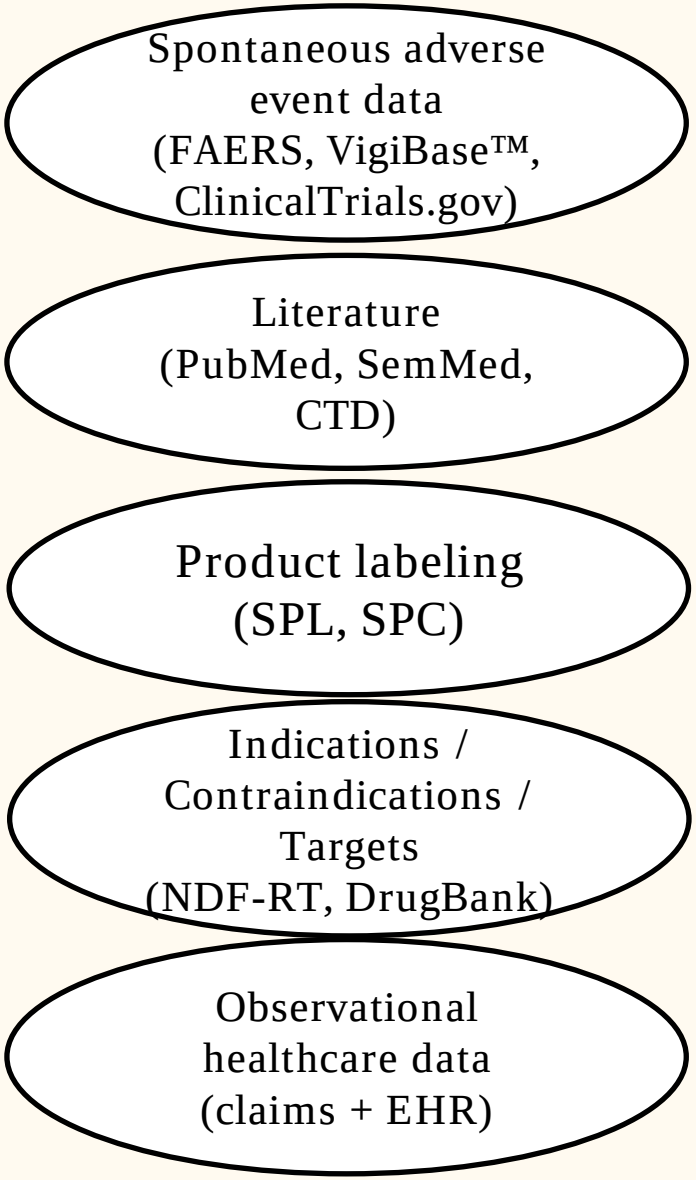
- **Objective:**
  - The objective of this workgroup (WG) is to establish an open-source standardized knowledge base for the effects of medical products and an efficient procedure for maintaining and expanding it.
- See:
  - Drug Safety paper: http://goo.gl/dswKCJ
  - Project homepage: http://goo.gl/Dz1zx9
  - Github project: https://github.com/OHDSI/KnowledgeBase
    - Known issues and bugs: https://goo.gl/grcka7
  - WebAPI documentation: http://goo.gl/2JCb92

The LAERTES evidence base

- Synthesizes adverse drug event evidence within a standard framework for clinical research
  - Standardizes drugs and HOIs
    - RxNorm and SNOMED
  - Summarized across evidence sources
  - Enables “drilling down” to examine specific evidence items

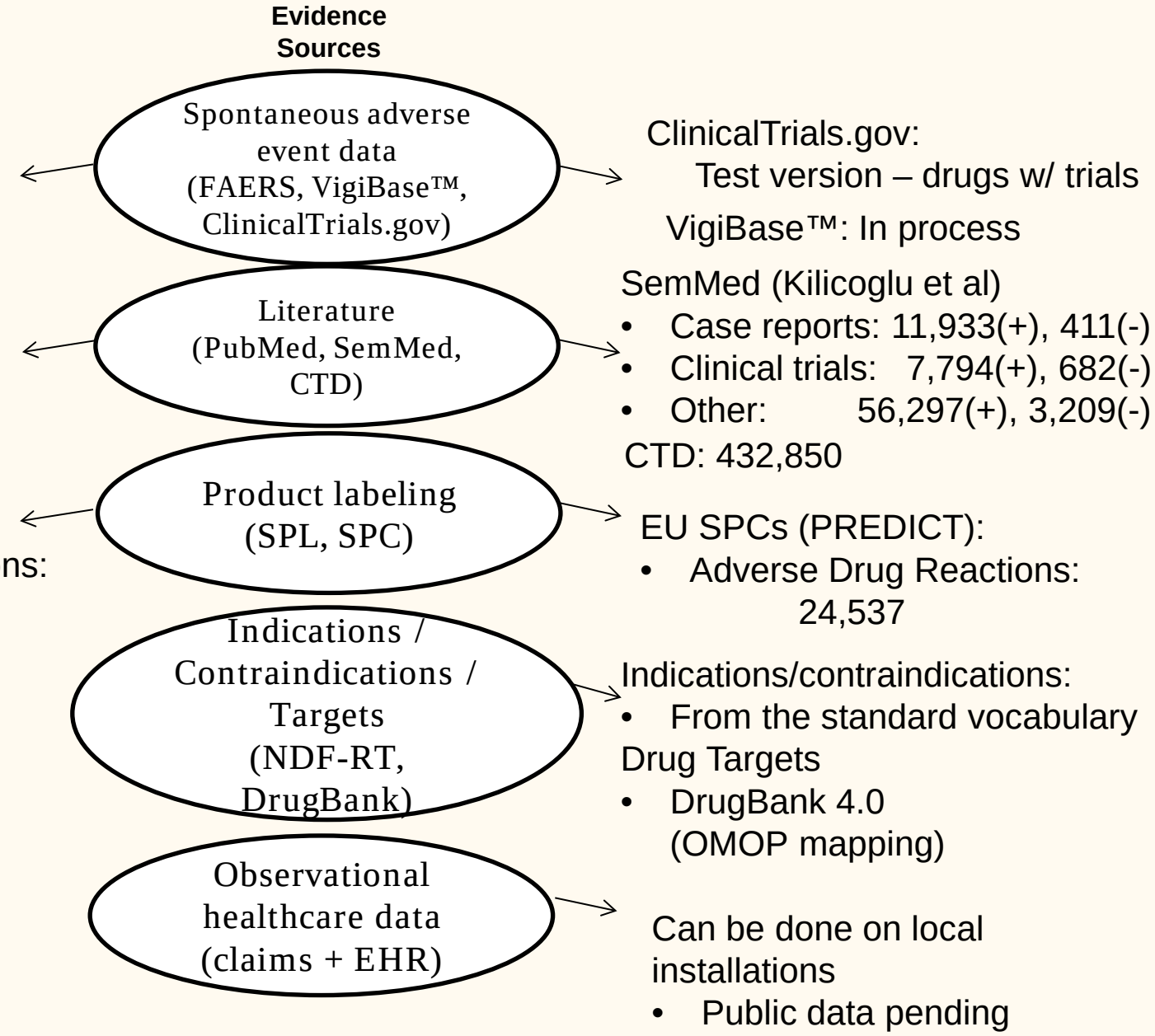

Spontaneous reporting
Electronic Health Records Data
Product labeling
Biomedical ontologies
Scientific literature
Newly selected information sources and integration methods
Other sources (to be decided)
An integrated evidence base of information sources relevant to medication safety investigations
User feedback
Predicted/inferred drug-HOI associations
Evaluation
Expert:
• Panel of medication safety experts
External:
• New drug-HOI associations from regulatory bodies
• Published case reports evaluated for causality
A "silver standard" reference set – the drug-HOI "Knowledge Base"

Current status of the evidence base



FAERS – FDA Adverse Event Reporting System; SPL – Structured Produce Labeling; SPC – Summary of Product Characteristics; FDB™ - First Databank™ EHR – Electronic Health Record; NDF-RT – National Drug File Reference Terminology, CTD – Comparative Toxicogenomics Database

Distinct drug-HOI pairs by source – 4/2016



Timeliness and number of evidence items

title	coverage_end_date	record count
PubMed	2016-03-16	307,392
CTD Chemical-Disease	2016-03-01	1,867,644
US Product Labeling	2016-01-25	985,730
SemMedDB	2015-06-30	327,483
FAERS	2014-12-31	53,671,628
EU Product labeling	2013-12-31	41,100

Basic web-based exploration of the evidence base

- Video demo:
  - Exploring LAERTES demo video
- Live site:
  - http://goo.gl/swZ6jf

Ways to access the LAERTES Evidence Base

	Database (AWS or custom)	RDF store (AWS or custom)	Linked Data (custom)	Web API	Atlas	KB Web
HOIs by Drug (RxNorm ingredient, clinical drug, branded drug)	**Yes**	**Yes**	**Yes**	**Yes**	No	**Yes**
Drugs by specific HOI SNOMED concept	**Yes**	*Part*	*Part*	**Yes**	No	**Yes**
Drugs by HOI SNOMED concept similarity	**Yes**	No	No	No	No	No
Source drug and HOI concepts (MeSH, MedDRA)	**Yes**	**Yes**	**Yes**	No	No	No
'Drill down' to proxy for the source documents	**Yes**	**Yes**	**Yes**	**Yes**	No	*Part*
Negative Controls	**Yes**	No	No	**Yes**	**Yes**	No
DrugBank	**Yes**	No	**Yes**	No	No	No
CT.gov	**Yes**	No	**Yes**	No	No	No

Roadmap

- Main goals this year
  1. A fully tested release version of the LAERTES drug-HOI evidence base
     - similar to how the OHDSI vocabulary is released
     - known issues addressed
  2. A more user friendly interface for exploring the evidence base
     - Better support for the safety investigator user scenario
  3. Develop explicit criteria for going from the drug-HOI evidence base to drug-HOI knowledge
     - Progress on the comprehensive knowledge base
  4. Funding to conduct further research and development

Acknowledgements

- Funding:
  - National Library of Medicine (1R01LM011838-01)
  - National Institute of Aging (K01AG044433-01)

How to get Involved

- Join our meetings:
  - http://www.ohdsi.org/web/wiki/doku.php?id=projects:workgroups:kb-wg

- Test and create issues:
  - https://github.com/OHDSI/KnowledgeBase