Development of a Curated Database of *In Vivo* Estrogenic Activity

P Ceger¹, N Kleinstreuer¹, X Chang¹, B Jones¹, J Hamm¹, L Rinckel¹, D Rotroff², W Casey³

¹ILS, RTP, NC, USA; ²North Carolina State University, Raleigh, NC, USA; ³NICEATM/DNTP/NIEHS/NIH/HH, RTP, NC, USA

Abstract

Currently mandated testing for potential estrogenic activity will involve thousands of chemicals, cost millions of dollars, and take decades to complete using current validated tests. High throughput screening and computational toxicology tools may streamline this process by the quick and cost-effective identification of endocrine active chemicals (EACs). Access to a comprehensive database of high-quality *in vivo* EAC toxicology data is critical for validating *in vitro* and *in silico* models of estrogenic activity and supporting the prioritization of chemicals for further testing. Accordingly, we reviewed the current scientific literature, identified high-quality *in vivo* EAC data, and compiled the data into a single database. The initial review focused on 52 reference chemicals selected by the EPA and NTP. Selected studies included data for these 52 chemicals for a number of different estrogenic endpoints (uterotrophic, pubertal, multigenerational, etc.). Data were extracted and compiled using a standardized ontology. An R script is under development to evaluate the quality of the data in an efficient and standardized manner by modified Klimisch criteria. Data that were classified as reliable were added to the database, which will be publicly available on the NTP website (http://ntp.niehs.nih.gov/go/40658).
Introduction

- U.S. (7 U.S.C. 136, 110 Stat 1613) and international regulations require the testing of chemicals for the detection of potential endocrine activity.
- As many as 10,000 chemicals may lack sufficient testing data, with several hundred new chemicals being added each year (EPA 2011).
- The U.S. Environmental Protection Agency (EPA) has developed a two-tiered strategy for identification of endocrine active chemicals (EACs).
- Tier 1 testing (Figure 1) consists of *in vitro* and *in vivo* screens. Such testing could cost millions of dollars per chemical (Martin 2012), take years to complete, and utilize many animals.
Figure 1  EPA Tier 1 Battery*

*Assays in shaded boxes are capable of detecting estrogenic activity.
• High throughput screening and computational toxicology tools are being developed to identify potential EACs and prioritize further screening efforts.

• Availability of a comprehensive, curated database of in vivo reference data will be critical for successful acceptance and implementation of these tools. Accordingly, the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) has assembled a comprehensive database of high-quality in vivo EAC data.

• Potential uses of this database include:
  - Linkage of in vivo effects to specific pathway perturbations
  - Evaluation of how biological responses are effected by exposures of different duration
  - Evaluation of species-specific responses to chemicals
  - Development and evaluation of physiologically-based pharmacokinetic models
  - Validation of in vitro and in silico models of estrogenic activity
  - Prioritization of chemicals for further testing

Scope of the Database

• EACs may affect the estrogen, androgen, and thyroid systems. Of the three systems, more in vivo and in vitro studies describe chemical effects on estrogenic signaling than on androgen and thyroid signaling. Thus, we focused on collecting data on estrogenic effects of EACs.

• Initial review focused on 52 chemicals selected by EPA and the NTP. Factors considered in chemical selection included availability of data from validated assays for estrogenic activity and evidence for in vitro interaction with the estrogen receptor, such as receptor binding or receptor transactivation. Chemicals selected included known negatives and positives with a wide potency range.

• The review included EPA guideline studies with estrogenic endpoints (e.g., uterotropin, female pubertal, and the fish short-term reproduction assays), as well as non-guideline studies with estrogenic endpoints such as altered uterine weight, alterations in the day of vaginal opening, and altered menstrual cycling.

• Figure 2 outlines the process of the literature review and database development.
Figure 2  Process Overview

Select Keywords
• Chemical identifiers: chemical name(s), CAS RN
• Assay names, endpoints: uterotrophic, pubertal, etc.

Search the Literature
• Search PubMed, Scopus, EmBase™
• Remove duplicates, assign a unique identifier to each article

Develop the Database
• Ontology developed for Rotroff 2013
• Standardized set of study classifications and vocabulary
• Mutual and consistent data entry among multiple users

Data Review
• Data reviewed by independent reviewer
• Articles reviewed and assigned a Klimisch score
• Reliable articles (Klimisch score of "1" or "2") added to the database and made available
Review of the Literature

- We obtained articles for the EAC database from the PubMed, Scopus, and EmBase™ databases. Multiple searches were conducted within each database.
  - Searches were first conducted for each substance of interest using the substance name, known synonyms, and Chemical Abstract Service Registry Number (CAS RN). Synonyms and CAS RNs were obtained from the ChemID Plus website (National Library of Medicine 2013).
  - Searches were then conducted for each guideline study, followed by searches for each endpoint of interest.

- This process was simplified in PubMed by the use of PubMatrix, a tool for multiplex literature searches (Becker 2003), which is free and available to the public on the Internet. Table 1 lists results from an example PubMatrix search.
  - Each number in the table indicates the number of articles that has been found.
  - Each number is also a hyperlink. Clicking on a number takes the user to a PubMed results page for that particular keyword combination.

<table>
<thead>
<tr>
<th></th>
<th>Uterotrophic Assay</th>
<th>Uterus</th>
<th>Uterine Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apigenin</td>
<td>2</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td>Apigenol</td>
<td>3</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td>4',5,7-Trihydroxyflavone</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C.I. Natural Yellow</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>520-36-5</td>
<td>1</td>
<td>9</td>
<td>4</td>
</tr>
</tbody>
</table>

- Results of substance searches were then cross-referenced both internally and between the different literature search engines, and duplicate articles were removed.
  - Each article was identified using the PubMed Identifier (PMID), a unique identifier developed, assigned, and maintained by PubMed. Articles that were not indexed within PubMed were assigned an arbitrary unique identifier (uID), for example NICEATM_01.
  - Articles were saved as files named with their PMID/uID, allowing a direct link between a database entry and the file in which its data is contained.
Development of the Database

- A literature review ontology was developed based on Rotroff (2013), which allowed for standardization of data entry across multiple users.
- The same information was collected for all studies, including:
  - Species and strain of test animal
  - Test animal age at first dose
  - Route of dosing
  - Number, frequency, and duration of dosing
  - Type, degree, and direction of response
  - Target tissues, receptors, or genes

Data Review

Reliability Coding with Klimisch Categories

- After data entry was complete, articles were reviewed independently for accuracy.
- Each article was reviewed for data quality using modified Klimisch criteria:
  - Klimisch (1997) described a systematic approach to the review of toxicity studies for data quality and adequacy for risk assessment.
  - Criteria included whether:
    - The study was conducted under Good Laboratory Practices
    - Key information was provided such as identity of test substances, experimental conditions, and statistical evaluations
  - While the Klimisch categories are well established, they lack detailed criteria for assigning codes. We modified the Klimisch categories to incorporate reliability codes.
- Reviewers then assigned each paper a reliability code ranging from 1 to 4 (Table 2).

Table 2 Assignment of Codes to Klimisch Reliability Categories

<table>
<thead>
<tr>
<th>Code</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Reliable without restriction</td>
</tr>
<tr>
<td>2</td>
<td>Reliable with restriction</td>
</tr>
<tr>
<td>3</td>
<td>Not reliable</td>
</tr>
<tr>
<td>4</td>
<td>Not assignable</td>
</tr>
</tbody>
</table>
Standardization and Automation of Reliability Coding

- Schenider et al. (2009) developed Toxicological Data Reliability Assessment Tool (ToxRTool).
  - ToxRTool was developed to increase transparency and provide guidance for more harmonized approaches to data quality evaluations.
  - ToxRTool is a free Internet download that consists of an Excel® data file and instructions.

- Each article evaluation with ToxRTool produces a separate spreadsheet; evaluation of many articles requires collation of data in all the spreadsheets. Using the free and open-source statistical programming language R (R Development Core Team 2008) we developed a script to automate some steps of the evaluation and data collation process.
  - The R script provides a simple graphical user interface that asks the user a series of yes/no questions.
  - Each question addresses one of the relevant modified Klimisch requirements.
  - Once all of the questions are answered, the R script applies the modified Klimisch criteria consistently to assign a reliability code.
  - The R script produces a collated output similar to the ToxRTool output.

Application of Reliability Coding to the Estrogenic Activity Database

- The R script was used to collect and collate information for the database of in vivo estrogenic activity.
  - Articles that were classified as code “1” or “2” were considered to be reliable for use and added to the database.
  - Articles that were classified as code “3” or “4” were added to a separate database, which could be used as additional or supporting information on a case-by-case basis.

- Curation, review, and addition of relevant articles to the database is ongoing.

Conclusions

- Regulatory agencies require data on endocrine activity from thousands of chemicals that have not yet been evaluated. Using current methods, this task will take decades to complete and cost millions of dollars.
- High throughput screens and computational toxicology tools are being developed to identify potential EACs and prioritize further screening.
- Developing a comprehensive database of in vivo endocrine effects is critical to the success of this holistic approach. Data from such a database can be used to validate in
vitro and in silico models, develop physiologically based pharmacokinetic models, evaluate dose- and duration-specific effects, and link in vivo effects to specific pathway perturbations.

- NICEATM has assembled a comprehensive database of high-quality in vivo EAC data that is continuously expanding as more studies are curated and evaluated.
- Studies in the database are evaluated for data quality and reliability using modified Klimisch criteria. An R script was developed to assign reliability criteria in a standardized manner.
- The database and the R script will be made available to the public via the NTP website (http://ntp.niehs.nih.gov/go/40658).

References


Acknowledgements

The Intramural Research Program of the National Institute of Environmental Health Sciences (NIEHS) supported this poster. Technical support was provided by ILS under NIEHS contracts N01-ES 35504 and HHSN27320140003C.

The views expressed above do not necessarily represent the official positions of any Federal agency. Since the poster was written as part of the official duties of the authors, it can be freely copied.