**Webisode 2 – EPPI-Centre tools for collecting and using data**

Presenter: James Thomas (EPPI-Centre, UCL)

EPPI-Centre Evidence Tools, Products, and Projects – A series of webisodes from the Evidence for Policy and Practice Information and Co-ordinating (EPPI) Centre. Hosted by AIR’s Center on Knowledge Translation for Disability and Rehabilitation Research (KTDRR).

**Slide 1: Cover slide**

EPPI-Centre Evidence Tools, Products, and Projects. A series of webisodes from the Evidence for Policy and Practice Information and Co-ordinating (EPPI) Centre.

Hosted by AIR’s Center on Knowledge Translation for Disability and Rehabilitation Research (KTDRR).

Copyright © 2018 American Institutes for Research (AIR). All rights reserved. No part of this presentation may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from AIR. Submit copyright permissions requests to the AIR Publications Copyright and Permissions Help Desk at copyright\_PS@air.org. Users may need to secure additional permissions from copyright holders whose work AIR included after obtaining permission as noted to reproduce or adapt materials for this presentation.

**Cover slide template:** dark blue background with white text and gray text. Gray bar at bottom with AIR logo on the left (gray and blue column on left; letters in blue, AIR (R) on the right; words below in blue, American Institutes for Research (R)). To the left of AIR logo, EPPI-Centre logo: A large blue script letter e to the left, with smaller black letters PPI to the right. Below PPI, in a smaller black box, is the word CENTRE in white text.

**Slide 2: EPPI-Centre tools for collecting and using data**

February 2018. James Thomas (EPPI-Centre, UCL)

**Title slide template:** Blue bar at top. On far left, Institute of Education. On the far right, UCL Logo: White image of Main Building with large white letters UCL to the right. In the center background, a photograph of London with title text superimposed over the image. White bar at the bottom: On far right, EPPI-Centre logo: A large blue script letter E to the left, with smaller black letters PPI to the right. Below PPI, in a smaller black box, is the word CENTRE in white text.

**Slide 3: Outline**

• Data throughout the lifecycle of a review

• Data at specific stages of the review

* Identifying and selecting studies
* Capturing data about studies
* Synthesising study findings

• Further ahead: what tools will be available in the future?

* Automation technologies

**Slide 4: The common stages of a systematic review**

Two stages; Map and Synthesis

**Under the Map stage:**

Form review team (involve ‘users’)

Formulate review question, conceptual framework and inclusion criteria (develop ‘protocol’)

**Included in both Map and Synthesis stage:**

Search for and identify relevant studies

Describe studies

**Under the Synthesis stage:**

Appraise included studies

Synthesise and appraise findings

Communicate and engage

**Slide 5: Collecting Data**

**Slide 6: Searching**

• Many resources for searching, e.g.

* Bibliographic databases
* Citation ‘trails’

• Important to keep track of all potentially relevant records

**Slide 7: Recording each search that is run**

Screenshot of an EPPI Reviewer webpage formanagingreferences.

In the middle of the screen are the Boolean search components. Top right of the page is a list of sources that have been searched. The page also displays duplicate references. Fields displayed on this page include:

SEARCH DATE:

SOURCE DataBase

DESCRIPTION:

FILE FORMAT:
SEARCH STRING:
and NOTES:

 **Slide 8: Standardised code sets**

Screenshot of webpage “Setup CodeSets Wizard.” This screen allows you to select single Codesets to import into your review. You can select a Codeset from the list provided and see it displayed in the centre column. The Codesets that are already in your review are displayed in the third column under Codesets in the review.

This page displays standardized codesets that help understand the context of included studies, the participants and/or interventions, and the reliability of the studies.

**Slide 9: e.g. tool to assess quality of qualitative studies**

Screenshot of a web page from the EPPI Reviewer on Codes.

The screen shot displays a window with folders. The list of items displayed help users to assess the quality of qualitative studies. For example, were steps taken to strengthen the rigor of the data collected and analysis of the data.

Quality Assessment Qualitative (QAQ) questions listed are as follows:

QAQ1. Were steps taken to strengthen rigour in the sampling?

QAQ2. Were steps taken to strengthen rigour in the data collected?

QAQ3. Were steps taken to strengthen the rigour of the analysis of data?

QAQ4. Were the findings of the study grounded in / supported by the data?

QAQ5. Please rate the findings of the study in terms of their breadth and depth.

QAQ6. Privileges YP perspectives/experiences?

QAQ7. Reliability
QAQ8. Overall how relevant is the study for this review?

QAQ9. Usefulness

**Slide 10: Coding studies**

Screenshot of the document details from the EPPI Reviewer.

The screenshot displays a window where the right half is shows the text of an article and on the left a series of drop down menus. The question, “What was the design of the evaluation” highlighted. Below are four check boxes that read: Trial; Post-test only; Pre- and post-test only; Other (specify).

A series of other questions are available for selection.

**Slide 11: Analysing Data**

**Slide 12: Mapping research activity**

Screenshot of the web page“Interactive database of DFID programmes relating to the Strategic Vision for Girls and Women Empowerment and Accountability”

Three ways of displaying data/information are pictured on this slide.

* A multi-colored pie chart with a series of numbers pointing to different areas.
* Evidence map 10 is a 5x9 table with blue dots in each cell that vary in size.
* VOSviewer image is a heat map or “density plot” showing a color spectrum ranging from red to green across a series of topics.

**Slide 13: R + Metafor powered Meta-Analysis**

Screenshot of a web page from EPPI-Reviewer. Images of the data collection front end of the software and displays R code, a forest plot, and other Fit Statistics.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metaphor package. Journal of Statistical Software, 36(3), 1-48. This work is licensed under the licenses Paper: Creative Commons Attribution 3.0 Unported License, <http://creativecommons.org/licenses/by/3.0/>

**Slide 14: Automation in EPPI-Centre Tools**

**Slide 15: Citation screening**

* Has received most R&D attention
* Diverse evidence base; difficult to compare evaluations
* ‘semi-automated’ approaches are the most common
* Possible reductions in workload in excess of 30% (and up to 97%)

**Summary of conclusions**

* Screening prioritisation
* ‘safe to use’
* Machine as a ‘second screener’
* Use with care
* Automatic study exclusion
* Highly promising in many areas, but performance varies significantly depending on the domain of literature being screened

In upper right corner: image of an article from *Systematic Reviews* titled “Using text mining for study identification in systematic reviews: a systematic review of current approaches,” O’Mara-Eves et. al., 2015. (Article cover reproduced per CC BY 4.0 - <https://creativecommons.org/licenses/by/4.0/legalcode>)

**Slide 16: How the machine learns…**

1. Database searches: Citations entered into database

2. Manual screening: Initial set of relevant and irrelevant studies is identified from a random sample of citations

3. Machine learning: Machine is ‘trained’, learning from the manually screened citations. List of studies to be screened manually in subsequent step is generated.

4. Manual screening: The list of studies generated in previous step is screened manually. If the stopping criterion has not yet been reached, the previous step is re-run, incorporating the new screening decisions.

5. Screening completes: Classifier automatically assigns unscreened citation as being included or excluded.

Each stage progresses left to right starting with one on the left to five on the right. Stage 4 can return to Stage 3 in order for the machine to increase the number of cases that are being reviewed. And it can work quite well…

**Slide 17: Does it work? e.g. reviews from Cochrane Heart Group**

Six graphs showing six reviews. The Y-axis is the cumulative number of studies found, on the X-axis are the number of items screened. Reviews 1006, 1007, and 1125 take only a few screenings to max out the number of cumulative includes. Reviews 0902, 1004, and 1309 take a higher number of screened items in order to reach the same level. Used with permission.

**Slide 18: Finding RCTs….**

• A machine learning RCT classifier was built using more than 280,000 records from Cochrane Crowd

• 60% of the studies have scores < 0.1

• If we trust the machine, and automatically exclude these citations, we’re left with 99.897% of the RCTs (i.e. we lose 0.1%)

• Is that good enough?

• Systematic review community needs to discuss appropriate uses of automation

Two graphs on the right reflect a high frequency (Y-axis) in relation to the ScoredValue (X-axis)

**Slide 19: Using the classifiers in EPPI-Reviewer**

Screenshot of EPPI-Reviewer 4 page with an image of a graph of the Distribution of classifier scores – Model: RCT showing Number of Items (Y-axis) and Range (X-Axis).

**Slide 20: Extraction of data from graphs**

Screenshot of a graph depicting True Positive rate (Y-axis) ranging from 0 to 1 and false positive rate (X-axis) from 0 to 1. The graph depicts results of a systematic review with multi-colored vertical bars.

**Slide 21: Collaborations, connected services and data sources**

Pictured is a complex flow chart depicting EPPI-Reviewer 5 software in the middle surrounded by data services and data stores, an open access index of studies, services including machine learning, de-duplication, data extraction, prediction and recommendation services and more. Items that flow in and out of the center include:

EPPI-Reviewer software for annotation, curation, and evidence synthesis; Human Behaviour Change Project web portal; Evidence Mapping service and visualisations; Meta-analysis in education; Guideline development (e.g. MAGICapp); Cochrane Services (CRSR/RevMan, Crowd…)

**Slide 22: Coming soon…EPPI-Reviewer version 5**

Screenshot of EPPI-Reviewer5 (Alpha)

The image shows the List tab of the new interface along with a table with columns for study ID, Title, Authors, and Year.

**Slide 23: Thank you**

Website: EPPI-Centre Website [http://eppi.ioe.ac.uk](http://eppi.ioe.ac.uk/)

Twitter: @EPPICentre

Twitter: @james\_m\_thomas

Email: James.thomas@ucl.ac.uk

*All screenshots in this presentation are used with permission.*

An Introduction to Systematic Reviews (Cover reprinted with permission)

Top right: IOE London logo. Leading education and social research. Institute of Education, University of London.

Image on right: London at night - the London Eye on left, bridge over River Thames center, Westminster Palace in background on right.

Contact information below image:

EPPI-Centre

Social Science Research Unit

UCL Institute of Education

University College London

18 Woburn Square

London WC1H 0NR

Tel +44 (0)20 7612 6397

Fax +44 (0)20 7612 6400

Email eppi@ioe.ac.uk

Web eppi.ioe.ac.uk/

**Slide 24**: **Disclaimer**

The contents of this presentation were developed by the EPPI-Centre for grant number 90DP0027 from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this presentation do not necessarily represent the policy of NIDILRR, ACL, HHS, and you should not assume endorsement by the Federal Government.

Bottom left AIR logo: gray and blue column on left; letters in blue, AIR (R) on the right; words below in blue, American Institutes for Research (R).