Session 1: Facilitating review and data reuse across the research ecosystem

Systematic Review Data Sharing and Standards

Ian Saldanha
Johns Hopkins Bloomberg School of Public Health, USA
Systematic Review Data Sharing and Standards

Ian Saldanha, MBBS, MPH, PhD
Associate Professor, Johns Hopkins Bloomberg School of Public Health

September 2, 2023

isaldan1@jhu.edu
@saldanha_ian
Disclosures

Director, Systematic Review Data Repository Plus (SRDR+)

AHRQ has funded SRDR+ since platform launch (2012).

Disclaimer: No statement in my presentation should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services (DHHS).
Interoperability in the evidence ecosystem

Trustworthy, efficient and integrated Evidence Ecosystem

Synthesize evidence
Relevant, timely, and living systematic reviews and HTA incorporating new data within existing knowledge

Produced evidence
Relevant and high-quality primary research, real world evidence, and big data

Produce and disseminate guidance
Trustworthy decision aids, clinical practice guidelines and HTA reports for patients, clinicians and policy-makers

Implement and evaluate
Clinical decision support and quality improvement initiatives, linked to impact evaluation on practice and patient outcomes in dynamic registries, pragmatic trials etc.

Why share systematic review data?

- **Open science**
  - Allows data/code verification/replication
- **Good science**
  - Allows checking robustness of data
- **Efficient science**
  - Facilitates updates
  - Facilitates inclusion in overviews, guidelines

Page MJ, Nguyen P-Y, Hamilton DG, Haddaway NT, Kanukula R, Moher D, McKenzie JE. Data and code availability statements in systematic reviews of interventions were often missing or inaccurate: a content analysis. *J Clin Epidemiol* 2022. DOI: [10.1016/j.jclinepi.2022.03.003](https://doi.org/10.1016/j.jclinepi.2022.03.003)

The evidence synthesis community is not doing great in sharing our data!

Random sample of 300 systematic reviews of interventions published in 2020

Only 86 (29%) had data availability statements!

Only 12 (4%) had data available for download (from the journal, a website, or a repository)!

Page MJ, Nguyen P-Y, Hamilton DG, Haddaway NT, Kanukula R, Moher D, McKenzie JE. Data and code availability statements in systematic reviews of interventions were often missing or inaccurate: a content analysis. J Clin Epidemiol 2022. DOI: 10.1016/j.jclinepi.2022.03.003
Journals mandating data sharing or inclusion of data availability statements makes a difference!

<table>
<thead>
<tr>
<th>Reported item</th>
<th>Mandatory requirement (%)</th>
<th>No mandatory requirement (%)</th>
<th>Percentages</th>
<th>Risk ratio (95% CI)</th>
<th>Risk ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All search strategies</td>
<td>31/87 (36)</td>
<td>50/213 (23)</td>
<td></td>
<td>1.5 (1.0 to 2.2)</td>
<td></td>
</tr>
<tr>
<td>Sharing of other data/material</td>
<td>16/87 (18)</td>
<td>4/213 (2)</td>
<td></td>
<td>10.7 (3.7 to 31.5)</td>
<td></td>
</tr>
<tr>
<td>Unprocessed extracted data</td>
<td>7/87 (8)</td>
<td>2/213 (1)</td>
<td></td>
<td>7.9 (1.8 to 33.7)</td>
<td></td>
</tr>
<tr>
<td>Data conversions performed</td>
<td>1/87 (1)</td>
<td>0/213 (0)</td>
<td></td>
<td>7.4 (0.3 to 183.5)</td>
<td></td>
</tr>
<tr>
<td>Data used in analysis</td>
<td>11/87 (13)</td>
<td>1/213 (&lt;0.5)</td>
<td></td>
<td>21.3 (3.8 to 119.1)</td>
<td></td>
</tr>
<tr>
<td>Analytic code</td>
<td>1/87 (1)</td>
<td>1/213 (&lt;0.5)</td>
<td></td>
<td>2.5 (0.3 to 23.9)</td>
<td></td>
</tr>
<tr>
<td>Citations of included and excluded studies</td>
<td>2/87 (2)</td>
<td>0/213 (0)</td>
<td></td>
<td>12.5 (0.6 to 262.8)</td>
<td></td>
</tr>
<tr>
<td>Metadata of shared files</td>
<td>1/87 (1)</td>
<td>0/213 (0)</td>
<td></td>
<td>7.4 (0.3 to 183.5)</td>
<td></td>
</tr>
</tbody>
</table>

Fig 5 | Association between journals’ data sharing requirements and reported items. Mandatory requirement—a mandatory instruction for sharing of data and materials, or in the absence of such data, a data availability statement stating why data were not shared and whether data are available on request. Equivalence range=0.9-1.1

Barriers to data sharing

- Insufficient motivation/career incentives
- Insufficient time
- Insufficient technical expertise
- Insufficient templates to facilitate data sharing
- Concerns about data ownership
- Fear of criticism

“Data sharing via supplementary files or public repositories is an effective tool to improve reproducibility of systematic reviews and should be made a standard practice.”

**FigShare. The Systematic Review Data Repository (SRDR) is an example of a repository for sharing materials specific to systematic reviews.**

Page MJ, Nguyen P-Y, Hamilton DG, Haddaway NT, Kanukula R, Moher D, McKenzie JE. Data and code availability statements in systematic reviews of interventions were often missing or inaccurate: a content analysis. *J Clin Epidemiol* 2022. DOI: 10.1016/j.jclinepi.2022.03.003
Goals and considerations of a systematic review

• Systematic reviewers need to conduct reviews with the goal in mind
  • Clinical/public health practice guidelines
  • Policy decisions
  • Academic interests
  • Business interests
  • Numerous other goals

• Considerations
  • Efficiency of the systematic review process
  • Use of the data by guideline developers (and others in the evidence ecosystem, e.g., clinical decision support tool developers)
A data sharing platform specific to systematic reviews

**Free** platform with two main purposes:
1. Data management (screening, extraction)
2. Data archiving, sharing, and re-use

A community resource

User accounts = **11,006**

Data shared publicly by systematic review authors:

Projects = **246**
Studies = **21,174**

(As of July 31, 2023)

https://srdrplus.ahrq.gov
FYI – Long oral presentation on SRDR+

Presentation Title:
The improved Systematic Review Data Repository Plus (SRDR+): A free, “FHIR-ed up” tool for screening, data extraction, and data sharing

Date: Tuesday September 5
Time: 11:25 am to 11:45 am
Session Title: Evidence synthesis innovations and technology
Room: Churchill
**Typical traditional data export from SRDR+**

<table>
<thead>
<tr>
<th>Outcome Description</th>
<th>Type</th>
<th>Population</th>
<th>Digest</th>
<th>Timing</th>
<th>Outcome</th>
<th>Outcome Value</th>
<th>Total N</th>
<th>Events</th>
<th>Percent</th>
<th>Outcome</th>
<th>Outcome Value</th>
<th>Total N</th>
<th>Events</th>
<th>Percent</th>
<th>Total N</th>
<th>Events</th>
<th>Percent</th>
<th>Events</th>
<th>Percent</th>
<th>Outcome</th>
<th>Outcome Value</th>
<th>Total N</th>
<th>Events</th>
<th>Percent</th>
<th>Events</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breastfeeding initial Any breast Category All Partic e564e7/4</td>
<td>4 mo</td>
<td>Video arm</td>
<td>30 16</td>
<td>40</td>
<td>Peer col</td>
<td>23 12</td>
<td>32</td>
<td>Video arm</td>
<td>30 21</td>
<td>70 Peer coll</td>
<td>75</td>
<td>100</td>
<td>65</td>
<td>54</td>
<td>54 8.8</td>
<td>No home</td>
<td>100 65</td>
<td>76.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breastfeeding initial Any breast Category All Partic e564e7/4</td>
<td>2 mo</td>
<td>Video arm</td>
<td>35 25</td>
<td>80 Peer coll</td>
<td>32 23</td>
<td>72</td>
<td>Video arm</td>
<td>32 23</td>
<td>72</td>
<td>Video arm</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breastfeeding initial Any breast Category All Partic e564e7/4</td>
<td>1.4 wk</td>
<td>Video arm</td>
<td>35 25</td>
<td>80 Peer coll</td>
<td>32 23</td>
<td>72</td>
<td>Video arm</td>
<td>32 23</td>
<td>72</td>
<td>Video arm</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breastfeeding initial Breastfe Category All Partic 65c30c0</td>
<td>PP</td>
<td>Home visit</td>
<td>72 61</td>
<td>84.4 Home visit</td>
<td>65 54</td>
<td>54 8.8</td>
<td>Home visit</td>
<td>100 65</td>
<td>76.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive initial Continuation Category All Partic c53893t</td>
<td>3 mo</td>
<td>Immediate</td>
<td>15 13</td>
<td>86.7</td>
<td>PP levonorgestrel</td>
<td>100 65</td>
<td>76.5</td>
<td>Immediate</td>
<td>15 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive initial Continuation Category All Partic c53893t</td>
<td>6 mo</td>
<td>Immediate</td>
<td>15 13</td>
<td>86.7</td>
<td>PP levonorgestrel</td>
<td>100 65</td>
<td>76.5</td>
<td>Immediate</td>
<td>15 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive initial Any breast Category All Partic e564e7/4</td>
<td>6 mo</td>
<td>Immediate</td>
<td>15 13</td>
<td>86.7</td>
<td>PP levonorgestrel</td>
<td>100 65</td>
<td>76.5</td>
<td>Immediate</td>
<td>15 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td>16 14</td>
<td>93.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive initial Any breast Category All Partic e564e7/4</td>
<td>2 wk</td>
<td>Home visit</td>
<td>580 78</td>
<td>13 Pediatric</td>
<td>583 73</td>
<td>12</td>
<td>Home visit</td>
<td>583 73</td>
<td>12</td>
<td>Home visit</td>
<td>583 73</td>
<td>12</td>
<td>Home visit</td>
<td>583 73</td>
<td>12</td>
<td>Home visit</td>
<td>583 73</td>
<td>12</td>
<td>Home visit</td>
<td>583 73</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive initial Any breast Category All Partic e564e7/4</td>
<td>1.5 mo</td>
<td>Home visit</td>
<td>580 163</td>
<td>28 Pediatric</td>
<td>583 182</td>
<td>31</td>
<td>Home visit</td>
<td>583 182</td>
<td>31</td>
<td>Home visit</td>
<td>583 182</td>
<td>31</td>
<td>Home visit</td>
<td>583 182</td>
<td>31</td>
<td>Home visit</td>
<td>583 182</td>
<td>31</td>
<td>Home visit</td>
<td>583 182</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive initial Any breast Category All Partic e564e7/4</td>
<td>2 wk</td>
<td>Home visit</td>
<td>580 126</td>
<td>22 Pediatric</td>
<td>583 123</td>
<td>21</td>
<td>Home visit</td>
<td>583 123</td>
<td>21</td>
<td>Home visit</td>
<td>583 123</td>
<td>21</td>
<td>Home visit</td>
<td>583 123</td>
<td>21</td>
<td>Home visit</td>
<td>583 123</td>
<td>21</td>
<td>Home visit</td>
<td>583 123</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive initial Breastfe Category All Partic 73423f7t</td>
<td>3 mo</td>
<td>Home visit</td>
<td>507 227</td>
<td>45 Pediatric</td>
<td>498 239</td>
<td>48</td>
<td>Home visit</td>
<td>498 239</td>
<td>48</td>
<td>Home visit</td>
<td>498 239</td>
<td>48</td>
<td>Home visit</td>
<td>498 239</td>
<td>48</td>
<td>Home visit</td>
<td>498 239</td>
<td>48</td>
<td>Home visit</td>
<td>498 239</td>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive initial Breastfe Category All Partic 73423f7t</td>
<td>2 wk</td>
<td>Home visit</td>
<td>521 95</td>
<td>18 Pediatric</td>
<td>505 111</td>
<td>22</td>
<td>Home visit</td>
<td>505 111</td>
<td>22</td>
<td>Home visit</td>
<td>505 111</td>
<td>22</td>
<td>Home visit</td>
<td>505 111</td>
<td>22</td>
<td>Home visit</td>
<td>505 111</td>
<td>22</td>
<td>Home visit</td>
<td>505 111</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive initial Attendi Urgent care Category All Partic b16034f</td>
<td>2 wk</td>
<td>Home visit</td>
<td>508 64</td>
<td>12.6 Hospital</td>
<td>506 73</td>
<td>14.4</td>
<td>Home visit</td>
<td>506 73</td>
<td>14.4</td>
<td>Home visit</td>
<td>506 73</td>
<td>14.4</td>
<td>Home visit</td>
<td>506 73</td>
<td>14.4</td>
<td>Home visit</td>
<td>506 73</td>
<td>14.4</td>
<td>Home visit</td>
<td>506 73</td>
<td>14.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive initial Attendi Urgent care Category All Partic b16034f</td>
<td>1.5 mo</td>
<td>Home visit</td>
<td>508 131</td>
<td>26.8 Hospital</td>
<td>506 139</td>
<td>27.5</td>
<td>Home visit</td>
<td>506 139</td>
<td>27.5</td>
<td>Home visit</td>
<td>506 139</td>
<td>27.5</td>
<td>Home visit</td>
<td>506 139</td>
<td>27.5</td>
<td>Home visit</td>
<td>506 139</td>
<td>27.5</td>
<td>Home visit</td>
<td>506 139</td>
<td>27.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraceptive initial Attendi Urgent care Category All Partic e56183d</td>
<td>2 wk</td>
<td>Home visit</td>
<td>508 2</td>
<td>6.4 Hospital</td>
<td>506 1</td>
<td>0.2</td>
<td>Home visit</td>
<td>506 1</td>
<td>0.2</td>
<td>Home visit</td>
<td>506 1</td>
<td>0.2</td>
<td>Home visit</td>
<td>506 1</td>
<td>0.2</td>
<td>Home visit</td>
<td>506 1</td>
<td>0.2</td>
<td>Home visit</td>
<td>506 1</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental health symptomatic Significance Category All Partic 521672b</td>
<td>2 wk</td>
<td>Home visit</td>
<td>508 103</td>
<td>20.8 Hospital</td>
<td>506 86</td>
<td>17.9</td>
<td>Home visit</td>
<td>506 86</td>
<td>17.9</td>
<td>Home visit</td>
<td>506 86</td>
<td>17.9</td>
<td>Home visit</td>
<td>506 86</td>
<td>17.9</td>
<td>Home visit</td>
<td>506 86</td>
<td>17.9</td>
<td>Home visit</td>
<td>506 86</td>
<td>17.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breastfeeding initial Any breast Category All Partic 50481b0</td>
<td>2 wk</td>
<td>Home visit</td>
<td>508 423</td>
<td>83.3 Hospital</td>
<td>506 424</td>
<td>83.8</td>
<td>Home visit</td>
<td>506 424</td>
<td>83.8</td>
<td>Home visit</td>
<td>506 424</td>
<td>83.8</td>
<td>Home visit</td>
<td>506 424</td>
<td>83.8</td>
<td>Home visit</td>
<td>506 424</td>
<td>83.8</td>
<td>Home visit</td>
<td>506 424</td>
<td>83.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breastfeeding initial Exclusiv w Cat * Category All Partic 729689c</td>
<td>3 mo</td>
<td>Text message</td>
<td>67 22</td>
<td>32.8 Breastfeed</td>
<td>62 19</td>
<td>30.6</td>
<td>Text message</td>
<td>62 19</td>
<td>30.6</td>
<td>Text message</td>
<td>62 19</td>
<td>30.6</td>
<td>Text message</td>
<td>62 19</td>
<td>30.6</td>
<td>Text message</td>
<td>62 19</td>
<td>30.6</td>
<td>Text message</td>
<td>62 19</td>
<td>30.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breastfeeding initial Exclusiv w Cat * Category All Partic 729689c</td>
<td>2 wk</td>
<td>Text message</td>
<td>71 36</td>
<td>50.7 Breastfeed</td>
<td>56 26</td>
<td>44.8</td>
<td>Text message</td>
<td>56 26</td>
<td>44.8</td>
<td>Text message</td>
<td>56 26</td>
<td>44.8</td>
<td>Text message</td>
<td>56 26</td>
<td>44.8</td>
<td>Text message</td>
<td>56 26</td>
<td>44.8</td>
<td>Text message</td>
<td>56 26</td>
<td>44.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Various sources of data for systematic reviews

**Public data sources**
- Journal article
- Short report (e.g., letter, abstract)
- Trial registration
- Results on trial registry
- Information from regulators

**Non-public data sources**
- Unpublished manuscript
- Individual participant data
- Grant proposal
- Study protocol
- Case report form
- Memos and emails

*Slide adapted from Dr. Peter Doshi*
But what data items should be shared with others in the ecosystem?

AHRQ Evidence-based Practice Center Program-identified minimum data items for systematic review data sharing

- For all studies in AHRQ-funded systematic reviews of comparative effectiveness, comparative harms, or diagnostic accuracy
- Contractual deliverable for each systematic review contract
- Structured dataset that is separate from the systematic review report and journal articles
Minimum items for data sharing – Level 1: Study level

Study Citation Information (below)
- Title
- Authors
- Year of Publication (or of presentation if conference abstract only)
- Journal (or conference name if conference abstract only)
- Volume
- Issue
- Page Numbers (or abstract number, if conference abstract only)
- PubMed ID (if available)
- Digital Object Identifier (DOI)
- Registration number (e.g., NCT number for ClinicalTrials.gov)

Study Characteristics
- Study design (RCT or not RCT)
- Funder type

Arm Details
- Name of each arm (group), e.g., intervention, diagnostic test

Sample Characteristics
- Country
- Overall sample size at baseline
- Sample size at baseline for each arm (group)
- Age
- Sex
- Race/ethnicity

Risk of Bias
- Risk of bias rating – for each item/domain in tool
- Risk of bias rating – overall (for main outcome of the report)

Outcomes
- Name of each extracted outcome (only outcomes prioritized for Strength of Evidence tables)
Minimum items for data sharing – Level 2: Review level

**Meta-Data**
- Project Name
- Attribution (e.g., EPC name)
- Authors of report (when available)
- Methodology description (Systematic review vs. rapid review vs. technical brief vs. evidence map)
- PROSPERO Registration ID
- DOI of AHRQ report (when available)
- Funding source

**PICODS for each Key Question**
- Populations
- Interventions/Exposures
- Comparators
- Outcomes
- Study Designs
- Settings

**Summary of evidence (for each Key Question)**
- Strength of evidence tables (summary of prioritized outcomes, findings, and strength of evidence)

**Meta-analysis results (for each Key Question)**
- Estimates from pair-wise meta-analyses (only outcomes prioritized for Strength of Evidence tables) (if conducted)
- Estimates from network meta-analyses (only outcomes prioritized for Strength of Evidence tables) (if conducted)
Making evidence synthesis data interoperable with other platforms

For interoperability, the data must be computable/machine-readable
• For that, the data structure needs to follow a standard.

There are now rigorously developed standards (FHIR) for most information relevant to systematic reviews.
  (e.g., citation information, study design, risk of bias, results)

*Great opportunity for systematic review data to be interoperable among platforms!*
So, where is SRDR+ with FHIR standards?

We are making the data in SRDR+ fully compliant with FHIR standards

• Data will be usable by other platforms in a machine-readable way using “API endpoints”

• Preliminary version done → Being refined

• Other platforms are also working on this
Some broad opportunities for interoperability of structured data in the ecosystem

1. Exporting data among platforms (for systematic review updates)
   One data extraction platform $\rightarrow$ Another data extraction platform

2. Exporting risk of bias ratings and meta-analysis results
   Data extraction platform $\rightarrow$ MAGICapp (a guideline authoring platform)

3. Importing information from systematic review protocol registers
   PROSPERO $\rightarrow$ Data extraction platform

4. Importing study information from registries (*Lene Seidler’s talk will discuss this*)
   ClinicalTrials.gov $\rightarrow$ Data extraction platform
1. Systematic reviews are not done in a vacuum. They are part of an evidence ecosystem and should be done with that goal in mind.

2. Prioritize bidirectional communication (with guideline developers and trial data generators).

3. The future calls for structured data outputs and data sharing, which can
   • Maximize utility of the evidence
   • Facilitate guideline development
   • Help reduce research waste
   • Contribute to open science.
Thank you!

Ian Saldanha
isaldan1@jhu.edu
@saldanha_ian