Instruments for Assessing Risk of Bias and Other Methodological Criteria of Published Animal Studies: A Systematic Review

January 14, 2015

David Krauth¹, Tracey Woodruff², Lisa Bero¹, ³, ⁴

¹ University of California, San Francisco, Department of Clinical Pharmacy, San Francisco, CA
² University of California, San Francisco, Department of Obstetrics, Gynecology, and Reproductive Sciences, San Francisco, CA; Program on Reproductive Health and the Environment (PRHE), Oakland, CA
³ Institute of Health Policy Studies, UCSF School of Medicine
⁴ San Francisco Branch of United States Cochrane Center

Funding Source: National Institute of Environmental Health Sciences (Grant # R 21ES 021028)
Instruments for Assessing Risk of Bias and Other Methodological Criteria of Published Animal Studies: A Systematic Review

David Krauth, Tracey J. Woodruff and Lisa Bero

http://dx.doi.org/10.1289/ehp.1206389

Online 14 June 2013
Disclosure Statement

• All authors declare that they have no conflicts of interest to disclose.
Risks of Bias IS NOT Reporting or Quality

• **Risks of bias**  
  Methodological criteria that can introduce a systematic error in the magnitude or direction of the results (Higgins and Green 2008)

• **Quality**  
  Study criteria related to how a study is conducted (e.g., in compliance with human subjects guidelines)

• **Reporting**  
  Completeness of information (e.g. study population described)
Why Assess Risk of Bias?

- **Efficacy:**
  - Effect Size: ↑
  - (Schulz and Grimes, 2002)

- **Harm:**
  - Effect Size: ↓
  - (Nieto et al. 2007)

- Improves confidence in data

- Critical step in systematic review process
Example of High Risk of Bias

Reported drug efficacy was significantly lower in studies that reported measures taken to conceal treatment allocation from the time of cerebral ischemia up to the time of outcome assessment (25.1% versus 54.0%; \( P < 0.001 \))

Macleod et al. 2008
Systematic Review Protocol

1. State objective
2. Selection criteria
3. Search strategy
4. Apply selection criteria
   - In duplicate, reproducible, transparent
5. Assess risk of bias of included studies
6. Analyze results, using meta-analysis if appropriate
Study Objective

Identify and summarize existing instruments for animal studies

- Clinical Drug Trials
- Preclinical Drug Studies
- Environmental Toxicology
Methods

**Search Strategy***
- Medline (January 1966 - November 2011)
- Reference lists

**Inclusion Criteria**
- Instruments for assessing risk of bias in animal studies
- English

**Exclusion Criteria**
- Review articles
- Application of an instrument

Krauth et al, 2013

Methods

Data Extraction – Instrument Characteristics

• Animal model
• Number of criteria
• Date of publication
• Tested for reliability
• Tested for validity
Methods

We extracted risk of bias criteria, reporting criteria, and other methodological characteristics
Study Design Elements Aimed at Reducing Bias and other Methodological Characteristics

1. Treatment allocation/Randomization
2. Concealment of Allocation
3. Blinding of Investigators
4. Inclusion/Exclusion Criteria
5. Sample Size Calculation
6. Compliance with Animal Welfare Requirements
7. Financial Conflicts of Interest
8. Statistical Model Explained
9. Use of Animals with Comorbidity
10. Test Animal Descriptions
11. Dose/Response (D/R) Model
12. All Animals Accounted for
13. Optimal Time Window Investigated
Flow Chart for Study Inclusion

Search Results
(n = 3731)

Citations excluded for not meeting inclusion criteria (n = 3643)

Evaluated full text (n = 88)

Studies excluded for not meeting inclusion criteria (n = 60)

Additional studies added based on screening references (n = 2)

Articles meeting inclusion criteria for systematic review (n = 30)
Results: Risk of Bias (n=30)

Number of Criteria Assessed: 2 – 25  
Date of Publication: 1993-2011  

Note: Dark bars represent criteria with empirical basis
Results: Reporting and Other Methodological Criteria (n = 30)

Number of instruments containing each criterion

- Sample size calculation
- Control group
- Statistical model explained
- Monitoring of physiological parameters
- Assessment > 2 outcomes
- Compliance with animal welfare requirements
- Clear hypothesis/research objective
- Outcome assessment in chronic and acute phase
- Rationale for animal model used
- Animal housing/husbandry

Note: Dark bars represent criteria with empirical basis
### Criteria with Empirical Evidence

<table>
<thead>
<tr>
<th>Type of Bias</th>
<th>Risk of Bias Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Selection</strong></td>
<td><em>Empirically tested in animal models</em></td>
</tr>
</tbody>
</table>
  *Concealment of allocation* (Macleod et al. 2008)* |
| **Performance**              | *Empirically tested in animal models*                                                |
| Systematic difference between treatment and control groups with regard to care or other exposure besides the intervention (Higgins and Green, 2008). | *Blinding* (Bebarta et al. 2003, Sena et al. 2007, Vesterinen et al. 2010)  
  *Use of animals with identical co-morbid conditions* (Crossley et al. 2008; Macleod et al. 2004; Macleod et al. 2008; Sena et al. 2007)  
  *Identical housing/husbandry conditions between treatment groups* (Duke et al. 2001; Gerdin et al. 2012)* |
| **Detection**                | *Empirically tested in animal models*                                                |
| Systematic differences between treatment and control groups with regards to how outcomes are assessed | *Blinding* (Bebarta et al. 2003; Vesterinen et al. 2010)  
  *Optimal time window investigated for outcome assessment* (EPA 2009)* |
| **Exclusion**                | *Empirically tested in clinical trials*                                              |
| Systematic difference between treatment and control groups in the number of animals that were included in and completed the study. | *Data on whether all animals are accounted for* (Tierney and Stewart 2005)  
  *Intention-to-treat analysis performed* (Melander et al. 2003; Porta et al. 2007)* |
| **Other Bias**               | *Empirically tested in animal models*                                                |
| Sample size calculation (Vesterinen et al. 2010) | *Test animal details* (Macleod et al. 2004; Sniekers et al. 2008)  
  *Appropriateness of dose selection (validated by use of a dose/response model)* (Bucher et al. 1996)  
  *Timing of exposure* (Benatar 2007; van der Worp et al. 2010; Vesterinen et al. 2010)  
  *Measurement of outcomes that are sensitive to the exposure* (Wood 2000)* |
| Type of funding source (Lundh et al. 2012) | *Empirically tested in clinical trials*                                              |
| Financial conflicts of interest stated (Lundh et al. 2012) | *Selective outcome reporting* (Hart et al. 2012; Rising et al. 2008)* |
# Summary of Commonly Used Instruments

<table>
<thead>
<tr>
<th>CHECKLIST</th>
<th>INSTRUMENT DESCRIPTION</th>
</tr>
</thead>
</table>
| Agerstrand et al 2011            | • 25 item instrument  
• Not empirically tested  
• No methodological score is used  
• Intended use of instrument is environmental toxicology research                                                                                                  |
| Kilkenny et al, 2010 The ARRIVE Guidelines | • 13 item instrument  
• Not empirically tested  
• No methodological score is used is used  
• No specific disease modeled  
• Developed using the CONSORT criteria as a foundation, and consensus and consultation from scientists, statisticians, journal editors, and research funders |
| Sena et al, 2007                 | • 21 item instrument  
• No methodological score is used  
• Provide empirical data for randomization and blinding  
• Disease modeled is stroke  
• Instrument derived from 4 previous checklists: STAIR, Amsterdam Criteria (Horn et al. 2001), CAMARADES, Utrecht Criteria (van der Worp et al. 2005) 
• Instrument appears to have validity |
Limitations of the Instruments (n = 30)

- Few instruments developed for animal toxicology (4)
- Most instruments not tested for validity and reliability
- Most instruments mix reporting, risk of bias, and other methodological criteria
Limitations of our Study

• Searched Medline database and articles published in English
Recommendation

Use of *empirically* based criteria for assessing risk of bias in animal toxicology studies
Acknowledgments

• Our funding source, the National Institute of Environmental Health Sciences

• Gloria Won (UCSF Mount Zion Campus) for assistance with developing the search strategy

• Dr. Dorie Apollonio (UCSF), Dr. David Dorman (North Carolina State University), and Rose Philipps (UCSF) for reviewing the manuscript
Questions?
Supplemental Slides NEXT
SUPPLEMENTAL SLIDE #1
Randomization

• 25 of 30 instruments include random allocation of treatment

• A systematic review of multiple sclerosis interventions in animal research has shown that non-randomized studies report significantly higher treatment efficacy (41.6%, 95% CI 36.7-46.5%) than randomized studies (20.6%, 95% CI 11.4-29.7%)
  (Vesterinen et al. 2010)

• In emergency medicine, animal studies lacking randomization were over three times more likely to show a statistically significant result relative to studies that included these attributes
  (Bebarta et al. 2003)
SUPPLEMENTAL SLIDE #2
Blinding of Investigators

• 23 of 30 instruments include blinding

• Blinding in experimental stroke studies significantly alters the effectiveness of an intervention with effect sizes ranging by 10% in studies with or without this feature (Crossley et al. 2008)

• A systematic review of multiple sclerosis interventions has shown that studies performed without blinded assessment of outcome report higher efficacy estimates (41.0%, 95% CI 36.2–45.8%) compared to blinded studies (29.8%, 95% CI 19.8–39.8%) (Vesterinen et al. 2010)
SUPPLEMENTAL SLIDE #3
Financial Conflict of Interest

• 9 of 25 instruments include disclosure of conflicts of interest

• Reviews of clinical studies have shown that study funding sources and financial ties of investigators (including university or industry affiliated investigators) are associated with favorable research outcomes for the sponsors [efficacy results risk ratio (RR): 1.32; harm results RR: 1.87] even when controlling for other risks of bias.

(Lundh et al. 2012)
SUPPLEMENTAL SLIDE #4
Animals with Co-morbidity

- 6 of 30 instruments state the need to use animals with pre-existing co-morbidity.

- Using co-morbid animals in experimental stroke studies was found to significantly alter the effectiveness of an intervention with effect sizes ranging by 10% in studies with or without these features
  
  (Crossley et al. 2008)
SUPPLEMENTAL SLIDE #5

Test Animal Details

• 14 of 30 instruments state the need to include detailed reporting of test animal characteristics

• In a meta-analysis containing 14 animal studies, it was determined that the efficacy of using nicotinamide to treat stroke outcomes depends on animal species and sex. Drug efficacy was effective in rats but not mice ($p < 0.0001$) and male species performed better than females ($p = 0.012$).

(Macleod et al. 2004)
Was every animal accounted for?

- 7 of 30 instruments include assessing whether all animals were accounted for.

- In a study comparing clinical data from 14 meta-analyses that addressed therapeutic treatments for cancer, it was shown that not accounting for all patients leads to more favorable research outcomes \( (p\text{-value} = 0.03) \) relative to studies that do account for all patients.

  (Tierney and Stewart 2005)
SUPPLEMENTAL SLIDE #7: References Cited

SUPPLEMENTAL SLIDE #8: References Cited