The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomized Studies in Meta-Analysis


Dr Daniel Pope – Lecturer in Epidemiology
Dr Nigel Bruce – Reader in Public Health
University of Liverpool, UK

Background

• Observational studies – aetiological hypotheses (small RR but large PAF)
• Systematic review methodology (inc. meta-analysis) attractive – precise estimate when magnitude of risk is small
• Caution required (susceptibility to bias)!
Bias and Confounding

• “…thorough consideration of sources of heterogeneity between observational study results…” Egger et al, 2003

Newcastle-Ottowa Scale

“Easy, convenient tool for quality assessment of non-randomised studies”
Newcastle-Ottowa Scale

Case-Control Studies and Cohort Studies

Star system based on three domains:
1) Selection of Study Groups
2) Comparability of Groups
3) Ascertainment of exposure/outcome

Development: Grouping Items

• Cohort studies
  • Selection of cohorts
  • Comparability of cohorts
  • Assessment of outcome

• Case-Control studies
  • Selection of case and controls
  • Comparability of cases and controls
  • Ascertainment of exposure
Development: Identifying Items

- Identify ‘high’ quality choices with a ‘star’
- A maximum of one ‘star’ for each item within the ‘Selection’ and ‘Exposure/Outcome’ categories; maximum of two ‘stars’ for ‘Comparability’

Current Development: Validity

- Face/content validity
- Criterion validity
- Construct validity
- Inter and Intra-rater Reliability
Future Development: Scoring

- Identify threshold score distinguishing between ‘good’ and ‘poor’ quality studies

Newcastle - Ottawa Quality Assessment Scale
Case Control Studies

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection
1) Is the case definition adequate?
   a) yes, with independent validation
   b) yes, eg record linkage or based on self reports
   c) no description
2) Representativeness of the cases?
   a) consecutive or obviously representative series of cases
   b) potential for selection biases or not stated
3) Selection of Controls
   a) community controls
   b) hospital controls
   c) no description
4) Definition of Controls
   a) no history of disease (endpoint)
   b) no description of source

Comparability
1) Comparability of cases and controls on the basis of the design or analysis
   a) study controls for ________ ______ (Select the most important factor.)
   b) study controls for any additional factor

Exposure
1) Ascertainment of exposure?
   a) secure record (eg surgical records)
   b) structured interview where blind to case/control status
   c) interview not blinded to case/control status
   d) written self report or medical record only
   e) no description
2) Same method of ascertainment for cases and controls?
   a) yes
   b) no

Non-Response Rate
1) Non-Response rate?
   a) non-respondent for both groups
   b) non-respondent described
   c) rate different and no designation
Newcastle-Ottawa Quality Assessment Scale: Case-Control Studies

- Selection (4)

- Comparability (1)

- Exposure (3)
  - A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability

**Selection**

1. **Is the case definition adequate?**
   a) yes, with independent validation ♦
   b) yes, eg record linkage or based on self reports
   c) no description

2. **Representativeness of the cases**
   a) consecutive or obviously representative series of cases ♦
   b) potential for selection biases or not stated

3. **Selection of Controls**
   a) community controls ♦
   b) hospital controls
   c) no description

4. **Definition of Controls**
   a) no history of disease (endpoint) ♦
   b) no description of source
**Comparability**

1. **Comparability of cases and controls on the basis of the design or analysis**
   
a) study controls for ___________ (select the most important factor) ♦
   
b) study controls for any additional factor (This criteria could be modified to indicate specific control for a second important factor.) ♦

**Exposure**

1. **Ascertainment of exposure**
   
a) secure record (eg surgical records) ♦
   
b) structured interview where blind to case/control status ♦
   
c) interview not blinded to case/control status
   
d) written self report or medical record only
   
e) no description

2. **Same method of ascertainment for cases and controls**
   
a) yes ♦
   
b) no

3. **Non-Response Rate**
   
a) same rate for both groups ♦
   
b) non respondents described
   
c) rate different and no designation
Newcastle-Ottawa Quality Assessment Scale: Cohort Studies

- **Selection (4)**
  - Representativeness of the exposed cohort
    - a) truly representative of the average population in the community
    - b) somewhat representative of the average population in the community
    - c) selected group of users (nurses, volunteers)
    - d) no description
  - Selection of the non-exposed cohort
    - a) drawn from the same community as the exposed cohort
    - b) drawn from a different source
    - c) no description
  - Ascertainment of exposure
    - a) secure record (e.g., surgical records)
    - b) structured interview
    - c) written self-report
    - d) no description
  - Demonstration that outcome of interest was not present at start of study
    - a) yes
    - b) no

- **Comparability (1)**
  - Comparability of cohorts on the basis of the design or analysis
    - a) study controls for
    - b) study controls for any additional factor
    (This criterion could be modified to indicate specific control for a second important factor.)

- **Outcome (3)**
  - Assessment of outcome
    - a) independent blind assessment
    - b) record linkage
    - c) written self-report
    - d) no description
  - Was follow-up long enough for outcomes to occur
    - a) yes (select an adequate follow-up period for outcome of interest)
    - b) no
  - Adequacy of follow-up of cohorts
    - a) complete follow-up — all subjects accounted for
    - b) subjects lost to follow-up unlikely to introduce bias (small number lost — > 10% of enrollees)
    - c) follow-up rate < 10% (select an adequate %)
    - d) no statement

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability.
**Selection**

1. Representativeness of the exposed cohort
   a) truly representative of the average ________ (describe) in the community ♦
   b) somewhat representative of the average ________ in the community ♦
   c) selected group of users eg. nurses, volunteers
   d) no description of the derivation of the cohort

2. Selection of the non exposed cohort
   a) drawn from the same community as the exposed cohort ♦
   b) drawn from a different source
   c) no description of the derivation of the non exposed cohort

3. Ascertainment of exposure
   a) secure record (eg. surgical records) ♦
   b) structured interview ♦
   c) written self report
   d) no description

4. Demonstration that outcome of interest was not present at start of study
   a) yes ♦
   b) no

**Comparability**

1. Comparability of cohorts on the basis of the design or analysis
   a) study controls for ________ (select the most important factor) ♦
   b) study controls for any additional factor (This criteria could be modified to indicate specific control for a second important factor.) ♦
Outcome

1. Assessment of outcome
   a) independent blind assessment ✦
   b) record linkage ✦
   c) self report
   d) no description

2. Was follow up long enough for outcomes to occur
   a) yes (select an adequate follow up period for outcome of interest) ✦
   b) no

3. Adequacy of follow up of cohorts
   a) complete follow up - all subjects accounted for ✦
   b) subjects lost to follow up unlikely to introduce bias - small number lost - > ___ % (select an adequate %) follow up, or description of those lost ✦
   c) follow up rate < ___% (select an adequate %) and no description of those lost
   d) no statement

Risk of Low Birth Weight and Stillbirth Associated With Indoor Air Pollution From Solid Fuel Use in Developing Countries

Pope D, Epidemiologic Reviews, 2010
Steps of a Cochrane Systematic Review

- Clearly formulated question
- Comprehensive data search
- Unbiased selection and abstraction process
- Critical appraisal of data
- Synthesis of data
- Perform sensitivity and subgroup analyses if appropriate and possible
- Prepare a structured report

Objective

- Quantify the association between exposure to indoor air pollution and low birth weight
Inclusion Criteria

- Types of studies
  - All study designs (intervention; observational)
- Population
  - Live singleton births
- Exposure
  - Any reporting of exposure to IAP (including solid fuel use etc)
- Outcomes
  - Studies reporting actual birth weight or LBW (<2500g)

Steps of a Cochrane Systematic Review

- Clearly formulated question
- Comprehensive data search
- Unbiased selection and abstraction process
- Critical appraisal of data
- Synthesis of data
- Perform sensitivity and subgroup analyses if appropriate and possible
- Prepare a structured report
Search Strategy

• Electronic Search of:
  – MEDLINE
  – EMBASE
  – Cochrane Controlled Trials Register
  – CINAHL
  – LILACS

• Other Data Sources:
  – Grey literature (PASCAL, ICP)
  – Contact with experts, review of references cited in retrieved articles

Steps of a Cochrane Systematic Review

• Clearly formulated question
• Comprehensive data search
  • Unbiased selection and abstraction process
• Critical appraisal of data
• Synthesis of data
• Perform sensitivity and subgroup analyses if appropriate and possible
• Prepare a structured report
Data Extraction

• 2 independent reviewers selected studies
• 2 independent reviewers extracted data using pre-determined forms
  – study design
  – population characteristics
  – Exposure (IAP)
  – Outcomes (LBW)
  – results
• differences resolved by consensus

Results

982 (from database search)
29 (abstract review)
7 (article review)
4 (selected for review)
6 (included in review)

2 unpublished studies identified
**Steps of a Cochrane Systematic Review**

- Clearly formulated question
- Comprehensive data search
- Unbiased selection and abstraction process
- Critical appraisal of data
- Synthesis of data
- Perform sensitivity and subgroup analyses if appropriate and possible
- Prepare a structured report

**Studies included:**

- 6 studies for data extraction (from 982)
- 2 cohort
  - 2 cross-sectional
- 1 case-control
- 1 intervention study
Quality assessment:

Selection – 4 stars:
(representativeness; exposure assessment – cohort/cross-sectional; control selection – case-control)

Comparability – 2 stars:
(adjustment for main/additional confounders eg. active/passive maternal smoking, gestational age, nutrition etc)

Outcome/Exposure – 3 stars:
(adequacy of outcome (measured LBW) and exposure
(indoor air pollution – measured vs self-report)

Quality assessment:

<table>
<thead>
<tr>
<th>Study</th>
<th>Selection</th>
<th>Comparability</th>
<th>Outcome/Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boy, 2002 (CS)</td>
<td>★★★★</td>
<td>★</td>
<td>★★</td>
</tr>
<tr>
<td>Mavalankar, 1992 (CC)</td>
<td>★★★</td>
<td></td>
<td>★★★</td>
</tr>
<tr>
<td>Mishra, 2004 (CS)</td>
<td>★★</td>
<td>★</td>
<td>★★</td>
</tr>
<tr>
<td>Siddiqui, 2008 (C)</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Tielsch, 2009 (C)</td>
<td>★★★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Thompson, 2005 (RCT)</td>
<td>★★★</td>
<td>★★</td>
<td>★★</td>
</tr>
</tbody>
</table>
## Quality assessment:

<table>
<thead>
<tr>
<th>Study</th>
<th>Selection</th>
<th>Comparability</th>
<th>Outcome/ Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boy, 2002 (CS)</td>
<td>★★★★</td>
<td>★</td>
<td>★★★</td>
</tr>
<tr>
<td>Mavalankar, 1992 (CC)</td>
<td>★★★</td>
<td></td>
<td>★★★</td>
</tr>
<tr>
<td>Mishra, 2004 (CS)</td>
<td>★★</td>
<td>★</td>
<td>★★</td>
</tr>
<tr>
<td>Siddiqui, 2008 (C)</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Tielsch, 2009 (C)</td>
<td>★★★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Thompson, 2005 (RCT)</td>
<td>★★★</td>
<td>★★</td>
<td>★★</td>
</tr>
</tbody>
</table>

## Steps of a Cochrane Systematic Review

- Clearly formulated question
- Comprehensive data search
- Unbiased selection and abstraction process
- Critical appraisal of data
- Synthesis of data
- Perform sensitivity and subgroup analyses if appropriate and possible
- Prepare a structured report
Quantification of Effects

- Exposure (e.g. solid fuel vs clean fuel)
- Outcome (%LBW)
- Effect estimates (EE)
  - Relative Risk (RR)
  - Odds Ratio (OR)
- Fixed-effect meta-analysis in the absence of statistical heterogeneity

OR = 1.38 (1.25, 1.52), p<0.0001
% Low Birth Weight (<2500g): 6 studies, 8 estimates

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>log(Odds Ratio)</th>
<th>SE</th>
<th>Weight</th>
<th>N/Fixed, 95% CI</th>
<th>Odds Ratio</th>
<th>N/Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boy 2002</td>
<td>0.26236</td>
<td>0.17543</td>
<td>0.3%</td>
<td>1.38 [0.92, 1.98]</td>
<td>OR = 1.38 (1.25, 1.52)</td>
<td>p&lt;0.0001</td>
</tr>
<tr>
<td>Mowalanker 1992a</td>
<td>0.20701</td>
<td>0.10091</td>
<td>25.6%</td>
<td>1.23 [1.01, 1.50]</td>
<td>OR = 1.27 (1.10, 1.47)</td>
<td>p=0.001</td>
</tr>
<tr>
<td>Mowalanker 1992b</td>
<td>0.25878</td>
<td>0.10926</td>
<td>23.6%</td>
<td>1.49 [1.22, 1.82]</td>
<td>OR = 1.27 (1.10, 1.47)</td>
<td>p=0.001</td>
</tr>
<tr>
<td>Mrra 2004</td>
<td>0.11333</td>
<td>0.17036</td>
<td>0.6%</td>
<td>1.12 [0.90, 1.38]</td>
<td>OR = 1.12 (0.90, 1.38)</td>
<td>p=0.27</td>
</tr>
<tr>
<td>Siddiqui 2005</td>
<td>0.57058</td>
<td>0.24738</td>
<td>4.1%</td>
<td>1.77 [1.08, 2.90]</td>
<td>OR = 1.77 (1.08, 2.90)</td>
<td>p=0.02</td>
</tr>
<tr>
<td>Thompson 2006</td>
<td>0.26236</td>
<td>0.30286</td>
<td>2.8%</td>
<td>1.30 [0.72, 2.35]</td>
<td>OR = 1.30 (0.72, 2.35)</td>
<td>p=0.001</td>
</tr>
<tr>
<td>TimHich 2009(a)</td>
<td>0.36678</td>
<td>0.10204</td>
<td>24.4%</td>
<td>1.49 [1.22, 1.82]</td>
<td>OR = 1.49 (1.22, 1.82)</td>
<td>p=0.001</td>
</tr>
<tr>
<td>Tieleman 2009(b)</td>
<td>0.53683</td>
<td>0.30891</td>
<td>2.7%</td>
<td>1.73 [0.93, 3.19]</td>
<td>OR = 1.73 (0.93, 3.19)</td>
<td>p=0.09</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>0.26236</td>
<td>0.17543</td>
<td>0.3%</td>
<td>1.38 [1.25, 1.52]</td>
<td>OR = 1.38 (1.25, 1.52)</td>
<td>p&lt;0.0001</td>
</tr>
</tbody>
</table>

Heterogeneity: Ch^2 = 5.54, df= 7 (p = 0.56), I^2 = 0%
Test for overall effect: Z = 9.37 (p = 0.00001)

Decreased risk: Increased risk

Interpretation Crucial:

- Exclusion from sensitivity analysis based on (i) birth weight based on self-reports (50%), (ii) no information on gestational age and (iii) unadjusted analysis
Applications:

- Assess quality of nonrandomized studies
- Incorporate assessments in interpretation of meta-analytic results
- Valid, repeatable and simple
- Limitations:
  - Study Designs → Too Simplistic

The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomized Studies in Meta-Analysis

www.lri.ca

NOS Quality Assessment Scales:
Case-control studies
Cohort studies
Manual for NOS Scales
Recommended Reading....