STRengthening Analytical Thinking for Observational Studies: STRATOS initiative: Study Design

Presented by SJW Evans.
Lead author N Pearce
TG5 Group: M Gail, D Altman, G Collins, L Duchateau, P Sekula, M Woodward
Acknowledgments & Conflicts of Interest

European Commission appointed independent Member of Pharmacovigilance & Risk Assessment Committee of EU at the European Medicines Agency

Member of CIOMS Working Group on Meta-analysis for safety

LSHTM (but not SE) funded by several pharmaceutical companies

Member of expert working groups at the UK Medicines & Healthcare Products Regulatory Agency (MHRA)

All remarks are a personal viewpoint and have not been approved by other members of TG5
STRATOS

• STROBE has provided guidance on *Reporting*
• STRATOS is an attempt to improve conduct
• Guidance on reporting is like lighting up a room-not saying if it is clean
• Critical appraisal is saying if it is clean enough-
  – Operating theatre or coal shed?
• Guidance on conduct is to make sure the room is clean enough
Design
Planning
Execution
Analysis
Report Writing
Communication
Planning
Execution
Analysis
Report Writing
Communication
Design
Objectives
Some principles

- single observational studies rarely definitive (or perfect!)
- Assessing epidemiologic evidence -> a process of triangulation across studies, aim to contribute to the pool of knowledge
- different populations, variety of designs, investigators, and methods,
- often involving meta-analysis (not “top of the hierarchy”) & integration of data from variety of sources and study types
- obtain valid effect estimates in a particular population during a particular risk period
Purposes

1. Descriptive
   – Disease oriented
   – Intervention oriented
     • Intervention utilisation
       – E.g. Compliance with Summary of Product Characteristics (label)
     • Risk factor distribution
     • Spontaneous reports of adverse drug reactions

2. Comparative
   – Causal effects; benefits, comparative effectiveness, harms
Comparative studies
The main focus of epidemiology

• They will usually want to estimate causal effects; {Safety- demonstrated absence of harm}
  – Usually they focus on harms, but may also look at-
    • Benefits (often reduction in harm), comparative effectiveness,
  – Moves towards formal decision making for risk/benefit
    • Will require confirmation of benefit from RCTs in practice
Main Comparative designs

• Randomised Controlled Trials (RCTs)
  – Systematic reviews (SRs) of RCTs

• Cohort Studies
  – “Field” studies; registry-based; databases

• Case-control studies
  – “Field” studies; registry-based/aided; databases

• Distinguish incidence and prevalence in each
Time has only one direction

- In some senses all studies are based on a cohort
- RCTs have random allocation to treatment, then follow-up
- Observational studies have non-random allocation and follow-up (though if totally deterministic, then extra information outside study or strong, untestable, assumptions required to interpret them)
- The issues are
  - what is the source population?
  - what is the outcome?
    - (Incidence or prevalence)
  - How is the sampling done?
Study Design Options

• All epidemiological studies are (or should be) based on a particular population (the *source population*) followed over a particular period of time (the *risk period*)

• The different study design options differ only in how the source population is defined and how information is drawn from this population and time period
Birth

End of Follow up

Death

other death

lost to follow up

“non-diseased”
symptoms

severe disease
Cohort studies

• define a source population - exposed & unexposed
  – follow-up (FU) to event

• Not just in databases or even in large longitudinal cohort studies (ALSPAC, National Cohorts, 1958, 1970)
  – Some cohort studies (often registry-based) have no valid comparative groups

• `Self-Controlled Case Series’ is a special case

• Case-cohort design {sample <100% non-cases}
Incidence and Prevalence

- **Incidence** is the number of new cases of the condition over a specified period of time.
- **Prevalence** is the number of cases of the condition at a particular point in time.
# A Hypothetical Incidence Incidence Study

<table>
<thead>
<tr>
<th></th>
<th>Exposed</th>
<th>Non-exposed</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>1,813</td>
<td>952</td>
<td></td>
</tr>
<tr>
<td>Non-cases</td>
<td>8,187</td>
<td>9,048</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>10,000</td>
<td>10,000</td>
<td></td>
</tr>
<tr>
<td>Person-years</td>
<td>90,635</td>
<td>95,163</td>
<td></td>
</tr>
<tr>
<td>Incidence rate</td>
<td>0.0200</td>
<td>0.0100</td>
<td>2.00</td>
</tr>
<tr>
<td>Incidence proportion (risk)</td>
<td>0.1813</td>
<td>0.0952</td>
<td>1.90</td>
</tr>
<tr>
<td>Incidence odds</td>
<td>0.2214</td>
<td>0.1052</td>
<td>2.11</td>
</tr>
</tbody>
</table>
**A Hypothetical Case-Control Study**

<table>
<thead>
<tr>
<th></th>
<th>Exposed</th>
<th>Non-exposed</th>
<th>Odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>1,813</td>
<td>952</td>
<td>1,813/952</td>
</tr>
<tr>
<td>Controls</td>
<td>1,313</td>
<td>1,452</td>
<td>1,313/1,452</td>
</tr>
<tr>
<td>Odds</td>
<td>1,813/1,313</td>
<td>952/1,452</td>
<td></td>
</tr>
<tr>
<td>Odds ratio</td>
<td></td>
<td></td>
<td>2.11</td>
</tr>
</tbody>
</table>
Odds Ratio

• OR=(1813/1313)/(952/1452) = 2.11

• This *incidence case-control study* yields the same estimate as would have been obtained by an incidence study but with a much smaller number of participants because we include *all* of the cases but only a *sample* of the non-cases
Methods of Sampling Controls

• From *survivors* (non-cases at end of follow-up) = cumulative sampling

• From *source population* = case-cohort sampling

• From *person-years* = density sampling
## Methods of Sampling Controls

<table>
<thead>
<tr>
<th></th>
<th>Exposed</th>
<th>Non-exposed</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cases</strong></td>
<td>1,813</td>
<td>952</td>
<td></td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from survivors</td>
<td>1,313</td>
<td>1,452</td>
<td>2.11</td>
</tr>
<tr>
<td>from source population</td>
<td>1,383</td>
<td>1,383</td>
<td>1.90</td>
</tr>
<tr>
<td>from person-year</td>
<td>1,349</td>
<td>1,416</td>
<td>2.00</td>
</tr>
</tbody>
</table>
Misconceptions re: Case-Control Studies

• Proceeds from effect (disease) to cause (exposure), i.e. reverse causality
• Inherently more prone to bias than cohort studies
• Odds ratio only approximately estimates the relative risk
• Depends on a “rare disease” assumption
• book by Keogh & Cox for modern views (Cambridge UP 2014)
Prevalence Case-Control Studies

This *prevalence case-control study* yields the same estimate as would have been obtained by a prevalence study but with a much smaller number of participants because we include *all* of the prevalent cases but only a *sample* of the non-cases.
## Study Design Options

<table>
<thead>
<tr>
<th>Study outcome</th>
<th>Incidence</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sampling on outcome</strong></td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Incidence studies</td>
<td>Incidence case-control studies</td>
<td>Prevalence case-control studies</td>
</tr>
</tbody>
</table>
Strengths & weaknesses - RCTs

• RCTs strong for causal inference (but not perfect)
  – Might be done in registries or databases
    – Staa TP et al. Pragmatic randomised trials using routine electronic health records: putting them to the test. BMJ. 2012;344:e55. also Pharmacoepidemiology session

• Weaknesses
  – Many biases can arise – selective publication worst for pepi
  – Costly in time & resources; unrepresentative?
  – Short FU; small sample size for clinical harms (& benefits?)
  – Elderly, pregnant, co-morbidity & co-prescription limited
These slides are a selection from the OMOP symposium in June 2012

Used with permission
Thanks to Paul Stang & Patrick Ryan

See also special issue of Drug Safety 2013;36 Suppl 1:S3-4.
Ground truth for OMOP 2011/2012 experiments

<table>
<thead>
<tr>
<th></th>
<th>Positive controls</th>
<th>Negative controls</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute Liver Injury</td>
<td>81</td>
<td>37</td>
<td>118</td>
</tr>
<tr>
<td>Acute Myocardial Infarction</td>
<td>36</td>
<td>66</td>
<td>102</td>
</tr>
<tr>
<td>Acute Renal Failure</td>
<td>24</td>
<td>64</td>
<td>88</td>
</tr>
<tr>
<td>Upper Gastrointestinal Bleeding</td>
<td>24</td>
<td>67</td>
<td>91</td>
</tr>
<tr>
<td>Total</td>
<td>165</td>
<td>234</td>
<td>399</td>
</tr>
</tbody>
</table>

Criteria for positive controls:
- Event listed in Boxed Warning or Warnings/Precautions section of active FDA structured product label
- Drug listed as ‘causative agent’ in Tisdale et al. 2010: “Drug-Induced Diseases”
- Literature review identified no powered studies with evidence of potential positive association

Criteria for negative controls:
- Event not listed anywhere in any section of active FDA structured product label
- Drug not listed as ‘causative agent’ in Tisdale et al, 2010: “Drug-Induced Diseases”
- Literature review identified no powered studies with evidence of potential positive association

- Isoniazid
- Fluticasone
- Indomethacin
- Clindamycin
- Ibuprofen
- Loratadine
- Sertraline
- Pioglitazone
New user cohort design applied to all test cases

Isoniazid: RR 4.04 (2.70 – 6.04)

Positive controls with larger RR:
- ARF – mefanamate
- AMI – Factor VIIa
- ALI – sulfisoxazole
- ALI – posaconazole

Negative controls with larger RR:
- UGIB- simethicone
- ARF – simethicone
- ARF – paromomycin
- ARF – clozapine
- ALI – sodium phosphate

UGIB-simethicone could be considered a classic example of confounding by indication. Why did restriction by indication and propensity score adjustment not mitigate this ‘false positive’?
New user cohort from OMOP (Drug Safety 2013)

• Applied new user cohort design consistently across 9 data sources
• Range of estimates indicate variability in observations
• 6 of 25 pairs (24%) have at least one significantly positive estimate and at least one significantly negative estimate
Protocols

• Must have research question and objectives
• Should be registered e.g. [http://www.encepp.eu/encepp/studiesDatabase.jsp](http://www.encepp.eu/encepp/studiesDatabase.jsp)
  – Declaration of Helsinki requires registration
• Design specified in detail in protocol
• Statistical analysis plan should be included
Aspects of Design for comparison

• Structure

• Designs described- case-control, self-controlled risk interval, self-controlled case series method, case-crossover

• All used for vaccine safety surveillance

• Exploratory designs (& analysis) should be clearly described as exploratory
Selection criteria

- Showed exclusion criteria widely used in pharmacoepidemiology but not well studied.
  - Exclusion criteria relating to data quality and validation were the most commonly applied (87% of publications), followed by patient characteristics (75%), disease-related (69%), exposure-related (38%) and miscellaneous (3%)
Is there a crisis in epidemiology?

• Stan Young “Any claim coming from an observational study is most likely to be wrong.” Young, S. S. and Karr, A. (2011), Deming, data and observational studies. *Significance*, 8: 116–120.

• John Ioannidis- 2015. Video of lecture can be watched through this website-. www.lshtm.ac.uk/newsevents/events/2015/07/24th-bradford-hill-memorial-lecture

• Replication of studies is not done often enough
  – Open data less likely in epidemiology
STRATOS

• The challenge is to bring statistical *thinking* into observational research including design
• We cannot write a textbook on observational research
• We may have to look at where investigators go wrong but show how we can do it well