Post Hoc Analysis

Brandford Hill


People reading research reports want investigators to answer four questions:

- Why did you start?
- What did you do?
- What answer did you get?
- What does it mean anyway?

Critical appraisal (general)

- Are the results valid?
- What are the results?
- Will they help me care for my patients?

Archie Cochrane

“It is surely a great criticism of our profession that we have not organised a critical summary, by specialty or sub-specialty, adapted periodically, of all relevant randomized controlled trials.”
### 1971

#### Cochrane Collaboration
- An international organisation that aims to help people make well-informed decisions about healthcare by preparing, maintaining and promoting the accessibility of systematic reviews of the effects of health care interventions.

### Where we were....


*"Current medical reviews do not routinely use scientific methods to identify, assess, and synthesize information."*

### More serious now

### Harder than it looks

Luc de Clapiers Vauvenarques, 1715-47 (Réflexions et Maximes)

*"Il est plus aisé de dire des choses nouvelles que de concilier celles qui ont été dites".*

'It is easier to say something new than to reconcile things that have already been said.'

or

'Synthesis of ideas is harder than merely proposing a new idea.'

### Research synthesis

- The application, in practice, of the principle that science is cumulative
- As science is cumulative, scientists should cumulate scientifically
Systematic Review

A critical assessment and evaluation of research (not simply a summary) that attempts to address a focused clinical question using methods designed to reduce the likelihood of bias.

Meta-analysis

In 1976, Gene Glass, an American social scientist coined the term 'meta-analysis' to refer to:

“the statistical analysis of a large collection of analysis results from individual studies for purposes of integrating the findings”

Will reduce the play of chance in research synthesis and will assist in drawing valid conclusions about the body of research

Develop a clear protocol

Meta-analysis

- A particular type of systematic review that uses quantitative methods to combine the results from a number of studies
- Since meta-analysis is a retrospective look at data, it is important to make the process rigorous and well defined to prevent opportunities for bias to distort the results. Only in this way can it achieve the status of a scientific discipline. This necessitates blinding the selection of papers, extraction of data and quality assessment in duplicate following an established protocol at the start of the study.

The Key Elements

What does it mean?

- Quantitative: numbers
- Systematic: methodical
- Combining: putting together
- previous research/what’s already done
- conclusions: new knowledge

Popularity

The popularity of meta analyses

Number of Meta Analysis publications are steadily increasing since 1993. We gathered the counts of journal articles including “meta analysis” or “publication type” from Pubmed, from years 1993 through 2004

Quality of Reporting of Meta-analyses (QUOROM)

<table>
<thead>
<tr>
<th>Rating</th>
<th>Subheading</th>
<th>Description</th>
<th>Expected (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td></td>
<td>Identify the report as a meta-analysis (or systematic review) of RCTs</td>
<td>100</td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
<td>Use a structured format</td>
<td>100</td>
</tr>
<tr>
<td>Objective</td>
<td></td>
<td>The clinical question explicitly</td>
<td>100</td>
</tr>
<tr>
<td>Data source</td>
<td></td>
<td>The databases (i.e., and other information sources used</td>
<td>100</td>
</tr>
<tr>
<td>Review methods</td>
<td></td>
<td>The selection criteria, population, intervention outcomes, and study design methods to yield assessment into estimation and from characteristics and quantitative data synthesis in sufficient detail to permit replication</td>
<td>100</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td>Characteristics of the RCTs included and excluded, population and quantitative findings (i.e., point estimates and confidence intervals) and subgroup analyses</td>
<td>100</td>
</tr>
<tr>
<td>Conclusion</td>
<td></td>
<td>The main results</td>
<td>100</td>
</tr>
</tbody>
</table>
QUORUM

Checklist for observational meta-analyses

Reporting of methods should include:
- Description of methods and procedures of studies included in the review
- Details of the selection and coding of studies
- Description of the data extraction and management
- Assessment of the quality of studies
- Description of the statistical methods used

Reporting of results should include:
- Table showing the results of each study
- Overview of the results

Checklist for observational meta-analyses

Reporting of background should include:
- Description of the hypothesis
- Description of the study population
- Details of the study design

Reporting of methods should include:
- Description of the data extraction and management
- Assessment of the quality of studies
- Description of the statistical methods used

Checklist for observational meta-analyses

Reporting of results should include:
- Table showing the results of each study
- Overview of the results

Benefits of Individual Patient Data

• Carry out time-to-event analyses
• Only practical way to do subgroup analyses
• More flexible analysis of outcomes
• Carry out detailed data checking
• Ensure quality of randomisation and follow up
• Ensure appropriateness of analysis
• Update follow up information

Other IPD Benefits

• More complete identification of trials
• Better compliance in providing missing data
• More balanced interpretation of results
• Wider endorsement and dissemination of results
• Better clarification of further research
• Collaboration on further research

Systematic reviews reduce bias

1. Conducting a comprehensive search of relevant literature
2. Using unambiguous inclusion and exclusion criteria for studies
3. Summarizing findings using explicit statistical methodologies

Bias in Meta-analysis

Publication Bias

Publication Bias

Publication Bias

Funnel Plot: what and how to read
What model?

- Fixed effects versus random effects
- Within versus between study variability
- Random effects gives more weight to small studies
- Will lead to larger confidence intervals

Lord Rayleigh, 1884

> “If, as is sometimes supposed, science consisted in nothing but the laborious accumulation of facts, it would soon come to a standstill, crushed, as it were, under its own weight...The work which deserves, but I am afraid does not always receive, the most credit is that in which discovery and explanation go hand in hand, in which not only are new facts presented, but their relation to old ones is pointed out.”

Premise

The results of a particular research study cannot be interpreted with any confidence unless they have been considered, systematically, together with the results of other studies addressing the same or similar questions.

Real life

- How well is this premise reflected in papers published in major general medical journals?

Classification of Discussion sections in RCT reports published in *Annals, BMJ, JAMA, Lancet, and NEJM*

<table>
<thead>
<tr>
<th>First trial addressing the question</th>
<th>May '97 n=26</th>
<th>May '01 n=33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contains an updated systematic review integrating the new results</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Discussed a previous review but did not attempt to integrate new results</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>No apparent systematic attempt to set new results in context of other trials</td>
<td>19</td>
<td>19</td>
</tr>
</tbody>
</table>

Classification of Discussion sections in RCT reports published in *Annals, BMJ, JAMA, Lancet, and NEJM*

<table>
<thead>
<tr>
<th>First trial addressing the question</th>
<th>May '97 n=26</th>
<th>May '01 n=33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contains an updated systematic review integrating the new results</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Discussed a previous review but did not attempt to integrate new results</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>No apparent systematic attempt to set new results in context of other trials</td>
<td>19</td>
<td>27</td>
</tr>
</tbody>
</table>
A contemporary example

- A contemporary example of the consequences of not taking research synthesis seriously

**Drugs to prevent heart rhythm abnormalities (arrhythmias) after heart attacks**

The theory

- Patients with arrhythmias are at increased risk of early death after heart attack
- Anti-arrhythmic drugs reduce arrhythmias after heart attack
- These drugs should reduce early death after heart attack

The consequence:

- At the peak of their use in the late 1980s, it has been estimated that anti-arrhythmic drugs were causing every year comparable numbers of deaths to the total number of Americans who died in the Vietnam war (Moore 1995).

The evidence

- A 1993 systematic review of 51 randomized trials of anti-arrhythmic drugs in heart attack
- 660 deaths among 11,712 patients allocated drugs
- 571 deaths among 11,517 patients allocated to control

Teo et al. JAMA 1993.

The benefit

- The discovery that these drugs are lethal could have been made a decade earlier if the Discussion sections in each report of a new trial had set the new results in the context of a systematic review of the results of all previous trials – in other words, if scientists had cumulated evidence scientifically.

Planning

- Setting the results of new results in the context of a systematic review of the results of all other relevant studies would become straightforward if systematic reviews were always done before embarking on new research.
### Critical appraisal (general)
- Are the results valid?
- What are the results?
- Will they help me care for my patients?

### Critical appraisal of a meta-analysis (specifics)
- Selection bias
- Generalizability
- Combinability
- Consistency
- Statistics
- Sensitivity analysis

### Meta analysis questions about validity
- Bias free search & inclusion criteria? What is being compared?
- Are studies comparable?
- Appraisal of methodology of primary studies?
- Consistent results from all primary studies?
  - if not, are the differences sensibly explained?
- Are the conclusions supported by the data?

### Other validity issues
- Transparent declaration of funding of work?
- Who employs the authors?
  - Drug Company sponsorship of Reviews vs. Cochrane review
- Open discussion of existing controversy & commercial gain?

### If we believe it — does it apply to our patient?
- Is our patient (or population) so different from those in the primary studies that the results may not apply?
- consider differences in:
  - time — many things change.
  - culture — both treatments and values of outcomes can be different
  - stage of illness or prevalence can effect results.

### We believe it! but does it matter?
- Is the benefit worthwhile to our patient?
- Ask the patient about cultural values.
- Think about Relative Risk Reduction vs. Absolute Risk to our patient.
- Potential benefit is the Absolute risk avoided in our patient = Absolute Risk Reduction (ARR)!
Adding Heparin to Aspirin Reduces the Incidence of Myocardial Infarction and Death in Patients With Unstable Angina: A Meta-analysis

[Review]

Oler, Allison, MD; Whooley, Mary A., MD; Oler, Jacqueline, PhD; Grady, Deborah, MD, MPH

Results during therapy

<table>
<thead>
<tr>
<th>Source</th>
<th>Myocardial Infarction or Death, No. (%)</th>
<th>Aspirin</th>
<th>Aspirin Plus Heparin</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penrose et al., 1983[1]</td>
<td>4112 (3)</td>
<td>9/102 (2)</td>
<td>0.32 (0.18-0.55)</td>
<td></td>
</tr>
<tr>
<td>ESC Group, 1980[2]</td>
<td>71/59 (12)</td>
<td>9/266 (3)</td>
<td>0.36 (0.19-0.67)</td>
<td></td>
</tr>
<tr>
<td>Cohen et al., 1990[3]</td>
<td>112 (3)</td>
<td>0/107 (0)</td>
<td>0.29 (0.06-1.45)</td>
<td></td>
</tr>
<tr>
<td>Cohen et al., 1994[4]</td>
<td>5/100 (5)</td>
<td>4/105 (4)</td>
<td>0.46 (0.24-1.45)</td>
<td></td>
</tr>
<tr>
<td>Singleton et al., 1994[5]</td>
<td>40/130 (31)</td>
<td>62/554 (11)</td>
<td>0.36 (0.26-1.39)</td>
<td></td>
</tr>
<tr>
<td>Surkinkel et al., 1995[6]</td>
<td>7/73 (10)</td>
<td>6/97 (6)</td>
<td>0.80 (0.39-1.95)</td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td>65/555 (11)</td>
<td>59/608 (10)</td>
<td>0.97 (0.44-2.02)</td>
<td></td>
</tr>
</tbody>
</table>

Results at 2 to 12 weeks

<table>
<thead>
<tr>
<th>Source</th>
<th>RR (95% CI)</th>
<th>RR (95% CI)</th>
<th>RR (95% CI)</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penrose et al., 1983[1]</td>
<td>1.95 (0.91-1.41)</td>
<td>0.97 (0.41-1.01)</td>
<td>0.95 (0.40-2.10)</td>
<td></td>
</tr>
<tr>
<td>ESC Group, 1980[2]</td>
<td>1.76 (0.50-5.26)</td>
<td>NA</td>
<td>NA</td>
<td>0.77 (0.40-1.16)</td>
</tr>
<tr>
<td>Cohen et al., 1990[3]</td>
<td>1.59 (1.24-1.95)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Cohen et al., 1994[4]</td>
<td>0.78 (0.47-1.31)</td>
<td>0.78 (0.41-1.51)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Singleton et al., 1994[5]</td>
<td>1.65 (1.25-2.16)</td>
<td>0.78 (0.41-1.51)</td>
<td>0.97 (0.72-1.36)</td>
<td></td>
</tr>
<tr>
<td>Surkinkel et al., 1995[6]</td>
<td>1.65 (1.15-2.36)</td>
<td>0.81 (0.42-2.06)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Summary</td>
<td>0.80 (0.60-1.09)</td>
<td>1.02 (0.84-1.26)</td>
<td>0.62 (0.56-1.25)</td>
<td></td>
</tr>
</tbody>
</table>

Areas of concern

- Inappropriate statistical conclusions
- Most studies unblinded
- Composite endpoint, e.g. small MI = death
- Partial exclusion of patients and studies

Their conclusions

“Our findings are consistent with a 33% reduction in risk of MI or death in patients with unstable angina treated with aspirin plus heparin compared with those treated with aspirin alone.

The bulk of evidence suggests that most patients with unstable angina should be treated with both heparin and aspirin.”
My conclusions

- Data is also consistent with no additional benefit of heparin over ASA
- Conclusions are based on a difference of 13 events
- No effect on deaths, almost all events are nonfatal MI (big & small)
- Benefit of heparin is uncertain, risk of bleeding may not be worth the benefit of reduced MI in some patients

Example - Meta-analysis

Meta-analysis - conclusion

Interpretation: Glycoprotein IIb/IIIa inhibitors reduce the occurrence of death or myocardial infarction in patients with acute coronary syndromes not routinely scheduled for early revascularisation. The event reduction is greatest in patients at high risk of thrombotic complications. Treatment with a glycoprotein IIb/IIIa inhibitor might therefore be considered especially in such patients early after admission, and continued until a decision about early coronary revascularisation has been made.


Meta-analysis - results

Meta-analysis – The studies

Articles

Platelet glycoprotein IIb/IIIa inhibitors in acute coronary syndromes: a meta-analysis of all major randomised clinical trials


**Meta-analysis - whose method?**

<table>
<thead>
<tr>
<th>Study</th>
<th>Model</th>
<th>Time</th>
<th>Combined risk of death or MI (OR 95% CI)</th>
<th>Death OR (95% CI)</th>
<th>MI OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boersma</td>
<td>Fixed</td>
<td>30</td>
<td>0.91 (0.85-0.98)</td>
<td>0.91 (0.81-1.03)</td>
<td>0.92 (0.85-1.00)</td>
</tr>
<tr>
<td>Brophy</td>
<td>Fixed</td>
<td>30</td>
<td>0.90 (0.83-0.98)</td>
<td>0.92 (0.72-1.19)</td>
<td>0.93 (0.77-1.13)</td>
</tr>
<tr>
<td>Brophy</td>
<td>Random</td>
<td>30</td>
<td>0.91 (0.80-1.04)</td>
<td>0.93 (0.70-1.24)</td>
<td>0.91 (0.70-1.19)</td>
</tr>
</tbody>
</table>

**Meta-analysis - results**

Largest in high-risk patients. An unexpected and significant interaction was seen between sex and allocated treatment, with a treatment benefit in men (0.81 [0.75-0.89]) but not in women (1.15 [1.01-1.30]). However, once patients were stratified according to troponin concentration, there was no evidence of a sex difference in treatment response, and a risk reduction was seen in men and women with raised troponin.

**Cochrane review**

- Review of clozapine in schizophrenia: 29 RCTs in 2490 participants
- Odds of relapsing in favour of clozapine in the sponsored trials (odds ratio 0.5 (95% confidence interval 0.3 to 0.7); 13 trials, 980 patients). Non-sponsored studies reported equivocal findings (odds ratio 0.4 (0.1 to 1.4); 10 trials, 783 patients).
- Conclusion: beware of sponsored trials

**The difference between evidence and inference**

- Minimum Bayes factor = $e^{-2.22}$ where z is the number of std errors from the null effect
- $Z = 0.25 / 1.7 = 1.47$
- Bayes factor = .33
  - Prior Odds of Null Hypothesis $: \text{Bayes} = \text{Posterior Odds of Null Hypothesis}$
- If prior probability = .5 means odds = 1/1
- Posterior odds = (1/1) * (1/3) = 1/3
- Posterior prob = odds (1+odds) = (1/3) / (1+1/3) = 25% chance that sponsored trials have more favorable results

**Even the Cochrane review can be wrong!!!**

- To claim differences in subgroups on the basis of their P values is known to be flawed.
- A formal assessment would entail calculating the ratio of the odds ratios (0.5/0.4=1.25) and constructing a 95% confidence interval for this quantity (approximately 0.3 to 5.0).
- So the evidence from this study is weak to support the authors’ conclusions
The difference between evidence and inference

- Want post probability = .05
- Or posterior odds = 1/19
- Prior odds = (1/19) / (1/3) = 3/19
- Prior prob = odds / (1+odds)
  = (3/19) / (1+3/19) = 14%

Conclusions

- If you initially believe equal chance of both sponsored and non-sponsored trials give favorable results (Ho = .5), then this study decreases Ho to only 25%
- However, if you initially believed that there was only a 15% probability of sponsored trials not giving more favorable results then this new data is strong enough to conclude at 95% probability that sponsored trials do report more favorable results.
- So maybe the Cochrane group was right but for the wrong reasons!

Lord Rayleigh 1884

- “..... The work which deserves, but I am afraid does not always receive, the most credit is that in which discovery and explanation go hand in hand, in which not only are new facts presented, but their relation to old ones is pointed out.”

Conclusion

- Science is cumulative, and scientists must cumulate scientifically!

As the main funder of science, the public has a right to expect that this will be reflected more effectively in the way that science is conducted and reported.

It remains unclear how effectively academia will respond to this fundamental challenge to its traditional ways of working.

Clinical Trials - Interim analysis

- early - mid 1980’s - ACE shown to decrease mortality in patients with CHF
- late 1980’s - trials to test ACE in patients following AMI
- 1987 - enrollment begins SAVE
- 1989 - enrollment begins AIRE
- 1990 - enrollment begins TRACE
- Sept 1992 - SAVE results published NEJM
- August 1993 - TRACE interim analysis
- Oct 1993 - AIRE results published LANCET
- August 1994 - end follow-up TRACE
- Dec 1995 - TRACE results published NEJM
Clinical Trials - Interim analysis

TRACE interim
Dead Alive
ACE 184 692
Placebo 227 646

21% mortality ACE
26% placebo
20% reduction (p=0.03)

Clinical Trials - Interim analysis

- Incorporating SAVE & AIRE as our prior knowledge and updating this by the interim TRACE results using Bayes Theorem reveals that the best estimate for the difference in mortality between treatment with ace inhibitors and placebo is 4.9 lives saved per 100 patients treated (95% credible interval 3 to 7 lives saved per 100 patients)

Lancet 1997

Bayesian interim statistical analysis of randomised trials
James M Brophy, Lawrence Joseph

Bayesian interim statistical analysis of randomised trials: the case against
L Kober, C Torp-Pedersen, D Cole, J R Hampton, A J Camm

Lancet 2000

Long-term ACE-inhibitor therapy in patients with heart failure or left-ventricular dysfunction: a systematic overview of data from individual patients

Marcus D Flather, Salim Yusuf, Lars Kober, Marc Pfeffer, Alistair Hall, Gordon Murray, Christian Torp-Pedersen, Stephen Ball, Janice Pogue, Lemuel Moyé, Eugene Braunwald, for the ACE-Inhibitor Myocardial Infarction Collaborative Group