Critical Appraisal
Dave Abbott
Senior Medicines Information Pharmacist
Aims

• Identify key components of clinical trial design and apply these to a critical appraisal of the literature
• Be able to work out measures of effectiveness
In this noninferiority trial, we randomly assigned 18,113 patients who had atrial fibrillation and a risk of stroke to receive, in a blinded fashion, fixed doses of dabigatran - 110 mg or 150 mg twice daily - or, in an unblinded fashion, adjusted-dose warfarin. The median duration of the follow-up period was 2.0 years. The primary outcome was stroke or systemic embolism.

Rates of the primary outcome were 1.69% per year in the warfarin group, as compared with 1.53% per year in the group that received 110 mg of dabigatran (relative risk with dabigatran, 0.91; 95% confidence interval [CI], 0.74 to 1.11; P<0.001 for noninferiority) and 1.11% per year in the group that received 150 mg of dabigatran (relative risk, 0.66; 95% CI, 0.53 to 0.82; P<0.001 for superiority). The rate of major bleeding was 3.36% per year in the warfarin group, as compared with 2.71% per year in the group receiving 110 mg of dabigatran (P = 0.003) and 3.11% per year in the group receiving 150 mg of dabigatran (P = 0.31). The rate of hemorrhagic stroke was 0.38% per year in the warfarin group, as compared with 0.12% per year with 110 mg of dabigatran (P<0.001) and 0.10% per year with 150 mg of dabigatran (P<0.001). The mortality rate was 4.13% per year in the warfarin group, as compared with 3.75% per year with 110 mg of dabigatran (P = 0.13) and 3.64% per year with 150 mg of dabigatran (P = 0.051).
What is Evidence Based Medicine?

‘The conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients’

What is Critical Appraisal?

Systematic examination of evidence to assess its validity and relevance
Why do Critical Appraisal?

- Just because something is published doesn’t make it valid
- If the data is valid within the confines of the trial protocol, is it applicable to your patients?
- If the data is applicable to your patients, is it in a readily-understandable form for your audience?
How to critically appraise the evidence

- CASP tools very useful
  Different tools for different types of trials

- BMJ publishing – How to Read a Paper (Trisha Greenhalgh)
Types of study

Randomised Controlled Trial (RCT)
Cohort
Case-control
Cross-over
Meta-analysis

Case studies / case series

Discussion:
What do the above types of study actually mean? What type do you think is ideal? Is this true / feasible for all situations?
RCTs

Aspects that need critical appraisal
- Treatment selection and comparators
- Patient selection
- Treatment allocation
- Treatment protocol
- Data collection
- Data analysis
- Conclusions
RCTs – appraisal of trial design

Bias

Discussion:

• *Where can bias come from?*
• *How can unintentional bias occur?*
• *How can bias be minimised?*
Treatment Selection

What is the study treatment being compared against?

Atorvastatin 80mg vs …
Targinact (oxycodone + naloxone) vs …

Comparing against gold standard, or against something that will make the new treatment look good?
Patient selection and treatment allocation

- Are the group of patients recruited particularly unwell / healthy?
- Do they match your population? Exclusion criteria
- Are they allocated to treatment / control randomly? Stratified (random ≠ equal)
- Is the allocation truly random? Concealed allocation
- Are there enough patients? Power calculation
Treatment protocol / data collection

• Is the study open-label / blind / double-blind? Is this appropriate?
  • Balance between ideal and feasible

• Is the study length appropriate?
  • Balance between ideal and feasible

• How are patients followed up?

• How many patients make it to the final analysis? Is this reasonable?
Patient follow-up

Where did the patients who don’t make it to the final analysis go and why?

Moved house?
Adverse effects?
Ineffective?

How is this factored into the final analysis?
Different ways of managing loss to follow-up

Intention-to-treat (ITT) / modified ITT

Include everyone recruited
Fairly true to use in the wild

Per protocol (PP)

Include everyone who completes the trial as planned
Possibility of bias: only assessing highly motivated pill-takers / treatment successes (need to know why people left the trial early)
Different ways of managing loss to follow-up (continued)

Last observation carried forward (LOCF)

Useful for static conditions
If response is changeable over time (response to anti-Parkinson’s treatments?) may over-estimate effects
What are we proving?

- Superiority?
  - Null hypothesis – no difference – if disprove, new treatment is ‘better’

- Non-inferiority?
  - If new treatment is no worse than a specified margin ($\Delta$) then new treatment is non-inferior

- Equivalence?
What are we measuring?

Patient orientated outcomes
  Something the patient will notice (death, stroke etc)

Disease orientated outcomes
  Measurement of the disease (PSA etc)
What outcomes are we measuring?

Hard / soft?
Proxy?
Subjective?
Recognised measure?

<table>
<thead>
<tr>
<th>Stroke</th>
<th>Cholesterol level</th>
<th>QoL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood pressure</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PSA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MI</td>
</tr>
</tbody>
</table>
Primary / secondary outcomes

Primary outcomes

   The main focus of the trial
   What the power calculation tends to be focused on

Secondary outcomes

   Subsidiary measure
   Trial may not be powered to detect differences
Subgroups

Take care with subgroup analyses

Trials often not designed to reliably investigate sub-groups, particularly if not specified in advance (post-hoc analysis)
Data presentation

There are many different ways to represent data. All have particular advantages / uses and disadvantages

Odds
Relative Risk
Absolute Risk
Number needed to treat
Odds

Odds

Odds of something happening

Number of people something happened to
Number of people something didn’t happen to

Odds ratio

Odds in treatment group / odds in control group
Odds example

Treatment group: 24/100 had a DVT
Control group: 31/100 had a DVT

Odds of event in treatment group: \( \frac{24}{100-24} = 0.32 \)
Odds of event in control group: \( \frac{31}{100-31} = 0.45 \)

Odds ratio = \( \frac{0.32}{0.45} = 0.71 \)

When events are rare, odds are similar to risk
Absolute Risk

Absolute risk

Percentage of people where something happens

Number of people something happened to
Total number of people you looked at

Absolute risk reduction

Difference in absolute risk between treatment and control groups
Absolute risk example

Treatment group: 24/100 had a DVT
Control group: 31/100 had a DVT

Absolute risk in treatment group: \( \frac{24}{100} = 0.24 \) (24%)
Absolute risk in control group: \( \frac{31}{100} = 0.31 \) (31%)

Absolute risk reduction = 0.31 – 0.24 = 0.07 (7%)
Relative risk

Relative risk: absolute risk in treatment group expressed relative to control group risk

\[
\frac{\text{AR in treatment group}}{\text{AR in control group}}
\]

Relative risk reduction: ARR expressed relative to control group risk

\[
\frac{\text{AR in control group} - \text{AR in treatment group}}{\text{AR in control group}}
\]
Relative risk reduction example

Treatment group: 24/100 had a DVT
Control group: 31/100 had a DVT

Absolute risk in treatment group: 24/100 = 0.24
Absolute risk in control group: 31/100 = 0.31

Relative risk reduction = (0.31 – 0.24)/0.31 = 0.22 (22%)
Number Needed to Treat

Number needed to treat:

Number of people you need to treat for one person to avoid an event

1/ARR

Need to include the time period
NNT example

Treatment group: 24/100 had a DVT
Control group: 31/100 had a DVT

Absolute risk in treatment group: 24/100 = 0.24
Absolute risk in control group: 31/100 = 0.31

Absolute risk reduction = (0.31 – 0.24) = 0.07 (7%)

NNT = 1/0.07 ~ 15

You would need to treat 15 people with drug (x) instead of drug (y) for (z) years for 1 person to not have a DVT
Whole picture?

If we treated 100 people with drug (y) instead of drug (x), what is likely to happen?

69 people who wouldn’t have had a DVT on drug (x) anyway will not have a DVT on drug (y)
24 people who would have had a DVT on drug (x) will have a DVT on drug (y) as well
7 people who would have had a DVT on drug (x) will avoid that DVT on drug (y)
A new anticoagulant, shinyboxagatran, has been brought to market. In a trial (2 years) of high risk patients, 97 / 2,432 on shinyboxagatran had a stroke or thrombotic event, compared to 131 / 2,629 on cheaparmin (current gold standard treatment)

In groups, derive a (true) stat that fits your world view:

Group 1: Lead Pharmacist – product enthusiast
Group 2: Lead Pharmacist – cynicism
Group 3: Lead Pharmacist – patient understanding
Graphs

Number of people suffering a stroke

Control
Active

% Number of people suffering a stroke

53
53.5
54
54.5
55
55.5
56
56.5

The Leeds Teaching Hospitals
NHS Trust
Any trial only looks at a sample of the population, and we use this to estimate what would happen in the whole population. There is therefore a chance that the sample of the population we looked at were not representative. The p value is the probability that any difference seen between treatment and control groups was just by chance. The 95% confidence interval is the range of values that you are 95% sure the population value lies between.
Significance

The range you are 95% sure the population’s value lies in

Old treatment is better

Treatments are the same

What you found in your sample

New treatment is better

P value: the probability that the treatments are actually the same, and what you saw in your sample arose just from chance
Confidence Intervals

Treatments are the same

a.  

b.  

c.  

d.  

Old treatment is better

New treatment is better

Treatments are the same
Quick test:

OR 0.7 (95% CI 0.64 – 0.76, p<0.01)

The information above means that the OR for your sample is 0.7, and that you are 95% sure that the population’s OR is between 0.64 and 0.76. The p value of <0.01 means that there is a less than 1% chance that the difference between treatment and control groups occurred by chance (the smaller the value, the more confident you are in the result).
Statistical significance vs clinical significance

New antihypertensive drug, reduces blood pressure by 2mmHg, p<0.001

Statistically significant, but will the patient notice?
Relate the findings to your patient cohort (comes back to patient orientated outcomes vs disease orientated outcomes)
A final thing to think about…

How did this paper get picked for publication?

If you were a drug company, would you push for publication if a trial showed your drug was worse?

If you were an editor of a journal, would you prefer to publish a paper that showed an amazing leap forward, or one that showed that the new product is about the same / worse than the old one?

Why did you pick this paper?

Is it the only one you could get access to, or did you do a comprehensive literature review?
Apply skills to RELY

Draw the information covered in this session to critically appraise the randomised controlled trial
If this trial is reflective of the effectiveness of dabigatran, should we use dabigatran in our patients?

Efficacy?
Safety?
Patient benefits?
Cost?