Session 5: How to critically appraise a paper – an example
Appraising research – broad questions

- What is the research question?
- Is the study ethical?
- Is the study design valid and appropriate?
- What are the results?
- What are the implications of the findings for clinical practice?
Mortality after Fluid Bolus in African Children with Severe Infection

Kathryn Maitland, M.B., B.S., Ph.D., Sarah Kiguli, M.B., Ch.B., M.Med.,
Peter Olupot-Olupot, M.B., Ch.B., Samuel O. Akech, M.B., Ch.B.,
Richard Nyeko, M.B., Ch.B., M.Med., George Mtove, M.D., Hugh Reyburn, M.B., B.S.,
Trudie Lang, Ph.D., Bernadette Brent, M.B., B.S., Jennifer A. Evans, M.B., B.S.,
James K. Tibenderana, M.B., Ch.B., Ph.D., Jane Crawley, M.B., B.S., M.D.,
and Diana M. Gibb, M.B., Ch.B., M.D., for the FEAST Trial Group*

ABSTRACT

BACKGROUND
The role of fluid resuscitation in the treatment of children with shock and life-threatening infections who live in resource-limited settings is not established.

Maitland K et al. NEJM (2011); 364:2483-95.
Appraising research – broad questions

- What is the research question?
- Is the study ethical?
- Is the study design valid and appropriate?
- What are the results?
- What are the implications of the findings for clinical practice?
The research question – questions to ask

1. What is the scientific rationale?
2. Is there a pre-defined hypothesis?
3. Where was the study conducted?
4. What was the study population?
5. What intervention(s) were compared?
6. What outcomes were considered?
The research question – questions to ask

1. What is the scientific rationale?

“The role of fluid resuscitation in the treatment of children with shock and life-threatening infections who live in resource-limited settings is not established.”

“Although the approach is endorsed by pediatric life-support training programs...evidence regarding the criteria for intervention and the volume and type of fluid is lacking.”

2. Is there a pre-defined hypothesis?

Rapid, early fluid resuscitation will lead to improvements in the outcomes of children with shock.

Maitland K et al. *NEJM* (2011); 364:2483-95.
The research question – questions to ask

3. Where was the study conducted?
   Kenya, Tanzania and Uganda

4. What was the study population?
   Children aged 60 days – 12 years of age with shock

5. What intervention(s) were compared?
   Albumin bolus, saline bolus or no bolus (control)

6. What outcomes were considered?
   48-hour mortality; pulmonary edema; increased intracranial pressure; mortality or neurologic sequelae at 4 weeks

Maitland K et al. NEJM (2011); 364:2483-95.
Appraising research – broad questions

- What is the research question?
- Is the study ethical?
- Is the study design valid and appropriate?
- What are the results?
- What are the implications of the findings for clinical practice?
Ethical issues – questions to ask

1. Was there a real need to conduct the trial? Was there equipoise when the trial was initiated?
2. Was the study approved by an ethics committee?
3. Did patients (or carers) give informed consent?
4. Was the trial of sufficient size to reach a meaningful conclusion?
5. Was there an independent DSMB?
6. What role did the funder(s) play?
Ethical issues – informed consent

- All patients should give informed consent prior to participating in any RCT
- Should be sought before the patient is recruited into the trial and before the result of any randomisation is known
- In particular, knowledge of what treatment the patient is to receive should not have any impact on whether s/he is recruited into the trial
### Study power

**Power** of a study - probability of correctly detecting a difference between the arms of a trial when one really exists

<table>
<thead>
<tr>
<th></th>
<th>After carrying out RCT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P&lt;0.05$ Conclude that new regimen is different</td>
</tr>
<tr>
<td>New regimen really is different to existing regimen</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>POWER $(1-\beta)$</td>
</tr>
<tr>
<td>New regimen is no different to existing regimen</td>
<td>✗</td>
</tr>
<tr>
<td></td>
<td>Type I error ($\alpha$)</td>
</tr>
</tbody>
</table>
Study power

- The power of the trial will depend on:
  - outcome of interest (binary or continuous),
  - minimum treatment effect felt to be clinically important
  - threshold for statistical significance (usually 0.05)
  - size of the trial (small trial = low power)

- Small underpowered trials are unethical

- Optimal sample size must be calculated at the start of the trial; details of this must be reported in protocol and any manuscripts
Ethical issues – questions to ask

1. Was there a real need to conduct the trial? Was there equipoise when the trial was initiated?
   Apparently so.

2. Was the study approved by an ethics committee?
   Yes - Imperial College (London), Makerere University (Uganda), Medical Research Institute (Kenya) and National Medical Research Institute (Tanzania).

3. Did patients (or carers) give informed consent?
   “In cases in which prior written consent from parents and guardians could not be obtained, provision was made for oral assent from a legal surrogate, followed by delayed written informed consent as soon as practicable.”

Maitland K et al. NEJM (2011); 364:2483-95.
4. Was the trial of sufficient size to reach a meaningful conclusion?

“We estimated that with a sample size of 3600 children, the study would have 80% power to detect a 33% relative reduction in mortality with a saline bolus as compared with the control group and a 40% reduction with an albumin bolus as compared with a saline bolus, assuming a risk of death of 11% in the control group, at a two-sided alpha level of 0.05, adjusted for two comparisons with the use of a nominal alpha of 0.025.”

Maitland K et al. *NEJM* (2011); 364:2483-95.
5. Was there an independent DSMB?

“An independent data and safety monitoring committee reviewed the interim analyses from the study twice a year. The Haybittle-Peto criterion was the statistical guide that the committee used in considering a recommendation to stop or modify the trial…”

6. What role did the funder(s) play?

“Neither of those bodies, nor Imperial College, London... had any role in the design of the study, the collection, analysis, or interpretation of the data, or the writing of the manuscript.”

Maitland K et al. *NEJM* (2011); 364:2483-95.
Appraising research – broad questions

- What is the research question?
- Is the study ethical?
- Is the study design valid and appropriate?
- What are the results?
- What are the implications of the findings for clinical practice?
Study design – questions to ask

1. What is the trial design?
2. What ‘treatments’ are being studied?
3. Who is included in the trial?
4. Was randomisation performed? If so, how and could bias have been introduced?
5. Was the study blinded? If not, is the lack of blinding likely to introduce any bias?
6. Were individuals in the different treatment groups treated equally throughout the trial?
Types of RCTs

- **Superiority trial**: demonstrate that one drug is substantially better than another
- **Equivalence trial**: demonstrate that two drugs have the SAME effect (bio-equivalence studies, different formulations of same drug, etc.)
- **Non-inferiority trial**: demonstrate that new drug is not SUBSTANTIALLY WORSE than standard-of-care
- Equivalence and non-inferiority trials require emphasis on confidence intervals rather than statistical significance
Identifying the study population

- **Study population** - the entire group of patients that you would ideally like to study

- Rarely feasible to study the population – thus, we aim to recruit a **sample** of patients who are **representative** of (ie. have similar characteristics to) the population

- **Inclusion criteria**: define the characteristics of patients who will be eligible for the RCT

- **Exclusion criteria**: identify patients who should not be included, usually because they may be placed at harm by participating
Allocating treatments to patients

- Usually done using randomly generated randomisation lists
- These are concealed from investigators to ensure that a patient’s treatment allocation is unknown when s/he is invited to participate
- After a patient is formally recruited, the list holder (e.g. statistician, pharmacist) will notify the clinician of the treatment that has been allocated
- May use sealed opaque envelopes
1. What is the trial design?
   Multicenter, parallel, superiority trial with three arms.

2. What ‘treatments’ are being studied?
   0.9% saline bolus, 5% human albumin bolus, or no bolus (dose dependent on stratum)

3. Who is included in the trial?
   “Children ... between 60 days and 12 years of age and presented with a severe febrile illness complicated by impaired consciousness (prostration or coma), respiratory distress (increased work of breathing), or both, and with impaired perfusion...”

Maitland K et al. NEJM (2011); 364:2483-95.
4. Was randomisation performed? If so, how and could bias have been introduced?

Permuted blocks (random sizes) stratified by clinical centre. Trial statistician generated randomization lists; numbers kept in opaque, sealed envelopes, which were opened in numerical order by study clinician.

5. Was the study blinded? If not, is the lack of blinding likely to introduce any bias?

No - unlikely to introduce bias given objective endpoint.

6. Were individuals in the different treatment groups treated equally throughout the trial?

It appears so.

Maitland K et al. NEJM (2011); 364:2483-95.
Appraising research – broad questions

- What is the research question?
- Is the study ethical?
- Is the study design valid and appropriate?
- What are the results?
- What are the implications of the findings for clinical practice?
1. How many patients were included in the trial, compared to the planned trial size?

2. Did randomisation work?

3. Did any patients drop out of the trial – if so, did drop-out rates differ by treatment arm and what analytical approach was used to allow for loss-to-follow-up and treatment switches?
Describing patient flow

Maitland K et al. NEJM (2011); 364:2483-95.
Did randomisation work?

- Were the two groups broadly similar at the start of the trial? (usually shown in Table 1)
- Not appropriate to formally test whether differences are significant - assuming study isn’t fraudulent, any differences must be due to chance, and this approach is likely to lead to high Type I error rate
- Existing statistical methods can deal with small imbalances between groups; if large imbalances exist, then multivariable methods should be used
Maitland trial

- DSMB stopped trial early at fifth interim review owing to safety concerns in saline-bolus and albumin-bolus groups, and because it was very unlikely that superiority of the bolus strategy over the control strategy would be shown.

- Data were available from 2995 children at study discontinuation (planned numbers around 3600).

Maitland K et al. *NEJM* (2011); 364:2483-95.
Results – questions to ask (2)

1. What was the primary endpoint of the trial and what was the treatment effect for this endpoint? How precise was this effect?

2. What were the main results from the secondary endpoints?
Trial endpoints (1)

- Generally have one primary endpoint (usually efficacy-related) and several secondary endpoints (efficacy and/or safety)
  - Major decisions based on results from primary endpoint;
  - Secondary endpoints provide supportive data

- Endpoints should:
  - be defined in advance
  - should capture ‘most important’ aspects of disease progression
  - be clinically relevant
  - be appropriate for population in the trial
Estimating the treatment effect

- Should do more than simply summarise the outcomes in the different treatment arms
- Estimate the treatment effect – additional benefit provided by new drug over-and-above that which would have occurred using standard-of-care or by chance (in a placebo-controlled trial)
Estimating the treatment effect

<table>
<thead>
<tr>
<th>Arm</th>
<th>N (%) dying</th>
<th>N (%) not dying</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin bolus</td>
<td>111 (10.6)</td>
<td>939 (89.4)</td>
<td>1050</td>
</tr>
<tr>
<td>Saline bolus</td>
<td>110 (10.5)</td>
<td>937 (89.5)</td>
<td>1047</td>
</tr>
<tr>
<td>No bolus</td>
<td>76 (7.3)</td>
<td>968 (92.7)</td>
<td>1044</td>
</tr>
</tbody>
</table>

- Absolute risk reduction (saline vs. no bolus) = 3.2% (110/1047 – 76/1044)
- For every 100 children receiving a saline bolus, we would expect an additional 3.2 deaths at 48-hours, compared to the number that we would have expected had none of the children received a bolus

Maitland K et al. *NEJM* (2011); 364:2483-95.
Alternative measures of treatment effect

- Absolute risk reduction (saline vs. no bolus) = 3.2% \((110/1047 - 76/1044)\)
- Relative risk (saline vs. no bolus) = 1.44 \((110/1047 \div 76/1044)\)
- Odds ratio (saline vs. no bolus) = 1.50 \((110/937 \div 76/968)\)

<table>
<thead>
<tr>
<th>Arm</th>
<th>N (%) dying</th>
<th>N (%) not dying</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin bolus</td>
<td>111 (10.6)</td>
<td>939 (89.4)</td>
<td>1050</td>
</tr>
<tr>
<td>Saline bolus</td>
<td>110 (10.5)</td>
<td>937 (89.5)</td>
<td>1047</td>
</tr>
<tr>
<td>No bolus</td>
<td>76 (7.3)</td>
<td>968 (92.7)</td>
<td>1044</td>
</tr>
</tbody>
</table>

Maitland K \textit{et al.} \textit{NEJM} (2011); 364:2483-95.
Describing the precision of the effect

- Should provide some evidence of the precision of the treatment effect
  - RR 1.44, 95% CI (1.09, 1.90)

- Best estimate of relative risk is 1.44 (44% increase in risk), but true value could range from 1.09 (9% increase) to 1.90 (90% increase)

- Allows us to put the results from trial into clinical context; can weigh up benefits in light of any disadvantages of drug (e.g. increased cost or worse toxicity profile)

Maitland K et al. NEJM (2011); 364:2483-95.
Number needed to treat (NNT)

- Number of people who need to be treated with intervention over period of study to prevent/cause one adverse event

\[ NNT = \frac{100}{\text{absolute risk reduction} \%} = \frac{100}{3.2} = 31.25 \]

- For every 32 children who receive a saline bolus, there will be one additional death within 48 hours

- Useful for comparing effectiveness of treatments

Maitland K et al. NEJM (2011); 364:2483-95.
Appraising research – broad questions

- What is the research question?
- Is the study ethical?
- Is the study design valid and appropriate?
- What are the results?
- What are the implications of the findings for clinical practice?
What are the implications for clinical practice?

- Generalisability of population and setting and intervention
- Consideration of all important outcomes (reflecting effectiveness, safety, acceptability of treatment)
- Benefits and costs
- Application in clinical practice (patients’ treatment preferences, other clinical considerations)
Summary

- To maximise the benefits of the trial design, important that RCTs are designed and analysed appropriately.
- Treatment effects/confidence intervals should be provided as well as some discussion of clinical impact of findings.
- Particular care should be taken when interpreting RCTs that include many endpoints, or that involve large numbers of (often post-hoc) subgroup analyses.
Helping you out - RAMMbo

- Recruitment
  - Appropriate spectrum of patients, representative of target population, inclusion/exclusion criteria

- Allocation (randomisation)

- Maintenance
  - Were all patients treated equally except for intervention, was allocation concealed?

- Measurements
  - blinded
  - objective