Lecture 4: Data analysis
Overview

Interim analyses

Final - Descriptive analysis

Participation – the Consort diagram (Figure 1)

Study participants (Table 1)

Primary analysis (*reminder superiority vs non-inferiority*)

Effectiveness (Intention to treat)

Modified intention to treat

Efficacy – per protocol

Secondary analyses

Planned and Hypothesis generating
Interim Analysis and Stopping Rules

In large trials interim analyses commonly done.

• Adverse events –
• Primary outcomes -

Can the study be ended early – hypothesis answered.
Or,
Should the study be ended early – patient’s safety.

Must use more stringent rules (p<0.01, not p<0.05)
Usually reviewed by independent panel (DSMB)
Enrolment began in April 1995. By early 1997 four HIV positive patients had relapsed with Rifampin mono resistance among all occurred in those taking once weekly RPT-INH. The DSMB, CDC, and the investigators decided to stop enrolment of HIV positive patients. Those still taking once weekly RPT-INH were switched to standard treatment.

<table>
<thead>
<tr>
<th></th>
<th>Once weekly INH-RPT</th>
<th>Twice weekly INH-RIF</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>30</td>
<td>31</td>
<td>-</td>
</tr>
<tr>
<td>Relapse</td>
<td>5</td>
<td>3</td>
<td>.41</td>
</tr>
<tr>
<td>RIF-R</td>
<td>4</td>
<td>0</td>
<td>.05</td>
</tr>
</tbody>
</table>
Final Analysis: Step 1 – Accounting for all subjects

- The CONSORT statement - JAMA 1996
- (consolidated standards of reporting trials)

- *Revised CONSORT Statement.*
- *Ann Intern Med 2001; vol 134: p666*

- [www.consort-statement.org/](http://www.consort-statement.org/)
Consort diagram – general structure

Potential Eligible Patients Identified n=

Excluded - not eligible n=
Reasons

Refused n=

Included and Randomized n=

Intervention A
n= (ITT)

Post-randomization valid exclusions . Reasons
N= (MITT)

Did not complete Tx
Did not complete F-U

Included in per protocol analysis n=

Intervention B
n= (ITT)

Post-randomization valid exclusions . Reasons
N= (MITT)

Did not complete Tx
Did not complete F-U

Included in per protocol analysis n=
Consort diagram example

Mfinanga LID 2014

13,588 screened and assessed for eligibility

11,913 excluded
  3,612 CD4 cell count out of range
  195 declined to participate
  3,897 HIV negative
  37 died before enrolment
  1,553 HIV test refused or unknown
  1,596 recent tuberculosis or HIV treatment
  447 lost contact or unable to follow-up before randomisation
  24 smear negative at baseline
  552 no reasons given

1,675 enrolled

834 assigned to early ART group

67 excluded
  6 multidrug resistant at baseline
  52 negative tuberculosis culture at baseline
  9 missing

767 analysed

841 assigned to late ART group

70 excluded
  11 multidrug resistant at baseline
  53 negative tuberculosis culture at baseline
  6 missing

771 analysed
Consort diagram example

Moxi Gati
Step 1B: Analysis of non-participants

Subjects who are screened as potential participants, but were not eligible, or refused.

If not randomized do not impact the internal validity of the study.

But affect external validity (capacity to generalize). Especially important if high exclusion or refusal rate.
Step 2: Describing and comparing study participants (Table 1)

- This is a simple descriptive analysis comparing study participants randomized to the different interventions
  - Demographic characteristics (age and sex)
  - Major clinical characteristics (extent of disease, drug resistance)
  - Comorbidities (HIV, Diabetes etc)
- No statistical testing please
**Baseline characteristics – example**

*Swaminathan 2010  varying lengths of treatment in HIV TB*

<table>
<thead>
<tr>
<th>Characteristic of Study Subjects</th>
<th>Reg6M (n=167)</th>
<th>Reg9M (n=160)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age, years (IQR)</td>
<td>33 (29-38)</td>
<td>33 (29-39)</td>
</tr>
<tr>
<td>Median weight, kg (IQR)</td>
<td>44 (39-50)</td>
<td>44 (39-50)</td>
</tr>
<tr>
<td>Median CD4 cells/mm (IQR)</td>
<td>152 (80-304)</td>
<td>167 (88-280)</td>
</tr>
<tr>
<td>Median viral Load, (copies/ml)</td>
<td>94,300 (n=100)</td>
<td>168,000 (n=113)</td>
</tr>
<tr>
<td>Males N %</td>
<td>119 (79%)</td>
<td>112 (75%)</td>
</tr>
</tbody>
</table>

**Pulmonary TB (n=299)**

| Culture Positive                        | 117 (78%)           | 110 (74%)           |
| Susceptible to all first-line drugs,    | 99 (88%)            | 95 (88%)            |
| Culture Negative                        | 34 (22%)            | 38 (26%)            |

**Extrapulmonary TB (n=28)**

| Culture Positive                        | 4 (25%)             | 2 (16%)             |
| Culture Negative                        | 12 (75%)            | 10 (84%)            |
## Baseline characteristics – example
**Moxi and Gati**

<table>
<thead>
<tr>
<th>Patient Characteristics</th>
<th>Regimen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gatifloxacin n=136</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>103 (76%)</td>
</tr>
<tr>
<td>Age (years):</td>
<td></td>
</tr>
<tr>
<td>&lt;40</td>
<td>90 (66%)</td>
</tr>
<tr>
<td>Body weight (Kg):</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>43.7</td>
</tr>
<tr>
<td>Sputum culture</td>
<td></td>
</tr>
<tr>
<td>3+ growth</td>
<td>107 (79%)</td>
</tr>
<tr>
<td>X-ray Chest</td>
<td></td>
</tr>
<tr>
<td>&gt;2 Zones affected</td>
<td>107 (79%)</td>
</tr>
</tbody>
</table>

Did the randomization work?
Step 3: Primary analysis

The primary analysis addresses the primary objective.

Sample size calculations were based on this planned analysis.
Primary analysis

Ideally all randomized participants must be included in the primary analysis.

Withdrawals: Participants who sign consent, and are randomized. But withdraw consent – so ethically not included in the analysis. Can bias the results of the study (the 2 groups of participants remaining may not be comparable)

Drop-outs from therapy: Do not complete therapy, but do complete follow-up post therapy. Contribute fully to analysis

Lost – no idea of final outcome. More difficult
Superiority Studies (reminder)

- Test New Interventions against a standard or placebo.

- Hypothesis: New intervention is better.

- New intervention will be adopted if patients’ outcomes are better.
Superiority studies:
Results: \textbf{CANNOT} conclude superiority
Superiority studies:
Results: **CAN** conclude superiority
Non-inferiority Studies

If current therapy is effective
- But is very costly, or lengthy
- Or has major side effects

Alternate therapies must be cheaper, shorter, or safer.

Then we want to show that the new treatment is not worse.

This is called a Non-inferiority study.
Non-Inferiority studies - Results

**CAN** conclude non-inferiority

- 0: No effect
- 0.7: Least Acceptable Effect of New Therapy
- 1.0: Effect of Standard Therapy
Non-Inferiority studies - Results

CANNOT conclude non-inferiority

- No effect
- Least Acceptable Effect of New Therapy
- Effect of Standard Therapy
Efficacy and Effectiveness

Effectiveness (intention to treat)

The effect of a specific intervention, procedure, regimen, or service, *when deployed in the field in routine circumstances*. This accounts for non-compliance, dropouts and side effects.

All patients randomized (allocated to treatment) are analysed, whether or not they completed the prescribed regimen, and follow-up.

**Conservative estimate:** Answers the public health question “What is the overall effect of this treatment given to a population?”
Efficacy vs Effectiveness

Efficacy (per protocol):

The extent to which a specific intervention, procedure, regimen, or service produces a beneficial result under ideal conditions;

This means the patient actually took all doses of treatment,

And all elements of the protocol followed (ie full follow-up)

Optimal Estimate: Answers the patient’s question “What will this drug do…. if I take it?”
### Duration of INH Therapy and efficacy/effectiveness

**IUAT trial - Patients with Fibrotic Lesions**

<table>
<thead>
<tr>
<th>Population</th>
<th>Duration</th>
<th>Reduction in TB</th>
</tr>
</thead>
<tbody>
<tr>
<td>All participants (Effectiveness)</td>
<td>INH 12 mo.</td>
<td>75%</td>
</tr>
<tr>
<td></td>
<td>INH 6 mo.</td>
<td>65%</td>
</tr>
<tr>
<td></td>
<td>INH 3 mo.</td>
<td>21%</td>
</tr>
<tr>
<td>Completer/compliers (Efficacy)</td>
<td>INH 12 mo.</td>
<td>93%</td>
</tr>
<tr>
<td></td>
<td>INH 6 mo.</td>
<td>69%</td>
</tr>
<tr>
<td></td>
<td>INH 3 mo.</td>
<td>31%</td>
</tr>
</tbody>
</table>

*Bull WHO 1982;555-64*

Why is the difference biggest for those randomized to 12 months?
### ITT and MITT Analyses: example
*Sterling et al; 3HP vs INH; NEJM 2011*

<table>
<thead>
<tr>
<th>Study Group</th>
<th>N</th>
<th>Subjects with Active TB</th>
<th>Difference in Cumulative Rate percentage points</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no.</td>
<td>no. per patient yr</td>
<td>Cumulative rate</td>
</tr>
<tr>
<td><strong>Modified intention-to-treat analysis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isoniazid only</td>
<td>3745</td>
<td>15</td>
<td>0.16</td>
</tr>
<tr>
<td>Combination therapy</td>
<td>3986</td>
<td>7</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>Per-protocol analysis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isoniazid only</td>
<td>2585</td>
<td>8</td>
<td>0.11</td>
</tr>
<tr>
<td>Combination therapy</td>
<td>3273</td>
<td>4</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Mis-use of ITT analysis

The intention of ITT is to produce realistic estimates of what the treatment will achieve in real life.

Many RCT select subjects carefully on the basis of compliance

- Baseline characteristics (lifestyle, employment, etc)
- Run-in period – often 1-3 months to assess compliance

What effect does this have on ITT analysis
Modified Intention to treat Analysis (MITT)

- There may be instances where patients may need to be randomized before all information is available.
- Particularly common in TB trials when eligibility depends upon culture and/or drug susceptibility testing.
- In latent TB trials, household contacts may start L:TBI therapy before knowing the DST of the index cases.
- Protocol may specify valid exclusions post randomization.
- Because LTBI treatment initiation cannot wait
Secondary Analyses: Planned

Many studies pre-specify planned secondary analysis

This should be stated in the published study protocol

- Not all subjects must be included
- Different sub-groups – effect of age or gender
- Different end-points – Adverse events
- Efficacy analysis may be a planned secondary analysis
Planned Primary and Secondary analyses – example

Gler et al, Use of Delanamid for MDR-TB; NEJM, 2013
Planned Primary and Secondary analyses – example
Gler et al Delamanid for MDR TB NEJM 2012

• Primary endpoint – proportion with sputum culture conversion at 2 months – MITT
• Multiple secondary endpoints assessed. These included time to sputum culture conversion
• Safety performed in all patients randomized who received at least one dose of study medication (ITT)
• All endpoints pre-specified in formal statistical analysis plan. Plan finalized and filed before analysis begun.
Consort diagram

611 patients assessed for Eligibility

96 Excluded – not eligible (91) - other (5)

34 Refused

481 Randomized

161 DMD 100 BID
161 in ITT (Safety)

20 Excluded
Negative Culture
Not MDR
141 in MITT
(2 months culture conversion)

18 Excluded
14 withdrew
4 SAE
0 Lost
123 in efficacy
(time to culture conversion)

160 DMD 200 BID
160 in ITT (Safety)

24 Excluded
Negative Culture
Not MDR
136 in MITT
(2 months culture conversion)

14 Excluded
6Withdrawed
6 SAE
2 Lost
122 in Efficacy
(time to culture conversion)

160 Placebo
160 in ITT (Safety)

35 Excluded
Negative Culture
Not MDR
125 in MITT
(2 months culture conversion)

15 Excluded
8 Withdrawed
4 SAE
3 Lost
120 in efficacy
(time to culture conversion)
Planned secondary analysis: Incidence of Adverse Events
Uses ITT population (took at least 1 dose of study drug)

<table>
<thead>
<tr>
<th></th>
<th>Delamanid 100mg Twice Daily (N=161)</th>
<th>Delamanid 200mg Twice Daily (N=161)</th>
<th>Placebo (N=160)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anemia</td>
<td>18(11.2)</td>
<td>10(6.2)</td>
<td>14(8.8)</td>
</tr>
<tr>
<td>Nausea</td>
<td>58(36.0)</td>
<td>65(40.6)</td>
<td>53(33.1)</td>
</tr>
<tr>
<td>Prolonged QT interval on ECG</td>
<td>16(9.9)</td>
<td>21(13.1)</td>
<td>6(3.8)</td>
</tr>
<tr>
<td>Paresthesias</td>
<td>17(10.6)</td>
<td>20(12.5)</td>
<td>12(7.5)</td>
</tr>
<tr>
<td>Anorexia</td>
<td>23(14.3)</td>
<td>34(21.2)</td>
<td>24(15.0)</td>
</tr>
<tr>
<td>Hypokalemia</td>
<td>20(12.4)</td>
<td>31(19.4)</td>
<td>24(15.0)</td>
</tr>
</tbody>
</table>
Primary analysis -
Uses MITT population:
2 Month culture
Conversion on MGIT

Planned secondary
Analysis -
Uses MITT population:
2 mos conversion
on solid media
Planned secondary Analysis – Efficacy: Uses per protocol Population

Time to culture conversion
Planned Primary and Secondary analyses – example

Use of TMC-207 (Bedaquiline) for MDR TB

Diacon et al  NEJM 2009
Bedaquiline for MDR TB Diacon et al  NEJM 2009
Primary Analysis (MITT)
### Secondary Analysis (ITT)
#### Incidence of Adverse Events

<table>
<thead>
<tr>
<th>Adverse Event</th>
<th>TMC207 (N=23)</th>
<th>Placebo (N=24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nausea</td>
<td>6(26)</td>
<td>1(4)</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>3(13)</td>
<td>1(4)</td>
</tr>
<tr>
<td>Arthralgia</td>
<td>4(17)</td>
<td>3(12)</td>
</tr>
<tr>
<td>Rash</td>
<td>2(9)</td>
<td>4(17)</td>
</tr>
<tr>
<td>Dizziness</td>
<td>3(13)</td>
<td>2(8)</td>
</tr>
</tbody>
</table>
Planned Secondary Analyses: (Efficacy)
Rate of bacterial killing (per protocol)

<table>
<thead>
<tr>
<th>No. of Subjects</th>
<th>Placebo</th>
<th>TMC207</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>Wk 1</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Wk 2</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Wk 4</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Wk 6</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>Wk 8</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>
Secondary Analyses: Post hoc (Hypothesis generating)

Hypothesis generating vs data dredging

- Once the primary and planned secondary analyses are done,
- Then many exploratory analyses can be performed

**Risks**- If 20 tests are done, 1 will be significant at \( p < .05 \) by chance alone. Especially if not clearly driven by a priori hypotheses, but rather by a desire for a \( p < .05 \)!!

**Advantages**- RCT generate a wealth of data which can and should be used to address other questions

- but very important to describe these analyses clearly as such.
Post hoc Analyses – example
DMD Improves outcomes and reduces mortality in MDR TB
Skripconoka et al, ERJ 2013
Post hoc Analyses – example
DMD Improves outcomes and reduces mortality in MDR TB  
Skripconoka et al, ERJ 2013

What they wrote in Abstract - Results and Conclusions:

• “Mortality was reduced to 1% on those receiving long-term DMD vs short-terms no DMD (8.3% p>.001)”

• “Treatment benefit was also seen on patients with XDR TB”

• “This analysis suggests that treatment with DMD for 6 months in combination with optimized background regimen can improve outcomes and reduce mortality on patients with both MDR and XTR TB”
Post hoc Analyses – example

DMD Improves outcomes and reduces mortality in MDR TB  
*Skripconoka et al, ERJ 2013*

**Methods:**

- Follow-up study after conclusion of initial 2 month treatment study
- Study launched 2-12 months after end of first study
- Substantial intervals between initial 2 month treatment with DMD, and later treatment
- Patients not randomized. Less than half selected for DMD by provider or by themselves.
### 24 month outcomes after treatment with DMD plus OBR in patients with MDR or XDR

*Skripconoka et al, ERJ 2013*

<table>
<thead>
<tr>
<th>Treatment Outcome</th>
<th>6-8 months DMD N=192</th>
<th>0-2 Months DMD N=229</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cured</td>
<td>110 (57%; 50-64)</td>
<td>111 (48%; 42-55)</td>
</tr>
<tr>
<td>Completed</td>
<td>33 (17%; 12-23)</td>
<td>15 (7%; 4-11)</td>
</tr>
<tr>
<td>Died</td>
<td>2 (1%; 0.1-4)</td>
<td>19 (8%; 5-13)</td>
</tr>
<tr>
<td>Failed</td>
<td>32 (17%; 12-23)</td>
<td>26 (11%; 8-16)</td>
</tr>
<tr>
<td>Defaulted</td>
<td>15 (8%; 4-13)</td>
<td>58 (25%; 20-32)</td>
</tr>
</tbody>
</table>
24 month outcomes after treatment with DMD plus OBR in patients with XDR only. *Skripconoka et al, ERJ 2013*

<table>
<thead>
<tr>
<th>Treatment Outcome</th>
<th>6-8 months DMD N=44</th>
<th>0-2 Months DMD N=12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cured</td>
<td>11 (25%; 13-40)</td>
<td>5 (42%; 15-72)</td>
</tr>
<tr>
<td>Completed</td>
<td>16 (36%; 22-45)</td>
<td>1 (8%; 0.2-38)</td>
</tr>
<tr>
<td>Died</td>
<td>0 (0)</td>
<td>3 (24%; 5-57)</td>
</tr>
<tr>
<td>Failed</td>
<td>14 (32%; 19-48)</td>
<td>3 (25%; 5-57)</td>
</tr>
<tr>
<td>Defaulted</td>
<td>3 (7%; 1-19)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>
Does the abstract reflect the design of the study?

Does the abstract reflect the strength of the findings?
Thanks