Clinical Trials
Lecture 4: Data analysis

Dick Menzies, MD

Respiratory Epidemiology and Clinical Research Unit

Montreal Chest Institute

TB Research methods course

July 17, 2014

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 -**

<u>Can</u> the study be ended early – hypothesis answered.

Or,

<u>Should</u> 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)

Interim Analysis Example
Vernon et al Lancet 1999

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.

	Once weekly INH-RPT	Twice weekly INH-RIF	p value
Number	30	31	-
Relapse	5	3	.41
RIF-R	4	0	.05

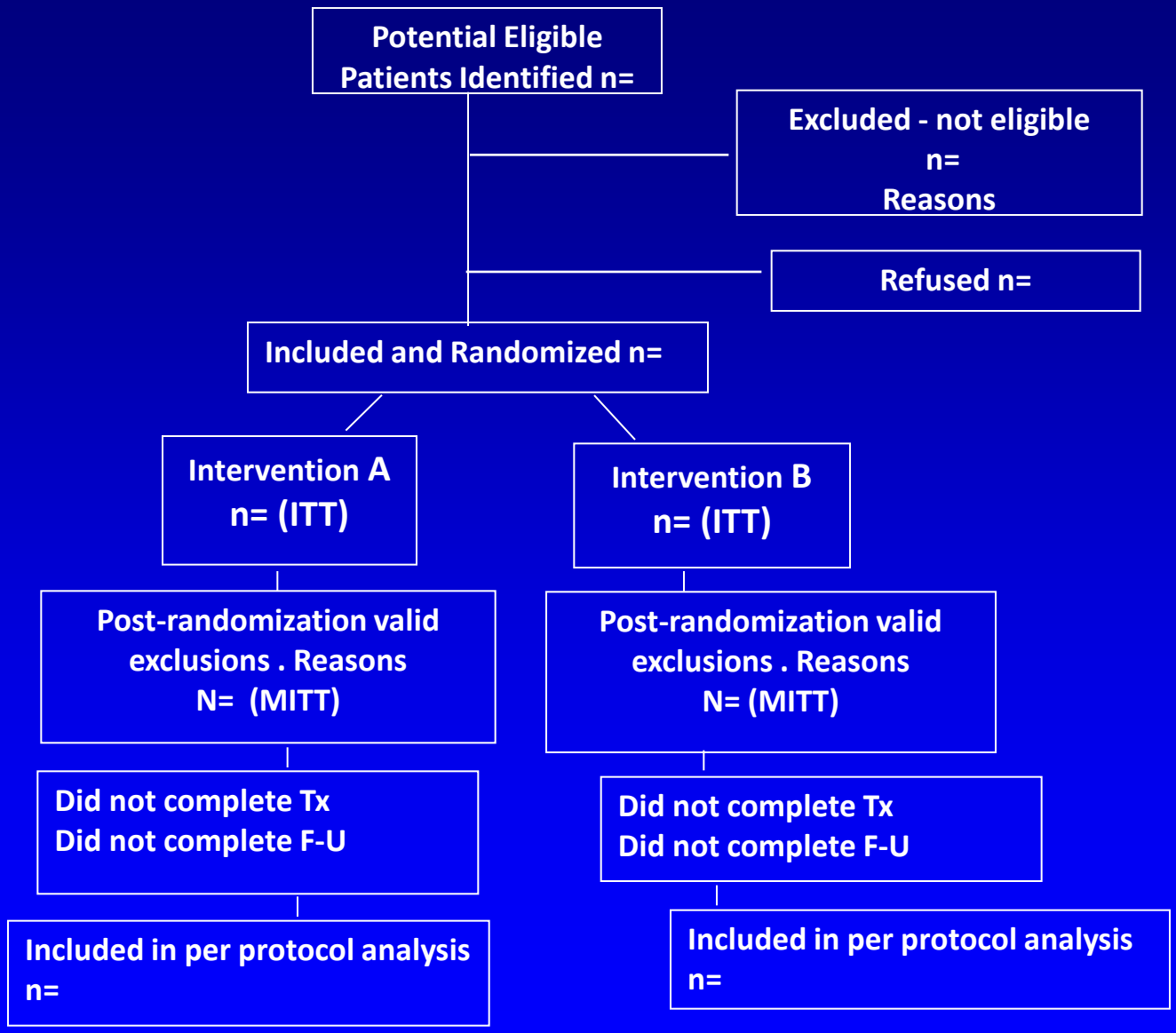
Final Analysis: Step 1 – Accounting for all subjects

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

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

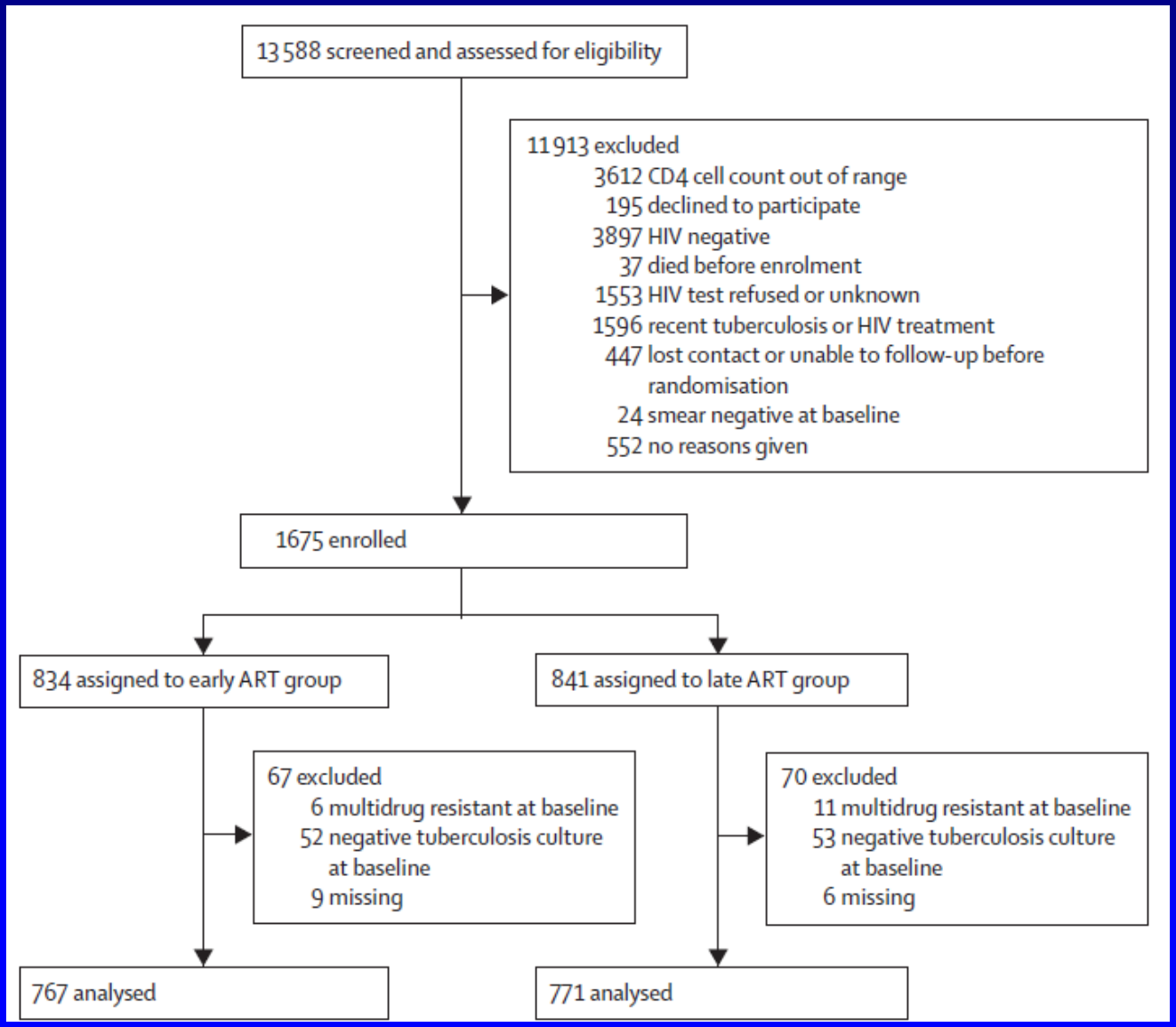
- *www.**consort-statement**.org/*

Consort diagram – general structure

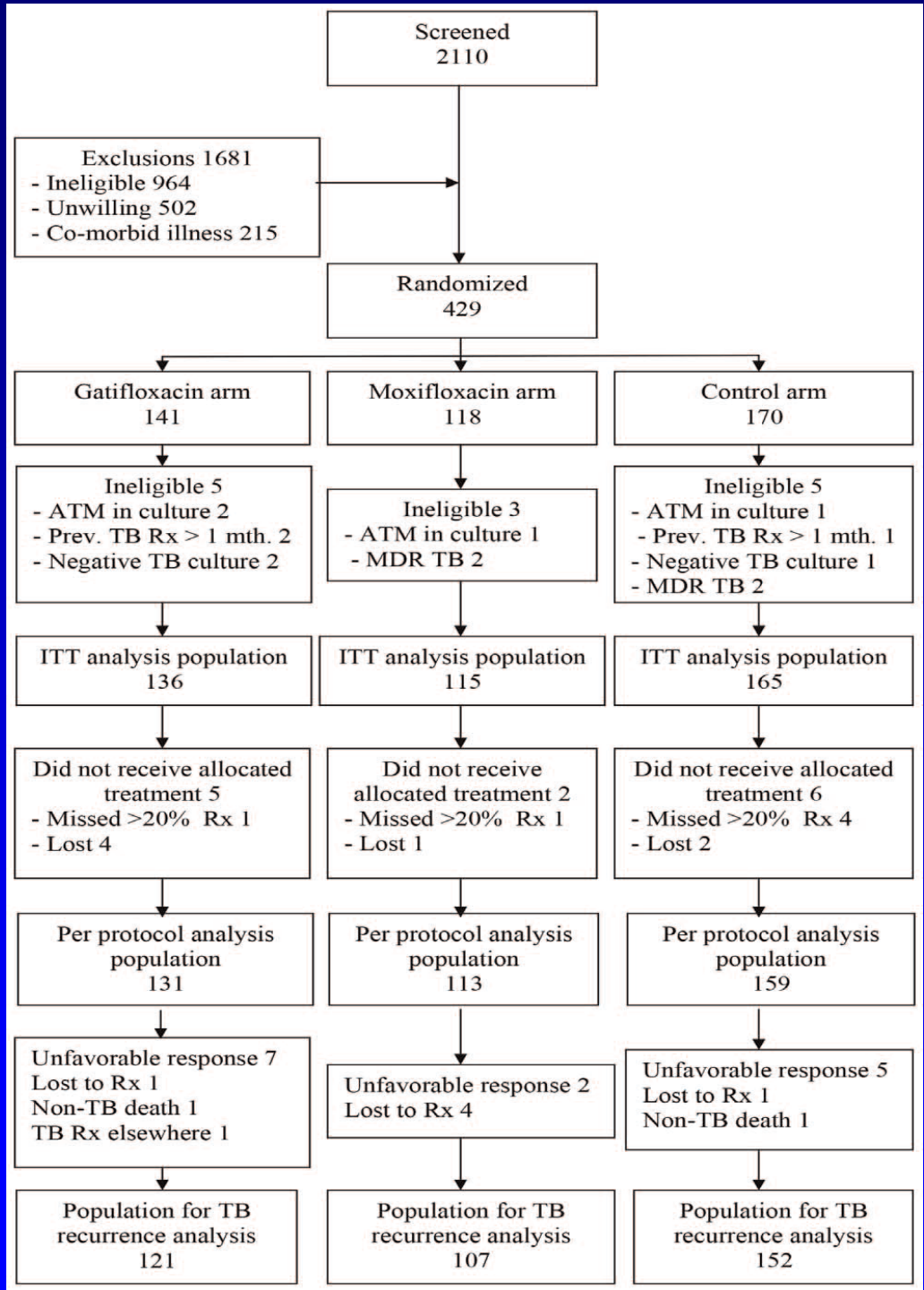


Consort diagram example

Mfinanga LID 2014



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

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

Patient Characteristics	Regimen		
	Gatifloxacin n=136	Moxifloxacin n=115	Control n=165
Sex			
Male	103 (76%)	83 (72%)	122 (74%)
Age (years):			
<40	90 (66%)	88 (77%)	120 (73%)
Body weight (Kg):			
Mean	43.7	44.2	43
Sputum culture			
3+ growth	107 (79%)	94 (82%)	127 (77%)
X-ray Chest			
>2 Zones affected	107 (79%)	94 (82%)	127 (77%)

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

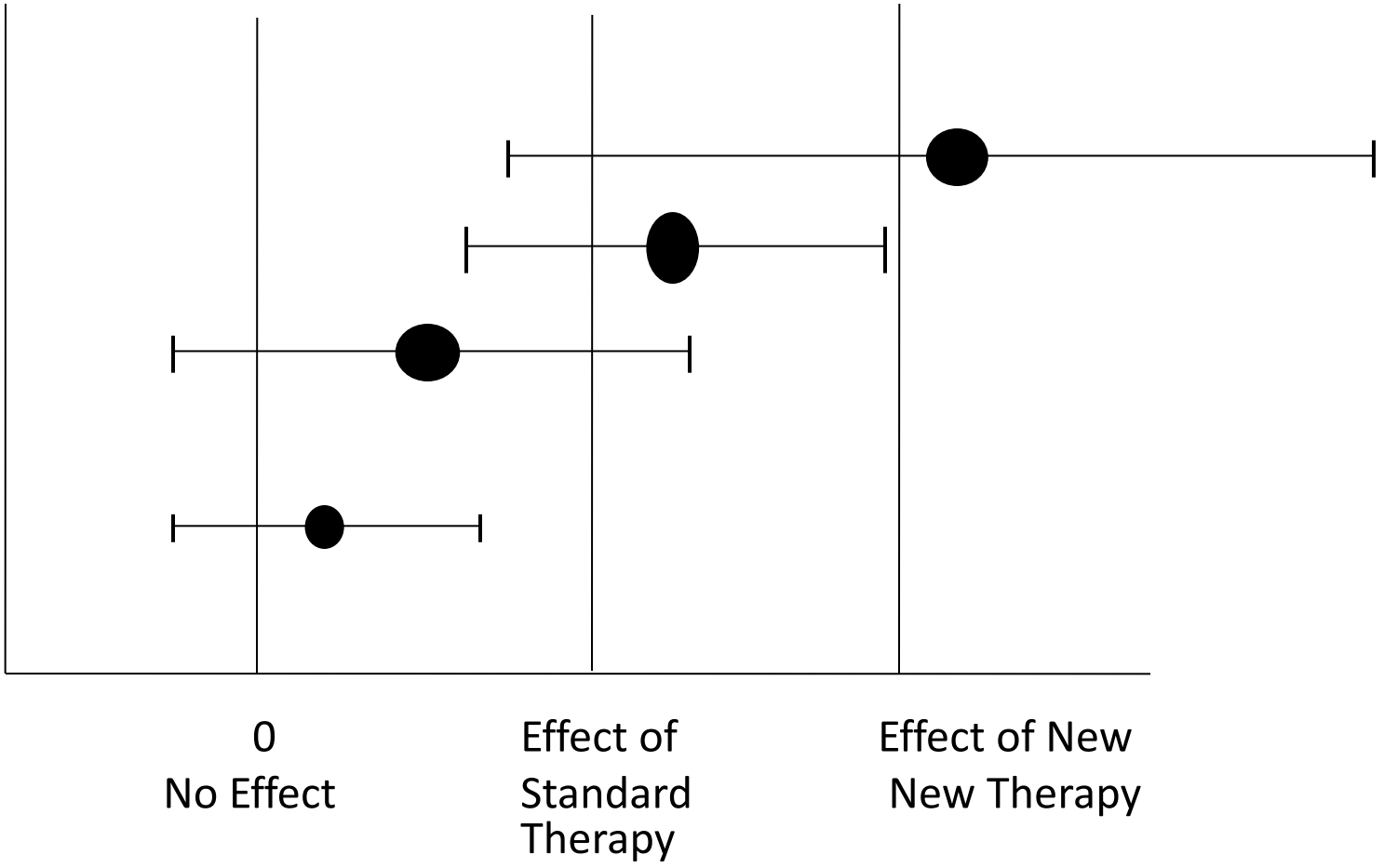
Superiority Studies (reminder)

- **Test New Interventions against a standard or placebo.**

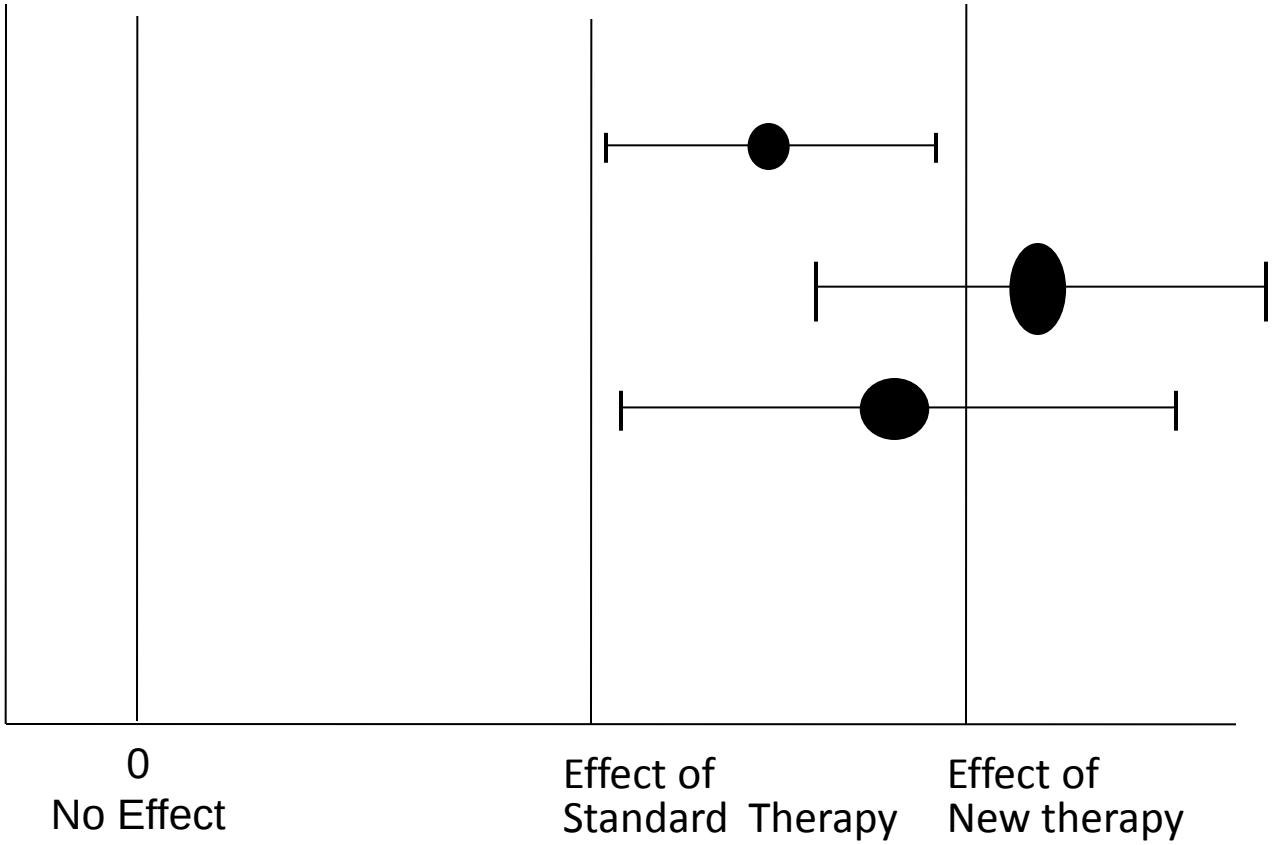
- **Hypothesis: New intervention is better.**

- **New intervention will be adopted if patients' outcomes are better.**

Superiority studies:
Results: 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

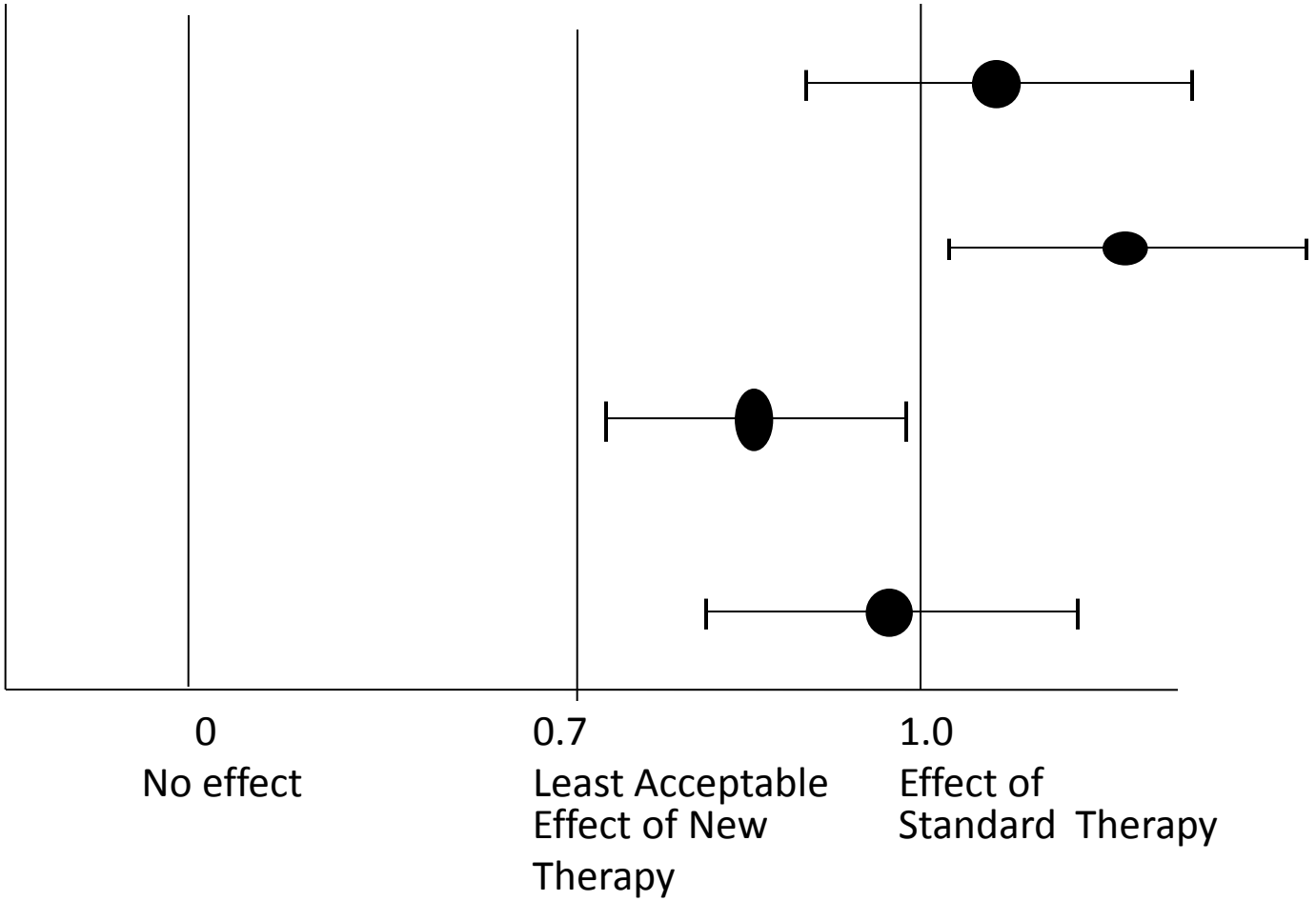
Alternate therapies must be cheaper, shorter, or safer.

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

This is called a Non-inferiority study.

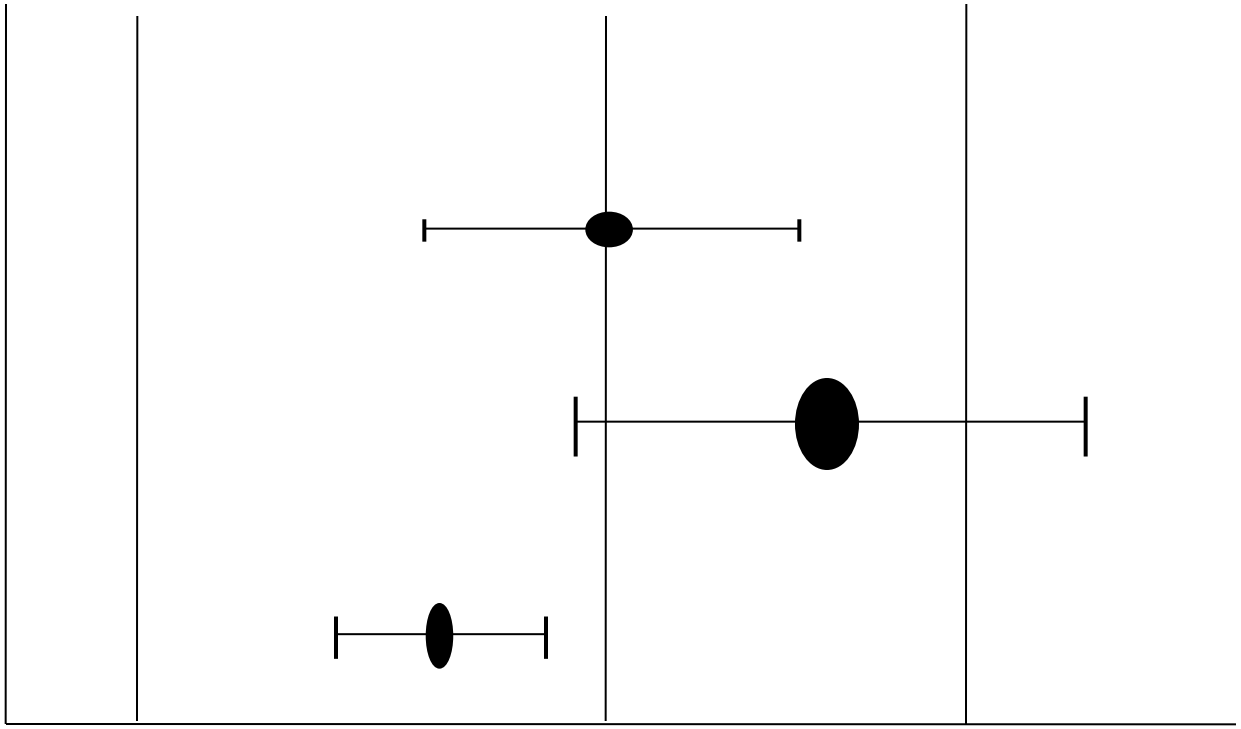
Non-Inferiority studies - Results

CAN conclude non-inferiority



Non-Inferiority studies - Results

CANNOT conclude non-inferiority



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)

Population	Duration	Reduction in TB
All participants (Effectiveness)	INH 12 mo.	75%
	INH 6 mo.	65%
	INH 3 mo.	21%
Completer/compliers (Efficacy)	INH 12 mo.	93%
	INH 6 mo.	69%
	INH 3 mo.	31%

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

Study Group	N	Subjects with Active TB			Difference in Cumulative Rate
		no.	no. per patient yr	Cumulative rate	percentage points
Modified intention-to-treat analysis					
Isoniazid only	3745	15	0.16	0.43	-0.24
Combination therapy	3986	7	0.07	0.19	
Per-protocol analysis					
Isoniazid only	2585	8	0.11	0.32	-0.19
Combination therapy	3273	4	0.05	0.13	

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

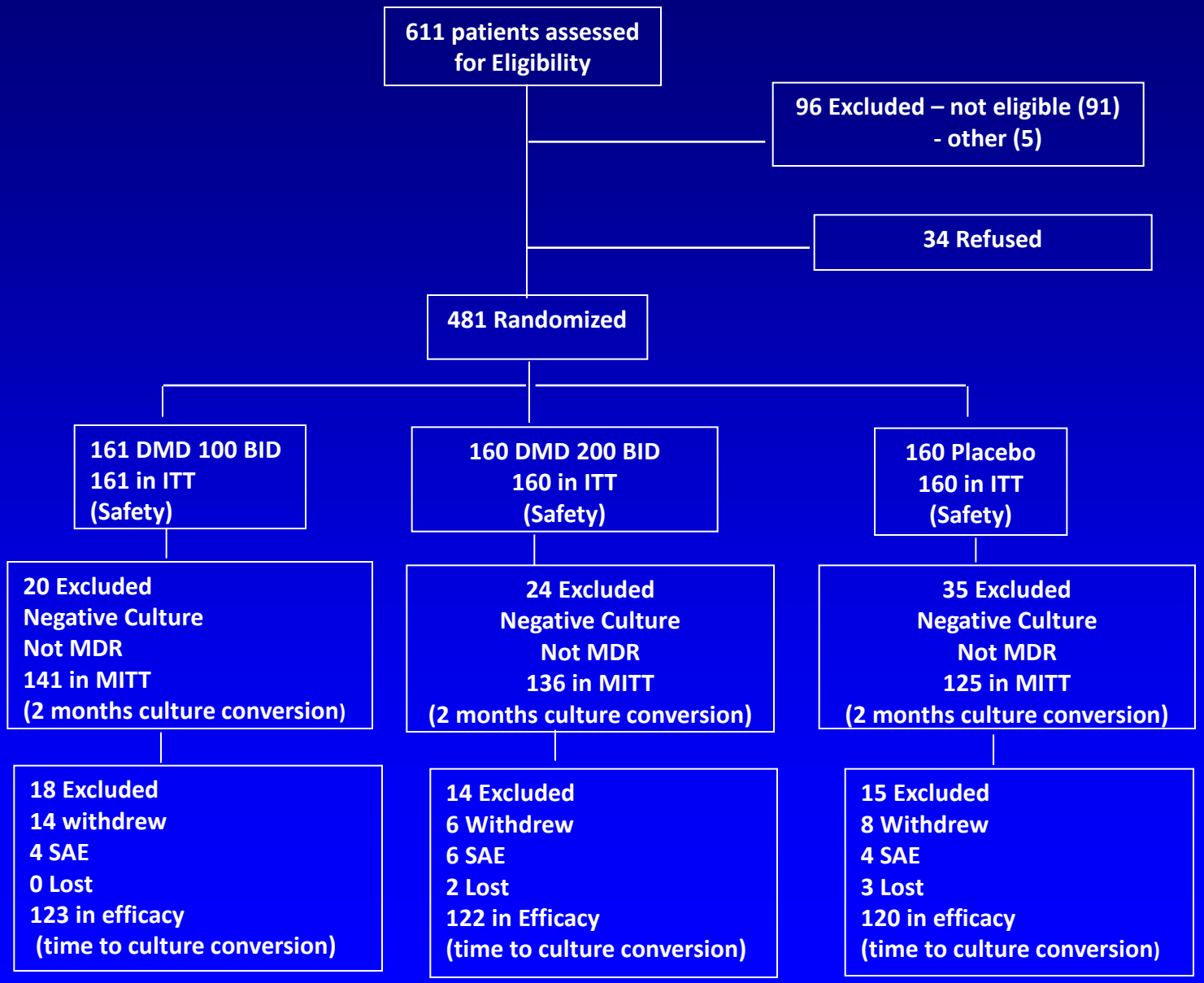
Planned Primary and Secondary analyses – example

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



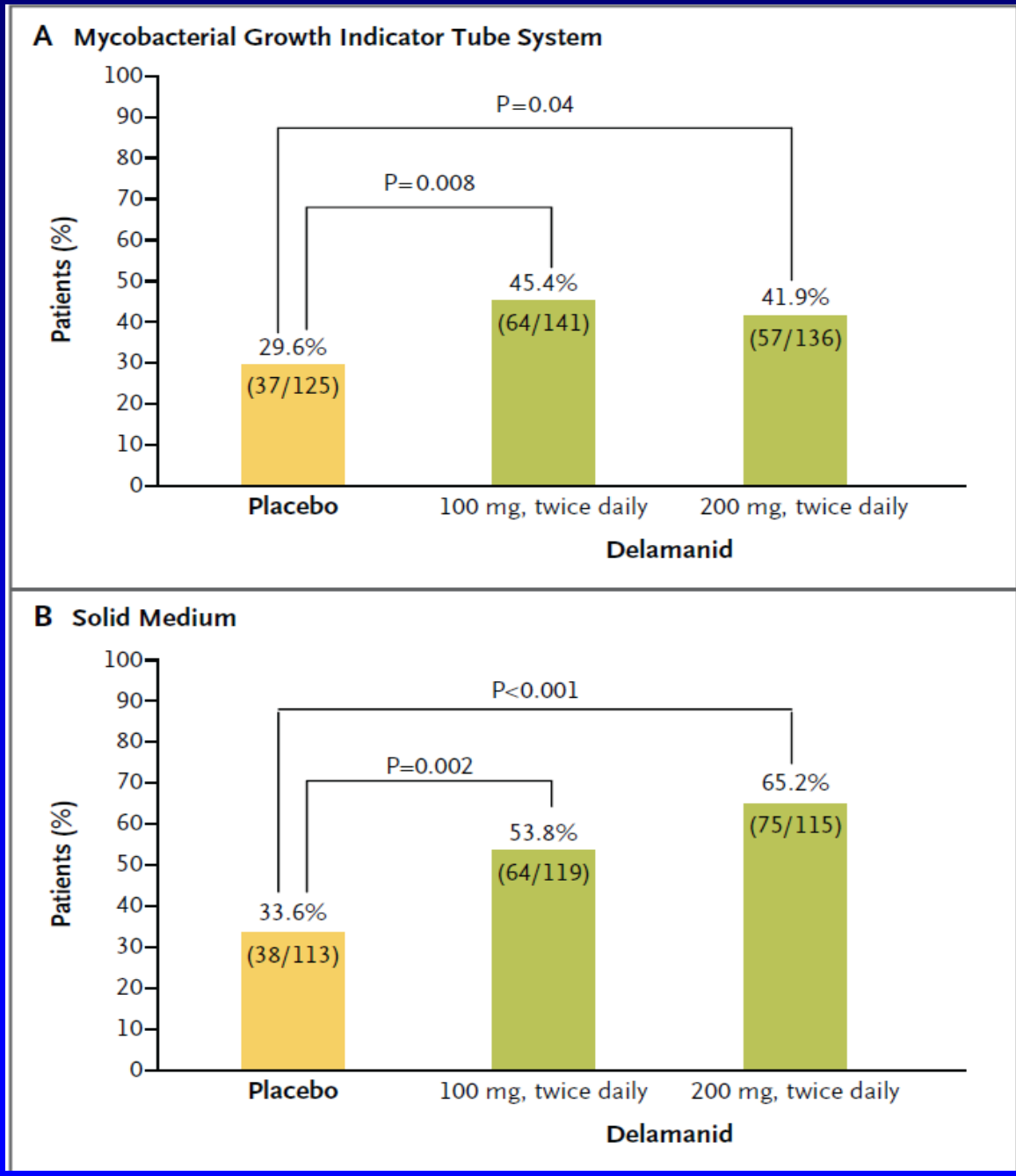
Planned secondary analysis: Incidence of Adverse Events

Uses ITT population (took at least 1 dose of study drug)

	Delamanid 100mg Twice Daily (N=161)	Delamanid 200mg Twice Daily (N=161)	Placebo (N=160)
Anemia	18(11.2)	10(6.2)	14(8.8)
Nausea	58(36.0)	65(40.6)	53(33.1)
Prolonged QT interval on ECG	16(9.9)	21(13.1)	6(3.8)
Paresthesias	17(10.6)	20(12.5)	12(7.5)
Anorexia	23(14.3)	34(21.2)	24(15.0)
Hypokalemia	20(12.4)	31(19.4)	24(15.0)

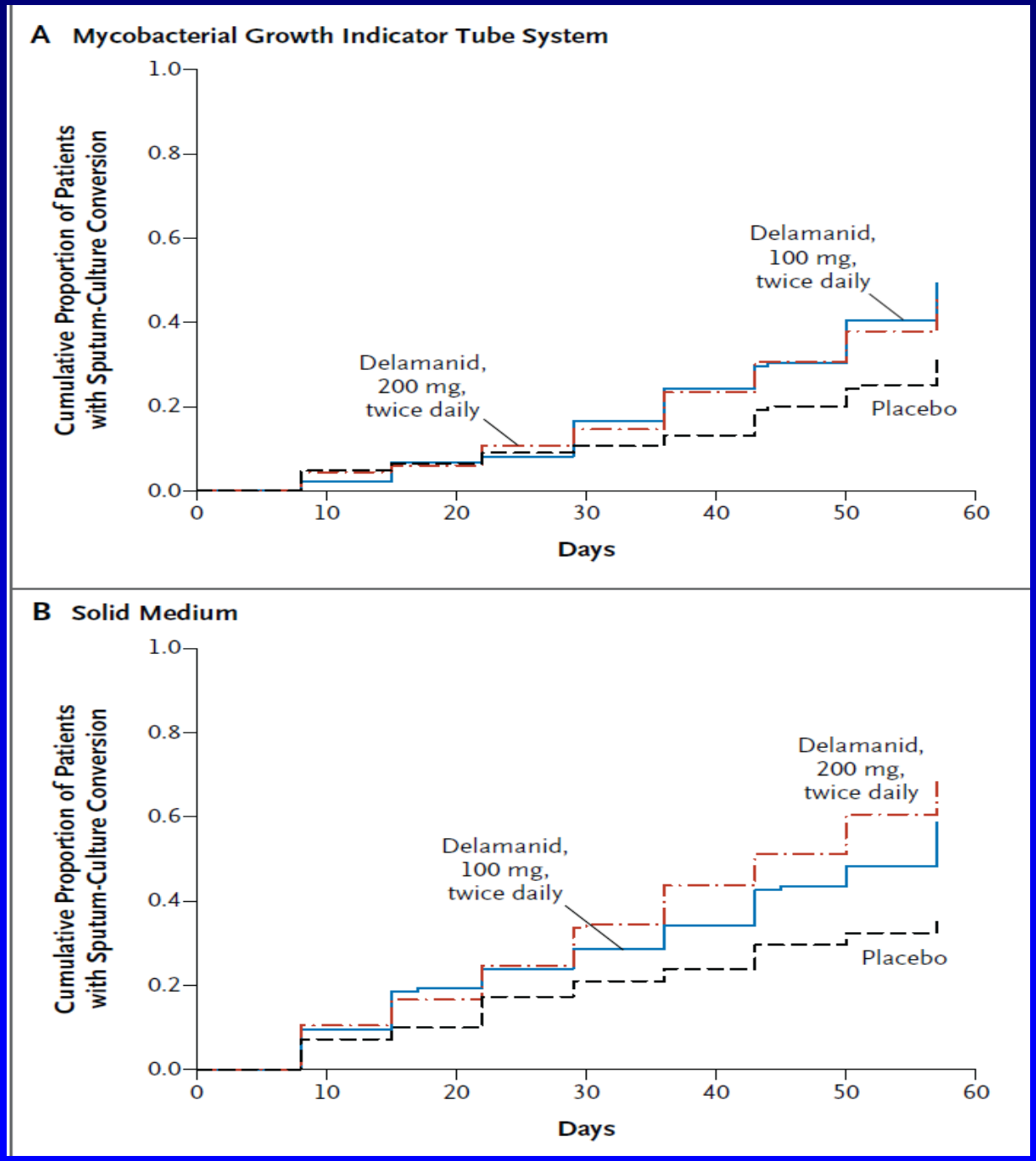
Primary analysis -
Uses MITT population:
2 Month culture
Conversion on MGIT

Planned secondary
Analysis -
Uses MITT population:
2 mos conversion
on solid media



Planned secondary Analysis – Efficacy:
Uses per protocol Population

Time to culture conversion

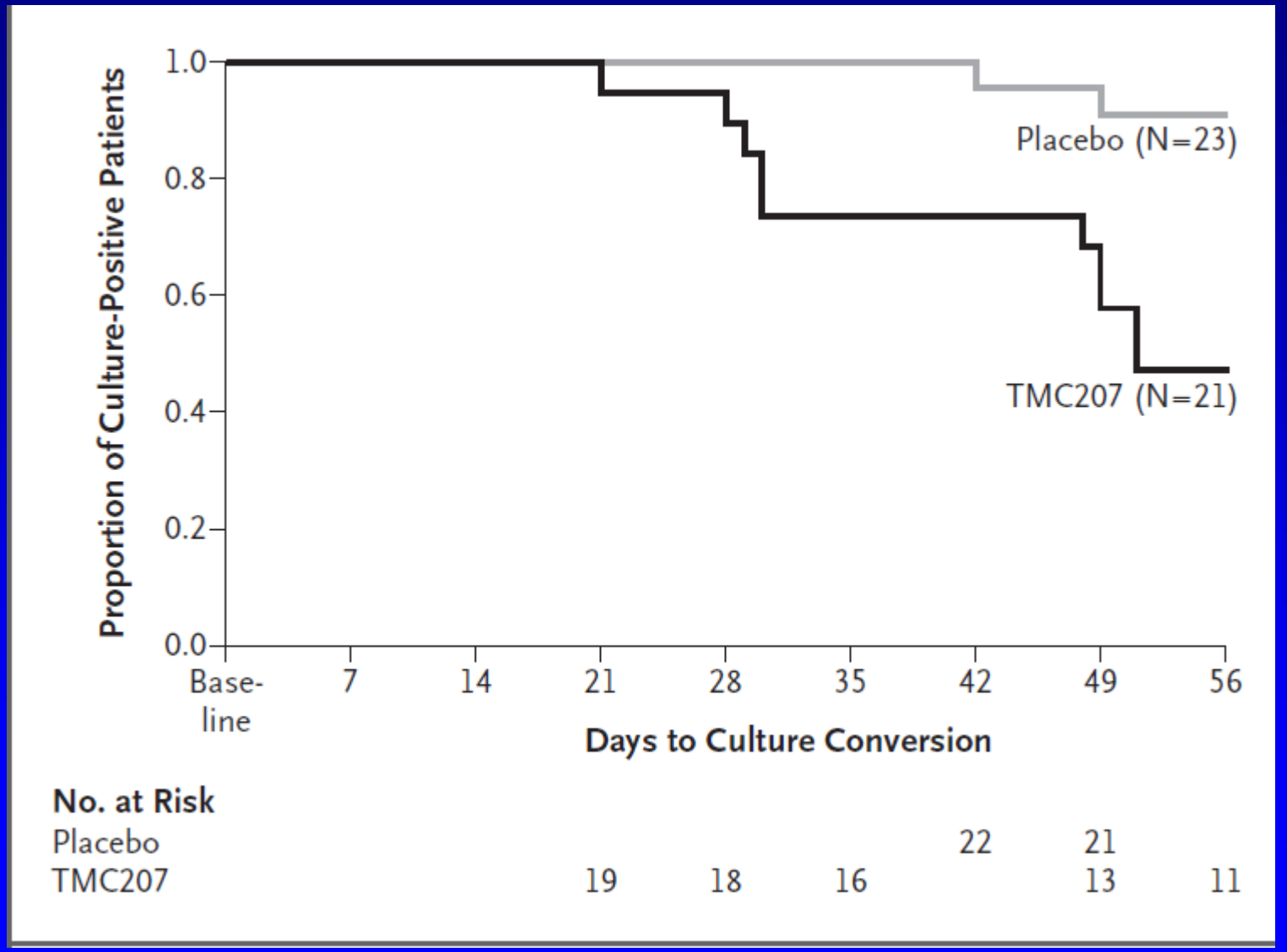


Planned Primary and Secondary analyses – example

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)

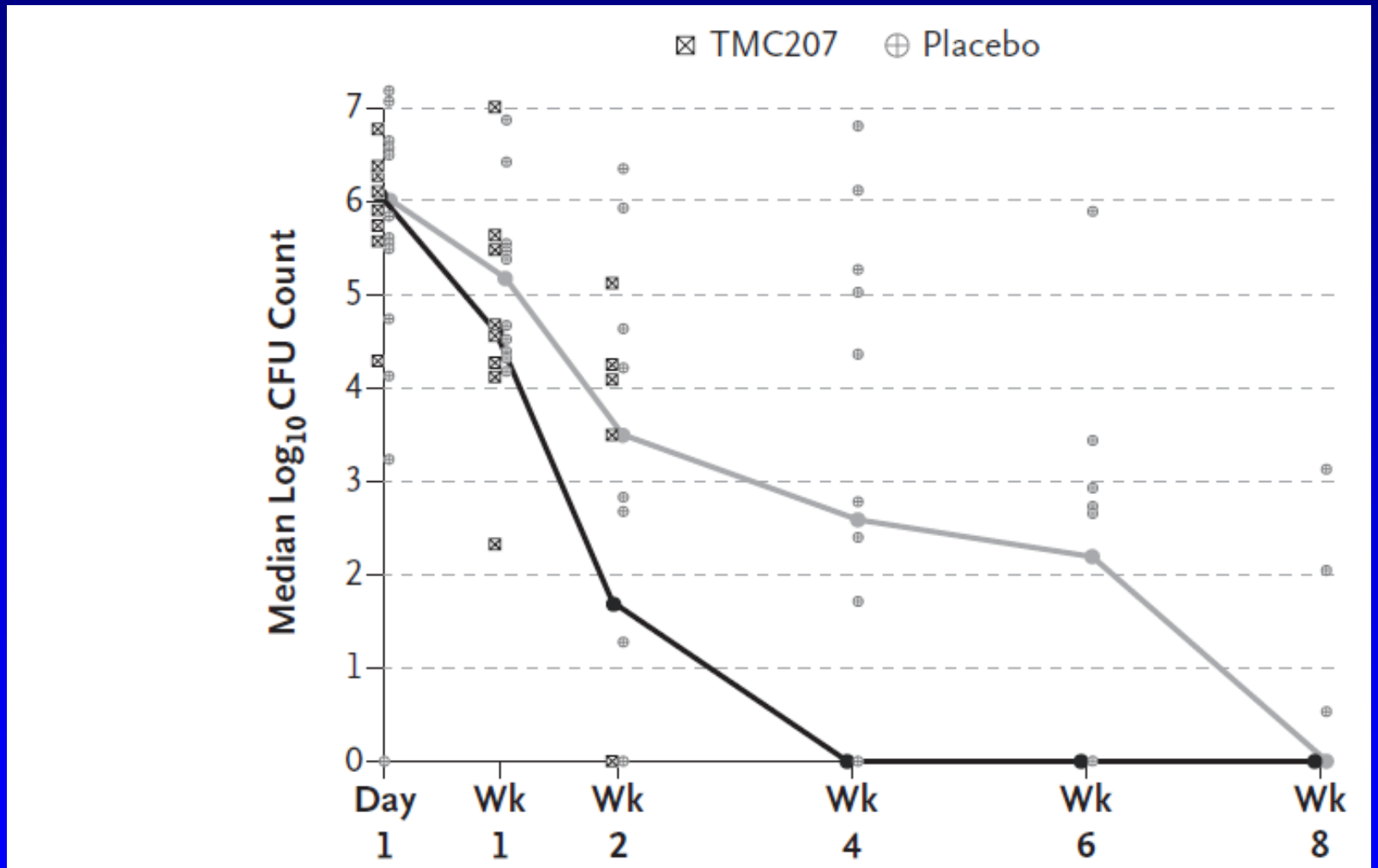


Secondary Analysis (ITT)

Incidence of Adverse Events

Adverse Event	TMC207 (N=23)	Placebo (N=24)
Nausea	6(26)	1(4)
Diarrhea	3(13)	1(4)
Arthralgia	4(17)	3(12)
Rash	2(9)	4(17)
Dizziness	3(13)	2(8)

Planned Secondary Analyses: (Efficacy)
Rate of bacterial killing (per protocol)



No. of Subjects						
Placebo	13	13	11	12	11	7
TMC207	9	8	9	6	5	7

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

Treatment Outcome	6-8 months DMD N=192	0-2 Months DMD N=229
Cured	110 (57%; 50-64)	111 (48%; 42-55)
Completed	33 (17%; 12-23)	15 (7%; 4-11)
Died	2 (1%; 0.1-4)	19 (8%; 5-13)
Failed	32 (17%; 12-23)	26 (11%; 8-16)
Defaulted	15 (8%; 4-13)	58 (25%; 20-32)

24 month outcomes after treatment with DMD plus OBR in patients with XDR only.

Skripconoka et al, ERJ 2013

Treatment Outcome	6-8 months DMD N=44	0-2 Months DMD N=12
Cured	11 (25%; 13-40)	5 (42%; 15-72)
Completed	16 (36%; 22-45)	1 (8%; 0.2-38)
Died	0 (0)	3 (24%;5-57)
Failed	14 (32%; 19-48)	3 (25%; 5-57)
Defaulted	3 (7%; 1-19)	0 (0)

Does the abstract reflect the design of the study?

Does the abstract reflect the strength of the findings?

Thanks