MRC Clinical Trials Unit

Smarter studies
Global impact
Better health

UCL

Trial design considerations for evaluation of new treatment regimens - an academic perspective

Andrew Nunn
MRC Clinical Trials Unit at UCL
25th November 2016

Outline

- Why we need new methods
- Multi-arm multi-stage designs
- STEP Phase IIc design
- Do we always need controls?
- Non-inferiority trials
  - the perils of biocreep
  - defining delta
- Endpoint definition

Assessing multiple regimens

- Of all the superiority RCTs registered in ClinicalTrials.gov between January 2010 and July 2012, 80% compared only one intervention with a control

Parmar, Carpenter, Sydes, Lancet, 2014

- How can we design future studies to maximise efficiency in the context of multiple new drugs and uncertainty about the optimum dosage for both efficacy and safety?

Reducing the time and cost in development of new TB regimens

Accelerating drug development

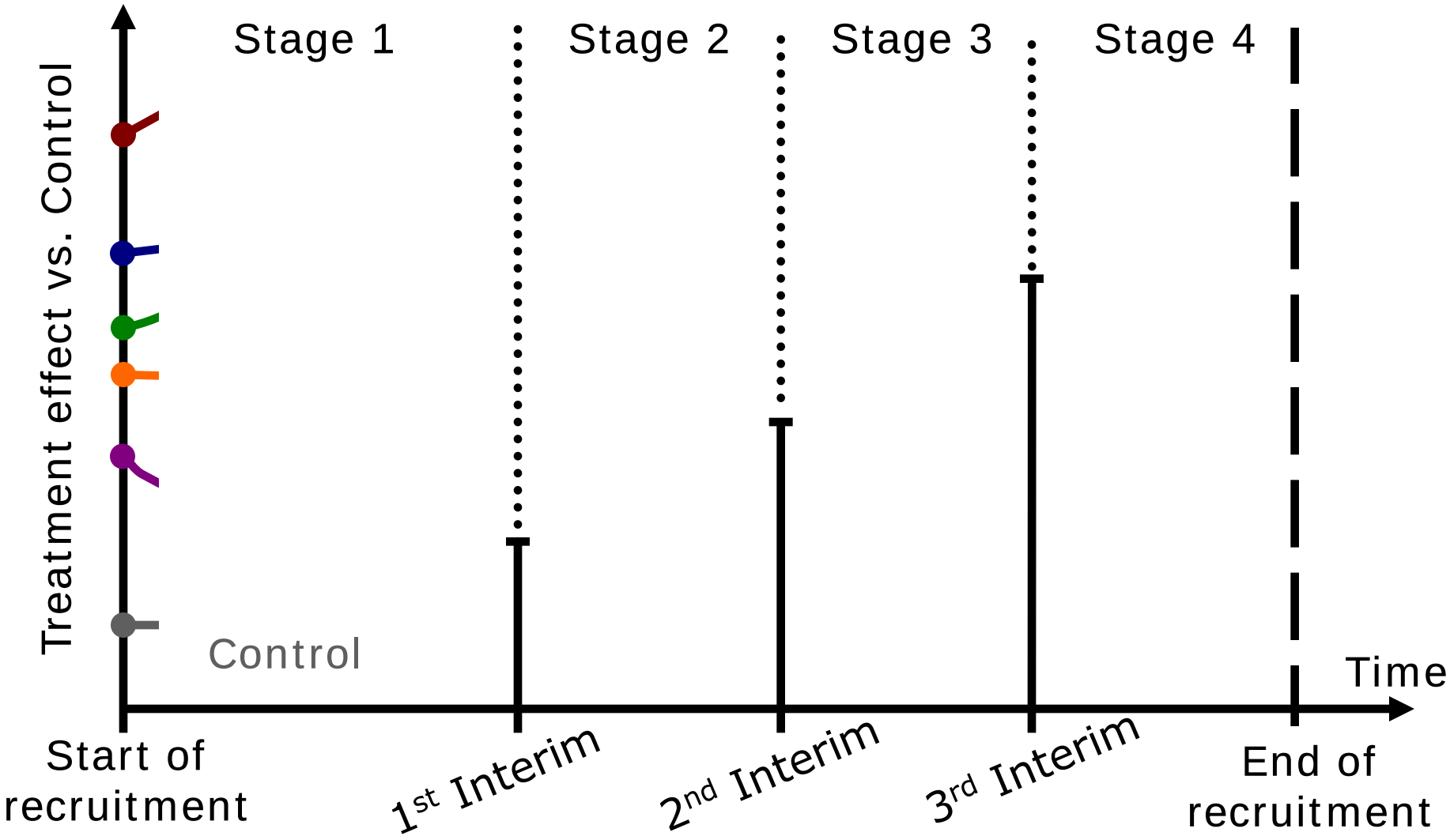
- Because there are now have several new candidate drugs in addition to the possibility of repurposing drugs like rifampicin it is becoming increasingly difficult to assess which combinations to take forward to phase III.

- The MAMS approach offers the opportunity to screen multiple regimens and drop those which are least promising, failing to achieve pre-specified targets.

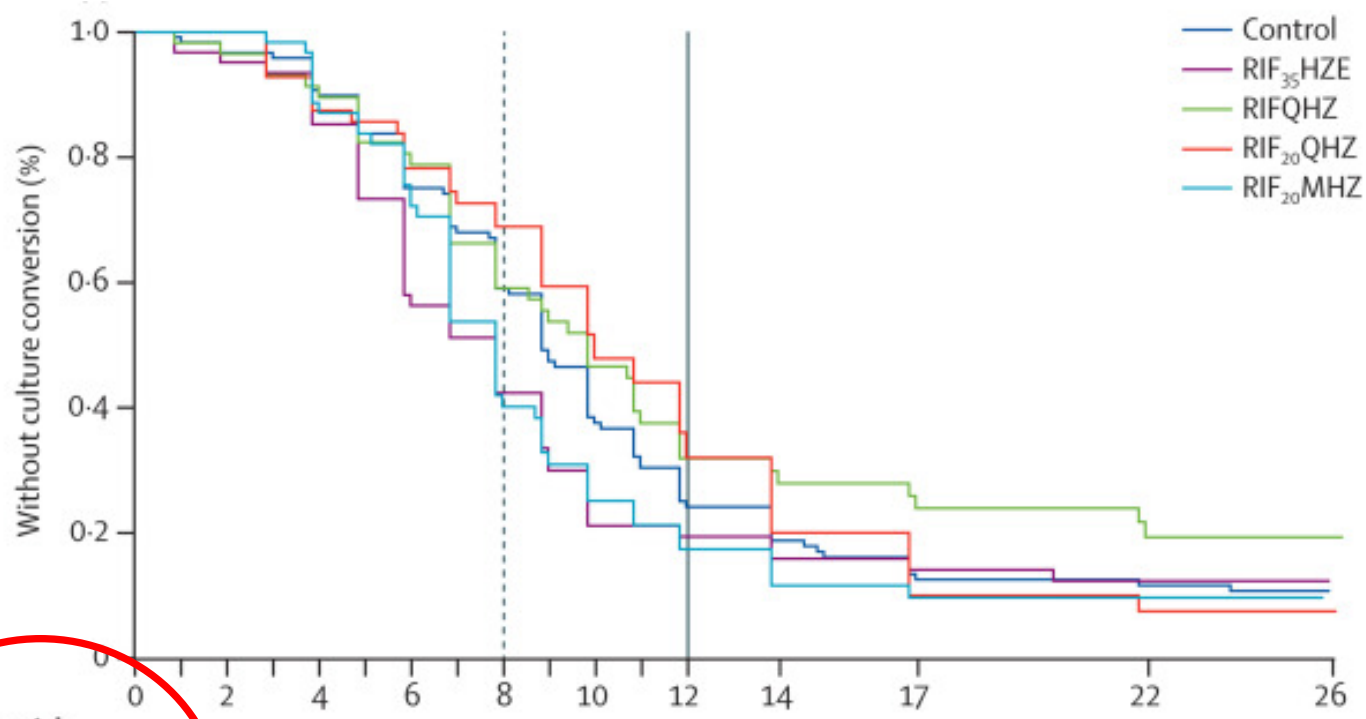
MAMS: multi-arm multi-stage designs

- Multi-arm phase II/III trials, originally developed in oncology, with planned interim analyses
- An **intermediate endpoint** used to compare each experimental arm with the common control
- Arms dropped if insufficient evidence of benefit using pre-specified critical values or hurdles
- The hurdles are progressively raised
- The final analysis is done on the **definitive endpoint** on the arms that remain

MAMS design example for 6-arm TB trial



PanACEA MAMS-TB: time to culture conversion (liquid medium)



Number at risk						
Control	123	105	66	27	14	13
$RIF_{35}HZE$	63	50	24	11	8	7
RIFQHZ	58	51	33	17	12	8
$RIF_{20}QHZ$	56	49	36	16	5	3
$RIF_{20}MHZ$	63	54	22	9	5	5

Boeree et al, Lancet Infec Dis, 2016

MAMS in TB

- Feasibility of MAMS design in TB demonstrated
- Arms without evidence of sufficient efficacy dropped early thereby reducing the sample size
- Slight risk of dropping an effective regimen
- Logistically challenging
- Culture results slow and not a good predictor
- Need for better and real time biomarkers measured earlier in treatment

- **Is MAMS enough?**

  Would limited data on relapse assist our decision making process?

STEP: accelerating development

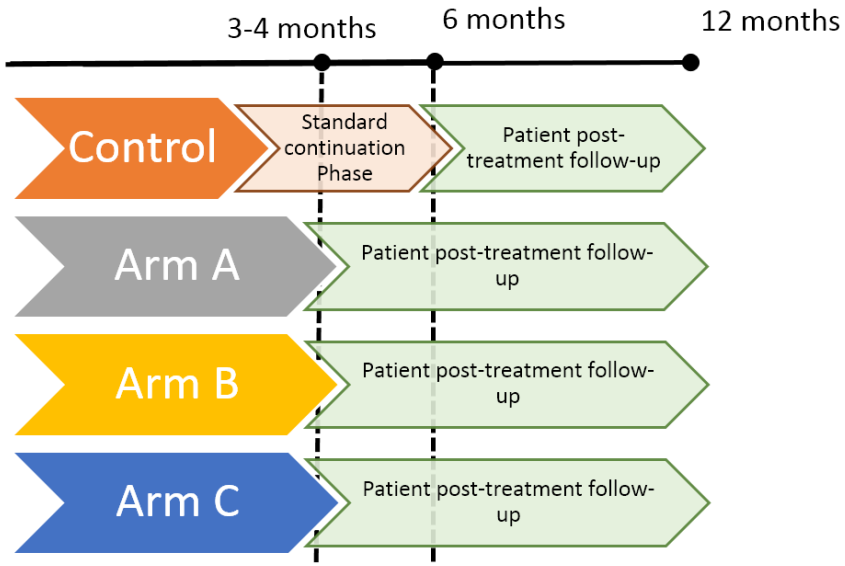
- Traditional design
  - **Phase I** safety
  - **Phase IIa** dose exploration (EBA)
  - **Phase IIb** early efficacy (culture conversion)
  - **Phase III** confirmatory non-inferiority (relapse)

- Culture conversion is of limited value in identifying regimens likely to be effective in Phase III

- A more informative Phase II study which included information on long term outcomes is desirable

STEP Phase IIc design

- sample size similar to Phase IIb study
- novel regimen(s) given for intended duration, 3 or 4m
- patients followed for 12 months post randomisation
- composite failure/relapse endpoint data collected



Decision-making for progression to phase III

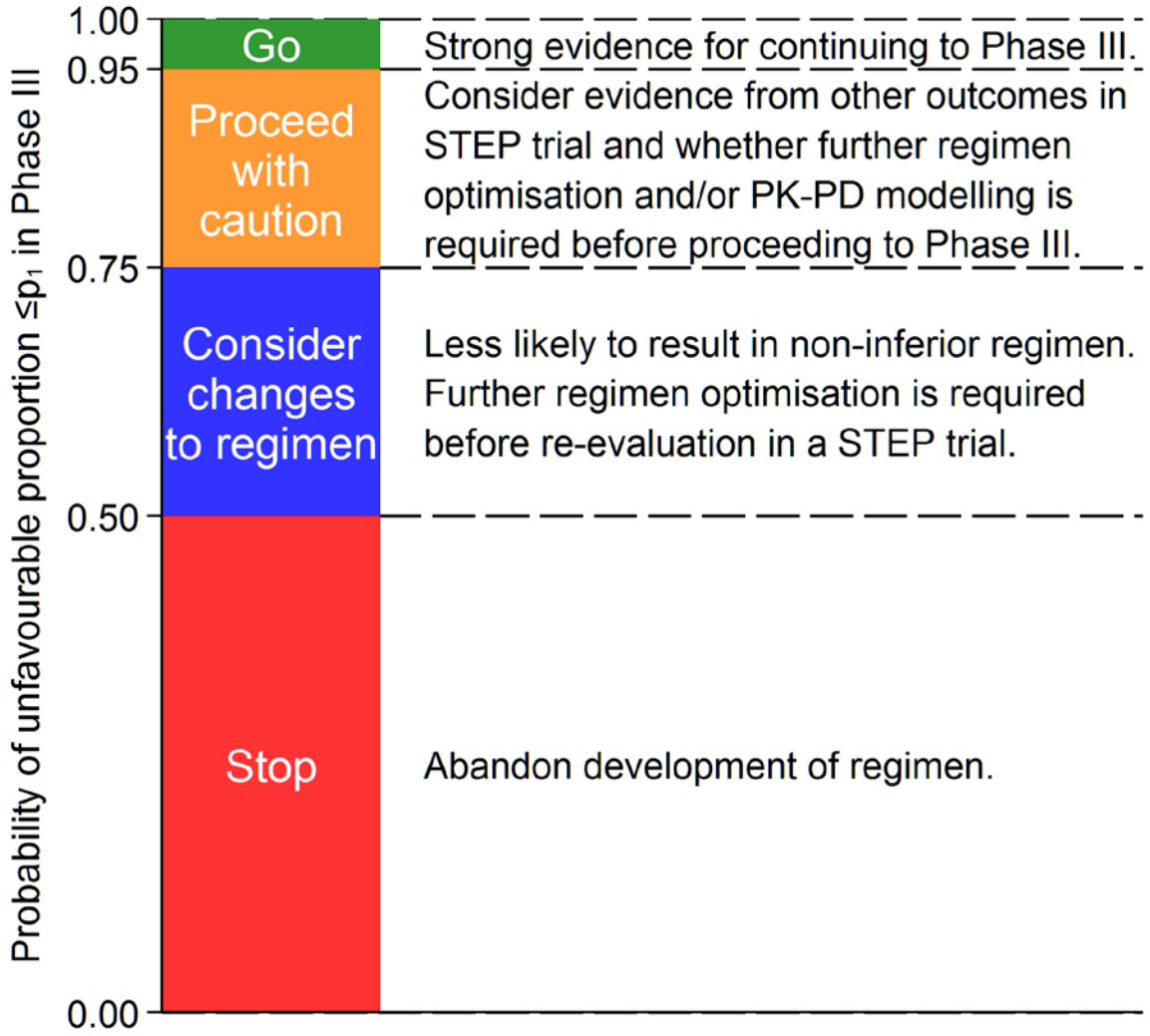
- **The key question:**

  What is the probability that new regimens will have efficacy at least as good as the 6-month control in a future phase III trial?

- The STEP design fits well into a **Bayesian framework** in which we assess the **predictive probability** that the unfavourable proportion is ≤ a pre-specified value p_1 in a hypothetical future phase III trial

Phillips et al, BMC Medicine, 2016

Traffic lights for Bayesian predictive probabilities, for given p_1



Uncontrolled trials – is there a place for them?

The penicillin experience

- Randomised clinical trials among those wounded in World War II in North Africa had been planned but the wish of surgeons not to withhold the penicillin from the severe cases led to disparity in the groups
  - The superiority of penicillin was so great that the benefit could still be demonstrated
- Is there a place for uncontrolled trials in TB, in particular in MDR-TB?

XDR-TB: do we need RCTs?

- Tugela Ferry, 52 of 53 patients died

  Gandhi et al, Lancet, 2006

- No agreed standard of care for XDR-TB

- "Outcome data show a treatment success rate of 28% for extensively drug-resistant TB (XDR-TB; 2013 cohort)"

  WHO Global TB Report, 2016

- **Nix-TB**: 42 patients enrolled to a single arm trial

  As of July 2016, there were no clinical or microbiological relapses

  47th Union conference, 2016, Liverpool

And MDR-TB?

- "Outcome data show a treatment success rate of 52% for MDR-TB (2013 cohort)"

  WHO Global TB Report, 2016

- Results from the last of six successive cohorts of MDR-TB patients in Bangladesh treated with a shortened regimen suggested that better options are available even without the introduction of new drugs

  Van Deun et al, AJRCM, 2010

- This has led to additional cohorts being studied and the development of the **STREAM** trial (2012) and has resulted in WHO recommendations for a shorter regimen (May 2016)

Non-inferiority trials – and in particular the choice of margin

Insufficient evidence?

- In 1993 WHO recommended an 8-month regimen, 2SHRZ/6TH, for DS-TB in which the duration of rifampicin was limited to 2 months

- In 2003 WHO modified this recommendation replacing streptomycin and thiacetazone by ethambutol

- It was not until 2010 that these recommendations were revoked based on results of The Union's Study A and a systematic review by Menzies which demonstrated the clear inferiority of the 8-month regimen

Jindani, Emerson, Nunn, Lancet, 2004

Menzies et al, PLOS Medicine, 2009

Non-inferiority or superiority?

- Non-inferiority trials are not necessarily larger than superiority trials

- **STREAM Stage 1** is designed as a non-inferiority trial on the assumption that although it is expected that the 9-month regimen will be more effective than the long 2011 WHO regimen the difference is expected to be small.

  9-month: 75% favourable

  WHO long regimen: 70%

- Given these assumptions:

  Total N for non-inferiority design: 400 (10% delta)

  superiority design: > 2500

Margin of non-inferiority

- How should delta be chosen?
- What should be done to minimise the possibility of biocreep, i.e. progressively inferior regimens being considered acceptable?

Choosing delta

- REMox TB substituted moxifloxacin for INH or EMB in the 2EHRZ/4HR regimen and assessed whether shortening treatment from 6m to 4m was possible.

- Delta was set at 6%*

  "this reflected consultation with clinicians in high-burden countries and reanalysis of previous trials showing the effect of shortening treatment to 4 months **without substituting a new drug**."

  Gillespie et al, NEJM, 2014

* The same margin was used in RIFAQUIN and OFLOTUB and a similar margin has been selected for TBTC Study 31 (6.6%).

The way forward?

- In the TB Alliance's NC006 study (**STAND**) the margin of non-inferiority is 12%
- Larger margins **=>** smaller, cheaper, faster studies
- But how do we avoid biocreep?
  - Smaller deltas?
  - Repeated studies?
  - Two controls?
- **STREAM Stage 2** includes two controls, the 9-month regimen from Stage 1 and the WHO 2011 recommended regimen

Endpoints

Mycobacterial definitions apart there is currently no clear consensus about other components of the primary efficacy endpoint

- All deaths: unfavourable? or only non-trauma deaths?
- Any change for drug-intolerance: unfavourable? or only more extensive change of treatment?
- Loss to follow-up: unfavourable? or unassessable?
- Reinfections: unfavourable? or unassessable?

Summary

- An increasing number of potential regimens to be assessed
- Need to be able to review multiple regimens together
- Culture conversion limited value for predicting long term outcome, an urgent need for new biomarkers
- MAMS and STEP designs enable more rapid differentiation between multiple candidate regimens
- Uncontrolled studies may have a place early in development
- Choice of the non-inferiority margin needs careful consideration as does the risk of bio-creep
- Endpoints need to be agreed

Thank you!